

30939

# KOMUNIKACJA NAUKOWA W ŚRODOWISKU CYFROWYM

Badania, zasoby, użytkownicy

NAUKA - DYDAKTYKA - PRAKTYKA





**KOMUNIKACJA NAUKOWA  
W ŚRODOWISKU CYFROWYM**

**Polish Librarians Association**  
SCIENCE-DIDACTICS-PRACTICE

**SCHOLARLY COMMUNICATION  
IN DIGITAL ENVIRONMENT**  
**Research, resources, tools**

Collective work edited by  
Agnieszka Korycińska-Huras, Małgorzata Janiak



UNIwersytet  
JAGIELLOŃSKI  
W KRAKOWIE

Warsaw 2014



**Stowarzyszenie Bibliotekarzy Polskich**  
NAUKA-DYDAKTYKA-PRAKTYKA

**KOMUNIKACJA NAUKOWA  
W ŚRODOWISKU CYFROWYM  
Badania, zasoby, użytkownicy**

Praca zbiorowa pod redakcją  
Agnieszki Korycińskiej-Huras i Małgorzaty Janiak



UNIwersytet  
JAGIELLOŃSKI  
W KRAKOWIE

Warszawa 2014



Komitet Redakcyjny serii wydawniczej  
«NAUKA – DYDAKTYKA – PRAKTYKA»

Jacek WOJCIECHOWSKI (przewodniczący), Stanisław CZAJKA, Artur JAZDON,  
Bożena KOREDZUK, Dariusz KUŹMINA, Mieczysław MURASZKIEWICZ,  
Janusz NOWICKI (sekretarz), Maria PRÓCHNICKA, Michał ROGOŹ, Barbara SOSIŃSKA-  
-KALATA, Elżbieta STEFAŃCZYK, Remigiusz SAPA, Anna TOKARSKA, Janusz TONDEL.

**Publikacja dofinansowana przez Wydział Zarządzania i Komunikacji Społecznej  
Uniwersytetu Jagiellońskiego oraz Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytetu Jagiellońskiego**

Recenzenci  
dr hab. Marek NAHOTKO  
dr hab. Remigiusz SAPA

Projekt okładki  
Funky Worky

Redakcja techniczna i korekta  
Jadwiga KRĘŻLEWICZ

© Copyright by Stowarzyszenie Bibliotekarzy Polskich  
© Copyright by Wydział Zarządzania i Komunikacji Społecznej UJ

ISBN 978-83-64203-37-4

CIP -- Biblioteka Narodowa

Komunikacja naukowa w środowisku cyfrowym :  
badania, zasoby, użytkownicy : praca zbiorowa /  
pod red. Agnieszki Korycińskiej-Huras i Małgorzaty  
Janiak ; Stowarzyszenie Bibliotekarzy Polskich. -  
Warszawa : Wydawnictwo SBB, 2014. - (Nauka,  
Dydaktyka, Praktyka ; 157)

# Spis treści

<b>Wstęp</b>	
Agnieszka Korycińska-Huras, Małgorzata Janiak .....	9
<b>Część I. Teoria, kontekst prawny</b>	
Marek Nahotko	
<i>Komunikacja naukowa jako proces organizacji wiedzy i informacji .....</i>	14
Diana Pietruch-Reizes	
<i>Otwarta nauka w świetle dokumentów Unii Europejskiej. Wybrane zagadnienia .....</i>	53
<b>Część II. Zielona droga OA. Zasoby obiektów cyfrowych</b>	
Sabina Cisek	
<i>Archiwa jakościowych danych badawczych w internecie .....</i>	68
Remigiusz Sapa	
<i>Wsparcie dla indywidualnego zarządzania informacją w polskich repozytoriach uczelnianych .....</i>	81
Agnieszka Korycińska-Huras	
<i>Archiwa i repozytoria Open Access we Francji. Analiza piśmiennictwa z lat 2005-2013 .....</i>	97
Krystyna Sanetra, Łukasz Mesek	
<i>Zasoby biblioteki cyfrowej a zasoby repozytorium instytucjonalnego w kontekście działania Jagiellońskiej Biblioteki Cyfrowej i Repozytorium UJ .....</i>	141
Cezary Mazurek, Tomasz Parkoła, Maciej Stroiński, Marcin Werla, Jan Węglarz	
<i>Infrastruktura dla eHumanistyki. Praktyczne aspekty gromadzenia i udostępniania źródeł w postaci cyfrowej na przykładzie repozytoriów i bibliotek cyfrowych w sieci PIONIER .....</i>	179

### **Część III. Zasoby bibliograficzne**

Leszek Szafrński <i>Rejestracja dorobku pracowników uczelni: Bibliografia Publikacji Pracowników Uniwersytetu Jagiellońskiego</i> . . . . .	192
Krystyna Sanetra <i>Baza e-PKA – rewitalizacja Starego Katalogu Biblioteki Jagiellońskiej</i> . . . . .	213
Krystyna Sanetra <i>KKZBUJ – OPAC nowej generacji?</i> . . . . .	241

### **Część IV. Użytkownicy, źródła, narzędzia, systemy**

Małgorzata Janiak <i>Użytkownicy bibliotek cyfrowych w piśmiennictwie anglojęzycznym: system versus użytkownik</i> . . . . .	270
Mirosław Górny, Jolanta Mazurek <i>Intensity of information system use as a factor in evaluations made by information users. The example of Polish digital library users</i> . . . . .	311
John Catlow, Mirosław Górny, Rafał Lewandowski <i>Digital libraries as a test bed for evaluating the effectiveness of information searching in OCR-processed texts</i> . . . . .	320
Małgorzata Janiak, Monika Krakowska <i>Wizualizacje modeli mentalnych bibliotek cyfrowych</i> . . . . .	341
Monika Krakowska <i>Źródła internetowe z zakresu nauk biologicznych i ich ocena w kontekście kształtowania kompetencji informacyjnych studentów</i> . . . . .	381
Paulina Motylińska <i>Wybrane zagadnienia wizualizacji w nauce o informacji</i> . . . . .	406
Agnieszka Korycińska-Huras <i>Benchmarking czasopism naukowych jako element strategii publikacyjnej naukowca. Przykład medycyny weterynaryjnej</i> . . . . .	428
Małgorzata Jaskowska, Magdalena Wójcik <i>Systemy informacyjne wspomagające podejmowanie decyzji w zakresie edukacji na poziomie wyższym</i> . . . . .	449



# Contents

## Introduction

Agnieszka Korycińska-Huras, Małgorzata Janiak .....	9
---	---

## Part One. Theory, legal context

Marek Nahotko <i>Scholarly communication as a process of knowledge and information organisation</i> ....	14
Diana Pietruch-Reizes <i>Open Science in light of European acts and regulations. Selected issues</i> .....	53

## Part Two. Green Road to OA. Collections of digital items

Sabina Cisek <i>Qualitative data repositories in Internet</i> .....	68
Remigiusz Sapa <i>Support for personal information management in Polish university repositories</i> .....	81
Agnieszka Korycińska-Huras <i>Open archives and repositories in France. State of the art 2005-2013</i> .....	97
Krystyna Sanetra, Łukasz Mesek <i>Collection of digital library versus resources of institutional repository in the context of the operation of the Jagiellonian Digital Library and Digital Repository of the Jagiellonian University</i> .....	141
Cezary Mazurek, Tomasz Parkoła, Maciej Stroiński, Marcin Werla, Jan Węglarz <i>Infrastructure for e-Humanities. Practical aspects of gathering and making available digital resources: example of repositories and digital libraries networked in PIONIER</i> .....	179

### **Part Three. Bibliographic resources**

Leszek Szafrński <i>Collecting information about academic staff's scholarly output: Bibliography of Publications of Jagiellonian University's Staff Members</i> . . . . .	192
Krystyna Sanetra <i>e-PKA – revitalization of the Jagiellonian Library's Old Catalog</i> . . . . .	213
Krystyna Sanetra <i>Jagiellonian Library's online catalog (KKZBUJ) – is it an OPAC of the new generation?</i> . . . . .	241

### **Part Four. Users, resources, tools, systems**

Małgorzata Janiak <i>Users of digital libraries in English-language scholarly literature: user versus system</i> . . . . .	270
Mirosław Górny, Jolanta Mazurek <i>Intensity of information system use as a factor in evaluations made by information users. The example of Polish digital library users</i> . . . . .	311
John Catlow, Mirosław Górny, Rafał Lewandowski <i>Digital libraries as a test bed for evaluating the effectiveness of information searching in OCR-processed texts</i> . . . . .	320
Małgorzata Janiak, Monika Krakowska <i>Visualizations of mental models representing digital libraries</i> . . . . .	341
Monika Krakowska <i>Evaluation of online resources in biological sciences in the context of developing information literacy among students</i> . . . . .	381
Paulina Motylińska <i>Selected problems of visualization in information science</i> . . . . .	406
Agnieszka Korycińska-Huras <i>Benchmarking of scholarly journals in veterinary medicine as a component of individual publishing strategy</i> . . . . .	428
Małgorzata Jaskowska, Magdalena Wójcik <i>E-advising systems in higher education</i> . . . . .	449

---

# Wstęp

Agnieszka Korycińska-Huras, Małgorzata Janiak

---

Przekazywana do rąk Czytelników monografia odzwierciedla zainteresowania naukowe teoretyków i praktyków, którzy w swojej działalności badawczej lub wdrożeniowej zajmują się różnymi aspektami komunikacji naukowej w środowisku cyfrowym. W zamyśle nie jest więc pełną panoramą wiedzy o wirtualnym obiegu przekazów naukowych, lecz uchwyconym w kadrze obrazem stanu badań i refleksji teoretycznej na ten temat, tak jak przedstawia się on w roku 2014. Trzeba bowiem zaznaczyć, że niemal wszystkie zamieszczone tu artykuły o charakterze badawczym zawierają relacje z badań przeprowadzonych w pierwszych trzech kwartałach bieżącego roku.

Książka dzieli się na cztery segmenty przedmiotowe. Część pierwszą, „Teoria i kontekst prawny”, otwiera artykuł Marka Nahotki przedstawiający nowatorskie ujęcie procesów publikowania naukowego w perspektywie językoznawczej teorii gatunków, z jednoczesnym podkreśleniem neurologicznej genezy struktur wiedzy wykształcanych w poznającym umyśle i psycholingwistycznych uwarunkowań praktyk informacyjnych. Ogólne spojrzenie na komunikację naukową zapośredniczoną przez media elektroniczne proponuje także Diana Pietruch-Reizes, która dokonuje przeglądu instrumentów prawnych regulujących lub wspierających urzeczywistnianie idei otwartej nauki, przewidzianych w dokumentach Unii Europejskiej.

W części drugiej, „Zielona droga OA. Zasoby obiektów cyfrowych”, skupione zostały teksty dotyczące przedsięwzięć repozytoryjnych realizowanych zgodnie z modelem zielonej drogi otwartego dostępu do treści i danych naukowych, oraz bibliotek cyfrowych zasilanych materiałem pochodzącym m.in. z digitalizacji dokumentów o wartości historycznej. Sabina Cisek rozpatruje metodologiczne, etyczne, organizacyjne, prawne oraz informacyjne i informatologiczne problemy związane z tworzeniem archiwów jakościowych danych badawczych (ang. *qualitative data repositories*), w których gromadzone są gotowe do ponowne-

go użycia empiryczne dane zastane i wywołane. Remigiusz Sapa kieruje swoją uwagę ku dziewięciu repozytoriom polskich uczelni, zarejestrowanym w Federacji Bibliotek Cyfrowych, które badał pod względem jakości oferowanych przez nie narzędzi wspierających użytkowników w realizowaniu określonego zestawu czynności z zakresu indywidualnego zarządzania informacją (ang. *personal information management* – PIM). Artykuł Agnieszki Korycińskiej-Huras, jedyny w tej części odnoszący się tylko do realiów zagranicznych, jest próbą odtworzenia, na podstawie analizy piśmiennictwa z lat 2005-2013, meandrów zielonej drogi OA we Francji, wynikających z konieczności pogodzenia centralnej archiwizacji dorobku naukowego w narodowym archiwum wielod dziedzinowym HAL i w repozytoriach lokalnych. Aspekt międzynarodowy uobecnia się także w pracy Krystyny Sanetry i Łukasza Meska, którzy charakteryzują politykę gromadzenia i obecny stan zasobów Jagiellońskiej Biblioteki Cyfrowej i Repozytorium Uniwersytetu Jagiellońskiego, zestawiając obie bazy z kilkunastoma przykładami podobnych inicjatyw zagranicznych. Z kolei Cezary Mazurek, Tomasz Parkoła, Maciej Stroiński, Marcin Werla i Jan Węglarz, reprezentujący Poznańskie Centrum Superkomputerowo-Sieciowe, przybliżają zasady funkcjonowania infrastruktury informatycznej służącej do zarządzania zasobami cyfrowymi polskich bibliotek i repozytoriów cyfrowych w ramach sieci Pionier, dowodząc że wysokowydajne rozwiązania w tej dziedzinie, oparte m.in. na wykorzystaniu technologii chmurowej, są równie potrzebne w humanistyce, co w naukach ścisłych.

Część trzecia poświęcona jest omówieniu trzech przedsięwzięć Biblioteki Jagiellońskiej, mających na celu usprawnienie zdalnego dostępu do zasobów bibliograficznych. Leszek Szafranski przedstawia stan zaawansowania prac nad Bibliografią Publikacji Pracowników Uniwersytetu Jagiellońskiego, zaś Krystyna Sanetra opisuje, w dwóch odrębnych, bogato ilustrowanych tekstach, efekty rewitalizacji tzw. Starego Katalogu Biblioteki Jagiellońskiej oraz postępy w doskonaleniu katalogu OPAC.

W części czwartej dominują zagadnienia związane z użytkownikami zasobów cyfrowych, choć znalazły tu miejsce również teksty dotyczące specyficznych kategorii źródeł informacji naukowej oraz narzędzi i systemów wspomagających uczestnictwo w komunikacji naukowej w środowisku cyfrowym. Małgorzata Janiak śledzi tendencje w anglojęzycznym piśmiennictwie naukowym na temat bibliotek cyfrowych z lat 1996-2014, lokalizując moment zwrotny, w którym badacze, wcześniej zaabsorbowani głównie kwestiami projektowania systemów informacyjno-wyszukiwawczych, zaczęli interesować się także zachowaniami i preferencjami użytkowników kolekcji. W badaniach opisanych w swoim artykule Mirosław Górny i Jolanta Mazurek sprawdzali hipotezę zakładającą, że jednym z istotnych czynników oceny biblioteki cyfrowej przez użytkowników jest

intensywność korzystania z jej systemu informacyjno-wyszukiwawczego. Autorzy analizują przy tym metodologiczne problemy związane z ustaleniem stopnia wiarygodności wyników oceny jakości bibliotek cyfrowych, uzyskiwanych od użytkowników, podkreślając że stanowią oni trudny do skalibrowania „instrument pomiarowy” o parametrach zmiennych w czasie. Użytkownicy bibliotek cyfrowych, a dokładniej studenci dwóch roczników studiów 1. i 2. stopnia w Instytucie Informacji Naukowej i Bibliotekoznawstwa UJ, byli przedmiotem badań Małgorzaty Janiak i Moniki Krakowskiej, które analizowały modele mentalne bibliotek cyfrowych, odzwierciedlone na wykonanych przez studentów rysunkach, poszukując w nich podobieństw do dwóch referencyjnych modeli teoretycznych: DELOS i 5S. Użytkownicy internetowych źródeł informacji naukowej, tym razem z dziedziny nauk biologicznych, są obecni również w artykule Moniki Krakowskiej, która ocenia przydatność tych zasobów w procesie kształtowaniu kompetencji informacyjno-komunikacyjnych (ang. *information literacy*) członków polskiej społeczności akademickiej – naukowców prowadzących dydaktykę i studentów. W oparciu o przegląd najnowszego piśmiennictwa Paulina Motylińska pokazuje użyteczność rozmaitych narzędzi wizualizacji informacji w usprawnianiu eksternalizacji wiedzy naukowej. Przyjmując punkt widzenia uczestnika komunikacji naukowej, Agnieszka Korycińska-Huras prezentuje propozycję praktycznego przewodnika dla naukowców, mającego ułatwiać wybór czasopisma naukowego na podstawie wyników indywidualnej oceny rentowności publikacji, na przykładzie periodyków z dziedziny medycyny weterynaryjnej. Potrzebami studenta, jako pełnoprawnego podmiotu komunikacji naukowej, a równolegle odbiorcy i nadawcy informacji związanych z tokiem studiów, zajmują się Małgorzata Jaskowska i Magdalena Wójcik, które w swoim artykule omawiają przykłady dobrych praktyk w zakresie projektowania i doskonalenia uczelnianych systemów wspomagających podejmowanie decyzji, przeznaczonych właśnie dla studentów.

Mamy nadzieję, że szerokie spektrum poruszonych w monografii tematów oraz różnorodność ujęć znajdzie uznanie u odbiorców zainteresowanych przeobrażeniami komunikacji naukowej w internecie oraz rozwojem oferty zasobów naukowych w postaci cyfrowej – uczonych, praktyków, studentów oraz obecnych i przyszłych użytkowników tych zasobów. Dziękujemy wszystkim Autorom, którzy zechcieli przyłączyć się do naszego przedsięwzięcia, oraz Recenzentom, bez których wyteżonej pracy książka nie mogłaby ukazać się tak szybko.



---

Część I

# **Teoria, kontekst prawny**

---

---

# Komunikacja naukowa jako proces organizacji wiedzy i informacji

Marek Nahotko

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

## Wprowadzenie

Komunikacja naukowa, w tym publikowanie naukowe, służy transferowi wiedzy pomiędzy uczestnikami procesu komunikowania. Wiedza, zakodowana w strukturach neuronalnych, będących materialną podstawą umysłu<sup>1</sup>, nie może jednak być przekazywana bezpośrednio „z mózgu do mózgu”. Transfer ten odbywa się za pośrednictwem informacji, do tej postaci bowiem wiedza jest konwertowana w celu jej komunikowania. W najprostszy sposób komunikacja naukowa, podobnie jak każda inna, może być postrzegana jako przekazywanie myśli, idei i znaczeń, będących treścią informacji, z jednego umysłu do innego za pomocą języka, co plasuje ją w nurcie empirystycznym [Piaget 1973, p. 10]. Zgodnie z tą koncepcją umysł ludzki jest rodzajem nieustrukturyzowanej pustki, niezapisanej kartki, *tabula rasa* (empiryzm genetyczny), a człowiek pozyskuje wiedzę o świecie, bezpośrednio poznając go poprzez zmysły i pośrednio przez stosowanie języka (a więc także za pomocą zmysłów – wzroku lub słuchu<sup>2</sup>). Łącznie wiedza oparta na percepcji i języku tworzy kopię funkcjonalną obiektów i procesów, składających się na świat obiektywny.

Wraz z rozwojem myśli filozoficznej i badań psychologicznych w XX w. empiryzm został porzucony i zastąpiony przez inny paradygmat, ogólnie zwany kon-

---

<sup>1</sup> Umysł jest tu rozumiany jako ogół aktywności mózgu. Według psychologów mózg jest materialnym podłożem działania umysłu jako systemu poznawczego, co oznacza, że funkcje umysłowe zachodzą na materialnym podłożu tkanki nerwowej i w ścisłej zależności od procesów przebiegających w tej tkance [Nęcka, Orzechowski, Szymura 2012, s. 24].

<sup>2</sup> Pomijam tu na razie funkcjonowanie tzw. wewnętrznego języka myśli [Fodor 2011].



strukturyzmem, na którym opiera się mój wywód. Sanna Talja i in. przytaczają definicję konstruktywizmu, określając go jako pogląd, według którego rzeczywistość jest konstruowana przez indywidualny umysł, ale w ścisłych relacjach ze światem zewnętrznym [Talja, Tuominen, Savolainen 2005, p. 81]. W informatologii konstruktywizm zwykle jest umieszczany w obrębie tzw. ujęcia kognitywnego, które ma swoje podstawy w psychologii kognitywnej i metodach badawczych nauk społecznych [Cisek 2002, s. 109]. Ujęcie to od kognitywizmu różni się tym, że główny nacisk kładzie się w nim na sposób, w jaki wiedza jest aktywnie tworzona przez podmiot poznający, czyli indywidualny umysł, w celu organizowania rzeczywistości wewnętrznej i zewnętrznej. Relacje pomiędzy wiedzą i informacją, czyli to, jak informacja zapisana w dokumentach wchodzi w relacje z wiedzą pojedynczych ludzi – odbiorców, oraz jak wiedza jednostek konwertowana jest w informację zdatną do przekazywania w procesie komunikacji, jest sednem dociekań informatologii [Brookes 1980, p. 128]. Obecnie informatologia zajmuje się także relacjami między wiedzą i danymi, czyli tym, jak dane z badań wchodzi w relacje z wiedzą; będzie jeszcze o tym mowa w dalszej części tekstu.

Istnieje wiele teorii i modeli opisujących procesy wiedzy. Tworzenie poprawnych teorii i modeli jest jednym z najważniejszych celów nauki, a jednocześnie jednym z największych stojących przed nią wyzwań. Modelowanie w informacji naukowej jest częstym sposobem przedstawiania jednostek działających (encji) oraz zachodzących między nimi procesów i relacji. Powstają liczne modele odwzorowujące zarówno procesy publikowania naukowego, procesy badawcze, jak i procesy wiedzy; w rzeczy samej, są to procesy tego samego rodzaju, a ich modele różnią się doбором encji. John Mackenzie Owen wskazuje dwa rodzaje takich modeli: po pierwsze te, w których stosuje się metaforę przewodu, przyjmując że informacja przepływa od nadawcy do odbiorcy przez przewód lub kanał (w swoich pracach opisał kilkanaście takich modeli [Mackenzie, Owen 2007, p. 80]). Do drugiej grupy zalicza natomiast modele kognitywne, w których komunikacja naukowa postrzegana jest jako proces polegający na wykorzystywaniu informacji, będącej odzwierciedleniem stanu mentalnego (wiedzy) autora, w celu zmiany stanu wiedzy odbiorcy; modeli takich skonstruowano już dziesiątki. Samych modeli i teorii w obszarze zachowań informacyjnych naliczono ponad 70 [Fisher, Erdelez, McKechnie 2005], a z pewnością nie są to wszystkie istniejące.

Według Donalda Case'a tworzenie modeli bywa związane z tworzeniem formalnej teorii [Case 2007, p. 120]. Modele, dzięki temu, że przedstawiane są często w formie graficznej, są łatwiejsze do zrozumienia od teorii; mogą także odnosić się do kilku teorii jednocześnie. Często dotyczą także problemów bardziej szczegółowych niż teorie. Jedne i drugie są uproszczonymi wersjami rzeczywistości.

W dalszej części tekstu przedstawię model publikowania naukowego jako procesu organizacji wiedzy i informacji. Łączy on wyniki dotychczasowych badań dotyczących modelowania publikowania naukowego jako części procesu badawczego oraz procesów wiedzy, w tym zachowań informacyjnych. W modelu tym uwzględnione zostały wyniki badań prowadzonych na gruncie wielu dziedzin wiedzy, przede wszystkim neurologii, psychologii i socjologii. Przedstawiony zostanie pogląd na miejsce organizacji wiedzy i informacji w tych procesach oraz rolę systemów organizacji wiedzy (SOW) i informacji (SOI), odmienny od dotychczas prezentowanych w literaturze.

Model ten ma ułatwić zrozumienie funkcjonowania struktur wiedzy i informacji w procesie komunikacji, w szczególności naukowej. Modelowane struktury funkcjonują w komunikacji obsługiwanej przez wszystkie technologie, zarówno analogowe (druk), jak i cyfrowe. W przypadku tych drugich powstają i znajdują zastosowanie także nowe rodzaje struktur. Funkcjonuje tu sprzężenie zwrotne: zmiany w środowisku (technologie informacyjne) powodują zmiany w strukturach wiedzy, a te pociągają za sobą tworzenie nowych elementów środowiska (nowych gatunków wypowiedzi), wpływających na struktury wiedzy. W opisie tych skomplikowanych zjawisk, cechujących się obecnie dużą dynamiką, zastosowane zostały teorie psychologii kognitywnej oraz teoria gatunków tekstu, wykorzystywana m.in. w językoznawstwie i socjologii. Pozwoliło to wykazać związki pomiędzy umysłową reprezentacją wiedzy a konwencjonalnymi strukturami gatunkowymi tekstów, m.in. publikacji naukowych.

## **1. Model organizacji wiedzy w procesie publikowania**

Innowacyjność procesów publikowania naukowego, wynikająca z ich przemierzania do sieciowego środowiska cyfrowego, ogranicza przydatność modeli „przewodowych” ze względu na znaczny wzrost liczby niezbędnych do uwzględnienia jednostek i złożoność relacji między nimi. Dlatego też w prezentowanym modelu konieczne stało się przejście na wyższy poziom ogólności, co osiągnąłem, przedstawiając procesy twórcze, służące przygotowaniu tekstu dzieła naukowego, jako procesy organizacji wiedzy (OW), a narzędzia językowe wykorzystywane w tych procesach jako system organizacji wiedzy (SOW). Produktem tych procesów jest informacja, która w dalszej kolejności organizowana jest z wykorzystaniem systemów organizacji informacji (SOI).

Pierwszym krokiem w kierunku utworzenia modelu było podjęcie decyzji co do wyboru i określenia podstawowych encji niezbędnych do ujęcia w modelu. Na ten temat również istnieje bardzo bogata literatura (por. [Bates 2005; Zins 2007]); zwy-

kle uwzględniane są takie jednostki, jak dane, informacja i wiedza, choć niektórzy autorzy uznają za konieczne wprowadzenie jeszcze innych, na przykład mądrości lub olśnienia [Świgoń 2012, s. 40]. Nie było jednak moim celem tworzenie nowych jednostek. Tam, gdzie to możliwe, starałem się wykorzystać elementy wskazane w opublikowanym piśmiennictwie, o których piszę w dalszej części tekstu.

Dla poprawnego modelowania procesów organizacji wiedzy w publikowaniu naukowym przyjąłem następujące podstawowe założenia:

- **Wiedza** (w tym naukowa) to utrwalona w pamięci reprezentacja rzeczywistości, powstała jako wynik funkcjonowania ludzkiego umysłu [Nęcka, Przechlewski, Szymura 2012, s. 137], czyli pracy centralnego układu nerwowego, która nie funkcjonuje w oderwaniu od niego; stan wiedzy jest indywidualną cechą każdej jednostki i zmienia się w każdej chwili w wyniku jej (jednostki) interakcji z otoczeniem, czyli uczenia się (pozyskiwania doświadczeń). Interakcje te polegają na wymianie bodźców, które trafiają do mózgu, poprzez narządy zmysłów, w formie danych lub informacji. Wiedzę należy więc traktować w sposób dynamiczny, raczej jako proces niż stan [Kron, Sofos 2008, p. 98]. Zgodnie z teorią Davida Kolba (Experiential Learning Model, ELT) w trakcie wykonywania zadania polegającego na rozwiązywaniu problemu lub osiągnięciu celu uzyskuje się doświadczenia [Kolb, Kolb 2005, p. 195]. Są one następnie przedmiotem refleksji (przemysłów) dokonywanych z wykorzystaniem istniejących struktur wiedzy. Dokonana refleksja umożliwia asymilację nabytych doświadczeń do istniejących struktur wiedzy i akomodację do nowej sytuacji. Zmiana struktur wiedzy powoduje kształtowanie się nowych zachowań, lepiej dostosowanych do potrzeb rozwiązania problemu, i testowanie tych zachowań. Wynika z tego, że wiedza jest tworzona w działaniu, przez co determinuje kolejne działania, które powodują modyfikację wiedzy jednostki. Jednostka tworzy więc swoją wiedzę o faktach, wartościach i procedurach podczas ciągłych interakcji swojej wiedzy z otoczeniem (światem zewnętrznym).

- Pozyskiwanie **danych** przebiega w interakcji ze środowiskiem, nie zaś w ramach realizacji procesów komunikacji, w tym komunikacji pomiędzy naukowcami. Bodźce pochodzące ze środowiska, w tym z przyrządów w laboratoriach badawczych i z eksperymentów, trafiają do naukowca w postaci danych. Dane nie zawierają więc subiektywnych odniesień typowych dla informacji<sup>3</sup>, bo

---

<sup>3</sup> Warto zwrócić uwagę, że według współczesnych opinii czysty empiryczny fakt jest iluzją, gdyż poznanie nie jest biernym doświadczeniem zmysłowym; umysł w trakcie poznawania jest aktywny [Ablewicz 1998, s. 27]. W naukach ścisłych, i w dużym stopniu także w przyrodniczych, jest to umysł badacza. W naukach społecznych natomiast, gdzie obiektem badań jest człowiek traktowany indywidualnie lub jako jednostka w społeczeństwie, w relację wchodzi dwa aktywne umysły: podmiotu i przedmiotu badań, przez co czasem trudno jest odróżnić dane od informacji, np. za pomocą badań ankietowych uzyskuje się dane, ale pochodzą one od respondenta – człowieka, są

nie reprezentują niczych struktur wiedzy. Nie można zatem wyróżnić procesu „tworzenia” danych, można je tylko zbierać i przetwarzać<sup>4</sup>. Natomiast ich odbiór jest subiektywny (podobnie jak w przypadku informacji) w takim sensie, że dane uzupełniają struktury umysłowe odbiorcy danych, modyfikując je w sposób indywidualny, zależny od wcześniej ukształtowanych struktur wiedzy odbiorcy (wynikających z jego wcześniejszych doświadczeń) oraz indywidualnych, kognitywnych zdolności i preferencji jednostki [Bergen 2012, p. 151].

- **Informacja** znajduje się w komunikatach przekazywanych między naukowcami w procesach komunikacji. Zawsze istnieje nadawca i odbiorca informacji (czasem potencjalny, wyobrażony). Ze względu na zakres moich badań najbardziej interesujące są dla mnie informacje zawarte w komunikatach publikowanych: publikacjach naukowych, stanowiących formę komunikacji formalnej. Istnieje jednak również komunikacja naukowa nieformalna, także polegająca na przekazywaniu informacji, na przykład ustnie, podczas dyskusji. Informacje zawarte w komunikatach są odzwierciedleniem (reprezentacją) stanu wiedzy, a więc stanu struktur umysłowych nadawcy. Podobnie jak dane, są one asymilowane w sposób indywidualny przez odbiorcę, zależnie od stanu struktur jego umysłu. Wynika z tego subiektywizm odbioru informacji: dla każdego odbiorcy ma ona nieco inne znaczenie. Jako że informacja tworzona jest świadomie, jej powstawaniu zawsze towarzyszy intencja przekazywania [Osika 2011, s. 51], a więc także przewidywanie odbioru, łącznie z tworzeniem przez nadawcę modelu odbiorcy, nawet gdyby przekaz miał być przesunięty w czasie i w przestrzeni. Przekazywanie wiedzy może odbywać się wyłącznie po jej uprzednim skonwertowaniu do postaci informacji, która po przekazaniu (odpowiednim zorganizowaniu) przekształcana jest w wiedzę odbiorcy. W wyniku przyswojenia informacji odbiorca nigdy nie uzyska struktury wiedzy tożsamej ze strukturą wiedzy nadawcy, ponieważ oznaczałoby to identyczność struktur wyjściowych, a nie ma dwóch ludzi o identycznych strukturach wiedzy [Krzysztofek 2014, s. 22]. Dodatkowo na sposób odbioru komunikatu i zawartych w nim informacji wpływa bardzo wiele czynników zewnętrznych, tworzących tzw. kontekst aktu komunikacji naukowej; jest on praktycznie zawsze różny dla autora i dla odbiorcy informacji.

---

więc obarczone subiektywizmem typowym dla informacji. To właśnie jest źródłem krytyki pod adresem metodologii nauk społecznych i wątpliwości co do ich „naukowości” [Ossowski 2001]. Kazimierz Krzysztofek twierdzi, że sytuacja ta może się zmienić dzięki wykorzystaniu olbrzymich zasobów danych cyfrowych o ludzkich zachowaniach i interakcjach, dostępnych głównie w internecie [Krzysztofek 2014, s. 27].

<sup>4</sup> Zgodnie z definicją twórczości, według której jest to proces prowadzący do nowego wytworu, akceptowanego przez pewną grupę w pewnym okresie [Nęcka 2012, s. 18], twórczy może być proces pozyskiwania danych, np. organizacja eksperymentu, ale same dane nie są tworzone w procesie twórczym.

- Z powyższych definicji wynika następujące rozumienie **SOW**: są to narzędzia tzw. metadyskursu, służące konwersji wiedzy do postaci informacji, w której wiedza staje się komunikowalna, a więc służące do przygotowania tekstu, który następnie może być publikowany. Dzięki stosowaniu tak rozumianych SOW wiedza indywidualna może być konwertowana, rozpowszechniana i asymilowana, w tym sensie stając się wiedzą społeczną. Tak definiowana OW należy do obszaru badań, dla którego nie ma w języku polskim dobrej nazwy, a który w języku angielskim dotychczas nazywano *scientific writing* [Hyland, Salager-Meyer 2008].

- Analogicznie **SOI** są narzędziami służącymi do organizowania informacji zawartych w komunikatach, obsługującymi procesy powielania (publikowania) i archiwizowania tekstów oraz wyszukiwania i udostępniania informacji zawartych w publikacjach. Są więc częścią systemu pośredniczenia pomiędzy nadawcą i odbiorcą informacji. Coraz częściej zresztą systemy tego rodzaju powstają nie tylko dla informacji, ale także dla nieprzetworzonych danych z badań (systemy organizacji danych – SOD).

Przedstawiona w moim modelu propozycja definicji danych, informacji i wiedzy najbliższa jest modelowi opisanemu przez Chaima Zinsa [Zins 2007, p. 498], w którym wiedza traktowana jest indywidualnie jako wewnętrzny produkt umysłu, a dane i informacja należą do świata zewnętrznego. Główna różnica pomiędzy tym modelem a moim polega na usunięciu bezpośredniej relacji między danymi i informacją, gdyż informacja nie musi pochodzić z przetwarzania (w umyśle) danych; może stanowić przetworzone informacje. Dane nie mogą być bezpośrednio komunikowane; muszą być przetworzone w umyśle, w wyniku czego stają się wiedzą, a więc informacja nie jest bezpośrednio tworzona z danych. Zarówno dane, jak i informacja bezpośrednio wpływają na stan wiedzy, poszerzając zakres doświadczeń jednostki, przy czym tylko informacja służy komunikowaniu wiedzy, jako że każdy akt komunikacji wywołuje potrzebę tworzenia informacji. Interpretacja danych (oraz informacji) prowadzi do tworzenia wiedzy, a nie informacji, chociaż oczywiście interpretacja danych może prowadzić do komunikowania nowo powstałej wiedzy (np. publikowane są wyniki badań).

Takie rozumienie danych i informacji jest w zasadzie zgodne z definicjami informacji spotykanymi w literaturze, wskazującymi że informacja (dane) jest „wykrytą (zarejestrowaną) różnicą, która jest istotna dla funkcjonowania organizmu” lub że „jest to treść oddziaływania: jego wielkość, zmiany w czasie i przestrzeni – traktowane w oderwaniu od pierwotnego fizycznego nośnika i wykorzystywane jako środki łączności systemów złożonych” [Nawrocki 2003, s. 49]. Również tak ogólna definicja, jak „informacja to wzorzec organizacji materii i energii” [Bates 2006, p. 1033] jest użyteczna, pod warunkiem że poczyni się rozróżnienie między informacją i danymi, przedstawione wcześniej w tym artykule. Przedstawione

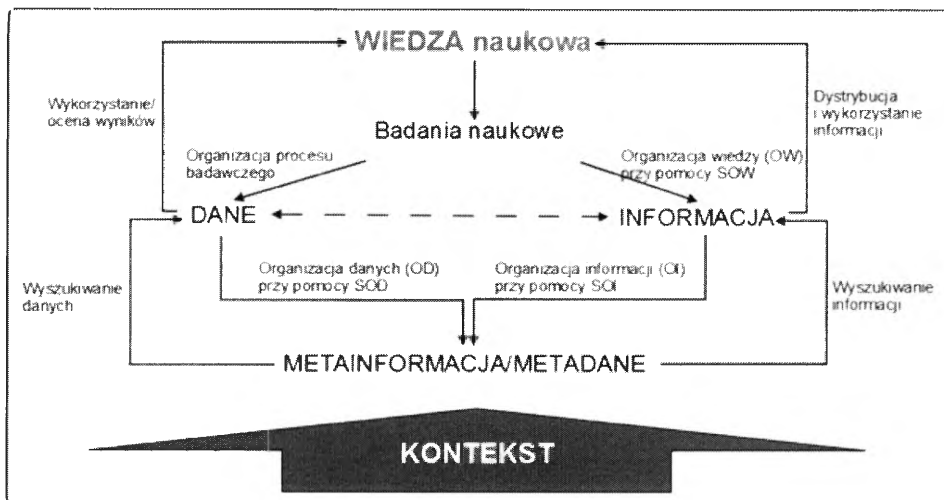
przeze mnie rozumienie informacji dotyczy wówczas jej szczególnego przypadku: informacji w komunikacji naukowej, czyli informacji wyższego poziomu, przetworzonej w umyśle, przez Andrzeja Chmieleckiego zwanej metainformacją<sup>5</sup> [Chmielecki 2001, s. 40]. Oczywiście nie jest moim zamiarem odnoszenie się do wszystkich istniejących definicji informacji, choćby ze względu na ich wielką liczbę (por. [Bawden, Robinson 2012, p. 65]).

Przedstawione w modelu kryterium służące rozróżnieniu między danymi i informacją jest bardzo zbliżone do tego, co postuluje Patrick Wilson, twórca teorii autorytetów kognitywnych [Wilson 1983]. Według niego ludzie tworzą wiedzę na podstawie dwojakiego rodzaju źródeł poznania: w oparciu o własne doświadczenia (czyli dane w mojej nomenklaturze) oraz na podstawie tego, czego dowiedzieli się od innych (czyli informacji). Wiedza uzyskana na podstawie doświadczenia (danych) zależy od zasobu poglądów (czyli wiedzy), które służą do interpretacji i zrozumienia otaczającego świata. Obszar wiedzy wynikającej z własnego doświadczenia uzupełniany jest informacją zdobywaną w procesach komunikacji z innymi osobami. Realizacja tych procesów powoduje powstawanie wielu relacji i zależności pomiędzy komunikującymi się stronami. Duża część wiedzy jednostki o otaczającym świecie ma źródło w informacjach uzyskiwanych od ludzi bezpośrednio lub za pośrednictwem nieosobowych źródeł informacji.

Przedstawiony przeze mnie model transferu danych, informacji i wiedzy w procesie komunikacji naukowej jest krokiem w kierunku porzucenia liniowego odwzorowywania procesów OW zachodzących w publikowaniu naukowym. Zostało ono potraktowane jako zespół procesów realizowanych w wyniku kodyfikacji (eksternalizacji) wiedzy, której rezultatem jest powstanie komunikatu zawierającego informację. W trakcie tych działań naukowcy muszą decydować o celach, jakim będzie służyć kodyfikowana wiedza, ewaluować tę wiedzę ze względu na jej przydatność z punktu widzenia przyjętych celów oraz identyfikować możliwe formy informacji, przydatne do osiągnięcia zamierzonych efektów. Muszą też identyfikować medium odpowiednie dla kodyfikacji wiedzy i dystrybucji informacji oraz stosować SOW przyjęte dla tego medium. Wszystkie te procesy i podejmowane w ich trakcie decyzje można uznać za efekt stosowania wybranego SOW w eksternalizacji wiedzy. Wybór SOW na tym etapie ma także bezpośredni wpływ na rodzaj SOI stosowanych na dalszych etapach publikowania naukowego.

---

<sup>5</sup> Konieczne jest odróżnienie metainformacji w rozumieniu Chmieleckiego od metainformacji w rozumieniu modelu przedstawionego przeze mnie na rysunku 1. W pierwszym przypadku metainformacja (informacja zapamiętana w pamięci osobniczej) określana jest przeze mnie mianem wiedzy; w drugim jest to informacja przetworzona z wykorzystaniem SOI, gromadzona w zasobach pamięci społecznej (np. w bazach danych).



Rysunek 1. Model: komunikacja naukowa w procesach OW i OI  
 Źródło: opracowanie własne

Na rysunku 1 przedstawione są procesy związane z organizacją danych, informacji oraz wiedzy w trakcie realizacji procesu naukowego, w tym głównie publikowania, jako części komunikacji naukowej. W modelu odrębnie potraktowane zostały procesy związane z tworzeniem, przekazywaniem i odbiorem – odpowiednio – informacji i danych oraz metainformacji i metadanych. Przerywana strzałka symbolizuje płynność granicy między informacją i danymi w niektórych okolicznościach, szczególnie w naukach społecznych, o czym była już mowa. Nie oznacza to jednak, że to, co dla kogoś jest informacją, ktoś inny może traktować jak dane. Takie błędne twierdzenia mogą wynikać z faktu, że zarówno dane, jak i informacja wpływają na stan wiedzy odbiorcy, wobec czego muszą mieć dla niego znaczenie. Przystawianie danych i informacji naukowej jest więc procesem, który Karl Popper nazywał „przywołaniem do świadomości” lub przekształcaniem „nieuświadomianego w uświadomiane”, przejściem ze świata pierwszego i świata trzeciego do świata drugiego [Popper, Eccles 1985, p. 36]. Zauważmy, że przedstawiony przeze mnie model w pewien sposób koresponduje z modelem trzech światów Poppera: po lewej stronie rysunku 1 znajduje się świat pierwszy, pośrodku świat drugi, a z prawej strony świat trzeci. Przy czym dla poprawnego umieszczenia informacji w takim modelu należy uwzględnić wieloaspektowość artefaktów informacyjnych: książka jako przedmiot należy do świata pierwszego, a jako treść do świata trzeciego; podobnie biblioteka jako gmach należy do świata pierwszego, a jako organizacja informacji do świata trzeciego.

Z tego, co powyżej, wynika również, że mówienie o danych lub informacji w informatologii ma sens dopiero wtedy, gdy istnieje ktoś, kto jest świadomy procesów związanych z ich przepływem. Co prawda znaczny wpływ na sposób działania każdej jednostki ma nieświadomy przepływ danych oraz ich przetwarzanie [Krzysztofek 2014, s. 21], jednak dane i informacje pochodzące z różnych źródeł są centralnie integrowane, dzięki czemu tworzone są nieustannie zmieniające się i wzajemnie oddziałujące wzorce oceny sytuacji i reagowania przez jednostkę na złożone czynniki wewnętrzne i zewnętrzne. Świadomość w tych warunkach może być uważana za jedno ze źródeł informacji. Traktowana jest wówczas jako wielorako uwarunkowany proces konstytuowania przez podmiot wiedzy o tym, co i jak się wydarza w dostępnej mu rzeczywistości, dający mu możliwość wpływania na nią.

W dalszych partiach tekstu szczegółowo omówię zjawiska przedstawione w prawej części rysunku 1, dotyczące konwersji typu wiedza / informacja / wiedza, a więc procesów związanych z tworzeniem, publikowaniem i odbiorem tekstu. W modelu uwzględnione zostały liczne transfery danych, informacji oraz metainformacji / metadanych. Zauważmy przy tym, że podobnie jak w przypadku SOW i SOI, termin metainformacja / metadane jest tu użyty w znaczeniu innym niż zazwyczaj przyjmowane w literaturze. Metainformacja to informacja przetworzona za pomocą SOI, dzięki której m.in. istnieje możliwość odszukania informacji pierwotnej w zbiorach informacji. Analogicznie dane przetwarzane są za pomocą SOD, co prowadzi do tworzenia metadanych. Jednak zarówno tworzenie metainformacji, jak i metadanych wymaga wkładu wiedzy indeksatora. W wyniku jej stosowania powstaje tekst specjalnego rodzaju – paratekst w terminologii Gerarda Genette'a [Genette 1991, p. 262], służący komunikowaniu przetworzonej wiedzy indeksatora, a więc, jak każdy tekst, zawierający informację. Jest to informacja wyższego poziomu (pochodna), służąca opisowi danych / informacji niższego poziomu, dlatego też nazywam ją metainformacją / metadanymi. Metadane w takim sensie także stanowią metainformację: jest to informacja o danych. Uwzględniłem także rolę kontekstu w realizacji tych transferów.

Wiedza naukowa stanowi centrum proponowanego modelu. Transfery informacji i danych służą zmianie stanu wiedzy indywidualnej naukowców (głównie jej przyrostowi) i jej dystrybucji w postaci skonwertowanej do informacji, co powoduje powstawanie licznych sprzężeń zwrotnych w modelu. Nie jest to więc model hierarchiczny, jak większość dotychczas przedstawianych w literaturze.

W dalszej części artykułu bardziej szczegółowo opisane będą części modelu dotyczące wiedzy i jej organizacji. Ogólnie wspomniane zostaną też procesy organizacji informacji i jej systemy.



## 2. Umysłowe struktury wiedzy

Mózg, jego budowa i funkcje zostały dotychczas w różnym stopniu poznane przez naukowców, bowiem najłatwiej jest badać obszary, które otrzymują bodźce wejściowe z zewnątrz organizmu (pierwszorzędowe pola sensoryczne przetwarzające obrazy, zapachy, dźwięki, smaki, dotyk). W tym przypadku dane wejściowe są znane i stosunkowo łatwe do opisanie. Nowe techniki badawcze, takie jak czynnościowy rezonans magnetyczny (fMRI) i pozytonowa tomografia emisyjna (PET), bardzo ułatwiły badania w tym zakresie, ujawniając wiele funkcji struktur mózgowych i procesy pozostające poza świadomością. Natomiast wszystkie niepierwszorzędowe<sup>6</sup> obszary kory otrzymują dane wejściowe z innych obszarów korowych, a więc z wnętrza mózgu, dlatego nie są tak dobrze poznane. W tym zakresie wciąż uczeni posługują się raczej teoriami niż wiedzą szczegółową.

### 2.1. Struktury pierwszorzędowe

Ludzki mózg zmienia się w ciągu całego życia w wyniku uczenia się, czyli nabywania doświadczeń<sup>7</sup> [Kaczmarek 2013, s. 67]. Narząd ten funkcjonuje w sposób zapewniający nieprzerwany odbiór danych i informacji zmysłowych. Uczenie się, towarzyszące każdej interakcji z otoczeniem, powoduje zmiany w sile połączeń synaptycznych w sieci neuronów [Spitzer 2012, p. 53], co sprawia, że każdy mózg jest indywidualnym tworem o niepowtarzalnej strukturze połączeń. Ludzki mózg jest mniej „upakowany” neuronami niż mózgi innych gatunków, co oznacza, że pomiędzy komórkami znajduje się więcej miejsca. Dzięki temu możliwe jest tworzenie znacznie większej liczby połączeń pomiędzy neuronami, co stanowi o znacznym potencjale rozwojowym tej struktury [DeFilippe 2011, p. 9].

Uczenie się jest ściśle związane z zapamiętywaniem i przypominaniem, a więc z używaniem pamięci. Jest to zupełnie inny rodzaj pamięci niż np. pamięć komputerowa. Każde przypominanie (powtórzenie / wzmocnienie bodźca) powoduje, że engramy<sup>8</sup> nie są już identyczne z tymi, które pozostały podczas poprzedniego procesu uczenia się i zapamiętywania [Kron, Sofos 2008, p. 88]. Każde

---

<sup>6</sup> Ze względu na pełnione funkcje pola korowe dzielą się na: pierwszorzędowe (odbior informacji i danych zewnętrznych), drugorzędowe (analiza i synteza, czyli rozpoznawanie sygnałów dochodzących z pól pierwszorzędowych, spostrzeganie) i trzeciorzędowe (integracja informacji z niższych pól, precyzja ruchów, orientacja przestrzenna, rozumienie złożonych konstrukcji językowych i dobór właściwych słów).

<sup>7</sup> Jan Kaczmarek w cytowanej publikacji oddziela wiedzę (intencjonalny składnik świadomego umysłu) od składnika subiektywnego, afektywnego, wynikającego z uczuć i emocji. Wydaje się jednak, że oba te elementy łącznie składają się na wiedzę, gdyż oba wynikają z funkcji mózgu, np. stres (czynnik afektywny) ma silny wpływ na jądro migdałowe [Spitzer 2012, p. 125].

<sup>8</sup> Engramy to ślady pamięciowe, umysłowe odbicie doświadczeń, przeżyć, które odciskają się w pamięci, stając się jej elementem.

nowe przypominanie powoduje tworzenie nowych śladów zmieniających engramy, wskutek czego powstają nowe struktury neuronowe. Również kontekst jest uaktualniany, co powoduje zmianę stanu wiedzy. Z drugiej strony zanik silnych i powtarzalnych bodźców powoduje eliminację połączeń neuronalnych – czyli zapomnienie. Wynika stąd, że w pamięci nie da się na zawsze zachować żadnego obrazu w niezmienionym kształcie.

Mózg jest przystosowany do budowania reguł na podstawie wielkiej liczby pojedynczych doświadczeń. Uczenie się polega na przetwarzaniu tych doświadczeń i budowaniu na ich podstawie reguł przydatnych w przyszłym postępowaniu z podobnymi przypadkami. Neurony kory wraz z połączeniami synaptycznymi o określonej sile tworzą reprezentacje docierających do nich sygnałów wejściowych, czyli wewnętrzne odwzorowania cech i struktur otoczenia. W mózgu reprezentowane są zarówno pojedyncze wydarzenia, jak i ogólne zasady, organizm, którego mózg jest częścią, oraz rzeczywistość zewnętrzna. Reprezentacje te przybierają formę strukturalnej mapy<sup>9</sup>, co oznacza ich odpowiednie uporządkowanie: podobne sygnały reprezentowane są w obszarach położonych blisko siebie, a im częściej pojawia się sygnał wejściowy, tym więcej miejsca (neuronów) przeznaczają się na utrwalenie jego reprezentacji, zgodnie z zasadą podobieństwa i częstości. Mapy powstają samoistnie, wskutek przetwarzania wzorców określonego rodzaju przez sieci neuronalne. Stymulacje i zadania w zakresie poznania funkcjonują jako perturbacje w istniejącej dynamice sieci. W wyniku spontanicznej aktywności neuronalnej wzorce połączeń funkcjonalnych są rekonfigurowane w odpowiedzi na zmiany sygnałów sensorycznych lub wymogów środowiskowych.

Sieci te mają prostą budowę, a ich praca podlega trzem regułom:

1. Synapsy są plastyczne, dzięki czemu zachodzi zjawisko neuroplastyczności, polegające na ciągłym tworzeniu i likwidowaniu połączeń między neuronami.
2. W mózgu istnieje wielka liczba wewnętrznych połączeń.
3. Sposób łączenia neuronów powoduje, że podczas pobudzania jednej komórki pobudzone zostają także komórki sąsiednie, natomiast bardziej odległe są hamowane [Spitzer 2012, p. 85].

Według Olafa Sporns wszystkie procesy poznawcze odbywają się w obrębie sieci anatomicznych, a topologia tych sieci nakłada silne ograniczenia na architektury poznawcze [Sporns 2011, p. 205]. Małe światy<sup>10</sup>, będące składnikiem

---

<sup>9</sup> Mapowanie kognitywne jest procesem, za pomocą którego osoba pozyskuje, gromadzi i koduje informacje o świecie, by później móc się do nich odwoływać. W efekcie mapowania powstają modele i schematy mentalne. Różnice pomiędzy tymi ostatnimi nie są jasne; zwykle autorzy piszą o statyczności i sztywności schematów (dynamice i elastyczności modeli) lub nadrzędności hierarchicznej schematów nad modelami [Jones et al. 2011].

<sup>10</sup> W sieciach „małych światów” powiązane siecią jednostki, oprócz wielu połączeń z najbliższymi sąsiadami w sieci, posiadają również mniejszą liczbę połączeń dalekiego zasięgu. Dzięki istnie-

wielkich uporządkowanych sieci neuronalnych, mają strukturę modułową<sup>11</sup> i hierarchiczną<sup>12</sup> zarazem (tzw. hierarchiczna modułowość), dzięki czemu w naturalny sposób wspomagają segregację i integrację funkcjonalną w obrębie mózgu. Funkcja poznawcza mózgu nie sprowadza się jednak tylko do działania tych wyspecjalizowanych modułów, jako że aktywność pojedynczych modułów musi być koordynowana w celu zapewnienia spójności systemu i integracji informacji. Wiele czynności poznawczych może być przedstawionych jako dynamiczna integracja rozproszonych (segregowanych) zasobów.

Realizacja funkcji mózgu odbywa się w wyspecjalizowanych obszarach, w połączeniu z innymi ośrodkami. Połączenia ośrodków mózgowych często przedstawiane są na przykładzie posługiwania się językiem, gdyż procesy te są jednymi z najbardziej złożonych z punktu widzenia percepcji i motoryki. Tak więc za artykulację mowy odpowiada ośrodek Broki, umiejscowiony w lewym płacie czołowym. Jednak możliwość przeklinania i śpiewania obsługiwana jest przez podobny obszar w prawej półkuli. W lewym płacie skroniowym znajduje się natomiast ośrodek Wernickego, służący rozumieniu mowy. Ośrodek ten odpowiada jednak tylko za pojmowanie słów i czasowych sekwencji mowy, natomiast melodia języka i emocjonalne zabarwienie wypowiedzi oraz jej kontekst odwzorowywane są w prawej półkuli<sup>13</sup> [Joseph 2011, p. 6]. Wskazanie na te dwa ośrodki byłoby jednak dużym uproszczeniem. Jak pisze Manfred Spitzer, podczas czytania informacje wejściowe, wstępnie przetworzone w obszarach wzrokowych, przetwarzane są dalej w obszarze wzrokowej formy słów (ang. *visual word form area*, VWFA), odpowiadającym za identyfikację graficznego kształtu wyrazów i, ze względu na podobieństwo funkcji, znajdującym się obok obszaru rozpoznawania twarzy. Dalej impulsy biegną w kierunku przedniej części płata skroniowego (długość pokonywanej trasy zależy od rodzaju przetwarzanego słowa, najdalej przedostają

---

niu tych połączeń dwa dowolne obiekty w sieci powiązane są krótkimi łańcuchami relacji – „szścioma stopniami oddzielenia”. Sieci te mogą być konstruowane w różny sposób. Obecnie uważa się, że w mózgu największą rolę odgrywają grupy obszarów mózgu, czyli węzły podstawowe, oraz ważne węzły centralne (huby), łączące te grupy z sobą. Taka architektura połączeń korowych jest charakterystyczna dla wszystkich ssaków.

<sup>11</sup> Modułem nazywane jest skupisko węzłów podstawowych, w którym występuje więcej dwustronnych połączeń wewnętrznych niż zewnętrznych. Często neurony tworzące moduł są powiązane funkcjonalnie.

<sup>12</sup> W każdej sieci neuronalnej funkcjonuje także swego rodzaju zagnieżdżona hierarchia, począwszy od bardzo powierzchownego podziału na dwie półkule mózgowe, po grupowanie bardziej szczegółowe, np. na funkcjonalne systemy mózgu (np. kora wzrokowa, ruchowa, słuchowa, czuciowa), pojedyncze obszary wyróżniane pod względem anatomicznym, jądra istoty szarej lub według kolumnowej organizacji komórek kory.

<sup>13</sup> Osoba z deficytem w tej części mózgu nie jest w stanie zrozumieć, czy np. pytanie zadane w barze piwnym: „wyjdziemy na zewnątrz?” oznacza propozycję odetchnięcia świeżym powietrzem, wypalenia papierosa, porozmawiania bez świadków, czy bójki.

się impulsy odwzorowujące imiona), by trafić do ośrodka Wernickego, który skomunikowany jest szybkimi połączeniami z ośrodkiem Broki. Stamtąd impulsy przechodzą do wspomagającej kory ruchowej i kory przedruchowej oraz ruchowej, gdzie wybierane są, aktywowane i dostrajane adekwatne do nich programy ruchowe. W efekcie można głośno odczytać tekst brzmiący jak spontaniczna wypowiedź [Spitzer 2012, p. 179].

Jak widać z tego przykładu, integracja zasobów umysłowych wymaga dynamicznej koordynacji (synchronizacji, spójności, sprzężenia liniowego i nieliniowego) oraz konwergencji. Powtarzalność połączeń pozwala na powstawanie wzorców połączeń funkcjonalnych, stosowanych w całym systemie, podczas gdy wysoko scentralizowane węzły sieci (huby) odgrywają wyspecjalizowaną rolę w zakresie koordynacji przepływu informacji. Scentralizowane węzły sieci neuronów wykorzystywane są w procesach tworzenia asocjacji, przetwarzania transmodalnego oraz dynamicznej konwergencji.

Z perspektywy sposobu działania sieci neuronalnej poznanie jest tworzeniem szczególnego rodzaju wzorców, wzajemnym oddziaływaniem funkcjonalnej segregacji i integracji oraz ciągłym powstawaniem dynamicznych struktur, które są formowane przez połączenia neuronalne i ulegają subtelnym modyfikacjom zarówno pod wpływem działania bodźców zewnętrznych, jak i wskutek endogennych zmian w stanie samych połączeń. Zdolności poznawcze, natura danych i informacji, które mogą być zgromadzone i przetworzone, są determinowane przez architekturę sieci mózgu. Procesy poznawcze są wynikiem zmiennych w czasie, dynamicznych i rozległych sekwencji zdarzeń, powodujących zmiany w strukturze powiązań neuronalnych. Różnorodność poznawcza, niemal nieskończona różnorodność stanów mentalnych i osobistych doświadczeń, jest odbiciem złożoności struktur mózgu. Perspektywa sieciowa umożliwia budowę ram teoretycznych pozwalających na uznanie procesów poznawczych za bezpośrednio wynikające z biologii mózgu.

## **2.2. Obszary wyższego rzędu**

Regularne powtarzanie się bodźców wejściowych powoduje tworzenie reprezentacji, czyli mapy każdego bodźca, z użyciem wspomnianych reguł. Mapy te są już dość dobrze poznane dla niższych obszarów korowych, w których powstają reprezentacje bodźców zewnętrznych działających na zmysły. Jako że cała kora mózgowa działa jednak w oparciu o trzy wcześniej wymienione reguły, uprawnione jest założenie, że gorzej poznane obszary korowe wyższego rzędu także zawierają reprezentacje w formie strukturalnych odwzorowań – map, które można nazwać „schematami poznawczymi”. Podczas każdej czynności umysłowej, w tym pisania i czytania, na różne sposoby aktywowane są dziesiątki różnych map (struktur) korowych.

Funkcjonowanie wyższych poziomów struktur mózgu opisywane jest za pomocą modeli uwzględniających problem mapowania. Według Richarda Andersona wiedza jest włączana w abstrakcyjne struktury zwane schematami poznawczymi<sup>14</sup> (łac. *schemata*) [Anderson 1977, p. 2]. Zawierają one wiedzę ogólną, czyli reprezentują to, co jest ogólnie uważane za prawdziwe dla danej klasy rzeczy, wydarzeń lub sytuacji, których konkretne wystąpienia zwykle różnią się w szczegółach. Ustrukturyzowanie służy reprezentacji relacji zachodzących pomiędzy częściami schematu [Anderson, Pearson 1984, p. 10]. Zwykle uważa się, że schematy budowane są hierarchicznie, składają się z „przegródek” przeznaczonych na części składowe kodowanego obiektu. Kodowanie obiektu polega na wypełnianiu przegródek w schemacie konkretnymi cechami obiektu. Częścią uschematyzowanej wiedzy jest określenie ograniczeń dotyczących możliwości wypełniania poszczególnych przegródek. Ograniczenia te muszą być dość elastyczne, tak aby możliwe było reprezentowanie różnych wariantów obiektu lub wydarzenia. Podobnie schematy rozumiane są w gramatyce kognitywnej, szczególnie na poziomie semantycznym [Taylor 2007, p. 145-146].

Schematy są ustrukturyzowaną wiedzą, gdyż wskazują na typowe relacje pomiędzy jej składnikami. Poza tym funkcjonują na różnych poziomach abstrakcji i mogą zawierać się jedne w drugich [Rumelhart, Ortony 1977]. Zakłada się, że możliwe jest zastosowanie głównego schematu bez uwzględniania wiedzy zawartej w schematach podporządkowanych. Z drugiej strony, gdy sytuacja tego wymaga, pobrane może być znaczenie wszystkich podschematów, co umożliwia pogłębienie interpretacji.

Tworzenie schematów jest związane z umysłową kategoryzacją świata. W swojej teorii prototypu Eleanor Rosch zakłada, że kategorie powstają w wyniku tworzenia pojęć, są to więc kategorie poznawcze. Cechy kodowanego obiektu określa się podczas procesu poznawczego przez dokonywanie porównań, poszukiwanie podobieństw i różnic, w wyniku czego uzyskiwany jest prototyp – najbardziej wyrazisty okaz danej kategorii (schematu poznawczego), wokół którego skupiają się podobne okazy. Funkcje prototypu pełni zestaw cech charakteryzujących właściwe, „dobre” elementy danego schematu [Tabakowska 1995, s. 40-42].

W świetle teorii schematów praca z tekstem polega na tworzeniu / wydobywaniu znaczeń tekstu podczas interakcji pomiędzy cechami komunikatu, wiedzą posiadaną przez autora / czytelnika a analizą kontekstu. Autorzy umieszczają w tekście językowe wskazówki pozwalające dotrzeć czytelnikowi do ważnych miejsc. Schematy aktywowane podczas lektury tekstu zawierają schematy pod-

---

<sup>14</sup> Termin ten do psychologii wprowadzili Jean Piaget [Piaget 1926] i Frederic Bartlett, według którego schemat oznacza aktywną organizację przeszłych reakcji lub doświadczeń [Bartlett 1932, p. 201].

rzędne, tworzące układy hierarchiczne. Miejsce w tej hierarchii uzależnione jest od ważności podschematu. Oznacza to, że ważność elementu tekstu będzie zmieniać się w zależności od roli, jaką odgrywa on w wykorzystanym schemacie.

Don Tucker i Phan Luu opisują model kognitywny ATA (ang. *Attention-to-Action*) autorstwa Dona Normana i Tima Shallice'a, który obejmuje kilka złożonych podsystemów regulujących zachowanie [Tucker, Luu 2012, p. 92]. Pierwszym z nich jest system planowania rywalizacji (ang. *contention scheduling*), regulujący zachowania przez wybór i uruchomienie kolejnych „schematów”. Działają one gwałtownie, automatycznie i bez udziału świadomości, przez sieć połączeń hamujących aktywację schematów, których wyjścia mogą ze sobą kolidować. Za sprawą sieci schemat działania o wysokim poziomie aktywacji może powstrzymać aktywację innych, potencjalnie konfliktowych schematów.

Schemat jest wewnętrznym modelem świata zewnętrznego, zbudowanym na podstawie wcześniejszych doświadczeń. Każdy schemat posiada bieżący wskaźnik poziomu aktywacji, który może być podwyższony przez specjalną percepcyjną stymulację aktywacji lub przez impulsy wychodzące z innych schematów powiązanych. Za pomocą tego mechanizmu można ustalać także precyzyjną kolejność aktywacji schematów w czasie. Gdy poziom aktywacji przekracza ustalony próg, schemat może pokierować zachowaniem natychmiastowo i stereotypowo w celu wykonania określonego zadania. Jest on wybierany przez system planowania rywalizacji, który blokuje lokalnie schematy konkurujące. Po dokonaniu wyboru działania wykonywane są automatycznie, gdyż są częścią schematu. Wiele jednoczesnych stymulacji może powodować aktywację dużej liczby schematów jednocześnie, co w razie pojawienia się wzajemnie wykluczających się odpowiedzi powoduje podatność na błędy [Rubinstein, Meyer, Evans 2001, p. 764].

W przypadku zaistnienia konfliktu pomiędzy wybranym schematem a wymaganiami środowiskowymi działa umieszczony w modelu system nadzorczy (ang. *supervisory attentional system* – SAS), którego interwencja powoduje wybór lepszego schematu. System ten ma dostęp do całościowej reprezentacji środowiska oraz do celów podmiotu. Działa powoli, elastycznie i pod kontrolą świadomości. Określone zostały także sytuacje, w których system ten może być uruchomiony; są to: planowanie i podejmowanie decyzji, poprawianie błędów, reakcja na nowe sytuacje oraz na niebezpieczeństwo, a także przewyciężanie reakcji nawykowych. W sytuacjach tych należy podejmować decyzje o dalszym postępowaniu, gdyż dotychczasowe działania (wyczone lub automatyczne) przestają być odpowiednie [Norman, Shallice 1986, p. 7].

Podjęcie wysiłku służącego poznaniu kreatywnemu zawsze powoduje tworzenie nowych lub rozbudowę istniejących umysłowych struktur konceptualnych.

Skutkiem jest wprowadzenie nowych, lub powiązanie dotąd niezależnych, cech. Gdy sieć przetwarzania semantycznego mózgu zaangażowana jest w rozszerzenie konceptualne<sup>15</sup>, pracuje bardzo intensywnie, szczególnie w swoich wyżej zorganizowanych obszarach, które biorą udział w selekcji leksykalnej, kontrolowanym wydobywaniu z pamięci i procesach integracji informacji. Ludzki system poznawczy stale stara się wykorzystywać posiadaną wiedzę („standardową”) w przetwarzaniu nowej informacji. Podczas procesu refleksji twórczej konieczne jest więc uwolnienie umysłu od ograniczającego wpływu wiedzy „standardowej”. Wcześniej pozyskana informacja, choć skądinąd potrzebna i relewantna, jest jednak czynnikiem zakłócającym, ponieważ zagraża dalszemu postępowi czynności twórczych przez ograniczanie możliwości tworzenia nowych, oryginalnych idei. Dlatego sieć sterowania procesami poznawczymi mózgu stara się usunąć te zakłócenia z umysłu. Można tego dokonać na dwa sposoby – przez powstrzymanie lub przez ignorowanie pierwotnie pozyskanej relewantnej informacji. Obie te czynności są bardzo trudne, gdyż mózg jest stworzony do skutecznego i biegłego wykonywania działań celowych. Podczas standardowej, codziennej aktywności umysł ludzki funkcjonuje w kontekście normatywnym, gdzie zakłócenia, które trzeba przezwyciężać, mogą być bezbłędnie rozpoznawane i łatwo pomijane jako nieistotne lub nierелеwantne w stosunku do bieżących zadań. Jednak podczas tworzenia kreatywnych idei informacja zakłócająca może być nadzwyczaj pertynentna do wykonywanego zadania, toteż nie może być pomijana w podobny sposób. W takich warunkach nierównowagi w układzie czołowo-prążkowym mózgu powstają szczególnie korzystne warunki do twórczego poznania, prawdopodobnie dzięki brakowi kognitywnych zahamowań oraz wzrostowi możliwości abstrahowania, co ułatwia ignorowanie istotnych czynników rozpraszających [Abraham 2014, p. 4].

Wymienione rodzaje schematów łączą się blisko z funkcjami schematów, przedstawionymi przez Jerzego Trzebińskiego [Trzebiński 2001, s. 91]. Są to: funkcja reprezentacji elementu rzeczywistości oraz funkcje reguł przetwarzania danych i informacji o tej rzeczywistości. Ta pierwsza dotyczy stałych cech modelowanego w umyśle obiektu, pozostałe stanowią zespół reguł operowania poznawczego na danych i informacjach, zgromadzonych w ramach schematu lub asymilowanych przez schemat z zewnątrz.

Jerzy Trzebiński zajmował się schematami, w których dominują relacje przyczynowo-skutkowe lub celowościowe, nazywając je schematami narracyjnymi [Trzebiński 2001, s. 93]. Narracja jest sposobem rozumienia świata przez ludzi,

---

<sup>15</sup> Rozszerzenie konceptualne (ang. *conceptual expansion*) polega na poszerzaniu ram istniejących struktur konceptualnych.

determinowanym przez strukturę ich wiedzy. Ludzie rozumieją i zapamiętują rzeczywistość, w szczególności rzeczywistość faktów zaobserwowanych w dłuższym okresie, w wyniku interpretacji tej rzeczywistości według reguł zawartych w schematach narracyjnych. Narracje są specyficznym rodzajem rozumienia świata; strukturalizują doświadczenie człowieka w formie opowiadania, ujmując je w kategoriach ludzkich intencji oraz problemów powstających w trakcie realizacji tych intencji. W tym sensie umysł interpretuje postrzegane zdarzenia jako powiązane historie, w których rozumieniu zostaje odzwierciedlona struktura narracji [Szalkowska, Francuz 2003, s. 90]. W konsekwencji w trakcie narracyjnej interpretacji danego faktu człowiek łatwiej przetwarza informacje zindywidualizowane i kontekstowe [Trzebiński 2002].

Narracje ułatwiają rozumienie zdarzeń rozciągniętych w czasie, więc również proces rozumienia jest rozciągnięty w czasie. Według Trzebińskiego można go porównać do procesu czytania, a bardziej nawet – pisania tekstu, który cechuje się tym, że treść dotychczas napisanych / przeczytanych części przechowywana jest w pamięci operacyjnej i służy rozumieniu bieżąco pisanych / czytanych zdań. Do czynności tych można wracać w dowolnym czasie, łatwo aktualizując treści i narrację przygotowującą do korzystania z kolejnych części. W związku z tym Trzebiński porównuje schemat narracyjny do wiedzy o gatunkach tekstów, funkcjonującej w określonej kulturze. Posiadanie takiej wiedzy umożliwia zrozumienie czytanych tekstów powstałych w określonym kręgu kulturowym. Wynika stąd, że potrzebna jest bardzo specjalistyczna wiedza, aby dobrze zrozumieć teksty naukowe funkcjonujące w dowolnej dziedzinie. Wiedzę tę trudno zdobyć inaczej, niż obcując z tekstami.

Schemat narracyjny służy przede wszystkim tworzeniu narracji (opowieści) przez jednostkę, co odzwierciedla fakt, że wiedza jest zjawiskiem indywidualnym. Jednak równie ważną funkcją schematów narracyjnych jest rozumienie wypowiedzi i zachowań innych ludzi, a przez to także własnej kultury. Oznacza to, że narracje, w ramach których podmiot rozumie świat, są najczęściej tworzone społecznie [Trzebiński 2001, s. 100]. W trakcie międzyludzkich negocjacji uzgadniane są wspólne interpretacje napływających informacji i danych, które stanowią podstawę tworzenia treści narracji.

Narracje przedstawiane w formie tekstów naukowych są zazwyczaj autonarracjami, czyli takimi narracjami, w których podmiot jest ich ważnym bohaterem. Zazwyczaj autonarracje są opowieściami zarówno o podmiocie, jak i o jego partnerach związanych jednym wątkiem narracyjnym. Przykładowo w artykułach naukowych temu celowi służy przedstawienie stanu badań. Autonarracje stanowią poznawczy kontekst działań jednostki. Wpływają na jej zachowania, kształtując rozumienie własnych celów, planów oraz otoczenia, w którym te cele i plany



się realizuje. Podmiot konstruuje autonarracje, by nadać sens ważnym dla siebie zdarzeniom i swojej roli / pozycji w tych zdarzeniach.

### 2.3. Psycholingwistyczne teorie produkcji i rozumienia mowy

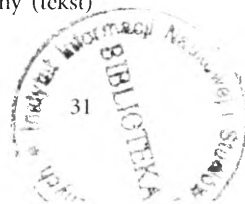
W komunikacji werbalnej<sup>16</sup> zazwyczaj stosowane są formy tekstowe – oralne i pisemne<sup>17</sup> – które, w obu przypadkach, mogą korzystać z technologii elektronicznej. Trzeba jednak zaznaczyć, że bez względu na właściwości medium (elektroniczne lub analogowe) opisywana wcześniej konwersja wiedza / informacja / wiedza dokonuje się zawsze za pomocą środków językowych.

Wiele modeli przetwarzania struktur wiedzy na język mowy (oralny) wypracowano na gruncie psycholingwistyki [Nowak 2011, s. 357]. Opisują one produkcję (ekspresję) i percepcję (recepcję) mowy i tekstu<sup>18</sup>. Modele produkcji mowy budowane są na podstawie analizy błędów językowych i zaburzeń płynności mówienia. Obejmują procesy mentalne i neuronalne zachodzące podczas tworzenia pojedynczych wypowiedzi oraz całość procesu komunikacji, w tym analizę dyskursu. Powstają z inspiracji lingwistycznych (np. modele Victorii Fromkin oraz Merrilla Garretta) lub psychologicznych (np. modele Kathryn Bock i Willema Levelta oraz Gary'ego Della). Modele percepcji mowy powstają w wyniku przeprowadzanych eksperymentów. Wyróżnia się w nich podejście wstępujące, autonomiczne (etapy: akustyczny, fonologiczny i leksykalny) lub zstępujące, interakcyjne (poziomy: leksykalny i poznawczy). Modelując mowę i jej odbiór, poszukuje się jednostek i etapów ich przetwarzania. Powstają modele mentalne, nadrzędne reprezentacje kognitywne, które obejmują wiedzę niezależną od tekstu i mogą być zmieniane na podstawie każdej informacji językowej. Wprowadzane jest też pojęcie sytuacji mentalnych, uwzględniających wszystkie elementy sytuacji (czasu, miejsca, relacji, intencji, uczestników) [Mazurkiewicz-Sokołowska 2010, s. 132]. Interesujący wydaje się nadal rozwijany model TRACE, którego podstawą są analogie sieci neuronalnych [McClelland, Elman 1986, p. 8]. Informacja przetwarzana jest w węzłach (jednostkach przetwarzających), które są ze sobą powiązane gęstą siecią połączeń. Każdy węzeł może osiągać dwa stany (poziomy): spoczynkowy i aktywacji oraz charakteryzuje się określonym progiem pobudzenia.

<sup>16</sup> Można pozwolić sobie na pominięcie niewerbalnych, pozajęzykowych elementów komunikacji, pomimo ich istotnej roli wynikającej z wyrażania afektywnych elementów komunikacji [Kaczmarek 2013, s. 67].

<sup>17</sup> Michał Bachtin nazywa formy mówione pierwotnymi (prostymi), wykształconymi w bezpośrednim językowym porozumiewaniu się. Formy wtórne (złożone) – literackie i naukowe, głównie pisane, powstają w warunkach skomplikowanego, rozwiniętego i zorganizowanego porozumiewania się w kulturze [Bachtin 1986, s. 350].

<sup>18</sup> Ida Kurcz proponuje odpowiednio określenia: dyskurs mówiony i dyskurs pisany (tekst) [Kurcz 2000, s. 148].



Rozumienie, definiowane jako czynny proces konstruowania reprezentacji, polegający na interpretacji odbieranych danych / informacji zgodnie z systemem posiadanej wiedzy i na ich włączaniu w ów istniejący już system [Kurcz 1987, s. 308], może także dotyczyć konstruowania reprezentacji struktury sądów zawartych w tekście. Można przyjąć, że struktura zbudowanej reprezentacji odzwierciedla określoną strukturę tekstu (por. [Szalkowska, Francuz 2003, s. 81]), na przykład strukturę artykułu naukowego. Takie podejście reprezentowali m.in. Walter Kintsch i Teun van Dijk, twórcy autonomicznej teorii cyklicznego przetwarzania [Kintsch, Dijk 1978, p. 365]. Autorzy ci przyjmują, że makrostruktura stanowiąca semantyczną strukturę tekstu jest odkrywana w wyniku interpretacji zdań logicznych (mikrostruktury). W stosunku do zdań logicznych jednostka stosuje różne operatory: dane z pamięci sensorycznej, wiedzę ogólną i osobistą zawartą w pamięci długotrwałej, wiedzę leksykalną, a także emocje, pragnienia, zainteresowania i przekonania. Wykorzystanie tych modalnych operatorów sprawia, że interpretacja tekstu jest pochodną zawartych w nich danych. Rozumienie większych partii tekstu zależy również od interpretacji mniejszych fragmentów wchodzących w ich skład. Dla rozumienia tekstu istotna jest więc jego spójność, bowiem interpretacja następuje linearnie przez ujmowanie relacji między kolejnymi zdaniami logicznymi. Gdy weryfikacja bazy tekstu jest niewystarczająca dla zbudowania spójności, włączana jest wiedza o świecie i dokonywana jest inferencja<sup>19</sup>. Rezultatem procesu interpretacji jest skonstruowanie makrostruktury tekstu.

W psychologii kognitywnej pojawiło się także przekonanie, że komunikacja międzyludzka jest produktem ubocznym ludzkich zdolności do reprezentowania w umyśle przekonań, pragnień, intencji, jakie żywią inne osoby [Gut 2010, s. 322]. Język jest wtedy wtórny do inferencji, czyli w tym przypadku: wnioskowania o intencjach komunikacyjnych innych osób. Intencje te wyznaczają znaczenie komunikatu. Dzięki językowi wzrasta jedynie efektywność aktu komunikacji.

Badania psycholingwistyczne wydają się wartościowe, gdyż tworzą ciekawą perspektywę poznawczą – pozwalają na zlokalizowanie źródła zakłóceń w akcie porozumiewania się. Psycholingwistyka stanowi źródło inspiracji i punkt wyjścia dla wszelkich rozważań o języku. Mimo że wymienione modele percepcji wydają się niemożliwe do zweryfikowania w praktyce, stanowią szkielet procesu percypowania mowy [Król 2006, s. 61].

---

<sup>19</sup> Inferencja (wnioskowanie psycholingwistyczne) polega na wyciąganiu wniosków na podstawie całości wiedzy odbiorcy komunikatu. Inferencja różni się od wnioskowania logicznego: jest znacznie bardziej zależna od kontekstu, w jakim jest realizowana [Kliś 2002, s. 103].

### 3. Systemy organizacji wiedzy

Dotychczas omówione zostały procesy mentalne związane z powstawaniem wiedzy w ogóle, w tym wiedzy naukowej, która zgodnie z teorią genploracji [Finke, Ward, Smith 1992] nie stanowi odrębnej jakości, gdyż dla tworzenia wiedzy każdego rodzaju stosowane są te same procedury. Zgodnie z tym modelem różnica pomiędzy poznaniem kreatywnym a tym, które potocznie uznawane jest za nietwórcze, czyli normatywnym, nie polega na wykonywaniu odmiennych czynności umysłowych, ale na tym, że inny jest kontekst przetwarzania informacji<sup>20</sup>. Sytuacja pobudzająca do rozwiązywania problemu twórczego (kontekst) jest bardziej otwarta, wieloznaczna, nieliniowa, abstrakcyjna i nieprzewidywalna w stosunku do tej wymaganej dla poznania normatywnego. W tej części artykułu zajmę się systemami organizacji wiedzy, czyli narzędziami umożliwiającymi konwersję indywidualnej wiedzy w informację (lub odwrotnie), co umożliwia eksternalizację wiedzy i jej komunikowanie. Konwersja ta polega na przekładzie schematów umysłowych nadawcy na odpowiednie struktury tekstu, które uzupełniają schematy umysłowe odbiorcy [Carley, Palmquist 1992, p. 603].

W procesie interpretacji tekstu stosowane są dwa rodzaje schematów [Anderson, Pichert, Shirey 1979, p. 2]. Pierwszy z nich zawiera wiedzę o konwencjach dyskursu i pozwala odbiorcy rozpoznać ogólną organizację tekstu. Istnieją konwencje specjalne, charakterystyczne dla poszczególnych form tekstu, a także konwencje o szerszym zakresie stosowania, wspólne dla większości z nich; można więc mówić o schemacie opowieści, schemacie listu, schemacie artykułu prasowego, schemacie raportu z badań naukowych (w określonej dziedzinie), artykule naukowego i wielu podobnych. Takie schematy, zawierające wiedzę o konwencjach na poziomie dyskursu, są nazywane schematami tekstowymi.

Schematy innego rodzaju, nazywane schematami treści, zawierają wiedzę autora / czytelnika o świecie rzeczywistym i wymyślnym. Posiadane przez nich schematy istotnie wspomagają strukturyzację interpretacji nowych informacji i wnioskowania w tym zakresie. Według Andersona, Picherta i Shireya dla rozumienia czytanego tekstu schematy treści są ważniejsze od schematów tekstowych.

Przekształcenie (konwersja) struktur wiedzy w struktury dokumentów piśmienniczych dokonywana jest z wykorzystaniem konwencjonalnych struktur tekstu, zwanych gatunkami. Pojęcie gatunku (fr. *genre*) definiowane i stosowane było przez badaczy z różnych dyscyplin, takich jak literaturoznawstwo, socjolo-

---

<sup>20</sup> Podobną myśl można znaleźć już u Jeana Piageta, który pisał o rozwoju poznawczym człowieka jako analogicznym do procesu badań naukowych ze względu na stałe tworzenie i modyfikację schematów poznawczych [Nęcka, Orzechowski, Szymura 2012, s. 34].

gia, kognitywistyka, etnografia, kulturoznawstwo zorientowane na badanie kultury popularnej, językoznawstwo, retoryka, w których tworzone były odrębne teorie gatunków. Teoria ta wykorzystywana jest także w informatologii, w zakresie reprezentacji i organizacji wiedzy oraz wyszukiwania informacji [Hajibayova, Elin 2014, p. 391]. Ta multidyscyplinarność wymaga specjalnego, wielokierunkowego podejścia do badań i analizy zjawiska gatunków tekstu, w szczególności tekstu naukowego. W tym celu podejmowane są próby integracji metodologii etnograficznych, socjokognitywnych i innych, dostosowanych do potrzeb analizy tekstu. Różnorodność ta jest przyczyną istnienia wielu definicji odzwierciedlających różne sposoby ujmowania przedmiotu badań [Bhatia 2008, p. 166]. Mi-sha Vaughan i Andrew Dillon podają definicję, która uwzględnia różne punkty widzenia [Vaughan, Dillon 2006, p. 503]. Według nich gatunek to klasa wydarzeń komunikacyjnych, które cechują się wspólnym zestawem konwencji i zasad, służących ułatwianiu interakcji komunikujących się stron (użytkowników) przez tworzenie i obsługę (np. modyfikację) oczekiwań w obrębie społeczności twórców (autorów) i odbiorców (czytelników). Definicja ta wskazuje, że poziom jednolitości konwencji w społeczności użytkowników języka może być badany zarówno z punktu widzenia postrzegania przez użytkowników jej fizycznej manifestacji (tekstu), jak i z punktu widzenia interpretacji jej znaczeń społecznych, historycznych i kulturowych. Jednak oprócz społeczno-kulturowych uwarunkowań stosowania gatunków tekstu na rzeczywistą aktywność językową podmiotu wpływają także indywidualne czynniki kognitywne, w związku z czym w analizie tekstów aktualizujących daną konwencję gatunkową brane są pod uwagę nie tylko cele komunikacyjne całej społeczności dyskursywnej, ale również strategie poznawcze stosowane przez jej członków dla osiągnięcia celów komunikacyjnych [Bhatia 1997, p. 629].

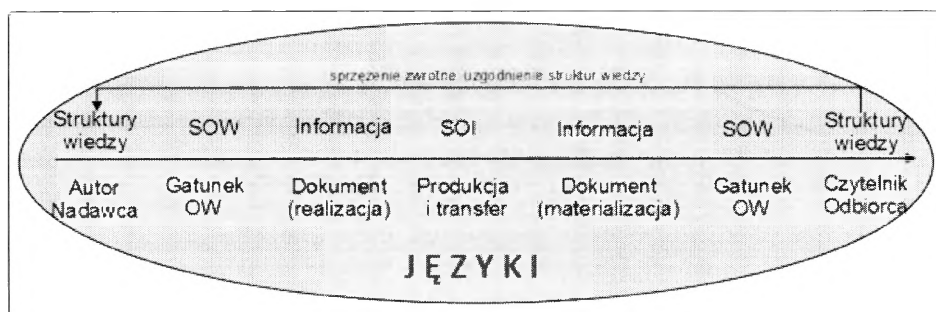
Według Voughan i Dillon gatunki spełniają dwie role. Pierwszą z nich jest „obsługiwanie” dyskursu, polegające na wspomagananiu dwóch stron komunikacji: tworzenia i recepcji tekstów. Każdy akt wymiany informacji powoduje przywoływanie struktur tekstu odzwierciedlających struktury umysłowe wiedzy stosowanej w tym akcie. W umysłach mówiącego (piszącego) i słuchającego (czytającego) istnieje nieświadomiony schemat organizacji wypowiedzi, który ze względu na swój uniwersalny charakter występuje u wszystkich uczestników aktu komunikacji [Wyrwas 2002, s. 22]. Znajomość schematów gatunkowych, będąca częścią wiedzy człowieka, pozwala mu, adekwatnie do różnych sytuacji, tworzyć i interpretować teksty spełniające różne cele komunikacyjne [Duszak 1998, s. 209]. W tym sensie schematy te uczestniczą w tworzeniu zarówno indywidualnych narracji, jak i praktyk komunikacyjnych w skali społecznej [Tonkin 1992, p. 51]. Już Bachtin uważał, że każdy tekst jest odpowiedzią na poprzednie teksty, z któ-

rymi wchodzi w relacje [Bachtin 1986, s. 361]. W tym znaczeniu gatunek tworzony jest przez wszystkie teksty, które powstały i powstają według danego wzorca, jawi się jako dynamiczny konstrukt semiotyczny, dający jednolity pogląd na treść i formę tekstu oraz pozwalający na osadzenie go w określonym kontekście społecznym, z uwzględnieniem ról użytkowników w tej społeczności. W ten sposób tworzona jest jednolita teoria pisania i czytania [Devitt 1993, p. 573], służąca wyjaśnieniu procesu powstawania gatunków rozumianych jako formy aktywności (uzyskiwania doświadczeń) człowieka, która sama jest historycznie formowana przez gatunki. Jednocześnie jest to więc aktywność wiedzotwórcza, gdyż wiedza powstaje w procesach zbierania doświadczeń [Kolb, Kolb 2005, p. 194]. Stosowanie wzorców gatunkowych ułatwia tym samym wybór tekstów, które odbiorca traktuje jako tzw. autorytety kognitywne [Wilson 1983]. Do elementów, według których wyznacza się teksty mające status autorytetów kognitywnych, należą – oprócz wcześniej zdobytej wiedzy i charakterystyk źródła (reputacja, autorstwo) – również cechy obiektu informacyjnego (tytuł, treść, struktura, prezentacja, grafika, funkcjonalność) [Rieh 2002, p. 150], czyli cechy gatunkowe tekstu.

Drugim celem stosowania gatunków jest satysfakcja emocjonalna, co oznacza, że nie są one pozbawione swego rodzaju roli afektywnej [Berger 1992]. Obcowanie z tekstem reprezentującym znany czytelnikowi gatunek budzi poczucie emocjonalnego bezpieczeństwa. Lista emocjonalnych gratyfikacji uzyskiwanych przez użytkowników gatunków, przedstawiona przez Arthura Bergera na przykładzie odbioru przekazów audiowizualnych, jest długa; w zależności od stosowanego gatunku może to być zaspokojenie ciekawości i bycie poinformowanym (wiadomości, teleturnieje, *talk shows*), przywrócenie wiary w sprawiedliwość (programy prawne) lub romantyczną miłość (komedie romantyczne), odczuwanie piękna (koncert symfoniczny) i inne. Gratyfikacje emocjonalne są ważne także w publikowaniu naukowym (np. wzrost prestiżu) [Borgman 2007, p. 72].

Przegląd definicji gatunków tekstu i poglądów naukowych na ich temat, prezentowanych zarówno przez autorów polskich, jak i obcych, przedstawił Marcin Maciejewski [Maciejewski 2009, s. 79 i nast.]. Ten autor także, podobnie jak cytowani poprzednio, zauważył sytuowanie tej koncepcji pomiędzy abstrakcyjną kompetencją ogółu użytkowników tekstu w danej społeczności a kompetencją pojedynczego podmiotu (autora lub czytelnika), posługującego się konkretnymi realizacjami i materializacjami tekstu. Gatunków jest wiele i ulegają one stałym zmianom, również ze względu na zachodzące między nimi relacje. Dynamiczny zbiór wszystkich stosowanych gatunków tekstu obrazuje praktyki komunikacyjne danej społeczności [Maciejewski 2009, s. 87]. Mają one duże znaczenie w procesach osobistych interpretacji świata przez jednostki i uzgadniania tych interpretacji w społecznościach użytkowników, gdyż oba procesy, a szczególnie

uzgadnianie interpretacji, odbywają się za pośrednictwem tekstów. Gatunki są odpowiednikami, odbiciami kognitywnych schematów, porządkującymi formy produkcji językowej, zarówno pisemnej, jak i ustnej. Jak stwierdza cytowany autor, teksty (w tym również hiperteksty) nie występują jednostkowo, jako niepowtarzalne komunikaty, lecz są tworzone przez użytkowników języka wielokrotnie, w podobnych sytuacjach, dla realizacji podobnych funkcji<sup>21</sup>, choć ich postać ulega modyfikacjom ze względu na różnice kontekstowe. Stosowanie gatunków jest sposobem na tworzenie językowych regularności, wykorzystywanych w procesach komunikacyjnych. Powtarzalność taka, jak stwierdziłem w poprzedniej części artykułu, stanowi podstawę do budowy schematów umysłowych.



Rysunek 2. Gatunki tekstu jako SOW

Źródło: opracowanie własne

Na rysunku 2 przedstawione zostało miejsce i rola gatunków tekstu w procesach wiedzy. W górnej części (ponad dolną strzałką) przedstawione są elementy modelu z rysunku 1, natomiast w dolnej umieszczeni zostali aktorzy, artefakty i procesy związane z elementami modelu będącego formalizacją procesu komunikacji. Gatunki stosowane są we wszystkich sytuacjach, w których dokonywana jest konwersja typu wiedza / informacja / wiedza. Należy więc traktować je jako systemy organizacji wiedzy, gdyż są to narzędzia językowe, za pomocą których dokonywana jest konwersja indywidualnej wiedzy nadawcy w informację o formie i strukturze zrozumiałej dla odbiorcy (konwertowalnej na jego wiedzę). Dzięki obustronnej znajomości gatunku możliwa jest komunikacja i na tyle efektywne uzgadnianie struktur wiedzy, aby możliwe było porozumienie między nadawcą i odbiorcą. W szczególności pomiędzy autorem a czytelnikiem zawiązuje się specyficzna relacja (transakcja); autor z jednej strony tworzy tekst, chcąc zrealizo-

<sup>21</sup> Taki sposób funkcjonowania gatunków podkreślała Carolyn Miller, twierdząc że gatunki są typizowanymi działaniami retorycznymi tworzonymi w sytuacjach powtarzalnych [Miller 1984, s. 159]. Gatunki traktowane są więc jako konstrukty dynamiczne (zmienne), o czym była już mowa.

wać własne cele retoryczne, z drugiej jednak strony musi brać pod uwagę reakcję potencjalnego czytelnika i antycypować znaczenia, które ten będzie wyznaczał [Rosenblatt 1988, p. 10]. Nie jest to łatwe zadanie, gdyż pełna zgodność struktur wiedzy nie jest możliwa ani przed aktem komunikacji, ani po nim, chociażby z tego powodu, że rozumienie gatunku także zależy od stanu indywidualnych struktur wiedzy.

Widoczne jest podobieństwo modelu przedstawionego na rysunku 2 do tradycyjnych schematów komunikacji typu nadawca → kodowanie → komunikat → dekodowanie → odbiorca. Różnica polega na tym, że w proponowanym modelu cały proces komunikacji „zanurzony” jest w różnego rodzaju środkach językowych. Stosowany jest język myśli (*mentalese*) [Przywara 2010, s. 100], ewentualnie język mowy, język pisma, jak również narzędzia językowe używane jako SOI. Kodowanie ma więc tu nieco inne znaczenie niż w modelu tradycyjnym, toteż lepiej jest mówić raczej o tłumaczeniu z jednego języka na inny. Na rysunku 2 wskazana została rola językowych narzędzi służących organizacji wiedzy, których zadaniem jest konwertowanie wiedzy nadawcy w informację i informacji w wiedzę odbiorcy.

W modelu przedstawionym na rysunku 2 wykorzystane zostały elementy modelu FRBR [IFLA 1998]. Dzieło, jako jednostka abstrakcyjna, nieistniejąca fizycznie, powstaje na poziomie struktur umysłowych. Autor, stosując zasady odpowiedniego gatunku tekstu, tworzy pierwszy egzemplarz, „rękopis” dzieła, czyli jego realizację. Realizacjami są także różne wersje dzieła, w tym jego tłumaczenia. Składają się one z określonych słów, zdań, akapitów itd., łącznie budujących tekst. Na tym poziomie kodowana jest więc informacja. Materializacja, czyli fizyczne (na wybranym nośniku) utrwalenie realizacji dzieła, obejmuje wszystkie obiekty fizyczne posiadające jednakowe cechy treściowe i formalne, czyli łącznie – cechy gatunkowe. Jest ona wynikiem organizacji informacji przeprowadzonej na poprzednim etapie. Organizacja informacji związana z materializacją dzieła zazwyczaj skutkuje stworzeniem wielu egzemplarzy materializacji<sup>22</sup>. Odbiorca, za pomocą jednego z nich (a dokładnie informacji, którą zawiera dokument), dokonuje modyfikacji swoich struktur umysłowych, pozyskując nową wiedzę.

Z punktu widzenia autora komunikatu konwersja wiedza / informacja / wiedza jest zawsze stratna: powoduje straty znaczeń, w tradycyjnych modelach symbolizowane przez „szumy informacyjne” związane z oddziaływaniem elementów kontekstu, w jakim odbywa się konkretny proces komunikacji, głównie zakłóceń w kanale informacyjnym. Działania w obszarze realizacji tekstu należałoby uznać

---

<sup>22</sup> Wiele egzemplarzy w wyniku materializacji dzieła powstaje w przypadku publikowania analogowego (druku); w przypadku publikowania cyfrowego różnica między materializacją a egzemplarzem zanika [Nahotko 2006, s. 138].

za proces „otwierania” tego kanału. Autor w tworzonej przez siebie informacji nigdy nie jest w stanie przedstawić własnych struktur wiedzy w sposób wyczerpujący (część z nich nie jest nawet dostępna dla jego świadomości [Trzópek 2010, s. 133]). Z punktu widzenia odbiorcy konwersja ta powoduje kolejne straty znaczeń, gdyż uwzględnia on (poddaje interpretacji) tylko część udostępnionej mu informacji, mianowicie tę, którą uważa za istotnie nową (nieoczekiwaną) [Spitzer 2012, p. 134], ją właśnie lepiej zapamiętuje od pozostałej, czyli w sposób bardziej trwały wpisuje w swoje struktury wiedzy. Z drugiej strony pozyskiwana informacja musi być asymilowana z wcześniej istniejącymi strukturami wiedzy, a w trakcie tych procesów mogą być budowane nowe znaczenia, których autor nigdy nie może w całości przewidzieć. Oprócz strat znaczeń można więc mówić także o ich naddatku, który jednak z punktu widzenia autora może mieć charakter „szumu”. Różnice w reprezentacji wiedzy nadawcy i odbiorcy wynikają także z samej natury gatunków tekstów, które są narzędziami językowymi. Każdy użytkownik języka naturalnego używa własnej odmiany tego języka (i własnych odmian gatunków), nieco różnej od wersji innych użytkowników. Elizabeth Bates porównuje nawet zbiór wszystkich znaczeń do góry lodowej, której widoczny wierzchołek, stanowiący powszechnie uznane znaczenia, spoczywa na podstawie zbudowanej ze znaczeń osobistych [Bates et al. 1979, p. 66]. Jednocześnie zindywidualizowane wersje języka są do siebie podobne ze względu na ich społeczne współtworzenie [Devitt 1993, p. 579].

W tych warunkach szczególnego znaczenia nabiera stosowanie gatunku jako konwencji, zgodnie z poglądem o społecznym zdeterminowaniu wszystkich form dyskursu, a szczególnie tych, które są stosowane w kontekście instytucjonalnym. Pomiędzy członkami określonej społeczności badaczy, związanej z określoną dyscypliną nauki, istnieje zawsze konsensus lub porozumienie w sprawie sposobu stosowania struktur wiedzy w wybranych formach dyskursywnych [Bhatia 1997, p. 641]. Podobnie jak inne formy dyskursu, gatunki są także konstruktami społecznymi, bardzo rygorystycznie kontrolowanymi przez praktykę społeczną. Jest to medium, przez które komunikują się ze sobą członkowie społeczności uczonych. Jak twierdzą Carol Berkenkotter i Thomas Huckin, gatunki są ściśle powiązane z metodologią dyscypliny, przez co pozwalają na tworzenie i przekazywanie informacji w sposób zgodny z obowiązującymi w niej normami, wartościami i ideologią [Berkenkotter, Huckin 1995, p. 1].

Konsensus ten jest uzyskiwany i negocjowany przez poinformowanych i znających praktykę badawczą członków społeczności naukowców w trakcie prowadzonych konwersacji i czynności badawczych. Z jednej strony interakcje i konwersacje umożliwiają osiągnięcie konsensusu, a z drugiej wywierają wpływ regulujący (czyli ograniczający) na inne gatunki, gdyż określają dopuszczalne elementy za-



sobu wiedzy społeczności. Tak jak wszystkie schematy mentalne, gatunki są pochodną powtarzających się zdarzeń lub sytuacji. Konkretnie chodzi o powtarzalne sytuacje retoryczne, pobudzające do jednakowej reakcji autorów i czytelników. W taki sposób tworzone są i ustalane formy retoryczne, cechujące się własnym słownictwem, gramatyką i stylem oraz formy, będące odpowiedzią na nie. Znamość gatunku oznacza więc nie tylko umiejętne posługiwanie się konwencją, lecz w równej mierze także umiejętność poprawnego reagowania na zaistniałą sytuację retoryczną. Jednocześnie wykształcenie się określonego gatunku działa hamująco na tworzenie innych reakcji na podobne sytuacje. Konstruowanie gatunku jest więc jednocześnie konstruowaniem sytuacji [Devitt 1993, p. 578].

Gatunki są więc społecznie uzgadniane poprzez tworzenie konwencji i włączane w dyskursywne praktyki członków społeczności badawczych. W znacznym stopniu zależą od intertekstualności dyskursu, czyli od tzw. interdyskursywności<sup>23</sup>. Praktyki dyskursywne w dużej mierze odzwierciedlają nie tylko konwencje stosowane w obrębie określonej społeczności badaczy, ale także konwencje społeczne, takie jak zmiany społeczne, instytucje społeczne, między innymi biblioteki i archiwa [Andersen 2004, p. 93], i wiedzę społeczną. Wszystkie wymienione czynniki mogą być uważane za ważne elementy tego, co w teorii gatunków nazywane jest „wiedzą gatunkową”. Gatunki, będące wynikiem nabywania wiedzy o konwencjach gatunkowych i posługiwania się nimi, regulują konstrukty gatunkowe, podtrzymując ich integralność.

Integralność ta, szczególnie w przypadku takich gatunków dyskursu naukowego, jak artykuły naukowe, osiągnięta jest w środowisku tradycyjnym dwoma sposobami: przez recenzowanie i interwencje redakcyjne. W środowisku elektronicznym dodatkowo stosowane mogą być nowe metody, takie jak komentowanie i punktowanie przez użytkowników. Ważnym aspektem jest nowość tworzonej informacji, jako reprezentacji oryginalnej wiedzy naukowej, do której należy przekonać potencjalnych (wyobrażonych przez autora) odbiorców [Bazerman 1988, p. 25]. Warto zwrócić także uwagę na fakt, że recenzja naukowa (a także np. komentarz) powinna być uznana za odrębny gatunek, charakteryzujący się własnymi celami i zasadami stosowania.

Na koniec należy wspomnieć jeszcze o coraz bardziej rozpowszechniających się gatunkach tekstów elektronicznych, czyli tak zwanych cybergatunkach [Dillon, Gushrowski 2000, p. 202]. Ich stosowanie w nauce jest coraz częstsze, między innymi za sprawą wykorzystywania nowych form publikowania elektroniczne-

---

<sup>23</sup> Interdyskursywność polega na tworzeniu hybrydowych lub nowych gatunków przez przystosowanie lub wykorzystanie ustalonych konwencji lub zasobów związanych z innymi gatunkami. Związana jest ona więc z różnymi procesami dyskursywnymi, takimi jak mieszanie, włączanie lub dostosowywanie norm gatunkowych w praktyce [Bhatia 2008, p. 175].

go. Jedną z najwcześniej powstałych form tego typu były repozytoria. Zgodnie z przedstawioną powyżej teorią schematów takie struktury, powstające dla nowych sposobów prowadzenia dyskursu, tworzone są z wykorzystaniem dotychczas istniejących schematów przez ich stopniową modyfikację aż do utworzenia wzorców wcześniej nieistniejących, charakterystycznych wyłącznie dla zmienionych warunków. W trakcie tych modyfikacji stosowane są metafory, rozumiane jako modele mentalne. Ułatwiają one dostosowanie schematów mentalnych do zachodzących zmian. Określenia takie jak „książka elektroniczna” lub „czasopismo elektroniczne” należy traktować jako metafory, podobne do wcześniej stosowanych, których przykładem może być nazywanie pulpitem obszaru wyświetlanego na monitorze komputerowym [Kotuła 2013, s. 111]. Wymienione formy publikowania elektronicznego są bowiem związane z powstawaniem nowych gatunków, za pomocą których realizowane są cele komunikacyjne odmienne od realizowanych za pomocą gatunków tradycyjnych.

Model analizy gatunków funkcjonujących w world wide web zaproponowali Inger Askehave i Anne Nielsen [Askehave, Nielsen 2005, p. 129]. Środowisko sieci globalnych jest integralną częścią stosowanych tam gatunków, więc jego cechy, takie jak nawigacyjność i hipertekstualność, muszą zostać uwzględnione podczas analizy tych gatunków. Dokumenty internetowe muszą być rozpatrywane z dwóch perspektyw: jako tradycyjny tekst oraz jako medium pozwalające użytkownikowi na nawigowanie w zasobach. Ten dualizm natury dokumentów sieciowych utrudnia prowadzenie tradycyjnej analizy.

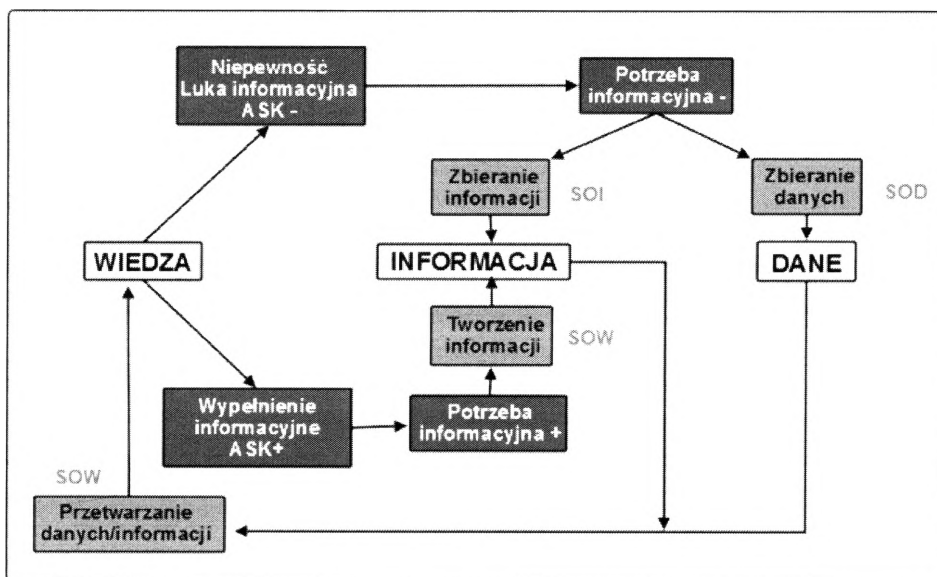
Michael Shepherd i Carolyn Watters w swoim artykule [Shepherd, Watters 1998, p. 100] zaproponowali, aby opisywać cybergatunki za pomocą trzech cech: treści, formy i funkcjonalności, w odróżnieniu od gatunków tradycyjnych, opisywanych jedynie za pomocą dwóch pierwszych atrybutów. Funkcjonalność jako cecha dodatkowa jest bowiem charakterystyczna dla środowiska sieciowego i umożliwia tworzenie zupełnie nowych form gatunkowych. Dzięki temu może dojść do rozwoju innowacji, pozwalających na przejście od prostych elektronicznych kopii (wręcz skanów w formie plików graficznych) publikacji tradycyjnych o bardzo ograniczonej funkcjonalności, zapełniających w dużej części pierwsze i współczesne repozytoria, do zupełnie nowych form publikowania, wykorzystujących w pełni funkcjonalność środowiska elektronicznego, czego często wymienianym przykładem są blogi (w tym tworzone przez naukowców) [McNeill 2005] oraz formy wciąż mające charakter eksperymentalny, takie jak publikacje rozwinięte (ang. *enhanced*), na przykład Rich Internet Publications [Bardi, Manghi 2014, p. 253].

## 4. Organizacja informacji

Organizacja informacji (OI), rozumiana tak, jak w modelu przedstawionym na rysunku 1, stanowiła zawsze przedmiot zainteresowania bibliotekoznawstwa i informatologii, nauk pokrewnych oraz dziedzin działalności praktycznej: bibliotekarstwa, księgarstwa, edytorstwa oraz archiwistyki itp. Opisowi procesów OI poświęcano grube tomy publikacji, w tym tekście skupię się więc jedynie na ich przedstawieniu z punktu widzenia opisanych wcześniej teorii, szczególnie teorii gatunków tekstu.

Na rysunku 3 przedstawione zostały procesy wiedzy i informacji w celu wskazania w nich miejsca SOW i SOI. Rysunek ten powstał na podstawie wcześniej opublikowanych modeli zachowań informacyjnych, modeli opracowanych m.in. przez Toma Wilsona [Wilson 1997; Wilson 1999], Brendę Dervin [Naumer, Fisher, Dervin 2008], Barbarę Niedźwiedzka [Niedźwiedzka 2003] i Natalię Godbold [Godbold 2006]. Wykazuje on także podobieństwo do cyklu DIK [Tian, Nakamori, Wierzbiński 2009, p. 79]. W przedstawionych procesach ważną rolę odgrywa także teoria anomalnych stanów wiedzy (ang. *anomalous states of knowledge* – ASK) Nicholasa Belkina [Belkin, Oddy, Brooks 1982, p. 65].

Teoria ASK wyjaśnia powstawanie potrzeby informacyjnej. Zgodnie z nią wyszukiwanie informacji następuje, gdy użytkownik stwierdzi istnienie problemu,



Rysunek 3. Organizacja danych, informacji i wiedzy  
Źródło: opracowanie własne

wskazującego na nieadekwatność jego stanu wiedzy. Jeśli stan wiedzy użytkownika jest adekwatny do rozwiązywanych przez niego problemów, to nie pojawia się potrzeba informacyjna. W podobny sposób można także wyjaśnić genezę potrzeby tworzenia informacji naukowej. Pewnego rodzaju anomalia cechuje stan wiedzy naukowca, który postanawia zostać autorem, anomalia ta ma jednak przeciwny znak. O ile u naukowca-użytkownika informacji ma ona znak ujemny (ASK-) i oznacza niedostatek – lukę wiedzy, którą użytkownik próbuje zniwelować, poszukując nowej informacji, o tyle u naukowca-autora ma znak dodatni (ASK+), gdyż posiada on nową wiedzę, możliwą do skonwertowania do postaci informacji tekstowej i rozpowszechnienia. Jeżeli wiedza ma walor nowości bezwzględnej<sup>24</sup>, czyli jest nowa nie tylko dla jej posiadacza, ale także w skali dyscypliny (a więc dla potencjalnych odbiorców), to naukowiec podejmuje decyzje o jej konwersji w informację i rozpowszechnieniu. W ten sposób struktury wiedzy jednostek ulegają ujednoczeniu w stopniu umożliwiającym współpracę.

Rola SOI w wyszukiwaniu informacji jest dobrze opisana w literaturze; narzędzia te mają służyć odszukaniu tekstów (publikacji) relewantnych do potrzeb osoby zbierającej informację oraz umożliwić do niej dostęp. Z punktu widzenia teorii gatunków tekstów rola tych narzędzi jest bardziej zróżnicowana. Jak pisze Jack Andersen, organizacja informacji jest typizowanym działaniem retorycznym, czyli również gatunkiem. Organizacja informacji w systemach informacyjnych może być traktowana jako działanie retoryczne, mające wpływ na rejestrowane zasoby dokumentów (a przez to ich gatunków), zarówno podczas wyboru dokumentów do umieszczenia w zasobie, jak i sposobu ich opracowania. Indeksujący tworzą nowe rodzaje tekstów – reprezentacje dokumentów – stając się autorami odpowiedzialnymi za ich treść [Andersen 2004, p. 93]. Organizacja informacji nie może także funkcjonować bez uwzględnienia społecznych działań komunikacyjnych, w ramach których, na różnych poziomach, używa się zarejestrowanych dokumentów tekstowych. Organizacja informacji zawartej w dokumentach, wspomaga te działania. Część z nich służy dystrybucji dokumentów w czasie. Archiwa, biblioteki, katalogi i bibliografie są gatunkami tworzonymi i sukcesywnie modyfikowanymi w celu wspomaganie i zarządzania procesami twórczości piśmiennej i dokumentowania, pełniącymi istotne funkcje w organizacji społeczeństwa [Andersen 2008, p. 360]. W ten sposób biblioteka jawi się

---

<sup>24</sup> O tym, czy tak rzeczywiście jest, decyduje (przyszły) autor indywidualnie, na podstawie własnej wiedzy. Istnieją mechanizmy naprawy ewentualnych pomyłek, np. wspomniane już recenzowanie lub komentowanie publikacji naukowych i ingerencje redakcji. Środowisko elektroniczne daje tu nowe możliwości, ale z drugiej strony bywa też krytykowane, np. za niestałość, łatwość mnożenia tekstów sposobem „kopiuj i wklej”, częsty brak autorstwa w tradycyjnym sensie i niszczenie indywidualnej kreatywności, funkcjonowanie treści ważnych obok zupełnie bezwartościowych (np. [Waters 2009, p. 29]).

jako wielki tekst, czy raczej – gatunek tekstu, który składa się z podgatunków (czasopisma i artykuły w nich, książki, mapy, normy, patenty i inne rodzaje dokumentów), tworzących system gatunków, czyli, według Charlesa Bazermana, zbiór gatunków powiązanych relacjami, które współdziałają w określonych warunkach [Bazerman 1994, p. 97].

Organizacja informacji tekstowej i jej wyszukiwanie w systemach informacyjnych jest równie złożona, jak czynności związane z pisaniem i czytaniem. Obydwie grupy działań opierają się na znajomości odpowiednich, różnorodnych gatunków tekstów. Są to powiązane ze sobą aktywności człowieka, wspierające się nawzajem. Wiedza o gatunkach stosowana w procesach pisania i czytania w określonych warunkach nie różni się od tej, która musi być użyta podczas organizowania i wyszukiwania informacji w systemach informacyjnych. Tych problemów dotyczą badania nad zwiększeniem efektywności wyszukiwania przez umieszczenie w opisie dokumentu elementu zawierającego informację o gatunku opisywanego tekstu (por. np. [Kwasnik et al. 2001, p. 24]).

W trakcie organizowania informacji przez pracowników informacji stosowane są kryteria przynależności gatunkowej tekstów. Wyróżnić można trzy gatunki tekstów, stosowanych w tych procesach, dla których funkcjonują różne podgatunki. Pierwszy obejmuje publikacje dostarczane do systemów informacyjnych z zewnątrz, dwa kolejne tworzone są w ramach systemów:

- Teksty zwane pierwotnymi, czyli publikacje naukowe. Ich autorami są naukowcy, którzy tworzą je dla innych naukowców. Bibliotekarze pełnią rolę pośredników, a wykazywana przez nich znajomość gatunków, którymi posługują się ich użytkownicy, może być niedostateczna. Stąd wynikają narzekania na absolwentów bibliotekoznawstwa jako na osoby nieznające specyfiki obsługiwanych przez siebie dziedzin [Rusek 1995, s. 60]. Nie są oni specjalistami w zakresie gatunków tekstów używanych przez naukowców do opisywania swoich struktur wiedzy, gdyż nie oni są głównymi adresatami tych tekstów. Powstaje więc problem nieadekwatnych struktur wiedzy bibliotekarzy. Problem ten może być pogłębiany przez zjawisko tworzenia lokalnych gatunków tekstów, funkcjonujących w obrębie małych grup, na przykład wąskich specjalności naukowych. W takich środowiskach uzgadnianie wiedzy jest utrudnione, a dostęp do wiedzy gatunkowej z zewnątrz prawie niemożliwy. Proces ten ułatwić mają np. instrukcje dla autorów, zamieszczane przez redakcje czasopism, kodyfikujące wiedzę gatunkową. Publikowanie pierwotnych tekstów naukowych przenosi się do internetu, co wymaga stosowania nowych sposobów organizacji informacji. W środowisku cyfrowym wzrasta rola relacji intertekstualnych, symbolizowanych przez hipertekst, aż do traktowania internetu jako jednego, wielkiego tekstu. Standardowy tekst może być rozłożony na mniejsze jednostki lub powiększany aż do rozmia-

rów *docuverse* (ang. *document+Universe*) powiązanych tekstów [Delany, Landow 2008, p. 94]. Tylko w obszarze stosowania gatunków tej grupy istnieje wymienność ról: czytelnicy publikacji naukowych bywają jednocześnie ich autorami, co zapewnia społeczne ujednolicanie struktur wiedzy gatunkowej.

- Teksty pochodne, czyli opisy tworzone przez bibliotekarzy, zarówno jako efekt opracowania formalnego (opis bibliograficzny), jak i rzeczowego, należące do tzw. paratekstu. Gatunki są tu zwykle dobrze opisane. Na przykład gatunek opisu bibliograficznego jest tak ściśle zdefiniowany, że aż znormalizowany<sup>25</sup>. Zbiory tych tekstów nazywane są danymi bibliograficznymi, chociaż w rzeczywistości są informacją zawartą w tekstach powstałych w wyniku konwersji wiedzy indeksującego. Twórcą tekstu opisu bibliograficznego, podobnie jak rzeczowego, jest bibliotekarz, który dobrze zna stosowany przez siebie gatunek tekstu. Tekst zawiera informację, będącą reprezentacją jego wiedzy, dotyczącej zarówno zasad opracowania (wiedza specjalistyczna), jak i zagadnienia opisywanego w tekście (wiedza czytelnika, odwzorowująca wiedzę autora). Zauważmy, że teksty tego gatunku, pomimo normalizacji, powstają z wykorzystaniem indywidualnych struktur wiedzy bibliotekarzy, stąd różnice w opisach tworzonych na podstawie tych samych norm (wiedzy gatunkowej) i tekstów pierwotnych. Odbiorcami tekstów opisów bibliograficznych są zarówno inni bibliotekarze, znający ten gatunek, jak i inni użytkownicy SOI (czytelnicy poszukujący literatury), posiadający różną wiedzę gatunkową, często niedostateczną. O ile znajomość gatunku tekstu pochodnego uważana jest za część wiedzy obowiązkowej naukowca, o tyle nie można jej wymagać od studentów, szczególnie pierwszych lat studiów. Zdobywanie wiedzy o gatunkach tekstów pochodnych stosowanych w określonej dziedzinie powinno być więc równorzędnym celem dydaktyki każdego kierunku studiów, umożliwia bowiem samodzielne korzystanie z literatury naukowej. Innym problemem są trudności w uzgadnianiu między bibliotekarzami struktur wiedzy dotyczących stosowanego gatunku opisów<sup>26</sup>, gdyż tworzą oni małe, rozproszone środowisko, w którym społeczne uzgadnianie wiedzy jest utrudnione. Duży wpływ na rozpowszechnienie wiedzy gatunkowej w tym zakresie miała komputeryzacja tworzenia tekstów opisów i wymuszona przez nią współpraca globalna.

- Trzeci gatunek tekstów tradycyjnie stosowany jest w obrębie narzędzi zwanych systemami organizacji wiedzy. W tym obszarze powstają teksty zawarte na przykład w tablicach klasyfikacyjnych lub słownikach haseł przedmiotowych, tezaursach oraz wielu innych, wymienionych m.in. w pracy Gail Hodge [Hodge 2000]. W tym zakresie sytuacja jest bardzo podobna do przedstawionej

---

<sup>25</sup> Por. np. normę PN-N-01152 *Opis bibliograficzny*.

<sup>26</sup> Stąd wynikała duża różnorodność instrukcji katalogowania stosowanych na przełomie XIX i XX wieku.

w punkcie poprzednim. Tworzone są narzędzia tekstowe (SOI według modelu z rysunku 1), których teksty należą do gatunków znanych wąskiej grupie osób stosujących je na co dzień (bibliotekarzom). Dodatkowo teksty należące do tych gatunków są reprezentacją wiedzy ich autorów (specjalistów z różnych dziedzin, w tym bibliotekarzy) i muszą podlegać stałej modyfikacji (to różni je od tekstów poprzednich grup), co często pociąga za sobą konieczność włączenia w te procesy innych autorów, z natury rzeczy posiadających odmienne struktury wiedzy. Za pomocą tych tekstów tworzone są kolejne teksty należące do innych gatunków (pochodne). Teksty tej grupy, tak jak wszystkie inne teksty, są interpretowane indywidualnie przez ich czytelników (bibliotekarzy i użytkowników) na podstawie odmiennych struktur wiedzy, co może powodować spadek efektywności wyszukiwania informacji. Społeczne ujednolicanie wiedzy gatunkowej jest bardzo utrudnione ze względu na lokalne ich stosowanie i zasadnicze różnice międzygatunkowe. W przypadku dwóch ostatnich grup nie występuje wymiennieść ról: nie wszyscy czytelnicy tekstów (użytkownicy gatunków) mogą być ich autorami, co utrudnia ujednolicanie struktur wiedzy.

## Zakończenie

Głównym wnioskiem wyłaniającym się z powyższych rozważań jest konstatacja, że wszystkie problemy związane z zagadnieniami wiedzy, informacji i danych, a w związku z tym także organizacji wiedzy i informacji, wraz z systemami tworzonymi dla usprawnienia tej organizacji, muszą być rozwiązywane w oparciu o wiedzę neurologiczną, dotyczącą budowy i funkcji mózgu. Struktury wiedzy i informacji oraz możliwości ich konwersji w procesach komunikacji zależą od funkcji centralnego układu nerwowego. Wiedza jest wynikiem funkcjonowania umysłu ludzkiego, poza umysłem nie istnieje, przez co nie może być bezpośrednio przekazywana. Wyjaśnia to rolę informacji i systemów informacyjnych jako elementu OI.

Struktury umysłu jednostki nie działają jednak oczywiście w próżni, lecz są ściśle związane z otaczającym światem, gdyż tylko taki związek daje szansę na właściwe dostosowanie struktur umysłowych do wymagań środowiska, co z kolei pozwala na optymalizację warunków życia, zwiększającą szansę na przeżycie. Ważnym elementem otoczenia człowieka od zawsze były różnego rodzaju struktury sieciowe, np. sieci kontaktów międzyludzkich, sieci dróg, szkół, placówek pocztowych, posterunków policji i inne, w których dokonywano m.in. organizacji informacji. Od lat 90. XX wieku zasadniczą rolę w tym zakresie zaczęły odgrywać globalne sieci informacyjne, głównie internet. Sieci te poważnie zmieniają

sposób funkcjonowania społecznie konstruowanych procesów organizacji wiedzy i informacji.

Pojawianie się w środowisku społecznym nowych możliwości, wymuszających zdobywanie nowych doświadczeń, takich jak stosowanie narzędzi komunikacji elektronicznej, powoduje zmiany struktur umysłowych, za którymi podążają zmiany w organizacji wiedzy i informacji, w tym powstawanie nowych gatunków tekstów cyfrowych. Publikowanie elektroniczne przybiera nowe formy, wcześniej nieznanne; jedną z nich jest tworzenie repozytoriów obiektów cyfrowych i samopublikowanie. Nowo powstające gatunki dokumentów tekstowych stymulują dalszy rozwój publikowania elektronicznego. Początkowo dla tworzenia charakterystyk nowych gatunków wykorzystywane są metafory odnoszące się do schematów znanych i od dawna stosowanych. W wyniku ich ewolucji powstają konwencje dotyczące formy i treści specyficzne dla środowiska elektronicznego, najlepiej dostosowane do istniejących możliwości technologicznych i potrzeb komunikacyjnych.

Organizacja wiedzy i informacji w komunikowaniu naukowym jest odzwierciedleniem poglądów na otaczający świat [Andersen 2004, p. 219]. Działalność ta silnie wpływa na sposób transmisji w społeczeństwie wiedzy skonwertowanej do formy informacji i jej archiwizację. Znajomość zagadnień dotyczących organizacji wiedzy i informacji wymaga zrozumienia społecznych i historycznych determinantów tych procesów. Od nich zależą ograniczenia, jakim podlegają procesy i struktury stosowane w komunikacji naukowej. Ich usprawnienie zależy natomiast od lepszego zrozumienia społecznych zasad funkcjonowania społeczności dyskursywnej, która jest specyficzną formą organizacji społecznej. Dla zrozumienia społeczności dyskursywnej duże znaczenie mogą mieć badania nad gatunkami tekstu. Tak jak gatunki budują sytuacje, a sytuacje tworzą gatunki, podobnie dyskurs tworzy społeczność, a społeczność – dyskurs. Być może więc społeczność można lepiej zdefiniować przez opis stosowanych w niej gatunków niż przez charakterystykę tworzących ją grup i ich członków.

Gatunek pełni bardzo istotną rolę zarówno z punktu widzenia funkcjonowania jednostki i jej procesów kognitywnych [Vaughan, Dillon 1998, p. 563], jak i procesów społecznych, w których jednostki uczestniczą. Jest on elementem wiążącym indywidualną wiedzę, zakodowaną w schematach umysłowych, ze społecznymi procesami komunikacyjnymi służącymi wymianie informacji. Gatunki kształtują aktywność komunikacyjną jednostek tworzących społeczności dyskursu i jednocześnie same są przez tę aktywność kształtowane. Stanowią pomost pomiędzy wiedzą autora i wiedzą czytelnika, pozwalają badać relacje pomiędzy działaniem twórczym (np. pisaniem tekstu) a wysiłkiem odbiorcy (np. czytaniem), stając się zasadniczym elementem tworzenia znaczeń. Autor, przez wybór



określonego gatunku, przystaje na ograniczenie kręgu swoich odbiorców do społeczności posiadającej odpowiednią wiedzę gatunkową, co powoduje ograniczenie możliwości dotarcia do szerszego grona odbiorców i realizacji większej liczby heterogenicznych celów. W zamian jednak autorowi, podążającemu za konwencją gatunku, znacznie łatwiej jest dotrzeć do wybranych odbiorców i zrealizować dokładnie wyselekcjonowane cele.

## BIBLIOGRAFIA

- Ablewicz, Krystyna (1998). *Hermeneutyka i fenomenologia w badaniach pedagogicznych*. W: Palka, Stanisław red. *Orientacje w metodologii badań pedagogicznych*. Kraków: Wydaw. Uniwersytetu Jagiellońskiego, s. 25-41.
- Abraham, Anna (2014). *Creative thinking as orchestrated by semantic processing vs. cognitive control brain networks* [dok. elektr.]. *Frontiers in Human Neuroscience* Vol. 8, Art. 95. <http://journal.frontiersin.org/Journal/10.3389/fnhum.2014.00095/full> [odczyt: 15.10.2015].
- Andersen, Jack (2008). *The concept of genre in information studies*. *Annual Review of Information Science and Technology* Vol. 42, No. 1, p. 339-367.
- Andersen, Jack (2004). *Analyzing the role of knowledge organization in scholarly communication: an inquiry into the intellectual foundation of knowledge organization*. Copenhagen: Royal School of Library and Information Science, 257 p.
- Anderson, Richard (1977). *Schema-directed process in language comprehension*. Urbana: University of Illinois, 28 p.
- Anderson, Richard; Pearson, David (1988). *A schema-theoretic view of basic progress in reading comprehension*. In: Carrell, Patricia L.; Devine, Joanne; Eskey, David E. eds. *Interactive Approaches to Second Language Reading*. Cambridge; New York: Cambridge University Press, p. 37-55.
- Anderson, Richard; Pichert, James; Shirey, Larry (1979). Effects of the reader's schema at different points in time. Technical Report no. 119. Champaign: University of Illinois, 36 p.
- Askehave, Inger; Nielsen, Anne (2005). *Digital genres: a challenge to traditional genre theory*. *Information Technology & People* Vol. 18, No. 2, p. 120-141.
- Bachtin, Michał (1986). *Estetyka twórczości słownej*. Warszawa: Państwowy Instytut Wydawniczy, 576 s.
- Bardi, Alessia; Manghi, Paolo (2014). *Enhanced publications: data models and information systems*. *LIBER Quarterly* Vol. 23, No. 4, p. 240-273.
- Bartlett, Frederic (1932). *Remembering*. Cambridge: Cambridge University Press, 317 p.
- Bates, Elizabeth et al. (1979). *The emergence of symbols: cognition and communication in infancy*. New York: Academic Press, 387 p.
- Bates, Marcia (2006). *Fundamental forms of information*. *Journal of the Association for Information Science and Technology* Vol. 57, No. 8, p. 1033-1045.

- Bates, Marcia (2005). *Information and knowledge: an evolutionary framework for information science* [dok. elektr.]. Information Research Vol. 10, No. 4. <http://InformationR.net/ir/10-4/paper239.html> [odczyt: 11.09.2014].
- Bawden, David; Robinson, Lynn (2012). *Introduction to information science*. London: Facet Publishing, 351 p.
- Bazerman, Charles (1994). *Systems of genre and the enactment of social intentions*. In: Freedman, Aviva; Medway, Peter ed. *Genre and the new rhetoric*. London: Taylor & Francis, p. 79-101.
- Bazerman, Charles (1988). *Shaping written knowledge. The genre and activity of the experiential article in science*. Madison: University of Wisconsin Press, 356 p..
- Belkin, Nicholas; Oddy, Robert; Brooks, Helen (1982). *ASK for information retrieval: Part I. Background and theory*. Journal of Documentation Vol. 38, No. 2, p. 61-71.
- Bergen, Benjamin (2012). *Louder than words: the new science of how the mind makes meaning*. New York: Basic Books, 296 p.
- Berger, Arthur (1992). *Popular culture genres. Theories and texts*. Newsbury Park: Sage Publications, 171 p.
- Berkenkotter, Carol; Huckin, Thomas (1995). *Genre knowledge in disciplinary communication – cognition – culture – power*. New Jersey: Erlbaum, 190 p.
- Bhatia, Vijay (2008). *Towards critical genre analysis*. In: Bhatia, Vijay; Flowerdew, John; Jones, Rodney ed. *Advances in discourse studies*. London: Routledge, p. 166-177.
- Bhatia, Vijay (1997). *Genre analysis today*. Revue belge de philologie et d'histoire Vol. 75, No. 3, p. 629-652.
- Borgman, Christine (2007). *Scholarship in the digital age. Information, infrastructure, and the Internet*. Cambridge, Mass: MIT Press, 336 p.
- Brookes, Bertram (1980). *The foundations of information science. Part I. Philosophical aspects*. Journal of Information Science Vol. 2, No. 3 / 4, p. 125-133.
- Carley, Kathleen; Michael Palmquist (1992). *Extracting, representing, and analyzing mental models*. Social Forces Vol. 70, No. 3, p. 601-636.
- Case, Donald (2007). *Looking for information. A survey of research on information seeking, needs, and behavior*. Amsterdam: Academic Press, 423 p.
- Chmielecki, Andrzej (2001). *Między mózgiem i świadomością. Próba rozwiązania problemu psychofizycznego*. Warszawa: Wydaw. Instytutu Filozofii i Socjologii PAN, 257 s.
- Cisek, Sabina (2002). *Filozoficzne aspekty informacji naukowej*. Kraków: Wydaw. Uniwersytetu Jagiellońskiego, 140 s.
- DeFilippe, Javier (2011). *The evolution of the brain, the human nature of cortical circuits, and intellectual creativity* [dok. elektr.]. Frontiers in Neuroanatomy Vol. 5, Art. <http://journal.frontiersin.org/Journal/10.3389/fnana.2011.00029/full> [odczyt: 15.10.2014].
- Delany, Paul; Landow, George (2008). *Zarządzanie cyfrowym słowem. Tekst w epoce reprodukcji elektronicznej*. W: Gwóźdź, Andrzej red. *Ekrany piśmiennosci*. Warszawa: Wydaw. Akademickie i Profesjonalne, s. 75-101.
- Devitt, Amy (1993). *Generalizing about genre: new conceptions of an old concept*. College Composition and Communication Vol. 44, No. 4, p. 573-586.
- Dillon, Andrew; Gushrowski, Barbara (2000). *Genres and the Web: is the personal home page the first uniquely digital genre?* Journal of the Association for Information Science and Technology Vol. 51, No. 2, p. 202-205.

- Duszak, Anna (1998). *Tekst, dyskurs, komunikacja międzykulturowa*. Warszawa: Wydaw. Naukowe PWN, 385 s.
- Finke, Ronald; Ward, Thomas; Smith, Steven (1992). *Creative cognition: theory, research and applications*. Cambridge, Mass: MIT Press, 239 p.
- Fisher, Karen; Erdelez, Sandra; McKechnie, Lynne (2005). *Theories of information behavior*. Medford, N.J.: American Society for Information Science and Technology, 431 p.
- Fodor, Jerry (2011). *Język myśli. LOT 2*. Warszawa: Wydaw. Naukowe PWN, 316 s.
- Genette, Gerard (1991). *Introduction to paratext*. New Literary History Vol. 22, No. 2, p. 261-272.
- Godbold, Natalya (2006). *Beyond information seeking: towards a general model of information behavior* [dok. elektr.]. Information Research Vol. 11, No. 4, Art. 269. <http://InformationR.net/ir/11-4/paper269.html> [odczyt: 7.09.2014].
- Gut, Arkadiusz (2010). *Kognitywne podstawy komunikacji*. Przegląd Filozoficzny R. 19, Nr 3, s. 321-337.
- Hajibayova, Lala; Elin, Jacob (2014). *User-generated genre tags through the lens of genre theories*. In: Babik, Wiesław ed. *Knowledge organization in the 21<sup>st</sup> century: between historical patterns and future prospects*. Würzburg: Ergon Verlag, p. 390-396.
- Hodge, Gail (2000). *Systems of knowledge organization for digital libraries: beyond traditional authority files* [dok. elektr.]. Washington: Digital Library Federation, 43 p. <http://www.clir.org/pubs/reports/pub91/pub91.pdf> [odczyt: 18.08.2014].
- Hyland, Ken; Salager-Meyer, Françoise (2008). *Scientific writing*. Annual Review of Information Science and Technology Vol. 42, No. 1, p. 297-338.
- IFLA Study Group on the Functional Requirements for Bibliographic Records (1998). *Functional Requirements for Bibliographic Record: Final Report*. München: Saur, 144 p.
- Jones, Natalie et al. (2011). *Mental models: an interdisciplinary synthesis of theory and methods* [dok. elektr.]. Ecology and Society Vol. 16, No. 1, Art. 46. <http://www.ecologyandsociety.org/vol16/iss1/art46> [odczyt: 15.10.2014].
- Joseph, Rhawn (2011). *Neuroscience, neuropsychology, neuropsychiatry, brain & mind: introduction, primer, & overview*. Indie: University Press Science Publishers, 778 p.
- Kaczmarek, Jan (2013). *Affective conception of information and affect representation in information systems*. Zagadnienia Informacji Naukowej Vol. 51, Nr 2, s. 64-77.
- Kintsch, Walter; van Dijk, Teun (1978). *Toward a model of text comprehension and production*. Psychological Review Vol. 85, No. 5, p. 363-394.
- Kliś, Maria (2002). *Wnioskowanie lingwistyczne a proces recepcji tekstu narracyjnego*. Roczniki Psychologiczne T. 5, s. 99-113.
- Kolb, Alice; Kolb, David (2005). *Learning styles and learning spaces: enhancing experiential learning in higher education*. Academy of Management Learning & Education Vol. 4, No. 2, p. 193-212.
- Kotuła, Sebastian (2013). *Komunikacja bibliologiczna wobec World Wide Web*. Lublin: Wydaw. Uniwersytetu Marii Curie-Skłodowskiej, 170 s.
- Kron, Friedrich; Sofos, Alivisios (2008). *Dydaktyka mediów*. Gdańsk: Gdańskie Wydaw. Psychologiczne, 165 s.
- Król, Anna (2006). *Artykulacja i percepcja w perspektywie psycholingwistycznej*. Homo Communicativus Nr 1, s. 54-61.

- Krzysztofek, Kazimierz (2014). *Obszary i konteksty informatologii w epoce cyfrowej: sieci – informacja – dane – software*. Zagadnienia Informatologii Vol. 52, Nr 1, s. 19-42.
- Kurcz, Ida (2000). *Psychologia języka i komunikacji*. Warszawa: Wydaw. Naukowe Scholar, 243 s.
- Kurcz, Ida (1987). *Język a reprezentacja świata w umyśle*. Warszawa: Wydaw. Naukowe PWN, 364 s.
- Kwasnik, Barbara et al. (2001). *Identifying document genre to improve Web search effectiveness*. Bulletin of the Association for Information Science and Technology Vol. 27, No. 2, p. 23-26.
- Maciejewski, Marcin (2009). *Gatunki hipertekstu w perspektywie tekstologicznej. Analiza na przykładzie internetowych prezentacji przedsiębiorstw*. Poznań: Wydaw. Naukowe Uniwersytetu im. Adama Mickiewicza, 371 s.
- Mackenzie Owen, John Stuart (2007). *The scientific article in the age of digitization*. Berlin: Springer Verlag, 263 p.
- Mazurkiewicz-Sokołowska, Jolanta (2010). *Lingwistyka mentalna w zarysie. O zdolności językowej w ujęciu integrującym*. Kraków: Universitas, 261 s.
- McClelland, James; Elman, Jeffrey (1986). *The TRACE model of speech perception*. Cognitive Psychology Vol. 18, No. 1, p. 1-86.
- McNeill, Laurie (2005). *Genre under construction: the diary on the Internet* [dok. elektr.]. Language@Internet Vol. 2, Art. 1. <http://www.languageatinternet.org/articles/2005/120> [odczyt 14.10.2014].
- Miller, Carolyn (1984). *Genre as social action*. Quarterly Journal of Speech Vol. 70, p. 151-167.
- Nahotko, Marek (2006). *Opis dokumentów elektronicznych. Teoretyczny model i możliwości jego aplikacji*. Kraków: Wydaw. Uniwersytetu Jagiellońskiego, 193 s.
- Naumer, Charles; Fisher, Karen; Dervin, Brenda (2008). *Sense-making: a methodological perspective*. In: *Sensemaking workshop* [dok. elektr.]. ACM SIGCHI Conference. Florence, 6 April 2008. <http://www.dmrussell.googlepages.com/sensemakingworkshop-papers> [odczyt: 12.08.2014].
- Nawrocki, Witold (2003). *W poszukiwaniu istoty informacji*. W: Jadacki, Jacek red. *Analiza pojęcia informacji*. Warszawa: Wydaw. Naukowe Semper, s. 37-62.
- Nęcka, Edward (2012). *Psychologia twórczości*. Sopot: Gdańskie Wydaw. Psychologiczne, 256 s.
- Nęcka, Edward; Orzechowski, Jarosław; Szymura, Błażej (2012). *Psychologia poznawcza*. Warszawa: Wydaw. Naukowe PWN, 740 s.
- Niedzwiedzka, Barbara (2003). *A proposed general model of information behavior* [dok. elektr.]. Information Research Vol. 9, No. 1. <http://InformationR.net/ir/9-1/paper164.html> [odczyt: 8.09.2014].
- Norman, Donald; Shallice, Tim (1986). *Attention to action. Willed and automatic control of behavior*. In: Davidson, Richard; Schwartz, Gary; Shapiro, David ed. *Consciousness and self-regulation*. New York: Plenum, p. 1-18.
- Nowak, Tomasz (2011). *Język w świetle odkryć nauki*. Kraków: Wydaw. Petrus, 452 s.
- Osika, Grażyna (2011). *Procesy i akty komunikacyjne. Koncepcje klasyczne i współczesne*. Kraków: Universitas, 214 s.
- Ossowski, Stanisław (2001). *O osobliwościach nauk społecznych*. Warszawa: Wydaw. Naukowe PWN, 226 s.

- Piaget, Jean (1973). *To understand is to invent. The future of education*. New York: Grosman Publishers, 148 p.
- Piaget, Jean (1926). *The language and thought of the child*. New York: Harcourt, 246 p.
- Popper, Karl; Eccles, John (1985). *The self and its brain*. Berlin, New York, London: Springer Verlag, 597 p.
- Przywara, Paweł (2010). *Problem mentalesi*. Roczniki Filozoficzne T. LVIII, Nr 2, s. 97-124.
- Rieh, Soo Young (2002). *Judgement of information quality and cognitive authority in the Web*. Journal of the Association for Information Science and Technology Vol. 53, No. 2, p. 145-161.
- Rosenblatt, Louise (1988). *Writing and reading: the transactional theory*. Champaign: University of Illinois at Urbana-Champaign, 17 p.
- Rubinstein, Joshua; Meyer, David; Evans, Jeffrey (2001). *Executive control of cognitive processes in task switching*. Journal of Experimental Psychology: Human Perception and Performance Vol. 27, No. 4, p. 736-797.
- Rumelhart, David; Ortony, Andrew (1977). *The representation of knowledge in memory*. In: Anderson, Richard; Spiro, Rand; Montague, William ed. *Schooling and the acquisition of knowledge*. Hillsdale: Erlbaum, p. 99-135.
- Rusek, Adam (1995). *Miła praca blisko domu. Osoby bez przygotowania zawodowego zatrudnione w bibliotekach o pracy i zawodzie bibliotekarza*. Warszawa: Biblioteka Narodowa, 84 s.
- Shepherd, Michael; Watters, Carolyn (1998). *The evolution of cybergenres*. In: *Proceedings of the 31<sup>st</sup> Annual Hawaii International Conference on Systems Sciences (HICSS'98)*. Los Alamitos: IEEE Press, p. 97-109.
- Spitzer, Manfred (2012). *Jak uczy się mózg*. Warszawa: Wydaw. Naukowe PWN, 350 s.
- Sporns, Olaf (2011). *Networks of the brain*. Cambridge, Mass: MIT Press, 412 p.
- Szalkowska, Anna; Francuz, Piotr (2003). *Rozumienie informacji prasowej: badania psychologiczne*. Przegląd Psychologiczny T. 46, Nr 1, s. 79-99.
- Świgoń, Marzena (2012). *Zarządzanie wiedzą i informacją. Podstawy teoretyczne. Badania w wymiarze indywidualnym*. Olsztyn: Wydaw. Uniwersytetu Warmińsko-Mazurskiego, 367 s.
- Tabakowska, Elżbieta (1995). *Gramatyka i obrazowanie. Wprowadzenie do językoznawstwa kognitywnego*. Kraków: Wydaw. Oddziału PAN, 103 s.
- Talja, Sanna; Tuominen, Kimmo; Savolainen, Reijo (2005). *"Isms" in information science: constructivism, collectivism and constructionism*. Journal of Documentation Vol. 61, No. 1, p. 79-101.
- Taylor, John (2007). *Gramatyka kognitywna*. Kraków: Universitas, 755 s.
- Tian, Jing; Nakamori, Yoshiteru; Wierzbicki, Andrzej (2009). *Knowledge management and knowledge creation in academia: a study based on surveys in a Japanese research university*. Journal of Knowledge Management Vol. 13, No. 2, p. 76-92.
- Tonkin, Elizabeth (1992). *Narrating our pasts: the social construction of oral history*. Cambridge: Cambridge University Press, 171 p.
- Tucker, Don; Luu, Phan (2012). *Cognition and neural development*. Oxford: Oxford University Press, 276 p.
- Trzebiński, Jerzy (2002). *Narracyjne konstruowanie rzeczywistości*. W: Trzebiński, Jerzy red. *Narracja jako sposób rozumienia świata*. Gdańsk: Gdańskie Wydaw. Psychologiczne, s. 17-24.

- Trzebiński, Jerzy (2001). *Narracja jako sposób rozumienia świata*. W: Owczarek, Bogdan; Mitosek, Zofia; Grajewski, Wincenty red. *Praktyki opowiadania*. Kraków: Universitas, s. 87-126.
- Trzópek, Joanna (2010). „Kto jest kontrolerem?” *Problem podmiotu świadomego „ja” we współczesnej psychologii kognitywnej*. Roczniki Psychologiczne T. XIII, Nr 1, s. 129-154.
- Vaughan, Misha; Dillon, Andrew (2006). *Why structure and genre matter for users of digital information: a longitudinal experiment with readers of a web-based newspaper*. *International Journal of Human-Computer Studies* Vol. 64, No. 6, p. 502-526.
- Vaughan, Misha; Andrew Dillon (1998). *The role of genre in shaping our understanding of digital documents*. In: *Proceedings of 61<sup>st</sup> Annual Meeting of the American Society for Information Science (ASIS)*. Medford: Information Today Inc., p. 559-566.
- Waters, Lindsay (2009). *Scholarship and silence*. In: Greco, Albert ed. *The state of scholarly publishing. Challenges and opportunities*. New Brunswick: Transaction Publishers, p. 27-33.
- Wilson, Patrick (1983). *Second-hand knowledge: an inquiry into cognitive authority*. Westport: Greenwood Press, 210 p.
- Wilson, Thomas (1999). *Models in information behaviour research*. *Journal of Documentation* Vol. 55, No. 3, p. 249-270.
- Wilson, Thomas (1997). *Information behaviour: an interdisciplinary perspective*. *Information Processing & Management* Vol. 33, No. 4, p. 551-572.
- Wyrwas, Katarzyna (2002). *Skarga jako gatunek mowy*. Katowice: Wydaw. Uniwersytetu Śląskiego, 135 s.
- Zins, Chaim (2007). *Conceptual approaches for defining data, information, and knowledge*. *Journal of the Association for Information Science and Technology* Vol. 58, No. 4, p. 479-493.

---

# Otwarta nauka w świetle dokumentów Unii Europejskiej. Wybrane zagadnienia

Diana Pietruch-Reizes

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

## Wprowadzenie

We współczesnym myśleniu o nauce nie sposób pominąć rozważań dotyczących otwartego dostępu do treści naukowych i badawczych. Wielowymiarowe ujmowanie tego zagadnienia pozwala uwzględnić, uporządkować i wzajemnie powiązać wszystkie charakterystyczne dla niego procesy. Peter Suber, angielski filozof i znawca zagadnień otwartości w nauce, w kompendium „Open Access” pisze o swobodnym przekazywaniu wiedzy za pośrednictwem tekstów w postaci cyfrowej i globalnej sieci, o „rewolucji w dostępie” [Suber 2014, s. 15]. Nasuwa się w tym miejscu refleksja naukoznawcza wybitnego filozofa i socjologa Floriana Znanieckiego, wyrażona przed sześćdziesięciu laty, w 1954 r., w związku z obchodami 200-lecia Uniwersytetu Columbia, które poświęcono tematowi „Prawo człowieka do wiedzy i swobodne z niego korzystanie” [Znaniecki 1984, s. 505]. Znaniecki, pisząc o ewolucji twórczej i dyfuzji wiedzy, stwierdza: „Kiedy mówimy o przekazywaniu nowej wiedzy, musimy pamiętać, że tylko w ciągu ostatnich stu lat, wraz z rozwojem zorganizowanego szkolnictwa – od uniwersytetów do przedszkoli – stała się ona dostępna dla całego młodego pokolenia. W ten sposób ideał <prawa człowieka do wiedzy> staje się rzeczywistością” [Znaniecki 1984, s. 516].

W dzisiejszym cyfrowym świecie to właściwe wykorzystywanie potencjału technologii informacyjno-komunikacyjnej umożliwia udostępnianie wiedzy szerzej niż kilka dekad temu.

## Kontekst Europejskiej Przestrzeni Badawczej

W 2000 r. w Lizbonie podczas specjalnego posiedzenia Rady Europejskiej opracowano nową strategię dla Unii, uwzględniającą wyzwania gospodarki opartej na wiedzy. Podkreślając ważną rolę badań i rozwoju technologii w gospodarce, zwrócono uwagę na działania w kierunku ustanowienia Europejskiej Przestrzeni Badawczej (EPB), m.in. przez „ułatwienie utworzenia do końca 2001 r. bardzo szybkiej transeuropejskiej sieci do celów elektronicznej komunikacji naukowej, (...), łączącej instytucje badawcze i uniwersytety, jak również biblioteki naukowe, ośrodki naukowo-badawcze i – stopniowo – szkoły” [Posiedzenie Rady Europejskiej w Lizbonie 2000, pkt 13]. W ramach szóstego programu ramowego Wspólnoty Europejskiej w dziedzinie badań w obrębie priorytetu tematycznego Technologie społeczeństwa informacyjnego przewidywano działania ukierunkowane na rozwój narzędzi „tworzenia, organizowania, nawigacji, wyszukiwania, dzielenia, zachowywania i upowszechniania zawartości cyfrowej” [Decyzja nr 1513/2002/]. Europejska Przestrzeń Badawcza została wpisana do art. 163 ust. 1 traktatu lizbońskiego z 2007 r. w następującym brzmieniu: „1. Unia ma na celu wzmacnianie swojej bazy naukowej i technologicznej przez utworzenie europejskiej przestrzeni badawczej, w której naukowcy, wiedza naukowa i technologie podlegają swobodnej wymianie (...)” [Traktat z Lizbony..., dok. elektr.].

Komisja Europejska (KE), publikując w 2007 r. Zieloną Księgę na temat Europejskiej Przestrzeni Badawczej, przedstawiła nowe spojrzenie na EPB i postawiła „pytanie o zdolność Europy do utrzymania przewagi konkurencyjnej w dziedzinie wiedzy i innowacji, która stanowi trzon odnowionej strategii lizbońskiej na rzecz wzrostu gospodarczego i zatrudnienia” [Zielona Księga, s. 5]. Jednym z kluczowych wymiarów funkcjonowania EPB – obok światowej klasy infrastruktur naukowo-badawczych czy wzmocnienia instytucji naukowo-badawczych – jest zapewnienie skutecznego przepływu wiedzy. W tym zakresie zwrócono uwagę na otwarcie dostępu do wiedzy w całej Europie przez wykorzystanie potencjału technologii informacyjno-komunikacyjnych. Przede wszystkim chodzi o „niezawodny, niedrogi i stały dostęp do wyników badań naukowych, a także ich szerokie rozpowszechnianie” [Zielona Księga, s. 19]. W Zielonej Księdze KE postulowała tworzenie bibliotek internetowych, repozytoriów wiedzy naukowej, także baz publikacji i wyników badań finansowanych ze środków publicznych, zwracając uwagę na konieczność zintegrowania tych elementów na poziomie europejskim i połączenia z podobnymi bazami danych poza Unią. Szczególne znaczenie dla poziomu prowadzonych prac naukowo-badawczych ma system publikowania informacji naukowej, „UE powinna wspierać tworzenie «kontinuum» dostępnych i wzajemnie powiązanych informacji naukowych, od danych pierwotnych



po publikację, wewnątrz poszczególnych społeczności i państw oraz pomiędzy nimi” [Zielona Księga, s. 19]. Te kwestie poruszono we wcześniejszym Komunikacie KE dotyczącym informacji naukowej w epoce cyfrowej z lutego 2007 r. [Komunikat Komisji... w sprawie informacji naukowej 2007, dok. elektr.].

Pisząc o znaczeniu skutecznego przepływu wiedzy, KE wskazywała na konieczność zharmonizowania regulacji w odniesieniu do praw własności intelektualnej, w tym przyjęcia wspólnych zasad transferu wiedzy i współpracy pomiędzy publicznym sektorem naukowym a sektorem przemysłowym, ponieważ „główną przeszkodę stanowią (...) niespójne, często nieodpowiednie regulacje i sposoby zarządzania prawami własności intelektualnej powstałej za pieniądze publiczne” [Zielona Księga, s. 19]. W tym kontekście w Zielonej Księdze postawiono pytania:

„Czy istnieje potrzeba realizacji na szczeblu UE polityki i rozwiązań praktycznych w zakresie zapewnienia otwartego dostępu do danych pierwotnych i wzajemnie weryfikowanych publikacji wyników badań finansowanych ze środków publicznych oraz ich rozpowszechniania?

Jakie przepisy powinny się składać na europejskie regulacje dotyczące przepływu wiedzy pomiędzy instytucjami naukowo-badawczymi a sektorem przemysłowym w oparciu o dobrą praktykę i modele?

Jakie warunki należy stworzyć do propagowania nowatorskich sposobów komunikacji, nauczania, dyskusji i oceny osiągnięć naukowych i technicznych przez Europejczyków oraz absorpcji wiedzy w celu podejmowania decyzji w oparciu o dowody?” [Zielona Księga, s. 21].

## **Dostęp do informacji naukowych – nowe tendencje**

Przywołany już wcześniej Komunikat KE w sprawie informacji naukowej w epoce cyfrowej stanowił pierwszy, bardzo ważny krok w kierunku „stworzenia jednolitego obszaru informacji naukowej, obejmującego szereg społeczności naukowych i krajów oraz sięgającego od danych pierwotnych po gotowe publikacje” [Komunikat Komisji... w sprawie informacji naukowej... 2007, s. 3]. Zauważono nowe tendencje w zakresie dostępu do publikacji naukowych i do danych badawczych, „rozwój ruchu otwartego dostępu, wychodzącego z założenia, że w epoce Internetu dostęp do publikacji i danych powinien zostać udoskonalszy. Ruch ten ma na celu zapewnienie bezpośredniego i swobodnego dostępu internetowego do publikacji naukowych” [Komunikat Komisji... w sprawie informacji naukowej... 2007, s. 3]. W przyszłych działaniach Komisji przewidywano wspieranie lepszego dostępu do wyników badań finansowanych przez Wspólnotę, w ten sposób, że „koszty projektów związane z publikacją, w tym publikacją otwartego

dostępu, kwalifikować się będą do wspólnotowej pomocy finansowej”. Podkreślono kluczowe znaczenie Deklaracji Berlińskiej w sprawie otwartego dostępu do wiedzy w naukach ścisłych i humanistyce z 2003 r.

Działania UE na rzecz podwyższenia konkurencyjności gospodarki opartej na wiedzy wymagają nie tylko usprawnienia generowania wiedzy, ale także zwiększenia zasięgu rozpowszechniania i wymiany wyników badań naukowych finansowanych ze środków publicznych. Polityka KE w odniesieniu do dostępu do informacji naukowej, zapoczątkowana opublikowaniem komunikatu w lutym 2007 r., znalazła odzwierciedlenie w konkluzjach Rady w formie przyjętej na posiedzeniu Rady ds. Konkurencyjności w dniu 23 listopada 2007 r. Rada uznała, że „inicjatywy, służące opracowaniu trwałych modeli gwarantujących otwarty dostęp do informacji naukowych, mają strategiczne znaczenie dla rozwoju nauki w Europie” [Konkluzje Rady w sprawie informacji... 2007, s. 4]. Zaapelowano do państw członkowskich o „wzmocnienie krajowych strategii i struktur służących udostępnianiu i rozpowszechnianiu informacji naukowych” [Konkluzje Rady w sprawie informacji... 2007, s. 8] przez przyjęcie jasnych strategii rozpowszechniania informacji naukowych i ich udostępniania, „zwiększenie koordynacji między państwami członkowskimi w zakresie strategii i praktyk udostępniania i rozpowszechniania informacji naukowych (...)” [Konkluzje Rady w sprawie informacji... 2007, s. 9] oraz uwzględnienie w procesie przygotowywania ram prawnych specyficznych cech informacji naukowych.

Rada apelowała do KE o „eksperymentowanie z zapewnianiem otwartego dostępu do publikacji naukowych będących rezultatem projektów finansowanych na mocy ramowych programów badawczych UE (...)” [Konkluzje Rady w sprawie informacji... 2007, s. 10]. Także w komunikacie z 2009 r. dotyczącym strategicznej roli infrastruktury TIK [Komunikat Komisji... Infrastruktury TIK... 2009] jako ważnego elementu europejskiej polityki badań naukowych wskazano na poprawę dostępu do informacji naukowych, m.in. przez europejskie i krajowe e-infrastruktury, które muszą sprostać nowym wyzwaniom współczesnej nauki opartej na danych, poprzez utworzenie spójnego, dobrze zarządzanego ekosystemu repozytoriów informacji naukowej [Komunikat Komisji... Infrastruktury TIK... 2009, s. 11].

W kolejnych dokumentach stanowiących sedno strategii Europa 2020, przyjętej podczas szczytu Rady Europejskiej w czerwcu 2010 r., mianowicie w komunikatach „Unia innowacji” czy „Europejska agenda cyfrowa”, podkreślono znaczenie tych projektów dla stworzenia lepszych warunków dla innowacji, „upowszechniania, transferu i wykorzystania wyników badań, między innymi poprzez otwarty dostęp do publikacji i danych pochodzących z badań naukowych finansowanych ze środków publicznych” [Komunikat Komisji... Europa 2020 Unia innowacji, s. 12].

W ramach działań związanych z promowaniem otwartości i korzystaniem z twórczego potencjału Europy szczególne znaczenie ma tzw. piąta swoboda, czyli zagwarantowanie swobodnego przemieszczania się nie tylko pracowników naukowych, ale także innowacyjnych pomysłów. „Prawdziwie otwarte innowacje wymagają brokerstwa, pośrednictwa oraz tworzenia sieci (...). Konkurencyjne w skali międzynarodowej klastry odgrywają kluczową rolę w skupianiu w jednym miejscu – fizycznie i wirtualnie – wielkich przedsiębiorstw i MŚP, uczelni, ośrodków badawczych oraz społeczności naukowców i praktyków w celu wymiany wiedzy i pomysłów. Trzeba zintensyfikować transfer wiedzy pomiędzy przedsiębiorstwami a ośrodkami akademickimi i sprawić, by odbywał się on ponad granicami” [Komunikat Komisji... Europa 2020 Unia innowacji, s. 20]. Dlatego KE będzie promować otwarty dostęp do wyników badań finansowanych ze środków publicznych, jak również wspierać rozwój „inteligentnych usług informacji w dziedzinie badań naukowych” [Komunikat Komisji... Europa 2020 Unia innowacji, s. 22]. W tym kontekście zwrócono uwagę na wzmocnienie ochrony prawa własności intelektualnej przez opracowanie przez zainteresowane podmioty „zestawu modelowych umów o ustanowieniu konsorcjum, obejmujących różne opcje, od tradycyjnych, chroniących prawa własności intelektualnej, po bardziej otwarte” [Komunikat Komisji... Europa 2020 Unia innowacji, s. 22].

## **Prawa dostępu do istniejącej wiedzy i wyników**

Przyjęty 30 listopada 2011 r. unijny program finansowania badań i innowacji „Horyzont 2020” na lata 2014-2020 wprowadza otwarty dostęp do publikacji naukowych:

Art. 18

„1. Zapewnia się otwarty dostęp do publikacji naukowych będących wynikiem badań finansowanych ze środków publicznych w ramach programu «Horyzont 2020». Jest on realizowany zgodnie z rozporządzeniem (UE) nr 1290/2013.

2. Promuje się otwarty dostęp do danych naukowych będących wynikiem badań finansowanych ze środków publicznych w ramach programu «Horyzont 2020». Jest on realizowany zgodnie z rozporządzeniem (UE) nr 1290/2013.” [Rozporządzenie Parlamentu Europejskiego... ustanawiające „Horyzont 2020”].

W zamierzeniu KE takie rozwiązanie powinno służyć zwiększeniu potencjału innowacyjnego Europy. Zdaniem Neelie Kroes, wiceprzewodniczącej Komisji Europejskiej odpowiedzialnej za agendę cyfrową, „podatnicy nie powinni dwukrotnie płacić za badania naukowe i potrzebny im jest łatwy dostęp do danych pierwotnych. Chcemy pójść o krok dalej pod względem rozpowszechniania i wy-

korzystywania wyników badań naukowych. Informacja to nowa ropa”. Natomiast unijna komisarz ds. badań, innowacji i nauki, Máire Geoghegan-Quinn stwierdziła, że „podatnikom należy się coś więcej. Otwarty dostęp do danych i opracowań naukowych sprawi, że istotne przełomy w nauce i biznesie będą dokonywać się szybciej, a wiedza i konkurencyjność w Europie zostaną wzmocnione” [Dane naukowe... 2012].

Rozporządzenie dotyczące zasad wykorzystywania i upowszechniania wyników działań podejmowanych w ramach programu „Horyzont 2020” [Rozporządzenie Parlamentu Europejskiego...ustanawiające zasady uczestnictwa...dla programu „Horyzont 2020”] określa prawa dostępu do istniejącej wiedzy i wyników (Sekcja II, art. 45-49). Uczestnicy określają istniejącą wiedzę, którą wykorzystują w swym działaniu, w dowolny sposób, w formie umowy pisemnej. Przepisy regulują kwestie zasad w zakresie praw dostępu (art. 46), w tym:

- zastrzeżenia formy pisemnej w odniesieniu do każdego wniosku o skorzystanie z praw dostępu bądź każdego zrzeczenia się praw dostępu;
- prawa dostępu nie obejmują prawa do udzielania sublicencji, chyba że właściciel wyników lub istniejącej wiedzy, do których dostępu dotyczy wniosek, wyrazi zgodę;
- uczestnik – przed przystąpieniem do umowy o udzielenie dotacji – jest zobowiązany poinformować pozostałych uczestników tego samego działania o istniejących ograniczeniach prawnych bądź limitach w zakresie udzielania dostępu do posiadanej przez niego istniejącej wiedzy;
- sytuacja, w której uczestnik kończy uczestnictwo w działaniu, nie ma wpływu na jego zobowiązanie do udzielania dostępu na warunkach określonych w umowie o udzielenie dotacji;
- umowa konsorcyjna może przewidywać, że w sytuacji, w której uczestnik narusza zobowiązania i nie naprawia takiego uchybienia, traci prawa dostępu.

Przepisy rozporządzenia regulują także prawa dostępu w zakresie realizacji (art. 47), w tym:

- przysługiwanie uczestnikowi praw dostępu do wyników innego uczestnika tego samego działania, gdy jest to uwarunkowane potrzebą dostępu do takich wyników w związku z celami wykonania swoich zadań w ramach działania, przy czym dostęp w takim przypadku udzielany jest nieodpłatnie;
- przysługiwanie uczestnikowi praw dostępu do istniejącej wiedzy innego uczestnika tego samego działania, gdy jest to uwarunkowane potrzebą dostępu do takiej wiedzy w związku z celami wykonania swoich zadań w ramach działania – przepis stanowi, że taki dostęp udzielany jest nieodpłatnie, chyba że uczestnicy jeszcze przed zawarciem umowy o udzielenie dotacji przyjęli inne uzgodnienia.

W dalszych przepisach znajdujemy regulacje prawa dostępu w zakresie wykorzystania (art. 48), w tym:

- przysługiwanie uczestnikowi praw dostępu do wyników innego uczestnika tego samego działania, gdy jest to uwarunkowane potrzebą dostępu do takich wyników w związku z celami wykorzystania swoich wyników – przepis stanowi, że „z zastrzeżeniem umowy, takiego dostępu udziela się na sprawiedliwych i rozsądnych warunkach”;
- przysługiwanie uczestnikowi praw dostępu do istniejącej wiedzy innego uczestnika tego samego działania, gdy jest to uwarunkowane potrzebą dostępu do takiej wiedzy w związku z celami wykorzystania swoich wyników;
- przysługiwanie podmiotowi powiązanemu, którego siedziba znajduje się w państwie członkowskim lub w państwie stowarzyszonym, praw dostępu do wyników i – z zastrzeżeniem wszelkich ograniczeń na mocy art. 46 ust. 3 – do istniejącej wiedzy, na sprawiedliwych i rozsądnych warunkach, gdy jest to uwarunkowane potrzebą dostępu do wyników i wiedzy w związku z korzystaniem z wyników uzyskanych przez uczestnika, z którym dany podmiot jest powiązany, chyba że umowa konsorcyjna stanowi inaczej; przepisy stanowią, że „o uzyskanie takich praw dostępu należy się zwrócić bezpośrednio do uczestnika będącego właścicielem wyników lub istniejącej wiedzy, chyba że uzgodniono inaczej zgodnie z art. 46 ust. 2.”, natomiast wniosek o dostęp można złożyć w okresie do jednego roku od zakończenia działania, chyba że uzgodniono inny termin.

Ostatnią grupę przepisów omawianego rozporządzenia stanowią prawa dostępu dla Unii i państw członkowskich (art. 49). W tym przypadku prawa przysługują instytucjom, organom, biurom lub agencjom UE w związku z tworzeniem, wdrażaniem i monitorowaniem polityki lub programów Unii. Wymienione instytucje otrzymują nieodpłatnie ograniczone do zastosowania niehandlowego i niekonkurencyjnego prawa dostępu wyłącznie do wyników uczestnika, który otrzymał finansowanie unijne.

## **Poprawa dostępu do informacji naukowej**

W komunikacie pt. „W stronę lepszego dostępu do informacji naukowej – zwiększanie korzyści z inwestowania środków publicznych w badania naukowe” [Komunikat Komisji ... W stronę lepszego dostępu... 2012] określono działania KE zmierzające do poprawy dostępu do informacji naukowej i wzmocnienia korzyści wynikających z inwestowania środków publicznych w badania naukowe. Wskazano także sposoby wdrażania polityki otwartego dostępu w ramach

programu „Horyzont 2020”. Procesy te wymagają współdziałania państw członkowskich, instytucji finansujących badania naukowe, instytucji realizujących te badania, wydawnictw naukowych, uczelni wraz z bibliotekami akademickimi, innowacyjnych gałęzi przemysłu oraz społeczeństw.

Prowadzenie nowoczesnych badań wymaga intensywnej wymiany wiedzy naukowej, pełniejszego i szerszego dostępu do informacji naukowych. To wpływa korzystnie na wzrost innowacyjności, rozszerzenie i wzmocnienie współpracy naukowej, a zarazem zmniejsza ryzyko powielania badań. Wzrasta także poziom otwartości i przejrzystości procesu naukowo-badawczego.

Przyjęta przez KE strategia w odniesieniu do otwartych danych i obiegu wiedzy zakłada, że „za informację opłaconą już z kiesy publicznej nie powinno się płacić ponownie za każdym razem, gdy się ją pozyskuje lub wykorzystuje, oraz że powinna ona w pełni służyć europejskim przedsiębiorstwom i obywatelom” [Komunikat Komisji... W stronę lepszego dostępu... 2012, s. 4]. Innymi słowy, otwarcie dostępu dla badaczy i obywateli do informacji naukowej uzyskanej dzięki finansowaniu ze środków publicznych w sieci powinno następować bez dodatkowych opłat, z wykorzystaniem trwałych e-infrastruktur.

Rosnące ceny czasopism naukowych i ograniczenie budżetów bibliotek uniwersyteckich i instytucji badawczych stały się impulsem do stworzenia rozwiązań idących w kierunku otwartego dostępu – „modelu, który zapewnia czytelnikom bezpłatny dostęp oraz możliwość wykorzystywania i ponownego wykorzystywania danych pobranych z internetu” [Komunikat Komisji... W stronę lepszego dostępu... 2012, s. 5]. Wyróżnia się dwa podstawowe modele otwartego dostępu:

1. Otwarty dostęp do publikacji naukowych – droga złota (otwarta publikacja); w tym modelu koszty publikacji zostają przeniesione z czytelników na autorów, najczęściej pokrywają je uczelnie lub instytucje badawcze, z którymi związany jest dany naukowiec, bądź inne podmioty odpowiedzialne za finansowanie określonych badań.

2. Otwarty dostęp do publikacji naukowych – droga zielona (samodzielna archiwizacja); w tym modelu opublikowany artykuł lub wersja artykułu po ostatecznej recenzji jest zamieszczana przez autora w otwartym internetowym repozytorium przed tradycyjną publikacją lub równoległe z nią. W modelu tym występują liczne warianty, będące wynikiem np. porozumień między autorem i wydawcą, przewidzianych w umowie wydawniczej.

W ostatnich latach coraz częściej można zaobserwować tendencje, których wyrazicielami są coraz liczniejsze instytucje finansujące badania i uczelnie, w kierunku zapewnienia przez badaczy otwartego dostępu do wyników badań finansowanych ze środków publicznych. Na podstawie danych KE „obecnie ok. 20% wszystkich artykułów naukowych jest dostępnych w trybie otwartego dostępu,

z czego 60% objętych jest modelem zielonej drogi. Niektórzy wydawcy oferują «czasopisma hybrydowe», zawierające nie tylko artykuły, za których publikację autor uiszczył opłatę (i z których w związku z tym czytelnicy mogą korzystać za darmo), ale także artykuły dostępne jedynie dla prenumeratorów lub na zasadzie *pay-per-view* (zapłać, aby oglądać)” [Komunikat Komisji ... W stronę lepszego dostępu... 2012, s. 5; Björk et al. 2009, dok. elektr.].

## Dostęp do wyników badań naukowych

Kwestie dostępu do wyników badań naukowych i danych badawczych wiążą się z poważnym problemem długoterminowej konserwacji zasobów cyfrowych. Dynamicznie postępujące zmiany w technologii informacyjnej, nowe rozwiązania w zakresie infrastruktury sprzętowej i sieciowej, trendy w rozwoju rozwiązań dostępowych i komunikacji elektronicznej wpływają na niestabilność informacji cyfrowej. W przypadku publikacji i danych badawczych jest to kwestia kluczowa, np. ze względu na możliwości odtwarzania i powtarzalność eksperymentów. Dla innowacji, stanowiących podstawę rozwoju gospodarczego, decydujące znaczenie mają szybkie postępy nauki. Dlatego potrzebne są wydajne systemy obliczeniowe, za pomocą których można prowadzić obliczenia złożonych symulacji dotyczących np. prognozowania pogody, modelowania procesów zachodzących w mózgu czy przetwarzać wyniki doświadczeń naukowych. Jako przykład może posłużyć działalność Centrum Nauk Obliczeniowych, stanowiącego część Interdyscyplinarnego Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego.

W kontekście rozwoju e-nauki opartej na dużej ilości danych, będących skutkiem wzrostu wartości surowych danych obserwacyjnych i doświadczalnych w prawie wszystkich dziedzinach nauki (nauki humanistyczne, bioróżnorodność, fizyka cząstek elementarnych, astronomia itd.), konserwacja cyfrowej informacji naukowej, począwszy od danych pierwotnych poprzez analizę aż do publikacji wyników badań, jest nie do przecenienia.

Chodzi przecież o czytelność danych i ich prawidłową interpretację przez przyszłe pokolenia.

Na poziomie Unii Europejskiej badanie zagadnień dostępu do informacji naukowej, jej rozpowszechniania i konserwacji zainicjowano m.in. poprzez projekty współfinansowane w ramach szóstego programu ramowego (6PR), np. CASPAR – Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval (<http://www.casparpreserves.eu/caspar-project.html>), dotyczący ochrony naukowych zasobów cyfrowych, usług wspierających wirtualizację

długoterminowej ochrony tych zasobów, pomimo zmian w stosowanym sprzęcie i oprogramowaniu; DRIVER – Digital Repository Infrastructure Vision for European Research ([http://www.driver-repository.eu/PublicDocs/D8.4\\_1.pdf](http://www.driver-repository.eu/PublicDocs/D8.4_1.pdf)), poświęcony sposobom łączenia repozytoriów informacji naukowej; SEADATANET – a Pan-European Infrastructure for Ocean and Marine Data Management (<http://www.seadatanet.org>), projekt mający na celu utworzenie paneuropejskiej infrastruktury zarządzania danymi morskimi, opracowanie systemu integrującego krajowe repozytoria danych morskich.

Wyniki badań przeprowadzonych w ramach projektu PARSE.Insight: Permanent Access to the Records of Science in Europe, dotyczące publikowania danych „wykazały, że jedynie 25% prowadzących badania udostępnia publicznie swoje dane podstawowe, 11% udostępnia je badaczom zajmującym się tą samą dyscypliną, a 58% – jedynie w ramach własnej grupy badawczej” [Komunikat Komisji... W stronę lepszego dostępu... 2012, s. 6].

Z analizy dokumentów sprawozdawczych KE dotyczących Europejskiej Przestrzeni Badawczej wynika, że europejskie uniwersytety, instytucje badawcze oraz podmioty finansujące badania wyraziły aprobatę dla otwartego dostępu do publikacji wyników badań. Wykazywały duże zainteresowanie wykorzystaniem internetowej platformy informacyjnej ERAWATCH (<http://erawatch.jrc.ec.europa.eu/>), stworzonej w celu wspierania trwającego procesu tworzenia Europejskiej Przestrzeni Badawczej, platformy udostępniającej informacje o najnowszych dokumentach, programach badawczych, agencjach finansujących badania, wynikach badawczych, także informacje w formie wskaźników dotyczących np. publikacji czy patentów. Wśród dokumentów zawierających m.in. informacje na temat otwartego dostępu ERAWATCH udostępnia raporty krajowe zawierające analizy i oceny krajowych polityk w zakresie badań i innowacji na tle strategii i celów UE, ze szczególnym uwzględnieniem priorytetów Europejskiej Przestrzeni Badawczej. Przykładem może być raport na temat sytuacji w Polsce, w którym wskazano na główne zmiany w odniesieniu do otwartego dostępu, mianowicie: przygotowanie i opublikowanie w 2012 r. założeń do projektu ustawy o otwartych zasobach publicznych<sup>1</sup>; projekt SYNAT (System Nauki i Techniki) realizowany w latach 2010–2013 przez sieć naukową zrzeszającą 16 instytucji krajowych,

---

<sup>1</sup> Projekt został wycofany ze względu na przyjęcie „(...) nowego kierunku pracy w związku z uchwaleniem przez Parlament Europejski i Radę Dyrektywy 2013/37/UE z dn. 26 czerwca 2013 r. zmieniającej Dyrektywę 2003/98/WE w sprawie ponownego wykorzystania informacji sektora publicznego. Zasadnicza zmiana dyrektywy polega na poszerzeniu zakresu informacji sektora publicznego podlegających ponownemu wykorzystywaniu, tj. będących w posiadaniu bibliotek, w tym uniwersyteckich, archiwów państwowych i muzeów, a więc tych, które m.in. miały być udostępniane do wykorzystywania na podstawie przepisów w projektowanych założeniach ustawy o otwartych zasobach publicznych” [Biuletyn Informacji Publicznej Rady Ministrów... , dok. elektr.].



finansowany przez Narodowe Centrum Badań i Rozwoju (celem projektu było utworzenie uniwersalnej, otwartej, repozytoryjnej platformy hostingowej i komunikacyjnej dla sieciowych zasobów wiedzy dla nauki, edukacji i otwartego społeczeństwa wiedzy); Wirtualna Biblioteka Nauki, projekt realizowany przez Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego ze środków Ministerstwa Nauki i Szkolnictwa Wyższego; program „Index Plus”, którego celem jest nadanie międzynarodowego charakteru polskim czasopismom naukowym; liczne oddolne inicjatywy instytucji naukowych na rzecz konsolidacji i modernizacji infrastruktury naukowo-badawczej i informatycznej jednostek naukowych poprzez tworzenie repozytoriów otwartego dostępu [Klincewicz 2014, s. 58].

## Zakończenie

Optymalizacja przepływu wiedzy naukowej, dostępu do niej i jej przekazywania, w tym za pośrednictwem cyfrowej Europejskiej Przestrzeni Badawczej, wymaga uruchomienia przez państwa członkowskie określonych procedur w celu wsparcia otwartego dostępu do publikacji naukowych. Według danych KE „20 państw członkowskich podjęło konkretne środki w celu wsparcia otwartego dostępu do publikacji naukowych, lecz tylko pięć ma konkretne przepisy dotyczące otwartego dostępu do danych naukowych” [Komunikat... Europejska przestrzeń badawcza 2014, s. 7-8]. Istnieją bariery o charakterze prawnym, technicznym, finansowym, społeczno-kulturowym i związanym z zaufaniem. Występują różnice w krajowych politykach, inicjatywach, a także praktykach związanych z otwartym dostępem, „niektóre z nich nie odzwierciedlają we właściwy sposób unijnej definicji otwartego dostępu” [Komunikat... Europejska przestrzeń badawcza 2014, s. 7]. W kontekście badań i rozwoju otwarty dostęp najczęściej odnosi się do dostępu do informacji naukowych, w szczególności do:

- recenzowanych w ramach wzajemnej weryfikacji artykułów naukowych (publikowanych w czasopismach akademickich) oraz
- danych stanowiących wyniki badań naukowych (danych leżących u podstaw publikacji lub surowych danych).

Badania naukowe wymagają coraz częściej bardzo dużych zbiorów danych, ogromnych mocy obliczeniowych i odpowiedniej wizualizacji danych oraz wyników badań. Rozwój e-nauki (otwarta nauka, Nauka 2.0) wykorzystuje e-infrastruktury, czyli środowiska umożliwiające dzielenie się zasobami naukowymi, gwarantujące dostęp do danych z dowolnego miejsca, wspierające współpracę naukową w wymiarze międzynarodowym oraz jakość i wydajność badań naukowych.

## BIBLIOGRAFIA

- Biuletyn Informacji Publicznej Rady Ministrów. Wykaz prac legislacyjnych i programowych Rady Ministrów, 9 maja 2014 [dok. elektr.]. <http://bip.kprm.gov.pl/kpr/form/r228,dok.html> [odczyt 5.10.2014].
- Björk, Bo-Christer et al. (2009). *Open Access to Scientific Journal Literature: Situation 2009* [dok. elektr.]. <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0011273> [odczyt 5.10.2014].
- Dane naukowe: otwarty dostęp do wyników badań przyczyni się do zwiększenia potencjału innowacyjnego Europy*. Komisja Europejska. Komunikat prasowy, Bruksela, dnia 17 lipca 2012 r. [dok. elektr.]. [http://europa.eu/rapid/press-release\\_IP-12-790\\_pl.htm](http://europa.eu/rapid/press-release_IP-12-790_pl.htm) [odczyt 5.10.2014].
- Decyzja nr 1513/2002/WE Parlamentu Europejskiego i Rady z dnia 27 czerwca 2002 r. dotycząca szóstego programu ramowego Wspólnoty Europejskiej w dziedzinie badań, rozwoju technologicznego i demonstracji, przyczyniającego się do utworzenia Europejskiej Przestrzeni Badawczej i innowacji (2002-2006). Załącznik 1. Cele naukowe i technologiczne, ogólne kierunki działań i priorytety. 1.1.2. Technologie społeczeństwa informacyjnego (IST). Zarządzanie informacją i interfejs informacyjny pkt a) [dok. elektr.]. <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:32002D1513&rid=1> [odczyt 5.10.2014].
- Dokument roboczy służb Komisji. Streszczenie oceny skutków towarzyszące dokumentowi: Zalecenie w sprawie dostępu do informacji naukowej oraz jej ochrony {C(2012) 4890 final}{SWD(2012) 222 final}(2012) [dok. elektr.]. <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52012SC0221&from=PL> [odczyt 5.10.2014].
- Klincewicz, Krzysztof (2014). *ERAWATCH Country Reports 2013. Poland* [dok. elektr.]. [http://erawatch.jrc.ec.europa.eu/erawatch/export/sites/default/galleries/generic\\_files/file\\_0532.pdf](http://erawatch.jrc.ec.europa.eu/erawatch/export/sites/default/galleries/generic_files/file_0532.pdf) [odczyt 5.10.2014].
- Komunikat Komisji dla Parlamentu Europejskiego, Rady i Europejskiego Komitetu Ekonomiczno-Społecznego w sprawie informacji naukowej w epoce cyfrowej: dostęp, rozpowszechnianie i konserwacja{SEC(2007)181} /\*COM/2007/0056 końcowy\*/ (2007) [dok. elektr.]. <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52007DC0056&from=PL> [odczyt 5.10.2014].
- Komunikat Komisji dla Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów. W stronę lepszego dostępu do informacji naukowej. Zwiększanie korzyści z inwestowania środków publicznych w badania naukowe /\*COM/2012/0401 final\*/ (2012) [dok. elektr.]. <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52012DC0401&qid=1414268686139&from=PL> [odczyt 5.10.2014].
- Komunikat Komisji do Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów – Infrastruktury TIK dla e-nauki (2009) [dok. elektr.]. <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52009DC0108&qid=1414256471081&from=PL> [odczyt 5.10.2014].
- Komunikat Komisji do Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów. Projekt przewodni strategii Europa 2020. Unia innowacji SEC(2010) 1161/\*COM/2010/0546 końcowy\*/ (2010) [dok.

- elektr.]. <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52010D-C0546&qid=1414257666966&from=PL> [odczyt 5.10.2014]
- Komunikat Komisji do Rady i Parlamentu Europejskiego. Europejska przestrzeń badawcza. Sprawozdanie z postępu prac za 2014 r. COM(2014) 575 final [dok. elektr.]. <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52014DC0575&qid=1414316742254&from=PL> [odczyt 5.10.2014].
- Komunikat Komisji do Rady i Parlamentu Europejskiego. Europejska przestrzeń badawcza. Sprawozdanie z postępu prac za 2014 r. [dok. elektr.]. [http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52014DC0575&qid=1414316742254&from=PL/\\*COM/2014/0575 final \\*/](http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52014DC0575&qid=1414316742254&from=PL/*COM/2014/0575%20final%20*/) [odczyt 5.10.2014].
- Konkluzje Rady w sprawie informacji naukowych w erze cyfrowej w formie przyjętej na posiedzeniu Rady ds. Konkurencyjności w dniu 23 listopada 2007 r. [dok. elektr.]. <http://register.consilium.europa.eu/doc/srv?l=PL&f=ST%2015362%202007%20INIT> [odczyt 5.10.2014]
- Posiedzenie Rady Europejskiej w Lizbonie, 23 i 24 marca 2000 r. Wnioski Prezydencji [dok. elektr.]. [http://archiwum-ukie.polskawue.gov.pl/HLP/files.nsf/0/A6EEAD-75D21463E1C1256F4F0038454C/\\$file/1PL.pdf](http://archiwum-ukie.polskawue.gov.pl/HLP/files.nsf/0/A6EEAD-75D21463E1C1256F4F0038454C/$file/1PL.pdf) [odczyt 5.10.2014].
- Przewodnik po otwartej nauce* (2009) [dok. elektr.]. <https://otwartanauka.pl/images/PDFs/przewodnik-po-otwartej-nauce.pdf> [odczyt 5.10.2014].
- Rozporządzenie Parlamentu Europejskiego i Rady (UE) nr 1290/2013 z dnia 11 grudnia 2013 r. ustanawiające zasady uczestnictwa i upowszechniania dla programu „Horyzont 2020” – programu ramowego w zakresie badań naukowych i innowacji (2014-2020) oraz uchylające rozporządzenie (WE) nr 1906/2006. Tekst mający znaczenie dla EOG Dz.U. L 347 z 20.12.2013, str. 81-103 [dok. elektr.]. <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:32013R1290&qid=1414266458702&from=PL> [odczyt 5.10.2014].
- Rozporządzenie Parlamentu Europejskiego i Rady nr 1291/2013 z dnia 11 grudnia 2013 r. ustanawiające „Horyzont 2020” – program ramowy w zakresie badań naukowych i innowacji (2014-2020) oraz uchylające decyzję nr 1982/2006/WE [dok. elektr.]. <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:32013R1291&from=PL> [odczyt 5.10.2014].
- Siewicz, Krzysztof (2012). *Otwarty dostęp do publikacji naukowych. Kwestie prawne* [dok. elektr.]. [http://pon.edu.pl/images/plon\\_publications/files/12\\_K\\_Siewicz\\_Otwarty\\_dostep\\_do\\_publicacji\\_naukowych.pdf](http://pon.edu.pl/images/plon_publications/files/12_K_Siewicz_Otwarty_dostep_do_publicacji_naukowych.pdf) [odczyt 5.10.2014].
- Siewicz, Krzysztof (2013). *Wolne oprogramowanie w instytucjach publicznych. Kwestie prawne* [dok. elektr.]. [http://otwartanauka.pl/images/PDFs/Wolne\\_oprogramowanie\\_w\\_instytucjach\\_publicznych.pdf](http://otwartanauka.pl/images/PDFs/Wolne_oprogramowanie_w_instytucjach_publicznych.pdf) [odczyt 5.10.2014].
- Suber, Peter (2014). *Otwarty dostęp* [dok. elektr.]. Warszawa: Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego. [http://pon.edu.pl/images/plon\\_publications/files/14\\_Peter\\_Suber\\_-\\_Otwarty\\_dostep.pdf](http://pon.edu.pl/images/plon_publications/files/14_Peter_Suber_-_Otwarty_dostep.pdf) [odczyt 5.10.2014].
- Szprot, Janusz red. (2014). *Otwarta nauka w Polsce 2014. Diagnoza* [dok. elektr.]. [http://pon.edu.pl/images/plon\\_publications/files/13\\_Otwarta%20nauka%20w%20Polsce%202014\\_Diagnoza.pdf](http://pon.edu.pl/images/plon_publications/files/13_Otwarta%20nauka%20w%20Polsce%202014_Diagnoza.pdf) [odczyt 5.10.2014].

Traktat z Lizbony zmieniający Traktat o Unii Europejskiej i Traktat ustanawiający Wspólnotę Europejską, sporządzony w Lizbonie dnia 13 grudnia 2007 r. Dz.U. 2009 nr 203 poz. 1569 [dok. elektr.]. <http://isap.sejm.gov.pl/DetailsServlet?id=WDU20092031569> [odczyt 5.10.2014].

Zielona Księga – Europejska Przestrzeń Badawcza: Nowe perspektywy KOM(2007) 161 wersja ostateczna [dok. elektr.]. <http://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52007DC0161&rid=4> [odczyt 5.10.2014].

Znanięcki, Florian (1984). *Spoleczne role uczonych*. Warszawa: PWN, 618 s.

---

Część II

**Zielona droga OA.  
Zasoby obiektów cyfrowych**

---

---

# Archiwa jakościowych danych badawczych w internecie

Sabina Cisek

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

## Wprowadzenie

Artykuł poświęcony jest archiwom jakościowych danych badawczych (ang. *qualitative data archives*), które stanowią specyficzny typ zasobów naukowych w internecie. Takie archiwa mogą oczywiście mieć również postać tradycyjną i być przechowywane „gdziekolwiek”, np. w szufladzie biurka badacza, rzecz jednak w tym, by standardowo stały się dostępne *online* w formie cyfrowej. Albo przynajmniej, w trudniejszych przypadkach, by informacja o ich istnieniu i miejscu przechowywania zamieszczona została w sieci, w odpowiednich specjalistycznych serwisach. Autorzy piszący o kolekcjach danych z reguły mają na myśli nowoczesną, zdigitalizowaną ich postać, takie znaczenie obowiązuje również w niniejszym tekście.

Archiwa danych jakościowych, zwane też czasami repozytoriami (ang. *qualitative data repositories*), stanowią podgrupę archiwów danych badawczych (ang. *research data archives / repositories*), dla których z kolei kategorią nadrzędną są zbiory danych *online* (ang. *data collections / datasets*). Ogólna charakterystyka i typologia kolekcji danych w internecie, stanowiąca kontekst dla opisu archiwów jakościowych danych badawczych, znajduje się w części pierwszej tego artykułu. Tam też zamieszczono wykaz dedykowanych tym zasobom serwisów wyszukiwawczych. Część druga dotyczy ściśle danych jakościowych i ich zbiorów, a także towarzyszących im wybranych problemów, zwłaszcza metodologicznych. Inne zagadnienia (etyczne, prawne, techniczne, z zakresu organizacji wiedzy) są jedynie zasygnalizowane. Podano również przykłady rzeczywiście funkcjonujących archiwów jakościowych danych badawczych. Zakończenie zawiera m.in. propo-

zycję wykorzystania tychże repozytoriów w badaniach i edukacji z zakresu bibliologii i informatologii.

Zagadnienie archiwizowania danych jakościowych oraz dzielenia się nimi (ang. *data sharing problem*) zainteresowało badaczy na szerszą skalę już w latach 90. XX wieku. Początkowo na ten temat pisali głównie – i nadal to robią – przedstawiciele nurtu jakościowego w naukach humanistycznych i społecznych, zwłaszcza socjologowie. W XXI wieku do grona autorów we wskazanym zakresie dołączyli informatolodzy i bibliotekoznawcy. Obecnie problematyka archiwów danych, w tym jakościowych, jest stosunkowo często dyskutowana w zagranicznej literaturze przedmiotu, m.in. informatologiczne czasopismo *IASSIST Quarterly* poświęciło temu zagadnieniu poczwórny numer specjalny [IASSIST 2010/2011]. Interesujący i wielostronny przegląd kwestii związanych z archiwizowaniem danych badawczych (dowolnych) przedstawiły także, na łamach *Annual Review of Information Science and Technology*, Stacy Kowalczyk i Kalpana Shankar [Kowalczyk, Shankar 2011]. W Polsce natomiast wybrane aspekty funkcjonowania archiwów danych jakościowych, w tym związane z nimi dylematy metodologiczne, rozważali w czasopiśmie *ASK* badacze związani z Instytutem Filozofii i Socjologii PAN [Filipkowski 2005; Kościański 2005; Palska 2005]. Z kolei przedstawicielka środowiska bibliotekarskiego, Bożena Bednarek-Michalska [Bednarek-Michalska 2012], wypowiedziała się na temat repozytoriów surowych danych badawczych w kontekście m.in. otwartej nauki, publicznego dostępu, a także spodziewanej w przyszłości aktywnej roli bibliotekarzy w tym obszarze.

Archiwizowanie badawczych danych jakościowych oznacza celowe, przemyślane ich gromadzenie, opracowanie, przechowywanie i udostępnianie, a także informowanie o ich istnieniu i zachęcanie do wykorzystywania. Są to typowe czynności z zakresu bibliotekarstwa i działalności informacyjnej, co więcej, dotyczą treści poznawczych utrwalonych na nośnikach materialnych (czyli informacji naukowej), w związku z czym bibliotekarze i informatolodzy, także w Polsce, powinni się tą sprawą zainteresować. Niniejszy artykuł ma się do tego przyczynić. Ponieważ jednak wskazana problematyka jest rozległa i wielowątkowa, całościowe, wyczerpujące jej omówienie w jednym, stosunkowo krótkim tekście nie jest możliwe. Skupiono się zatem na aspektach metodologicznym oraz naukowym (poznawczym) archiwów danych jakościowych, pozostałe jedynie sygnalizując i pozostawiając do odrębnych, pogłębianych rozważań.

Warto również zauważyć, iż słowo „jakościowy” ma kilka odmiennych znaczeń. Mianowicie, może m.in. wskazywać na pewien specyficzny sposób poznawania świata w nauce (badanie jakościowe, metodologia jakościowa), może też posiadać zupełnie inny, aksjologiczny, wartościujący charakter, związany z oceną jakości, np. archiwów albo serwisów www.

## Dane badawcze i ich zbiory w internecie – charakterystyka, typologia, wyszukiwanie

Problematyka zasobów danych w internecie jest bardzo obszerna, wiąże się m.in. z zagadnieniami *Big Data* i *data mining*, kompleksowa charakterystyka tematu wymagałaby zatem osobnej, jemu tylko poświęconej publikacji. W artykule pominięta została także ogólna, wielowymiarowa dyskusja nad różnymi sposobami rozumienia terminu „dane” oraz jego relacjami do pokrewnych „informacji” i „wiedzy”, wskazano jedynie pewne podstawowe wątpliwości. Natomiast pojęcie danych **jakościowych**, jako centralne dla bieżących rozważań, jest szczegółowo omówione w następnym rozdziale.

Na użytek tego artykułu przyjmuje się, zgodnie zresztą z duchem literatury przedmiotu [Bednarek-Michalska 2012; Corti 2011], iż obiektem zainteresowania są tzw. **dane surowe** (ang. *raw data*), czyli w niewielkim stopniu przetworzone, jak np. dokumenty finansowe, filmy, fotografie, nagrania audio, notatki, pamiętniki, rysunki, statystyki demograficzne i ekonomiczne, transkrypcje wywiadów, wykazy wydatków organów państwowych i samorządowych, wypełnione ankiety, zapisy z instrumentów astronomicznych, laboratoryjnych albo monitoringu, zestawienia numeryczne i inne. Tak rozumiane dane przeciwstawione są opracowaniom, czyli tradycyjnym publikacjom naukowym, fachowym, publicystycznym, jak artykuły, eseje, książki, podręczniki, raporty, referaty itp.

Oczywiście „surowe dane” to metafora, określenie nieprecyzyjne. Z metodologicznego punktu widzenia „czyste”, całkowicie nieprzetworzone dane nie istnieją, co więcej, można wyróżnić kolejne „stopnie” ich postępującego przetworzenia [Kowalczyk, Shankar 2011, p. 255].

Dodatkowo, określenie tego, co może być kwalifikowane jako „dane”, jest względne. Zasoby statystyczne, np. takie, jakie oferuje europejski serwis Eurostat <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home>, bez wzbudzenia kontrowersji można nazwać danymi, zawartość Google Maps albo kolekcje fotografii w Narodowym Archiwum Cyfrowym <http://www.nac.gov.pl> – zapewne też, natomiast jak potraktować, przykładowo, pełne teksty *Monitora Sądowego i Gospodarczego* <https://ems.ms.gov.pl>, obejmujące bieżące ogłoszenia spółek handlowych, postanowienia o upadłości itp.? Przy odpowiednio sformułowanym problemie badawczym właściwie **wszystko** może stać się danymi, m.in. wpisy na forach dyskusyjnych, posty w blogach, a nawet publikacje naukowe (np. dla kogoś, kto prowadzi badania bibliometryczne). Kowalczyk i Shankar [Kowalczyk, Shankar 2011, p. 250-251] posługują się następującym sformułowaniem, określając dane jako „każdą informację, która może być przechowywana w formie elektronicznej, w tym teksty, dane liczbowe, grafikę, nagrania wideo i filmy, nagrania



audio, oprogramowanie, algorytmy, równania, animacje, modele, symulacje etc. Tego rodzaju dane mogą być tworzone na różne sposoby, w tym metodą obserwacji, kalkulacji lub eksperymentu”<sup>1</sup>. Generalnie w internecie istnieje ogromna liczba zróżnicowanych zbiorów danych, dających się wyodrębnić spośród innych zasobów jako mniej lub bardziej spójne całości. Niektóre z nich powstały **nieintencjonalnie**, „przy okazji”, np. usługa Google Książki stanowi właściwie ogromny korpus językowy, który dodatkowo kilka lata temu „dorobił się” własnego narzędzia analitycznego o nazwie Ngram Viewer <https://books.google.com/ngram>. Inne zostały zbudowane **celowo**, jak – by pozostać w tym samym kręgu – Narodowy Korpus Języka Polskiego <http://www.nkjp.pl>.

Z innego punktu widzenia zbiory danych w internecie można podzielić na **otwarte i zamknięte**. Te pierwsze są udostępnione publicznie, bez ograniczeń, drugie natomiast cechują się dostępem mniej lub bardziej restrykcyjnym.

Jeżeli chodzi o **typy** kolekcji danych, to najbardziej widoczne są trzy:

- wspomniane już korpusy językowe,
- (otwarte) dane badawcze / naukowe,
- (otwarte) dane publiczne (często produkowane przez rządy państw, instytucje i urzędy krajowe oraz organizacje międzynarodowe, takie jak Bank Światowy, OECD, Światowa Organizacja Handlu czy Unia Europejska).

To prawdopodobnie nie jest wyczerpująca typologia, ale stworzenie takowej nie jest celem niniejszych dociekań. Na marginesie warto jednak zauważyć, iż we współczesnym piśmiennictwie informatologicznym generalnie brakuje przemyślanej, rzetelnej typologii źródeł informacji w internecie, która obejmowałaby m.in. także archiwa „surowych” danych i odróżniała je od innych zasobów *online*.

W Polsce bodaj najwięcej uwagi poświęca się zbiorom danych publicznych, prawdopodobnie dlatego, iż mają one także wymiar obywatelski i polityczny (zob. np. <http://mojepanstwo.pl> albo <http://otwartedane.pl>), a poza tym w dużej mierze podlegają polskim i unijnym regulacjom prawnym, w tym *Ustawie o dostępie do informacji publicznej* [Obwieszczenie... 2014].

**Archiwa / repozytoria danych badawczych** z kolei są to intencjonalnie stworzone zbiory danych / materiałów, powstałych w wyniku badań empirycznych w różnych przedsięwzięciach naukowych lub też zgromadzonych na potrzeby konkretnych projektów badawczych. Innymi słowy, termin ten nie określa wszystkich istniejących kolekcji danych, które ewentualnie mogłyby „przydać się” w dociekania naukowych (zob. też wcześniejsze uwagi), lecz tylko te, które obejmują ma-

---

<sup>1</sup> Tekst oryginalny: *any information that can be stored in digital form, including text, numbers, images, video or movies, audio, software, algorithms, equations, animations, models, simulations, etc. Such data may be generated by various means including observation, computation, or experiment* [Kowalczyk, Shankar 2011, p. 250-251].

teriały uzyskane w rezultacie badań, a także zostały zorganizowane celowo, odpowiednio uporządkowane, opisane metadanymi i przygotowane do udostępniania [Archiwum... 2014; Brady 2006; Kowalczyk, Shankar 2011; IASSIST 2010 / 2011]. Cytowane już Kowalczyk i Shankar [Kowalczyk, Shankar 2011, p. 256-257] wyróżniają trzy rodzaje archiwów danych badawczych, biorąc pod uwagę ich wielkość i zasięg. Są to kolekcje danych pochodzących z jednego projektu naukowego (ang. *research data collections*), z jednego kierunku bądź obszaru badawczego (ang. *community collections*) i wreszcie wielod dziedzinowe, często międzynarodowe, oparte na dobrze zorganizowanej infrastrukturze technologicznej (ang. *reference collections*).

Archiwa danych badawczych mogą zawierać, po pierwsze, materiał empiryczny pochodzący z badań nauk matematycznych, przyrodniczych i z nurtu pozytywistycznego w dyscyplinach społecznych, czyli tzw. **dane ilościowe**, chociaż w rzeczywistości są to nie tylko liczby, ale również zdjęcia, np. galaktyk albo dna morskiego. Dobry przykład takiego zbioru w otwartym dostępie stanowi NASA / IPAC Extragalactic Database <http://ned.ipac.caltech.edu>, a w Polsce – Archiwum Danych Społecznych <http://www.ads.org.pl> (dostęp „półotwarty”, możliwy po zarejestrowaniu, nie dotyczy zbiorów zastrzeżonych). Po drugie, archiwa mogą obejmować tzw. **dane jakościowe**, zgromadzone w ramach podejścia interpretatywnego / rozumiejącego w naukach humanistycznych i społecznych. Widać, że rozróżnienie archiwów danych ilościowych i jakościowych jest właściwie pochodną stosowanej metodologii badań.

W ostatnich kilku latach pojawiły się interesujące inicjatywy, a także globalne internetowe serwisy wyszukiwawcze, promujące (otwarty) dostęp do „surowych” danych badawczych. Są to m.in.:

- Databib <http://databib.org>, serwis wyszukiwawczy archiwów danych badawczych, obejmujący obecnie (lipiec 2014) 978 repozytoriów z wszystkich dziedzin nauki, w tym humanistycznych i społecznych,
- DataCite <http://www.datacite.org>, założona w 2009 roku międzynarodowa organizacja *non-profit*, mająca na celu ułatwienie dostępu do danych badawczych w internecie,
- Data repositories [http://oad.simmons.edu/oadwiki/Data\\_repositories](http://oad.simmons.edu/oadwiki/Data_repositories) (serwis dotyczy danych otwartych, jest mocno niekompletny),
- ICSU World Data System <http://www.icsu-wds.org>, istniejące od 2008 roku przedsięwzięcie Międzynarodowej Rady Nauki (International Council for Science), z dostępem do zbiorów danych badawczych poprzez Data Portal <http://www.icsu-wds.org/services/data-portal>,
- re3data.org – Registry of research data repositories <http://www.re3data.org/>, rejestr powstały w 2012 roku w Niemczech, w połowie roku 2014 indeksujący ok. 650 archiwów danych badawczych z całego świata.

W marcu 2014 roku Databib, DataCite i re3data.org ogłosiły połączenie sił do końca roku 2015, z zamiarem stworzenia uniwersalnego serwisu wyszukiwawczego dla kolekcji ilościowych i jakościowych danych badawczych w internecie [re3data.org team 2014].

## **Archiwizowanie i ponowne wykorzystywanie danych jakościowych – wybrane aspekty**

Dane ilościowe są archiwizowane standardowo. Piotr Filipkowski pisze: „Potrzeba archiwizowania danych ilościowych jest oczywistością. Nikt nie podaje jej w wątpliwość. Specjalizacja sprzyja wręcz temu, by jedni zbierali i archiwizowali te dane, inni je porządkowali, zestawiali w tabele i wykresy, uczeni analizowali i omawiali, a jeszcze inni – z większą już swobodą – uogólniali” [Filipkowski 2005, s. 32]. W naukach przyrodniczych wiodące czasopisma często wymagają podania, gdzie przechowywane są dane empiryczne stanowiące podstawę publikacji, a ich opracowanie i udostępnienie stanowi warunek wstępny przyjęcia artykułu do druku [Kowalczyk, Shankar 2011, p. 248]. „Wśród badaczy jakościowych (...) tendencja jest dokładnie odwrotna: dane zebrane przez jednych badaczy niezwykle rzadko są traktowane jako źródło w badaniach drugich” [Filipkowski 2005, s. 32]. Jak na wpół żartobliwie zauważa Arja Kuula [Kuula 2010 / 2011], materiały jakościowe często „kończą” w szufladzie, na strychu albo w domku letniskowym badacza. Zrozumienie potrzeby archiwizowania, udostępniania i powtórnego wykorzystywania danych **jakościowych** na szerszą skalę dopiero się rodzi.

Tym niemniej archiwizacja i udostępnianie jakościowych danych badawczych stanowi dzisiaj dynamicznie rozwijającą się dziedzinę zarówno działalności, jak i refleksji teoretycznej. W ostatnich latach pojawiły się różne krajowe i międzynarodowe inicjatywy w tym zakresie, powstają kolejne repozytoria, pojawiła się też odpowiednia oferta edukacyjna, m.in. w Wielkiej Brytanii [Bishop 2014; Corti 2011; IASSIST 2010/2011; Perez-Hernandez 2014; Secondary 2014]. Także biblioteki akademickie na świecie zaczęły umieszczać na swoich stronach przewodniki po zbiorach danych jakościowych [Identifying... 2014]. Wynika to, po pierwsze, z rosnącej popularności badań jakościowych, po drugie – z coraz silniejszego oddziaływania ideologii otwartej nauki [zob. m.in. Pantou 2010], po trzecie wreszcie – z możliwości, które oferuje współczesna technologia informacyjna i sieciowa.

Długą tradycję w archiwizowaniu i udostępnianiu danych badawczych ma Wielka Brytania, gdzie już w roku 1994 powstał pierwszy oficjalny narodowy serwis o zbiorach danych jakościowych ESDS Qualidata, który w 2012 roku zo-

stał włączony do systemu z zakresu nauk ekonomicznych i społecznych UK Data Service <http://ukdataservice.ac.uk>. Od 2014 roku działa w USA portal Qualitative Data Repository QDR <https://qdr.syr.edu>, finansowany przez amerykańską Narodową Fundację Nauki. Zainteresowanie archiwizowaniem danych jakościowych widać również w kilkunastu krajach europejskich, stan obecny i perspektywy w tym zakresie zostały przedyskutowane m.in. podczas warsztatów w Bremie w 2009 roku [IASSIST 2010/2011]. W Polsce wspomniany już Instytut Filozofii i Socjologii PAN prowadzi obecnie projekt pod nazwą Archiwum Danych Jakościowych, finansowany ze środków Narodowego Programu Rozwoju Humanistyki [Archiwum... 2014].

Czym jednak są dane jakościowe, które mają być udostępniane poprzez repozytoria *online*? Jest to materiał empiryczny zgromadzony z wykorzystaniem technik jakościowych w ramach podejścia interpretatywnego / rozumiejącego w naukach o człowieku. Źródłami danych mogą tutaj być: 1) dokumenty, teksty już istniejące lub stworzone na prośbę badacza, oficjalne i nieoficjalne, 2) fokus, zogniskowany wywiad grupowy, 3) obserwacja w różnych wariantach (etnograficzna, naturalistyczna, samoobserwacja, *shadowing*, uczestnicząca lub nie), 4) wywiad indywidualny (jakościowy, pogłębiony, narracyjny, *Sense-Making* i in.) [Cisek 2013].

Empiryczne dane jakościowe dzieli się czasami na niewywołane (zastane) i wywołane (wytworzone). Pierwsze z nich powstają „naturalnie”, bez ingerencji badacza (np. listy, posty w blogach, wpisy w portalach społecznościowych, zdjęcia z wakacji), drugie – w efekcie czynności uczonego, w „sztucznych”, stworzonych na potrzeby badania sytuacjach, np. podczas obserwacji uczestniczącej albo wywiadu pogłębionego. Jednakże, jak trafnie zauważa Tim Rapley, „w rzeczywistości w obu przypadkach działania badacza odgrywają podstawową rolę w procesie wytwarzania materiałów i ich przekształcania w dane. W obu przypadkach trzeba je przecież odkryć, fizycznie zgromadzić, wybrać te, które zatrzymamy, i odrzucić te, które pominiemy. (...) Najważniejsza decyzja wiąże się z uznaniem konkretnego zestawu materiałów za zbiór danych” [Rapley 2010, p. 34].

W piśmiennictwie dotyczącym archiwizowania danych jakościowych autorzy najczęściej mają na myśli udostępnianie i powtórne wykorzystanie materiałów **wywołanych**, jednakże zbiory materiałów niewywołanych również mogą być cenne, zwłaszcza że granica między nimi bywa płynna, jak wskazano w poprzednim akapicie. Doskonały przykład stanowią tutaj kolekcje cyfrowe Ośrodka Karta, w tym Archiwum Fotografii <http://foto.karta.org.pl> i Archiwum Historii Mówionej <http://www.audiohistoria.pl>.

W repozytoriach badawczych danych jakościowych znajduje się zatem bogaty, z reguły słabo ustrukturyzowany materiał empiryczny w postaci werbalnej (dzien-

niki, nagrania audio, notatki terenowe, ogłoszenia, opowieści, pamiętniki, regu-  
laminy, transkrypcje, zapisy wywiadów, zarządzenia itp.), wizualnej (fotografie,  
rysunki) lub multimedialnej (filmy, nagrania wideo, zawartość serwisów www).  
Także w tym przypadku, podobnie jak generalnie w odniesieniu do archiwów  
danych badawczych, trzeba podjąć decyzję dotyczącą stopnia przetworzenia da-  
nych zamieszczonych w zbiorze, np. czy będą tam oryginalne nagrania przepro-  
wadzonych wywiadów, czy tylko ich transkrypcje, czy jedno i drugie (najlepiej).

Podstawowe pytanie brzmi – czemu mają służyć archiwa danych jakości-  
owych, po co właściwie je organizować? Otóż przede wszystkim po to, by umożli-  
wić **powtórne wykorzystanie** istniejącego, już raz wytworzonego materiału em-  
pirycznego, przede wszystkim w celach naukowych, ale również edukacyjnych,  
a nawet gospodarczych i społecznych (kontrola wydatkowania publicznych pie-  
niędzy, transfer wiedzy z nauki do gospodarki). Dzięki temu można osiągnąć  
m.in. korzyści poznawcze, o których mowa w dalszej części rozważań, dydak-  
tyczne, zwłaszcza w szkolnictwie wyższym, gdzie dane empiryczne zgromadzone  
w autentycznych badaniach mogą stanowić znakomitą pomoc w nauczaniu me-  
tologii i nie tylko, oraz finansowe i organizacyjne. Badania jakościowe często  
są czas- i pracochłonne, a więc drogie. Powtórne wykorzystanie raz zgromadzo-  
nego materiału empirycznego przynosi oszczędność czasu, pieniędzy i wysiłku.  
Poza tym dane jakościowe mogą np. dotyczyć grup ludzi, do których dotarcie  
stało się utrudnione bądź niemożliwe. Wówczas istniejące, „gotowe” dane są je-  
dynym źródłem dla przyszłych badaczy.

Tworzeniu archiwów danych jakościowych towarzyszą problemy metodolo-  
giczne, etyczne, organizacyjne, prawne, a także informacyjne / informatologiczne  
(z zakresu organizacji wiedzy, zarządzania informacją i udostępniania).

Zagadnienia **metodologiczne** wiążą się z poznawczymi celami tworzenia ar-  
chiwów danych jakościowych. Obejmują dwa podstawowe wymiary, odnoszące  
się do wykorzystania udostępnionego materiału empirycznego: dla zapewnienia  
intersubiektywnej kontrolowalności i sprawdzalności badań jakościowych oraz  
dla przeprowadzenia nowych projektów badawczych.

Intersubiektywna sprawdzalność stanowi uniwersalny warunek naukowości,  
może jednak być osiągnięta na różne sposoby. W metodologii ilościowej w na-  
ukach przyrodniczych często można „po prostu” powtórzyć badania poprzed-  
ników (sprawdzić poprawność obliczeń, wykonać ten sam eksperyment itp.).  
W dociekaniach jakościowych zazwyczaj nie jest to możliwe, nie da się drugi raz  
zrobić „tej samej” obserwacji etnograficznej czy wywiadu narracyjnego, dlatego  
należy pozostawić tzw. ścieżkę sprawdzenia (ang. *audit trial*), czyli rygorystycz-  
nie i wyczerpująco opisać własne postępowanie badawcze. Szczegółowo scha-  
rakteryzowane powinny zostać nie tylko: kontekst, uwarunkowania i założenia

badania oraz sam tok rozumowania, ale również – właśnie – dane empiryczne, na których oparte zostały wnioski [Cisek 2013]. Jeszcze lepiej byłoby, gdyby czytelnicy, inni naukowcy, mogli sięgnąć do oryginalnego materiału empirycznego i stwierdzić, czy został on odpowiednio i poprawnie wykorzystany, zapewniając trafność i rzetelność badań. Z drugiej strony niektórzy obawiają się niewłaściwego wykorzystania „swoich” danych przez „konkurencję”, np. do uzyskania grantów, albo złośliwej krytyki. Jednakże transparentność procesu badawczego jest wartością nadrzędną [Kuula 2010/2011, p. 14].

Dane jakościowe zgromadzone w archiwach można też wykorzystać w nowych projektach badawczych. Potencjał jest znaczny i obejmuje m.in.:

- badania długookresowe, np. powrót po latach do tych samych badanych albo przynajmniej do tej samej grupy społecznej;
- dociekania porównawcze, np. zbadanie istniejących danych empirycznych inną metodą / techniką (komparatystyka metodologiczna). Hanna Palska zauważa: „Użytkownicy nowych metod interpretacyjnych raz po raz apelują o reanalizę starych materiałów, by skonfrontować własne techniki interpretacyjne z technikami niegdyś używanymi” [Palska 2005, s. 14];
- powtórne poszukiwanie odpowiedzi na pytania postawione w oryginalnych dociekaniach, co może zaowocować uchwyceniem innego aspektu badanego wycinka rzeczywistości. Kolejna analiza istniejących materiałów, przeprowadzona w odmiennych ramach pojęciowych, często pozwala odnaleźć relacje i treści, których dotąd nikt nie zauważył [Kuula 2010 / 2011, p. 14];
- wykorzystanie tego samego, „starego” materiału empirycznego do poszukiwania rozwiązań zupełnie nowych, innych zagadnień.

Jednakże czy można (w obydwu znaczeniach – czy się da i czy wolno) ponownie wykorzystywać badawcze dane jakościowe? Ich powtórne użycie wzbudza niejakie kontrowersje metodologiczne, ponieważ dla dociekań jakościowych często charakterystyczne jest **osobiste** zaangażowanie badacza, nawiązanie emocjonalnej i intelektualnej więzi z badanymi, pozwalające ich zrozumieć. Gdy inni uczeni mają do dyspozycji tylko sam zapis materiału empirycznego, tego ważnego czynnika może im brakować. To zastrzeżenie oczywiście dotyczy tylko danych wywołanych, powstałych w interakcji z badanymi lub na prośbę badacza. W odniesieniu do danych zastanych, np. istniejących dokumentów, nie ma ono większego sensu.

Z kolei kontrargumenty wobec wskazanego zastrzeżenia są następujące:

- W repozytoriach *de facto* archiwizuje się jakościowy **materiał empiryczny**, z którego dopiero należy „zrobić” różne **konkretne dane** na potrzeby kolejnych, odmiennych projektów naukowych. Co więcej, z reguły jest on redundantny wobec zagadnienia badawczego, dla którego został zgromadzony.

Jak zauważa Piotr Filipkowski: „Każde dane wymagają interpretacji. Dane jakościowe – dokumenty osobiste, relacje biograficzne, notatki z obserwacji, zapisy fokusów itp. – otwierają szczególne możliwości reinterpretacji. Są to bowiem prawie zawsze dane nadmiarowe w stosunku do pytań, jakie postawił wobec nich badacz, który je zgromadził i zanalizował” [Filipkowski 2005, s. 34]. Materiały takie można wykorzystać, jak już wcześniej wspomniano, do rozwiązania innego, zupełnie nowego problemu poznawczego.

- W dociekaniach jakościowych nie istnieje pojedynczy i „jedynie słuszny” sposób analizy i interpretacji zebranego materiału badawczego, więc nie można twierdzić, że tylko „oryginalny” badacz ma prawo to zrobić [Kuula 2010/2011, p. 14].
- Czasami i w „tradycyjnych” badaniach jakościowych bywa tak, że kto inny np. przeprowadza wywiady, a kto inny je analizuje i to jest akceptowane [Kuula 2010/2011, p. 14].

Aby powtórne wykorzystanie jakościowego materiału empirycznego było uzasadnione i wykonalne, konieczne jest zamieszczenie w archiwach, oprócz danych, także informacji kontekstowych, m.in. dla jakiego przedsięwzięcia badawczego i za pomocą jakich narzędzi / technik materiał został zgromadzony, jakie są jego relacje do innych zbiorów danych, czy wykorzystano konkretne oprogramowanie itp. Należy również określić nośnik, dostępność techniczną i, pod względem prawnym, kto jest właścicielem oraz podać inne potrzebne metadane [Kowalczyk, Shankar 2011, p. 253; Kuula 2010 / 2011, p. 14]. Cytowany już Filipkowski pisze: „Archiwizowanym danym powinna towarzyszyć rzetelna i wyczerpująca dokumentacja określająca założenia, cele, metody badawcze i wyniki projektu, w ramach którego te dane zebrano. Jest to warunek ich kompetentnej reanalizy” [Filipkowski 2005, s. 37-38]. Ważne jest także, by materiał zgromadzony i udostępniany poprzez archiwa był dobrej jakości, zarówno w wymiarze naukowym, jak i informacyjnym, by cechował się kompletnością intelektualną i techniczną, niezmiennością / trwałością, poprawnością i spójnością [Kowalczyk, Shankar 2011, p. 254]. Wskazane kwestie łączą w sobie aspekty **metodologiczne** i z zakresu **zarządzania informacją**.

Problemy **etyczne** dotyczą ochrony prywatności, a także emocji, psychiki i społecznego funkcjonowania badanych. W dociekaniach jakościowych badacz często spędza z nimi wiele czasu, są to konkretne osoby (a nie – jak w metodologii ilościowej – „elementy” próby reprezentatywnej), poruszane bywają wrażliwe kwestie. W związku z tym, oczywiście, badani powinni podpisać zgodę na archiwizację, ich dane osobowe należy utajnić, a notatki i wypowiedzi poddać anonimizacji. Z drugiej strony, jak się okazuje, wbrew licznym i zdaje się nadmiernym obawom badaczy, zdecydowana większość badanych nie ma nic prze-

ciwko powtórnemu wykorzystaniu związanych z nimi materiałów empirycznych, często wręcz tego oczekują, wychodząc z założenia, że skoro poświęcili swój czas, emocje, uwagę „dla dobra nauki”, to powinien być z tego daleko idący pożytek [Kuula 2010/2011, p. 15-16].

Wymiar **organizacyjny i prawny** jest równie istotny. Należy m.in. rozwiązać kwestie otwartego bądź restrykcyjnego dostępu do archiwów, prawa własności do materiału empirycznego (pozostające w gestii badacza, badanych lub instytucji finansującej), a także wspomnianej w poprzednim akapicie ochrony prywatności (np. udostępnianie informacji kontaktowych do badanych). Należy także rozważyć, kto ma się zajmować tworzeniem archiwów danych jakościowych: każda zainteresowana placówka badawcza, biblioteki akademickie i naukowe, organizacje społeczne lub uczelnie oraz czy powinny powstawać centralne archiwa narodowe / państwowe albo nawet europejskie. Są już próby formułowania pewnych zasad, jedną z nich stanowią tzw. *Panton Principles* [Bednarek-Michalska 2012; Panton 2014].

Aspekt **informacyjny** (informatologiczny) wiąże się z organizacją wiedzy i zarządzaniem informacją w repozytoriach jakościowych danych badawczych. Chodzi tutaj m.in. o formaty danych, standardy opisu i katalogowania jakościowego materiału empirycznego, adekwatny system metadanych, który umożliwiłby sprawny dostęp, wyszukiwanie i wymianę informacji między archiwami / instytucjami, a także o kwestie związane ze starzeniem się nośników, potrzebą digitalizacji itd. Jest to kolejny, ogromny obszar spraw do rozwiązania, który należałoby rozważyć w odrębnej publikacji.

## Zakończenie

W artykule zarysowano obszar zagadnień związanych z tworzeniem archiwów / repozytoriów danych badawczych, w szczególności jakościowych. Obecnie podejmowane są różnorodne działania, których celem jest zbudowanie infrastruktury i kultury wielokrotnego wykorzystywania jakościowego materiału empirycznego.

Otwiera się tu szerokie pole do badań i działalności, także dla praktyków i teoretyków bibliotekoznawstwa i nauki o informacji. Po pierwsze, interesujące są wymienione już kwestie związane z aspektami informacyjnym i organizacyjnym tworzenia repozytoriów danych jakościowych. Po drugie, można sprawdzić, czy w istniejących archiwach danych nie ma materiałów, które dałoby się wykorzystać również i w dociekaniach bibliologicznych i informatologicznych, zamiast, powiedzmy, przeprowadzać kolejną standardową ankietę. Po trzecie, archiwa da-



nych badawczych to rodzaj naukowych źródeł informacji, należy więc włączyć je do treści kształcenia z zakresu informacji naukowej, zarządzania informacją itp. Po czwarte wreszcie, w literaturze przedmiotu nie ma publikacji poświęconych użytkownikom archiwów / repozytoriów danych jakościowych, ich zachowaniom informacyjnym itd., a warto byłoby to zbadać.

To tylko kilka propozycji, zapewne interesujących możliwości jest więcej.

## BIBLIOGRAFIA

- Archiwum Danych Jakościowych. Instytut Filozofii i Socjologii Polskiej Akademii Nauk (2014). [http://www.ifispan.waw.pl/index.php?lang=pl&m=page&pg\\_id=55](http://www.ifispan.waw.pl/index.php?lang=pl&m=page&pg_id=55) [odczyt: 06.07.2014].
- Bednarek-Michalska, Bożena (2012). *Repozytoria surowych danych – dlaczego biblioteki powinny je znać?* [dok. elektr.] Biuletyn EBIB nr 8 (135). <http://open.ebib.pl/ojs/index.php/ebib/article/view/152/292> [odczyt: 02.07.2014].
- Bishop, Libby (2014). *An invitation to secondary analysis of qualitative data. Secondary analysis of qualitative data: an advanced course, University of Essex, 22-23 May 2014* [dok. elektr.]. <http://ukdataservice.ac.uk/media/440302/secondaryanalysis22-23may14.pdf> [odczyt: 11.07.2014].
- Brady, Angela (2006). *Qualitative data archives*. W: *The SAGE Dictionary of Social Research Methods* [dok. elektr.]. <http://srmo.sagepub.com/view/the-sage-dictionary-of-social-research-methods/n164.xml> [odczyt: 08.07.2014].
- Cisek, Sabina (2013). *Metodologia jakościowa we współczesnej informatologii. Wybrane aspekty*. Przegląd Biblioteczny R. 81, z. 3, s. 299-310.
- Corti, Louise (2011). *The European Landscape of Qualitative Social Research Archives: Methodological and Practical Issues* [dok. elektr.]. Forum: Qualitative Social Research Vol. 12, Issue 3. <http://nbn-resolving.de/urn:nbn:de:0114-fqs1103117> [odczyt: 06.07.2014].
- Filipkowski, Piotr (2005). *Po co archiwizować dane jakościowe i jak robią to inni*. ASK nr 14, s. 31-52.
- IASSIST Quarterly Vol. 34, No. 3-4; Vol. 35., No. 1-2. *Special Issue: Qualitative and Qualitative Longitudinal Resources in Europe. Mapping the field and exploring strategies for development (2010 / 2011)* [dok. elektr.]. [http://www.iassistdata.org/downloads/iqvol34\\_35.pdf](http://www.iassistdata.org/downloads/iqvol34_35.pdf), <http://www.iassistdata.org/iq/issue/34/3> [odczyt: 02.07.2014].
- Identifying Repositories for Data Sharing* (2014) [dok. elektr.]. <http://research.library.gsu.edu/content.php?pid=348991&sid=2855062> [odczyt: 10.07.2014].
- Kościański, Artur (2005). *Archiwum społecznych danych jakościowych: potrzeby, kontrowersje, propozycje praktyczne*. ASK nr 14, s. 19-29.
- Kowalczyk, Stacy; Shankar, Kalpana (2011). *Data sharing in the sciences*. Annual Review of Information Science and Technology Vol. 45, Issue 1, p. 247-294.

- Kuula, Arja (2010/2011). *Methodological and Ethical Dilemmas of Archiving Qualitative Data* [dok. elektr.]. IASSIST Quarterly Vol. 34-35, p. 12-17. <http://www.iassistdata.org/iq/methodological-and-ethical-dilemmas-archiving-qualitative-data> [odczyt: 06.07.2014].
- Obwieszczenie Marszałka Sejmu Rzeczypospolitej Polskiej z dnia 14 kwietnia 2014 r. w sprawie ogłoszenia jednolitego tekstu ustawy o dostępie do informacji publicznej. Dziennik Ustaw 2014 poz. 782 [dok. elektr.]. <http://www.dziennikustaw.gov.pl/du/2014/782/1> [odczyt: 01.09.2014].
- Palska, Hanna (2005). *O potrzebie ochrony danych jakościowych. Z doświadczeń socjologa-humanisty*. ASK nr 14, s. 7-17.
- Pampel, Heinz (2013). *How to find an appropriate research data repository?* PLOS Tech Blog [dok. elektr.]. <http://blogs.plos.org/tech/how-to-find-an-appropriate-research-data-repository/> [odczyt: 02.07.2014].
- Pampel, Heinz et al. (2013). *Making Research Data Repositories Visible: The re3data.org Registry* [dok. elektr.]. PLOS ONE Vol. 8 (11). <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0078080> [odczyt: 02.07.2014].
- Panton Principles. Principles for Open Data in Science* (2010) [dok. elektr.]. <http://panton-principles.org> [odczyt: 07.07.2014].
- Perez-Hernandez, Danya (2014). *New Repository Offers a Home for Data That Aren't Numbers* [dok. elektr.]. <http://chronicle.com/blogs/wiredcampus/new-repository-offers-a-home-for-data-that-arent-numbers/50865> [odczyt: 10.07.2014].
- Rapley, Tim (2010). *Analiza konwersacji, dyskursu i dokumentów*. Warszawa: Wydawnictwo Naukowe PWN, 238 s.
- re3data.org team (2014). *DataCite, re3data.org, and Databib announce Collaboration* [dok. elektr.]. <http://www.re3data.org/2014/03/datacite-re3data-org-databib-collaboration/> [odczyt: 12.07.2014].
- Secondary Qualitative Data Sources and How to Find Them* (2014) [dok. elektr.]. <http://medanth.wikispaces.com/Secondary+Qualitative+Data+Sources+and+How+to+Find+Them> [odczyt: 06.07.2014].
- What is DataCite?* (2014) [dok. elektr.]. <http://www.datacite.org/whatisdatacite> [odczyt: 10.07.2014].

---

# Wsparcie dla indywidualnego zarządzania informacją w polskich repozytoriach uczelnianych

Remigiusz Sapa

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

## Wprowadzenie

Zasadniczym celem przyjętym w artykule jest rozpoznanie różnych form wsparcia oferowanych przez repozytoria uczelniane, mających ułatwiać ich użytkownikom realizowanie procesów indywidualnego zarządzania informacją. W pierwszej kolejności wskazano kluczowe konteksty działań podejmowanych w ramach indywidualnego zarządzania informacją w relacjach z zewnętrznymi zasobami informacji. Następnie omówiono wyniki poszukiwania dostępnych w repozytoriach narzędzi i rozwiązań, wspierających użytkowników w tych działaniach. Badaniami, które zostały przeprowadzone w lipcu i sierpniu 2014 roku, objęto dziewięć repozytoriów polskich uczelni zarejestrowanych w Federacji Bibliotek Cyfrowych i jednocześnie używających terminu „repozytorium” w swojej nazwie. Poszukiwania były prowadzone z perspektywy użytkownika dążącego do zrealizowania następujących zadań indywidualnego zarządzania informacją: przejęcia dokumentu lub jego fragmentów do własnej kolekcji w celu późniejszego wykorzystania, przejęcia opisu bibliograficznego całego dokumentu lub jego fragmentu do własnych zasobów w celu późniejszego wykorzystania lub uzyskania ponownego dostępu, utrzymania dokumentu lub jego opisu w obszarze percepcji użytkownika.

Przyjmując, że jednym z celów budowania repozytoriów uczelnianych jest umożliwienie i ułatwienie dostępu do zasobów informacji generowanych przez daną uczelnię i jej pracowników, konieczne wydaje się dążenie do sytuacji, w której między działaniami podejmowanymi przez użytkowników w ramach indywidualnego zarządzania informacją (ang. *personal information management* – PIM)

a rozwiązaniami stosowanymi w repozytoriach zachodzi jakaś forma synergii. Chodzi o to, by tworząc repozytorium, nie koncentrować się wyłącznie na liczbie oferowanych obiektów informacyjnych czy merytorycznej wartości ich treści, ale także by zwrócić uwagę na potencjalną efektywność i sprawność ich wykorzystania przez użytkowników, która, jak zostało tutaj przyjęte, wynika między innymi z „przyjazności” repozytorium właśnie dla procesów indywidualnego zarządzania informacją.

Indywidualne zarządzanie informacją bywa różnie rozumiane [Jones 2007, p. 453-455; Świąg 2012, s. 192-200]. Podczas gdy jedni autorzy postrzegają tę problematykę szeroko, jako obejmującą w zasadzie wszelkie działania, narzędzia i rozwiązania stosowane przez człowieka zarządzającego na własny użytek informacją będącą w jego dyspozycji [np. Jones 2008, p. 5], w tym też przez niego tworzoną [Whittaker 2011, p. 10], inni wyraźnie zawężają rozumienie PIM do wybranych aspektów, w tym np. wyłącznie do relacji między człowiekiem i komputerem, osadzonych w tradycji badań w obszarze znanym jako *human-computer interactions* [np. Bergman et al. 2013, p. 1995], bądź tylko do działań podejmowanych przez człowieka w kontekście wykonywania zadań zawodowych [np. Barreau 1995, p. 327]. Punktem wyjścia dla prowadzonych tutaj rozważań było szerokie spojrzenie na problematykę PIM. Dokonując jednak jej operacjonalizacji na potrzeby badań, skupiono się tylko na wybranych działaniach, precyzyjnie określonych w części poświęconej metodologii.

Termin „repozytorium” też bywa różnie rozumiany, a dyskusje na ten temat toczą się również w piśmiennictwie polskim [szerzej: Sapa 2013, s. 118-121]. W przekonaniu autora w dobie konwergencji narzędzi i usług informacyjnych dążenie do tworzenia radykalnie rozłącznej typologii przedsięwzięć podejmowanych w obszarze zarządzania cyfrowymi zasobami informacji musi być skazane na niepowodzenie. Tak samo jak nie da się wskazać ostrej granicy między telefonem a komputerem (np. w przypadku smartfonów lub fabletów) lub usługą i produktem informacyjnym (np. indywidualizowany druk publikacji na życzenie), tak też trudno byłoby zdecydowanie odróżnić repozytoria od np. bibliotek cyfrowych. Można raczej mówić o realizacji funkcji repozytoryjnych przez różne przedsięwzięcia. W dalszym toku wywodu uwaga zostanie skupiona na tych inicjatywach, których twórcy jednoznacznie deklarują nadrzędność funkcji repozytoryjnych w swoim działaniu.

Problematyka relacji między PIM użytkowników a repozytoriami uczelnianymi nie była jak dotąd szerzej poruszana w piśmiennictwie. Owszem, podejmowano bezpośrednio zagadnienia relacji między PIM użytkownika a biblioteką i oferowanymi przez nią usługami i produktami (co ze względu na duże zaangażowanie bibliotek akademickich w budowanie repozytoriów jest z pewnością

interesujące), ale jedynie sporadycznie [np. Arth et al. 2011] lub wycinkowo [np. Fourie 2011b; Newton-Smith 2000; Otopah, Dadzie 2013; Copeland 2011]. Choć nie brakuje odrębnych publikacji zarówno na temat indywidualnego zarządzania informacją, jak i różnych aspektów funkcjonowania repozytoriów uczelnianych, do tej pory nie łączono tych dwóch obszarów zainteresowania w ramach badań możliwości wsparcia PIM użytkowników przez rozwiązania stosowane w repozytoriach.

## **Rola indywidualnego zarządzania informacją w relacjach z zewnętrznymi zasobami informacji**

Koncepcja PIM zakłada, że człowiek podejmuje wysiłek intelektualny, emocjonalny, a czasem i fizyczny, poświęca czas i środki, by samodzielnie gromadzić, klasyfikować i kategoryzować, opisywać, archiwizować i przetwarzać informacje w imię spodziewanych w przyszłości korzyści.

Jako przyczynę tych działań zazwyczaj wskazuje się chęć zapewnienia sobie sprawnego ponownego dostępu do informacji wcześniej pozyskanych: wytworzonych lub znalezionych, zarówno celowo, jak i przypadkowo [np. Bruce, Jones, Dumais 2004; Jones 2008]. Cel ten można osiągać zasadniczo dwiema drogami: budując własne kolekcje zasobów informacyjnych lub tworząc narzędzia ułatwiające ich ponowne pozyskanie (spisy linków, „zakładki” do odwiedzonych wcześniej stron www, wykazy bibliograficzne z sygnaturami bibliotecznymi itp.).

Współcześnie może wydawać się, że sprawne i powszechnie dostępne narzędzia wyszukiwawcze (np. wyszukiwarki z „rodziny” Google’a) mogą eliminować potrzebę rozwijania własnego PIM, skoro przydatne zasoby można w dowolnej chwili ponownie odnaleźć w internecie. Jest to pomysł o tyle atrakcyjny, że może pozwolić jednostce na zrezygnowanie z czasochłonnego samodzielnego zarządzania informacją, a także, a może przede wszystkim, z trudnych również z intelektualnego i emocjonalnego punktu widzenia decyzji na temat tego, co się w przyszłości przyda, a co nie, co usunąć już z własnych zbiorów, a co zachować [Robinson, Johnson 2012, p. 9-12]. Nie wydaje się jednak, by traktowanie ponownego wyszukiwania jako alternatywy dla PIM było właściwe w odniesieniu do zachowań informacyjnych środowiska akademickiego. Z pewnością konieczne są odrębne badania roli ponownego wyszukiwania w dostępie do informacji [Whittaker 2011], ale, choćby ze względu na trudności w ponownym odszukaniu, temporalność części zasobów czy wreszcie fakt, że nie wszystko jest dostępne w internecie, należy je traktować raczej jako strategię uzupełniającą PIM w wybranych sytuacjach.

PIM ma także wspomagać pamięć przez przypominanie o zadaniach do wykonania lub utrzymywanie informacji w „polu widzenia” tak długo, aż zdoła się ją w sposób właściwy wykorzystać – np. załatwić reprezentowaną przez tę informację sprawę. Dobrym przykładem są przychodzące maile, z których część zawiera treści wymagające działań odbiorcy nie od razu, ale w niedalekiej przyszłości. Takie wiadomości muszą w jakiś sposób cały czas pozostawać dla adrestata „aktywne” (w przeciwieństwie do „załatwionych”), by o nich nie zapomnieć [Whittaker 2011, p. 19-20]. Funkcje „przypominania” można realizować w różny sposób: np. pozostawiając „niezałatwione” maile w folderze „odebrane”, przez umieszczenie pozyskanego, ważnego pliku (np. skanu) na pulpicie komputera (a nie w folderze, który dla tego typu plików przewidziało się w ramach własnego PIM), zaopatrzenie się w narzędzia automatycznego przypominania lub powiadamiania, ale też przez pozostawienie książki do przeczytania na biurku, zrobienie notatki i przyklejenie jej w widocznym miejscu czy przygotowywanie różnych list zadań do wykonania [Robinson, Johnson 2012, p. 7].

Nie wchodząc głębiej w analizę samej koncepcji indywidualnego zarządzania informacją oraz rozwiązań i narzędzi stosowanych w ramach PIM, warto przyjrzeć się sytuacji, kiedy indywidualna przestrzeń informacyjna człowieka<sup>1</sup> „styka” się z zewnętrznym wobec niej systemem informacji (takim jest też repozytorium uczelniane).

Zachowania człowieka w ramach PIM w takiej sytuacji mogą być kształtowane przez różne, czasem bardzo specyficzne czynniki. Z jednej strony można przypuszczać, że będzie to subiektywna ocena przydatności informacji pozyskanej w zewnętrznym systemie (użytkownik określa, czy zamierza z niej korzystać na pewno, często i intensywnie, czy raczej tylko „być może” i co najwyżej sporadycznie) i jej potencjalnej ulotności. Z drugiej strony wpływ na decyzje użytkownika będą też miały zapewne cechy samego systemu, a szczególnie zaufanie do jego stabilności i trwałości w czasie, które pozwalają zakładać pewność uzyskania ponownego dostępu w przyszłości – przeświadczenie, że znaleziony obiekt informacyjny „nie zginie”, nie musi być zatem dodatkowo zabezpieczony w ramach własnej kolekcji. Nie można też apriorycznie wykluczyć czynników psychologicznych, które mogą kazać jednym ludziom gromadzić różne zasoby zdecydowanie bardziej intensywnie niż innym bądź zmuszać ich do nadawania napotykanym zasobom własnego porządku i organizowania ich według własnych preferencji. Wymienienie wszystkich czynników kształtujących relacje między

---

<sup>1</sup> Indywidualna przestrzeń informacyjna człowieka (ang. *personal space of information* – PSI) – termin wprowadzony przez Williama Jonesa na określenie zasobów i kolekcji informacji kontrolowanych przez jednostkę wraz z narzędziami i konstrukcjami wykorzystywanymi do zarządzania tymi zasobami [Jones 2007, p. 462-463].

PIM a zewnętrznym systemem informacji wymagałoby oczywiście inaczej zaprojektowanych badań i rozważań – chodzi tutaj tylko o podkreślenie złożoności takiej sytuacji i specyficzności podejmowanych każdorazowo działań, a przyjęty tryb przypuszczający wynika z niedostatku empirycznych dowodów na istnienie sygnalizowanych zależności.

Właśnie ze względu na tę różnorodność możliwych sytuacji, a co za tym idzie – zachowań i oczekiwań, nie dążono do oceny poszczególnych rozwiązań stosowanych w repozytoriach, co wymagałoby zastosowania innych metod, w tym przede wszystkim badania samych użytkowników w różnych sytuacjach, ale do zidentyfikowania form wsparcia dla PIM użytkowników, oferowanych przez repozytoria.

## **Zakres i metoda badań**

Punktem wyjścia do projektowania badań było wskazanie kluczowych działań (zadań), które mogą być podejmowane w ramach indywidualnego zarządzania informacją w sytuacji kontaktu z zasobami repozytoriów uczelnianych, stanowiącymi kolekcje informacji zewnętrzne w stosunku do indywidualnej przestrzeni informacyjnej człowieka. Katalog tych działań z jednej strony wynika z zakresu samej koncepcji PIM w postaci przyjętej na użytek niniejszych rozważań, a z drugiej – ze specyfiki potrzeb pracowników naukowo-dydaktycznych uczelni, dla których, między innymi, tworzone są repozytoria uczelniane. Dlatego trzeba podkreślić, że przyjęcie innego sposobu rozumienia PIM lub innej perspektywy użytkownika (np. osoby odpowiedzialnej za PR uczelni) mogłoby doprowadzić do wytypowania nieco innego zestawu działań, a uzyskane i przedstawione w dalszej części artykułu wyniki należy interpretować nie w odniesieniu do PIM w ogóle, lecz do działań wymienionych poniżej:

- przejęcie dokumentu lub jego fragmentu do własnej kolekcji,
- przejęcie opisów bibliograficznych całych dokumentów lub ich fragmentów,
- uzyskanie ponownego dostępu (w ramach danego systemu i poza nim) do wcześniej wyszukanych dokumentów lub ich fragmentów,
- utrzymywanie dokumentu lub jego opisu w obszarze percepcji użytkownika.

Celem badań nie było uzyskanie wyników ilościowych i w związku z tym nie miało znaczenia, czy jakieś rozwiązanie jest częste, czy rzadkie. Nie chodziło bowiem o naszkicowanie statystycznego obrazu badanego zjawiska, ale o rozpoznanie możliwie wielu różnych rozwiązań oferowanych przez repozytoria objęte badaniami, a także o dogłębną analizę ich specyfiki pod kątem potencjału w zakresie wspierania działań i rozwiązań, które mogą być stosowane przez użytkow-

ników w ramach indywidualnego zarządzania informacją. Rozwiązań, z których każde, niezależnie od tego, jak często się pojawia, może być wykorzystane w doskonaleniu samych repozytoriów.

Z tego względu, a także z powodu niemożności ostrego rozdzielenia repozytoriów od choćby bibliotek cyfrowych, uwzględniono w badaniach tylko te przedsięwzięcia, które w swojej nazwie sygnalizują koncentrację na funkcjach repozytoryjnych i jednocześnie pełnią faktyczną rolę repozytoriów konkretnych uczelni. Wyboru dokonano spośród przedsięwzięć zarejestrowanych w Federacji Bibliotek Cyfrowych (wykaz z adresami www i kodami używanymi na oznaczenie poszczególnych repozytoriów w dalszej części tekstu – zob. Aneks 1). Badania zostały przeprowadzone w lipcu i sierpniu 2014 roku i polegały na poszukiwaniu rozwiązań oferowanych w serwisach www tych repozytoriów z perspektywy pracownika naukowo-dydaktycznego, dążącego do zrealizowania wymienionych wcześniej działań PIM. Poszczególne rozwiązania testowano na pięciu dokumentach ostatnio dodanych do repozytorium lub, w przypadkach gdy takie wykazy były niedostępne, na pięciu dokumentach wydanych w roku 2014, wyświetlanych na początku listy uporządkowanej w układzie zstępującym.

## **Rozwiązania wspierające realizację działań w ramach PIM użytkowników**

### **Przejęcie dokumentu lub jego fragmentu do własnej kolekcji**

Przejęcie znalezionej kopii dokumentu oznacza powielenie danego pliku i zapisanie kopii w PSI użytkownika. Zidentyfikowano dwa zasadnicze rozwiązania stosowane w repozytoriach, które można interpretować jako próby wspierania PIM użytkownika w tym zakresie. Nie brano pod uwagę opcji zapisania pliku z poziomu narzędzi aplikacji pozwalającej na wyświetlenie dokumentu (np. Adobe Reader), jako niewynikającej z możliwości samego repozytorium.

Pierwsze ze wspomnianych rozwiązań, stosowane w repozytoriach budowanych w oparciu o dLibrę (Repolis, RCUTP, RUW), sprowadza się do zaproponowania funkcji „pobierz”, która pojawia się w menu z lewej strony, gdy wyświetlony jest opis znalezionej kopii dokumentu. W RUW jej dostępność jest uzależniona od zarejestrowania się w systemie, co stanowi wyraźne i zbędne utrudnienie, gdyż i tak każdy może, z poziomu aplikacji, zapisać na swoim dysku pliki po ich wyświetleniu, nie jest bowiem do tego wymagana rejestracja. Aktywacja funkcji „pobierz” powoduje zapisanie wersji automatycznie kompresowanej w formacie zip, której towarzyszy dodatkowy plik informacyjny „readme”, zawierający z reguły tytuł i nazwę pliku dokumentu oraz różne informacje dodatkowe (np.



informację o zastosowanym standardzie kodowania znaków) w języku polskim i angielskim. Nie wnosi on jednak żadnych dodatkowych treści w stosunku do głównego dokumentu, nie stanowi też gotowego pliku z pełnym opisem bibliograficznym, który można w prosty sposób przejść do własnych zbiorów bibliograficznych. Z punktu widzenia PIM wydaje się całkowicie zbędny. Całość zapisuje się automatycznie w nowym folderze o nic nieznaczącej dla użytkownika nazwie numerycznej. Sam plik dokumentu głównego automatycznie zapisany jest pod nazwą składającą się z różnych elementów (brak standardu nazywania plików dostrzeżono nawet w obrębie jednego repozytorium), np. nazwiska i inicjału autora oraz początku tytułu, samego początku tytułu, samego nazwiska itp. Podsumowując: skorzystanie z funkcji „pobierz” wymaga od użytkownika chcącego umieścić dany dokument w swojej kolekcji tytułu umiejętności i czynności (zapisanie całości na swoim dysku, rozpakowanie, przeniesienie pliku do odpowiedniego folderu w swoim PSI, usunięcie pliku dodatkowego i pustego już folderu, dokonanie zmiany nazwy pliku pod kątem własnych potrzeb), że jej używanie raczej utrudnia proces przejścia dokumentu, niż go ułatwia.

Drugie podejście, typowe dla repozytoriów budowanych w oparciu o system DSpace (eRIKA, AMUR, RUKW, RUŁ, RUMK) oraz RPK, polega na umieszczeniu w opisie bibliograficznym danego dokumentu linków odsyłających bezpośrednio do zawierających go plików (pole „pliki”, „pliki tej pozycji”, „plik zasobu”). Ich kliknięcie nie powoduje jednak automatycznego zapisania dokumentu we wskazanej przez użytkownika lokalizacji, a jedynie wyświetlenie dokumentu. Tym samym opcja ta nie stanowi dodatkowego wsparcia, ponieważ użytkownik i tak musi najpierw wyświetlić dokument, a dopiero potem może go zapisać, korzystając przy tym z narzędzi aplikacji, a nie samego repozytorium. Pewnym plusem tego rozwiązania, w przypadku repozytoriów opartych na systemie DSpace, jest możliwość wykorzystania narzędzi po stronie użytkownika i przejścia od razu do zapisywania pliku (bez wyświetlania dokumentu) po kliknięciu prawym przyciskiem myszy na nazwie pliku, ikonie pliku lub poleceniach w rodzaju „oglądaj / otwórz”, „zobacz / otwórz” czy „przejrzyj / otwórz” i następnie wybraniu polecenia „zapisz element docelowy jako...”. Warunkiem skorzystania z tej funkcji jest jednak odpowiednia wiedza i umiejętności użytkownika (repozytoria nie podpowiadają tej możliwości na odpowiednich ekranach). Operacji tej nie ułatwia fakt braku czytelnego, standardowego formatu nazw plików, stosowanego w poszczególnych repozytoriach.

Istotnym wsparciem dla osób pobierających pliki byłaby precyzyjna informacja o prawnych uwarunkowaniach przejmowania poszczególnych dokumentów i ich późniejszego wykorzystania. Przy czym nie chodzi o ogólne informacje zamieszczone w regulaminach, generalnych zasadach udostępniania czy dokumentach

informacyjnych na temat otwartego dostępu, które wymagają od użytkownika „zejścia” ze swojej ścieżki postępowania i wykonywania szeregu czynności wykraczających zdecydowanie poza najkrótszą drogę do celu (przejścia dokumentu), a o precyzyjną informację na temat tego, czy zapisanie danego pliku na własnym dysku jest legalne, i co zgodnie z prawem można z tym obiektem dalej zrobić. Chociaż w opisach bibliograficznych pojawiają się odpowiednie pola, to znajdujące się w nich wpisy bywają niejasne (szczególnie dla osób nieznających dobrze prawa autorskiego), np. „porozumienie UTP” (pole „licencja”, RCTUP) czy „dla wszystkich w zakresie dozwolonego użytku” (pole „prawa dostępu”, RCTUP). Z kolei w RUW opis zawiera wprawdzie pole „treść licencji”, ale umieszczony w nim link odsyła jedynie do ogólnego regulaminu. Bywa też, że w opisie w ogóle nie przewidziano takich pól. Tak jest np. w AMUR, gdzie pod opisem bibliograficznym jest tylko ogólnikowy zapis: „Wszystkie pozycje w AMUR są chronione prawem autorskim chyba, że zostało wskazane inaczej” (błąd interpunkcyjny pochodzi z oryginału) czy w RUKW. Interesujące rozwiązania znaleziono natomiast w repozytoriach eRIKA, RUŁ i RUMK, gdzie w opisie znajdują się informacje na temat typu licencji Creative Commons, na której dany dokument jest udostępniany (podobnie w RPK), wraz z linkiem do szerszego opisu tej licencji w języku polskim na stronach Creative Commons (<http://creativecommons.org>).

Nie znaleziono żadnych rozwiązań wspierających przejmowanie fragmentów tekstów. Repozytoria operują wyłącznie całymi dokumentami. Oznacza to, że nawet w wyniku wyszukiwania pełnotekstowego znajdowane są całe dokumenty, w których występują zastosowane terminy wyszukiwawcze, jednak bez wskazania miejsca w tych dokumentach, w którym wspomniane terminy się znajdują (jak to się dzieje np. w Google Books). Gorzej, jeśli wyszukiwanie pełnotekstowe na poziomie repozytorium nie zawsze działa, co w kilku przypadkach (AMUR, RUŁ, RUKW) stwierdzono eksperymentalnie – do dostępnego formularza wyszukiwawczego wprowadzano przypadkowe słowa znalezione we wcześniej wybranych w danym repozytorium dokumentach i próbowano te dokumenty ponownie wyszukać, jednak kwerenda nie przyniosła oczekiwanego rezultatu. Weryfikacja tego wniosku oczywiście wymagałaby znacznie szerzej zakrojonych i inaczej zaprojektowanych badań, ale sygnał płynący z tych kilku przypadków jest niepokojący.

### **Przejęcie opisów bibliograficznych całych dokumentów lub ich fragmentów**

Stwierdzono istnienie kilku rozwiązań, które można uznać za próby udzielenia pewnego wsparcia użytkownikowi w zakresie przejmowania opisów bibliograficznych na własne potrzeby (np. tworzenia własnych bibliografii, przygotowywania bibliografii załącznikowych do publikacji itp.). Można dostrzec wyraźny

bezpośredni związek ich dostępności ze stosowanym oprogramowaniem, co może sugerować, że ich udostępnienie nie wynika z woli i dążeń twórców repozytoriów, a z rozstrzygnięć podjętych przez autorów oprogramowania.

Po pierwsze, stwierdzono możliwość skorzystania z funkcji „eksport metadanych” (repozytoria budowane w oparciu o dLibrę), która oferuje 4 formaty opisu (OAI-PMH, RDF, RIS, BIBTEX). Żaden z nich nie jest jednak opisem bibliograficznym sformatowanym do bezpośredniego wykorzystania przez użytkownika akademickiego – wymagają one zastosowania odpowiedniego oprogramowania do zarządzania bibliografiami (np. EndNote). Dla porównania Google Scholar, w ramach funkcji „cytuj”, także proponuje kilka formatów metadanych gotowych do przejścia przy zastosowaniu odpowiedniego oprogramowania, ale jednocześnie wyświetla opisy w kilku wybranych stylach, w pełni przygotowane do wykonania operacji „kopiuj / wklej” i nie wymagające żadnej dodatkowej ingerencji użytkownika. W repozytoriach opartych na dLibrze po zalogowaniu się jest też możliwość wyeksportowania skonstruowanych wcześniej przez siebie całych wykazów bibliograficznych (w ramach funkcji „ulubione publikacje” / „wyeksportuj listę publikacji”). Opisy zawierają jednak tylko tytuły i adresy www, a wyeksportowany dokument jest *de facto* jedynie stroną www na której dopiero trzeba wykonywać operację „kopiuj / wklej” lub przetwarzać treści w inny sposób, by opisy przejść do własnych zasobów.

Po drugie, pojawiają się rozwiązania pozwalające na zapisanie jakiejś formy opisu bibliograficznego bezpośrednio w różnych serwisach społecznościowych, w tym typu *social bookmarking* i przeznaczonych dla środowiska akademickiego (np. Mendeley), przydatne dla tych użytkowników, którzy w którymś z nich budują swoje bibliografie. Przykład opisu przejmowanego do Twittera pokazuje jednak, że mogą to być opisy dalekie od kompletności wymaganej od naukowca lub studenta i w formie mało przydatnej w jego zadaniach akademickich (np. w repozytorium AMUR: nazwa repozytorium, tytuł dokumentu, adres www strony z opisem danego obiektu w repozytorium).

Kolejnym rozwiązaniem jest możliwość przesłania informacji o danym dokumencie mailem na własne konto pocztowe. W repozytoriach opartych na systemie DSpace operacja ta wymusza jednak otwarcie własnego programu pocztowego, skonfigurowanego na danym komputerze, co jest niewygodne choćby z tej racji, że na komputerze, z którego użytkownik sięga do zasobów repozytorium, nie musi być skonfigurowany żaden program pocztowy. Z kolei w repozytoriach opartych na dLibrze wymagane jest wcześniejsze zalogowanie – tylko wówczas pojawia się polecenie „powiadom znajomego”, pozwalające przesłać jakąś postać opisu na własne konto. Niestety, opis jest zapisywany dokładnie w takiej formie, w jakiej widoczny jest na ekranie (a zatem wymagającej dużego nakładu pracy, by

uzyskać z niego opis w postaci wykorzystywanej w komunikacji naukowej) i, co więcej, jeszcze „wzbogacony” nazwą właściciela konta i tytułem „Ciekawe publikacje”. W przypadku Repolis operacja wysłania opisu na własne konto osoby zarejestrowanej jednak nie powiodła się (mimo wielu prób), bez żadnej informacji ze strony systemu o przyczynach problemu.

Istnieje też możliwość przejścia opisu przy pomocy polecenia „zapisz element docelowy jako”, dostępnego po kliknięciu prawym przyciskiem myszy na aktywnym tytule-linku na liście wyników lub w innym wykazie publikacji. W efekcie zapisuje się plik html z informacjami o dokumencie w postaci bardziej lub mniej sformatowanej i czytelnej dla użytkownika. W repozytoriach opartych na dLibrze jest to w zasadzie kopia ekranu z opisem bibliograficznym dokumentu, natomiast w repozytoriach budowanych na bazie systemu DSpace jest to „zrzut” z tego ekranu, ale różnie sformatowany. Bywa nawet, że sam opis bibliograficzny ginie wśród różnych informacji przejętych z tego ekranu (np. eRIKA, RUŁ i RUMK). Co więcej, w obu przypadkach automatyczna nazwa pliku jest raczej nieprzydatna dla PIM użytkownika, gdyż nie identyfikuje treści dokumentu: w pierwszym jest zawsze taka sama (docmetadada.html), a w drugim jest to nazwa numeryczna.

Zgodnie z informacjami dostępnymi w repozytoriach, w niektórych przypadkach po zalogowaniu się można subskrybować informacje o nowych publikacjach. Niestety, próby dokonania takiej subskrypcji w repozytoriach zbudowanych w oparciu o dLibrę nie do końca się powiodły. Otrzymano tylko plik z wykazem nowości z RUW (dotarł do autora dopiero 1.10.2014 roku). Natomiast w repozytoriach zbudowanych na bazie systemu DSpace i w RPK nie każdy może się zalogować (konieczne jest np. podanie numeru karty bibliotecznej, adresu poczty elektronicznej z domeny danej uczelni itp.). Jeśli wierzyć informacji np. w repozytorium eRIKA: „Jakie korzyści daje zarejestrowanie się w systemie? (...) Zaprenumerowanie kolekcji i otrzymywanie e-maili z wiadomością o nowych pozycjach dodanych do kolekcji” lub w repozytorium: AMUR „Użytkownicy mogą otrzymywać e-mailem informacje na temat pozycji nowo dodanych do kolekcji. Użytkownicy mogą zapisać się do nieograniczonej liczby kolekcji”, to taka możliwość istnieje, a przynajmniej jest (była) rozważana przez twórców oprogramowania lub samego repozytorium. Otrzymany z RUW wykaz zawierał opisy dokumentów w zaskakującej postaci: Tytuł – Nazwisko, Imię (kto i gdzie używa opisów bibliograficznych w takiej formie: np. „Pan Tadeusz – Mickiewicz, Adam?”), wzbogacone słowami kluczowymi i informacją na temat praw autorskich (wskazanie właściciela praw i odesłanie do regulaminu). Żeby faktycznie przejść takie opisy do własnych zbiorów bibliograficznych, użytkownik i tak musi je przetworzyć, a następnie uzupełnić o informacje znalezione w innych źródłach, osobno dla każdego dokumentu.

Niektóre repozytoria oferują również możliwość prenumerowania kanałów RSS informujących o nowościach lub planach, jednak zawartość i struktura oferowanych w nich opisów są różne i raczej nieodpowiednie dla naukowca czy studenta. Zawierają wybrane elementy opisu, niewystarczające dla celów komunikacji naukowej, np. tytuł (link do dokumentu), nazwisko i imię autora (jeśli jest), datę wprowadzenia do repozytorium czy abstrakt (w repozytoriach bazujących na systemie DSpace).

Niezależnie od rozwiązań wspierających (lub pozornie wspierających) przejmowanie opisów bibliograficznych, zawsze pozostaje możliwość wydawałoby się najprostsza: „zamalowanie” opisu wyświetlonego na ekranie po wyszukaniu danego dokumentu i przejścia (wklejenia) go do tworzonej przez siebie w ramach PIM bibliografii (np. do analizy stanu badań). Jednak format i postać opisów w każdym z badanych przypadków są tak zaprojektowane, że zdecydowanie mniejszego wysiłku i nakładu pracy wymaga ich ręczne przepisanie z ekranu i następnie wprowadzenie do własnych dokumentów niż kopiowanie i przetwarzanie do pożądanego postaci, co wymagałoby wielu dodatkowych czynności: usuwania zbędnych treści i znaków formatowania, zmiany w kolejności elementów, ujednolicenia czcionek, usuwania elementów graficznych czy konsekwencji umieszczania opisów w tabelach, dołączania brakujących elementów z innych ekranów (np. z ekranów poszerzonego opisu) itp. Żeby przekonać się o skali trudności i zbędnej pracy (przecież ten opis już ktoś w repozytorium raz zrobił!) wystarczy wykonać prosty eksperyment: skopiować do nowego dokumentu („kopiuj / wklej”) opis publikacji „Zmierzch państw narodowych, czyli finał klasycznej Europy” dostępnej w RUŁ, a następnie przetwarzać go tak długo, aż uzyska się postać wymaganą od autorów publikacji np. w „Przeglądzie Bibliotecznym”.

W pracy akademickiej często zachodzi konieczność opatrzenia opisem bibliograficznym fragmentu tekstu (np. przypisy do cytatów, informacja o źródle informacji w trakcie robienia „notatek” metodą „kopiuj / wklej” fragmentów publikacji itp.). Niestety, w badanych repozytoriach nie znaleziono żadnych narzędzi pozwalających na dołączanie opisów (w jakiegokolwiek postaci i formie) do przejmowanych fragmentów tekstu. Wiąże się to z sygnalizowaną wcześniej zasadą, że repozytoria operują na całych obiektach (książka, artykuł itd.) i nie oferują narzędzi zarządzania informacją niezależnie od stosowanych „kontenerów wiedzy”. Pod tym względem ich funkcjonalność niewiele różni się od kolekcji zasobów drukowanych, gdzie z oczywistych względów nie udostępniano poszczególnych kartek książki czy wybranych fragmentów artykułów z własnymi opisami bibliograficznymi. Nie znaleziono rozwiązań pozwalających na sprawne funkcjonowanie repozytoriów w środowisku informacyjnym, charakteryzującym się dez-

integracją tradycyjnych kontenerów wiedzy na rzecz swobodnego zarządzania wybranymi fragmentami dokumentów, „remiksowania” ich treści i rekreowania kontenerów przez użytkowników według własnych potrzeb.

### **Uzyskanie ponownego dostępu (w ramach danego systemu i poza nim) do wcześniej wyszukanych dokumentów**

Można założyć, że w niektórych sytuacjach użytkownik nie będzie chciał przejmować całych dokumentów lub ich opisów bibliograficznych do własnego PSI, a zdecyduje się jedynie na zapewnienie sobie szybkiego ponownego dostępu do wcześniej wyszukanych zasobów. Warto przyjrzeć się dwóm sytuacjom: pierwszej, w której użytkownik buduje poza repozytorium własną kolekcję linków do tych zasobów, i drugiej, gdy tworzy indywidualną kolekcję linków bezpośrednio w systemie repozytorium. W obu przypadkach musi wprowadzić każdy link jakoś opisać, ale nie potrzebuje pełnych i odpowiednio sformatowanych opisów bibliograficznych wykorzystywanych w komunikacji naukowej, bo celem nie jest tworzenie własnej bibliografii czy przypisów, a tylko identyfikacja danego obiektu.

Wsparcie w przypadku zaistnienia pierwszej z wymienionych sytuacji mogą być te wszystkie funkcje, które zostały omówione przy okazji analizy rozwiązań umożliwiających przejmowanie opisów bibliograficznych (np. zapisywanie informacji w serwisach społecznościowych, przesyłanie mailem na własny adres). Jeśli w tamtej roli raczej się nie sprawdzają, to już jako narzędzia organizowania ponownego dostępu mogą okazać się przydatne, szczególnie jeśli użytkownik opiszze pozyskiwane informacje swoimi nazwami, umieści je w odpowiednich własnych kolekcjach itd. W przypadku repozytoriów budowanych w oparciu o dLibrę można też kliknąć prawym przyciskiem myszy na obrazku reprezentującym dokument (na ekranie, gdzie jest jego opis) i wybrać polecenie „zapisz element docelowy jako”. W efekcie zapisany zostanie plik html o nazwie „info” (dla każdego obiektu zawsze ta sama nazwa pliku), który nie zawiera samego dokumentu, ale jego otwarcie powoduje automatyczne załadowanie pliku tego dokumentu.

Druga sytuacja zachodzi tylko wtedy, gdy repozytorium umożliwia swoim użytkownikom tworzenie własnych kolekcji linków w systemie. Opcja ta jest z całą pewnością dostępna w trzech repozytoriach opartych na dLibrze, a z informacji wynika, że także w RPK. Natomiast nic nie wskazuje na to, by taką funkcję oferowały repozytoria oparte na systemie DSpace (wniosek na podstawie informacji dostępnej w badanych systemach – utworzenie własnego konta wymaga zarejestrowania się i jest dostępne tylko dla uprawnionych użytkowników). W repozytoriach budowanych z wykorzystaniem dLibry własne konto pozwala na tworzenie wykazów tych dokumentów, które w trakcie przeszukiwania zasobów

użytkownik uzna za wartę ponownego przejrzania (polecenie „dodaj do ulubionych”). Do automatycznie zapisywanego tytułu takiego dokumentu można dodać własne tagi i w przyszłości wykorzystać je do ponownego odszukania danej publikacji. Tytuły w tak skonstruowanym wykazie są automatycznie linkami do konkretnych dokumentów.

Wskutek przyjęcia wspomnianej wcześniej zasady operowania wyłącznie na poziomie całych kontenerów wiedzy repozytoria nie oferują narzędzi umożliwiających sprawny ponowny dostęp do fragmentów wcześniej wyszukanych obiektów.

### **Utrzymywanie dokumentu lub jego opisu w obszarze percepcji użytkownika**

Nie udało się znaleźć narzędzi wyraźnie zaprojektowanych dla utrzymania konkretnych dokumentów lub ich opisów w obszarze percepcji użytkownika, choć z pewnością niektóre rozwiązania mogą być wykorzystywane właśnie w tym celu.

Badane repozytoria w zasadzie nie oferują możliwości szybkiego i prostego wydruku czytelnie sformatowanej strony zawierającej informację o dokumencie, którą można by umieścić w miejscu widocznym dla użytkownika. Użycie poleceń plik / drukuj w trakcie przeglądania ekranu z opisem bibliograficznym prowadzi do bardzo różnych efektów: od wydruku ekranu w postaci takiej, jaka jest widoczna w komputerze (w repozytoriach budowanych w oparciu o dLibrę), przez częściowo sformatowane wydruki w repozytorium AMUR, do bardzo „zaśmieconego” wydruku (aż czterostronicowego, w którym najważniejsze informacje są dopiero na czwartej stronie) w repozytorium eRIKA. W RPK dostępna jest wprawdzie ikonka drukarki, ale jej kliknięcie prowadzi do prostego wydrukowania całej strony w postaci takiej, w jakiej jest ona dostępna na ekranie.

W tych repozytoriach, w których na stronach opisu dostępne są też pliki graficzne identyfikujące dany dokument (z reguły pierwsza strona, okładka) – częściowo w RCUTP, RUW i w niektórych opisach w RPK – można ich użyć do zrobienia tapety lub zamieszczenia na pulpicie, wykorzystując opcje dostępne pod prawym przyciskiem myszy. Oczywiście pobrany plik z dokumentem lub jakąś formą opisu (w postaci omawianej wcześniej) też można zapisać na pulpicie pod znaczącą dla użytkownika nazwą.

Również wysłanie maila na własny adres z jakąś formą informacji o dokumencie było już przedstawione wcześniej. Z punktu widzenia funkcji przypominania interesującym rozwiązaniem wydaje się automatyczne narzucanie tematu maila, zbudowanego z nazwy repozytorium i tytułu dokumentu (AMUR, RUMK). Z kolei w RUŁ taki mail zyskuje temat zbudowany z samego adresu www, który – z oczywistych względów – nie ma w sobie takiego potencjału dla wspomaganie pamięci użytkownika o danym dokumencie, jak tytuł.

W repozytoriach budowanych w oparciu o dLibrę można po zalogowaniu dodatkowo skorzystać z funkcji „udostępnij ulubione publikacje swoim znajomym”, która między innymi pozwala za pomocą „komponentu DLibry” umieścić informacje o dokumencie na swojej witrynie www (RCUTP, RUW).

## Wnioski

Przeprowadzone rozpoznanie form wsparcia oferowanych przez repozytoria uczelniane dla wybranych, podstawowych działań podejmowanych przez użytkownika w ramach indywidualnego zarządzania informacją ujawniło zarówno wyraźne niedostatki, jak i przykłady interesujących rozwiązań, choć być może w niektórych przypadkach wymagających jeszcze dopracowania.

Nawet jeśli repozytoria zostały wyposażone w funkcje z założenia mające ułatwiać przejmowanie dokumentów czy opisów bibliograficznych użytkownikom, to zastosowane rozwiązania szczegółowe (np. sposób przejścia pliku, format opisu) czynią z nich narzędzia niezbyt przydatne, a w niektórych przypadkach wręcz utrudniające wykonanie takich operacji. Co więcej, zastosowane rozwiązania w zakresie prezentowania opisów bibliograficznych nie tylko nie ułatwiają, ale i zdecydowanie utrudniają ich przejście w ramach najprostszej procedury „kopiuj / wklej”. Można odnieść wrażenie, że tak jak wcześniej twórcy bibliografii dorobku pracowników uczelni [Sapa 2011], tak teraz twórcy repozytoriów wyraźnie nie dążą do dostarczania użytkownikom gotowych (poprawnych z punktu widzenia potrzeb komunikacji naukowej) opisów bibliograficznych. Natomiast z pewnością na uwagę zasługuje umożliwienie – w niektórych przypadkach – przejmowania opisów bezpośrednio do odpowiedniego oprogramowania służącego do zarządzania bibliografiami (np. EndNote) oraz w miarę precyzyjne informowanie o charakterze licencji, zgodnie z którą dany dokument został udostępniony. Wyraźnie tradycyjne podejście do zarządzania informacją, bazujące na zarządzaniu jej standardowymi „kontenerami” (dokumentami), a nie na zarządzaniu treścią, prowadzi do braku narzędzi wspierających użytkownika chcącego operować na ich mniejszych fragmentach.

Interesujące są różne rozwiązania wspierające proces uzyskiwania ponownego dostępu do zasobów odnalezionych w repozytoriach, szczególnie te dobrze osadzone we współczesnym środowisku informacyjnym i organizowane w ramach systemu repozytorium. Można jednak odnieść wrażenie, że brakuje celowego podejścia do ich projektowania, określenia konkretnych działań i zachowań, które mają je ułatwiać, co przekłada się często na ich wyraźne niedopracowanie. Ten brak jest może jeszcze bardziej widoczny w przypadku rozwiązań służących



utrzymywaniu dokumentu lub jego opisu w obszarze percepcji użytkownika – nie znaleziono żadnych narzędzi dedykowanych tego typu funkcjom, więc użytkownikowi pozostaje wykorzystanie rozwiązań zaprojektowanych przede wszystkim do innych zadań lub tych o niejasnym przeznaczeniu (np. wydruk ekranu z opisem bibliograficznym sformatowanym w sposób utrudniający odczytanie tego opisu).

Wydaje się, że mimo różnych ograniczeń, w badanych repozytoriach uczelnianych istnieje duży potencjał dla wspierania indywidualnego zarządzania informacją. By go w pełni wykorzystać, konieczne jest jednak uwzględnienie rzeczywistych potrzeb w tym zakresie, a szczególnie bardziej celowe podejmowanie decyzji w oparciu o precyzyjne określenie, komu i do czego konkretnie mają służyć oferowane rozwiązania, by mogły sprzyjać uzyskaniu synergii między systemem repozytorium a PIM użytkownika.

## BIBLIOGRAFIA

- Arth, Janet et al. (2011). *Report of the current awareness / personal information management (CA/PIM) subgroup of the Web Services Steering Committee* [dok. elektr.]. [http://conservancy.umn.edu/bitstream/11299/102255/1/Current%20awareness\\_personal%20info%20mgmt%202009.pdf](http://conservancy.umn.edu/bitstream/11299/102255/1/Current%20awareness_personal%20info%20mgmt%202009.pdf) [odczyt 03.10.2014].
- Barreau, Deborah K. (1995). *Context as a factor in personal information management systems*. *Journal of the American Society for Information Science* Vol. 46, No. 5, p. 327-339.
- Bergman, Ofer et al. (2013). *Folder versus tag preference in personal information management*. *Journal of the American Society for Information Science and Technology* Vol. 64, No 10, p. 1995-2012.
- Bruce, Harry; Jones, William; Dumais, Susan (2004). *Information behaviour that keeps found things found* [dok. elektr.]. *Information Research* Vol. 10, No. 1, paper 207. <http://InformationR.net/ir/10-1/paper207.html> [odczyt 03.10.2014].
- Copeland, Andrea J. (2011). *Analysis of public library users' digital preservation practices*. *Journal of the American Society for Information Science and Technology* Vol. 62, No. 2, p. 1288-1300.
- Fourie, Ina (2011). *Personal information and reference management: librarians' increasing creativity*. *Library Hi Tech* Vol. 29, No. 2, p. 387-393.
- Jones, William (2007). *Personal information management*. *Annual Review of Information Science and Technology* Vol. 41, No. 1, p. 453-504.
- Jones, William (2008). *Keeping found things found: the study and practice of personal information management*. Amsterdam, Boston: Morgan Kaufmann Publishers, 448 p.
- Newton-Smith, Carol J. (2000). *The personal information management behaviour of academics: implications for librarians' support*. *Master thesis* [dok. elektr.]. [https://repository.uwa.edu.au/R/-?func=dbin-jump-full&object\\_id=9364&local\\_base=GEN01-INS01](https://repository.uwa.edu.au/R/-?func=dbin-jump-full&object_id=9364&local_base=GEN01-INS01) [odczyt 03.10.2014].

- Otopah, Francis O.; Dadzie, Perpetua (2013). *Personal information management practices of students and its implications for library services*. Aslib Proceedings Vol. 65, No. 2, p. 143-160.
- Robinson, Sara; Johnson, Frances (2012). *The process and affective environment of students' personal information management* [dok. elektr.]. *Enhancing Learning in the Social Sciences* Vol. 4, No. 2. <http://journals.heacademy.ac.uk/doi/abs/10.11120/elss.2012.04020005> [odczyt 03.10.2014].
- Sapa, Remigiusz (2011). *Warsztat naukowca a problem formatu informacji bibliograficznej generowanej przez systemy informacyjne*. W: Pietruch-Reizes, Diana; Babik, Wiesław; Frączek, Renata red. *Bezpieczna, innowacyjna i dostępna informacja: perspektywy dla sektora usług informacyjnych w społeczeństwie wiedzy*. Katowice: Polskie Towarzystwo Informacji Naukowej, s. 175-183.
- Sapa, Remigiusz (2013). *Realizacja funkcji repozytoryjnych przez największe przedsięwzięcia zarejestrowane w Federacji Bibliotek Cyfrowych tworzone i współtworzone przez uczelnie*. *Przegląd Biblioteczny* nr 2, s. 117-132.
- Świgoń, Marzena (2012). *Zarządzanie wiedzą i informacją: podstawy teoretyczne, badania w wymiarze indywidualnym*. Olsztyn: Wydaw. Uniwersytetu Warmińsko-Mazurskiego, 367 s.
- Whittaker, Steve (2011). *Personal information management: from information consumption to curation*. *Annual Review of Information Science and Technology* Vol. 45, No. 1, p. 3-62.

#### ANEKS 1.

#### WYKAZ REPOZYTORIÓW UWZGLĘDNIONYCH W BADANIACH

L.p.	Repozytorium	Adres www	Kod	System
1	RePolis. Repozytorium Publikacji Naukowych Politechniki Śląskiej / Repozytorium Cyfrowe Politechniki Śląskiej Repolis (na str. głównej jest kilka nazw!)	<a href="http://repolis.bg.polsl.pl/dlibra">http://repolis.bg.polsl.pl/dlibra</a>	RePolis	dLibra
2	Repozytorium Cyfrowe UTP w Bydgoszczy (NIE DZIAŁA 21.08, 24.08, 26.08, zadziałało 26.08 po 12.00)	<a href="http://dlibra.utp.edu.pl/dlibra">http://dlibra.utp.edu.pl/dlibra</a>	RCUTP	dLibra
3	eRIKA. Repozytorium Instytucjonalne Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego	<a href="http://repozytorium.ka.edu.pl">http://repozytorium.ka.edu.pl</a>	eRIKA	DSpace
4	Repozytorium Politechniki Krakowskiej	<a href="http://suw.biblos.pk.edu.pl">http://suw.biblos.pk.edu.pl</a>	RPK	własny
5	AMUR. Repozytorium Uniwersytetu im. Adama Mickiewicza	<a href="https://repozytorium.amu.edu.pl/jspui">https://repozytorium.amu.edu.pl/jspui</a>	AMUR	DSpace
6	Repozytorium Uniwersytetu Kazimierza Wielkiego w Bydgoszczy	<a href="http://repozytorium.ukw.edu.pl">http://repozytorium.ukw.edu.pl</a>	RUKW	DSpace
7	Repozytorium Uniwersytetu Łódzkiego	<a href="http://dspace.uni.lodz.pl:8080/xmlui">http://dspace.uni.lodz.pl:8080/xmlui</a>	RUŁ	DSpace
8	Repozytorium Uniwersytetu Mikołaja Kopernika	<a href="http://repozytorium.umk.pl">http://repozytorium.umk.pl</a>	RUMK	DSpace
9	Repozytorium Uniwersytetu Wrocławskiego	<a href="http://www.repozytorium.uni.wroc.pl/dlibra">http://www.repozytorium.uni.wroc.pl/dlibra</a>	RUW	dLibra

---

# Archiwa i repozytoria Open Access we Francji. Analiza piśmiennictwa z lat 2005-2013

Agnieszka Korycińska-Huras

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

Artykuł zawiera syntetyczne omówienie wyników badań nad stanem rozwoju francuskich archiwów dziedzicznych i repozytoriów instytucjonalnych w świetle piśmiennictwa opublikowanego w latach 2005-2013. Kwerendy bibliograficzne, skoncentrowane głównie na zasobach oferowanych w swobodnym dostępie, prowadzone były w serwisach OAister, DRIVER, OpenAIRE, Google Scholar, w bazach Francis, Emerald Insight, Academic Search Complete, w bibliografii BABIN oraz w licznych webografiach dostępnych w internecie. W pierwszym etapie wyszukiwań zastosowałam strategię formowania klas z użyciem różnorodnych kombinacji argumentów w języku francuskim i angielskim: *libre accès*, *archives ouvertes*, *archives institutionnelles*, *autoarchivage*, *voie verte*, *valorisation de la recherche*, *open access*, *open archives*, *open repositories*, *self-archiving*. Jako główny termin wyszukiwawczy przyjąłam *archive*, z przymiotnikami *thématique* (dziedziczne) i *institutionnelle* (instytucjonalne), ponieważ on właśnie przyjął się powszechnie we francuskiej nomenklaturze. W piśmiennictwie sporadycznie pojawiają się synonimy lub *quasi-synonimy* *répertoire*, *dépôt*, *entrepôt*, oznaczające składnicę / zbiornicę / bazę. W kolejnym kroku, po wyselekcjonowaniu wyjściowego zbioru relewantnych dokumentów, dopełniałam wyniki kwerendy, korzystając ze strategii indeksów cytowań [Szczepańska 2007, dok. elektr.]. W wyborze publikacji, które ostatecznie poddałam analizie, preferowałam z rozmysłem prace relacjonujące wyniki badań empirycznych na dużych próbach oraz artykuły o charakterze podsumowującym i przeglądowym, pod względem formalnym przeważnie dokumenty niesamoistne piśmienniczo.

Na początek kilka najważniejszych faktów z historii. Pierwsze francuskie inicjatywy w zakresie autoarchiwizacji przypadają na koniec lat 90. XX wieku: w 1998 r. na Uniwersytecie Lumière Lyon 2 utworzono bazę pełnotekstową prac doktorskich Cyberthèses. Sygnałem rosnącego zainteresowania środowisk akademickich i naukowych ideą Otwartego Dostępu była konferencja pod hasłem „Publication électronique des résultats de la recherche” („Elektroniczne publikowanie wyników badań naukowych”), zorganizowana w Paryżu w marcu 2000 r. Już w czerwcu tego samego roku w ramach CNRS (Centre national de recherche scientifique, Krajowe Centrum Badań Naukowych) utworzono odrębną jednostkę CCSD (Centre pour la Communication Scientifique Directe, Centrum Bezpośredniej Komunikacji Naukowej), które w listopadzie 2000 r. inaugurowało serwis HAL (HyperArticles en Ligne). Pierwotnie funkcjonował on jako *mirror* serwisu ArXiv, potem z inicjatywy Francka Laloë przekształcił się w archiwum dziedzinowe przeznaczone dla francuskich fizyków, by ostatecznie zyskać rangę centralnego (narodowego) wielodyscyplinarnego archiwum e-printów, gromadzącego publikacje wszystkich uczonych, niezależnie od obywatelstwa, afiliowanych we francuskich placówkach naukowych i akademickich, prywatnych i publicznych [Foulonneau et al. 2008, dok. elektr.].

Niebawem wdrożono kolejne projekty: Uniwersytet Lyon 2 stowarzyszył się z kilkoma innymi uczelniami, by wspólnie otworzyć repozytorium Cyberdoc, z pomocą CCSD powstały: repozytorium paryskiego Instytutu Jean Nicod i archiwum dziedzinowe @rchiveSIC skupiające piśmiennictwo z zakresu informatologii i bibliologii. Własne bazy e-printów założyły: Institut français d'exploitation de la mer (IFREMER) – Archimer, Institut national de la recherche agronomique (INRA) – ProdINRA, Institut de recherche pour le développement (IRD) – Horizon/Pleins textes. W roku 2006 sześć uczelni i samodzielnych instytutów naukowo-badawczych podpisało Deklarację Berlińską: Lumière Lyon 2, Strasbourg, Lille 1, ParisTech (Institut des Sciences et Technologies de Paris), Institut national polytechnique de Toulouse (INPT) i INSA Lyon (Institut national des sciences appliquées de Lyon), a następnie utworzyło repozytoria prac doktorskich [Bourrion et al. 2006, dok. elektr.; Rigeade 2012, dok. elektr.; Suber 2009, dok. elektr.; Chronologie... 2014, dok. elektr.].

Lurent Romary miał rację, gdy w przedmowie do wydanego w 2005 r. tomu zbiorowego „Les archives ouvertes. Enjeux et pratiques. Guide à l'usage des professionnels d'information” pisał, że praca ta daje „zatrzymany w kadrze obraz wybitnie rozwojowej sytuacji, dlatego za kilka lat będzie cennym narzędziem dla badaczy historii ruchu Open Access” [Aubry, Janik eds. 2005, p. 8]. Rzeczywiście, tom pod redakcją Martine Aubry i Joanny Janik był pierwszym we Francji opracowaniem podsumowującym tamtejsze dokonania praktyczne i stan badań

nad inicjatywami na rzecz otwartego dostępu do nauki, pomyślanym jako przewodnik dla bibliotekarzy i pracowników informacji. W kontekście odtwarzania rozwoju otwartych archiwów we Francji interesujący jest w szczególności artykuł Gabriela Gallezota „Les archives ouvertes en France” [Gallezot 2005]. Na wstępie autor definiuje i typologizuje otwarte archiwa, proponując przyjęcie konsensusu, w myśl którego *archives ouvertes* będą rozumiane jako: 1) w ujęciu wąskim – składnice artykułów (pre-printów i post-printów) umieszczanych tam na zasadzie autoarchiwizacji, dostępne swobodnie, bez ograniczeń prawnych i ekonomicznych, oparte na protokole OAI-PMH oraz 2) w ujęciu bardziej elastycznym – jako zbiory wszelkich materiałów służących tworzeniu wiedzy naukowej, niekoniecznie pochodzących z autoarchiwizacji, pod warunkiem że dostępu do nich nie obwarowuje się żadnymi restrykcjami, a transfer danych obsługiwany jest przez interoperacyjny protokół OAI-PMH.

Zgodnie z tą szerszą definicją w dalszych rozważaniach Gallezot zalicza do kategorii otwartych archiwów również biblioteki cyfrowe i platformy publikacji czasopism elektronicznych, jako że za nadrzędne kryterium włączające uznaje wolny dostęp do treści oraz kompatybilność systemu z agregatorami pobierającymi metadane. W ramach dalszych ustaleń porządkujących dzieli archiwa na instytucjonalne, dziedziczne, tematyczne oraz metaarchiwa, czyli serwisy rozproszone i agregatory (fr. *moissonneurs*, dokładny odpowiednik ang. *harvesters*). W uwagach poprzedzających ilościowy opis stanu realizacji projektów OA we Francji zaznacza, że zadanie to jest trudne z racji znacznego rozproszenia podmiotów odpowiedzialnych za tworzenie otwartych archiwów oraz braku jakiegokolwiek struktury scalającej, która umożliwiłaby im porozumienie i koordynację działań. Gallezot ubolewał, że w 2005 r., czyli 7 lat od uruchomienia pierwszego repozytorium prac doktorskich Cyberthèses, francuskie środowisko zwolenników OA nie doczekało się żadnej organizacji federującej, ani nawet wspólnego serwisu internetowego lub blogu. Wobec braku innych źródeł autorytatywnej informacji o liczbie i kondycji francuskich archiwów OA Gallezot sporządził ich listę tylko na podstawie skrzyżowanych wyszukiwań w serwisach OAForum, OpenArchive.org i OAister. Zastrzegł przy tym, że nie można uznać jej za kompletną, ponieważ nie znalazły się na niej archiwa dopiero powstające, projekty nigdzie nie zgłoszone przez twórców, wewnętrzne, zamknięte repozytoria różnych instytucji oraz platformy publikacji czasopism naukowych OA, których wykorzystane w kwerendzie serwisy nie indeksowały. Zestawiony przez Gallezota wykaz obejmował 19 jednostek, w tym bibliotekę cyfrową Gallica, z największą kolekcją pełnych tekstów w liczbie 10 000. Najwięcej, bo aż 9, podlegało formalnie ośrodkowi CCSD. We wnioskach końcowych Gallezot stwierdzał wprawdzie, że z uwagi na niemałą liczbę działających archiwów i wspieranie idei Otwartego

Dostępu przez kilka ważnych instytucji krajowych „Francja nie pozostaje w tyle”, jednak zarazem apelował o stanowcze wsparcie polityczne dla inicjatyw tego rodzaju, a przede wszystkim o pilne wypracowanie narzędzi służących wymianie pionierskich doświadczeń i propagowaniu autoarchiwizacji. Proponował w tym celu stworzenie oficjalnej strony internetowej oraz wspólnego agregatora z bramką dostępową, który penetrowałby wszystkie krajowe archiwa dokumentujące działalność podmiotów finansowanych z kiesy publicznej.

Kolejną próbę porachowania francuskich archiwów OA podjęła rok później grupa studentów ENSSIB (Ecole nationale supérieure des sciences de l'information et des bibliothèques), którzy posiłkując się danymi z archives.eprints.org i DOAR, doliczyli się ich, odpowiednio, 28 i 17. Niemożność jednoznacznego ustalenia stanu faktycznego autorzy tego studium przypisywali, podobnie jak Gallezot, brakowi oficjalnych statystyk krajowych, wynikającemu z dotychczasowej biernej postawy centralnych instytucji odpowiadających za organizację nauki [Bourrion et al. 2006, p. 17].

W połowie I dekady XXI w. obawy o skutki nadmiernej fragmentacji i nieuporządkowania działań realizujących założenia otwartego dostępu zaczęto wyrażać także w kręgu francuskich instytucji naukowych wysokiego szczebla. We wrześniu 2005 r. CNRS wraz z trzema krajowymi instytutami naukowo-badawczymi (INRA, INSERM, INRIA) i Konferencją Rektorów Uniwersytetów (Conférence des Présidents d'Université, CPU) zawarły porozumienie o rozpoczęciu prac przygotowawczych w celu uruchomienia wspólnego portalu repozytoryjnego opartego na platformie softwarowej HAL, która gwarantowałaby interoperacyjność z innymi archiwami OA. Zamierzenie to zostało skonkretyzowane w „Memorandum na rzecz skoordynowanego podejścia na poziomie krajowym do otwartej archiwizacji dorobku nauki” („Protocole d'accord en vue d'une approche coordonnée, au niveau national, pour l'archivage ouvert de la production scientifique”) podpisanym w lipcu 2006 r. przez pierwotną grupę sygnatariuszy, do których dołączyły wkrótce kolejne cztery placówki badawcze oraz Konferencja Rektorów *Grandes Ecoles* (CGE), a wiosną 2007 r. jeszcze czterech nowych partnerów. W „Memorandum”, mającym obowiązywać przez dwa lata, przewidziano utworzenie specjalnego komitetu sterującego (COSTRAO), w którym resort rządowy reprezentowali przedstawiciele Ministerstwa Szkolnictwa Wyższego i Badań Naukowych (dalej: MESR), oraz komitetu naukowego o nazwie COST. W dokumencie podkreślano, że wdrożenie portalu repozytoryjnego HAL ma służyć przede wszystkim autoarchiwizacji naukowego dorobku piśmienniczego francuskich uczonych, a w następstwie zwiększeniu widoczności i kapitalizacji wyników badań w międzynarodowej komunikacji naukowej [Zespół ICM 2011, dok. elektr.; Creppy 2007, dok. elektr.; Bourrion et al. 2006, dok. elektr.].

Na efekty „Memorandum”, w ostateczności zresztą dość ograniczone, trzeba było czekać. W rok po jego przyjęciu określenie dokładnej liczby francuskich baz piśmiennictwa w wolnym dostępie nadal nastęczało trudności. Pierre Baruch skompilował je na podstawie danych z DOAR, ROAR, listy opublikowanej przez MESR oraz serwisu archives-ouvertes.fr. Zidentyfikował w ten sposób 40 kolekcji, które podzielił na: a) repozytoria instytucjonalne – 24, b) repozytoria prac magisterskich, doktorskich i habilitacyjnych – 10, c) biblioteki cyfrowe udostępniające zdigitalizowane wersje dokumentów i artefaktów o wartości historycznej – 9, d) platformy publikacji czasopism elektronicznych – 5, e) bazy danych – 2. W grupie czwartej, obejmującej serwisy publikacji i udostępniania czasopism elektronicznych, umieścił także CNR INIST i CAIRN, mimo że ich oferta była wówczas (i jest nadal) w większości odpłatna. Wysilek Barucha był cenny z uwagi na metodyczność i wielostronność przeprowadzonych kwerend, niemniej uzyskanym przez niego wynikiem trudno jest w pełni zaufać, gdyż sam autor zastrzega, że dostępne źródła nie pozwoliły mu dokładnie ustalić np. udziału pełnych tekstów w poszczególnych kolekcjach. Rachel Creppy podaje, że w roku 2006 wyszukiwanie w bazie systemu Driver wykazywało 20 repozytoriów instytucjonalnych we Francji, z których 86% zawierało pełne teksty prac magisterskich i doktorskich, a 71% także pełne teksty artykułów naukowych, w większości pre-printów (71%). Wszystkie zarejestrowane bazy oferowały swobodny dostęp do materiału, a w 28% z nich istniały również wydzielone segmenty zasobu zastrzeżone hasłem [Creppy 2007, dok. elektr.].

W kwietniu 2007 r. podjęto pierwszą próbę empirycznej oceny wpływu postanowień „Memorandum” na tworzenie repozytoriów przez francuskie szkoły wyższe i placówki naukowo-badawcze. Inicjatorem badania było konsorcjum Couperin (Consortium universitaire de publication numérique, Akademickie Konsorcjum Elektronicznego Publikowania) zrzeszające wówczas wszystkie francuskie uniwersytety, 70 szkół wyższych o innym statusie i liczne instytuty badawcze. Na ankietę skierowaną do wszystkich 210 instytucji naukowych we Francji odpowiedziała tylko jedna trzecia badanych, tj. 74 placówki, wśród których 62% stanowiły uniwersytety i 21% uczelnie typu *grande école*. Po kwantyfikacji wyników okazało się, że 53% instytucji dopiero rozważało możliwość stworzenia repozytorium lub pracowało nad jego wstępnym projektem, 18% było w fazie implementacji, a 29% już posiadało działające repozytorium. W ankiecie proszono o wskazanie wszystkich przedsięwzięć OA wdrażanych lub planowanych w danej placówce, zakładając że może być ich kilka. W ten sposób uzyskano łącznie 87 zgłoszeń projektów, z których najwięcej, 48%, dotyczyło repozytoriów prac magisterskich, doktorskich i habilitacyjnych, 10% zakładało tworzenie baz przeznaczonych *stricte* do autoarchiwizacji bieżącego dorobku kadry naukowej (tylko

pre- i post-rinty publikacji naukowych), pozostałe zaś miały charakter mieszany [Bruley et al. 2008, dok. elektr.]. Tylko 16 respondentów zadeklarowało planowany lub rzeczywisty udział w projektach międzyinstytucjonalnych. Dla 33 repozytoriów uściślono liczbę dokumentów pełnotekstowych zdeponowanych w ciągu minionych 12 miesięcy. W połowie przypadków skumulowany roczny wpływ wynosił poniżej 100 jednostek, 5 repozytoriów odnotowało wynik „0”, a w 4 nie było żadnego pełnego tekstu. Wśród 39 projektów, dla których w odpowiedzi na ankietę Couperin podano ogólną liczbę zarchiwizowanych pełnych tekstów, próg 1500 osiągnęły 4 repozytoria, a największy zbiór sięgał 19 545 jednostek.

W odniesieniu do 19 projektów zasygnalizowano, że w stosownych regulaminach obowiązujących w podmiocie odpowiedzialnym przewidziany został nakaz autoarchiwizacji (fr. *obligation de dépôt*, ang. *deposit mandate*, otwarty mandat<sup>1</sup>) określonych typów dokumentów (ciążący na autorach i/lub bibliotekach bądź ośrodkach dokumentacji). Jedynie dla 9 repozytoriów ankietowani przedstawili statystyki odwiedzin za ostatni rok, z rozpiętością między 200 a 130 000 połączeń, niemniej przedstawiciele Couperin komentujący wyniki badania uznali te dane za wystarczające do stwierdzenia, że potencjalny popyt na zawartość repozytorium jest „bardzo znaczny”. Przytłaczająca większość zwróconych ankiet (67 z 87) została wypełniona przez pracowników bibliotek i ośrodków dokumentacji, mimo że formularze były jawnie adresowane do wszystkich komórek organizacyjnych danej instytucji, włączonych w jakimkolwiek stopniu w projekt budowy repozytorium. Równoległe na pytanie o główne przeszkody w jego implementacji najwięcej odpowiedzi wskazywało na trudności w komunikacji wewnątrzinstytucjonalnej, niedostateczne zaangażowanie naukowców i brak koordynacji między jednostkami organizacyjnymi, którym powierzono pieczę nad projektem. Autorów raportu z badań Couperin skłoniło to do przypuszczenia, że bibliotekarze i dokumentaliści, będący najwyraźniej kołem zamachowym rozwoju OA we francuskim środowisku akademickim i naukowym, pozostają na tym polu osamotnieni. Prócz omówionych już aspektów w raporcie Couperin przedstawiono statystyki i zestawienia obrazujące rozkład respondentów według typu instytucji i reprezentowanych dyscyplin naukowych, hierarchię deklarowanych celów tworzenia repozytorium (na pierwszym miejscu wymieniano rozpowszechnianie i kapitalizację wartości wyników badań naukowych, fr. *valorisation*), sposoby

---

<sup>1</sup> Określenie „otwarty mandat”, wypromowane w kolejnych publikacjach przez Krzysztofa Siewicza [Siewicz 2012, 2013, dok. elektr.], pojawia się także w najnowszym raporcie ICM UW „Otwarta nauka w Polsce 2014. Diagnoza” [Szprot red. 2014, dok. elektr.]. Krzysztof Siewicz uściśla, że wyrażenie „otwarty mandat” jest wprawdzie kalką językową angielskiego *open-access mandate*, jednak decyduje się go używać, gdyż „funkcjonuje [ono] dość powszechnie w praktyce” [Siewicz 2012, dok. elektr., s. 13].



promowania praktyki autoarchiwizacji w lokalnej społeczności naukowej, jednostki organizacyjne podmiotu współodpowiedzialne za wykonanie projektu. Opisano przyjęte w poszczególnych projektach procedury autoarchiwizacji (kto, jak, kiedy) oraz zagadnienia techniczno-informatyczne dotyczące transferu danych, kosztów sprzętu i ewentualnych prac programistycznych, metod zabezpieczenia danych oraz skomunikowania repozytoriów z innymi serwisami podobnego rodzaju. W konkluzji raportu konsorcjum zapowiadało wolę powtarzania badania w cyklach półtorarocznych lub dwuletnich.

Wartość poznawcza i prognostyczna tego badania była krytykowana przez francuskich informatologów. Niezależnie od tego, że sami autorzy raportu Couperin powątpiewali w wiarygodność niektórych odpowiedzi ankietowych, zarzucano mu przede wszystkim niereprezentatywność oraz brak perspektywy porównawczej, która uwierzytelniałaby wysuwane na jego podstawie wnioski uogólniające lub przewidywania przyszłych tendencji rozwojowych [Schöpfel, Prost 2010, dok. elektr.]. Przykładem mylności takich ewentualnych uogólnień jest fakt, że choć z badania Couperin zdawały się wyłaniać obiecujące widoki dla rozwoju archiwizacji prac doktorskich, to według statystyk zbiorczych za lata 2005-2008 repozytorium prac doktorskich TEL miało stabilne wpływy na poziomie 1600 jednostek rocznie, co niezmiennie stanowiło tylko 16% produkcji rocznej (w omawianym okresie we Francji przyznawano 10 000 tytułów doktora rocznie) [Bosc 2008, dok. elektr.]. Podsumowanie raportu Couperin, w którym napisano, że inicjatywy tworzenia otwartych repozytoriów przez francuskie instytucje nauki wykazują „liczne cechy typowe dla nowej, intensywnie rozwijającej się dziedziny o znacznym potencjale”, w innym miejscu określonej przymiotnikiem „wschodząca”, brzmi rzeczywiście nieco na wyrost, zwłaszcza w świetle niskich wskaźników wpływu materiałów do już istniejących baz i wyrażanej we wcześniejszych partiach tego dokumentu hipotezy, że nieudzielenie odpowiedzi na ankietę przez niektóre placówki (dopowiedzmy: 70%) było oznaką braku zaangażowania w projekty repozytoryjne. Warto też zaznaczyć, że w raporcie Couperin nie opublikowano listy w pełni operacyjnych repozytoriów instytucjonalnych działających we Francji w roku 2007. Wiadomo tylko, że według autorów w badanej próbie było ich 25.

Ocenę liczebności francuskich archiwów OA powtórzyła w 2008 r. Anna Wojciechowska, która w serwisie OpenDOAR zidentyfikowała ich wówczas 40, a w ROAR – 44. Zbiór ten posegmentowała na: a) biblioteki cyfrowe – 2, b) platformy publikacji i remediacji czasopism naukowych – 3, c) serwisy publikacji i remediacji wydawnictw zwartych – 1, d) zdecentralizowane repozytoria instytucjonalne (samodzielne bazy niezintegrowane z większymi strukturami) – 16, e) repozytoria scentralizowane: tu wymieniła bazy posadowione na platformie

HAL, a wśród nich (według stanu na rok 2008) wyodrębnione portale 30 instytucji oraz kolekcje 86 uniwersytetów, 63 uczelni typu *grande école* i 49 placówek naukowych innego rodzaju, f) archiwa dziedzinowe – 10 oraz e) 3 tzw. portale generyczne HAL (fr. *portails génériques*), czyli HyperArticles en Ligne (HAL), Cours en ligne (CEL) i Thèses en Ligne (TEL) [Wojciechowska 2008, dok. elektr.].

W roku 2008, wyznaczonym jako kres i data bilansu pierwszego dwulecia „Memorandum”, niezależną próbę podsumowania jego owoców podjęła Hélène Bosc. W swoim artykule porównywała efektywność działania wspólnej platformy repozytoryjnej HAL z osiągnięciami twórców autonomicznego repozytorium o nazwie Archimer, otwartego w 2005 r. w IFREMER (Institut français de recherche pour l'exploitation de la mer), państwowym instytucie badawczym specjalizującym się w biologii ekosystemów morskich, oceanografii i hodowli ryb. Za główne parametry przesądzające o racji bytu repozytorium instytucjonalnego jako składnika systemu komunikacji naukowej uznała, po pierwsze, tzw. *taux de remplissage*, czyli iloraz dwóch wielkości: całkowitej liczby publikacji ogłoszonych przez pracowników instytucji naukowej w danym roku kalendarzowym (dzielnik) oraz liczby pre- i post-printów tych publikacji, zamieszczonych legalnie przez autorów w macierzystym repozytorium (dzielna), zaś po drugie, udział – w całym zasobie – post-printów z naukowych czasopism recenzowanych. Dowodziła, że właśnie te dwa mierniki, nie zaś tempo przyrostu ogólnego wolumenu zbiorów, powinny być uznawane za naczelną kryteria szacowania korzyści płynących z utrzymywania otwartych repozytoriów przez publiczne instytucje naukowe. Według ustaleń poczynionych przez Bosc na podstawie ogólnie dostępnych statystyk w 2008 roku kondycja platformy HAL przedstawiała się następująco: 1) zgodnie z zapisami „Memorandum” w portalu HAL powinny były funkcjonować odrębne archiwa 197 instytucji naukowych, podczas gdy w rzeczywistości było ich zaledwie 25, co oznaczało, że większość instytucji reprezentowanych przez sygnatariuszy dokumentu nie wypełniło jego postanowień, 2) 31,8% zasobu stanowiły post-rinty artykułów ogłoszonych w recenzowanych czasopismach naukowych, w liczbie 20 000, 3) wbrew diagnozom Couperin odsetek prac magisterskich i doktorskich wynosił nie więcej niż 14,6%, 4) analiza e-printów pod względem przynależności do poszczególnych dziedzin wiedzy wykazała przewagę nauk ścisłych (fizyka, informatyka, matematyka, astronomia) – 64,4%, kosztem nauk humanistycznych i społecznych – 18,61%, z marginalną pozycją nauk o życiu – niecałe 5%, 5) miesięczny wpływ e-printów od października 2007 do maja 2008 r. wynosił około 1500, czyli mniej niż w swoich najbardziej pesymistycznych prognozach zakładał główny pomysłodawca HAL, Franck Laloë [Laloë et al. 2007, dok. elektr.], 6) ogólna liczba nowych oryginalnych prac naukowych zarchiwizowanych w HAL w latach 2006-2008 osiągnęła 11 000 dokumentów na

120 000 publikowanych rocznie we Francji, czyli 10%, nie zaś jedną czwartą, jak informowano na stronie startowej portalu.

W tym samym czasie sytuacja repozytorium Archimer była kontrastowo odmienna, jako że w roku 2007 zawierało ono 77% z 358 artykułów opublikowanych przez 300 badaczy Instytutu IFREMER w roku 2007 i zaindeksowanych w bazie Web of Science, a ogółem 3400 pełnych tekstów, z czego 94% w wolnym dostępie. Uśredniony wskaźnik *taux de remplissage* za cały okres istnienia repozytorium (2005-2008) wynosił 80%. Hélène Bosc określa Archimer jako wzorcowy przykład samodzielnego repozytorium instytucjonalnego, które bardzo wysoki wskaźnik „ściągłości” e-printów zawdzięcza decyzji władz IFREMER o powierzeniu obsługi procesu archiwizacji specjalnie zatrudnionemu pracownikowi własnego ośrodka informacji i dokumentacji. Jednocześnie z rezerwą odnosi się do scentralizowanego modelu HAL, w którym „zapomniano”, jak mówi, o możliwościach protokołu OAI-PMH, ufając że podmioty zrzeszone w koalicji wokół „Memorandum” będą współzarządzać (fr. *co-piloter*) portalem sprawnie i bez zwady. Zastanawia się, dlaczego nie uznano za priorytet wyposażenia HAL w funkcjonalność agregatora (której, *nota bene*, nie ma do dziś), co umożliwiłoby ewolucję w kierunku archiwów typu *distributed* lub *harvested* (według terminologii Swan et al.) i mimo że jednym z kluczowych haseł „Memorandum” była „interoperacyjność systemów” [Swan et al. 2005, dok. elektr.]. W poszukiwaniu przyczyn powolnego rozwoju platformy wspomina o tarciach i konfliktach interesów między partnerami, sygnalizowanych oddolnie przez francuskie środowisko bibliotekarskie. Zauważa, że od 2006 r. komitet sterujący COSTRAO odbył tylko 6 posiedzeń, a zawartą w protokole COST informację o zakupie 50 000 artykułów z dziedziny fizyki za kwotę 60 000 euro od komercyjnego wydawcy EDP Sciences komentuje słowami: „Badacze mogą się cieszyć, (...) to jest duży postęp, niemniej można również stwierdzić, że za sprawą takich wyborów HAL nie odpowiada dziś wizji wszystkich francuskich partnerów” tego przedsięwzięcia. Dla podkreślenia kruchości sojuszu przypieczętowanego „Memorandum” i niespójności postaw poszczególnych jego członków przypominała, że instytut naukowo-badawczy INRA (Institut national de la recherche agronomique), który w 2004 r. wraz z INSERM jako pierwszy zgłosił akces do przyszłej platformy HAL w odpowiedzi na apel Francka Laloë, podpisał wprawdzie „Memorandum”, lecz jednocześnie zapowiedział, że stworzy własne repozytorium niezintegrowane z HAL, godząc się na transferowanie jego zawartości *ex post* do HAL. Powstało w ten sposób repozytorium o nazwie Prod-INRA, które w 2008 r. było jednak nadal tylko bibliograficzną bazą danych, mimo że badacze INRA publikowali w tamtym okresie około 6 000 artykułów rocznie.

Głównym postulatem Bosc, wyrażonym w konkluzji artykułu, jest wprowadzenie odgórnego nakazu archiwizacji dorobku przez pracowników placówek

naukowych. Tezę o celowości tego rozwiązania uzasadniała zresztą nie raz i nie była w tym osamotniona. Już w 2006 r. w komunikacie prasowym podsumowującym IV. ogólnokrajową konferencję dokumentalistów w Nancy („Journées des Professionnels de l’IST”) stwierdzano, że cały wysiłek mobilizowania naukowców do autoarchiwizacji w repozytoriach spoczywa wyłącznie na barkach bibliotekarzy i pracowników informacji, którym – gdy działają bez oparcia w oficjalnych uregulowaniach prawnych nakładających przymus – udaje się sprowokować wpływy co najwyżej rzędu 12%. Oceniano, że wprowadzenie polityki „kija i marchewki”, przewidującej dodatkowe punkty w ocenie pracowniczej za zdeponowanie e-printu w macierzystym lub centralnym repozytorium OA, pozwoliłoby zwiększyć ten odsetek do maksimum 20%. Jako złoty środek sugerowano wobec tego zadekretowanie obowiązku autoarchiwizacji, wzorem dwóch pionierskich instytucji: CIMAGREF i INERIS, którym posunięcie to umożliwiło zreplikowanie 100% dorobku naukowego pracowników we własnych repozytoriach [Joint Draft Agreement on Open Access in France, dok. elektr.].

W tym samym roku 2006, po części w odpowiedzi na zacytowany apel środowiska bibliotekarzy i dokumentalistów o pilne podbudowanie zapisów „Memorandum” stosownymi regulacjami w prawie francuskim, H  l  ne Bosc wsp  łtworzy  la, na zlecenie w  adz INRA, pierwsze we Francji studium zale  zno  ci mi  dzy wprowadzeniem obowi  zku autoarchiwizacji a poziomem wype  lnienia repozytori  w instytucjonalnych [L’Hostis, Aventurier, Bosc 2006, dok. elektr.]. W oparciu o przegl  d najnowszych bada  n zagranicznych i analiz   skutk  w   wczesnej polityki zach  t dla autor  w, nazwanej „rozwi  zaniem    la CNRS”, L’Hostis, Aventurier i Bosc przedstawili autorsk   propozycj   uregulowa  n (tabela 1), kt  ra spotka  la si   jednak z niemal zupe  n   oboj  tno  ci   ze strony praktyk  w.

Tabela 1. Zasady obligatoryjnej autoarchiwizacji e-print  w w repozytoriach instytucji naukowych wedlug L’Hostis, Aventuriera i Bosc

Zgłaszanie	Obowi��zkowa autoarchiwizacja metadanych wszystkich publikacji (pierwszy autor wprowadza do zasobu wype��lniony arkusz opisu bibliograficznego za pomoc�� specjalnego interfejsu)
Archiwizacja pe��nego tekstu	Obowi��zkowa, tak��e w odniesieniu do pre-print��w
Wolny dost��p / restrykcje dost��pu	O ile to mo��liwe, udost��pnienie bez ogranicze��n dla wszystkich u��zytkownik��w od razu. Dokumenty obj��te embargiem pozostaj�� ukryte tylko do czasu ustania przeszk��d formalnych.
Zwi��zek z ewaluacj��	W formularzu opisu e-printu nale��y stworzy�� pole, w kt��rym autor zgłaszaj��cy oznacza dokument zgodnie z klasyfikacj�� przyj��t�� do cel��w oceny parametrycznej przez instytucj�� kontroluj��c��, kt��rej podlega jego macierzysta plac��wka. Komisje i inne gremia dokonuj��ce ewaluacji uwzgl��dniaj�� wy��cznie dorobek zarchiwizowany w repozytorium (maj�� prawo dost��pu do wszystkich pe��nych tekst��w).

Źr  dło: [L’Hostis, Aventurier, Bosc 2006, dok. elektr.]

W roku 2010 MESR wspólnie z konsorcjum Couperin, CNRS i INIST wydało raport „Open Access in France. A State of the Art Report – April 2010” [Open Access in France... 2010, dok. elektr.]. Omówiono w nim syntetycznie stan zaawansowania programów digitalizacji zasobów historycznych oraz rozwoju obu dróg wdrażania otwartego dostępu do piśmiennictwa naukowego: zielonej i złotej. W części wstępnej przypomniano najważniejsze dane charakteryzujące rynek publikacji naukowych we Francji. W ślad za statystykami OST (Observatoire des Sciences et Techniques, Obserwatorium Nauki i Techniki) z 2008 r. [Filliatreau dir. 2008, dok. elektr.], które uznano za wciąż aktualne, podano że całkowita liczba aktywnych badaczy zatrudnionych przez różne kategorie placówek naukowych sięgała 98 000. Na podstawie ekstrakcji opisów bibliograficznych z bazy Scopus przedstawiono również krzywą francuskiej produkcji wydawniczej w latach 2003-2009, ukazującą stały przyrost rzędu 2-4 tysięcy dokumentów rocznie, ze szczytem aktywności w roku 2008 i nagłym spadkiem w roku następnym:

Tabela 2. Wielkość produkcji publikacyjnej instytucji naukowych afiliowanych we Francji

2003	2004	2005	2006	2007	2008	2009	ogółem
54 194	56 115	60 824	64 168	67 911	71 085	68 843	427 821

Źródło: [Open Access in France... 2010, dok. elektr.]

Zastrzegano jednocześnie, że dane te są prawdopodobnie zaniżone, gdyż pochodzą wyłącznie z bazy Scopus, która po pierwsze, wybiórczo rejestruje francuskojęzyczne czasopisma z dziedziny nauk społecznych i humanistycznych oraz po drugie, uwzględnia tylko artykuły ogłoszone w recenzowanych czasopismach naukowych i materiałach konferencyjnych. W komentarzu do tych ustaleń Hélène Bosc twierdziła, że szacunki wielkościowe oparte na kwerendach w bazie Scopus, podobnie jak w każdej innej bazie o ograniczonym zasięgu formalno-wydawniczym, nie mogą stanowić punktu odniesienia dla miarodajnych wyliczeń poziomu autoarchiwizacji w repozytoriach ze względu na dużo większą różnorodność deponowanych tam materiałów [Bosc 2012, dok. elektr.]. W roku 2008 sama oceniła realną produkcję piśmienniczą francuskich uczonych na 120 000 dokumentów rocznie [Bosc 2008, dok. elektr.], natomiast w 2012 r. zredukowała tę liczbę do 100 000, wyrażając pogląd, że aby zyskać autorytatywne źródło informacji w tym zakresie, wystarczyłoby wprowadzić obowiązek autoarchiwizacji. Byłoby wówczas wiadomo, że wszystko lub niemal wszystko, co opublikowali francuscy badacze, znajduje się w krajowym systemie repozytoriów.

Raport ministerialny z 2010 r. zawiera pierwszą oficjalnie zatwierdzoną przez MESR listę repozytoriów instytucjonalnych i archiwów dziedzinowych, z wy-

szczególnieniem zakresu przedmiotowego, oprogramowania, języka dokumentów, liczby pełnych tekstów oraz liczby rekordów zawierających tylko opisy bibliograficzne. Opracowany na podstawie wskazań OpenDOAR i ROAR spis liczył 68 pozycji. Rozkład zidentyfikowanych serwisów według typu gromadzonych dokumentów i typu instytucji odpowiedzialnej obrazują, odpowiednio, tabele 3 i 4.

Tabela 3. Typy gromadzonych dokumentów, N = 68

Typy gromadzonych dokumentów		Liczba repozytoriów	
Piśmiennictwo naukowe (artykuły z czasopism, materiały konferencyjne itd.)	Repozytoria centralne i archiwa dziedziczne	6	56
	Repozytoria instytucjonalne	50	
Prace magisterskie i doktorskie		9	
Wydawnictwa zwarte		1	
Raporty z badań		2	
Ogółem		68	

Źródło: [Open Access in France... 2010, dok. elektr.]

Tabela 4. Repozytoria instytucjonalne według rodzaju instytucji odpowiedzialnej, N = 50

Typ instytucji	Liczba instytucji	Liczba repozytoriów	Procent instytucji utrzymujących repozytorium
Uniwersytety	86	17	20 %
<i>Grandes écoles</i>	około 230	11	5%
Instytuty naukowo-badawcze	około 36	22	61 %

Źródło: [Open Access in France... 2010, dok. elektr.]

Do pozytywnych wniosków płynących z powyższej statystyki autorzy raportu ministerialnego zaliczyli dość wysoki poziom zaangażowania ośrodków naukowo-badawczych (61%), z których większość przeprowadziła migrację swoich zbiorów do portalu HAL lub postanowiła stworzyć osobne repozytoria, całkiem izolowane, jak w przypadku Archimer, lub skomunikowanie z HAL przez system służący do okresowego pobierania metadanych bibliograficznych, scalanych następnie z rekordami e-printów we własnym zasobie. Ocena działań podjętych przez uniwersytety i *grandes écoles* była surowa. Powiedziano, że „liczby mówią same za siebie“, świadcząc o „znacznym rozdźwięku“ (ang. *important gap*) między zamiarami zgłaszanymi przez uczelnie w odpowiedzi na ankietę Couperin z 2007 r. a tym, co faktycznie udało się im osiągnąć we wdrażaniu zielonej drogi.

Jako jeden z czynników mogących warunkować taki stan rzeczy wskazano brak jednoznacznej polityki wspierania inicjatyw OA na poziomie władz poszczególnych instytucji. Zaproponowano następujące kierunki działań naprawczych, mających pobudzić francuskie instytucje nauki do większej operatywności: 1) ostateczne ustalenie zasad zarządzania platformą HAL, która, według słów raportu, zyskała już środowiskową aprobatę jako centralne repozytorium państwa, lecz zarazem utraciła instytucjonalne oparcie przewidziane w „Memorandum”; gdy dokument ten wygasł bez przedłużenia w roku 2008, rozwiązano bowiem oba ciała odpowiedzialne za bieżące sterowanie HAL (COSTRAO i COST), zawieszając portal w administracyjnej próżni; autorzy raportu ministerialnego sugerowali, że impulsem do wypracowania nowego trybu koordynacji HAL powinny stać się konkluzje niedawnych obrad 7. Konferencji Berlińskiej zorganizowanej w Paryżu w grudniu 2009 r. (wkrótce okazało się, że nadzieje te były płonne [Zespół ICM 2011, dok. elektr.], 2) wprowadzenie obowiązku autoarchiwizacji przez instytucje zatrudniające uczonych lub przez rządowe agencje finansujące badania naukowe; stojący za raportem eksperci MESR odnotowali, że we Francji ciągle nikt jeszcze nie korzysta z tego narzędzia, mimo że, jak dobitnie wskazują międzynarodowe badania i przykłady uczelni europejskich (m.in. Uniwersytetu w Liège), nakaz *immediate deposit* ma zasadnicze znaczenie dla efektywności pozyskiwania e-printów; z niejakim zawodem przypomnieli, że w 2007 r. agencja ANR (Agence nationale de recherche), będąca głównym gestorem publicznych środków na badania naukowe, wydała oświadczenie, w którym stanowczo życzyła sobie, choć nie nakazywała, aby naukowcy realizujący opłacane przez nią projekty badawcze systematycznie składali w repozytoriach lokalnych lub w portalu HAL elektroniczne kopie publikacji, o ile sprawozdają one wyniki tych projektów; dla ścisłości dodają, że do roku 2010 zalecenia ANR znalazły odzew tylko u władz Wydziału Nauk Humanistycznych i Społecznych CNRS, którego politykę w tym zakresie omówiła w cytowanym wcześniej artykule H el ene Bosc [Bosc 2008, dok. elektr.], 3) wykorzystanie szans rozwoju zielonej drogi zwi azanych z udziałem w inicjatywach międzynarodowych, takich jak NEEO, sieć uczelni ekonomicznych pilotująca projekt Economists Online, w której w 2010 r. zrzeszone były 3 uniwersytety francuskie, i OpenAIRE, zaplanowana przez Komisj  Europejsk a międzynarodowa infrastruktura mająca zapewnić archiwizację co najmniej 20% opublikowanych wyników prac badawczych sfinansowanych ze środków 7. Ramowego Programu Badań i Rozwoju [Prost 2010, dok. elektr.; Vathonne 2010, p. 8-9]. Dodam, że według informacji z października 2014 r. w OpenAIRE uczestniczy czternaście francuskich repozytoriów instytucjonalnych, w tym Archimer i OATAO, oraz narodowe archiwum wielod dziedzinowe HAL [OpenAIRE, dok. elektr.].

Rok 2010 przyniósł pierwsze dogłębne badanie stanu rozwoju zielonej drogi, przeprowadzone przez Joachima Schöpfela i Héléne Prost w ramach projektu realizowanego przez laboratorium GERIICO (Groupe d'études et de recherche interdisciplinaire en information et communication) [Schöpfel, Prost 2009, dok. elektr.]. Miało ono scalić i uzupełnić wyniki dwóch wcześniejszych analiz, których Schöpfel był współautorem, a mianowicie: badania udziału szarej literatury we francuskich archiwach OA [Schöpfel, Stock 2009a] i oceny możliwości wykorzystania francuskich zasobów OA do celów edukacji na odległość [Schöpfel, Stock 2009b]. W uzasadnieniu celowości swojego przedsięwzięcia badawczego Schöpfel i Prost podkreślili, że wcześniejsze rozpoznania empiryczne były po pierwsze, nieliczne, a po drugie, fragmentaryczne i wzajemnie niespójne, co wynikało z płytkości kwerend podejmowanych w celu zidentyfikowania istniejących archiwów OA i stosowania niejednorodnych kryteriów ich klasyfikacji, a w rezultacie operowania na niereprezentatywnych próbach. Zastanawiającą rozbieżność danych, która ujawniała się zresztą w piśmiennictwie francuskim już od 2005 r., autorzy demonstrują m.in. na przykładzie dwóch badań ogłoszonych w kilkumiesięcznym odstępie: analizy Bertignac i Gac, ustalającej liczbę aktywnych francuskich archiwów OA na 80, z czego 54 instytucjonalne, z ogólnym zasobem 500 000 pełnych tekstów [Bertignac, Gac 2009], oraz teoretycznie przynajmniej autorytatywnego raportu ministerialnego, w którym liczba ta spada do 68 [Open Archives... 2010, dok. elektr.].

Dlatego też w swoim badaniu Schöpfel i Prost położyli szczególny nacisk na skorygowanie słabości uprzednio stosowanych metod wyszukiwania i kwalifikowania archiwów do próby, tak aby ostateczny wykaz obiektów wiernie zdawał sprawę z zastanego stanu rzeczy. Autorska procedura, stworzona na potrzeby dwóch pierwszych badań pilotażowych [Schöpfel, Stock 2009a i 2009b] i stosownie udoskonalona, opierała się na kilku krokach: 1) maj 2009 r. – zestawienie listy internetowych źródeł metainformacji o kolekcjach obiektów cyfrowych we Francji i na świecie: katalog ten, liczący 19 pozycji, obejmuje wyłącznie źródła instytucjonalne, z których większość nie była wykorzystywana w poprzednich badaniach tego rodzaju, np. ABHATOO, portal informacyjny Krajowego Ośrodka Dokumentacji Królestwa Maroka, 2) maj / czerwiec 2009 r. – wysortowanie z tych źródeł francuskich archiwów OA, odpowiadających trzem kryteriom: francuska afiliacja, czynny adres URL i udokumentowane niedawne wpływy e-printów, 3) lipiec / październik 2009 r. – charakteryzowanie wybranych archiwów według siatki 58 kryteriów podzielonych na kategorie tematyczne. Każda jednostka z wyłonionego w ten sposób zbioru 151 otwartych archiwów została opisana pod względem cech ujętych w pięć kategorii tematycznych: informacje ogólne (nazwa, instytucja odpowiedzialna itd.), informacje szczegółowe (typ ar-



chiwum, zawartość, oprogramowanie, wielkość zasobu itd.), typ gromadzonych dokumentów (szara literatura, raporty z badań itd., ogółem 26 opcji), dane jakościowe (kto archiwizuje materiał, zasady kontroli jakości, schemat metadanych itd.), komentarze (pole uwag przeznaczone dla autorów badania). Prócz tego dla każdego z archiwów poszukiwano statystyk wykorzystania, za optymalny wskaźnik przyjmując liczbę pobranych pełnych tekstów (ang. *downloads*), tak jak zalecano w jeszcze wówczas nieprzełożonej na język francuski niemieckiej specyfikacji kryteriów audytu i oceny jakości archiwów OA DINI [Schöpfel, Müller 2014, dok. elektr.].

Obróbka danych, prowadzona w październiku i grudniu 2009 r., miała charakter trójwarstwowy i przewidywała analizę statystyczną, porównanie danych z wynikami z badań pilotażowych z roku 2008 oraz ocenę jakościową. Przed zaprezentowaniem jej wyników należy zaznaczyć, że podczas kompilowania spisu otwartych archiwów autorzy badania z rozmysłem postanowili nagiąć się do kryteriów kwalifikacyjnych stosowanych w poszczególnych źródłach metainformacji, a że kryteria te były niejednolite, od dość restrykcyjnych po bardzo liberalne, to ostateczny wykaz okazał się równie heterogeniczny. Schöpfel i Prost piszą, że obok „prawdziwych” archiwów OA znalazło się tam również 15 bibliotek cyfrowych, 12 serwisów internetowych z ofertą swobodnego dostępu do materiałów „różnego autoramentu” oraz obiekty „przypominające raczej bazy danych”. Dla wyjaśnienia tej decyzji dokonują najpierw przeglądu funkcjonujących w piśmiennictwie francuskim i anglosaskim definicji *open archives / archives ouvertes* [m.in. Carr, Brody, Swan 2008, dok. elektr.; Smith 2008, dok. elektr.; Armbruster, Romary 2009, dok. elektr.], wydobywają z nich zestaw cech konstytutywnych, a następnie podważają ich wartość typizującą, dochodząc do wniosku, że nie da się stworzyć normatywnej definicji otwartego archiwum (repozytorium), która byłaby przydatna w badawczym porządkowaniu spotykanych w sieci form podaży informacji i danych w ramach zielonej drogi OA. Wobec tego, jeżeli chce się uzyskać kompletny obraz francuskich osiągnięć na tym polu, najlepiej jest przyjąć podejście indukcyjne, czyli stworzyć repertorium serwisów, które według różnych podmiotów klasyfikujących zasługują na miano otwartych archiwów, a dopiero potem podjąć próbę ich kategoryzowania. W myśl tej zasady autorzy dzielą uzyskany zbiór 151 archiwów dwojako: już to według nomenklatury przejętej od Armbrustera i Romary’ego – na repozytoria instytucjonalne (ang. *institutional repository*), archiwa dziedzinowe (ang. *subject-based repository*), repozytoria narodowe (ang. *national repository*), archiwa naukowe (ang. *research repository*) [Armbruster, Romary 2009, dok. elektr.], już to ze względu na rodzaj archiwizowanych dokumentów, wyróżniając w ten sposób:

Tabela 5. Typologia baz OA według Schöpfela i Prost

1. Archiwa przyjmujące materiał wytwarzany na bieżąco	1.1. Pre-printy i post-printy	
	1.2. Surowe dane	1.2.1. Wyniki badań 1.2.2. Obiekty badań
2. Bazy dziedzictwa kultury	2.1. Komunikacja naukowa	2.1.1. Pre-printy i post-printy
	2.2. Utrwalanie dorobku kultury	2.2.1. Obiekty badań

Źródło: [Schöpfel, Prost 2010, dok. elektr.]

Omawiana publikacja Schöpfela i Prost dostępna jest tylko w języku francuskim, dlatego warto dość szczegółowo przytoczyć zawarte w niej spostrzeżenia. W aspekcie ilościowym autorzy stwierdzili, w porównaniu do roku 2008, spektakularny wzrost liczby samych archiwów OA i wielkości ich zbiorów. Lista zestawiona w roku 2008 wykazywała 56 archiwów z łącznym zbiorem około 703 tys. rekordów kontra 150 archiwów (+170%) z prawie 1,9 miliona rekordów (+164%) w roku 2009. Datując początek zielonej drogi we Francji na połowę lat 90. XX wieku, wyodrębnili w jej późniejszym rozwoju cztery etapy, z punktem przełomowym w postaci podpisania Deklaracji Berlińskiej: 1) 1995-2000, okres prekursorski, utworzenie pierwszych archiwów dziedzinowych i repozytoriów, np. CITHER, CNUM-CNAM i LACITO Archive, 2) 2001-2004, pierwsze decyzje polityczne wyrażające poparcie dla idei OA, mobilizacja środowisk naukowych wokół baz OA zarządzanych centralnie, takich jak HAL, TEL, NUMDAM, Cyberthèses, EduTICE i @rchiveSIC, 3) 2005-2007, dynamiczny rozwój, powołanie do życia licznych samodzielnych repozytoriów instytucjonalnych, m.in. Archimer, LARA (raporty z badań naukowych) lub transfer kolekcji do portalu HAL (INSERM, IRD, Institut Pasteur), 4) od 2008, konsolidacja, nieznaczne spowolnienie tempa liczebnego przyrostu nowych archiwów, wiele dużych projektów w toku, m.in. SPIRE (Sciences Po Paryż), OATAO (Tuluza), DUMAS (repozytorium prac magisterskich).

Jeśli chodzi o liczbę nowo stworzonych repozytoriów w podziale według instytucji odpowiedzialnej, autorzy zaobserwowali równomierny przyrost tak po stronie instytutów naukowo-badawczych, jak i placówek kształcenia wyższego (tabela 6), przy czym po zbadaniu statystyk wpływów materiału okazało się, że uczelnie nadrabiają wprawdzie zapóźnienia i budują infrastruktury bazodanowe, lecz mają trudności, by zapewnić je treścią.

Tabela 6. Przyrost liczby repozytoriów instytucjonalnych we Francji w latach 2008-2009

Repozytoria według instytucji	Liczba w roku 2008, N = 55		Liczba w roku 2009, N = 150 <sup>2</sup>	
Uczelnie wyższe	27	50%	61	41%
Instytuty naukowo-badawcze	26	47%	66	44%
Inne instytucje	2	7%	23	15%

Źródło: [Schöpfel, Prost 2010, dok. elektr.]

Posegmentowanie archiwów według klucza zaproponowanego przez Armbrustera i Romary'ego (2009) dało następujący wynik:

Tabela 7. Przyrost liczby baz OA we Francji w latach 2008-2009

	Liczba w roku 2008, N = 55	Liczba w roku 2009, N = 150
Repozytoria instytucjonalne	31	81
Archiwa dziedzinowe	13	19
Archiwa centralne	7	11
Archiwa naukowe	4	23
Inne	1	16

Źródło: [Schöpfel, Prost 2010, dok. elektr.]

Autorzy sprawdzili także stosowane w poszczególnych instytucjach systemy zachęty do autoarchiwizacji, znajdując obok kilku takich, które zdawały się nie korzystać z żadnych instrumentów tego rodzaju, a przynajmniej nigdzie o tym nie informowały, aż 18 placówek optujących za rozwiązaniem obowiązkowym (tabela 8 poniżej). Warto zauważyć, że akurat w tym względzie tryb pozyskiwania danych przyjęty przez Schöpfela i Prost, skądinąd dyskusyjny, przyniósł pożytek, ponieważ w innych badaniach, czasowo niemal równoległych, liczba ta była niewyrażniejsza niż zaniżona. W swojej publikacji z 2009 r. Héléne Bosc, niestrudzona orędowniczka modelu Immediate Deposit/Optional Access (ID/OA), zidentyfikowała we Francji, wyłącznie na podstawie kwerendy w ROAR z 19.10.2008 r., tylko dwa podmioty stosujące nakaz autoarchiwizacji (jeden wydział uniwersytecki, jedną fundację) oraz jeden ośrodek naukowo-badawczy, który zadeklarował wolę

<sup>2</sup> Na stronie 13 cytowanej publikacji Schöpfel i Prost podają, że przeanalizowali ogółem 151 francuskich archiwów OA (według ich autorskiej definicji). Tyle też pozycji liczy nominalny spis badanych baz, zamieszczony w załączniku „Annexe C. Sites analysés” (s. 43-46). Na stronie 27 pada jednak liczba 150. Autorzy nie wyjaśniają przyczyn tej rozbieżności. Wartości liczbowe widniejące w wierszach tabel 6, 7 i 8 dla roku 2009 (które przejęłam z oryginalnego tekstu *verbatim*) sumują się do 150. Podobna nieregularność dotyczy także danych dla roku 2008, których suma – w tabelach 6, 7 i 8 – wynosi już to 55, już to 56, mimo że na stronie 27 swojego opracowania Schöpfel i Prost stwierdzają, że liczba archiwów OA działających we Francji w roku 2008 wynosiła 56 [Schöpfel, Prost 2010, dok. elektr.].

wprowadzenia takiego nakazu w najbliższej przyszłości [Bosc 2009, dok. elektr.]. Co ciekawe, także cztery lata później, w roku 2012, we współautorskim artykule pod znamionym tytułem „Le libre accès en France en 2012: Entre immobilisme et innovation” („Otwarty dostęp we Francji w 2012 r.: między beczynnością a innowacyjnością”) podtrzymywała w domyśle tezę o wręcz nieobecności narzędzi typu *deposit mandate* w krajowym środowisku instytucji nauki, najpierw pisząc o udanym stosowaniu ich przez belgijski Université de Liège, by następnie dorzucić, że „przykład bliskiego francuskojęzycznego sąsiada nie doczekał się jeszcze naśladowców we Francji”. Donosiła też, że zapowiedzi wprowadzenia obowiązku autoarchiwizacji ogłoszone w ROARMAP przez instytut INRA nie zostały zrealizowane [Bosc, Dillaerts 2012, dok. elektr.].

Tabela 8. Polityka instytucji w dziedzinie autoarchiwizacji w latach 2008-2009

	Liczba w roku 2008, N = 56	Liczba w roku 2009, N = 150
Zachęta bez obligacji	26	81
Obowiązek	12	16
Digitalizacja zasobów historycznych	7	18
Typ mieszany lub nieokreślony	11	35

Źródło: [Schöpfel, Prost 2010, dok. elektr.]

Schöpfel i Prost chcieli również precyzyjnie ustalić *taux de remplissage* (według nazewnictwa Héléne Bosc), czyli odsetek bieżącej produkcji publikacyjnej trafiający rokrocznie do wszystkich repozytoriów instytucjonalnych razem wziętych, w zawężeniu do artykułów z recenzowanych czasopism naukowych. Zamiar ten zniweczyła niejednorodność schematów metadanych, stosowanych w poszczególnych systemach informacyjno-wyszukiwawczych, wskutek której w wielu bazach wysortowanie artykułów za pomocą filtra „typ dokumentu” było niemożliwe, a zliczanie ręczne nie wchodziło w grę z racji wielkości zasobu. Z relewantnych danych, które udało im się uzyskać zaledwie dla 72 serwisów (52% badanej stawki), często kosztem żmudnych rekonstrukcji na podstawie rozmaitych źródeł rozproszonych w sieci, ekstrapolowali wskaźnik *taux de remplissage* rzędu 35%, z zastrzeżeniem, że wartość ta może być zaniżona. W celach poglądowych i bez roszczeń do reprezentatywności zdecydowali się powtórzyć analizę tylko w odniesieniu do zawartości platformy repozytoryjnej HAL, którą porównali z zasobami baz Scopus i Web of Science. Ustalili w ten sposób, że w latach 2006-2008 we wszystkich kolekcjach HAL zaindeksowano 15% artykułów z recenzowanych czasopism naukowych, jakie w analogicznym okresie wykazano w bazie Scopus. Natomiast po ograniczeniu analizy do artykułów z recenzo-

wanych czasopism naukowych o zasięgu międzynarodowym odsetek ten spadł do 7-8% zarówno w porównaniu do bazy Scopus, jak i Web of Science. Autorzy uznali te wyniki za realistyczne.

Przypadkowi platformy HAL poświęcili zresztą w swoim opracowaniu osobny podrozdział, wskazując kilka czynników utrudniających rzetelną ocenę wolumetryczną tego zasobu: 1) niemożność ustalenia dokładnej liczby rekordów w momencie N; aktualna liczba podana na stronie startowej serwisu nie zgadzała się z sumą wyników zwróconych przez system po zastosowaniu różnych filtrów (np. tylko opisy bibliograficzne, tylko rekordy z pełnym tekstem, tylko rekordy bez restrzykcji dostępu, tylko artykuły itd.), 2) niejednorodność struktury formularzy wyszukiwania zaawansowanego w poszczególnych portalach instytucjonalnych HAL (np. w portalu INSERM nie było opcji pozwalającej wysortować tylko rekordy z pełnym tekstem), 3) nieprecyzyjność metadanych opisu bibliograficznego, zwłaszcza w odniesieniu do materiałów nieprzeznaczonych do szerokiego upubliczniania (dokumenty do użytku wewnętrznego, pomoce dydaktyczne itd.), które charakteryzowane były albo bardzo szczegółowo, albo bardzo ogólnie, np. jako „inne”, 4) obecność dubletów, spowodowana po pierwsze tym, że każdy autor może umieścić swój e-print w dowolnej liczbie repozytoriów oraz po drugie, podjętą przez CCSD decyzją o systematycznym transferowaniu do HAL zawartości innych archiwów pozostających w jego (CCSD) gestii; na tej zasadzie w HAL powielają się np. prace magisterskie i doktorskie zdeponowane wcześniej w PASTEL i TEL, a z kolei e-prints z centralnego archiwum dziedzinowego @rchiveSIC (informatologia i bibliologia) trafiają równolegle i do HAL generycznego (HyperArticles en Ligne), i do segmentu dziedzinowego HAL-SHS (nauki humanistyczne i społeczne). Schöpfel i Prost zastanawiają się, jak w tych okolicznościach skorygować sumaryczną liczbę rekordów we wszystkich bazach uznanych (przez siebie) za otwarte archiwa, i dochodzą do wniosku, że jest to niewykonalne. Można by wprawdzie odjąć z 1,87 miliona rekordów całą zawartość generycznego HAL, co zmniejszyłoby tę liczbę o około 7%, do 1,73 miliona, jednak wynik takiej operacji wciąż nie byłby wiarygodny, bo w HAL generycznym składane są również materiały, których autorzy nie archiwizują nigdzie indziej. Według Schöpfela i Prost najrozsądniej byłoby założyć, że rzeczywista liczba mieści się między 1,9 a 1,7 miliona rekordów.

W dyskusji wyników autorzy ocenili trafność dwóch z czterech hipotez przyjętych przed rozpoczęciem badania, a wynikających z dwóch poprzednich analiz pilotażowych [Schöpfel, Stock 2009a i 2009b]. Uściślił od razu, że hipoteza trzecia i czwarta odnosiły się do dwóch ważnych aspektów użytkowania otwartych archiwów we Francji i głosiły, odpowiednio, że w roku 2009 3) poprawi się dostępność automatycznie generowanych statystyk liczby odwiedzin i pobrań

abstraktów / pełnych tekstów dokumentów oraz 4) dane te nadal pozostaną niepełne i nieznormalizowane. Zgodnie z harmonogramem projektu badawczego realizowanego przez Schöpfela i Prost wyniki badań dotyczących wykorzystania zasobów OA i podstawowych parametrów zachowań użytkowników miały zostać opublikowane – i zostały – w postaci osobnego opracowania [Schöpfel, Stock 2009b].

Hipoteza pierwsza wyrażała się w domniemaniu, że w roku 2009 r. liczba otwartych archiwów i wolumen ich zasobów znacząco zwiększy się w stosunku do roku 2008. Tak też się stało. Omawiając zaobserwowaną tendencję wzrostową, autorzy wycofali się z krytycyzmu wobec wyników pierwszego badania Couperin z 2007 r. [Bruley et al. 2007, dok. elektr.; Bruley et al. 2008, dok. elektr.], do których Schöpfel pierwotnie dystansował się ze względu na niereprezentatywność próby. Francuskie otwarte archiwa (w znaczeniu lansowanym przez autorów badania, czyli wraz z bibliotekami cyfrowymi i platformami publikacji i remediacji czasopism naukowych) zaliczały się, ich zdaniem, do największych na świecie. W lipcu 2009 r. archiwum HAL zajmowało trzecie, a HAL-INRIA piąte miejsce w RanKing Web of World Repositories (Webometrics). HAL-INRIA zostało też ogłoszone największym repozytorium instytucjonalnym w skali międzynarodowej. Wbrew wskazaniom ROAR z czerwca 2010 r., które przedstawiały francuskie archiwa w niekorzystnym świetle na tle Niemiec, Zjednoczonego Królestwa i USA, badania Schöpfela i Prost dowodziły, w ich opinii, że Francja miała pozycję porównywalną do tych trzech państw, z dużą liczbą istniejących archiwów, z podobnym średnim wolumenem zbiorów i z takim samym *taux de remplissage* jak Zjednoczone Królestwo i USA, a znacznie większym niż Niemcy. Autorzy uważali, że w raporcie MESR niesłusznie umniejszono wagę tego faktu [Open Access in France... 2010, dok. elektr.].

Potwierdzenie znalazła również druga hipoteza badawcza przewidująca wzrost udziału w zbiorach otwartych archiwów szarej literatury definiowanej jako publikacje nieprzeznaczone do obiegu handlowego (prace magisterskie, doktorskie, referaty wygłoszone na konferencjach, nieopublikowane rękopisy itd.): w 2009 r. materiały tego rodzaju zawierało ogółem 74% francuskich baz OA, w tym 93% repozytoriów instytucjonalnych. Autorzy przyjmowali ten fakt z zadowoleniem, niemniej przypominali, że głównym zamierzeniem pionierów zielonej drogi we Francji, skupionych wokół założyciela HAL Francka Laloë, było szybkie i bezpośrednio rozpowszechnianie pre-printów, wzorem prekursorskiego serwisu ArXiv. Natomiast z jawną dezaprobatą odnieśli się do praktyki zamieszczania w otwartych archiwach zdigitalizowanych dokumentów o wartości historycznej, stwierdzając że użytkownikom trudno jest wówczas rozeznaczyć cel istnienia i zakres przedmiotowy zbioru, a po wtóre stawia to pod znakiem zapytania pertinent-

ność wyników wyszukiwań. Odnośnie do struktury zbiorów w repozytoriach instytucjonalnych niepokoił ich również wysoki odsetek rekordów bibliograficznych bez pełnych tekstów. Tendencję do łączenia autoarchiwizacji ze zwykłą rejestracją bibliograficzną przypisywali temu, że francuskie instytucje naukowe, nadmiernie wpatrzone w statystyki cytowań i indeksy wpływu, od których zależy ministerialna ocena ich pracy i przydział funduszy, wołają stawić na szybkie powiadomianie o publikacjach, nie zaś na szybkie ich udostępnianie. Wszystkie te spostrzeżenia są zrozumiałe.

Polskiemu czytelnikowi, przyzwyczajonemu do demarkacji pojęć „biblioteka cyfrowa” i „archiwum / repozytorium OA”<sup>3</sup>, niełatwo jest natomiast odcyfrować sens słów Schöpfela i Prost, gdy mówią oni, że „kolejnym ograniczeniem koncepcji otwartych archiwów jest mnożenie się kolekcji zawierających obiekty dziedzictwa kulturowego”, a następnie dodają, że według ich rachunków w roku 2009 biblioteki takie, działające we Francji w liczbie 30, stanowiły 20% wszystkich archiwów OA i skupiały 46% wszystkich zarchiwizowanych dokumentów. Nie dość, że w zdaniu tym pada liczba 20, podczas gdy we wcześniejszych partiach tekstu badacze sygnalizowali włączenie do analizowanego spisu 15 bibliotek cyfrowych, to jeszcze powstaje wrażenie, że postrzegają je oni jako swego rodzaju twory pasożytnicze, zagarniające pole zastrzeżone dla autoarchiwizacji i wypaczające założenia otwartego dostępu. Takie stawianie sprawy musi zdumiewać, bowiem autorzy najpierw sami zdekonstruowali i odrzucili funkcjonujące w piśmiennictwie definicje *open archives*, po czym z własnej woli postanowili uznać biblioteki cyfrowe za podkategorię otwartych archiwów, choć przecież mogli odwrócić tę hierarchię (przyjmując, że to biblioteki cyfrowe są pojęciem nadrzędnym) lub, skoro już unieważnili cudze taksonomie, stworzyć własny system porządkujący, z którego biblioteki cyfrowe zostałyby wykluczone. Posunięcie takie mogło przyjść im tym łatwiej, że ENSSIB lansowała wówczas pogląd, że archiwa OA są specyficznym podzbiorem bibliotek cyfrowych (fr. *bibliothèques numériques*) wyszczególnionym ze względu na typ gromadzonych dokumentów [Quels sont... 2010, dok. elektr.].

---

<sup>3</sup> Pisząc o demarkacji, mam na myśli nie ostre rozgraniczanie pojęć „biblioteka cyfrowa” i „archiwum / repozytorium OA”, a tylko świadomość, że oba te rodzaje baz różnią się rodowodem i pierwotnym przeznaczeniem, choć jednocześnie zgadzam się z Remigiuszem Sapa, który w niniejszym tomie stwierdza: „(...) w dobie konwergencji narzędzi i usług informacyjnych dążenie do tworzenia radykalnie rozłącznej typologii przedsięwzięć podejmowanych w obszarze zarządzania cyfrowymi zasobami informacji musi być skazane na niepowodzenie. Tak samo jak nie da się wskazać ostrej granicy między telefonem a komputerem (np. w przypadku smartfonów lub fabletów) czy usługą i produktem informacyjnym (np. indywidualizowany druk publikacji na życzenie), tak też trudno byłoby zdecydowanie rozróżnić repozytoria od np. bibliotek cyfrowych. Można raczej mówić o realizacji funkcji repozytoryjnych przez różne przedsięwzięcia” (s. 82).

W podsumowaniu swojego raportu Schöpfel i Prost stwierdzili, że szybki rozwój usług na rzecz swobodnego dostępu do dorobku nauki we Francji plasuje ją pod tym względem w czołówce krajów europejskich (fr. *au coeur de ce mouvement sur le vieux continent*), przy czym źródłem słabości może być utrzymujący się model „dwóch prędkości”, z prężnym scentralizowanym ośrodkiem HAL i mgławicą wzajemnie niepowiązanych lokalnych inicjatyw na peryferiach, których koordynacją mogłyby z powodzeniem zająć się co najmniej trzy instytucje: CCSD, COUPERIN/GTAO i MESR. W 2009 r. koordynacja ta nie była jednak zapewniowana. Podczas obrad pierwszej ogólnokrajowej konferencji Couperin, poświęconej zagadnieniom Otwartego Dostępu we Francji – „Journées d'étude sur les archives ouvertes” – zorganizowanej w Paryżu w kwietniu 2009 r., podkreślano wprowadzenie dynamicznego rozwoju repozytoriów instytucjonalnych, wymieniano plany HAL co do konkretnych rozwiązań technicznych mających zapewnić interoperacyjność, nawoływano do poddania się rygorom normalizacji i zgadzano się, że sprawnemu funkcjonowaniu platformy Francja zawdzięcza swoje „miejsce w gronie państw, które osiągnęły zaawansowanie w dziedzinie otwartych archiwów”, jednak sposób zarządzania HAL określano jako „polityczną porażkę”. Przedstawiciele Couperin i MESR twierdzili, że przyszłość CCSD w strukturze organizacyjnej CNRS jest niepewna i że jednostka ta nie ma żadnej siły przebicia w negocjacjach z wydawcami, ponieważ nie udało jej się dotychczas ani spowodować masowego napływu e-printów do HAL, ani skutecznie zadbać o wyposażenie platformy w narzędzia generowania statystyk frekwencyjnych, co w następstwie wytrąca jej z ręki najważniejszy argument przetargowy mogący skłonić koncerny wydawnicze do polubownej współpracy. W imieniu środowiska naukowego związanego z informatologią i bibliotekoznawstwem Jean-Claude Guédon, autor tekstu Deklaracji Budapesztańskiej, wyraził opinię, że poszczególne ośrodki bądź, ogólnie, podmioty odpowiedzialne za definiowanie i wdrażanie polityki informacyjnej we Francji niedostatecznie się ze sobą komunikują, jakkolwiek można mieć nadzieję, że jest to zjawisko przejściowe [Schöpfel, Prost 2009c, dok. elektr.].

Schöpfel i Prost uzależniali dalszy rozwój francuskich archiwów (przypomnę: według ich własnej, wszechogarniającej definicji) od dwóch czynników: 1) decyzji politycznych dotyczących wprowadzenia (lub nie) obowiązku autoarchiwizacji, 2) scalenia pojedynczych repozytoriów instytucjonalnych, w szczególności uniwersytetów i *grandes écoles*, w większe jednostki, np. w ramach PRES (Pôle de recherche et d'enseignement supérieur), czyli tworzonych od 2006 r. aliansów uczelni mających sprzyjać bardziej efektywnemu wykorzystaniu potencjału badawczo-dydaktycznego i usuwaniu sztucznych barier w komunikacji naukowej. Zapowiedzią takich ruchów integrujących miało być utworzenie nowego ogólnokrajowego systemu archiwizowania prac doktorskich, planowane przez Agence



bibliographique de l'enseignement supérieur (ABES) na przełom lat 2012/2013. Z perspektywy roku 2014 dodam, że agencja ABES dotrzymała terminów i w kwietniu 2013 r. uruchomiła wspólny system przesyłu rozpraw doktorskich, składanych w pełnym tekście w archiwum STAR, oraz informacji o otwartych przewodach doktorskich, gromadzonych w bazie STEP, z funkcją pobierania metadanych z kilku repozytoriów lokalnych (np. APOGEE, ADUM) [Jordan 2014, dok. elektr.]. Nie sprawdziła się natomiast formuła aliansów uczelni PRES, które na mocy nowej ustawy o szkolnictwie wyższym i badaniach naukowych z 2013 r. zostały (z wyjątkiem trzech) zastąpione strukturami o nazwie „wspólnoty uczelni i instytucji naukowo-badawczych” (fr. *communautés des universités et des établissements*) [Loi (Ustawa) n° 2013-660..., dok. elektr.].

Ostatnie z dużych, głęboko drążących opracowań poświęconych rozwojowi repozytoriów instytucjonalnych we Francji pochodzi z roku 2012 i jest to praca doktorska Marine Rigeade [Rigeade 2012, dok. elektr.]. Autorka krytycznie osądza wyniki badań Schöpfela i Prost, wytykając im przede wszystkim celowe (a jej zdaniem nieuzasadnione) uwzględnienie nie tylko bibliotek cyfrowych, ale też baz bibliograficznych niezawierających wcale pełnych tekstów (Cristallography Open Database), co sztucznie powiększyło rozpatrywany zbiór „otwartych archiwów” i ujemnie wpłynęło na wiarygodność wyliczonych przez autorów statystyk. Do opracowania własnej listy repozytoriów instytucjonalnych we Francji Rigeade wykorzystwała ROAR, OpenDOAR, oficjalne wykazy udostępniane przez MESR, ABES i CCSD oraz wyniki samodzielnie przeprowadzonego sondażu obejmującego wszystkie francuskie uniwersytety, *grandes écoles* i instytuty naukowo-badawcze (poszukiwanie informacji o istnieniu lub planach powołania repozytorium na stronach internetowych instytucji, sondaż e-mailowy i telefoniczny). Zidentyfikowała w tym trybie 36 repozytoriów, z których 23 (64%) należały do szkół wyższych. Większość z nich była zaindeksowana w katalogach OpenDOAR i ROAR, choć autorka zauważyła jednocześnie, że tylko 10 miało interfejs anglojęzyczny, co ewidentnie ogranicza korzyści płynące z rejestracji w międzynarodowych *reference sites*. W wielu serwisach brakowało informacji o liczbie przechowywanych dokumentów, a dane dostępne w OpenDOAR i ROAR autorka uznała, po ręcznej weryfikacji, za często nieaktualne. Po sprawdzeniu stanu zasobów za pomocą narzędzi dostępnych w poszczególnych systemach informacyjno-wyszukiwawczych stwierdziła, że tylko 14 repozytoriów utrzymuje zbiór powyżej 1000 dokumentów, z czego 6 przekracza pułap 5000. Rigeade próbuje też wyjaśnić przyczyny dwóch praktyk niezgodnych z ideą otwartego dostępu, tj. masowej rejestracji bibliograficznej bez pełnych tekstów (w bazie ProdINRA w 2012 r. pełne teksty dołączono do zaledwie 4% ze 125 000 rekordów) oraz udostępniania dużej części materiałów pełnotekstowych tylko użytkownikom auto-

ryzowanym. Utrzymuje mianowicie, że wiele instytucji naukowych godzi się na przyjmowanie samych opisów bibliograficznych i tekstów z restrykcjami dostępu, gdyż w pierwszym rzędzie zależy im na choćby nominalnym powiększeniu zasobu, na osiągnięciu „masy krytycznej”, która zapewni repozytorium odpowiednią rangę i widoczność w środowisku naukowym, w nadziei, że to z kolei ułatwi starania o zgodę wydawców i autorów na sukcesywne przenoszenie publikacji do strefy swobodnego dostępu. W tym kontekście Rigeade wspomina również o postulowanym przez Stevana Harnada, a jego śladem we Francji przez H  l  ne Bosc, wyposażeniu rekord  w bibliograficznych w repozytoriach w funkcj   „Request a copy”, kt  ra, w przypadku tekst  w ob  oonych przez wydawc   embargiem, umo  liwiałaby wysłanie do autora wiadomo  ci elektronicznej z pro  b   o udost  pnienie pre-printu na u  ytek wlasny [Harnad et al. 2009, dok. elektr.; Bosc, Dillaerts 2012, dok. elektr.]. Sygnalizuje,   e w roku 2012 funkcja „Request a copy”, pod nazw   „Fair Deal Button”, była dost  pna w co najmniej jednym repozytorium instytucjonalnym we Francji, tj. OATAO w Tuluzie.

Opr  cz 36 repozytori  w samodzielnych autorka wymienia r  wnie   65 „portali instytucjonalnych” stworzonych przez r  zne podmioty na wsp  lnej platformie repozytoryjnej HAL. Celowo nie zlicza tych kolekcji jako odr  bnych obiekt  w (a fakt,   e Sch  pfel i Prost post  pili odwrotnie, poczytuje za jeszcze jeden mankament ich badania), podaj  c nast  puj  ce powody: 1) „portale instytucjonalne” w HAL tworzone s   w ten spos  b,   e w charakterystykach metadanych publikacji autor  w afiliowanych w danej instytucji umieszcza si  , r  cznie lub automatycznie, specjalne oznaczenie (zwane „piecz  tk  ” – fr. *tampon*), dzi  ki kt  remu pliki te zostaj   przyporz  dkowane do danej kolekcji, 2) wlasiciel „piecz  tki”, tj. instytucja, kt  ra zdecydowała si   skorzysta  c z HAL, rezygnuj  c z zało  zenia odr  bnego repozytorium, tworzy wlasn   podstron   w portalu g  wnym HAL, przy czym swoboda jej personalizacji ogranicza si   do wyboru szaty graficznej, 3) interfejsy wyszukiwawcze wszystkich „portali” musz   by  c dost  pne w dw  ch wersjach j  zykowych (angielskiej, francuskiej) i umo  liwia  c przegł  danie materiału za pomoc   okre  szonej liczby indeks  w, 4) wlasiciel „piecz  tki” podporz  dkowuje si   zasadom polityki gromadzenia i udost  pniania przyj  tej centralnie dla całej platformy HAL, zgodnie z kt  r   „w wirtualnych kolekcjach HAL (...) mog   by  c zamieszczone tylko artykuły lub dokumenty naukowe podobne do artykułu” [Foulonneau et al. 2007, dok. elektr.]. Bior  c pod uwag  ,   e autonomia wlaszcicieli „piecz  tek” sprowadza si   wył  cznie do deponowania e-print  w i wyboru element  w estetycznych podstrony w HAL, Rigeade uwa  za,   e odr  bne rachowanie „portali instytucjonalnych” byłoby nadu  życiem.

Wielk   zalet   pracy Marine Rigeade jest wnikliwe, poparte metodycznym przegł  dem pi  smiennictwa om  wienie mocnych stron i niedostatk  w francu-

skiego modelu środek-peryferie, którego filarem ma być centralne repozytorium państwowe HAL. Autorka porządkuje przeciwstawne argumenty wysuwane, odpowiednio, przez zwolenników repozytoriów centralnych i obrońców polityki premiującej repozytoria instytucjonalne, przywołując m.in. polemikę między Franckiem Laloë i Stevanem Harnadem podczas Open Access Forum w październiku 2004 r. [Laloë 2006, dok. elektr.]. Wydaje się, że szczegółowe relacjonowanie sporu o komplementarność lub wzajemne wykluczanie się obu tych modeli nie jest potrzebne, gdyż nie odnosi się on tylko do Francji. Dość przypomnieć, że Franck Laloë obstawał za archiwizacją scentralizowaną przede wszystkim dlatego, ażeby zapewnić naukowcom możliwość deponowania e-printu zawsze w jednym miejscu (fr. *dépôt unique*), niezależnie od afiliacji autora lub autorów, kontrolę jakości zasobu repozytoryjnego dzięki przyjęciu ostrych kryteriów selekcyjnych oraz bezpieczeństwo długoterminowego przechowywania treści. Stevan Harnad oponował zaś, jak zwykle, mówiąc że głównym problemem zielonej drogi OA nie jest wybór miejsca i trybu autoarchiwizacji, ale skłonienie autorów do tego, by ją praktykowali. Dla polskiego odbiorcy najbardziej interesujące są te fragmenty opracowania Rigeade, w których – na podstawie najświeższych wówczas informacji, pozyskanych również z pierwszej ręki od źródeł osobowych – opisuje ona już zapoczątkowane lub dopiero planowane działania zmierzające do zapewnienia interoperacyjności platformy HAL z samodzielnymi repozytoriami instytucji nauki.

Otóż w roku 2012 platforma HAL nadal nie miała funkcji agregowania metadanych z innych archiwów (ang. *harvesting*), w związku z czym każdy (autor lub reprezentująca go instytucja), kto chciał zdeponować jakiegokolwiek treści w tym zasobie, musiał wprowadzić je bezpośrednio do HAL i tylko tam, chyba że jego macierzyste repozytorium instytucjonalne zdołało już wyposażyć się w system spakietowanego przesyłu danych, skomunikowanego z HAL. Jedynie w takim przypadku zamieszczenie treści w repozytorium instytucjonalnym gwarantowało ich późniejsze automatyczne przekazanie do HAL. Warto dodać, że w świetle obowiązujących przepisów żadna placówka niepodlegająca CNRS nie miała ani obowiązku transferowania swoich zasobów do HAL, ani obowiązku implementowania w tym celu specjalnych systemów przesyłowych. Rigeade podziela opinię Sandrine Malotaux, dyrektor służby biblioteczno-dokumentacyjnej Institut national polytechnique w Tuluzie, że nie ma sensu kruszyć kopii o to, który z modeli archiwizacji, centralny czy lokalny, cechuje się wyższym wskaźnikiem *taux de remplissage* i lepiej nastraja uczonych do udziału w zielonej drodze, skoro w kraju brakuje rozwiązań zapewniających wszystkim repozytoriom instytucjonalnym interoperacyjność, kompatybilność z agregatorami, indeksowanie przez różne wyszukiwarki oraz możliwość dwukierunkowego transferowania danych,

z repozytoriów lokalnych do HAL i odwrotnie. Sandrine Malotaux należy do zwolenników uprzywilejowania, a przynajmniej równego traktowania repozytoriów instytucjonalnych, głównie (choć nie tylko) ze względu na bardzo wąskie spektrum treści akceptowanych w HAL, niemniej najważniejszą i bezdyskusyjną koniecznością jest dla niej wyjście z dotychczasowego impasu w kwestii interoperacyjności. W modelu scentralizowanym to obiekty peryferyjne muszą dostosować się do ośrodka, a nie ośrodek do peryferii, skutkiem czego *gros* wysiłków związanych z budową wzajemnych powiązań komunikacyjnych spoczywa we Francji na repozytoriach instytucjonalnych. Nakład pracy jest znaczny, podobnie jak skala koniecznych do usunięcia trudności technicznych. Po pierwsze, repozytoria lokalne muszą skorelować z HAL swoje schematy metadanych, tak aby osiągnąć pełną zgodność z normami CCSD w zakresie haseł wzorcowych, tytułów czasopism, nazw geograficznych, języków i instytucji. W roku 2012 wspólną pracę nad stworzeniem odpowiedniej do tego celu usługi webowej rozpoczęły ABES, INRA, Sciences-Po i Institut national polytechnique w Tuluzie. Po drugie, aby móc myśleć o interoperacyjności systemów, trzeba uporać się z różnicą języków programowania: PSP w HAL, a w repozytoriach lokalnych najczęściej JAVA, oraz z faktem, że stosowany w HAL język opisu usług webowych (WSDL) nie działa sprawnie w środowisku JAVA [Rigeade 2012, dok. elektr.].

Kilka miesięcy po opublikowaniu pracy Marine Rigeade konsorcjum Couperin wraz ze stowarzyszeniem dyrektorów bibliotek akademickich ADBU oraz instytucjami INRA i IRIA ogłosiły wyniki kolejnego okresowego badania kondycji repozytoriów instytucjonalnych [Synthèse... 2012, dok. elektr.]. Ponownie wykorzystano technikę ankiety. Stopa zwrotu kwestionariuszy była dużo wyższa niż w roku 2010 (178 na około 210 wysłanych), co skłoniło autorów raportu z badania do konstatacji, że uzyskane wyniki „można uznać za reprezentatywne”. 92% ankiet wypełniły biblioteki i ośrodki dokumentacji. Z informacji udzielonych przez respondentów wynikało, że 50% z nich miało już działające repozytoria, a 25% namyślało się nad takim projektem lub rozpoczynało jego wdrażanie. Komentarze do wyników badań zawarte w oficjalnym raporcie Couperin zostały sformułowane zwięźle, lecz dobitnie. Stwierdzono w nich na przykład, że „50% placówek z czynnym repozytorium we Francji w 2012 r. to ostatecznie dość mało. Trzeba zastanowić się, dlaczego 15% instytucji nie planuje ‘żadnego projektu’ i, stosownie do rezultatów tej analizy, wprowadzić narzędzia zachęty i wsparcia” [Synthèse... 2012, dok. elektr., p. 2]. Zakres przedmiotowy zbiorów gromadzonych w repozytoriach obejmował nauki ścisłe (78%), prawo, nauki ekonomiczne i nauki o zarządzaniu (52%), nauki humanistyczne (48%) oraz medycynę (32%). Nie odnotowano istotnej zmiany na lepsze jeśli chodzi o stopień rozproszenia infrastruktury baz: 80% repozytoriów funkcjonowało jako jednostki odosobnione,

zabezpieczające wyłącznie potrzeby własnej instytucji. Poza nielicznymi wyjątkami, np. uniwersytetów łączących się w drodze fuzji administracyjnej, większość placówek wolała zarządzać swoimi archiwami samodzielnie, aniżeli przystępować do większych struktur zbiorowych. Autorzy raportu ocenili, że największą trwałość rokuje łączenie repozytoriów instytucji sąsiadujących geograficznie, zwłaszcza jeśli dodatkowo są one ze sobą powiązane pod względem organizacyjnym lub administracyjnym. Zachęty, aby przystępować do PRES lub tzw. grup dziedzinowych (fr. *pôles thématiques*), nie trafiły na podatny grunt.

Ciekawe wyniki dała kwantyfikacja odpowiedzi ankietowych na pytanie o pobudki tworzenia repozytoriów. Wbrew temu, co sugerowała Marine Rigeade, próbując dociec, dlaczego tak duża część rekordów w repozytoriach zawiera tylko metadane bibliograficzne [Rigeade 2012, dok. elektr.], aż 20% respondentów uznało, że możliwość zwiększenia liczby cytowań i indeksów wpływu własnej produkcji naukowej nie gra dla nich żadnej roli (fr. *sans objet*). Na pierwszych miejscach w hierarchii celów znalazły się natomiast rozpowszechnianie i kapitalizacja wyników badań własnych (bez troski o indeksy wpływu i inne wskaźniki bibliometryczne?) oraz digitalizacja zasobów o wartości historycznej. Choć autorzy raportu ograniczają się w tym miejscu do neutralnej uwagi, że motywacje tworzenia repozytoriów „są bardzo zróżnicowane”, to jednak trudno nie dostrzec, że pozostają one odległe od założeń zielonej drogi, szczególnie gdy chodzi o sposób rozumienia terminu „kapitalizacja dorobku naukowego” (fr. *valorisation*) w abstrakcji od cytowań, pomijanie roli *peer-reviewing* w podnoszeniu jakości autoarchiwizowanych pre-printów oraz dość zaskakującą (trzeba oddać część racji Schöpfelowi i Prost...) preferencję dla digitalizacji dokumentów historycznych.

Kolejne pytanie ankiety dotyczyło profili kompetencyjnych i zakresu obowiązków pracowników zaangażowanych w tworzenie lub bieżącą obsługę repozytorium. Także i ta edycja badania potwierdziła, że największe obciążenia z tego tytułu ponoszą biblioteki ankietowanych instytucji, w większości przypadków odpowiadające za cały proces – od projektowania do rozruchu bazy, bez przydziału dodatkowych etatów. Rozkład czynności wykonywanych przez bibliotekarzy i dokumentalistów był następujący: wprowadzanie e-printów do bazy – 59%, kontrola metadanych – 74%, moderacja naukowa – 9%. 59% instytucji zgłosiło, że za składanie e-printów do repozytorium odpowiadają u nich bezpośrednio naukowcy, a 20% – że zadanie to wypełniały sekretariaty poszczególnych jednostek organizacyjnych. Na marginesie dodam, że w odróżnieniu od poprzedniego badania jako największą bolączkę w bieżącej pracy bibliotekarze wskazali już nie wewnętrzne kłopoty komunikacyjne z innymi komórkami organizacji, ale trudności we współdziałaniu z naukowcami (83% głosów). 40% bibliotekarzy określiło stopień zaangażowania i zainteresowania kadry naukowej jako „słaby”

podkreślając jednocześnie potrzebę kształtowania pozytywnego wizerunku repozytorium za pomocą szerokiej gamy narzędzi edukacyjno-promocyjnych.

Dla większej czytelności dalsze wnioski sformułowane w raporcie Couperin z 2012 r. ujmę w punktach.

1. Wybór platformy softwarowej. 56% instytucji posadowiło swoje zbiory na platformie HAL, przy czym niektóre (w raporcie nie wyszczególniono, ile) nie stworzyły własnego „portalu instytucjonalnego”, czyli nie stosowały „pieczętki”. Należy chyba przez to rozumieć, że po prostu zlecały odpowiednim służbom (np. swoim bibliotekom) deponowanie e-printów bezpośrednio w HAL generycznym (HyperArticles en Ligne). Wśród instytucji, które utrzymywały własne repozytoria poza środowiskiem HAL, najwięcej, 44%, stosowało oprogramowanie ORI-OAI, jednak tylko 6 deklarowało udział w prowadzonych wówczas testach usług webowych zabezpieczających przesył danych do HAL, a 7 przyznało się w tym zakresie do zupełnej bierności, stwierdzając że nie interesowały się dotąd możliwościami przełamania informatycznej izolacji swoich repozytoriów. W 56% instytucji nie były one powiązane ani z żadnym zewnętrznym systemem informacyjno-wyszukiawczym, ani z platformą intranetową lub lokalnym systemem bibliotecznym. W grupie respondentów, którzy twierdzili, że ich repozytorium jest całkowicie (?) lub w zadowalającym stopniu sprzężone z serwisami funkcjonującymi w ich bliższym i dalszym otoczeniu, a stanowili oni 30% ogółu ankietowanych, 9 instytucji zgłosiło, że naukowcy odprowadzają swoje e-rinty do bazy za pośrednictwem intranetu, z poziomu indywidualnego konta, 2 importowały metadane z archiwum prac doktorskich Apogée, a kolejne 5 przysyłało je do systemu zarządzania informacją o badaniach naukowych GRAAL [Ollendorff 2013, dok. elektr.]. Autorzy raportu podkreślili konieczność przyspieszenia tempa prac nad interoperacyjnością systemów, uznając że główną rolę w tym zakresie powinni odegrać specjaliści HAL. Za dobry prognostyk uznali fakt, że – jak wynikało z odpowiedzi ankietowych – większość instytucji, także tych jeszcze nieskomunikowanych z platformą centralną, miała pracowników zdolnych sprawnie posługiwać się usługami webowymi proponowanymi przez HAL.

2. Polityka instytucji w dziedzinie autoarchiwizacji. Według autorów raportu w lokalnych repozytoriach instytucjonalnych łatwiej jest bezkonfliktowo wprowadzać selektywny obowiązek autoarchiwizacji określonych typów dokumentów, np. pełnych tekstów prac doktorskich, które w wielu szkołach wyższych są obligatoryjnie przekazywane do krajowego repozytorium STAR. Wydaje się, że decyzje tego rodzaju powinny jednak zapadać na szczeblu senatu uczelni, obejmując odgórnie wszystkie jednostki podrzędne, nie zaś, jak dotychczas, na poziomie rad poszczególnych instytutów lub katedr. Instytucje korzystające z platformy HAL stosują natomiast raczej politykę zachęty.

3. Gromadzone typy dokumentów i proporcja pełnych tekstów. W lokalnych repozytoriach instytucjonalnych dominują prace doktorskie (72%) oraz post-printy artykułów z czasopism i referatów z materiałów konferencyjnych. 40% repozytoriów nie odmawia przyjmowania żadnego typu piśmiennictwa, 40% nie przyjmuje dokumentacji administracyjno-księgowej. 80% instytucji utrzymujących własne repozytoria lokalne deklaruje, że odsetek rekordów zawierających pełne teksty wnosi 100%. W „portalach instytucjonalnych” na platformie HAL odsetek ten nie przekracza 37%.

4. Statystyki i inne usługi. Na odnośne pytanie ankietowe uzyskano nieliczne odpowiedzi, szczególnie ze strony instytucji korzystających z platformy HAL. Spośród pozostałych 16 zawiadomiło, że prowadzi zliczanie statystyk, jednak tylko w 10 przypadkach dostarczone dane okazały się zdadne do obróbki. Zdaniem autorów raportu wynikało stąd, że oferta usług służących monitorowaniu ruchu i zachowań użytkowników była miernie rozwinięta zarówno w serwisie centralnym HAL, jak i w repozytoriach lokalnych. W związku z tym postulowali oni z naciskiem, że „konieczne jest wyposażenie repozytoriów w kompletny zestaw parametrowalnych, intuicyjnych w użyciu narzędzi kontroli statystycznej, które będą interoperacyjne z innymi systemami (eksport) i oparte na prostych kryteriach akceptowanych przez wszystkie podmioty (liczba odsłon rekordu, liczba pobrań pełnego tekstu itd.)”. Wyniki ankiety Couperin uzmysłowiły również, że w wielu instytucjach, w których planowano przeniesienie własnych zbiorów na platformę HAL lub już z niej korzystano, pracownicy oddelegowanych do tego służb nie wiedzieli, jakie usługi statystyczne są tam dostępne. Autorzy raportu potraktowali to jako oznakę „jaskrawego niedostatku szkoleń dla użytkowników portali instytucjonalnych w HAL” i zaapelowali do koordynatorów platformy (w 2012 r. funkcję tę sprawowała opisana niżej Bibliothèque scientifique nationale), aby ustalić kompromisową listę usług, na które klienci platformy zgłaszają zapotrzebowanie, zająć się ich wdrożeniem, a następnie przeszkolić zainteresowanych, aby potrafili efektywnie się nimi posługiwać.

Raport z badania Couperin z 2012 r. kończy się wnioskiem, że mimo „niezaprzeczalnych postępów” widocznych w porównaniu z rokiem 2008, wciąż nierozwiązana pozostaje kwestia jednorodności i spójności krajowego systemu repozytoriów instytucjonalnych, którego podstawą powinno być, a nie jest, stosowanie wspólnych praktyk i standardów. Za pilną potrzebę uznano systematyczne angażowanie uczonych w proces autoarchiwizacji, pod warunkiem jednak, że w organizacji pracy repozytorium uwzględni się ich dezyderaty: krótsze i mniej szczegółowe formularze rejestracji e-printu, generowanie statystyk pobrań i wskaźników bibliometrycznych, interoperacyjność systemów, dzięki której e-print złożony w jednej bazie zostanie automatycznie wyeksportowany do kilku

następnych. Powiedziano też, że opcja stworzenia własnego portalu na platformie HAL jest dla wielu instytucji atrakcyjna głównie dlatego, że wymaga niewielkich nakładów ludzkich i finansowych, a zarazem – w subiektywnym odczuciu – zdaje się zwalniać z odpowiedzialności za jakość budowanego zasobu, skutkując m.in. rozluźnieniem dyscypliny w zakresie deponowania pełnych tekstów. Tu właśnie, w opinii autorów raportu, należałoby szukać przyczyn nieproporcjonalnie dużego udziału samych rekordów bibliograficznych w portalach instytucjonalnych HAL. Uważają oni, że nadeszła najwyższa pora, aby BNS, jako organ koordynujący HAL, uregulowała i zacieśniła współpracę między użytkownikami platformy przez zrzeczenie ich w sieci umożliwiającej wymianę dobrych praktyk [Synthèse... 2012, dok. elektr.].

Odnosząc się do diagnozy zawartej w raporcie Couperin, współbrzmiejącej zresztą z wynikami badań własnych i z postulatami Marine Rigeade, można wysunąć przypuszczenie, że sygnalizowane trudności z przekształceniem rozproszonej galaktyki francuskich repozytoriów instytucjonalnych w sprawnie działającą sieć, której podwaliną byłaby interoperacyjność i przekształcenie HAL w dostawcę usług (ang. *service provider*), a tym samym agregowanie metadanych (ang. *harvesting*) i metawyszukiwanie, były przynajmniej po części następstwem braku odpowiednich ram organizacyjnych, przed czym, przypomnę, już w 2005 r. przestrzegał Gallezot [Gallezot 2005, p. 129]. Impas w tej dziedzinie pogłębiał się począwszy od roku 2008, gdy siłą inercji wygasło „Memorandum”, a jednocześnie okazało się, że wystosowany rok wcześniej przez ANR apel do instytucji nauki o zabezpieczenie stałego dopływu treści do HAL pozostał w zasadzie martwą literą [Bosc, Dillaerts 2012, dok. elektr.]. W roku 2009 na zlecenie MESR zespół badawczy Komitetu Informacji Naukowej i Technicznej CNRS pod kierunkiem Jeana Salençonu sporządził raport oceniający m.in. stan rozwoju komunikacji naukowej w środowisku elektronicznym we Francji [Salençon dir. 2008, dok. elektr.]. Zawarte w nim rekomendacje zostały częściowo uwzględnione w rządowych dokumentach programowych określających kierunki krajowej polityki badań naukowych i kształcenia wyższego, dołączonych jako aneks do projektu ustawy budżetowej na rok 2009. W ślad za opracowaniem Salençonu jako jeden z priorytetów wyznaczono w nich „przywrócenie dynamiki rozwoju otwartych archiwów przez uznanie HAL za referencyjną platformę krajową; w tym celu koordynacja HAL będzie opierać się na zorganizowanym współdziałaniu między twórcą serwisu, CNRS, a instytucjami naukowo-badawczymi i szkołami wyższymi”, na przykład w ramach odrębnie powołanej specjalnej jednostki operacyjnej zwanej *unité mixte de service* [Bosc, Dillaerts 2012, dok. elektr.].

W roku 2010 MESR postanowiło zreformować model zarządzania platformą HAL, powierzając to zadanie nowo utworzonemu ciału federacyjnemu o nazwie



Bibliothèque Scientifique Numérique (BSN), zrzeszającemu, jak kiedyś koalicja sygnatariuszy „Memorandum”, wszystkie liczące się organizacje francuskiego świata nauki, w tym obie konferencje rektorów szkół wyższych (CPU, CGE). Wzmianki o jego powołaniu pojawiły się już w roku 2006 w raporcie Salençon [Salençon dir. 2006, dok. elektr.], a następnie w roku 2008 w dokumencie MESR pt. „Les très grandes infrastructures de recherche” („Wielkie infrastruktury badawcze”), choć dokładny opis powierzonych mu zadań, w rozbiciu na 9 bloków tematycznych, przedstawiono dopiero w roku 2011 [PAP 2010, p. 94; Groudiev 2012, dok. elektr.]. W praktyce działaniem platformy niezmiennie kierowało CCSD, tyle tylko, że odtąd jako operator delegowany przez BSN. Wspomniana w dokumentach rządowych specjalna jednostka operacyjna, *unité mixte de service* (dosłownie: międzyinstytucjonalna jednostka usługowa, UMS), nadal nie istniała, nie zawarto też żadnej umowy bądź porozumienia, w którym zostałyby skodyfikowane relacje między repozytorium centralnym HAL a lokalnymi repozytoriami uczelni i ośrodków naukowo-badawczych [Bosc, Dillaerts 2012, dok. elektr.; PAP 2010, p. 94].

Sytuacja zmieniła się dopiero w drugim kwartale 2013 r., kiedy to z inicjatywy CNRS i MESR doszło do podpisania, z datą 2 kwietnia, nowej „Konwencji partnerstwa na rzecz otwartych repozytoriów i rozwoju interoperacyjnej platformy HAL”, do której przystąpiły: Bibliothèque Nationale de France, konsorcjum Couperin, CLEO (Centre pour l'édition électronique ouverte), platforma remediacji i publikacji czasopism elektronicznych PERSEE, stowarzyszenie ADBU oraz 13 ośrodków naukowo-badawczych. Na mocy „Konwencji” BSN jest naczelną instancją władną określać cele krajowej polityki w zakresie rozwoju repozytoriów OA, a CCSD występuje w roli podporządkowanego jej organu wykonawczego. CCSD ma w zamierzeniu przekształcić się w *unité mixte de service* (UMS) i, za pośrednictwem kolegialnego Komitetu Naukowo-Technicznego, zapewnić warunki wspólnego zarządzania platformą HAL przez wszystkich partnerów „Konwencji”. Kontrolę nad UMS mają sprawować łącznie: CNRS, organizacja reprezentująca wszystkie instytuty naukowo-badawcze oraz uczelnia reprezentująca wszystkie szkoły wyższe. Do zadań Komitetu Naukowo-Technicznego, działającego pod egidą BSN, należy: 1) tworzenie narzędzi i rozwiązań zapewniających interoperacyjność między współpracującymi repozytoriami, bezpieczeństwo długoterminowego przechowywania zasobów oraz opracowywanie i udostępnianie jednolitych reguł identyfikacji i opisu dokumentów, a także wspólnych strategii promowania francuskich repozytoriów w sieci, 2) prowadzenie szkoleń dla nowo przystępujących partnerów, 3) kodyfikowanie i rozpowszechnianie dobrych praktyk w gronie sygnatariuszy „Konwencji”. Zasięg formalny dokumentów kwalifikowanych do HAL ograniczono

do prac doktorskich oraz pre- i post-printów artykułów, co do których oczekuje się, że będą reprezentowały poziom naukowy właściwy tekstom przedkładanym do recenzji.

W myśl artykułu 4. „Konwencji” instytucje partnerskie mają do wyboru dwa równorzędne sposoby uczestnictwa w krajowym systemie autoarchiwizacji dorobku nauki: 1) albo korzystają z platformy HAL, w której CCSD tworzy dla nich odrębny portal („pieczętka”), 2) albo zakładają własne repozytorium zdecentralizowane, którego zasoby (opisy bibliograficzne i pełne teksty) mają zostać zreplikowane z HAL generycznym, pod warunkiem że spełniają szereg naukowych i technicznych kryteriów kompatybilności z HAL, wymienionych w załączniku do „Konwencji”. Jako podstawowy model eksportu metadanych wewnątrz systemu HAL-repozytoria lokalne i na zewnątrz przyjęto OAI\_DC. Wszystkie rekordy w HAL generycznym i w portalach instytucjonalnych wyposażone są również w pliki sitemap umożliwiające ich zaindeksowanie przez roboty wyszukiwarek horyzontalnych. Instytucje, które decydują się na utrzymywanie własnych repozytoriów, mogą zażyczyć sobie, aby metadane były pobierane w formacie OAI cache. Za wykonywanie kopii bezpieczeństwa i długoczasowe archiwizowanie zasobów w modelu OAIS odpowiada CINES (Centre informatique national de l'enseignement supérieur, Krajowe Centrum Informatyczne Szkolnictwa Wyższego). HAL generyczny skomunikowany jest z bazami ArXiv, PubMed Central i Repec [Berthaud 2013, dok. elektr.; Convention... 2013, dok. elektr.].

Podpisanie „Konwencji” można uznać za duży krok naprzód po kilku latach stagnacji. Szkoda tylko, że ani na stronie internetowej Bibliothèque scientifique nationale, ani w dostępnym *online* piśmiennictwie fachowym nie sposób znaleźć informacji o tym, czy CCSD przeobraził się już w *unité mixte de service* i czy na jego czele ukonstytuował się Komitet Naukowo-Techniczny mający trzymać pieczę nad funkcjonowaniem platformy HAL i godzić interesy wszystkich partnerów, a jeśli tak, to jakie decyzje podjął i co zamierza. Mimo że „Konwencja” nakazuje przewodniczącemu Komitetu przedstawianie corocznych sprawozdań z działalności, w październiku 2014 r., czyli ponad rok od daty jej wejścia w życie, nigdzie w dostępnych źródłach nie ma wzmianki o ewentualnym opublikowaniu takiego raportu. Do dyspozycji pozostaje jedynie dokument z 2012 r., w którym BSN wyklada swoją strategię działania do roku 2015, przewidującą – w odniesieniu do bloku tematycznego 4. „Otwarte archiwa” – zrealizowanie trzech priorytetów: 1) odnowienie porozumienia określającego warunki partnerskiego współzarządzania platformą HAL, 2) reaktywowanie portalu archives.ouvertes.fr w wersji dwu-, a docelowo trójjęzycznej, 3) utworzenie analogicznego do Sherpa/Romeo serwisu Héloïse, informującego o polityce wydawców komercyjnych w dziedzinie autoarchiwizacji [Stratégie 2012-2015, dok. elektr.].

Wszystkie trzy priorytety zostały wypełnione, serwis Héloïse funkcjonuje od marca 2012 r. Sukcesem, którego część można zaliczyć na konto BSN jako organu będącego formalnym następcą CCSD, jest natomiast serwis Isidore (pełna nazwa: *Intégration de Services, Interconnexion de DOnnées de la Recherche et de l'Enseignement*, Integracja serwisów, agregacja danych wytwarzanych przez ośrodki badawcze i uczelnie), agregator z metawyszukiwarką penetrujący około 2650 elektronicznych źródeł informacji naukowej z zakresu nauk humanistycznych i społecznych. Serwis, zaprojektowany przez kompetentną jednostkę CNRS, a obecnie osadzony na platformie cyfrowej Huma-Num, został uruchomiony przez CCSD w roku 2011, przy współpracy trzech komercyjnych firm: Antidot, Sword i Mondeca. W architekturze systemu opartego na wykorzystaniu protokołu OAI-PMH, RDFa (dla ustrukturyzowanych stron www) oraz RSS i Atom (głównie dla serwisów społecznościowych) Isidore zajmuje pozycję dostawcy usług, współdziałając z trzema grupami dostawców treści: platformami elektronicznego publikowania (np. Cairn.info, Persée, Revues.org, Erudit), bibliotekami cyfrowymi (np. Gallica, Bibliothèque Sainte-Geneviève, Bibliothèque Inter-universitaire de médecine) oraz archiwami i repozytoriami instytucjonalnymi OA (HAL-SHS, TEL) [Isidore 2014, dok. elektr.; Plateforme Isidore 2014, dok. elektr.].

Ogólny wydzźwięk „Konwencji” wskazuje, że francuskie władze ministerialne, jak też najbardziej wpływowe instytucje nauki i świata akademickiego, chcą odejść od akcentowania dualizmu HAL-repozytoria instytucjonalne na rzecz wizji sprzężonego systemu. Tym niemniej nadal bez odpowiedzi pozostają pytania o to: 1) co stoi na przeszkodzie, aby docelowo wzbogacić HAL w funkcję agregowania metadanych, 2) w jaki sposób planuje się zapewnić szerokie udostępnianie i widoczność tych zasobów francuskich, które chętnie gromadzone są w wielu repozytoriach instytucjonalnych, a nie kwalifikują się do replikacji w HAL, bo nie są pracami doktorskimi ani dokumentami odpowiadającymi kryteriom artykułu gotowego do zrecenzowania. To drugie pytanie wydaje się szczególnie ważne, jako że liczne duże instytucje naukowe we Francji postanowiły stworzyć odrębne repozytoria właśnie po to, aby ominąć w ich opinii nadmiernie restrykcyjne zasady akwizycji zbiorów w HAL [Van der Graaf, Van Eijndhoven 2008, dok. elektr.; Rigeade 2012, dok. elektr.].

Przegląd piśmiennictwa naukowego dotyczącego rozwoju zielonej drogi OA we Francji, które wyselekcjonowałam według kryteriów podanych na wstępie tego artykułu, z zawężeniem zasięgu chronologicznego do lat 2005-2013 i z celową preferencją dla publikacji dostępnych w pełnym tekście *online*, nasuwa następujące wnioski:

**Po pierwsze**, dokładna ilościowa ocena stanu francuskich archiwów OA, jeśli zgodzić się, że ta ogólna nazwa oznacza, w tamtejszym kontekście, zarówno ar-

chiwa dziedzinowe, jak i repozytoria instytucjonalne oraz heterogeniczny zespół kolekcji budowanych na platformie HAL, wraz z narodowym archiwum wielodziedzinowym HyperArticles en Ligne, wydaje się utrudniona ze względu na: a) brak jednolitej oficjalnej listy archiwów, wskutek czego autorzy podejmujący się badań w tym zakresie musieli (i nadal muszą) samodzielnie kompilować wykazy archiwów na podstawie wielu rozproszonych źródeł [Synthèse... 2012, dok. elektr.; Schöpfel 2013, dok. elektr.], dla uzupełnienia danych posiłkując się też techniką sondażu bezpośredniego [Rigeade 2012, dok. elektr.]; zadania tego nie ułatwia bynajmniej fakt, że według badań Couperin z 2012 r. tylko 40% francuskich repozytoriów było zarejestrowanych w międzynarodowych serwisach typu *reference site*, b) brak konsensusu w sprawie kryteriów typologizowania otwartych archiwów, prowadzący do tego, że wyniki poszczególnych kwantyfikacji odzwierciedlają nie tyle stan faktyczny, ile osobiste przeświadczenia i wybory badacza [vide Schöpfel, Prost 2010, dok. elektr.].

W odniesieniu do punktu a) diagnoza wynikająca z przeglądu piśmiennictwa wydaje się trafna, skoro 2 października 2013 r. konsorcjum Couperin oznajmiło powołanie nowej grupy roboczej do spraw Otwartego Dostępu (Groupe de travail pour l'accès ouvert aux publications de la recherche, GTAQ), której jako priorytet wyznaczono, oprócz przeprowadzenia kolejnej edycji badań rozpoczętych w roku 2007 [Bruley et al. 2007, dok. elektr.; Bruley 2008, dok. elektr.; Synthèse... 2012, dok. elektr.], także stworzenie ogólnodostępnej bazy danych rejestrującej wszystkie otwarte archiwa we Francji. Pierwsze posiedzenie GTAQ odbyło się 24 października 2013 r. [Couperin lance... 2013, dok. elektr.].

Jeśli zaś chodzi o punkt b), to warto wspomnieć o najnowszym artykule tandemu badawczego Schöpfel i Prost, ogłoszonym w 2013 r. na łamach „Revue canadienne des sciences de l'information et de bibliothéconomie”, w którym najpierw pośrednio przyznali oni, że rekonesans rozwoju zielonej drogi we Francji według podejścia indukcyjnego dał mało zadowalające rezultaty [Schöpfel, Prost 2010, dok. elektr.], a następnie, tym razem drogą dedukcji, stworzyli propozycję uniwersalnej typologii repozytoriów instytucjonalnych, różnicując je ze względu na główny cel: gromadzenie materiału *versus* kontrola jakości publikacji. W ten sposób wyodrębnili cztery podstawowe modele (według ich nazewnictwa: „archetypy”), a mianowicie: 1) platformy elektronicznego publikowania (fr. *plateformes de publication*), tj. repozytoria gromadzące wyłącznie bieżące publikacje naukowe pracowników, ewentualnie z ograniczeniem do wybranej formy piśmienniczej, np. Archive électronique de l'Institut Jean Nicod, PASTEL, 2) repozytoria typu *container*, przyjmujące materiały wszelkiego rodzaju, w tym szarą literaturę, pomoce dydaktyczne i zbiory danych (ang. *raw data*), np. HAL-IN2P3, 3) repozytoria z elementami profilu biblioteki cyfrowej (fr. *sites de patrimoine*),

w których obok publikacji bieżących typu *born digital* archiwizowane są również wybiórczo zdigitalizowane treści o charakterze historycznym, np. Publications et Travaux Académiques de Lorraine oraz 4) repozytoria działające pod rządami *deposit mandate*, mające rejestrować całość bieżącej produkcji publikacyjnej danej instytucji (fr. *dépôt*), np. HAL-Bioemco.

**Po drugie**, uwagę francuskich badaczy zajmujących się przedmiotową problematyką niezmiennie przykuwa kwestia uregulowań dotyczących obowiązku autoarchiwizacji. W świetle piśmiennictwa dzielą się oni na trzy obozy: zwolenników modelu Immediate Deposit/Optional Access (ID/OA) na czele z Héléne Bosc [Bosc 2008, dok. elektr.; Bosc 2009, dok. elektr.; Bosc, Dillaerts 2012; Dillaerts 2011, dok. elektr.], zwolenników pozyskiwania e-printów dzięki umowom zawieranim z wydawcami [Romary, Armbruster 2009, dok. elektr.] oraz takich, którzy nie podzielają wprawdzie przekonania, że samo narzędzie ID/OA jest w stanie zapewnić stuprocentowe pokrycie (fr. *exhaustivité*) bieżącej produkcji naukowej w repozytoriach, lecz uważają, że Francja bezwzględnie potrzebuje czytelnych rozstrzygnięć prawnych w tej materii [Schöpfel, Prost 2010, dok. elektr.].

W latach 2012-2013 CNRS kilkakrotnie zabierał głos, by opowiedzieć się za wprowadzeniem centralnego nakazu autoarchiwizacji: w czerwcu 2012 r. Komitet Etyczny CNRS (Comets) wzywał CNRS do opracowania, na użytek wszystkich członków francuskiej społeczności naukowej, wiążących wytycznych, które przewidywałyby systematyczne archiwizowanie e-printów na platformie HAL, jako że ona „jedyna we Francji ma środki, aby zagwarantować m.in. długoczasowe przechowywanie zbiorów cyfrowych” [Jonchère 2013, dok. elektr.]. W listopadzie 2012 r. CNRS złągodził nieco swoje stanowisko, rekomendując tym razem, aby polityka *deposit mandate* ustalana była indywidualnie przez uczelnie, instytuty naukowo-badawcze i agencje finansujące badania (fr. *bailleurs de fonds*), z dopuszczeniem obowiązkowej autoarchiwizacji także w lokalnych repozytoriach, nie zaś wyłącznie w HAL [Assises de la recherche 2012, dok. elektr.]. W marcu 2013 r. oświadczenie w tej sprawie wydało także konsorcjum Couperin, zwracając uwagę, że o ile kilka dużych instytutów naukowo-badawczych (np. INRIA, IFREMER) postanowiło wyartykułować i wdrożyć zdecydowaną politykę w zakresie *deposit mandate*, o tyle szkoły wyższe, mimo życzeń wyrażanych przez konferencje rektorów (CPU, CGE), są tym zainteresowane w nikłym stopniu [Bien universel... 2013, dok. elektr.].

Wydaje się jednak, że ożywionej dyskusji na łamach naukowych i deklaracjom gremiów wysokiego szczebla nie towarzyszą równie aktywne działania w praktyce. Jest to tylko przypuszczenie, gdyż dostępne piśmiennictwo, przynajmniej to z lat 2009-2012, daje w tym względzie obraz dość chaotyczny i nacechowany sprzecznościami. I tak Joachim Schöpfel i Héléne Prost donoszą w swoim komu-

nikacie z 2011 r., że w kwietniu 2010 r. podczas specjalnie zwołanej konferencji dyrekcja INRA ogłosiła, że z początkiem roku 2011 wprowadzi obowiązek archiwizacji e-printów w zmodernizowanym repozytorium ProdiNRA2 dla wszystkich zatrudnionych badaczy [Schöpfel, Prost 2011, dok. elektr.]. W roku 2012 H el ene Bosc pisze z kolei, że nie znalazła we Francji  adnej instytucji odwo ujacej si e do *deposit mandate* (co oznacza oby, że plany INRA spaliły na panewce?). Niemal w tym samym czasie w serwisie Couperin.org ukazują si e dwa dokumenty b edące efektem drugiego badania stanu rozwoju zielonej drogi we Francji, tj. synteza liczbowa odpowiedzi na pytania ankiety (dane ilo ciowe brutto)[R esum e des r esultats... 2012, dok. elektr.] oraz oficjalne om owienie wyników [Synth ese 2012, dok. elektr.]. W pierwszym z nich przy pytaniu „Jaka jest polityka państwa instytucji w sprawie autoarchiwizacji?” widnieje liczba 31 instytucji, kt ore o wiadczyły, że stosują model obowi zkowy. W drugim dokumencie o zagadnieniu *deposit mandate* nie ma natomiast wcale mowy [Synth ese 2012, dok. elektr.]. Gdy w styczniu 2013 r. Christine Ollendorff relacjonuje wyniki badania Couperin podczas konferencji „G en eraliser l'acc es ouvert aux r esultats de la recherche”, stwierdza że zidentyfikowano w nim 6 instytucji, kt orych w lady narzuciły pracownikom obowi zek autoarchiwizacji, przyjmując że sprawozdania z dzia alno ci naukowej na potrzeby ocen parametrycznych b edą opracowywane wy lacznie na podstawie metadanych ekstrahowanych z repozytorium [Ollendorff 2013, dok. elektr.]. Trzy lata wcześnie, w roku 2009, Schöpfel i Prost naliczyli 16 repozytori w instytucjonalnych funkcjonujących w ramach *deposit mandate*.

Wi kszą jasno c przynoszą badania Laurent Jonch ere'a z lutego 2013 r. [Jonch ere 2013, dok. elektr.]. W swoim por ownawczym studium, opracowanym na potrzeby projektu implementacji repozytorium Uniwersytetu w Rennes, szczeg olowo scharakteryzowa  on polityk  francuskich instytucji naukowych i uczelni wobec *deposit mandate*, konfrontując rodzime rozwi zania z zaleceniami Unii Europejskiej oraz praktykami stosowanymi w Belgii i w Wielkiej Brytanii. Na podstawie samodzielnie przeprowadzonych wywiad w, analizy pi smiennictwa oraz informacji dost epnych w serwisach ROARMAP, Juliet i na stronach domowych instytucji zidentyfikowa  og olem 12 o rodk w naukowo-badawczych i szk l wy szych, nie licząc ANR i CNRS, w kt orych w jakikolwiek sp os b sformalizowano powinno ci badaczy w zakresie autoarchiwizacji lub przynajmniej do takich krok w si e zobowi zано. Model Immediate Deposit/Optional Access (ID/OA) obowi zuywa  w 6 instytutach naukowo-badawczych: IRSTEA, dawniej CIMAGREF (tylko post-printy), IFREMER, CIRAD, INRA, Laboratoire de psychologie et neurosciences cognitives (post-printy i szara literatura), INERIS (obowi zkowe zg aszanie przynajmniej opis w bibliograficznych). Przyj e a go r ownież europejska agencja finansujaca badania naukowe EUR-Oceans Cons-

ortium (EOC), mająca wśród członków 8 instytucji francuskich (m.in. INSU-CNRS, UBO, UPMC, Université Aix-Marseille 2). Dwa inne instytuty, INRIA i INSERM, ograniczały się do stymulowania autoarchiwizacji przez „silną zachęćę” (fr. *forte incitation*). Jedyńą uczelnią, która – według informacji z lutego 2013 r. – rozważała opcję ID/OA, był Uniwersytet w Nancy. Stwierdzając, że szkołom wyższym z dużym trudem przychodzi regulowanie zasad autoarchiwizacji do robku naukowego, Jonchère wyjaśniał, że w dużej mierze wynika to ze specyficznej organizacji pracy uczonych w ramach międzyinstytucjonalnych zespołów badawczych (fr. *laboratoires, groupes de recherche*), bowiem dla wydawanych przez te zespoły publikacji niełatwo jest ustalić afiliację pierwszego autora, a więc też wskazać, która z dwóch lub więcej zatrudniających go instytucji (uczelnia? zespół badawczy?) powinna wyegzekwować obowiązek autoarchiwizacji. Na wypadek, gdyby czytelnicy dziwiliby się rozbieżności jego ustaleń z wynikami zawartymi we wcześniejszej o rok syntezie liczbowej wyników ankiety Couperin [Résumé des résultats... 2012, dok. elektr.], uściślał, że wymieniona tam liczba 31 instytucji stosujących rzekomo *deposit mandate* mieściła w sobie również wszystkie uczelnie wymagające tylko autoarchiwizacji prac doktorskich.

**Po trzecie**, na podstawie piśmiennictwa trudno jest zrekonstruować rzeczywiste postępy zarówno w integrowaniu krajowego systemu repozytoryjnego skupionego wokół HAL, jak i w rozbudowie funkcjonalności poszczególnych repozytoriów instytucjonalnych. Na pewno jest ku temu wola polityczna, o czym świadczy choćby wystąpienie Komitetu Etycznego CNRS (Comets) z 2012 r., w którym stwierdzono: „Konieczne jest zharmonizowanie zasad opisu dokumentów w HAL z tymi, jakie zaczyna się właśnie stosować w europejskim serwisie OpenAIRE (...)” [Comité d'éthique du CNRS... 2012, dok. elektr.]. Komunikaty naukowe z lat 2005-2012 obfitują w zapowiedzi ambitnych działań, niemniej o ich realizacji ginie potem słuch. Wystarczy przytoczyć przykład repozytorium Prod-INRA, które zgodnie z deklaracjami ogłoszonymi na cytowanej już konferencji w kwietniu 2010 r. miało zostać ściśle zintegrowane z innymi wewnętrznymi systemami informacyjno-wyszukiwawczymi Instytutu, m.in. z tezaurusami terminologicznymi opracowanymi przez służby biblioteczno-dokumentacyjne oraz z systemem księgowo-budżetowym. Planowano wdrożenie usługi generowania wskaźników bibliometrycznych i rozmaitych zestawień, zgodnie z kryteriami ewaluacji stosowanymi przez AERES (Agence d'évaluation de la recherche et de l'enseignement supérieur, Agencja oceny badań naukowych i szkół wyższych). Zasilanie zasobu miało opierać się na agregowaniu pełnych tekstów i lub metadanych z HAL, PubMed (identyfikator PMID), z baz wydawców komercyjnych (DOI) i z innych repozytoriów, m.in. ośrodka badań agronomicznych CIRAD (Recherche agronomique pour le développement), z użyciem specjalnego for-

matu, o którym w 2011 r. mówiono, że jest „w trakcie opracowywania” [Schöpfel, Prost 2011, dok. elektr.]. Informacji o tym, czy opisane działania miały miejsce, nie ma ani w literaturze, ani w serwisie ProdiNRA, ani nawet w blogu OpenINRA stworzonym, aby powiadamiać zainteresowanych o polityce OA Instytutu, a w którym ostatni wpis w zakładce „Interoperacyjność” pochodzi z lipca 2012 r. [Le Hénaff 2012, dok. elektr.].

**Po czwarte**, w pracach przeglądowych i podsumowujących, bo takie były głównym przedmiotem mojego zainteresowania, stale przewija się wątek dyskusji nad optymalnym wariantem współistnienia centralnego repozytorium państwowego HAL i satelitarnych repozytoriów uczelni oraz innych instytucji naukowych, które, jak wynikałoby z niedawnych badań Joachima Schöpfela i Hélène Prost, cechują się we Francji nadzwyczajną dynamiką rozwojową [Schöpfel, Prost 2013, dok. elektr.]. Niespieszność, z jaką francuskie władze wdrażają rozwiązania mające zapewnić interoperacyjność poszczególnych systemów w ramach jednej sieci krajowej, zdaje się jednak sugerować, że drogowskazem wciąż pozostają dla nich kategoryczne słowa autorów Raportu Komitetu ds. Informacji Naukowej i Technicznej CNRS, wydane w roku 2009 pod redakcją Jeana Salençon, w którym napisano: „Rozkwit rozproszonych baz danych (serwerów uniwersyteckich, serwerów instytutów naukowo-badawczych), w których gromadzi się publikacje zatrudnionych tam uczonych, ale też inne, bardzo zróżnicowane materiały, stoi w sprzeczności z celami HAL. Komfort, jaki daje uczelni lub innej placówce naukowej ‘trzymanie wszystkiego w zasięgu ręki’, ściera się z wymogiem jakości (jednolitość dokumentów przyjmowanych do repozytorium) i eksponowania francuskiego dorobku nauki (...). Z tych właśnie przyczyn wykluczony jest rozwój w kierunku systemu, w którym HAL stałby się zwykłym agregatorem tych baz w ramach dystrybuowanej architektury OAI. Ogólniej rzecz biorąc, istnieje ryzyko rozbieżności między wymogami funkcjonowania repozytorium instytucjonalnego (które z definicji gromadziłoby wszystkie rodzaje dokumentów tworzonych przez daną instytucję, niekoniecznie wszystkie je udostępniając, choćby z racji poufności, np. w przypadku sprawozdań z wykonania kontraktów) i archiwum w pełni otwartego” [Salençon dir. 2009, p. 22].

Zdrowy rozsądek każe domniemywać, że francuską koncepcję zielonej drogi OA, której trzonem i najbardziej hołubionym osiągnięciem jest centralne archiwum HAL, można z powodzeniem pogodzić z równoległym rozwojem repozytoriów instytucjonalnych i że, jak żarliwie dowodziły m.in. Marine Rigeade i Muriel Foulonneau [Rigeade 2012, dok. elektr.; Foulonneau et al. 2008], nie ma tu żadnej antynomii. Patrząc na nieciągły i do pewnego czasu nieskoordynowany charakter prac nad uelastycznieniem tej koncepcji, hamowanych przez dogmat nadrzędności HAL, które nie może zostać zdegradowane do funkcji „zwykłego agregatora”,



i długo bronioną zasadę *dépôt direct*, wykluczającą import e-printów z innych lokalizacji, można odnieść wrażenie, że model francuski, rzeczywiście unikalny w Europie, pielęgnowany był jako swego rodzaju *French exception*.

Oznaką przełomu w tej logice myślenia może być jednak wystąpienie programowe minister Szkolnictwa Wyższego i Badań Naukowych Geneviève Fioraso z 24 stycznia 2013 r., określające oficjalne stanowisko rządu francuskiego w sprawie Otwartego Dostępu. Minister deklaruje w nim „jednoznaczne poparcie dla otwartych archiwów”, uzasadniając decyzję o reformie trybu zarządzania platformą HAL „w połączeniu z repozytoriami instytucjonalnymi”. Wyraża przekonanie, że w dziedzinie organizacji swobodnego dostępu do dorobku nauki „różnorodność możliwych do zastosowania modeli jest czynnikiem sukcesu”, trzeba tylko „zadbać o ich komplementarność”, wdrażając „zintegrowany program, w którym w pełni wyzyskane i zsumowane będą zalety każdego z tych modeli”. Zapowiada siedem kierunków działań mających stanowić dowód na to, że „francuski rząd chce uczestniczyć w tworzeniu europejskiej przestrzeni badań naukowych”, której fundamentem jest spójność [Discours... 2013, dok. elektr.].

## BIBLIOGRAFIA

- Armbruster, Chris; Romary, Laurent (2009). *Comparing Repository Types: Challenges and Barriers for Subject-Based Repositories, Research Repositories, National Repository Systems and Institutional Repositories in Serving Scholarly Communication* [dok. elektr.]. [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1506905](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1506905) [odczyt: 01.07.2014].
- Assises de la recherche et de l'enseignement supérieur. Compte-rendu par thème des contributions reçues à la date du 24/11/12* (2012) [dok. elektr.]. <http://intranet.cnrs.fr/intranet/documents/assises-contributions.pdf> [odczyt: 01.07.2014].
- Aubry, Christine; Janik, Joanna dir. (2005). *Les archives ouvertes. Enjeux pratiques. Guide à l'usage des professionnels de l'information*. Paris: ADBS, 332 p.
- Baruch, Pierre (2007). *La diffusion libre du savoir. Accès libre et archives ouvertes* [dok. elektr.]. L'Archicube No. 3, p. 77-95. [http://archivesic.ccsd.cnrs.fr/sic\\_00169330/en/](http://archivesic.ccsd.cnrs.fr/sic_00169330/en/) [odczyt: 01.07.2014].
- Berthaud, Christine (2013). *Éléments de la politique nationale pour les archives ouvertes* [dok. elektr.]. <http://couperin.sciencesconf.org/> [odczyt: 01.07.2014].
- Bertignac, Catherine; Gac, Dominique (2009). *La voie verte: les archives ouvertes*. In: Open Access Week. 19 October 2009. Brest: Institut Universitaire Européen de la Mer.
- Bibliothèque scientifique numérique (2014) [dok. elektr.]. <http://www.bibliothequescientifiquenumerique.fr/> [odczyt: 10.10.2014].
- Bien universel, par essence, l'article scientifique, n'est pas un bien comme les autres* (2013) [dok. elektr.]. In: Couperin.org. <http://www.couperin.org/groupes-de-travail-et-projets-deap/open-access/libre-acces/257-libre-acces/1052-article-scientifique> [odczyt: 01.07.2014].

- Bosc, Hélène (2008). *L'auto-archivage en France: deux exemples de politiques différentes et leurs résultats* [dok. elektr.]. Liinc em Revista Vol. 4, No. 2, p. 196-217. [http://archive-sic.ccsd.cnrs.fr/sic\\_00340885/en/](http://archive-sic.ccsd.cnrs.fr/sic_00340885/en/) [odczyt: 01.07.2014].
- Bosc, Hélène (2009). *Le droit des chercheurs à mettre leurs résultats de recherche en libre accès: appropriation des archives ouvertes par différentes communautés dans le monde* [dok. elektr.]. Terminal Vol. 102, p. 93-113. [http://halshs.archives-ouvertes.fr/sic\\_00340784/](http://halshs.archives-ouvertes.fr/sic_00340784/) [odczyt: 01.07.2014].
- Bosc, Hélène; Dillaerts, Hans (2012). *Le libre accès en France en 2012: Entre immobilisme et innovation* [dok. elektr.]. Liinc em Revista Vol. 8, No. 2, p. 323-347. [http://hal.archives-ouvertes.fr/docs/00/76/74/70/PDF/bosc\\_dillaerts\\_libre\\_acces2012.pdf](http://hal.archives-ouvertes.fr/docs/00/76/74/70/PDF/bosc_dillaerts_libre_acces2012.pdf) [odczyt: 01.07.2014].
- Bourrion, Daniel et al. (2006). *Les chercheurs en lettres et sciences humaines et les archives ouvertes* [dok. elektr.]. [http://halshs.archives-ouvertes.fr/docs/00/08/60/84/PDF/chercheurs\\_LSH\\_AO\\_v1.0.pdf](http://halshs.archives-ouvertes.fr/docs/00/08/60/84/PDF/chercheurs_LSH_AO_v1.0.pdf) [odczyt: 01.07.2014].
- Bruley, Caroline et al. (2007). *Résultats de l'enquête sur les projets d'archives ouvertes de la recherche dans les établissements du consortium Couperin* [dok. elektr.]. [http://www.couperin.org/archivesouvertes/IMG/pdf/resultats\\_enquete\\_version2.pdf](http://www.couperin.org/archivesouvertes/IMG/pdf/resultats_enquete_version2.pdf) [odczyt: 01.07.2014].
- Bruley, Caroline et al. (2008). *Bilan d'une enquête sur les archives ouvertes dans les établissements d'enseignement supérieur et de recherche* [dok. elektr.]. Ametist No. 2. <http://ametist.inist.fr/document.php?id=417> [odczyt: 01.07.2014].
- Carr, Leslie; Brody, Tim; Swan, Alma (2008). *Repository Statistics: What Do We Want to Know?* [dok. elektr.]. [http://pubs.or08.ecs.soton.ac.uk/30/1/OR08\\_submission\\_94.pdf](http://pubs.or08.ecs.soton.ac.uk/30/1/OR08_submission_94.pdf) [odczyt: 01.07.2014].
- Chronologie du LA en France* (2014) [dok. elektr.]. In: *Libre Accès à la communication scientifique*. <http://open-access.infodocs.eu/chronologie-du-la-en-france/> [odczyt: 01.07.2014].
- Comité d'éthique du CNRS (Comets) (2012) [dok. elektr.]. *Pour un renforcement des archives ouvertes HAL. Suivi de l'avis sur l'open access publié en juin 2012*. <http://www.cnrs.fr/comets/spip.php?article73> [odczyt: 01.07.2014].
- Convention de partenariat en faveur des archives ouvertes et de la plateforme mutualisée HAL* (2013) [dok. elektr.]. [http://www.bibliothequescientifiquenumerique.fr/IMG/pdf/Convention\\_HAL\\_AO\\_26-03-2013.pdf](http://www.bibliothequescientifiquenumerique.fr/IMG/pdf/Convention_HAL_AO_26-03-2013.pdf) [odczyt:01.07.2014].
- Couperin lance un nouveau GTAO* (2013) [dok. elektr.]. In: Couperin.org. <http://www.couperin.org/component/content/article/261-couperin/nos-activites/groupes-de-travail-et-projets-deap/a-la-une/1097-couperin-lance-un-nouveau-gtao> [odczyt: 01.07.2014].
- Creppy, Rachel (2007). *Archives ouvertes, archives institutionnelles et protocole français* [dok. elektr.]. Bulletin des Bibliothèques de France Vol. 52, No. 6, p. 42-45. <http://bbf.enssib.fr/consulter/bbf-2007-06-0042-008> [odczyt: 01.07.2014].
- Dillaerts, Hans (2010). *Analyse prospective du libre accès en France* [dok. elektr.]. In: Brodoux, Evelyne; Charton, Ghislaine dir. *Enjeux politiques du document numérique*. Paris: ADBS, p. 47-70. <http://www.adbs.fr/enjeux-politiques-du-document-numerique-92125.htm?RH=OUVRAGEADBS> [odczyt: 01.07.2014].
- Dillaerts, Hans (2012). *Libre accès à la communication scientifique et contexte français: prospective, développement et enjeux pour la créativité et l'interdisciplinarité?* [dok. elektr.]. <http://tel.archives-ouvertes.fr/tel-00768432> [odczyt: 01.07.2014].

- Discours de Geneviève Fioraso lors des 5e journées Open Access* (2013) [dok. elektr.]. <http://www.enseignementsup-recherche.gouv.fr/cid66992/discours-de-genevieve-fioraso-lors-des-5e-journees-open-access.html> [odczyt: 01.07.2014].
- Filliatreau, Ghislaine dir. (2008). *Rapport de l'Observatoire des Sciences et des Techniques* [dok. elektr.]. [http://www.obs-ost.fr/sites/default/files/BDC\\_R08\\_OST.pdf](http://www.obs-ost.fr/sites/default/files/BDC_R08_OST.pdf) [odczyt: 01.07.2014].
- Foulonneau, Muriel et al. (2007). *Réseaux d'archives institutionnelles en Europe: logiques de développement et convergences* [dok. elektr.]. <http://archivesic.ccsd.cnrs.fr/docs/00/20/50/49/PDF/foulonneaurevision2.pdf73> [odczyt: 01.07.2014].
- Foulonneau, Muriel et al. (2008). *Réseaux d'archives institutionnelles en Europe: logiques de développement et convergences* [dok. elektr.]. *Ametist* Vol. 2, No. 2. <http://archivesic.ccsd.cnrs.fr/docs/00/20/50/49/HTML/> [odczyt: 01.07.2014].
- Gallezot, Gabriel (2005). *Les archives ouvertes en France*. In: Aubry, Christine; Janik, Joanna dir. *Les archives ouvertes. Enjeux pratiques. Guide à l'usage des professionnels de l'information*. Paris: ADBS, 332 p. 97-115.
- Groudiev, Stéphanie (2012). *La Bibliothèque scientifique numérique. Un cadre politique de coordination des actions en faveur de l'IST* [dok. elektr.]. <http://bbf.enssib.fr/consulter/bbf-2013-01-0061-014> [odczyt: 01.07.2014].
- Harnad, Stevan et al. (2009). *Maximizing and Measuring Research Impact Through University and Research-Funder Open-Access Self-Archiving Mandates* [dok. elektr.]. *Wissenschaftsmanagement* Vol. 15, No. 4. <http://eprints.ecs.soton.ac.uk/16616/1/16-Harnad-Carr.pdf> [odczyt: 01.07.2014].
- Isidore (2104) [dok. elektr.]. <http://www.rechercheisidore.fr/> [odczyt: 01.07.2014].
- Joint Draft Agreement on Open Access in France* (2006) [dok. elektr.]. In: *Open Access Archivangelism*. <http://openaccess.eprints.org/index.php?/archives/105-Joint-Draft-Agreement-on-Open-Access-in-France.html> [odczyt: 01.07.2014].
- Jonchère, Laurent (2013). *Projet d'archive ouverte. Université de Rennes. Synthèse sur les politiques institutionnelles de libre accès à la recherche* [dok. elektr.]. [http://www.academia.edu/3079617/Synthese\\_sur\\_les\\_politiques\\_institutionnelles\\_de\\_libre\\_acces\\_a\\_la\\_recherche](http://www.academia.edu/3079617/Synthese_sur_les_politiques_institutionnelles_de_libre_acces_a_la_recherche) [odczyt: 01.07.2014].
- Jordan, Brett (2014). *20 ans d'ABES* [dok. elektr.]. <http://cdn.knightlab.com/libs/timeline/latest/embed/index.html?source=0Ak9U3b5-Ym1wdGJOWU1pckk4bVU3UIRpVW1iV0VSRHc&font=Bevan-PotanoSans&maptype=toner&lang=fr&height=650> [odczyt: 01.07.2014].
- L'Hostis, Dominique; Aventurier, Pascal; Bosc, Hélène (2006). *Archives ouvertes – Vers une obligation de dépôt? Synthèse sur les réalisations existantes, les pratiques des chercheurs et le rôle des institutions* [dok. elektr.]. <http://archivesic.ccsd.cnrs.fr/docs/00/11/86/46/PDF/Note-AO-version-211106-diff-Externe.pdf> [odczyt: 01.07.2014].
- Laloë, Franck (2006). *France's HAL, OAI interoperability, and Central vs Institutional Repositories. Open Access Archivangelism* [dok. elektr.]. <http://openaccess.eprints.org/index.php?/archives/139-Frances-HAL,-OAI-interoperability,-and-Central-vs-Institutional-Repositories.html> [odczyt: 01.07.2014].
- Laloë, Franck et al. (2007). *Open Archives (OA) the expectations Open Archives (OA), the expectations of the scientific communities. Open archives in all disciplines: the Hal project* [dok. elektr.]. CERN, OAI-5 meeting, 20<sup>th</sup> April 2007. <http://indico.cern.ch/event/5710/session/20/contribution/27/material/slides/1.pdf> [odczyt: 01.07.2014].

- Le Hénaff, Diane (2012). *Les idées pour ProdINRA 3* [dok. elektr.]. <http://ist.blogs.inra.fr/openinra/2012/07/03/des-idees-pour-prodinra-3/> [odczyt: 10.10.2014].
- Loi n° 2013-660 du 22 juillet 2013 relative à l'enseignement supérieur et à la recherche. NOR: ESRJ1304228L [dok. elektr.]. <http://www.legifrance.gouv.fr/affichTexte.do?cid-Texte=JORFTEXT000027735009&dateTexte=20130730> [odczyt: 01.07.2014].
- Malotau, Sandrine (2011). *Pour des bibliothèques engagées dans la diffusion des savoirs de l'université: l'exemple de l'Institut national polytechnique de Toulouse* [dok. elektr.]. Bulletin des bibliothèques de France Vol. 56, No. 1. <http://bbf.enssib.fr/consulter/bbf-2011-01-0054-011> [odczyt: 01.07.2014].
- Ollendorff, Christine (2013). *Enquête Archives ouvertes* [dok. elektr.]. <http://webcast.in2p3.fr/videos-ollendorff> [odczyt: 01.07.2014].
- Open Access in France. A State of the Art Report* (2010) [dok. elektr.]. [http://openaccess.inist.fr/sites/openaccess/IMG/pdf/Open\\_Access\\_in\\_France\\_SELLreport\\_april\\_2010.pdf](http://openaccess.inist.fr/sites/openaccess/IMG/pdf/Open_Access_in_France_SELLreport_april_2010.pdf) [odczyt: 01.07.2014].
- OpenAIRE (2014) [dok. elektr.]. <https://www.openaire.eu/search/openaire-data-providers> [odczyt: 10.10.2014].
- Plateforme Isidore (2014) [dok. elektr.]. In: *Wikipedia*. [http://fr.wikipedia.org/wiki/Plateforme\\_Isidore](http://fr.wikipedia.org/wiki/Plateforme_Isidore) [odczyt: 01.07.2014].
- Pour une politique nationale de l'IST. Proposition commune Couperin-ADBU-AURA* (2009) [dok. elektr.]. [http://www.couperin.org/IMG/pdf/Politique\\_nationale\\_IST\\_Couperin-ADBU-AURA.pdf](http://www.couperin.org/IMG/pdf/Politique_nationale_IST_Couperin-ADBU-AURA.pdf) [odczyt: 01.07.2014].
- Quels sont les critères qui permettent de différencier une archive ouverte d'une bibliothèque numérique?* (2010) [dok. elektr.]. <http://www.enssib.fr/content/quels-sont-les-criteres-qui-permettent-de-differencier-une-archive-ouverte-dune-bibliotheque> [odczyt: 01.07.2014].
- Résumé des résultats. Enquête archives ouvertes 2012* (2012) [dok. elektr.]. In: Couperin.org. <http://www.couperin.org/fr/news/261/1004> [odczyt: 01.07.2014].
- Rigeade, Marine (2012). *Les archives ouvertes institutionnelles en France: état des lieux et perspectives* [dok. elektr.]. <http://www.enssib.fr/bibliotheque-numerique/notices/5-6708-les-archives-ouvertes-institutionnelles-en-france-etat-des-lieux-et-perspectives> [odczyt: 01.07.2014].
- Salençon, Jean dir. (2008). *Rapport du Comité IST. Information Scientifique et Technique* [dok. elektr.]. [http://cache.media.enseignementsup-recherche.gouv.fr/file/2008/65/8/Rapport\\_IST-Juin\\_2008\\_31658.pdf](http://cache.media.enseignementsup-recherche.gouv.fr/file/2008/65/8/Rapport_IST-Juin_2008_31658.pdf) [odczyt: 01.07.2014].
- Siewicz, Krzysztof (2012). *Otwarty dostęp do publikacji naukowych. Kwestie prawne* [dok. elektr.]. [http://repozytorium.ceon.pl/bitstream/handle/123456789/335/K\\_Siewicz\\_Otwarty\\_dostep\\_do\\_publicacji\\_naukowych.pdf?sequence=4](http://repozytorium.ceon.pl/bitstream/handle/123456789/335/K_Siewicz_Otwarty_dostep_do_publicacji_naukowych.pdf?sequence=4) [odczyt: 01.07.2014].
- Siewicz, Krzysztof (2013). *Prawne możliwości wprowadzenia otwartego mandatu wobec publikacji naukowych* [dok. elektr.]. <http://ceon.pl/images/dspace/ceon-memorandum-otwarty-mandat.pdf> [dostęp: 01.07.2014].
- Schöpfel, Joachim; Müller, Uwe (2014). *Evaluer la qualité des archives ouvertes: le certificat DINI* [dok. elektr.]. Partnership: the Canadian Journal of Library and Information Practice and Research Vol. 9, No. 1. <http://microblogging.infodocs.eu/wp-content/uploads/2014/04/2733-16365-3-PB.pdf> [odczyt: 01.07.2014].
- Schöpfel, Joachim; Prost, Hélène (2009). *L'accès libre en mouvement. Journées d'étude sur les archives ouvertes du consortium COUPERIN*, 2-3 avril 2009 [dok. elektr.].

- [http://hal.archives-ouvertes.fr/docs/00/39/69/57/PDF/CR\\_JEAO\\_COUPERIN\\_2009.pdf](http://hal.archives-ouvertes.fr/docs/00/39/69/57/PDF/CR_JEAO_COUPERIN_2009.pdf) [odczyt: 01.07.2014].
- Schöpfel, Joachim; Prost, H  l  ne (2010). *D  veloppement et Usage des Archives Ouvertes en France Archives Ouvertes en France. Rapport. 1e partie partie: D  veloppement* [dok. elektr.]. [http://hal.archives-ouvertes.fr/docs/00/49/73/89/PDF/DUAO\\_Partie1.pdf](http://hal.archives-ouvertes.fr/docs/00/49/73/89/PDF/DUAO_Partie1.pdf) [odczyt: 01.07.2014].
- Schöpfel, Joachim; Prost, H  l  ne (2011). *Le Libre Acc  s en 2010 – Nice, Metz, Paris et ailleurs* [dok. elektr.]. Bulletin des biblioth  ques de France Vol. 56, No. 2, p. 85–88. [http://archivesic.ccsd.cnrs.fr/docs/00/74/48/39/PDF/AO\\_France\\_2010\\_5.1\\_Version\\_ArchiveSIC.pdf](http://archivesic.ccsd.cnrs.fr/docs/00/74/48/39/PDF/AO_France_2010_5.1_Version_ArchiveSIC.pdf) [odczyt: 01.07.2014].
- Schöpfel, Joachim; Prost, H  l  ne (2013). *Archives institutionnelles. Observations sur un nouveau mode d'information scientifique* [dok. elektr.]. Revue canadienne des sciences de l'information et de biblioth  conomie Vol. 37, No. 2, p. 122-136. [http://muse.jhu.edu/login?auth=0&type=summary&url=/journals/canadian\\_journal\\_of\\_information\\_and\\_library\\_science/v037/37.2.schopfel.html](http://muse.jhu.edu/login?auth=0&type=summary&url=/journals/canadian_journal_of_information_and_library_science/v037/37.2.schopfel.html) [odczyt: 01.07.2014].
- Schöpfel, Joachim; Stock, Catherine (2009a). *Grey literature in French digital repositories: a survey* [dok. elektr.]. The Grey Journal Vol. 5, No. 3, p. 147-161. [http://archivesic.ccsd.cnrs.fr/docs/00/37/92/32/PDF/GL10\\_Schopfel\\_Stock\\_final.pdf](http://archivesic.ccsd.cnrs.fr/docs/00/37/92/32/PDF/GL10_Schopfel_Stock_final.pdf) [odczyt: 01.07.2014].
- Schöpfel, Joachim; Stock, Catherine (2009b). *Les archives ouvertes en France – Un potentiel documentaire pour la formation    distance* [dok. elektr.]. Distances et Savoirs Vol. 7, No. 4, p. 443-456. [http://hal.inria.fr/docs/00/48/15/97/PDF/Schopfel\\_Stock\\_2009.pdf](http://hal.inria.fr/docs/00/48/15/97/PDF/Schopfel_Stock_2009.pdf) [odczyt: 01.07.2014].
- Smith, Kathlin (2008). *Institutional Repositories and E-Journal Archiving: What Are We Learning?* [dok. elektr.]. Journal of Electronic Publishing Vol. 11, No. 1. <http://quod.lib.umich.edu/jjep/3336451.0011.107?view=text;rgn=main> [odczyt: 01.07.2014].
- Strategie 2012-2015. Biblioth  que scientifique num  rique* (2012) [dok. elektr.]. <http://www.bibliothequescientifiquenumerique.fr/IMG/pdf/BSN-strategie-2012-2015.pdf> [odczyt: 01.07.2014].
- Suber, Peter (2009). *Timeline of the Open Access Movement* [dok. elektr.]. <http://legacy.earlham.edu/~peters/fos/timeline.htm> [odczyt: 01.07.2014].
- Swan, Alma et al. (2005). *Developing a model for e-prints and open access journal content in UK further and higher education* [dok. elektr.]. Learned Publishing Vol. 18, No. 1, p. 25-40. <http://eprints.ecs.soton.ac.uk/11000/> [odczyt: 01.07.2014].
- Synth  se sur les r  sultats de l'enqu  te Archives ouvertes Couperin-ADBU-INRA-INRIA 2012* (2012) [dok. elektr.]. <http://adbu.fr/wp-content/uploads/2012/09/Archives-ouvertes-2012-final.doc> [odczyt: 01.07.2014].
- Szczepańska, Anna (2007). *Podstawowe strategie wyszukiwania informacji i ich wykorzystywanie w praktyce*. Przegląd Biblioteczny z. 2. [http://eprints.rclis.org/15040/1/PODSTAWOWE\\_STRATEGIE\\_WYSZUKIWANIA\\_INFORMACJI.pdf](http://eprints.rclis.org/15040/1/PODSTAWOWE_STRATEGIE_WYSZUKIWANIA_INFORMACJI.pdf) [odczyt: 01.07.2014].
- Szprot, Jakub red. (2014). *Otwarta nauka w Polsce 2014. Diagnoza* (dok. elektr.). <http://pon.edu.pl/index.php/nasze-publicacje?pubid=13> [odczyt: 05.11.2014].
- Van der Graaf, Maurits; Van Eijndhoven, Kwame (2008). *The European repository landscape: Inventory Study into the Present Type and Level of OAI-Compliant Digital Repository Activities in the EU* [dok. elektr.]. [http://scholar.google.com/citations?view\\_op=view\\_citation&hl=en&user=V7z81CYAAAAJ&citation\\_for\\_view=V7z81CYAAAAJ:u5H-HmVD\\_uO8C](http://scholar.google.com/citations?view_op=view_citation&hl=en&user=V7z81CYAAAAJ&citation_for_view=V7z81CYAAAAJ:u5H-HmVD_uO8C) [odczyt: 01.07.2014].

- Vathonne, Aurélie (2010). *Openaire.eu: les résultats de la recherche scientifique en Europe*. Bases No. 277, p. 8-9.
- Wojciechowska, Anna (2008). *Archives ouvertes: l'état des lieux et pratiques dans les domaines des mathématiques et de l'informatique* [dok. elektr.]. <http://tel.archives-ouvertes.fr/docs/00/39/20/81/PDF/these-AW.pdf> [odczyt: 01.07.2014].
- Zespół ICM pod kier. Marka Niezgódki (2011). *Wdrożenie i promocja otwartego dostępu do treści naukowych i edukacyjnych. Praktyki światowe a specyfika polska. Przewidywane koszty, narzędzia, zalety i wady* [dok. elektr.]. [https://www.academia.edu/1374779/Wdrozenie\\_i\\_promocja\\_otwartego\\_dostepu\\_do\\_tresci\\_naukowych\\_i\\_educacyjnych.\\_Praktyki\\_swiatowe\\_a\\_specyfika\\_polska.\\_Przewidywane\\_koszty\\_narzedzia\\_zalety\\_i\\_wady](https://www.academia.edu/1374779/Wdrozenie_i_promocja_otwartego_dostepu_do_tresci_naukowych_i_educacyjnych._Praktyki_swiatowe_a_specyfika_polska._Przewidywane_koszty_narzedzia_zalety_i_wady) [odczyt: 01.07.2014].

---

# Zasoby biblioteki cyfrowej a zasoby repozytorium instytucjonalnego w kontekście działania Jagiellońskiej Biblioteki Cyfrowej i Repozytorium UJ

Krystyna Sanetra, Łukasz Mesek

Biblioteka Jagiellońska

---

## 1. Biblioteka cyfrowa a repozytorium instytucjonalne

Według „Podręcznego słownika bibliotekarza” [Czapnik, Gruszka 2011, s. 37] biblioteka cyfrowa (biblioteka elektroniczna, biblioteka wirtualna) to platforma dostępu do opracowanego zbioru dokumentów cyfrowych – tekstu, grafiki, dźwięku, filmu itp. – przechowywanych i udostępnianych za pośrednictwem sieci komputerowych. Autorzy słownika za nazwę przyjętą uznali termin „biblioteka cyfrowa”, tworząc odsyłacz całkowity „zob.” od terminu „repozytorium cyfrowe”.

Zakres znaczeniowy terminu „biblioteka cyfrowa” oraz charakterystykę poszczególnych elementów wchodzących w skład tego typu systemów informacyjnych, szczegółowo, na podstawie piśmiennictwa polskiego i zagranicznego, w układzie chronologicznym i tematycznym, przedstawiła Małgorzata Janiak [Janiak 2012]. Autorka zaznaczyła, że termin ten ewoluował i zmieniał swój zakres semantyczny. Pojawiał się na zmianę z pojęciami „biblioteka elektroniczna”, „biblioteka wirtualna”, „biblioteka bioniczna”, „hiper-biblioteka”, które, w początkowej fazie tworzenia tego rodzaju zasobów, odnosiły się do różnych typów bibliotek. W fazie dalszego rozwoju systemów oraz konkretnych projektów, w piśmiennictwie angielskojęzycznym zaczęło dominować pojęcie *digital library*, a w polskim – „biblioteka cyfrowa” [Janiak 2012].

Termin „biblioteka cyfrowa” składa się z dwóch pojęć. Pierwsze wskazuje na wyraźną ciągłość roli i zadań biblioteki tradycyjnej, zachowanych w bibliotece cyfrowej [Nahotko 2004]. Biblioteka cyfrowa, podobnie jak tradycyjna, posiada funkcje umożliwiające gromadzenie zbiorów, opracowanie metadanych, udostęp-

nianie obiektów (w tym przypadku cyfrowych). Z kolei drugi człon nazwy odnosi się do aspektów technologicznych: przetwarzania informacji, tworzenia kolekcji, organizacji danych i ich prezentacji, wykorzystywania obiektów cyfrowych i usług elektronicznych – wszystko to w środowisku internetu, którego rozwój doprowadził do znaczących zmian w gromadzeniu i użytkowaniu informacji, a także jej udostępnianiu, niezależnemu od lokalizacji przechowywanych danych cyfrowych.

Na potrzeby niniejszego artykułu przyjęto definicję zawartą w dokumencie Komisji Europejskiej „i2010: Biblioteki Cyfrowe”, zgodnie z którą biblioteki cyfrowe są uporządkowanymi zbiorami zasobów cyfrowych udostępnianymi publicznie. Mogą się na nie składać obiekty powstałe w wyniku digitalizacji (cyfrowe kopie książek i innych materiałów analogowych), pochodzące z bibliotek i archiwów. Mogą one również być tworzone w oparciu o informacje, które powstały pierwotnie w formie elektronicznej. W celu rozszerzenia dostępu do informacji w bibliotekach cyfrowych wykorzystywany jest potencjał technologii informacyjnych w trzech priorytetowych obszarach: dostęp *online*, digitalizacja zbiorów analogowych, konserwacja i archiwizacja cyfrowych zasobów.

Precyzyjne zdefiniowanie terminu „repozytorium instytucjonalne” nie jest proste. W literaturze przedmiotu dość powszechnie występuje definicja Clifforda A. Lyncha, opisująca repozytorium instytucjonalne jako „(...) zestaw usług, które uczelnia oferuje członkom swojej społeczności w celu rozpowszechniania materiałów cyfrowych sporządzanych przez instytucję i jej członków oraz zarządzania tymi materiałami” [Lynch 2003, p. 2]. Definicję tę przytacza m.in. Susan Gibbson podkreślając, że jest ona szczególnie trafna, ponieważ akcent pada w niej na usługi, a nie na technologie. Uważa, że choć „zestaw usług” repozytorium instytucjonalnego musi być określony przez daną instytucję, to wszystkie repozytoria instytucjonalne mają pewne wspólne cechy i funkcje. Zalicza do nich: rejestrację zasobów tworzonych w formie cyfrowej, specyfikę społeczności instytucji, która podejmuje decyzje odnośnie do tego, co ma być deponowane i jest deponytariuszem zasobów, wsparcie organizacyjne i finansowe repozytorium przez instytucję, bieżącą archiwizację oraz stabilny i szeroki dostęp do zasobów [Gibbson 2004]. Jak stwierdza Remigiusz Sapa, nie ma jednak powszechnej zgody co do tego, jaki zestaw funkcji definiuje repozytorium. W podsumowaniu dokonanego przez siebie przeglądu piśmiennictwa autor wskazuje natomiast przykładowe funkcje repozytoriów uczelnianych, takie jak: archiwizowanie dorobku cyfrowego uczelni, wykorzystanie repozytorium uczelnianego jako narzędzia do formalnego publikowania oryginalnych prac, przechowywanie i udostępnianie niepublikowanych zasobów naukowych, edukacyjnych i związanych z funkcjonowaniem jednostki, umożliwienie alternatywnego, otwartego dostępu do treści dystrybuowanych w komercyjnym obiegu wydawniczym, promocję uczelni (zgromadzenie w jed-



nym miejscu cyfrowego dorobku pracowników, monitorowanie zainteresowania środowiska tym dorobkiem, poprawę jego dostępności w sieci), zarządzanie prawami do zasobów cyfrowych [Sapa 2013, s. 119-121].

Między biblioteką cyfrową a repozytorium instytucjonalnym brak jest wyraźnego rozróżnienia, zarówno w piśmiennictwie, jak i w praktyce. Często jeden system łączy funkcje obu źródeł informacji, jak np. Repozytorium Politechniki Krakowskiej, gromadzące zasoby naukowe i dydaktyczne, których autorami są pracownicy i studenci PK, a także cyfrowe wersje zbiorów Biblioteki PK (np. kopie cyfrowe 198 książek wydanych przed 1939 r.). W poniższej tabeli, na przykładzie Jagiellońskiej Biblioteki Cyfrowej (JBC) oraz Repozytorium Uniwersytetu Jagiellońskiego (RUJ), zestawiono najważniejsze cechy biblioteki cyfrowej i repozytorium instytucjonalnego.

Zważywszy, że zarówno biblioteka cyfrowa, jak i repozytorium instytucjonalne zawierają obiekty cyfrowe, obie platformy mogą być obsługiwane przez to samo

Tabela 1. Zestawienie najważniejszych cech Jagiellońskiej Biblioteki Cyfrowej i Repozytorium Uniwersytetu Jagiellońskiego

	JBC	RUJ
główny cel	udostępnianie obiektów cyfrowych wraz z metadanymi	
struktura zasobu	podział na kolekcje / fasety	
oprogramowanie	możliwe jest zastosowanie takiego samego oprogramowania	
	dLibra	DSPACE
udostępniane zasoby	zbiory biblioteczne w postaci cyfrowej: 1) analogowe poddane digitalizacji 2) oryginalnie / natywnie cyfrowe (ang. <i>born digital</i> )	bieżący dorobek naukowy, badawczy, dydaktyczny pracowników uczelni 1) materiały oficjalnie wydane 2) materiały nieopublikowane oficjalnie
uwierzytelnianie autorów	przez kartotekę haseł wzorcowych w NUKAT / KKZBUJ	przez zintegrowanie systemu autoryzacji repozytorium z systemem uwierzytelniania pracowników UJ (SAP)
metadane	konwertowane z KKZBUJ / NUKAT; uzupełniane, przydzielane do kolekcji	tworzone w repozytorium, integrowane z Bibliografią Publikacji Pracowników UJ (BPP UJ)
gromadzenie zasobów	bibliotekarze	autorzy jako depozytariusze (we współpracy z bibliotekarzami)
udostępnianie zasobów	zgodnie z prawem autorskim: 1) należące do domeny publicznej wszystkim użytkownikom internetu 2) w trybie Open Access 3) chronione prawem autorskim tylko z komputerów na terenie biblioteki / bibliotek	zgodnie z prawem autorskim: 1) wszystkim użytkownikom internetu 2) w trybie Open Access
archiwizacja	długoterminowa	
prowadzenie	bibliotekarze	

Źródło: opracowanie własne

oprogramowanie, a inicjatorem i realizatorem w obu przypadkach jest zwykle biblioteka. Różnica między biblioteką cyfrową i repozytorium wynikać może jedynie z różnych funkcji tych systemów informacyjno-wyszukiwawczych. Podział między nimi można więc określić jako funkcjonalny, wynikający z częściowo różnych funkcji, a dodatkowo z rodzaju gromadzonych zasobów. Przedmiotem dalszych rozważań będą właśnie gromadzone zasoby.

## **2. Analiza wybranych bibliotek cyfrowych i repozytoriów pod względem typów udostępnianych zasobów**

W poniższym rozdziale zostaną przedstawione przykłady bibliotek cyfrowych oraz repozytoriów, ze szczególnym uwzględnieniem typów dokumentów, jakie zostały w nich zgromadzone. Spośród bibliotek cyfrowych zostały wybrane największe i najpopularniejsze. Repozytoria cyfrowe to z reguły bazy ocenione najwyższej według rankingu Webometrics. Wydzielonych zostało także kilka bibliotek cyfrowych, które pełnią również rolę repozytoriów.

### **2.1. Biblioteki cyfrowe**

#### **2.1.1. Harvard University Library Digital Collections**

Biblioteka cyfrowa Uniwersytetu Harvarda w Cambridge zorganizowana jest w postaci linków prowadzących do kolekcji zdigitalizowanych dokumentów. Kolekcje te obejmują zbiory bibliotek, muzeów, archiwów, w tym także zbiory specjalne. Układ kolekcji jest tematyczny, niemniej niektóre z tematów wiążą się ściśle z formalnym podziałem dokumentów. Do najważniejszych kolekcji tematycznych należą: sztuka, architektura, religia, historia, kultura, botanika, biologia, muzyka, polityka, prawo. Podstawowe typy dokumentów to:

- obrazy, fotografie, grafika,
- dokumenty tekstowe,
- nagrania dźwiękowe,
- muzykalia,
- zbiory kartograficzne.

Ponadto w bibliotece Uniwersytetu Harvarda prowadzone są prace nad gromadzeniem i udostępnianiem informacji o zasobach internetu: odnośniki do interesujących, tematycznie związanych stron są włączane w obręb poszczególnych kolekcji. Każda kolekcja tematyczna ma swój specyficzny układ typów dokumentów, który zależy od jej charakterystyki i rejestrowanych zasobów. Na przykład

w kolekcji „Women Working 1800-1930” wyróżnione zostały następujące typy dokumentów:

- książki i broszury,
- dzienniki i wspomnienia,
- dokumenty urzędowe,
- kluczowe organizacje,
- czasopisma,
- rękopisy,
- fotografie,
- prospekty reklamowe.

### 2.1.2. Library of Congress Digital Library

Digitalizacja zasobów największej biblioteki świata została rozpoczęta już w latach 90. XX w. Różnice technologiczne występujące w procesie digitalizacji w tak długim okresie są widoczne podczas przeglądania dokumentów zgromadzonych w 9 kolekcjach tematycznych:

- American Memory – największa kolekcja, podzielona na 16 podkategorii tematycznych, które z kolei zawierają kolejne podkolekcje,
- Prints and Photographs – zawierająca materiały graficzne kolekcja obejmuje 70 kolekcji tematycznych,
- Historic Newspapers – kolekcja obejmuje prawie 1500 tytułów czasopism wydawanych w latach 1836-1922,
- Performing Arts – to rodzaj encyklopedii opierającej się na zdigitalizowanych materiałach źródłowych,
- Veterans History – kolekcja obejmująca materiały dotyczące wojen prowadzonych przez Stany Zjednoczone i żołnierzy w nich uczestniczących,
- Sound Recordings – zasoby dźwiękowe wydzielone z różnych kolekcji tematycznych,
- Film – kolekcja obejmuje 15 podkolekcji tematycznych,
- Maps – kolekcja obejmuje 16 podkolekcji tematycznych,
- Manuscripts – kolekcja obejmuje 27 podkolekcji tematycznych.

W przedstawionych kolekcjach znajdują się prawie wszystkie typy dokumentów, poczynając od ulotek, poprzez afisze teatralne, na filmach kończąc. Położenie nacisku w organizacji biblioteki cyfrowej na podział zasobów na kolekcje tematyczne jest coraz częściej spotykanym trendem. Wśród zasobów Biblioteki Kongresu kolekcje tematyczne mieszają się z podziałem według niektórych typów dokumentów, niemniej jednak jest to przykład przedstawienia treści, który wydaje się bardziej odpowiadający potrzebom użytkowników.

### 2.1.3. HathiTrust Digital Library

HathiTrust nie jest typową biblioteką cyfrową. Jak określają to sami twórcy [Guidelines... 2011], jest raczej szeroko zakrojonym projektem repozytorium do przechowywania obiektów cyfrowych, pochodzących z ponad 60 bibliotek uniwersyteckich. Gromadzi przeważnie kopie cyfrowe książek i czasopism, głównie z domeny publicznej, ale również cały szereg dokumentów publikowanych na licencji Creative Commons, wprowadzanych do akademickich bibliotek cyfrowych przez autorów związanych z poszczególnymi uniwersytetami. W bazie rejestrowane są także książki objęte prawem autorskim, ale w tym przypadku podawana jest jedynie informacja, która biblioteka posiada daną publikację w swoich zbiorach. HathiTrust jest swego rodzaju *backupem* – zabezpiecza treści już wcześniej zgromadzone i udostępnione w internecie – w którym zostały dodatkowo rozwinięte pewne narzędzia służące nawigacji w rozbudowanym zbiorze dokumentów. O jego wielkości świadczą poniższe zestawienia statystyczne: 11 481 506 woluminów, 5 922 738 tytułów książek, 298 349 tytułów czasopism, 4 018 527 100 stron, 515 terabajtów (dane z października 2014 r.).

### 2.1.4. Project Gutenberg

Projekt Gutenberg to najstarsza biblioteka cyfrowa (stworzona w 1971 r.), początkowo bazująca na ręcznie przepisywanych tekstach, intensywnie rozwijająca się po 1995 r., kiedy to na szerszą skalę zaczęto wykorzystywać OCR (Optical Character Recognition) [Kolasa 2012]. W ramach projektu zbudowano bibliotekę cyfrową gromadzącą prawie 47 tys. tytułów z zakresu literatury pięknej, dostępnych w domenie publicznej. Ciekawostką jest różnorodność formatów gromadzonych plików, odzwierciedlająca intencję twórców biblioteki, aby udostępniać zbiory na jak największym spektrum urządzeń. Zasoby nie są ograniczone jedynie do dokumentów tekstowych, można wśród nich odnaleźć także:

- książki audio z syntezatorem mowy,
- książki audio czytane przez lektora,
- kompilacje,
- surowe dane,
- nagrania muzyczne,
- muzykalnia,
- inne nagrania,
- obrazy ruchome,
- obrazy nieruchome.

Do podstawowych kategorii, według których uporządkowane zostały zgromadzone dokumenty, należą:

- autor – układ alfabetyczny,
- tytuł – układ alfabetyczny,
- język – podział na języki, w których opublikowano powyżej 50 i poniżej 50 dokumentów,
- ostatnio dodane – w ciągu: 24 godzin, 7 dni i 30 dni.

Dodatkowo została jeszcze udostępniona lista 100 najpopularniejszych dokumentów. Pierwszą z nich w ciągu jednego dnia pobrano ponad 1500 razy, co świadczy o dużej popularności całego projektu.

### 2.1.5. Universal Library – million book project

Universal Library to typowy projekt digitalizacyjny, zrealizowany przez Carnegie Mellon University we współpracy z instytucjami rządowymi i ośrodkami naukowymi z Indii oraz Chin. Projekt polegał na skanowaniu książek, opublikowanych w wielu językach, z zastosowaniem OCR oraz udostępnieniu wszystkich publikacji bez ograniczeń w internecie. Nie zostały tu wyróżnione żadne kolekcje tematyczne, a układ, według którego przegląda się zgromadzone dokumenty, jest bardzo prosty. Publikacje przypisane są do następujących kategorii:

- początek tytułu – układ alfabetyczny,
- nazwisko autora – układ alfabetyczny,
- zakresy lat – (ostatni zakres to lata 2001-2007),
- dziedziny nauki – 13 dziedzin,
- języki – 12 języków.

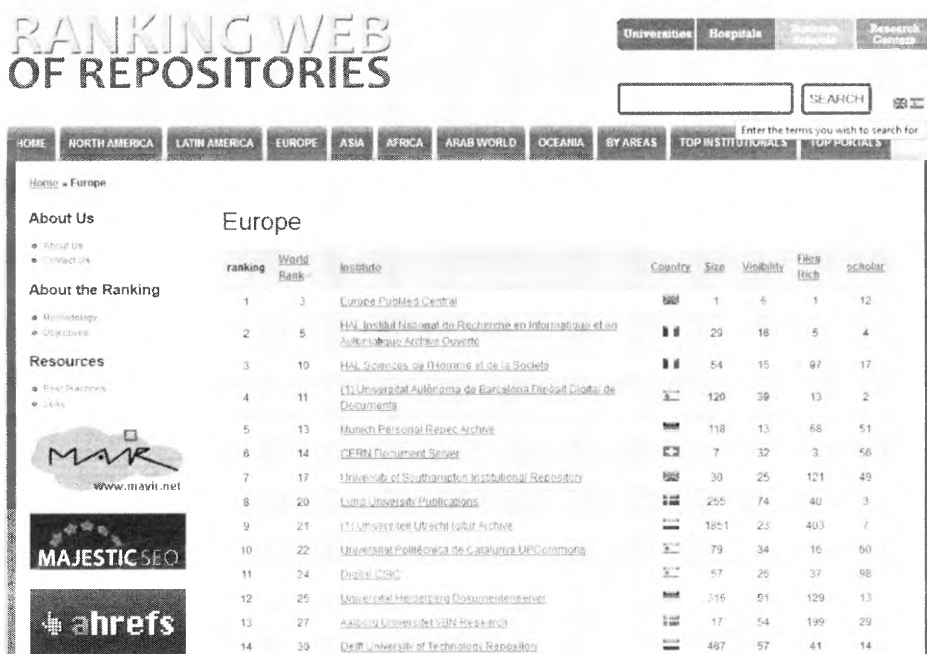
Należy zaznaczyć, że przeglądanie zasobów tej biblioteki jest bardzo niewygodne, a dostarczane wyniki często nie są relewantne, np. w przypadku książek w języku angielskim przedimek nieokreślony jest brany pod uwagę jako początek tytułu, co znacznie utrudnia przeglądanie alfabetyczne.

## 2.2. Repozytoria

Poza niektórymi wyjątkami podstawą doboru repozytoriów przedstawionych w tej części rozdziału jest ranking szpitali, uniwersytetów, ośrodków badawczych i repozytoriów światowych – „The Ranking Web of World Hospitals, Universities, Research Centers and Repositories”. Jest to jedna z głównych baz tego rodzaju, stworzona przez Cybermetrics Lab – ośrodek badawczy należący do Najwyższej Rady Badań Naukowych (hiszp. Consejo Superior de Investigaciones Científicas – CSIC) z siedzibą w Madrycie. Cybermetrics Lab zajmuje się ilościową analizą zasobów internetu (badania infometryczne), w szczególności udostępnionych w sieci publikacji akademickich i naukowych.

W omawianym rankingu (określanym w dalszej części artykułu jako „Webometrics”) znajdują się głównie repozytoria ośrodków naukowych, zestawione według złożonych indeksów określających obecność w internecie (ang. *web presence*) i widoczność linków (ang. *web impact*) odsyłających do zawartości tych repozytoriów. Dane te pochodzą z największych wyszukiwarek komercyjnych.

Przechodząc do omówienia poszczególnych repozytoriów, z reguły sklasyfikowanych najwyżej w poszczególnych regionach świata, przedstawiono na ilustracji 1 widok rankingu wraz z poszczególnymi wskaźnikami. Warto nadmienić, że istnieje możliwość sortowania według każdego z tych wskaźników, co pozwala na zawężenie oceny repozytoriów do jednego z wybranych kryteriów.



Ilustracja 1. Ranking repozytoriów europejskich

Źródło: [Ranking Web of Repositories. Europe, dok. elektr., stan na dzień 5.09.2014]

### 2.2.1. Dipòsit digital de Documents de la Universitat Autònoma de Barcelona

Repozytorium Autonomicznego Uniwersytetu w Barcelonie w rankingu Webometrics zajmuje 4 pozycję w Europie i 11 na świecie. Nie wyróżnia się szczególnie rozmiarem ani widocznością w internecie, a kryterium klasyfikującym je tak wysoko jest „naukowość”, oceniana na podstawie liczby odsyłaczy do tej bazy w Google Scholar.

Podstawowy podział dokumentów w repozytorium wyznaczony jest przez ich typ, choć klasyfikacja nie jest przeprowadzona ściśle i konsekwentnie według metodologii podziału na typy formalne dokumentów. Wprowadzony został podział odzwierciedlający raczej potrzeby użytkownika niż bibliotekarza. Przedstawia się on następująco:

- materiały dla studentów (programy, egzaminy, przewodniki, materiały do nauki),
- książki i kolekcje (zarówno całe książki, jak i wybrane rozdziały, dokumenty instytutów, przewodniki biblioteczne),
- literatura naukowa (m.in. patenty, rozprawy doktorskie, pre-printy),
- publikacje periodyczne (raporty roczne, czasopisma, newslettery),
- artykuły (opublikowane),
- materiały przygotowywane na konferencje i kongresy,
- dokumenty graficzne (plakaty, rysunki, pocztówki, mapy),
- dokumenty multimedialne (audio, wideo oraz wystawy biblioteczne),
- archiwa osobiste i instytucjonalne.

W repozytorium zgromadzono ok. 110 tys. dokumentów, z czego ponad 1/4 stanowią opublikowane artykuły. Najwięcej jest jednak materiałów dla studentów, co świadczy o nastawieniu twórców platformy na tę właśnie grupę odbiorców. Jest to niewątpliwie repozytorium instytucjonalne, którego zasoby stanowią publikacje i materiały pracowników i studentów uniwersytetu.

## 2.2.2. University of California eScholarship Repository

Repozytorium Uniwersytetu w Kalifornii jest dobrym przykładem rozwinięcia biblioteki cyfrowej i wzbogacenia jej o publikacje pracowników uniwersytetu. W rankingu Webometrics eScholarship zostało sklasyfikowane na 3 miejscu w Ameryce Północnej i na 6 miejscu wśród repozytoriów światowych.

W informacji na temat eScholarship twórcy przedstawiają repozytorium jako platformę dedykowaną zarówno instytutom i organizacjom naukowym, jak i osobom związanym z uniwersytetem, służącą do zarządzania ich dorobkiem naukowym [About... 2014]. eScholarship nie jest dużym repozytorium, zawiera tylko ok. 75 tys. publikacji. Jest jednak o tyle warte uwzględnienia, że bardzo aktywnie wspiera politykę Open Access poprzez bezpośredni kontakt z naukowcami i realizowanie sugerowanych przez nich udogodnień, a szczególnie poprzez przejrzyste zarządzanie prawami autorskimi.

Zasoby repozytorium można przeglądać za pomocą katalogu lub przeszukiwać, używając w tym celu okna wyszukiwarki. Dokumenty, podobnie jak w innych repozytoriach, są zaklasyfikowane do heterogenicznych kategorii, odpo-

wiadających raczej potrzebom użytkowników niż systemom klasyfikacyjnym. Zastosowana kategoryzacja obejmuje:

- dokumenty ostatnio dodane,
- dokumenty najbardziej popularne,
- dokumenty publikowane przez wydziały uniwersytetu,
- dokumenty publikowane przez inne jednostki uniwersytetu,
- czasopisma,
- książki,
- prace doktorskie i inne prace na stopień,
- prace już wcześniej opublikowane,
- autorów,
- daty publikacji.

Istnieje możliwość zawężenia otrzymanych wyników wyszukiwania według trzech kryteriów:

- artykuły recenzowane,
- dyscypliny naukowe,
- wydziały uniwersytetu.

Po otrzymaniu listy wyników wyszukiwania można podejrzeć abstrakt dokumentu oraz listę dokumentów podobnych do wybranej publikacji. Interesującą opcją jest podręczna statystyka każdego dokumentu, informująca o liczbie wyszukiwań i pobrań oraz pokazująca ich miesięczną średnią. eScholarship jest przykładem repozytorium instytucjonalnego tworzonych przez uniwersytet, w ramach którego funkcjonuje kilkanaście oddziałów w całej Kalifornii.

### **2.2.3. Universidade de São Paulo Biblioteca Digital de Teses e Dissertações**

Chociaż w nazwie pojawia się określenie „biblioteka cyfrowa”, to jednak platformę tę, ze względu na jej zawartość, można sklasyfikować jako typowe repozytorium instytucjonalne.

Repozytorium Uniwersytetu w São Paulo zostało sklasyfikowane jako pierwsze na terenie Ameryki Łacińskiej i jako 12 w rankingu światowym. Warto zwrócić uwagę, że w repozytorium zdeponowane są tylko prace magisterskie oraz rozprawy doktorskie i habilitacyjne, a ich liczba wynosi prawie 48 tys. Platforma repozytorium oparta jest na oprogramowaniu DSpace, zmodyfikowanym i przystosowanym do potrzeb publikacji tego jednego typu dokumentów. W pierwszej kolejności zwraca uwagę brak okna wyszukiwawczego oraz tak charakterystycznych dla podstawowej wersji DSpace kolekcji, co zrównuje w zasadzie wyszukiwanie z przeglądaniem. Aby użyć którejs z tych usług, trzeba kliknąć w menu po lewej stronie ekranu. Wyszukiwanie proste realizowane jest za pomocą zaimple-



mentowanej wyszukiwarki Google, natomiast w wyszukiwaniu zaawansowanym można filtrować wyniki według następujących kryteriów:

- autor,
- tytuł,
- słowa w abstrakcie,
- słowa kluczowe,
- promotor,
- jednostka uniwersytetu,
- rok obrony.

Podczas przeglądania zasobów można również zawęzić listę otrzymanych wyników według podobnych kryteriów jak te użyte w filtrowaniu podczas wyszukiwania zaawansowanego:

- roku obrony,
- dziedziny wiedzy – baza obejmuje 618 dziedzin,
- autora – w kolejności alfabetycznej,
- typu dokumentu – praca magisterska, rozprawa doktorska, rozprawa habilitacyjna,
- promotora – w kolejności alfabetycznej,
- jednostki uniwersytetu – w kolejności alfabetycznej.

Według danych statystycznych w repozytorium najczęściej jest prac magisterskich – prawie 60%, prac doktorskich jest niespełna 40%, a rozpraw habilitacyjnych – tylko 0,6%. Z przedstawionych zasad publikowania wynika, że deponowanie prac i rozpraw jest obowiązkowe, a związane z nim reguły, dotyczące m.in. formatów plików (pdf) i wymaganych metadanych, są ściśle określone. Dane statystyczne dotyczące repozytorium są przedstawiane bardzo precyzyjnie. Obejmują zarówno informacje o przeprowadzanych sesjach wyszukiwawczych, jak i liczbę pobrań plików (z rozróżnieniem tych dokonywanych przez użytkowników i przez roboty), a także całkowity rozmiar zdeponowanych plików oraz dane z Google Analytics, dotyczące zachowań użytkowników, używanych przez nich urządzeń i oprogramowania.

#### **2.2.4. University of Hong Kong Scholars Hub**

Repozytorium Uniwersytetu w Hong Kongu zostało sklasyfikowane na 2 miejscu wśród repozytoriów azjatyckich i na 68 miejscu w rankingu światowym.

W repozytorium gromadzone są: artykuły, książki oraz fragmenty książek, materiały z konferencji, czasopisma, prace na stopień, dokumentacja grantów, projekty studenckie oraz strony naukowców. W sumie zdeponowanych jest ponad 167 tys. obiektów, z czego najczęściej artykułów, materiałów konferencyjnych oraz prac na stopień.

Repozytorium oparte jest na platformie DSpace-CRIS (Current Research Information System), będącej nowym, dodatkowym, otwartym modułem DSpace. Moduł ten rozszerza funkcje podstawowej wersji DSpace, dając możliwość zarządzania, kolekcjonowania i udostępniania danych dotyczących różnych aspektów procesu naukowego, związanych m.in.: z naukowcami, organizacjami, jednostkami, nagrodami, grantami.

Dostęp do dodatkowych możliwości widoczny jest już z poziomu strony głównej repozytorium, a do uwzględnionych wśród nich kategorii należą:

- badania (zarówno zapiski laboratoryjne jak i raporty z badań),
- publikacje,
- dysertacje, prace dyplomowe,
- granty,
- patenty,
- usługi dla społeczności uniwersytetu.

Każda z wymienionych kategorii ma dostosowany do niej panel wyszukiwawczy, zawierający z jednej strony okno zaawansowanego wyszukiwania, a z drugiej – podkategorie, według których można przeglądać zgromadzone w danej kategorii dokumenty. Dla przykładu publikacje można przeglądać według:

- typów (wymienionych powyżej),
- kolekcji (do których należą dokumenty poszczególnych wydziałów uniwersytetu),
- tytułów (uszeregowanych alfabetycznie),
- autorów (uszeregowanych alfabetycznie).

Bardzo funkcjonalny wydaje się podział, według którego można przeglądać publikacje w kategorii „granty”:

- sponsor / instytucja przyznająca,
- dziedzina,
- typ – alfabetyczna lista grantów, w ramach których powstały publikacje zdeponowane w repozytorium.

Repozytorium Uniwersytetu w Hong Kongu dzięki zastosowaniu dodatkowego modułu DSpace zyskuje bardzo na funkcjonalności, potrzebnej w codziennej naukowej działalności uniwersytetu. Niestety publicznie dostępne statystyki są bardzo ubogie, ograniczają się jedynie do ogólnych informacji na temat liczby publikacji i ich pobrań w poszczególnych latach.

### 2.2.5. University of Pretoria Institutional Repository

Repozytorium Uniwersytetu w Pretorii zostało sklasyfikowane na 1 miejscu wśród repozytoriów afrykańskich i na 99 miejscu w rankingu światowym.

W repozytorium zdeponowane są głównie: artykuły, materiały z konferencji oraz prace dyplomowe i dysertacje – w sumie ok. 32 tys. obiektów.

Repozytorium posadowione jest na platformie DSpace, przy czym jest to raczej podstawowa wersja, w której zastosowano domyślny schemat podziału na kategorie wyszukiwawcze. W podstawowym podziale na kolekcje nie zostały wyróżnione dokumenty publikowane przez poszczególne wydziały i jednostki pozawydziałowe uniwersytetu oraz związane z różnymi dziedzinami nauki.

Przeoglądanie repozytorium jest ograniczone do kategorii:

- zbiorów i kolekcji wraz z podkolekcjami,
- dat (podanie daty oraz sortowanie według dat są dodane jako opcje zawężające),
- autorów – w kolejności alfabetycznej,
- tytułów – w kolejności alfabetycznej,
- tematów – określonych przez słowa kluczowe, których jest w bazie ponad 100 tys.,
- promotorów – w kolejności alfabetycznej.

Dodatkowo wprowadzone zostało narzędzie „Discover”, pozwalające na zawężanie wyników wyszukiwawczych poprzez wybór:

- autora,
- tematu (w rzeczywistości słowa kluczowego),
- daty.

Narzędzie „Discover” to jedna z funkcji, która została dodana do wersji 3.0 platformy DSpace, ale dotychczas uruchomiona została tylko w kilku repozytoriach. Jest to narzędzie szczególnie przydatne w zawężaniu przeglądania ograniczonego do poszczególnych dziedzin nauki, pozwalające m.in. na wyselekcjonowanie autorów, np. według dat publikacji ich tekstów, zajmujących się konkretnym tematem. Możliwe są także inne kombinacje w obrębie zastosowanych filtrów, których dobór zależy od twórców repozytorium.

### 2.2.6. King Saud University Repository

Repozytorium Uniwersytetu King Saud w Arabii Saudyjskiej zostało sklasyfikowane na 1 miejscu wśród repozytoriów krajów arabskich i na 569 miejscu w rankingu światowym.

Repozytorium gromadzi nie tylko dokumenty uniwersytetu, użycza również przestrzeni innym instytucjom i organizacjom naukowym. Rozwinęło się na bazie biblioteki cyfrowej, której zbiory, obejmujące dysertacje, rękopisy oraz rzadkie książki, wciąż są dostępne. Kolekcja dokumentów uniwersytetu podzielona jest na podkolekcje obejmujące poszczególne wydziały oraz wydawnictwo uniwersyteckie. W repozytorium znajduje się ok. 14 tys. obiektów.

Repozytorium posadowione jest na platformie DSpace, a w przyjętej strukturze danych nie zostały poczynione żadne modyfikacje. Dostępne są jedynie podstawowe sposoby przeglądania zawartości repozytorium poprzez:

- zbiory i kolekcje,
- daty wydania,
- autorów – w porządku alfabetycznym, przy czym można zauważyć pewną nieścisłość, gdyż do powtarzających się nazwisk autorów nie dodano żadnych dopowiedzeń rozróżniających,
- tytuły – większość w języku angielskim, w porządku alfabetycznym,
- tematy – a raczej, wnosząc na podstawie ich liczby (ponad 18 tys.), słowa kluczowe, w znakomitej większości w języku arabskim.

Ze względu na specyfikę języka repozytorium ma ograniczony zasięg. Dotyczy to zarówno udostępnianych metadanych (tylko tytuły i nazwiska autorów są podawane w języku angielskim) – a więc także możliwości wyszukiwania w serwisach światowych – jak i umieszczonych w repozytorium dokumentów, z których większość opublikowano w języku arabskim.

### 2.2.7. Queensland University of Technology Institutional Repository

Repozytorium Politechniki w Brisbane zostało sklasyfikowane na 1 miejscu wśród repozytoriów Oceanii i na 16 miejscu w rankingu światowym.

W repozytorium znajduje się ponad 55 tys. obiektów, do których należą: artykuły, książki i rozdziały książek, materiały konferencyjne, prace dyplomowe i dysertacje, raporty, recenzje, patenty, materiały niepublikowane.

Repozytorium posadowione jest na platformie ePrints, która, w porównaniu do platformy DSpace, wyróżnia się jeszcze bardziej uproszczonym interfejsem. Nie występuje tu podział na kolekcje i podkolekcje, a jedyną dostępną kategorią dokumentów jest mocno rozbudowana klasyfikacja przedmiotowa. Pozostałe dwa kryteria przeglądania zawartości repozytorium to alfabetyczna lista nazwisk autorów oraz lata publikacji. Przy takiej strukturze niezmiernie ważne jest odpowiednie opisanie poszczególnych publikacji. Co więcej, aby wyszukiwanie było skuteczne, każda z publikacji musi mieć taki sam zestaw metadanych.

Na przykładzie tego repozytorium wskazać można różnice w podejściu do budowania poszczególnych platform. W oprogramowaniu ePrints zdecydowanie położono nacisk na wyszukiwanie, a przeglądanie jest tylko funkcją dopełniającą. Najlepiej widać tu także różnorodność metadanych, traktowanych jako kryteria filtrowania w wyszukiwaniu zaawansowanym. Bardzo dobrze prezentowane są również statystyki, poczynając od ogólnych, dotyczących całego repozytorium,

poprzez statystyki dla poszczególnych jednostek uniwersytetu, a kończąc na statystykach poszczególnych autorów czy wręcz tytułów artykułów.

### **2.2.8. Repozytorium Uniwersytetu im. Adama Mickiewicza (AMUR)**

Z repozytoriów polskich w rankingu Webometrics najwyżej ocenione zostało repozytorium uniwersyteckie w Poznaniu, które w światowym rankingu zajmuje 226 miejsce. Zdeponowano w nim ponad 10 tys. dokumentów. Platformą, na której zostało oparte, jest DSpace, przy czym jest to wersja podstawowa, bez dodatkowych modyfikacji. Dostępne są również jedynie podstawowe sposoby przeglądania zawartości repozytorium poprzez:

- zbiory i kolekcje,
- datę wydania,
- autorów – w porządku alfabetycznym,
- tytuły – w porządku alfabetycznym,
- słowa kluczowe – w porządku alfabetycznym.

Podstawowe kolekcje odzwierciedlają strukturę organizacyjną uniwersytetu, dodatkowo obejmują także zbiór czasopism naukowych i prac doktorskich. W obrębie poszczególnych wydziałów, jako podkolekcje zostały wyróżnione:

- artykuły naukowe,
- doktoraty,
- książki / rozdziały,
- materiały dydaktyczne,
- materiały konferencyjne,
- materiały niepublikowane.

Jest to zestawienie typów dokumentów replikowane dla wszystkich kolekcji, co ułatwia zarządzanie całym zbiorem.

### **2.2.9. Repozytorium Politechniki Krakowskiej**

Drugim najlepiej ocenionym w Polsce jest Repozytorium Politechniki Krakowskiej, które w światowym rankingu zajmuje 327 miejsce. Zawartość repozytorium stanowią: artykuły z czasopism, książki, rozdziały lub fragmenty książek, wyniki badań i raporty naukowe, sprawozdania naukowe, prace dyplomowe, materiały i prezentacje konferencyjne, materiały do zajęć dydaktycznych oraz materiały dotyczące jednostek organizacyjnych Politechniki Krakowskiej. W zbiorach repozytorium zdeponowanych jest ponad 6 tys. obiektów.

Repozytorium działa na platformie opracowanej przez informatyków i bibliotekarzy Politechniki Krakowskiej, stąd jego niepowtarzalny charakter, jego funk-

cjonalność nie odbiega przy tym znacznie od repozytoriów opartych na platformie DSpace czy ePrints.

Warto zwrócić uwagę, że zasoby repozytorium stanowią tylko część większego zbioru baz danych, możliwych do jednoczesnego przeszukiwania. Do tych zintegrowanych na potrzeby wyszukiwania zasobów należą następujące kolekcje:

- Repozytorium Politechniki Krakowskiej,
- Bibliografia Publikacji Pracowników Politechniki Krakowskiej,
- Baza Przedmiotów,
- inne bazy bibliograficzne.

Typy zasobów, jakie zostały wyróżnione do wyszukiwania we wszystkich bazach danych, są rozszerzone w stosunku do typów dokumentów wyszczególnionych w zasobach repozytorium. Należą do nich:

- książka,
- fragment książki,
- rozprawa habilitacyjna,
- rękopis,
- podręcznik, skrypt,
- artykuł,
- pre-print,
- praca magisterska,
- rozprawa doktorska,
- prezentacja.

Dodatkowo zawężenie wyszukiwania jest możliwe jeszcze poprzez określenie wydziału Politechniki, formatu pliku oraz języka. Natomiast przeglądanie zasobów możliwe jest według następujących kryteriów:

- autorów – szeregowanie alfabetyczne,
- kolekcji – wymienionych powyżej,
- wydziałów Politechniki,
- ostatnio opublikowanych – z podziałem na wyżej wymienione kolekcje.

Nadanie temu zespołowi baz danych nazwy „repozytorium” może być mylące, jednak funkcjonalność takiego rozwiązania jest niezaprzeczalna. Zintegrowane przeszukiwanie całego dorobku publikacyjnego danej jednostki w sytuacji, w której część zbiorów stanowią jedynie informacje bibliograficzne, przynosi zdecydowanie więcej korzyści niż ograniczenie tylko do pełnych tekstów. Dotyczy to szczególnie tych repozytoriów, w których nie ma obowiązku deponowania dorobku pracowników uczelni.

## 2.3. Biblioteki cyfrowe i repozytoria

### 2.3.1. CERN Document Server

Platforma CERN-u to typowy przykład ewolucji biblioteki cyfrowej w stronę repozytorium, choć bez zatarcia granic pomiędzy obydwoma jednostkami. Ścisły podział został zachowany, co widać na stronie głównej, gdzie w lewej kolumnie znajdują się materiały właściwe dla biblioteki cyfrowej, zaś w prawej – dokumenty instytucjonalne. Platforma CERN-u zajmuje 6 miejsce w rankingu repozytoriów europejskich i 14 miejsce wśród światowych. Wpływ na tę ocenę mają głównie wielkość i liczba pełnych plików, a więc plików tekstowych zawierających treść naukową lub popularyzatorską.

Repozytorium CERN-u należy do największych repozytoriów na świecie, udostępnia prawie 1 400 tys. dokumentów. Jak już wspomniano wyżej, w strukturze platformy dokumenty te zostały wyraźnie rozdzielone na zasoby biblioteczne oraz te, za których publikację CERN odpowiada, będące najczęściej wynikiem prac związanych z nim naukowców. Do pierwszej grupy należą:

- artykuły i pre-printy,
- książki i materiały pokonferencyjne,
- prezentacje i przemówienia,
- publikacje periodyczne,
- multimedia.

Dokumenty należące do publikacji wydawanych przez CERN zostały pogrupowane w kategorii częściowo odpowiadające kategoriom w pierwszej grupie. Są to:

- artykuły i pre-printy,
- serie CERN,
- jednostki CERN,
- publikacje dotyczące akceleratorów CERN,
- eksperymenty CERN.

### 2.3.2. Virginia Tech University Digital Library and Archives

Biblioteka Cyfrowa i Archiwa Politechniki Stanowej Wirginii to kolejny przykład platformy łączącej zasoby biblioteki cyfrowej i repozytorium. Według Webometrics została ona sklasyfikowana na 7 miejscu w Ameryce Północnej i na 15 miejscu w rankingu światowym. Jedną ze składowych części Digital Library and Archives jest VtechWorks, będące repozytorium instytucjonalnym. VtechWorks posiadone jest na platformie DSpace, a jego struktura obejmuje kolekcje, któ-

re – jak w większości repozytoriów opartych na tej platformie – odzwierciedlają strukturę organizacyjną uniwersytetu.

Natomiast pewne kolekcje, np. czasopisma naukowe, nie zostały uwzględnione w repozytorium instytucjonalnym, tylko włączone, jako odrębna usługa, do Digital Library and Archives. W ten sposób udostępnianych jest również kilka kolekcji tematycznych, co świadczy o stopniowym przekształcaniu biblioteki cyfrowej w większą platformę obejmującą publikacje naukowe, czasopisma, repozytorium instytucjonalne i rejestrującą wydarzenia odbywające się na uniwersytecie. W całym Digital Library and Archives w 2012 r. zgromadzonych było ponad 628 tys. plików, co stanowi dużą różnicę w stosunku do 40 tys. dokumentów w Vtech-Works [DLA... 2012]. Stąd zapewne wysoka ocena w rankingu Webometrics, obejmująca Digital Library and Archives jako całość, a nie tylko repozytorium instytucjonalne.

The screenshot displays the CNKI website interface. At the top, there is a search bar with the text 'Full-Text Search:' and a 'CNKI Search' button. Below the search bar is a navigation menu with links: Home, Journal Papers, About CNKI, User Service, FAQ, Contact Us, and 中文. The main content area is divided into several sections, each with a category title and a grid of search results. The categories shown are: Science-Engineering (A), Literature/History/Philosophy, Science-Engineering (B), Politics/Military Affairs/Law, Science-Engineering (C), and Education & Social Sciences. Each section contains a grid with columns for 'Recent Updated Journal' and 'Recent Updated Papers', listing various academic titles and links.

Ilustracja 2. Strona główna Zintegrowanej Bazy Danych Zasobów Wiedzy Chin  
Źródło: [Tsinghua Tongfang Knowledge Network Technology Co. 2006, dok. elektr.]



### 2.3.3. China Knowledge Resource Integrated Database

Zintegrowana Baza Danych Zasobów Wiedzy Chin nie została uwzględniona w rankingu Webometrics, ale jest to na tyle ciekawa inicjatywa, że warto ją w tym rozdziale uwzględnić. Jest czymś więcej niż tylko repozytorium, choć to ostatnie stanowi jej istotną część składową. Baza ma układ tematyczny, od ogółu do szczegółu, według 10 grup dziedzinowych obejmujących m.in. „Science Engineering” czy „Literature / History / Philosophy”. Na stronie głównej, w dwóch kolumnach obejmujących poszczególne grupy dziedzin, z lewej strony przedstawione są informacje o dziesięciu aktualizacjach w poszczególnych czasopismach, a z prawej wyświetla się dziesięć ostatnio opublikowanych artykułów. Ilustracja 2 przedstawia układ bazy.

### 2.4. Podsumowanie analizy

Wybór przedstawionych baz i zasobów poddanych analizie odzwierciedla ich różnorodność i pokazuje, że nawet przy zastosowaniu tego samego oprogramowania struktura organizacji kolekcji bywa zupełnie odmienna. W przeglądzie tym chodziło właśnie o odzwierciedlenie tej różnorodności, ponieważ analiza poszczególnych przypadków miała znaczący wpływ na ustalenia dotyczące typów gromadzonych dokumentów oraz ich organizacji w Jagiellońskiej Bibliotece Cyfrowej i w Repozytorium Uniwersytetu Jagiellońskiego. Analizy nie ograniczono tylko do baz tworzonych przez pojedyncze jednostki uniwersyteckie, ale podjęto próbę szerszego spojrzenia na duże projekty, zakładające gromadzenie zarówno obiektów zdigitalizowanych, jak i *born digital*. Co więcej, nie poprzestano jedynie na przedstawieniu typów dokumentów, ale w przypadkach, w których wydawało się to uzasadnione, pokazano również kryteria wyszukiwawcze lub kolekcje tematyczne. Często zdarza się bowiem, że twórcy odpowiedzialni za organizację zasobów celowo przenoszą punkt ciężkości np. na kolekcje tematyczne, pomijając zupełnie typologię dokumentów. Przedstawienie dużych jednostek repozytoryjnych, tworzonych często przez wiele instytucji, z których zasobów korzystają miliony odbiorców, ukazuje, jakie trendy w udostępnianiu informacji, w kreowaniu sposobów dostępu do typów dokumentów i grupowaniu ich w kolekcje oraz wyświetlaniu pojedynczych dokumentów są obecnie propagowane. Zderzenie w rzeczywistości wirtualnej rozwiązań zastosowanych w tych projektach z rozwiązaniami stosowanymi w komercyjnych wyszukiwarkach może posłużyć głębszej analizie realizowania potrzeb informacyjnych użytkowników w narastającym natłoku informacji.

### **3. Zasoby cyfrowe publikowane w bibliotekach cyfrowych i repozytoriach**

Zasoby biblioteki cyfrowej tworzone są zasadniczo na bazie zbiorów drukowanych / analogowych biblioteki, opublikowanych uprzednio przez różnych wydawców. Polityka selekcji materiałów do digitalizacji oparta jest zwykle jednocześnie na kilku kryteriach, takich jak wartość dokumentu z punktu widzenia dziedzictwa kulturowego i historii rozwoju nauki, potrzeba bezpiecznego i nowoczesnego udostępnienia szerszemu gronu użytkowników, konieczność ochrony oryginałów przed zniszczeniem, zapotrzebowanie użytkowników. Z kolei zasoby repozytorium obrazują bieżące życie naukowe uczelni i tu najistotniejsza jest rola autorów, bo w ogromnej mierze zasoby te są wynikiem archiwizacji przez autorów prac opublikowanych w innych źródłach lub jedynie w repozytorium. W niniejszym rozdziale przedstawiono najważniejsze kryteria decydujące o tym, czy zasób cyfrowy zdeponowany będzie w bibliotece cyfrowej, czy w repozytorium.

#### **3.1. Digitalizacja jako sposób zabezpieczenia dziedzictwa narodowego – zbiory ważne historycznie, kulturowo i artystycznie**

Dokumenty mające dużą wartość historyczną, kulturową i artystyczną w znakomitej większości należą do domeny publicznej. Fakt ten znajduje odzwierciedlenie w zasobach zgromadzonych przez biblioteki zrzeszone w ramach Federacji Bibliotek Cyfrowych (FBC). Dokumenty, dla których wygasł 70-letni okres ochrony praw autorskich, cieszą się dużą popularnością jako obiekty wybierane w bibliotekach do digitalizacji. Czy cieszą się również popularnością wśród użytkowników bibliotek cyfrowych? Opierając się na zasobach FBC, dla ilustracji podano poniżej listę najczęściej wykorzystywanych dokumentów w okresie 20.08-20.09.2014 r. wraz z określeniem ich typu. W nawiasie podana została liczba odwiedzin:

- *Morze wiosną* – obraz [668]
- *Widok Warszawy z tarasu Zamku Królewskiego* – obraz [211]
- *Nowy Kurier Warszawski*, 1939, nr 21 (3 listopada) – czasopismo [151]
- *Wykaz dóbr ziemskich skonfiskowanych przez rządy zaborcze w latach 1773-1867* – wykaz, spis [149]
- *Polski starosta weselny: przemówienia, toasty, zwyczaje oraz pieśni i śpiewki weselne u ludu naszego w różnych okolicach kraju z dod. nadzwyczajnym: Wesele wiejskie* – książka [147]
- *Mapa samochodowa Polski (stan dróg) na rok 1939/1940* – mapa [141]

- *Atlas geograficzny ilustrowany Królestwa Polskiego* [Dokument kartograficzny] – atlas, mapa, książka [132]
- *Dresden. Kgl. Zwinger. Ostlicher Mittelpavillon* [Drezno. Zwinger. Pawilon Środkowy (Glockenspielpavillon)] – fotografia [130]
- *Podstawy automatyki: zbiór zadań z przykładowymi rozwiązaniami* – książka [122]
- *Herbarz polski* Kaspra Niesieckiego S. J., T. 8: [R-S] – książka [116]
- *Kodeks dyplomatyczny Małopolski*, T. 3, 1333-1386 – książka [111]
- *Wieczór autorski z pisarką niemiecką Juli Zeh* – plakat, afisz [108]
- [Portret Małgorzaty Kreczmar] – fotografia [105]
- [Portret Joanny Jurandot-Nawrockiej] – fotografia [104]
- *Monografie historyczno-genealogiczne niektórych rodzin polskich*, T.3 – książka [98]

Jak widać, wykorzystywane były wszystkie typy dokumentów, przy czym, choć 90% zasobów FBC stanowią czasopisma, czasopismo wśród użytkowanych dokumentów w testowanym okresie pojawiło się tylko raz. Dwa pierwsze najczęściej przeglądane dokumenty to obrazy o wysokiej wartości artystycznej, umieszczone w zasobach FBC przez Muzeum Narodowe w Warszawie. Obecnie najczęściej udostępnianymi obiektami artystycznymi w bibliotekach cyfrowych są: obrazy, ryciny, rysunki, mapy i fotografie, w tym także zdjęcia eksponatów trójwymiarowych.

Z czasem, kiedy muzea i galerie będą udostępniać w przestrzeni cyfrowej coraz większą liczbę zdigitalizowanych kopii eksponatów, szczególnie w technice 3D, będą tworzone nowe portale lub serwisy dołączane do już istniejących, funkcjonujące analogicznie do wirtualnych bibliotek. Takie przykłady można podać już dziś, np. zbiory Victoria and Albert Museum są prezentowane w ramach projektu 3D-Coform. Projekt nie ogranicza się jedynie do ekspozycji ciekawych artefaktów, ma również cele edukacyjne, związane zarówno z prezentowanymi obiektami, jak i ze sposobami ich przeniesienia w przestrzeń cyfrową. Liczba jednostek gromadzących zasoby kultury udostępniane na integrujących platformach, będzie zapewne stale wzrastać wraz z rozwojem technologii, pozwalającej na digitalizację najróżniejszych typów obiektów.

Dobrym przykładem platformy integrującej jest portal Europeana, który z założenia ma umożliwić wolny dostęp do dóbr nauki i kultury. Typy dokumentów uwzględnione w filtrach wyszukiwania tego portalu wyodrębnione zostały na podstawie rodzaju utrwalonych treści: obrazu, tekstu, nagrań wideo, dźwięku, grafiki / animacji 3D. W kolekcjach poszczególnych instytucji dostarczających zasoby do Europeany nie zawsze jednak można znaleźć metadane dokładnie określające typ dokumentu.

Repozytoria w ich obecnym rozumieniu i kształcie nie są raczej przystosowane do gromadzenia zasobów artystycznych czy dokumentów o dużym znaczeniu historycznym. Podstawową przesłanką tworzenia repozytoriów jest bowiem deponowanie dokumentów powstających pierwotnie w postaci cyfrowej, a nie ich wtórnych, cyfrowych kopii.

## **3.2. Zabezpieczenie zbiorów zagrożonych unicestwieniem**

### **3.2.1. Zabezpieczenie zbiorów tradycyjnych**

Bardzo ważnym kryterium wyboru dokumentów do digitalizacji w bibliotekach cyfrowych jest zabezpieczenie zbiorów, szczególnie druków z XIX i XX w., utrwalonych na tzw. kwaśnym papierze. W porównaniu z drukami sprzed 1850 r. część zbiorów późniejszych, szczególnie wielkoformatowych, znajduje się w katastrofalnym stanie. Obecnie biblioteki stoją zatem przed wielkim wyzwaniem, jakim jest zabezpieczenie treści ilustrujących życie codzienne z tego okresu, zgromadzonych w gazetach i czasopismach. Szacunkowa liczba dotychczas zdigitalizowanych czasopism wciąż nie przekracza 10% zasobów, a polityka udostępniania w wielu bibliotekach stoi w sprzeczności z polityką zabezpieczenia. Kwestia zabezpieczenia jak największego zasobu poszczególnych tytułów jest problemem na skalę kraju, a nie tylko poszczególnych bibliotek

Aktualnie nie ma w Polsce instytucji, która zajęłaby się koordynacją digitalizacji czasopism, których zasoby rozproszone są w wielu instytucjach na terenie kraju. Nie ma również narzędzia pozwalającego na przeglądanie dotychczas zdigitalizowanych zasobów oraz informowanie o brakach poszczególnych numerów. Wyszukiwarka FBC, wraz z listą zasobów planowanych do skanowania, jest tylko „protezą”, która ma ograniczone zastosowanie w praktyce. Jak pokazuje doświadczenie, koordynatorzy projektów wyznaczający zasoby do digitalizacji nie zawsze zapoznają się z tą listą, co w efekcie powoduje duplikowanie digitalizowanych zasobów, zamiast ich uzupełniania. Jeśli w najbliższym czasie nie powstanie instytucja powołana do koordynacji działań w tym zakresie, to istnieje ryzyko zmarnotrawienia dużych środków, jakie zostały przeznaczone na digitalizację w VIII Programie Ramowym „Horizon 2020”. Wstępem do działań koordynacyjnych mogłoby być wykorzystanie Centralnego Katalogu Czasopism Polskich i, przy współpracy z Poznańskim Centrum Superkomputerowo-Sieciowym zarządzającym Federacją Bibliotek Cyfrowych, stworzenie narzędzia służącego z jednej strony do przeglądania rozproszonych zasobów według przejrzystej struktury, a z drugiej – do pozyskiwania informacji o zasobach do zdigitalizowania, rozproszonych w wielu instytucjach w Polsce. Oczywiście koordynacja mia-

łaby objąć nie tylko czasopisma, ale całość zbiorów digitalizowanych we wszystkich bibliotekach w kraju.

Kolejnym typem dokumentów, które wymagają pilnego zeskanowania, są dokumenty życia społecznego, a szczególnie zbiory spoza oficjalnego obiegu wydawniczego. Bardzo zły jest stan fizyczny tych zasobów, wykonanych niejednokrotnie za pomocą prymitywnych metod, podczas gdy treści w nich zawarte często są jedynym świadectwem skomplikowanej historii naszego kraju. Pozostałe dokumenty zagrożone zniszczeniem z powodu kwasowej hydrolizy papieru to książki. Ponieważ zainteresowanie, jakie wzbudzają one u użytkowników, z reguły nie jest tak wielkie, aby prowadzić do destrukcji obiektu, dodatkowym kryterium przy ich digitalizacji jest stan zachowania danego egzemplarza oraz jego unikatowość.

Kolejną kategorią dokumentów, które mogą zawierać bardzo ważne treści, są zbiory audiowizualne: poczynając od płyt gramofonowych, poprzez taśmy szpulowe i kasety magnetofonowe (w przypadku dokumentów dźwiękowych) oraz taśmy 16 czy 32 mm, na kasetach wideo (w przypadku filmów) kończąc. Niejednokrotnie to właśnie w bibliotekach zostały także zdeponowane materiały dokumentujące życie codzienne, których trwałość, ze względu na amatorski charakter wykonania oraz zróżnicowaną jakość nośników i niekorzystne warunki przechowywania, może być zagrożona.

Zbiory specjalne, ze względu na jakość wykorzystanego do ich produkcji materiału (papier czerpany, pergamin), są generalnie w o wiele lepszym stanie niż publikacje drukowane na tzw. kwaśnym papierze, dlatego proces destrukcji w ich przypadku zachodzi wolniej, także ze względu na bardziej ograniczoną ekspozycję. Poza tym zbiory specjalne to wartościowy materiał źródłowy, którego digitalizacja jest bardzo często powiązana z większymi projektami badawczymi. Nad tymi zbiorami zwykle czuwają konserwatorzy, którzy powinni mieć udział w podejmowaniu decyzji o tym, kiedy i jaką technologią dany obiekt powinien zostać poddany skanowaniu. Zbiory specjalne to często dzieła unikatowe, należące do Narodowego Zasobu Bibliotecznego, stanowiące materiał źródłowy do badań, a więc ich miejsce na liście dokumentów priorytetowo wybieranych do digitalizacji jest dość wysokie.

Rodzaje dokumentów, które są uznawane za priorytetowe podczas wyboru zasobów do digitalizacji, to – poczynając od najważniejszych:

- gazety i czasopisma z XIX i XX w.,
- dokumenty życia społecznego, a szczególnie zbiory spoza oficjalnego obiegu wydawniczego,
- zbiory audiowizualne, szczególnie o charakterze dokumentalnym, zarejestrowane na nośnikach analogowych,

- pozostałe zbiory z XIX i XX w. z uwzględnieniem stanu zachowania,
- zbiory specjalne: inkunabuły, stare druki, rękopisy, zbiory graficzne i kartograficzne oraz muzykalia.

### 3.2.2. Zabezpieczenie publikacji cyfrowych

W epoce internetu i rosnącej popularności dokumentów cyfrowych liczba publikacji wytwarzanych rocznie na średniej wielkości uniwersytecie mieści się w granicach 10 tys. tytułów. Do tego należy doliczyć niepublikowane materiały naukowe i dydaktyczne. Łatwość upublicznienia materiału cyfrowego lub jego skopiowania jest równa łatwości jego usunięcia. Równocześnie łatwość upubliczniania materiałów cyfrowych sprzyja rozproszeniu publikacji elektronicznych na stronach www instytutów i wydziałów. Między innymi dlatego zostały powołane repozytoria instytucjonalne, zarządzane przez bibliotekarzy, mające na celu zabezpieczanie i archiwizowanie dorobku cyfrowego uczelni i agregowanie dostępu do niego. W repozytoriach gromadzi się najczęściej następujące typy dokumentów:

- artykuły z czasopism,
- książki i fragmenty książek (w tym podręczniki, skrypty, monografie),
- materiały konferencyjne,
- materiały edukacyjne dla studentów,
- prace dyplomowe i dysertacje,
- hasła z encyklopedii i słowników,
- recenzje,
- normy, patenty,
- materiały niepublikowane,
- dokumenty dźwiękowe,
- dokumenty filmowe,
- atlasy, mapy,
- bibliografie.

Zarządzanie tak zróżnicowanym zasobem, wytworzonym w wersji cyfrowej i w takiej postaci zdeponowanym, wymaga od zarządzających tego typu bazą danych nie tylko ścisłej współpracy z informatykami, ale także pewnej wiedzy informatycznej, a przy tym również – bibliotekarskiej skrupulatności i sumienności.

### 3.3. Zapotrzebowanie użytkowników (popularność zbiorów)

Dla użytkownika – zainteresowanego dotychczas głównie zasobami bibliotecznymi, przechowywanymi w magazynach i dostępnymi jedynie lokalnie, na

terenie biblioteki – rozwój bibliotek cyfrowych w efekcie digitalizacji jest wielkim krokiem naprzód. Dużym ograniczeniem tego rozwoju, powodującym udostępnianie w bibliotekach cyfrowych przede wszystkim dokumentów znajdujących się w domenie publicznej, jest jednak obowiązujące prawo autorskie. Obecnie przyjęty 70-letni okres ochrony utworów skutecznie ogranicza szeroki dostęp do nowszych tytułów, wydanych po II wojnie światowej. Z najnowszych publikacji udostępnianych w postaci cyfrowej użytkownicy mogą korzystać jedynie z komputerów znajdujących się w bibliotece. Obok chronionych prawem autorskim dostępne są też czasopisma oraz inne publikacje, ukazujące się w ramach wolnych licencji. Jeśli chodzi o zasoby JBC, do najpopularniejszych typów dokumentów należą (począwszy od najczęściej przeglądanych):

- zbiory ikonograficzne oraz mapy,
- czasopisma,
- stare druki,
- rękopisy,
- książki, w tym książki nowe na licencjach otwartych.

Natomiast w repozytoriach, zarówno polskich, jak i światowych, wśród najczęściej przeglądanych publikacji – w kolejności od najpopularniejszych – zdecydowanie dominują następujące typy dokumentów:

- artykuły z czasopism,
- materiały edukacyjne,
- prace dyplomowe i dysertacje,
- książki i rozdziały książek – przy czym są to w znakomitej większości tytuły publikowane na wolnych licencjach.

W tym miejscu należy zwrócić uwagę na ten szczególny typ dokumentów, jakim są wspomniane książki – publikowane na wolnych licencjach, występujące zarówno w bibliotekach cyfrowych, jak i w repozytoriach. W JBC są to książki typu *born digital*, opublikowane przez różnych autorów i różnych wydawców, pozyskane do zasobów BJ w ramach egzemplarza obowiązkowego, wymiany lub daru. Natomiast w RUJ gromadzone są książki powstające na wolnych licencjach, których autorzy należą do społeczności UJ i które są umieszczane w RUJ przez Bibliotekę Jagiellońską.

### **3.4. Archiwizowanie materiałów naukowych**

Naukowość to kryterium, według którego kwalifikowana jest do zamieszczenia w repozytoriach zdecydowana większość publikowanych w nich dokumentów. Jedynym odstępstwem od tej reguły są historyczne zasoby naukowe, gromadzone w bibliotekach cyfrowych. Sytuacja ta jest wynikiem rozdzielania

funkcji obydwu platform i przydzielenia repozytoriom roli skarbnicy publikowanych i niepublikowanych wyników badań, a także surowych danych, prezentacji z konferencji czy raportów z grantów. Co prawda w bibliotekach cyfrowych są również gromadzone zasoby naukowe, ale z reguły nie są one wystarczająco wyodrębnione z zasobów historycznych tych bibliotek. Natomiast w repozytoriach, z definicji gromadzących dorobek naukowy pracowników instytucji, jest on w większości przypadków podzielony na następujące grupy, obejmujące różne typy dokumentów:

- badania, w tym surowe dane,
- publikacje, w tym: książki / fragmenty książek, artykuły z czasopism,
- dysertacje, prace dyplomowe,
- raporty z grantów,
- patenty, normy,
- materiały konferencyjne,
- czasopisma publikowane przez jednostki uczelni.

Z jednej strony ruch Open Access, a z drugiej – polityka uczelni dążących do zgromadzenia w jednym miejscu dorobku naukowego swoich pracowników przyczyniają się do zwiększania ich aktywności publikacyjnej, zarówno na łamach czasopism naukowych, jak i w repozytoriach. To powoduje gwałtowny wzrost liczby publikacji naukowych dostępnych w internecie, przy czym nie zawsze jest to dostęp powszechny. Niekiedy polityka ochrony własności intelektualnej prowadzi wręcz do stosowania w tym zakresie działań podobnych do tych występujących na komercyjnym rynku wydawniczym.

### **3.5. Archiwizowanie materiałów dydaktycznych**

Do zasobów ważnych ze względów dydaktycznych można zaliczyć:

- wykłady, szczególnie te rejestrowane na wideo,
- materiały do ćwiczeń,
- skrypty,
- materiały do nauki języków obcych.

Pojawiają się one zarówno w bibliotekach cyfrowych (kopie cyfrowe wydanych materiałów), jak i w repozytoriach, przy czym w tych ostatnich są reprezentowane o wiele częściej. Wynika to głównie z charakteru repozytoriów, które są platformami usługowymi przeznaczonymi dla społeczności akademickich. W repozytoriach uczelnianych coraz częściej są również przechowywane zasoby tworzone w ramach platform e-learningowych, co z jednej strony zapewnia alternatywny sposób dostępu do nich, a z drugiej – stanowi dodatkowe zabezpieczenie treści.



### 3.6. Publikowanie i rozpowszechnianie

„Biblioteki cyfrowe są formą publikowania, realizowaną przez biblioteki tradycyjne, które umieszczają w sieci zdigitalizowane wersje publikacji drukowanych, zazwyczaj posiadanych w swoich zasobach” [Nahotko 2010, s. 92]. Jest to swoista reedycja tych dzieł z zadbaniami – poprzez metadane (często powiązane z metadanymi dotyczącymi danych druków w katalogach komputerowych) – o ich autentyczność i zgodność z pierwowzorem.

Podobny proces publikowania ma miejsce w repozytoriach, ale tylko w przypadku tzw. e-printów, czyli rozpowszechnianych tą drogą wersji artykułów publikowanych w czasopiśmie komercyjnym. Natomiast repozytoria instytucjonalne są serwisem, w którym publikowane są też oryginalne prace naukowe. Publikowane obiekty cyfrowe opatrywane są metadanymi.

Wśród coraz liczniej utrwalanych, przechowywanych i udostępnianych zasobów informacji powstających na uczelniach można wydzielić typy dokumentów najczęściej udostępniane w repozytoriach. Mogą to być:

- czasopisma wydawane przez jednostki uczelni,
- materiały konferencyjne,
- materiały edukacyjne / dydaktyczne,
- wykłady, wywiady, inne zapisy audio i wideo,
- wystawy biblioteczne,
- fotoreportaże z imprez itp.

Repozytoria uczelniane mogą też przyjąć na siebie funkcje wydawnicze, czyli zadanie formalnego publikowania opracowywanych tekstów (organizowanie procesu wydawniczego, w tym recenzowanie, opracowanie redakcyjne, skład, korektę, zarządzanie prawami autorskimi, promocję). Dotyczy to w szczególności takich typów dokumentów, jak czasopisma, książki, podręczniki, materiały konferencyjne. Działalność wydawnicza z założenia miałaby charakter *non-profit* i była prowadzona na licencji Open Access. Jest to tym bardziej możliwe, że naukowcy, otrzymując granty, mają często zapewnioną pulę środków na publikację wyników badań, z których finansowane jest opracowanie książek. Niestety takie rozwiązania, szczególnie w Polsce, są jeszcze dość mało popularne.

Warto też podkreślić istotne znaczenie repozytoriów jako platform dystrybucji publikacji. Jest to tzw. zielona droga Open Access, polegająca na niekomercyjnym udostępnianiu dokumentów za pomocą platform zarządzanych z reguły przez biblioteki uniwersyteckie. Autorzy tekstów mają możliwość starania się o publikację w wydawnictwach komercyjnych, nie zamykając sobie jednak drogi do szerokiego upowszechniania opracowanych treści poprzez zdeponowanie ich w repozytorium instytucjonalnym. Rozwijający się od lat 90. XX w. ruch Open

Access, sprzyjający wzrostowi świadomości autorów, odnośnie do ich praw oraz korzyści, jakie płyną z publikowania w oparciu o wolne licencje, powoduje, że dotychczasowy model wydawniczy ulega erozji, co widać na przykładzie działań największych wydawców, pozwalających na równoległe publikowanie artykułów zarówno w modelu otwartym (np. pre-printów), jak i tradycyjnym.

## 4. Dobór zasobów cyfrowych do JBC i RUJ

Tworząc w bibliotece uniwersyteckiej równoległe funkcjonujące platformy biblioteki cyfrowej i repozytorium, należy ustanowić kryteria doboru zasobów do tych dwóch baz. Kryteria te powinny uwzględniać wiele aspektów związanych zarówno z formalnymi, jak i treściowymi charakterystykami dokumentów. Ich zakresy mogą się krzyżować i zachodzić na siebie w różnych relacjach hierarchicznych. Natomiast podstawową i oczywistą cechą, wspólną dla dokumentów przeznaczonych zarówno do bibliotek cyfrowych, jak i do repozytoriów, jest ich cyfrowość.

### 4.1. Jagiellońska Biblioteka Cyfrowa

Podstawowym celem gromadzenia i kryterium doboru dokumentów w przypadku bibliotek cyfrowych jest zapewnienie możliwie najszerszego, powszechnego i bezpłatnego dostępu do zasobów własnych placówek bibliotecznych w nowoczesny sposób – w postaci cyfrowej. Dotyczy to zarówno obiektów powstałych w wyniku **digitalizacji zbiorów własnych**, jak i gromadzonych **zasobów oryginalnie cyfrowych** (ang. *born digital*), otrzymywanych od różnych wydawców w ramach egzemplarza obowiązkowego.

Dobór dokumentów do digitalizacji opiera się na kilku kryteriach, wśród których do najbardziej istotnych należą:

- zabezpieczanie dziedzictwa narodowego, dzieł szczególnie cennych, unikatowych, należących do Narodowego Zasobu Bibliotecznego,
- zabezpieczanie zbiorów zagrożonych zniszczeniem z powodu kwasowej degradacji papieru,
- realizacja przyjętych założeń związanych z tworzeniem kolekcji tematycznych i wystaw okolicznościowych,
- cyfrowe udostępnianie zasobów umożliwiających badania naukowe (posiadających wartość intelektualną i historyczną),
- cyfrowe udostępnianie zasobów wspomagających działalność kulturalną i oświatową,

- cyfrowe udostępnianie materiałów dydaktycznych i popularyzatorskich,
- cyfrowe udostępnianie materiałów związanych z poszczególnymi regionami,
- realizacja potrzeb użytkowników.

Podkreślić należy, że przy selekcji dokumentów do digitalizacji z analogowych zasobów biblioteki, gromadzonych przez wiele wieków i lat, priorytety ustala się zwykle w oparciu o więcej niż jedno z wymienionych wyżej kryteriów. W bibliotece cyfrowej umieszcza się natomiast wszystkie pozyskane dla biblioteki zbiory opublikowane oryginalnie w postaci elektronicznej (selekcję, o ile jest potrzebna, przeprowadza się na etapie gromadzenia). Udostępniane w bibliotece cyfrowej obiekty (wtórne – będące wynikiem digitalizacji oraz prymarne – natywnie cyfrowe) kategoryzowane są według kryteriów formalnych, analogicznie do tradycyjnych zbiorów bibliotecznych. Tak więc w odrębnych kolekcjach udostępniane są książki, czasopisma, rękopisy, stare druki, muzykalia, ikonografia, dokumenty życia społecznego, dokumenty audiowizualne i dokumenty natywnie cyfrowe. Obiekty cyfrowe deponowane są w bibliotece cyfrowej przez bibliotekarzy.

#### **4.2. Repozytorium instytucjonalne UJ**

Podstawowym celem repozytorium instytucjonalnego jest gromadzenie, przechowywanie i udostępnianie bieżącego dorobku intelektualnego społeczności naukowej jednej lub kilku uczelni [Crow 2002]. Repozytoria stają się obecnie jednym z głównych kanałów komunikacji naukowej – są związane z ideą otwartej nauki i otwartego dostępu.

W rozwoju ruchu otwartego dostępu do wiedzy istotne znaczenie miała Deklaracja Budapesztańska, która została opublikowana w 2002 r., a po 10 latach potwierdzono ponownie wagę wprowadzonych przez nią dwóch podstawowych strategii: otwartego dostępu do wiedzy w repozytoriach (zwanego modelem zielonym) oraz w czasopismach (zwanego modelem złotym). W zaktualizowanych rekomendacjach na kolejne dziesięciolecia znalazł się m.in. zapis: „Każda instytucja – mająca uprawnienia do nadawania stopni naukowych – powinna wprowadzić zasadę deponowania wyników badań w repozytorium instytucjonalnym Open Access (...)” [Bednarek-Michalska et al. 2012, dok. elektr., s. 1-2].

W cyfrowym repozytorium uczelnianym powinny więc znaleźć się:

- **materiały naukowe:**
  - opublikowane i niepublikowane, tworzone przez pracowników naukowych uczelni (w tym pre-printy, post-printy, materiały publikowane w komercyjnych wydawnictwach),
  - rozprawy doktorskie,
  - materiały z konferencji naukowych,

- wydawnictwa w formie cyfrowej, publikowane przez jednostki uczelni (np. czasopisma),
- dane badawcze do wielokrotnego wykorzystywania i weryfikacji (wyniki badań, pomiarów, statystyki, raporty, sprawozdania itp.);
- **niepublikowane materiały dydaktyczne**, wytworzone przez pracowników naukowych uczelni na potrzeby wykładów, ćwiczeń, egzaminów;
- **zasoby dokumentujące działalność uczelni**.

Za pośrednictwem repozytorium instytucjonalnego można też realizować funkcje wydawnicze – pozyskiwać wartościowe zrecenzowane i zredagowane teksty, zapewniać redakcję techniczną i skład, pozyskiwać numery ISBN / ISSN, formalnie wydawać opracowane teksty w postaci cyfrowej oraz dbać o ich promocję. Materiały cyfrowe są deponowane w repozytorium przez pracowników naukowych / autorów tekstów przy współpracy z bibliotekarzami i przechowywane bezterminowo.

## 5. Metadane w JBC i w RUJ

W Jagiellońskiej Bibliotece Cyfrowej i powstającym Repozytorium Uniwersytetu Jagiellońskiego do tworzenia opisów obiektów zastosowano ten sam standard metadanych, czyli Dublin Core. Wynika to z potrzeby zapewnienia jednolitego standardu danych, np. w przypadku publikacji pojawiających się w obu serwisach, szczególnie zaś w przypadku budowania multiwyszukiwarki obsługującej zarówno JBC, jak i RUJ.

Tabela 2. Metadane w Jagiellońskiej Bibliotece Cyfrowej i Repozytorium Uniwersytetu Jagiellońskiego

Dublin Core	JBC	RUJ	Opis
dc:abstract	description.abstract	abstract.en	Abstrakt w jęz. ang.
		abstract.other	Abstrakt w innym jęz.
		abstract.pl	Abstrakt w jęz. pol.
dc:coverage	coverage	–	Zakres
dc:contributor	degree.tutor	contributor.advisor	Promotor
	creator	contributor.author	Autor
	contributor	contributor.editor	Redaktor
		contributor.ilustrator	Ilustrator
		contributor.institution	Instytucja sprawcza
	degree.adviser	contributor.reviewer	Recenzent
contributor	contributor.translator	Tłumacz	

c.d. tabeli 2.

Dublin Core	JBC	RUJ	Opis
dc:date	date.access_date date	date.accession	Data dostępu
		date.created	Data powstania
		date.issued	Data wydania
dc:description	–	description.additional	Opis dodatkowy
	–	description.conftype	Typ konferencji
	–	description.edition	Wydanie
	–	description.govdocpl	Identyfikator lokalny patentu
	title	description.number	Numer (czasopisma)
	–	description.physical	Opis fizyczny
	–	description.points	Punkty
	–	description.publication	Liczba arkuszy wydawniczych
	description.scale	description.scale	Skala
	source.bibliographic_citation	description.series	Tytuł serii
	source.bibliographic_citation	description.seriesnumber	Numer serii
	–	description.sponsorship	Instytucja dotująca
	–	description.uri	Uniform Resource Identifier
	source.bibliographic_citation	description.volume	Numer tomu
dc:format	description.duration	format.extent	Czas trwania lub rozmiar pliku
	format	format.medium	Nośnik
dc:identifier	locationOf PhysicalObject	identifier.callnumber	Sygnatura
	–	identifier.reviewed	Cytata bibliograficzna dzieła recenzowanego
	–	identifier.doi	DOI
	–	identifier.eisbn	e-ISBN
	–	identifier.eissn	e-ISSN
	identifier.isnb	identifier.isnb	ISBN
	identifier.issb	identifier.issn	ISSN
	–	identifier.uri	URI
dc:language	language	language	Język dokumentu
	–	language.container	Język publikacji nadrzędnej
dc:participation	–	participation	Udział procentowy autorów

c.d. tabeli 2.

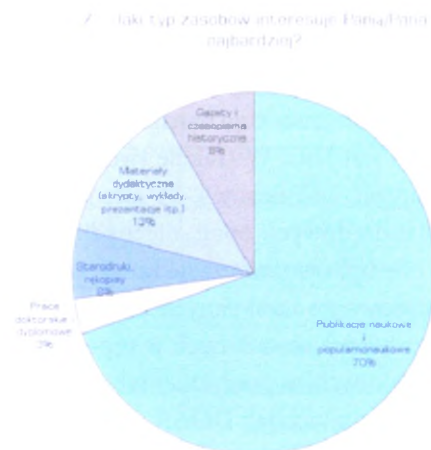
Dublin Core	JBC	RUJ	Opis
dc:place	degree.faculty	place.created	Miejsce powstania
	degree.institute	place.department	Instytut / Zakład / Katedra
	–	place.doct	Afiliacja doktoratu / pracy dyplomowej
	publisher.place	place	Miejsce wydania
dc:relation	relation.hasversion	relation.hasversion	Inne wydania
	relation.supplements	relation.isformatof	Dodatki
	relation.main_title	relation.ispartof	Wydawnictwo główne
	relation.previous	relation.isreplacedby	Tytuł poprzedni
	relation.next	relation.replaces	Tytuł następny
	relation.completion	relation.requires	Katalog komputerowy
dc:rights	rights	rights	Status prawny
	rights.holder	rights.holder	Właściciel praw
dc:source	source.dig	–	Identyfikator dokumentu cyfrowego
dc:subject	subject	subject.en	Temat i słowa kluczowe w jęz. ang.
		subject.pl	Temat i słowa kluczowe w jęz. pol.
dc:title	title	title	Tytuł właściwy
	title.version	title.alternative	Inny tytuł
	source.bibliographic_citation	title.container	Tytuł publikacji nadrzędnej
	source.bibliographic_citation	title.journal	Tytuł czasopisma
	title.original	title.original	Tytuł oryginału
	source.bibliographic_citation	title.volume	Tytuł numeru
dc:type	type	–	Typ dokumentu
	–	type.group	Kategoria publikacji (dla formularza oceny pracownika)
dc:locationOf PhysicalObject	locationOf PhysicalObject	–	Lokalizacja oryginału
dc:classifications	classifications.kbn	–	Klasyfikacja KBN
	classifications.ukd	–	Klasyfikacja UKD
dc:digitisation-sponsor	digitisationsponsor	–	Źródło finansowania
dc:tags	tags	–	Tagi

Sporządzone w tabeli 2 zestawienie metadanych wyraźnie pokazuje dużo większą ich szczegółowość w budowanym repozytorium UJ niż w Jagiellońskiej Bibliotece Cyfrowej. Wynika to z kilku przyczyn. Po pierwsze, szczegółowość metadanych wymusza platforma DSpace, znacznie bardziej rozwinięta pod tym względem niż dLibra, na której oparta jest JBC. W trakcie tworzenia biblioteki cyfrowej do jej funkcjonowania wystarczył podstawowy, 15-elementowy schemat, z czasem wzbogacony o kilka niestandardowych opcji, niezbędnych do wprowadzania dodatkowych metadanych. Nie było w tym czasie konieczności używania kwalifikatorów, które służą doprecyzowaniu metadanych używanych w repozytorium. Potrzeba ścisłego ustrukturyzowania metadanych w repozytorium jest drugą, ważną przyczyną różnic przedstawionych w powyższej tabeli. Mając na uwadze eksport danych do multiwyszukiwarek, takich jak DOAR, ROAR lub Google Scholar, istotne jest jak najbardziej precyzyjne opisanie dokumentu, chociażby pod względem ścisłego określenia zakresu tematycznego czy autorstwa. Nie mniej istotne, a może nawet ważniejsze, jest uwzględnienie danych jednostki organizacyjnej, w tym przypadku Uniwersytetu Jagiellońskiego. Bardzo istotnym powodem uszczegółowienia schematu metadanych RUJ była też konieczność dostosowania go do potrzeb oceny pracowników. Repozytorium tworzone jest bowiem również jako platforma mająca na celu zgromadzenie dorobku naukowego wszystkich pracowników Uniwersytetu w postaci Bibliograficznej Bazy Pracowników UJ. Zastosowanie rozwiązania, pozwalającego na umieszczanie obok opisów bibliograficznych również pełnych tekstów dokumentów, umożliwi z jednej strony prowadzenie dokładnej sprawozdawczości, a z drugiej – promocję osiągnięć naukowych UJ.

## **6. Podsumowanie – JBC i RUJ – model współdziałania**

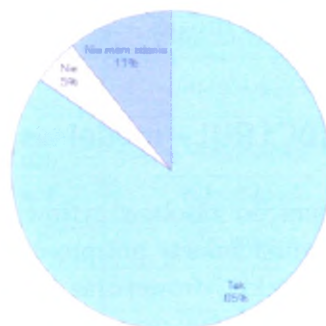
Zainteresowanie dostępem do zasobów cyfrowych w społeczności akademickiej UJ potwierdzają wyniki ankiety przeprowadzonej w 2012 r. w ramach ewaluacji Jagiellońskiej Biblioteki Cyfrowej oraz przygotowań do budowy Repozytorium UJ. Próba objęła ponad 1000 pracowników i studentów Uniwersytetu, a otrzymane rezultaty, przedstawione na wykresach 1, 2 i 3, potwierdziły hipotezy stawiane przez autorów badań.

Jak widać na wykresie 1, najwięcej, bo aż 70% ankietowanych, było zainteresowanych publikacjami naukowymi i popularnonaukowymi. To był jeden z głównych czynników mobilizujących bibliotekarzy do rozpoczęcia prac nad tworzeniem repozytorium instytucjonalnego UJ. Potrzebę udostępnienia dorobku naukowego pracowników Uniwersytetu w extranecie potwierdziły dodatkowo dane przedstawione na wykresie 2.



Wykres 1. Zapotrzebowanie na poszczególne typy dokumentów wśród pracowników i studentów UJ  
 Źródło: [Szafrński, Brawuski 2014, dok. elektr., s. 5]

8. Czy uważa Pan/Pani, że prace naukowe powstające na Uniwersytecie Jagiellońskim powinny być zamieszczane w repozytorium cyfrowym dostępnym dla pracowników i studentów UJ (poprzez ekran)?

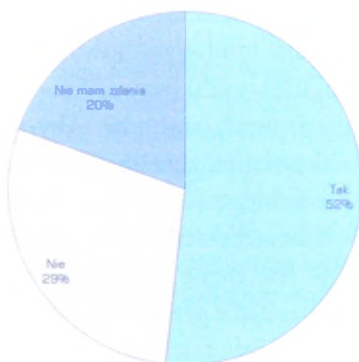


Wykres 2. Zapotrzebowanie społeczności UJ na prace naukowe dostępne lokalnie  
 Źródło: [Szafrński, Brawuski 2014, dok. elektr., s. 5]

O ile znakomita większość pracowników i studentów UJ widzi potrzebę dostępu do platformy, na której zdeponowane byłyby prace naukowe wytworzone na Uniwersytecie, to za powszechnym, otwartym dostępem w internecie opowiada się tylko połowa z nich. Dokładne dane przedstawione zostały na wykresie 3.



10 – Czy uważa Pani/Pan, że prace naukowe powstałe na Uniwersytecie Jagiellońskim powinny być zamieszczane w otwartym repozytorium cyfrowym dostępnym dla każdego użytkownika sieci Internet?



Wykres 3. Nastawienie pracowników i studentów UJ do polityki otwartości publikacji naukowych  
 Źródło: [Szafrński, Brawuski 2014, dok. elektr., s. 5]

Opierając się na przedstawionych powyżej wynikach badań, Biblioteka Jagiellońska przystąpiła do budowy, obok Jagiellońskiej Biblioteki Cyfrowej (dLibra), Repozytorium Uniwersytetu Jagiellońskiego (DSpace). Wychodząc z założenia, że „promowanie dorobku intelektualnego uczelni wymaga jego konsolidacji i wyeksponowania” [Sapa 2013, s. 127], Biblioteka Jagiellońska wybrała taki model dostarczania informacji w postaci cyfrowej, w którym zasoby cyfrowe pozyskane przez uczelnię są oddzielone od zasobów przez nią wytworzonych.

Wydaje się, że utrzymywanie odrębnych platform na potrzeby biblioteki cyfrowej i repozytorium instytucjonalnego to – zarówno dla użytkowników zewnętrznych, jak i pracowników uczelni – łatwiejszy dostęp do określonych treści adresowanych do poszczególnych grup odbiorców. Co więcej, zachowanie odrębności tych serwisów umożliwi promocję zasobów instytucjonalnych zgromadzonych w repozytorium przez rejestrację w wyszukiwarkach oraz agregatorach danych i treści naukowych, akceptujących tylko serwisy o charakterze otwartego repozytorium instytucjonalnego [Werla 2013].

Z drugiej strony bardzo istotną kwestią jest umożliwienie użytkownikowi zintegrowanego wyszukiwania zasobów w sposób, do jakiego w ciągu ostatniej dekady przyzwyczaili go serwisy komercyjne, w tym obejmująca 95% rynku wyszukiwarka Google. Biblioteki cyfrowe, repozytoria oraz platformy integrujące zasoby zarówno z bibliotecznych magazynów w formie cyfrowej, jak i dokumen-

tów oryginalnie cyfrowych są, podobnie jak serwisy komercyjne, dostępnymi w internecie rejestratorami rzeczywistości. Istotną różnicę stanowi natomiast jakość gromadzonych przez nie zasobów.

W przypadku zbiorów jednostek naukowych, zamieszczanych zarówno w bibliotekach cyfrowych, jak i w repozytoriach, do ich integracji i łatwiejszego wyszukiwania niezbędne stają się multiwyszukiwarki, oferujące użytkownikowi interfejs oparty na jednym oknie wyszukiwawczym, czasami wzbogacony o możliwość przeglądania zgromadzonych zasobów. Aby wyszukiwanie przebiegało sprawnie, potrzebna jest maksymalna zbieżność opisów bibliograficznych, osiągnięta np. przez zastosowanie jednego schematu metadanych. Biorąc pod uwagę pewne różnice, wynikające chociażby z okresu powstawania zasobów cyfrowych, ujednoczenie ich opisów nie jest wcale łatwym zadaniem. Pojawiają się różne rozwiązania w tym zakresie, bazujące przede wszystkim na tabelach konwersji.

Różne role, jakie odgrywają biblioteki cyfrowe i repozytoria w zaspokajaniu potrzeb informacyjnych użytkowników, mają wpływ na rodzaje udostępnianych w nich zasobów. Z zasady bazy biblioteki cyfrowej i repozytorium powinny być rozłączne pod względem gromadzonych zbiorów, jednakże nie wyklucza się dublowania niektórych treści i publikowania tych samych dokumentów w obydwu z nich.

Niedogodności związane z koniecznością prowadzenia odrębnych sesji wyszukiwawczych w JBC i w RUJ można zniwelować przez połączenie tych dwóch baz za pomocą narzędzia indeksującego metadane z obu źródeł. Zadanie to ułatwia fakt, że opis bibliograficzny zarówno w JBC, jak i w RUJ jest tworzony w oparciu o format Dublin Core, choć – jak wynika to z tabeli 2 – zastosowane metadane wymagają ujednoczenia. Spośród wielu dostępnych obecnie aplikacji służących integracji procesów wyszukiwania najbardziej funkcjonalny jest system Hydra, udostępniany na licencji Open Source (Apache 2). Jest on stosowany z powodzeniem w takich instytucjach, jak Stanford University, University of Notre Dame czy London School of Economics and Political Science. Hydra jest rozwiązaniem szerszym niż multiwyszukiwarka. Implementacja narzędzia na dwóch platformach skorelowałaby obie te bazy. Poza postawieniem aplikacji serwerowej konieczne jest przygotowanie modelu łączenia opisów [Drabik 2014].

Innym rozwiązaniem może być przeniesienie wszystkich zasobów cyfrowych na jedną platformę, np. DSpace. Decyzja taka zależeć będzie głównie od przyrostu zasobów oraz aktualizacji i nowych rozwiązań w oprogramowaniu dLibra. Tak czy inaczej, obserwując tendencje światowe, można spodziewać się, że biblioteki uniwersyteckie najprawdopodobniej będą obejmowały coraz szersze zasoby cyfrowe swoich jednostek macierzystych i udostępniały je na coraz lepszych platformach integrujących różne dane.

## BIBLIOGRAFIA

- About eScholarship* [dok. elektr.]. [http://escholarship.org/about\\_escholarship.html](http://escholarship.org/about_escholarship.html) [odczyt: 23.08.2014].
- Bednarek-Michalska, Bożena et al. (2012). *Budapeszteńska inicjatywa otwartego dostępu do wiedzy po dziesięciu latach: ustawienie domyślne – otwartość* [dok. elektr.]. <http://www.budapestopenaccessinitiative.org/boai-10-translations/polish> [odczyt: 10.10.2014].
- Czapnik, Grzegorz; Gruszka, Zbigniew (2011). *Podręczny słownik bibliotekarza*. Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, 612 s.
- DLA Annual Report 2011/12* [dok. elektr.]. <http://scholar.lib.vt.edu/vtpubs/anreport/dla/AR201112.pdf> [odczyt: 25.08.2014].
- Drabik, Adrian (2014). *Biblioteka cyfrowa a repozytorium – Jagiellońska Biblioteka Cyfrowa a Repozytorium UJ (Hydra)* [dok. elektr.]. W: *E-eksplozja: narzędzia – metody – użytkownicy: konferencja naukowa*. Kraków: Biblioteka Jagiellońska, 27.03.2014. <http://jbc.bj.uj.edu.pl/dlibra/publication?id=290555&from=&dirids=1&tab=1&lp=7&QI> [odczyt: 11.10.2014].
- Gibbons, Susan (2004). *Defining an institutional repository* [dok. elektr.]. *Library Technology Reports* Vol. 40, Issue 4, p. 6-10. <http://web.b.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=2&sid=bfebd563-3750-427b-a237-dd2c73f7f960%40sessionmgr111&hid=105> [odczyt: 9.09.2014].
- Guidelines for digital object deposit*. Version 1.2 (2011) [dok. elektr.]. [http://www.hathitrust.org/deposit\\_guidelines](http://www.hathitrust.org/deposit_guidelines) [odczyt: 23.08.2014].
- i2010: Biblioteki Cyfrowe* – Komunikat Komisji Parlamentu Europejskiego, Rady Europejskiego Komitetu Ekonomiczno-Społecznego oraz Komitetu Regionów (2005) [dok. elektr.]. [http://ec.europa.eu/prelex/detail\\_dossier\\_real.cfm?CL=pl&DosId=193359](http://ec.europa.eu/prelex/detail_dossier_real.cfm?CL=pl&DosId=193359) [odczyt: 08.09.2014].
- Janiak, Małgorzata (2012). *Biblioteka cyfrowa, biblioteka elektroniczna, biblioteka wirtualna*. W: Janiak, Małgorzata; Krakowska, Monika; Próchnicka, Maria red. *Biblioteki cyfrowe*. Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, s. 15-65.
- Kolasa, Władysław M. (2012). *Biblioteki cyfrowe na świecie – powstanie i rozwój*. W: Janiak, Małgorzata; Krakowska, Monika; Próchnicka, Maria red. *Biblioteki cyfrowe*. Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, s. 66-95.
- Lynch, Clifford A. (2003). *Institutional Repositories: essential infrastructure for scholarship in the digital age* [dok. elektr.]. *Association of Research Libraries Bimonthly Report on Research Library issues and actions from ARL, CNL and SPARC* No. 226. <http://www.arl.org/storage/documents/publications/arl-br-226.pdf> [odczyt: 08.09.2014].
- McEntyre, Johanna R. (2011). *UKPMC: a full text article resource for the life sciences* [dok. elektr.]. *Nucleic Acids Research* Vol. 39. <http://europepmc.org/articles/PMC3013671> [odczyt: 24.08.2014].
- Nahotko, Marek (2004). *Cyfrowa najmłodsza siostra bibliotek* [dok. elektr.]. <http://www.wsp.krakow.pl/konspekt/19/nahotko.html> [odczyt: 10.09.2014].
- Nahotko, Marek (2010). *Komunikacja naukowa w środowisku cyfrowym: globalna biblioteka cyfrowa w informatycznej infrastrukturze nauki*. Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, 283 s.

- Ranking Web of Repositories. Europe* [dok. elektr.]. <http://repositories.webometrics.info/en/Europe> [odczyt: 5.09.2014].
- Sapa, Remigiusz (2013). *Realizacja funkcji repozytoryjnych przez największe przedsięwzięcia zarejestrowane w Federacji Bibliotek Cyfrowych tworzone i współtworzone przez uczelnie*. Przegląd Biblioteczny R. 81, Z. 2, s. 117–132.
- Szafrąński, Leszek; Brawuski Bartłomiej (2014). *Bibliografia Pracowników UJ w Repozytorium UJ* [dok. elektr.]. Kraków: Biblioteka Jagiellońska. <http://jbc.bj.uj.edu.pl/dlibra/docmetadata?id=277960&from=publication> [odczyt: 6.10.2014].
- Tsinghua Tongfang Knowledge Network Technology Co.* (2006) [dok. elektr.]. <http://en.cnki.com.cn> [odczyt: 5.09.2014.]
- Werla, Marcin (2013). *Tworzenie repozytoriów instytucjonalnych w oparciu o system dLibra. Czym różnią się od siebie biblioteki cyfrowe i repozytoria instytucjonalne?* [dok. elektr.]. <https://confluence.man.poznan.pl/community/pages/viewpage.action?pageId=26051667> [odczyt: 12.09.2014].

---

# **Infrastruktura dla eHumanistyki. Praktyczne aspekty gromadzenia i udostępniania źródeł w postaci cyfrowej na przykładzie repozytoriów i bibliotek cyfrowych w sieci PIONIER**

Cezary Mazurek, Tomasz Parkoła, Maciej Stroiński,  
Marcin Werla, Jan Węglarz

Poznańskie Centrum Superkomputerowo-Sieciowe

---

## **Wprowadzenie**

W 2014 roku portal Federacji Bibliotek Cyfrowych w sieci PIONIER obejmował metadane opisujące ponad dwa miliony obiektów cyfrowych zebranych z serwisów internetowych bibliotek, archiwów i muzeów z całej Polski. To efekt ponad dziesięciu lat intensywnych prac realizowanych przez liczną grupę współpracujących ze sobą instytucji kultury i nauki. Wykorzystywanie w tym procesie bieżących wyników prac badawczo-rozwojowych Poznańskiego Centrum Superkomputerowo-Sieciowego doprowadziło do wytworzenia krajowej infrastruktury bibliotek cyfrowych, wspierającej powszechny dostęp do dziedzictwa kulturowego i ułatwiającej wykorzystanie tego dziedzictwa w nauce i edukacji. W artykule przedstawiono obecny stan rozwoju polskiej sieci repozytoriów i bibliotek cyfrowych w kontekście narzędzi dla eHumanistyki, takich jak Wirtualne Laboratorium Transkrypcji, wraz z podsumowaniem najważniejszych doświadczeń zgromadzonych w kolejnych latach rozwoju tej wartościowej infrastruktury. Ponadto na przykładzie trzech projektów rozwojowych (EuropeanaCloud, LoCloud oraz MAN-HA) omówiono potencjał zastosowania technologii chmurowych w sektorze kultury.

Od lat obserwować można dynamiczny wzrost roli technologii informatycznych w różnych dziedzinach nauki. Trend taki jest szczególnie widoczny w sferze nauk ścisłych, gdzie zaawansowana aparatura badawcza generuje w czasie eksperymentów olbrzymie ilości danych źródłowych, które trzeba szybko przechwycić i zapisać, a następnie przetworzyć do postaci przydatnej w analizie i interpretacji

uzyskanych wyników. Skala działania jest tu coraz częściej tak duża, że buduje się infrastruktury międzynarodowe, których nieodzownym elementem są szybkie sieci optyczne zapewniające możliwość rozpraszania danych i obliczeń.

Mogłoby się wydawać, że w naukach humanistycznych i społecznych, w których procesy badawcze znacznie trudniej poddają się algorytmizacji, nie ma potrzeby stosowania na dużą skalę technologii informatycznych, poza – być może – dostarczeniem niezbędnych narzędzi komunikacji naukowej. Warto jednak sprawdzić zasadność tego przypuszczenia. Czy humanistyka potrzebuje zatem infrastruktury informatycznej dużej skali, przystosowanej do składowania i przetwarzania dużych ilości danych?

Podstawą do udzielenia odpowiedzi na to pytanie i zarazem podstawą niniejszego artykułu jest potrzeba efektywnego dostępu do danych źródłowych będących przedmiotem badań humanistycznych. W przypadku nauk ścisłych przyczyną zaangażowania aparatury informatycznej dużej skali jest ogromna ilość danych źródłowych i wymagania względem wysokowydajnych systemów obliczeniowych.

Dla wielu dyscyplin humanistyki danymi źródłowymi są zbiory dziedzictwa kulturowego, a stopień ich dostępności w postaci cyfrowej determinuje rozmiar niezbędnej infrastruktury. W ostatnich latach obserwuje się także rozwój nowych technik służących pozyskiwaniu wiedzy w nauce, gospodarce lub nawet polityce, takich jak analiza, masowe przetwarzanie i wizualizacja danych. Istnieje wyraźny związek pomiędzy rozwojem tych technologii a możliwościami stawiania sobie nowych wyzwań również w dziedzinie infrastruktury służącej przetwarzaniu danych w dziedzinie kultury i humanistyki [Alvaro 2013, dok. elektr.].

Dostępne w ciągu kilku ostatnich lat fundusze krajowe i międzynarodowe umożliwiły wielu polskim instytucjom kultury i nauki cyfryzację ich zbiorów i udostępnianie ich *online*. Uruchomiona w 2002 r. Wielkopolska Biblioteka Cyfrowa była pierwszym w Polsce przykładem stworzenia regionalnej infrastruktury informatycznej do udostępniania zbiorów *online* [Górny et al. 2003]. Powielenie tego modelu w wielu regionach Polski doprowadziło do powstania dużej sieci bibliotek cyfrowych, zintegrowanej dzięki powołaniu w 2007 r. Federacji Bibliotek Cyfrowych, która funkcjonuje w sieci Polski Internet Optyczny PIONIER. Dalszy rozwój następował wielotorowo i obejmował zarówno tworzenie nowych bibliotek cyfrowych i repozytoriów, jak i rozbudowę lub podnoszenie jakości tych już istniejących, a także konstruowanie narzędzi ułatwiających korzystanie z dostępnych *online* zbiorów w celu prowadzenia badań.

W dalszej części artykułu omówiony został obecny kształt polskiej infrastruktury dla eHumanistyki. Zaprezentowano także wybrane doświadczenia wynikające z jej dotychczasowego rozwoju. Następnie przedstawiono potencjał wykorzystania technologii chmurowych w sektorze kultury, przez nakreślenie trzech

różnych problemów występujących obecnie w polskiej sieci bibliotek cyfrowych i wskazanie sposobów ich rozwiązania w ramach aktualnie wdrażanych krajowych i międzynarodowych projektów rozwojowych.

## 1. Wybrane elementy polskiej infrastruktury dla eHumanistyki

We wrześniu 2014 r. liczba obiektów dostępnych przez portal Federacji Bibliotek Cyfrowych (<http://fbc.pionier.net.pl>) przekroczyła dwa miliony. Obiekty te były udostępniane przez ponad 100 serwisów internetowych i pochodziły z kolekcji kilkuset instytucji kultury. Około 80% tych obiektów udostępnionych jest przez szesnaście największych bibliotek cyfrowych, z których jedenaście to inicjatywy regionalne (w nawiasach podano adresy www oraz liczbę udostępnianych obiektów cyfrowych na dzień 19 listopada 2014 r.):

- Cyfrowa Biblioteka Narodowa Polona (<http://www.polona.pl>, 302 tys. obiektów),
- Jagiellońska Biblioteka Cyfrowa (<http://jbc.bj.uj.edu.pl>, 255 tys. obiektów),
- Wielkopolska Biblioteka Cyfrowa (<http://www.wbc.poznan.pl>, 240 tys. obiektów),
- e-biblioteka Uniwersytetu Warszawskiego (<http://ebuw.uw.edu.pl>, 183 tys. obiektów),
- Śląska Biblioteka Cyfrowa (<http://www.sbc.org.pl>, 120 tys. obiektów),
- Małopolska Biblioteka Cyfrowa (<http://mbc.malopolska.pl>, 87 tys. obiektów),
- Kujawsko-Pomorska Biblioteka Cyfrowa (<http://kpbc.umk.pl>, 78 tys. obiektów),
- Biblioteka Cyfrowa - Regionalia Ziemi Łódzkiej (<http://bc.wimbp.lodz.pl>, 54 tys. obiektów),
- Biblioteka Cyfrowa Uniwersytetu Wrocławskiego (<http://www.bibliotekacyfrowa.pl>, 51 tys. obiektów),
- Repozytorium Cyfrowe Instytutów Naukowych (<http://rcin.org.pl>, 50 tys. obiektów),
- Elbląska Biblioteka Cyfrowa (<http://dlibra.bibliotekaelblaska.pl>, 44 tys. obiektów),
- Pomorska Biblioteka Cyfrowa (<http://pbc.gda.pl>, 41 tys. obiektów),
- Bałtycka Biblioteka Cyfrowa (<http://bibliotekacyfrowa.eu>, 41 tys. obiektów),
- Podlaska Biblioteka Cyfrowa (<http://pbc.biaman.pl>, 33 tys. obiektów),
- Zachodniopomorska Biblioteka Cyfrowa „Pomerania” (<http://zbc.ksiaznica.szczecin.pl>, 30 tys. obiektów),
- Radomska Biblioteka Cyfrowa (<http://bc.mbpradom.pl>, 30 tys. obiektów).

Duża liczba serwisów udostępniających zbiory dziedzictwa kulturowego jest oczywiście naturalnym odzwierciedleniem dużego rozproszenia samych zbiorów, ale też i efektem stosunkowo dużej łatwości budowania bibliotek cyfrowych. Łatwość ta w znacznej mierze wynika z popularności i dostępności rozwijanego przez Poznańskie Centrum Superkomputerowo-Sieciowe (PCSS) oprogramowania dLibra. To oprogramowanie jest technologiczną podstawą większości polskich bibliotek cyfrowych. Umożliwia tworzenie zarówno małych archiwów cyfrowych, takich jak np. Publiczne Archiwum Cyfrowe Agnieszki Osieckiej (<http://www.archiwumagnieszkosieckiej.pl/>), jak i dużych bibliotek czy repozytoriów, chociażby tych wymienionych na powyższej liście (z wyłączeniem CBN Polona). Nie mniejsze znaczenie ma też fakt, że usługi bibliotek cyfrowych były od dawna traktowane jako istotny i pożądany element polskiej infrastruktury informatycznej dla nauki – koncepcję ich budowy zawarto w zainicjowanym w roku 2000 programie pn. „PIONIER Polski Internet Optyczny – Zaawansowane Aplikacje, Usługi i Technologie dla Społeczeństwa Informacyjnego” realizowanym w latach 2001-2005 [Rychlewski et al. 2000].

Tak wysoka technologiczna homogeniczność środowiska polskich bibliotek cyfrowych, uzyskana dzięki skoordynowanemu (ale nie centralnemu) działaniu, w połączeniu z wysoką interoperacyjnością oprogramowania dLibra (m.in. dostęp do danych przez protokoły OAI-PMH i OAI-ORE) ułatwiły zintegrowanie polskich bibliotek cyfrowych już na bardzo wczesnym etapie ich rozwoju. Początkowo wszystkie biblioteki cyfrowe wymieniały się danymi między sobą i każda z nich posiadała również informacje o zbiorach wszystkich pozostałych bibliotek. Z czasem jednak, gdy łączny wolumen obiektów zaczął być liczony w dziesiątkach tysięcy, niezbędne okazało się uruchomienie jednego centralnego punktu integrującego rozproszone zasoby, którym stała się wspomniana wcześniej Federacja Bibliotek Cyfrowych. Inicjatywa ta była początkowo przede wszystkim wyszukiwarką dającą dostęp do indeksu metadanych rozproszonych zbiorów, z czasem jednak zaczęła się wzbogacać o kolejne funkcje, takie jak:

- serwis informacyjny dający dostęp do wiadomości i opracowań, związanych z tematyką polskich bibliotek cyfrowych;
- kursy e-learningowe przeznaczone dla małych instytucji kultury, mówiące o tym, jak budować biblioteki cyfrowe i jak współpracować z Europejską Biblioteką Cyfrową Europeana;
- serwis pytań i odpowiedzi umożliwiający każdemu internaucie zadawanie pytań związanych z tematyką bibliotek cyfrowych, na które odpowiadać mogą inni internauci oraz eksperci.

Widać więc, że Federacja Bibliotek Cyfrowych stale się rozwija, stając się czymś na kształt wirtualnego centrum kompetencji w zakresie bibliotek cyfrowych, bu-



dowanego wokół zintegrowanych zasobów polskich instytucji kultury i nauki [Mazurek et al. 2011].

Naturalnym kolejnym etapem rozwoju opisywanej tu infrastruktury jest tworzenie narzędzi pozwalających na prowadzenie w środowisku cyfrowym badań na zdigitalizowanych materiałach źródłowych. Bardzo dobrym przykładem w tym względzie jest Wirtualne Laboratorium Transkrypcji (WLT, <http://wlt.synat.pcss.pl/>) powstałe w PCSS w ramach projektu SYNAT. Laboratorium to umożliwia każdemu zarejestrowanemu użytkownikowi tworzenie i prowadzenie indywidualnych bądź zespołowych projektów transkrypcyjnych. Cykl życia takiego projektu jest następujący [Dudczak et al. 2013]:

1. Utworzenie projektu i opisanie go podstawowymi metadanymi.
2. Wgranie materiału do transkrypcji poprzez dostarczenie plików ze skanami bądź import obiektu bezpośrednio z biblioteki cyfrowej.
3. Uruchomienie procesu OCR w oparciu o dostępną w WLT bezpłatną usługę OCR.
4. Określenie uprawnień innych użytkowników do projektu (w przypadku projektów zespołowych).
5. Właściwa transkrypcja i adnotacja tekstu.
6. Korekta i akceptacja ostatecznych wyników (zwłaszcza w przypadku projektów zespołowych).
7. Eksport danych wynikowych (skany + OCR + adnotacje) do dalszego wykorzystania.

Ważnym elementem jest tutaj widoczna w kroku numer 2 możliwość importu danych bezpośrednio z bibliotek cyfrowych. Pokazuje to, że mamy do czynienia z szeregiem zintegrowanych narzędzi, które oferują udostępnianie zbiorów, konwersję między formatami (na potrzeby wykorzystania w WLT), a także usługę OCR i zespołowe środowisko transkrypcyjne. Co więcej, dzięki aplikacji Wycinanki (<http://wlt.synat.pcss.pl/cutouts>) istnieje możliwość trenowania silnika OCR i używania wypracowanych w ten sposób profili rozpoznawania znaków do poprawiania wyników procesu OCR dla wybranych dokumentów (np. starodruków).

Opisany w tym artykule fragment infrastruktury informatycznej dla eHumanistyki pokazuje, że w Polsce istnieje obecnie kompleksowa, wielowarstwowa aparatura informatyczna, która oferuje naukowcom humanistom wsparcie w dostępie do zbiorów, w przeszukiwaniu dostępnych materiałów i w ich dalszym wykorzystaniu w pracy naukowej. Oczywiście nie jest to infrastruktura zamknięta i z pewnością następować będzie jej dalszy rozwój zarówno w wymiarze horyzontalnym (zwiększanie spektrum dostępnych zasobów i usług), jak i wertykalnym (powstawanie nowych usług opartych na usługach już istniejących – np. usług pozwalających na dalsze analizy tekstów transkrypcji opracowanych w WLT).

Zebrane do tej pory doświadczenia wynikające z budowy opisanej powyżej infrastruktury pozwalają na wyciągnięcie pewnych wniosków i wskazanie pozytywnych czynników, które były kluczowe w dotychczasowym rozwoju, ale także tych efektów ubocznych, które pojawiły się w sposób zupełnie niepożądany. Jest to przedmiotem następnej części artykułu.

## **2. Praktyczne doświadczenia z rozwoju sieci bibliotek cyfrowych w sieci PIONIER**

Z pewnością jednym z najważniejszych czynników, które ukształtowały krajobraz polskich bibliotek cyfrowych, jest model regionalnej współpracy zainicjowany na początku bieżącego stulecia przez Poznańską Fundację Bibliotek Naukowych i PCSS podczas budowania Wielkopolskiej Biblioteki Cyfrowej. Model ten pojawił się w realiach, w których nie było właściwie możliwości uzyskania odpowiednich środków na sfinansowanie takich przedsięwzięć, jak masowa digitalizacja czy udostępnianie zbiorów *online*. Działania te były wówczas widziane raczej jako domena entuzjastów, a zarazem niekoniecznie istotny dodatek do podstawowych zadań instytucji kultury. Wdrożenie modelu, w którym grupa instytucji osadzona w tym samym lokalnym czy regionalnym kontekście kulturowym mogła współdzielić jedną infrastrukturę informatyczną do udostępniania zbiorów *online* i promować wspólnie jeden punkt dostępu do tych zbiorów, okazało się optymalnym sposobem działania [Mazurek et al. 2005]. Z punktu widzenia pojedynczej instytucji podejście takie pozwalało na obniżenie kosztów bieżącego utrzymania biblioteki cyfrowej, likwidowało właściwie barierę wejścia w postaci konieczności poczynienia inwestycji we własną infrastrukturę sprzętowo-programową i dawało naturalne środowisko do współpracy i wymiany wiedzy, pozytywnie wpływające na podnoszenie jakości realizowanych działań. Jedynym wymaganiem w tym modelu było zaakceptowanie wspólnej regionalnej „marki”, pod którą zbiory były udostępniane *online*, ale to okazało się czymś naturalnym i nierodzącym poważniejszych konfliktów. Ambicje poszczególnych instytucji były zaspokajane przez jasne wyrażanie w opisach poszczególnych obiektów, skąd dany obiekt pochodzi i kto odpowiada za jego udostępnienie.

To, iż w pierwszym konsorcjum regionalnym jednym z kluczowych partnerów był silny publiczny badawczo-rozwojowy ośrodek informatyczny, jakim jest PCSS, doprowadziło do wczesnego wypracowania platformy technologicznej, która była później wykorzystywana przez kolejne przedsięwzięcia regionalne i instytucjonalne. W początkowym okresie rozwoju bibliotek cyfrowych w Polsce (w przybliżeniu lata 2002-2005) nie było właściwie na rynku żadnego oprogramowania, które

byłoby rozwijane z myślą o potrzebach instytucji kultury, w modelu nastawionym na wiele potencjalnych wdrożeń. Wypracowany przez PCSS system dLibra był pod tym względem rozwiązaniem unikatowym. Dzięki temu, że powstawał w Polsce, użytkownicy – pracownicy polskich instytucji kultury i nauki – mogli liczyć na lokalne wsparcie techniczne oraz dostęp do dokumentacji w języku polskim i mieć realny wpływ na kształtowanie się tego systemu w miarę wydawania jego kolejnych wersji. W krajowym środowisku użytkowników, które skupiło się wokół systemu, możliwe było wypracowywanie wspólnych kierunków rozwojowych i współdzielenie kosztów rozwoju oprogramowania – funkcje realizowane na potrzeby jednej instytucji były włączane do nowej wersji systemu i mogły z nich korzystać wszystkie inne zainteresowane podmioty. Pozwoliło to też uniknąć typowych czynników ryzyka występujących w przedsięwzięciach realizowanych w trybie projektowym, takich jak:

- „uwięzienie” w oprogramowaniu, które opracowano samodzielnie wewnątrz instytucji kultury w ramach tego projektu, a którego rozwój po zakończeniu projektu został zatrzymany;
- „uwięzienie” w oprogramowaniu opracowanym przez komercyjną firmę na potrzeby danego projektu, w sytuacji gdy po sfinalizowaniu projektu firma przestaje rozwijać taki produkt, np. ze względu na zmianę głównego obszaru działalności lub mniejszą niż zakładana rentowność produktu;
- „uwięzienie” w komercyjnym oprogramowaniu o zamkniętym kodzie źródłowym, w sytuacji gdy dostawca oprogramowania po względnie tanim wdrożeniu systemu zaczyna podnosić ceny na wszelkie usługi dodatkowe, wiedząc że szansa na zmianę oprogramowania przez klienta jest bardzo niska.

Czynniki pozytywne stanowiące przeciwwagę dla powyższych czynników ryzyka mogły zaistnieć właściwie tylko dlatego, że kolejne biblioteki cyfrowe były inicjowane na zasadzie działań oddolnych, bez próby narzucania rozwiązań przez instytucje centralne, a koordynacja strony technologicznej prowadzona była – i nadal jest – w ramach działań związanych z budową ogólnopolskiej infrastruktury informatycznej na rzecz nauki przez konsorcjum złożone z 22 jednostek wiodących – operatorów sieci miejskich MAN i centrów komputerów dużej mocy (KDM). Co ciekawe, również po uruchomieniu centralnego finansowania digitalizacji (WPR Kultura+) rola instytucji centralnych ograniczyła się właściwie do zarządzania strumieniami finansowania, publikowania wytycznych dotyczących standardów realizacji projektów digitalizacyjnych i odbierania wyników tych projektów (obowiązkowo przekazywanej kopii plików wzorcowych). Było to zapewne po części spowodowane również tym, że system regionalnych bibliotek cyfrowych zdążył się mocno ustabilizować i dojrzeć, zanim pojawiły się pomysły centralnego zarządzania przedsięwzięciami z zakresu digitalizacji.

Mimo omówionych powyżej zalet, takie rozproszone podejście miało również pewne efekty uboczne, których dobrze byłoby uniknąć, gdyby to było możliwe. Jednym z takich efektów jest z pewnością różnorodność schematów metadanych w polskich bibliotekach cyfrowych. Początkowo wszystkie biblioteki oparte na oprogramowaniu dLibra zaczynały funkcjonowanie z 15-elementowym schematem Dublin Core Metadata Element Set, z czasem jednak wiele z nich rozbudowało ten schemat na własne potrzeby. Skutkiem było powstanie kilkudziesięciu odmian schematu Dublin Core, co z kolei utrudniło agregację metadanych na poziomie Federacji Bibliotek Cyfrowych. Obecnie problem ten próbuje się rozwiązać przez wdrażanie obszerniejszego schematu metadanych PLMET, który powstał na podstawie analizy schematów stosowanych w poszczególnych polskich bibliotekach cyfrowych. Innym problemem w dłuższej perspektywie okazała się popularność formatu DjVu. Format ten zaczęto wykorzystywać w początkowych latach funkcjonowania Wielkopolskiej Biblioteki Cyfrowej. W tym okresie był on faktycznie formatem oferującym unikalnej jakości kompresję obrazów i inne cechy, takie jak podział dokumentu na warstwy czy przechowywanie zarówno obrazów (skanów), jak i tekstu (wyników OCR). Z czasem jednak format PDF oferował coraz lepsze możliwości, a format DjVu był coraz słabiej wspierany przez głównego producenta oprogramowania do obsługi dokumentów w tym formacie. Do tego doszły jeszcze obostrzenia zabezpieczeń wprowadzane przez przeglądarki internetowe, które spowodowały znaczne utrudnienia w dostępie do publikowanych *online* materiałów w formacie DjVu. Niestety wiele z nowo powstających czy nawet już funkcjonujących bibliotek cyfrowych nie analizowało aktualnej sytuacji dotyczącej formatów udostępniana zbiorów, tylko decydowało się na format DjVu, ponieważ cieszył się on opinią najpopularniejszego w polskich bibliotekach cyfrowych. Obecnie na szczęście można zaobserwować coraz częstsze odchodzenie od tego formatu na rzecz publikowania dokumentów w formacie PDF lub po prostu w postaci plików graficznych z możliwością skalowania.

Przyjęty w Polsce model regionalnych bibliotek cyfrowych dobrze sprawdził się w przypadku instytucji kultury różnej skali – od małych, które mogą przyłączać się do działających już inicjatyw często praktycznie bez ponoszenia kosztów, aż po te największe, które budując regionalne konsorcja, mogą zaprosić innych do korzystania z własnej infrastruktury i obniżyć tym samym koszty funkcjonowania w sieci. Ponadto instytucje kultury zyskują nowe możliwości za sprawą coraz bardziej popularnych technologii chmurowych, które zostały omówione w następnej części artykułu.

### 3. Potencjał technologii chmurowych w sektorze kultury

Poniżej przedstawione są trzy wybrane kierunki rozwojowe realizowane w latach 2013-2015, które pokazują, jak różnorodne problemy sektora kultury można rozwiązywać dzięki technologiom chmurowym. Zróżnicowany zakres tych projektów, połączony aspektami technologicznymi, pokazuje że omówiona wcześniej zaawansowana infrastruktura udostępniania zbiorów ma cały czas znaczny potencjał rozbudowy i dzięki postępowi technologicznemu może być rozwijana w kierunkach, które jeszcze kilka lat temu zdawały się wymagać działań o nadmiernej złożoności.

#### 3.1. Usługi dla małych instytucji kultury – LoCloud

Celem projektu LoCloud (<http://locloud.eu>) jest opracowanie zestawu narzędzi i usług opartych na technologiach chmurowych, które umożliwiają bardzo małym instytucjom kultury udostępnianie zbiorów *online* i przekazywanie informacji o tych zbiorach do Europeany. Chodzi tu nie tylko o instytucje publiczne, ale również o projekty organizacji pozarządowych czy wręcz prywatne muzea i archiwa. Tak małe instytucje często nie są w stanie ponieść kosztów jakiegokolwiek infrastruktury informatycznej ani też kosztów zatrudnienia wykwalifikowanego personelu IT. Dlatego celem projektu jest dostarczenie bardzo prostych w obsłudze usług, które pozwolą zbiorom tych instytucji zaistnieć w sieci.

Jedną z takich usług jest LoCloud Collections (<http://locloud.pl/>), dzięki której możliwe jest uruchomienie w ciągu kilku minut własnego systemu kolekcji cyfrowych i udostępnianie *online* cyfrowych wizerunków własnych zbiorów. Użytkownicy LoCloud Collections mogą samodzielnie konfigurować swój system kolekcji cyfrowych, a także określać, jak duże jest ich zapotrzebowanie na przestrzeń dyskową w chmurze LoCloud. Dzięki temu ponoszą tylko koszty w zakresie niezbędnym do funkcjonowania w danym momencie. LoCloud Collections zostało opracowane tak, aby spełniać wszystkie podstawowe potrzeby związane z udostępnianiem zbiorów *online*, a jednocześnie być systemem łatwym w obsłudze. Dowodem na to może być fakt, iż pełne przeszkolenie z obsługi tego systemu zajmuje około godziny. Dla porównania pełne szkolenie z systemu dLibra trwa około dwóch dni. Ponadto w ramach projektu funkcjonuje dedykowane centrum wsparcia dla małych instytucji kultury (<http://support.locloud.eu/>).

Technologie chmurowe w tym projekcie dają dużą elastyczność w tworzeniu kolejnych kolekcji cyfrowych i pozwalają na dynamiczne przydzielanie zasobów (mocy obliczeniowej, przestrzeni dyskowej) zgodnie z zapotrzebowaniem poszczególnych użytkowników – małych instytucji kultury.

### **3.2. Usługi przetwarzania danych dla bibliotek cyfrowych – MAN-HA**

Celem projektu MAN-HA (<http://man-ha.pionier.net.pl/>) jest podniesienie niezawodności usług świadczonych w sieci PIONIER oraz zaoferowanie nowych rodzajów usług w trybie wysokiej dostępności. Usługi te są skierowane przede wszystkim do środowiska naukowego i obejmują m.in. utworzenie polskiej federacji zarządzania tożsamością (<http://aai.pionier.net.pl/>), udostępnienie chmurowej bazy danych czy stworzenie centralnego punktu dostępu do wybranych usług w sieci PIONIER. Jednym z zadań w ramach projektu MAN-HA jest też opracowanie i udostępnienie usług przetwarzania danych dla bibliotek cyfrowych. Usługi te mają wykorzystywać powstałą w ramach projektu PLATON chmurę dedykowaną do obliczeń kampusowych, w celu zaoferowania m.in. możliwości masowego konwertowania danych i metadanych czy przetwarzania OCR. Pierwszym obszarem wykorzystania tych usług będzie próba zbudowania pełnotekstowego indeksu wyszukiwawczego obejmującego zbiory Federacji Bibliotek Cyfrowych.

Wykorzystanie technologii chmurowych pozwoli na dynamiczne przydzielanie zasobów obliczeniowych z chmury obliczeń kampusowych w zależności od tego, jakie będzie w danym momencie zapotrzebowanie na obliczenia ze strony usług przetwarzania danych dla bibliotek cyfrowych i jaka będzie dostępność mocy obliczeniowej.

### **3.3. Usługi dla agregacji metadanych – Europeana Cloud**

Celem projektu EuropeanaCloud (<http://pro.europeana.eu/web/europeana-cloud>) jest opracowanie współdzielonego systemu składowania i przetwarzania danych, z którego korzystać będą mogli tzw. agregatorzy metadanych, czyli inicjatywy takie jak Federacja Bibliotek Cyfrowych, które gromadzą metadane obiektów z wielu różnych źródeł, przetwarzają je na własne potrzeby i przekazują do Europeany. Koncepcja projektu zakłada, że w początkowej jego fazie trzej agregatorzy (Europeana, The European Library oraz Federacja Bibliotek Cyfrowych) zbudują wspólną infrastrukturę informatyczną i zaczną z niej korzystać na własne potrzeby, a następnie będą promować wykorzystanie usługi wśród instytucji będących operatorami kolejnych serwisów agregujących metadane. Poza potencjalnym obniżeniem kosztów utrzymania infrastruktury informatycznej agregatorów znaczącym zyskiem z przystąpienia do takiego współdzielonego systemu będzie łatwość wymiany danych pomiędzy poszczególnymi partnerami. Będzie to znacznie prostsze, ponieważ zniknie konieczność kopiowania danych np. z Federacji Bibliotek Cyfrowych do Europeany – każdy z uczestników infrastruktury EuropeanaCloud będzie miał natychmiastowy dostęp do danych upubliczniczonych w chmurze przez innego użytkownika.

Dzięki zastosowaniu zaawansowanych technologii chmurowych stworzona w projekcie EuropeanaCloud infrastruktura będzie niezawodna i wysoko dostępna, rozproszona geograficznie, a także wysoce skalowalna. Są to niezbędne cechy systemu informatycznego, który ma potencjalnie funkcjonować w tak dużej skali.

#### 4. Podsumowanie

W niniejszym artykule omówione zostały wybrane elementy polskiej infrastruktury informatycznej dla eHumanistyki, które służą udostępnianiu zbiorów instytucji kultury, ułatwiają przeszukiwanie tych zbiorów i pozwalają na wykorzystanie ich w pracy naukowej. Przedstawiono również kluczowe czynniki, które w sposób pozytywny wpłynęły na rozwój tej infrastruktury i umożliwiły jej osiągnięcie obecnego stanu. Ponadto zasygnalizowany został istotny potencjał technologii chmurowych w sektorze kultury, które mogą w najbliższych latach przynieść znaczne udogodnienia i nowe możliwości dla różnej wielkości instytucji, od tych najmniejszych, takich jak prywatne lokalne muzea czy archiwa, poprzez średnie, poszukujące mocy obliczeniowej do przetwarzania danych i metadanych, po te największe o krajowej i międzynarodowej skali działania.

Główne wyzwania, jakie staną przed instytucjami kultury w najbliższej przyszłości, dotyczą z pewnością aspektów takich, jak zarządzanie jakością w masowych projektach digitalizacyjnych, długoterminowe składowanie danych, interoperacyjność na poziomie krajowym i międzynarodowym czy wreszcie kwestii prawnych związanych z licencjonowaniem i zapewnianiem otwartego dostępu do treści i metadanych czy udostępnianiem dzieł osieroconych. Z pewnością też coraz wyraźniej będzie widać trend projektowania i budowania bibliotek cyfrowych jako kompleksowych systemów do zarządzania kolekcjami obiektów cyfrowych, które nie pozostają skupione wyłącznie na prezentacji *online*, ale kompleksowo pokrywają cały proces digitalizacyjny i zarządzanie wynikami tego procesu po jego zakończeniu, uwzględniając w tym zarówno udostępnianie, jak i długoterminową archiwizację. Pierwsze wdrożenia tego typu systemów mają już miejsce w Polsce (por. Repozytorium Cyfrowe Instytutów Naukowych RCIN, <http://www.rcin.org.pl>).

Rozwój infrastruktury dla eHumanistyki to także działania mające na celu pobudzenie rozwoju sektora rynku treści. Komisja Europejska zwraca uwagę w jednym z projektów przewodnich pn. „Europejska Agenda Cyfrowa”, który jest integralną częścią „Strategii na rzecz inteligentnego i zrównoważonego rozwoju sprzyjającego włączeniu społecznemu – Europa 2020” [Komunikat Komisji Europa 2020..., dok. elektr.], na powszechne oczekiwanie, że dostęp do treści w inter-

necie będzie równie łatwy, jak w świecie rzeczywistym. Zdaniem Komisji Europejskiej brakuje jednolitego rynku sektora treści. Należy podkreślić, iż polska infrastruktura dla eHumanistyki, a w szczególności zasoby i usługi bibliotek cyfrowych, to także ogromny potencjał dla branż kreatywnych przedsiębiorstw i twórców.

## BIBLIOGRAFIA

- Alvaro, Sandro (2013). *Big Data and Digital Humanities: From social computing to the challenges of connected culture* [dok. elektr.]. [http://blogs.cccb.org/lab/en/article\\_big-data-i-humanitats-digitals-de-la-computacio-social-als-reptes-de-la-cultura-connectada/](http://blogs.cccb.org/lab/en/article_big-data-i-humanitats-digitals-de-la-computacio-social-als-reptes-de-la-cultura-connectada/) [odczyt: 10.10.2014].
- Dudczak, Adam et al. (2013). *Improving the Workflow for Creation of Textual Versions of Polish Historical Documents*. In: Bembenik, Robert et al. eds. *Intelligent Tools for Building a Scientific Information Platform: Advanced Architectures and Solutions*. Berlin, Heidelberg: Springer, p. 187-198.
- Górny, Mirosław et al. (2003). *Zastosowanie oprogramowania dLibra do budowy Wielkopolskiej Biblioteki Cyfrowej*. W: Zeszyty Naukowe Wydziału ETI Politechniki Gdańskiej. Technologie Informacyjne. Gdańsk: Wydawnictwo Politechniki Gdańskiej, 2003. I Krajowa Konferencja Technologie Informacyjne, Gdańsk, 18-21 maja 2003 roku. s. 109-117.
- Komunikat Komisji KOM(2010) 2020 wersja ostateczna. *EUROPA 2020. Strategia na rzecz inteligentnego i zrównoważonego rozwoju sprzyjającego włączeniu społecznemu* [dok. elektr.]. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:PL:PDF> [odczyt: 18.10.2014].
- Mazurek, Cezary et al. (2011). *Rola Federacji Bibliotek Cyfrowych w dostępie on-line do polskiego dziedzictwa kulturowego i naukowego*. W: [Nowakowski, Antoni red.] *Infobazy 2011 – nauka, projekty europejskie, społeczeństwo informacyjne: materiały VI krajowej konferencji naukowej*, Gdańsk–Sopot, 5-7 września 2011 roku. Gdańsk: Centrum Informatyczne TASK, s. 74-79.
- Mazurek, Cezary et al. (2005). *Wdrażanie regionalnych bibliotek cyfrowych w sieci PIONIER w oparciu o środowisko dLibra*. W: [Nowakowski, Antoni red.] *Infobazy 2005 – bazy danych dla nauki*. Gdańsk: Centrum Informatyczne TASK, s. 58-64.
- Rychlewski, Jacek et al. (2000). *PIONIER: Polish Optical Interent*. W: *ISThmus 2000. Materiały z konferencji (11-13 kwietnia 2000)*. Poznań: Instytut Informatyki Politechniki Poznańskiej, s. 19-28.



---

Część III

## **Zasoby bibliograficzne**

---

---

# Rejestracja dorobku pracowników uczelni: Bibliografia Publikacji Pracowników Uniwersytetu Jagiellońskiego

Leszek Szafrąński

Biblioteka Jagiellońska

---

## Wprowadzenie

Obecnie na świecie obserwuje się dynamiczne zmiany w procesie komunikacji naukowej i w dostępie do informacji. Proces obiegu informacji odbywa się w internecie i za pomocą nośników cyfrowych. Uniwersytety i ośrodki naukowe, dostosowując się do zachodzących zmian, tworzą bibliografie swoich pracowników dostępne *online*. W następnym etapie przekształcają się one w repozytoria z pełnymi tekstami dorobku naukowego pracowników.

Głównym celem tworzenia tego typu bibliografii specjalnych jest upowszechnianie, popularyzacja i informowanie o osiągnięciach naukowych uczelni oraz tworzenie dokumentacji działalności naukowej ich pracowników. Dorobek naukowy zatrudnionych badaczy jest także jednym z ważniejszych elementów oceny parametrycznej szkół wyższych. Bibliografia, a zwłaszcza repozytorium, wspomagają również edukację.

Artykuł podzielony jest na dwie części. W pierwszej z nich autor, na podstawie własnych badań i literatury przedmiotu, przeanalizował wszystkie bibliografie publikacji pracowników polskich uczelni wyższych, które udostępniane są *online*. W drugiej części opisał rozwiązania przyjęte przy tworzeniu Bibliografii Publikacji Pracowników Uniwersytetu Jagiellońskiego (BPP UJ).

Bezpośrednim powodem podjęcia analizy bibliografii pracowników uczelni w Polsce była budowa systemu do obsługi BPP UJ. Podobne badania prowadziła już między innymi Maria Garczyńska, która, na podstawie wyników ankiet, poddała analizie 53 bibliografie szkół wyższych i ustaliła, że: 62% z nich powstało na mocy wewnętrznych uregulowań prawnych władz uczelni, 74% opracowy-

wanych było przez bibliotekarzy, w większości w oddziałach informacji naukowej, 28 zostało udostępnionych *online*. Bazy tworzone były na oprogramowaniu: Expertus, ALEPH, Prolib, MAK BN, SOWA2, SQL [Garczyńska 2006, s. 64-67, 70-71]. Tym tematem zajmowała się również Aleksandra Legeżyńska (później Zawadzka), która w swoim opracowaniu z 2006 roku skupiła się na badaniu bibliografii 18 uniwersytetów polskich, między innymi pod kątem dostępu *online*, wykorzystywanego oprogramowania i typów dokumentów [Legeżyńska 2006, s. 61]. W kontynuacji badań z roku 2011 uwzględniała kryteria dostępu *online* i wykorzystanego oprogramowania. W 2005 r. siedem, a w 2011 r. już 12 uniwersytetów posiadało bibliografie pracowników dostępne *online*. Bazy były tworzone na oprogramowaniu: Expertus, ALEPH, Promax, KOHA, MAK BN, SKEP lub w oparciu o własną przeglądarkę. W tym czasie nastąpił wzrost liczby użytkowników systemu Expertus [Zawadzka 2011, s. 69].

Iwona Sójkowska i Barbara Barańska-Malinowska na podstawie ankiet zbadały natomiast 17 bibliografii uczelni technicznych. Z przeprowadzonej przez autorki analizy wynika, że w tego typu szkołach tworzone są własne systemy do rejestrowania danych i większość baz działa już od końca lat 80. XX wieku. Dane rejestrowane są również w systemach Expertus i ALEPH. W zastosowanej typologii dokumentów zostały wyróżnione materiały opublikowane i inne typy dokumentów, np.: rozprawy doktorskie, raporty i sprawozdania z badań, zgłoszenia patentowe [Sójkowska, Barańska-Malinowska 2011, s. 22-24].

Szersze analizy przeprowadził Remigiusz Sapa, który omówił w swojej pracy bibliografie pracowników czołowych polskich uczelni wyższych pod kątem dostępu do pełnych tekstów publikacji. Na 20 badanych bibliotek uczelnianych tylko 6 umożliwiało taki dostęp [Sapa 2010, s. 83]. Natomiast Barbara Woźniak w swojej prezentacji przedstawiła dynamikę rozwoju bibliografii instytucjonalnych. Ustaliła, że w większości baz spisy są niekompletne i zróżnicowana jest w nich szczegółowość informacji [Woźniak 2013, dok. elektr.].

## **Bibliografie uczelni w internecie**

W omówionych poniżej badaniach autor przeanalizował bibliografie publikacji pracowników na podstawie stron internetowych uczelni i bibliotek uczelnianych w Polsce. Do analizy porównawczej zostały wybrane bibliografie, które są udostępniane w internecie. W badaniu nie zostały uwzględnione bibliografie opracowywane przez pojedyncze jednostki uczelni, np.: wydziały, instytuty itp. oraz bibliografie zawieszane z powodu przeniesienia danych do nowszej wersji systemu, np. „Bibliografia Publikacji Pracowników Akademii Wychowania Fi-

zycznego we Wrocławiu do 2001 r.” Z analizy wyłączono także repozytoria uczelniane. Bazy były weryfikowane według sześciu kryteriów: wykorzystywanego oprogramowania, formatu prezentacji opisu bibliograficznego, typów dokumentów, możliwości wydruku, jednostki zajmującej się opracowywaniem bibliografii w danej uczelni, sposobu zapisu danych w alfabetycznych niełacińskich.

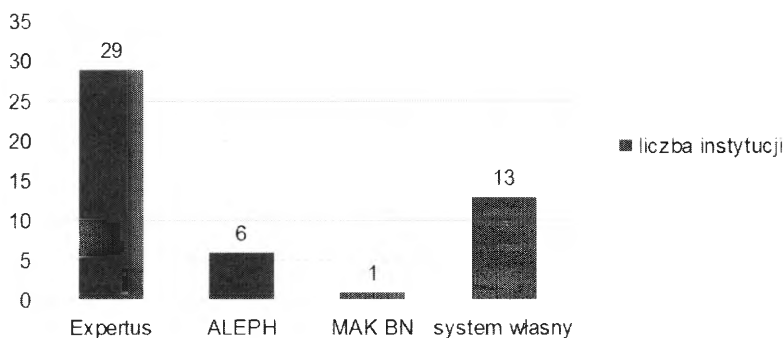
Badaniem zostało objętych 49 bibliografii, które są udostępniane w internecie. Były to bibliografie uniwersytetów, uczelni technicznych, uczelni medycznych, uczelni wojskowych, akademii, akademii wychowania fizycznego oraz po jednej bibliografii uczelni artystycznej i uczelni morskiej. Na podstawie zebranych danych i literatury przedmiotu autor ustalił, że tworzeniem bibliografii pracowników uczelni w 94% badanych instytucji zajmują się biblioteki. Tylko w kilku przypadkach bibliografię opracowują inne jednostki uczelni, np. na Politechnice Świętokrzyskiej dane wpisywane przez pracowników do systemu obsługującego bibliografię weryfikują osoby wyznaczone przez dziekanów poszczególnych wydziałów, na Uniwersytecie Warszawskim bibliografię tworzy Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego, na Politechnice Rzeszowskiej systemem administruje Samodzielna Sekcja Rozwoju Kadry Naukowej.

W 51% badanych bibliotek uzupełnianiem i tworzeniem bibliografii zajmują się oddziały informacji naukowej, działy bibliograficzne lub wydzielone specjalnie do tego zadania sekcje lub zespoły ds. bibliografii, dokumentacji lub rejestracji dorobku pracowników. Powierzenie tego typu prac oddziałom informacji wydaje się najbardziej uzasadnione. W czterech przypadkach oddziały informacji i bibliotekę główną wspomagają biblioteki poszczególnych wydziałów uczelni [Garczyńska 2006, s. 68]. W Bibliotece Politechniki Łódzkiej system bibliograficzny obsługuje Oddział Zbiorów Elektronicznych. Na Uniwersytecie Ekonomicznym w Poznaniu bibliografię tworzy Sekcja Szkoleń, Baz Danych i Digitalizacji, natomiast na Politechnice Wrocławskiej – Sekcja Analiz Dorobku Naukowego. Nie udało się uzyskać danych w Akademii im. Jana Długosza w Częstochowie, ponieważ w trakcie badania strona internetowa bibliografii była w budowie.

W analizowanych bibliografiach dane zapisane alfabetami niełacińskimi podawane są transkrypcji lub transliteracji na alfabet łański. Wyjątkami są Politechnika Lubelska, Politechnika Świętokrzyska i Uniwersytet Warszawski, gdzie dane bibliograficzne prac pisanych w cyrylicy nie są transliterowane, tylko podawane w oryginalnej formie. Nie jest to zasada konsekwentnie stosowana w bibliografiach tych uczelni, pozostałe języki są tam bowiem transliterowane.

Najczęściej wybieranym oprogramowaniem do obsługi bibliografii jest Expertus firmy Splendor. Dane na wykresie 1 wskazują, że ponad 59% uczelni korzysta z tego oprogramowania.

## Systemy do obsługi bibliografii



Wykres 1. Liczba instytucji użytkujących poszczególne systemy bibliograficzne

Źródło: opracowanie własne

Dodać jednak należy, że Uniwersytet Łódzki jest w fazie testowania Expertusa. Do roku 2012 bibliografię obsługiwał tam system MAK Biblioteki Narodowej, który użytkuje również Politechnika Koszalińska. Kolejnym popularnym programem do obsługi bibliografii pracowników jest system ALEPH. Pozostałe biblioteki korzystają z oprogramowania własnego.

W każdej bibliografii stosuje się inną typologię dokumentów. Jest to zależne od tradycji tworzenia spisów bibliograficznych, przyjętych celów i ustaleń na poszczególnych uczelniach, ale również od wytycznych Ministerstwa Nauki i Szkolnictwa Wyższego potrzebnych do oceny parametrycznej instytucji. Różnorodność typów dokumentów w zasadzie nie jest zależna od profilu danej uczelni (technicznej, medycznej, uniwersytetu), wynika raczej z przyjętych założeń na poziomie poszczególnych szkół wyższych. Podobne typy dokumentów mogą zatem wystąpić w bibliografiach pracowników uniwersytetu, politechniki lub uczelni medycznej.

Bardziej szczegółowe badania w tym zakresie przedstawiła Maria Garczyńska w artykule „Bibliograficzne bazy danych w ocenie pracowników i jednostek szkół wyższych w Polsce” [Garczyńska 2013, dok. elektr.]. Według jej opracowania w bibliografiach występuje średnio 28 typów dokumentów, ale ich liczba waha się od 3 do 73. Najbardziej rozbudowana typologia dokumentów jest stosowana w Katalogu Prac Naukowo-Badawczych Politechniki Gdańskiej. Znajdują się tam kategorie niespotykane w innych bibliografiach, między innymi: prawa autorskie do utworów będących wynikiem działalności twórczej w zakresie architektury, urbanistyki oraz sztuki, ekspozycje artystyczne, prezentacje projektów architektonicznych, naukowe publikacje instytutów badawczych i organizacji międzyna-

rodowych. W tabeli 1 przedstawione są wszystkie typy dokumentów, występujące w Katalogu Prac Naukowo-Badawczych Politechniki Gdańskiej.

Tabela 1. Typy dokumentów w Katalogu Prac Naukowo-Badawczych Politechniki Gdańskiej

1. artykuły w czasopismach
2. artykuły w czasopismach dostępnych w wersji elektronicznej [także <i>online</i> ]
3. artykuł w czasopiśmie z listy filadelfijskiej
4. artykuły w czasopismach recenzowanych i innych wydawnictwach ciągłych
5. prawa autorskie do utworów będących wynikiem działalności twórczej w zakresie architektury, urbanistyki oraz sztuki
6. naukowe publikacje instytutów badawczych i organizacji międzynarodowych
7. broszura
8. rozdział w książce – monografii autorskiej / podręczniku o zasięgu krajowym
9. rozdział w książce – monografii autorskiej / podręczniku w języku o zasięgu międzynarodowym
10. rozdział w skrypcie, podręczniku, poradniku, informatorze
11. praca doktorska pracowników zatrudnionych w PG oraz studentów studium doktoranckiego
12. doktorat osób niezatrudnionych w PG
13. publikacje elektroniczne
14. edycja / redakcja czasopism z listy filadelfijskiej, czasopism recenzowanych, wydawnictw zbiorowych oraz dzieł zbiorowych recenzowanych
15. opracowanie hasła encyklopedycznego
16. red. lub edytorstwo monografii lub podręcznika / poradnika o zasięgu krajowym
17. red. lub edytorstwo monografii lub podręczników o zasięgu międzynarodowym
18. redakcja abstraktów i posterów
19. publikacja elektroniczna recenzowana
20. redakcja skryptów
21. ekspertyza
22. katalogi wystaw, foldery, przewodniki i informatory turystyczne
23. artykuł w czasopiśmie z listy filadelfijskiej – wersja <i>online</i> lub CD
24. rola ekspercka, członkostwo w jury, funkcja specjalistyczna
25. grafika książkowa, dzieło artystyczne wydane z prawami autorskimi
26. rozprawa habilitacyjna
27. habilitacja uzyskana przez pracownika spoza PG
28. recenzja publikowana, wstęp, tłumaczenie, wywiad
29. recenzja monografii (dla wydawnictw) niepublikowana
30. artykuł w czasopiśmie indeksowanym TR Master Journal List
31. artykuł w czasopiśmie wyróżnionym w JCR
32. komunikaty, abstrakty, postery
33. komunikaty, abstrakty – wersja elektroniczna
34. know-how
35. druk zwarty (publikacja książkowa, raport, sprawozdanie itp.)
36. rozdział, artykuł w książce – dziele zbiorowym / podręczniku o zasięgu krajowym
37. rozdział, artykuł w książce – dziele zbiorowym / podręczniku w języku o zasięgu międzynarodowym

c.d. tabeli 1

38. książka – dzieło zbiorowe / podręcznik o zasięgu krajowym
39. książka – dzieło zbiorowe / podręcznik w języku o zasięgu międzynarodowym
40. książka – monografia autorska / podręcznik o zasięgu krajowym
41. książka – monografia autorska / podręcznik w języku o zasięgu międzynarodowym
42. norma
43. nostryfikacja doktoratu
44. nostryfikacja habilitacji
45. publikacje ulotne (w gazetach)
46. podręczniki akademickie
47. poradniki, informatory, przewodniki oraz materiały o charakterze informacyjno-szkoleniowym (teoretyczne), encyklopedie, słowniki, leksykony
48. prezentacja multimedialna, prelekcja
49. patent
50. międzynarodowy patent na wynalazek
51. referaty w materiałach konferencyjnych
52. członkostwo lub pełnienie funkcji: jury, ekspert, kuratora wystawy
53. raport końcowy, sprawozdanie zakończenia projektu badawczego
54. referat w materiałach konferencyjnych recenzowanych
55. recenzja naukowa monografii, wydawnictwa zbiorowego i ciągłego (opublikowana)
56. skrypt zawierający materiały o charakterze szkoleniowym (w tym ćwiczenia i zadania)
57. skrypt, konspekt dydaktyczny, materiały szkoleniowe
58. materiały pomocnicze: instrukcje, ćwiczenia, zadania – dostępne także w internecie
59. skrypt uczelniany recenzowany lub o zasięgu ogólnokrajowym, materiały do wykładów (ew.+ ćwicz.)
60. suplement, wydanie specjalne, dodatek
61. tłumaczenie artykułu w czasopiśmie, rozdziału w monografii bądź podręczniku
62. tłumaczenie monografii, podręcznika, skryptu
63. wzór użytkowy
64. redakcja materiałów konferencyjnych recenzowanych
65. wystawa, ekspozycja artystyczna, prezentacja projektów architekt.
66. redakcja materiałów konferencyjnych nierecenzowanych
67. zgłoszenie patentowe
68. opublikowane międzynarodowe zgłoszenie patentowe
69. wdrożenie, prototyp, aplikacja
70. raport, sprawozdanie, praca badawcza
71. prace nb i nu niepublikowane, raport/sprawozdanie niepublikowane w wersji elektronicznej (online, CD-ROM)
72. publikacja w wydawnictwie zbiorowym nierecenzowanym
73. publikacja w wydawnictwie zbiorowym recenzowanym (także w materiałach konferencyjnych)

Źródło: Katalog Prac Naukowo-Badawczych Politechniki Gdańskiej [dok. elektr.]. <http://prace-nb.pg.gda.pl> [odczyt: 10.08.2014]

Poszczególne uczelnie stosują różnorodne formy prezentacji opisu bibliograficznego publikacji. W systemie ALEPH standardowo dane można wyświetlać w pięciu formatach: format standardowy, karta katalogowa, cytata, etykiety nazw, etykiety MARC. W systemie Expertus natomiast można wybrać do 12 różnych formatów prezentacji opisu, co często jest wykorzystywane przez instytucje tworzące bibliografie. W systemach własnych, tworzonych indywidualnie przez uczelnie na potrzeby bibliografii, liczba formatów prezentacji nie przekracza trzech. Z wykresu 2 wynika, że jest to najczęściej spotykane rozwiązanie w analizowanych bibliografiach. Standardowo stosowane formaty to: opis ciągły lub podstawowy, opis tabelaryczny lub katalogowy, lista tytułów lub opis skrócony.



Wykres 2. Liczba form prezentacji opisu bibliograficznego  
w zależności od liczby wystąpień w bibliografiach  
Źródło: opracowanie własne

Bibliografie były analizowane również pod kątem możliwości wydruku. 84% badanych bibliografii posiada tę funkcję, bardziej lub mniej rozbudowaną. Pozostałe nie dają możliwości wydruku w obsługiwanym oprogramowaniu, ale w większości jest możliwość przekonwertowania wyświetlanych metadanych do formatów pdf i rtf, które posiadają opcje druku.

Niektóre bibliografie są połączone bezpośrednio z repozytoriami uczelni, w których umieszczane są pełne teksty publikacji. W ten sposób tworzone są bibliografie między innymi na Politechnice Wrocławskiej, Politechnice Lubelskiej i Politechnice Krakowskiej.



Obecnie uczelnie tworzą bibliografie na podstawie własnych wewnętrznych wytycznych. Jest to widoczne szczególnie przy wyborze typów dokumentów i form prezentacji danych. Najbardziej popularnym systemem do obsługi bibliografii jest Expertus, ale wiele uczelni tworzy systemy autorskie, które mogą być rozbudowywane według własnych potrzeb. Wadą większości programów jest to, że nie dają one możliwości tworzenia repozytoriów pełnych tekstów publikacji, co powoduje konieczność budowy dwóch systemów: jednego do obsługi bibliografii, a drugiego do repozytorium. Naraża to uczelnie na dodatkowe koszty utrzymywania dwóch niezależnych od siebie systemów.

## **Bibliografia Publikacji Pracowników Uniwersytetu Jagiellońskiego**

Uniwersytet Jagielloński rozpoczął w 2013 r. projekt HaZarD (Harmonizacja zarządzania dydaktyką na Uniwersytecie Jagiellońskim w Krakowie), w ramach którego w Bibliotece Jagiellońskiej jest budowany system do zarządzania i tworzenia bibliografii pracowników UJ dostępnej w internecie. Planowane uruchomienie Bibliografii Publikacji Pracowników Uniwersytetu Jagiellońskiego przewidziane jest na ostatni kwartał 2014 r. Główne cele, jakie zostały ustalone na początku projektu, związane były z integracją systemu obsługującego bibliografię z systemami już działającymi na Uniwersytecie i, w dalszym etapie, rozwinięciem funkcjonalności systemu o moduł do obsługi repozytorium. Natomiast cele szczegółowe powstania bibliografii to między innymi:

- ocena parametryczna pracowników i jednostek UJ,
- jeden system do obsługi bibliografii Uniwersytetu,
- zasób wykorzystywany na potrzeby ankietowe i statystyczne uczelni,
- wymiana informacji z systemami wspomagającymi zarządzanie uczelnią,
- widoczność publikacji pracowników w wyszukiwarkach naukowych (Google Scholar, BASE, Scientific Commons, OpenAIRE) i Polskiej Bibliografii Naukowej.

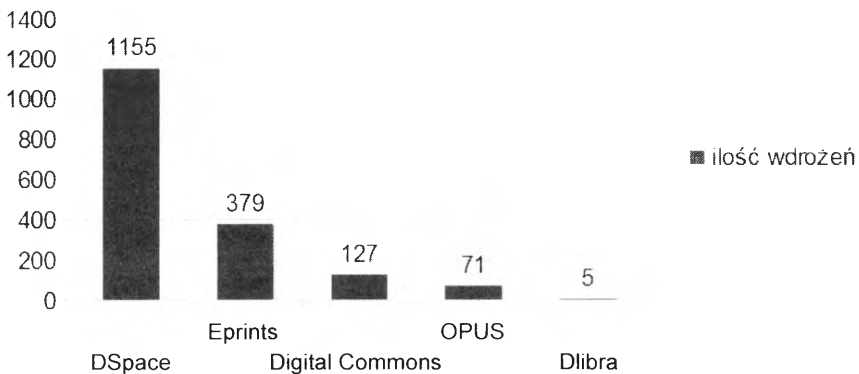
## **Oprogramowanie**

Wybór odpowiedniego systemu do obsługi bibliografii, a w późniejszym okresie całego repozytorium Uniwersytetu, był jedną z najważniejszych decyzji, jakie w ostatnim czasie musiała podjąć Biblioteka Jagiellońska. Spośród wielu systemów komercyjnych i bezpłatnych, które są dostępne obecnie na świecie, należało wybrać taki, który potrafiłby obsłużyć repozytorium i dodatkowo bibliografię pracowni-

ków całego Uniwersytetu, lub napisać własny system od podstaw. Ze względu na ograniczone możliwości finansowe i czasowe nie było warunków do zbudowania własnego systemu, dlatego po wielu konsultacjach z innymi jednostkami uczelni został wybrany system DSpace (<http://www.dspace.org>), do którego, we własnym zakresie, pracownicy biblioteki dobudowali moduł do obsługi bibliografii. System został wybrany z kilku powodów, między innymi ze względu na:

- możliwość rozbudowy i implementacji według własnych potrzeb (oprogramowanie *open source*),
- wydajność systemu (obsługa dużej ilości danych),
- funkcjonalność systemu,
- przechowywanie danych w różnych formatach,
- możliwość tworzenia hierarchicznej struktury bazy,
- możliwość importu i eksportu metadanych z / do baz zewnętrznych,
- zgodność z protokołem OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting),
- popularność na świecie – najbardziej popularne systemy przedstawia wykres 3.

repozytorium - oprogramowanie



Wykres 3. Liczba wdrożeń najbardziej popularnych systemów do obsługi repozytoriów na świecie  
 Źródło: opracowanie własne na podstawie badań własnych i Directory of Open Access Repositories – OpenDOAR [dok. elektr.]. <http://www.opendoar.org> [odczyt: 10.08.2014]

W Bibliotece Jagiellońskiej została zainstalowana wersja 4.2 oprogramowania DSpace. Minusem systemu jest brak wsparcia technicznego. Jest to podstawowa wada programów bezpłatnych. Ten brak rekompensuje strona internetowa projektu, gdzie można wymieniać się doświadczeniami na trzech listach dyskusyj-

nych, blogach i zapoznawać się z dostępnymi tam materiałami szkoleniowymi. Ze względu na to, że wdrożeń DSpace na świecie jest prawie 1200, społeczność użytkowników jest bardzo duża i wspiera się nawzajem w rozbudowie funkcjonalności systemu oraz w świadczeniu pomocy technicznej [DCAT Discussion Forum, dok. elektr.].

## Typy dokumentów

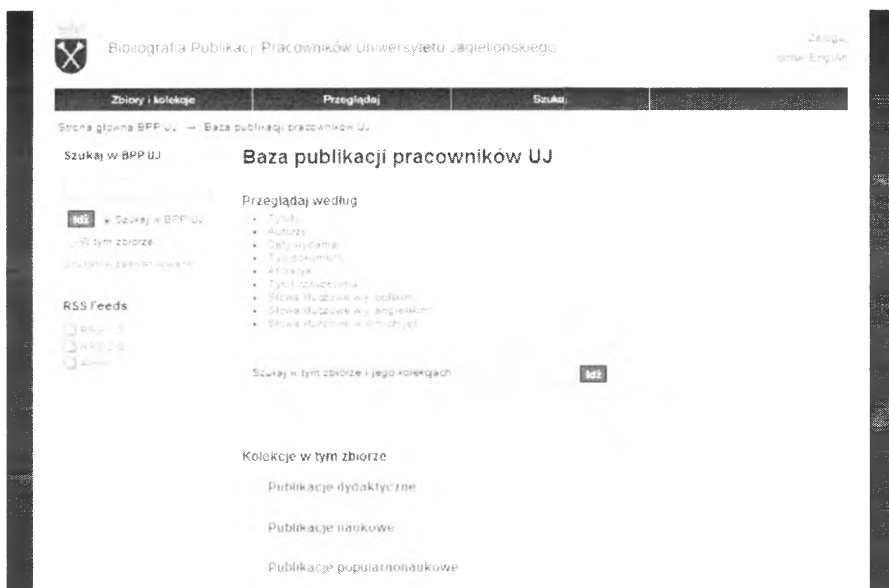
W pierwszym etapie prac nad powstaniem systemu, na podstawie analizy ankiet ministerialnych i wydziałowych, bibliografii oraz repozytoriów innych uczelni, zostały dobrane typy dokumentów i metadane opisujące każdy z nich. Na tym etapie pojawiły się pierwsze problemy terminologiczne związane z nazewnictwem, ponieważ ustalenia i normy biblioteczne różnią się od wytycznych ministerialnych. Przykładowo w parametryzacji na potrzeby jednostek używane są takie określenia dokumentów, jak: monografia, podręcznik, skrypt, redakcja monografii, natomiast w terminologii bibliotecznej wszystkie wymienione dokumenty to rodzaj wydawniczy: książka. Na potrzeby repozytorium czy bibliografii wystarczający jest typ „książka”, ale już do oceny parametrycznej jednostek wymagane są wszystkie wyżej wymienione rodzaje piśmiennicze dokumentów. Przy połączeniu bibliografii z repozytorium, a takie jest założenie przyszłościowe BPP UJ, również należało uwzględnić różnice w typach dokumentów. Potrzebny był kompromis oraz wybór optymalnej i najbardziej funkcjonalnej grupy dokumentów, jak również ich liczby. Typy dokumentów zastosowane w BPP UJ to:

- książka, monografia,
- skrypt,
- podręcznik,
- doktorat, praca dyplomowa,
- artykuł (rozdział) w książce,
- artykuł w materiałach konferencyjnych,
- hasło w encyklopedii, słowniku,
- artykuł w czasopiśmie,
- redakcja serii,
- redakcja czasopisma,
- redakcja materiałów konferencyjnych,
- redakcja publikacji zbiorowej,
- redakcja encyklopedii, słownika,
- recenzja,
- norma,

- zgłoszenie patentowe,
- patent,
- inne materiały niepublikowane,
- dokument dźwiękowy,
- dokument filmowy,
- atlas, mapa,
- bibliografia.

Autor już podczas pisania tego artykułu miał świadomość, że podana lista typów dokumentów nie jest zamknięta i prawdopodobnie wiele razy będzie się zmieniać już po oficjalnym uruchomieniu systemu. Jest to udowodnione na podstawie doświadczeń innych bibliotek tworzących systemy bibliograficzne. Obecnie jednak spełnia swoją rolę, ponieważ osoby uzupełniające dane w wersji testowej nie napotkały do tej pory problemów z zakwalifikowaniem wprowadzanych opisów do odpowiednich typów dokumentów.

Kolejnym etapem prac było określenie głównych kolekcji. Baza została podzielona na trzy kolekcje, odpowiadające podziałowi w ankietach parametrycznych, co powinno być najbardziej funkcjonalne z punktu widzenia przyszłych użytkowników (ilustracja 1). W każdej kolekcji używany jest ten sam zestaw typów dokumentów.



Ilustracja 1. Główne kolekcje w BPP UJ

Źródło: Bibliografia Publikacji Pracowników UJ [dok. elektr.].  
<http://bpp.uj.edu.pl/bpp/xmlui/community-list> [odczyt: 10.08.2014]

## Prezentacja metadanych

Podstawowym formatem metadanych w DSpace jest Dublin Core. Swoją popularność w bibliotekach, repozytoriach i archiwach cyfrowych prawdopodobnie zawdzięcza prostocie. W wersji podstawowej składa się z 15 elementów: tytuł, twórca, współtwórca, temat, opis, wydawca, data, rodzaj, format, identyfikator, źródło, język, powiązania, zasięg, prawa. Jednak Dublin Core w tej wersji dla użytkowników okazał się niewystarczający, dlatego powstała nowa wersja formatu, która ma zdefiniowanych 55 terminów [DCMI Usage Board 2012, dok. elektr.]. W nowej wersji można uszczegóławiać i rozszerzać znaczenie elementów podstawowych.

DSpace ma zdefiniowane 55 terminów Dublin Core. Na potrzeby BPP UJ zostały one rozbudowane do 106. Przykładowy zestaw metadanych dla książki z elementami Dublin Core przedstawia tabela 2.

Tabela 2. Przykładowy zestaw metadanych w BPP UJ

Typ	Książka		
Nazwa pola:	znacznik	Obow.	Powt.
tytuł:	dc.title	tak	nie
tom:	dc.description.volume *	nie	nie
wariant tytułu:	dc.title.alternative *	nie	tak
autor:	dc.contributor.author *	tak	tak
współtwórca:	dc.contributor.other *	nie	tak
tłumacz:	dc.contributor.translator *	nie	tak
redaktor:	dc.contributor.editor *	nie	tak
instytucja sprawcza:	dc.contributor.institution *	nie	tak
wydanie:	dc.description.edition *	nie	nie
miejsce wydania : wydawca:	dc.pubinfo *	tak	tak
data wydania:	dc.date.issued	tak	nie
seria:	dc.description.series *	nie	tak
ISBN:	dc.identifier.isbn *	nie	tak
eISBN:	dc.identifier.eisbn *	nie	tak
DOI:	dc.identifier.doi *	nie	nie
adres URL:	dc.identifier.weblink *	nie	tak
data dostępu:	dc.date.accession *	nie	tak
strony:	dc.description.physical *	tak	nie
uwagi:	dc.description.additional *	nie	nie
abstrakt w j. polskim:	dc.abstract.pl *	nie	nie
abstrakt w j. angielskim:	dc.abstract.en *	nie	nie
abstrakt w j. dokumentu:	dc.abstract.other *	nie	nie
słowa kluczowe w j. polskim:	dc.subject.pl *	nie	tak

c.d. tabeli 2

Typ	Książka		
słowa kluczowe w j. angielskim:	dc.subject.en *	nie	tak
słowa kluczowe w j. dokumentu:	dc.subject	nie	tak
punktacja wydziałowa:	dc.description.points *	nie	nie
liczba arkuszy wydawniczych:	dc.description.publication *	nie	nie
wydział:Instytut/Zakład/Katedra:	dc.affiliation *	nie	tak
język:	dc.language	tak	tak

\* elementy Dublin Core utworzone w BPP UJ

Źródło: [Drabik, Mesek, Szafranski, dok. elektr.]

Metadane w BPP UJ prezentowane są w pięciu formatach: opis skrócony do szybkiego przeglądania bibliografii (lista tytułów), opis ciągły do zestawień bibliograficznych, opis tabelaryczny do rozdzielania poszczególnych stref opisu w celu bardziej czytelnej prezentacji metadanych podobnie jak w katalogu bibliotecznym, opis pozycji z terminami Dublin Core przetłumaczonymi na język polski i metadane (Dublin Core) z oryginalnymi nazwami.

Z poziomu opisu skróconego, opisu ciągłego i opisu tabelarycznego jest możliwość wydruku i eksportu metadanych do formatów PDF, RIS i BibTeX.

## Funkcjonalność systemu

Wyszukiwanie danych w systemie zostało zorganizowane na dwa sposoby. Dla użytkownika zostały przygotowane ekrany: „przełóżaj” i „szukaj”. W zakładce „przełóżaj” użytkownik ma do wyboru następujące indeksy: autorów, dat wydania, typów dokumentów, afiliacji, tytułów czasopism, słów kluczowych w języku polskim, słów kluczowych w języku angielskim, słów kluczowych w innych językach. W panelu „szukaj” poza wyszukiwaniem prostym jest możliwość filtrowania zapytań przez dodawanie kolejnych filtrów i łączenie zapytań za pomocą operatorów Boole’a. Pozwala to na tworzenie zestawień bibliograficznych na poziomie wydziałów, instytutów, indywidualnych pracowników, za określony czas według dat wydania i według typów dokumentów. Wydziały, instytuty i typy dokumentów wybiera się z listy. W panelu „szukaj” znajduje się również funkcja „odkryj” (ang. *discover*), która pozwala na szybkie odnalezienie danych na wyszukanych stronach. Na ilustracji 2 przedstawiony jest widok zakładki „szukaj” z trzema filrami.



Ilustracja 2. Ekran „Szukaj” z dodanymi trzema filtrami  
 Źródło: Bibliografia Publikacji Pracowników UJ [dok. elektr.].  
<http://bpp.uj.edu.pl/bpp/xmlui/discover> [odczyt: 10.08.2014]

Pracownicy UJ nie muszą zakładać konta w BPP UJ, ponieważ logowanie do programu odbywa się przez punkt logowania UJ. Dodatkową korzyścią z połączenia tych dwóch systemów jest automatyczne pobieranie danych o pracowniku i rozpoznawanie go na poziomie wypełniania formularzy z danymi do bibliografii. W ten sposób w trakcie wprowadzania danych o publikacjach automatycznie z SAP UJ pobierane są imiona i nazwiska autorów, jeżeli są oni pracownikami UJ, oraz jednostka, w której pracują. W przyszłości system ma zostać połączony również z systemem USOS, skąd będą pobierane dane doktorantów UJ.

Do metadanych artykułów, które są publikowane w czasopiśmiech tzw. punktowanych, czyli takich, które znajdują się w Wykazie Czasopism Naukowych Ministra Nauki i Szkolnictwa Wyższego w częściach A, B, C, system na podstawie numeru ISSN, tytułu i roku wydania dodaje punkty, które wyświetlają się pod opisem publikacji. Jednak żeby ta opcja działała poprawnie, wykazy są sprawdzane przed importem do systemu. Podyktowane jest to brakiem ujednoczenia wpisywanych tytułów na poziomie Ministerstwa oraz błędami w nazwach i numerach ISSN. Po kontroli pliki konwertowane są do formatu CSV i importowane do BPP UJ.

Dodatkowo w systemie dodane zostały metadane: liczba arkuszy wydawniczych i punktacja wydziałowa, które wprowadzane są ręcznie przez redaktorów na podstawie informacji dostarczanych przez autorów, biblioteki i sekretariaty wydziałów.

## Model organizacji pracy w BPP UJ

W Bibliotece Jagiellońskiej, w Oddziale Zbiorów Cyfrowych została wyodrębniona Sekcja Repozytorium Uniwersytetu Jagiellońskiego (RUJ). Jest to zespół pracowników, który zajmuje się obsługą i koordynacją prac nad uzupełnianiem bibliografii. Po zakończeniu projektu HaZarD rozpocznie proces tworzenia Repozytorium UJ. Już na etapie budowy systemu została również nawiązana współpraca z bibliotekami wydziałowymi i instytutowymi UJ w celu pozyskiwania danych do bibliografii. W trakcie spotkań z przedstawicielami bibliotek uzyskane zostały informacje na temat funkcjonowania kilku baz indeksujących publikacje pracowników na Uniwersytecie. W Collegium Medicum UJ tworzona jest Bibliografia UJ CM w systemie Expertus. Baza opracowywana jest przez Oddział Informacji Naukowej Biblioteki Medycznej CM UJ. Na Wydziale Matematyki i Informatyki działa baza publikacji pracowników w oprogramowaniu APACZ [APACZ – Autorzy, Publikacje, Artykuły i Czasopisma, dok. elektr.]. Baza ta jest tworzona bez współudziału biblioteki wydziałowej. Uzupełnianiem danych zajmują się osobiście pracownicy Wydziału. W Instytucie Geografii i Gospodarki Przestrzennej funkcjonuje baza bibliograficzna, do której dane wprowadza oddelegowany do tego zadania pracownik. Również w tym przypadku biblioteka instytutowa nie uczestniczy w tworzeniu bazy. Według uzyskanych informacji bazy zostały wykonane na zlecenie jednostek przez firmy zewnętrzne. Zanalizowane zostały również strony internetowe wydziałów. Na podstawie przeprowadzonych badań stwierdzono, że na stronach większości wydziałów zamieszczone są listy z publikacjami pracowników. W nielicznych jednostkach UJ takich informacji nie ma lub podane są tylko wybrane, najciekawsze publikacje.

W celu dalszej współpracy z jednostkami i bibliotekami UJ w Sekcji RUJ został powołany pełnomocnik ds. kontaktów z pracownikami, którzy będą zbierać dane do BPP UJ. Dodatkowo przygotowywana została instrukcja wprowadzania danych oraz materiały szkoleniowe, które mają przygotować bibliotekarze do nowych zadań. W pierwszym etapie została nawiązana współpraca z trzema wydziałami UJ: Wydziałem Zarządzania i Komunikacji Społecznej, Wydziałem Matematyki i Informatyki oraz Wydziałem Studiów Międzynarodowych i Politycznych. Współpraca polega na dostarczaniu przez pracowników Wydziałów do RUJ wykazów publikacji: w pierwszej kolejności za lata 2012-2013, a następnie na udziale w testowej ocenie parametrycznej jednostek w nowym systemie. Przewidziany jest automatyczny eksport metadanych z BPP UJ do ankiet wydziałowych i ich uzupełnianie w systemie SAP oraz tworzenie listy publikacji za poszczególne lata. Na ilustracji 3 przedstawiony jest model organizacji pracy w BPP UJ, a w późniejszym etapie – w Repozytorium UJ.





Ilustracja 3. Model organizacji pracy w BPP UJ

Źródło: Opracowanie własne

Według wstępnych założeń metadane będą uzupełniane w bazie przez pracowników Sekcji RUJ, bibliotekarzy z bibliotek wydziałowych i instytutowych oraz pracowników UJ, którzy będą sami chcieli uzupełniać swoje konto. Po wpisaniu do formularzy dane, zanim zostaną udostępnione w internecie, będą pojawiać się na panelu kontrolnym redaktora, który po skorygowaniu ewentualnych błędów i sprawdzeniu kompletności będzie zatwierdzał je do publikacji. Ten sposób deponowania jest możliwy dzięki funkcjom panelu kontroli dostępu, który gwarantuje nadawanie uprawnień na poziomie pojedynczych osób i grup użytkowników.

## Podsumowanie

System budowany w ramach projektu HaZarD stanowi punkt wyjścia do tworzenia uczelnianego repozytorium publikacji pracowników oraz prac doktorskich i prac dyplomowych studentów UJ. Dla Biblioteki Jagiellońskiej jest to

szansa na stworzenie kompletnej bazy danych, zawierającej opisy bibliograficzne prac powstających na Uniwersytecie, która będzie wykorzystywana przez wszystkie jednostki UJ. Baza będzie również używana do oceny parametrycznej oraz różnego rodzaju ankiet, analiz i statystyk na poziomie całej uczelni, pojedynczego wydziału czy instytutu. Kompletność bibliografii uzależniona jest w pierwszej kolejności od odpowiednich zarządzeń władz uczelni. Zarządzenia nakładają na pracowników obowiązek dostarczania własnych prac do biblioteki. Brak odpowiednich zapisów i wsparcia finansowego może być przyczyną zahamowania rozwoju systemu.

Najcenniejszy wydaje się fakt, że Bibliografia Publikacji Pracowników powstaje w Bibliotece Jagiellońskiej. Jednak sukces tego przedsięwzięcia uzależniony jest od wielu czynników, między innymi od zaangażowania pracowników naukowych i bibliotekarzy z bibliotek wydziałowych i instytutowych we współtworzenie bibliografii.

Czy przyjęte założenia okazały się słuszne, czy wymagają poprawy okaże się po oficjalnym uruchomieniu bazy.

## BIBLIOGRAFIA

- APACZ – *Autorzy, Publikacje, Artykuły i Czasopisma* [dok. elektr.]. <http://apacz.matinf.uj.edu.pl> [odczyt: 17.08.2014].
- DCAT Discussion Forum (2013) [dok. elektr.]. <http://www.dspace.org/community> [odczyt: 18.08.2014].
- DCMI Usage Board (2012). *DCMI Metadata Terms* [dok. elektr.]. <http://dublincore.org/documents/dcmi-terms> [odczyt: 18.08.2014].
- Garczyńska, Maria (2006). *Bibliografie publikacji pracowników państwowych szkół wyższych w Polsce w świetle badań ankietowych*. Przegląd Biblioteczny R. 74, z. 1, s. 60-76.
- Garczyńska, Maria (2013). *Bibliograficzne bazy danych w ocenie pracowników i jednostek szkół wyższych w Polsce*. W: Derfert-Wolf, Lidia red. *Bibliograficzne bazy danych i ich rola w rozwoju nauki: II Konferencja naukowa Konsorcjum BazTech*, Poznań, 17-19 kwietnia 2013. Poznań, s. 1-12 [dok. elektr.]. <http://eprints.rclis.org/22539/1/garczynska.pdf> [odczyt: 17.08.2014].
- Legeżyńska, Aleksandra (2006). *Dorobek naukowy pracowników uniwersytetów w Polsce – problemy rejestracji elektronicznej*. Zagadnienia Informacji Naukowej nr 2, s. 59-75.
- Sapa, Remigiusz (2010). *Access to Scholarly Output of Academic Staff: Bibliographic Databases and Institutional Repositories in Polish Academic Libraries*. Libri Vol. 60, Iss. 1, p. 78-91.
- Sójkowska, Iwona; Barańska-Malinowska, Barbara (2011). *Bibliografia publikacji pracowników źródłem informacji wspomagającej przygotowanie oceny jednostek naukowych*. W: Szarski, Henryk; Dudziak, Danuta red. *III Wrocławskie Spotkania Bibliotekarzy*. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, s. 17-31.

- Szarski, Henryk; Dudziak, Danuta red. (2011). *III Wrocławskie Spotkania Bibliotekarzy*. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, 344 s.
- Woźniak, Barbara (2013). *Dokumentacja dorobku naukowego pracowników uczelni wyższych – rola bibliotek w budowaniu społeczeństwa wiedzy* [dok. elektr.]. [http://www.wsb.pwr.wroc.pl/files/prv/id105/IV%20WSB%20prezentacje/19%20czerwca%20%20sroda/III%20Sesja/04.%20Wozniak\\_Barbara\\_Dokumentowanie.pdf](http://www.wsb.pwr.wroc.pl/files/prv/id105/IV%20WSB%20prezentacje/19%20czerwca%20%20sroda/III%20Sesja/04.%20Wozniak_Barbara_Dokumentowanie.pdf) [odczyt: 18.08.2014].
- Zawadzka, Aleksandra (2011). *Rejestracja dorobku pracowników uniwersytetów. Dotychczasowe dokonania i spojrzenie w przyszłość*. W: Szarski, Henryk; Dudziak, Danuta red. *III Wrocławskie Spotkania Bibliotekarzy*. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, s. 61-71.

### Lista bibliografii dostępnych online

Nazwa uczelni	Nazwa bibliografii	Adres URL
Uniwersytet Wrocławski	Bibliografia Publikacji Pracowników i Doktorantów Uniwersytetu Wrocławskiego	<a href="http://www.bu.uni.wroc.pl/publikacje">http://www.bu.uni.wroc.pl/publikacje</a>
Uniwersytet Gdański	Bibliografia Publikacji Pracowników, Doktorantów i Studentów Uniwersytetu Gdańskiego	<a href="http://expertus.bg.univ.gda.pl">http://expertus.bg.univ.gda.pl</a>
Politechnika Krakowska	Bibliografia Publikacji Pracowników Politechniki Krakowskiej	<a href="http://suw.biblos.pk.edu.pl/browseByAuthor&amp;db=BPP">http://suw.biblos.pk.edu.pl/browseByAuthor&amp;db=BPP</a>
Politechnika Koszalińska	Bibliografia Prac Pracowników	<a href="http://biblioteka.tu.koszalin.pl/joomla/index.php?option=com_content&amp;view=article&amp;id=136&amp;Itemid=140">http://biblioteka.tu.koszalin.pl/joomla/index.php?option=com_content&amp;view=article&amp;id=136&amp;Itemid=140</a>
Uniwersytet Medyczny w Białymstoku	Bibliografia Publikacji Pracowników Uniwersytetu Medycznego w Białymstoku	<a href="http://biblioteka.amb.edu.pl/expertus/index.html">http://biblioteka.amb.edu.pl/expertus/index.html</a>
Politechnika Białostocka	Publikacje Pracowników PB	<a href="http://katalog.pb.edu.pl/F/?func=find-b-0&amp;local_base=bpb03">http://katalog.pb.edu.pl/F/?func=find-b-0&amp;local_base=bpb03</a>
Uniwersytet w Białymstoku	Publikacje Pracowników Naukowych UwB	<a href="http://katalogi.uwb.edu.pl/F/264GIV1Y776PHC78VNK4DV5I2AQ6F1BXJNYHJGDJLSJMEFSLU8-02804?func=find-d-0">http://katalogi.uwb.edu.pl/F/264GIV1Y776PHC78VNK4DV5I2AQ6F1BXJNYHJGDJLSJMEFSLU8-02804?func=find-d-0</a>
Akademia Techniczno-Humanistyczna w Bielsku-Białej	Bibliografia Dorobku Piśmienniczego Pracowników Akademii Techniczno-Humanistycznej w Bielsku-Białej od roku 1999	<a href="http://bibl.ath.bielsko.pl/~expertus/0/">http://bibl.ath.bielsko.pl/~expertus/0/</a>
Uniwersytet Kazimierza Wielkiego w Bydgoszczy	Baza Dorobku Naukowego Pracowników Uniwersytetu Kazimierza Wielkiego w Bydgoszczy	<a href="http://bibliografia.ukw.edu.pl/">http://bibliografia.ukw.edu.pl/</a>
Akademia im. Jana Długosza w Częstochowie	Bibliografia Publikacji Pracowników WSP w Częstochowie (1994-)	<a href="http://www.bg.ajd.czyst.pl/WSPCZ/uep.shtml">http://www.bg.ajd.czyst.pl/WSPCZ/uep.shtml</a>

Nazwa uczelni	Nazwa bibliografii	Adres URL
Gdański Uniwersytet Medyczny	Bibliografia Gdańskiego Uniwersytetu Medycznego	<a href="http://medline.amg.gda.pl/expertus/">http://medline.amg.gda.pl/expertus/</a>
Akademia Wychowania Fizycznego i Sportu w Gdańsku	Bibliografia Publikacji Pracowników Akademii Wychowania Fizycznego i Sportu w Gdańsku od r. 2000	<a href="http://www.awfgda.expertus.com.pl/e/">http://www.awfgda.expertus.com.pl/e/</a>
Politechnika Gdańska	Katalog Prac Naukowo-Badawczych Politechniki Gdańskiej	<a href="http://prace-nb.pg.gda.pl/">http://prace-nb.pg.gda.pl/</a>
Akademia Morska	Bibliografia Pracowników AM	<a href="http://bg.am.gdynia.pl/katalog/cdi/sz zam_search.html">http://bg.am.gdynia.pl/katalog/cdi/sz zam_search.html</a>
Politechnika Śląska	Bibliografia Publikacji Pracowników Politechniki Śląskiej od 1988	<a href="http://www.bg.polsl.pl/expertus/p/">http://www.bg.polsl.pl/expertus/p/</a>
Śląski Uniwersytet Medyczny	Bibliografia Publikacji Pracowników Śląskiego Uniwersytetu Medycznego w Katowicach (d. ŚAM) od 1990	<a href="http://213.227.100.63/expertus/">http://213.227.100.63/expertus/</a>
Politechnika Świętokrzyska	Publikacje Naukowe Pracowników	<a href="http://www.dorobek.tu.kielce.pl/publikacje/publikacje_szukaj.php?">http://www.dorobek.tu.kielce.pl/publikacje/publikacje_szukaj.php?</a>
Uniwersytet Pedagogiczny im. Komisji Edukacji Narodowej	Bibliografia Publikacji Pracowników Uniwersytetu Pedagogicznego im. Komisji Edukacji Narodowej w Krakowie	<a href="http://www.up.krakow.pl/biblio/bibliografia/index.php?base=bib">http://www.up.krakow.pl/biblio/bibliografia/index.php?base=bib</a>
Akademia Górniczo-Hutnicza	Bibliografia Publikacji Pracowników AGH	<a href="http://www.bpp.agh.edu.pl/new/">http://www.bpp.agh.edu.pl/new/</a>
Akademia Wychowania Fizycznego w Krakowie	Bibliografia Publikacji Pracowników Akademii Wychowania Fizycznego w Krakowie (od roku 2009)	<a href="http://biblioteka.awf.krakow.pl/bib/">http://biblioteka.awf.krakow.pl/bib/</a>
Uniwersytet Ekonomiczny w Krakowie	Dorobek	<a href="http://gollum.uek.krakow.pl/bazy_ac/dorobek_multi/index.cgi">http://gollum.uek.krakow.pl/bazy_ac/dorobek_multi/index.cgi</a>
Collegium Medicum Uniwersytet Jagielloński	Bibliografia UJ CM	<a href="http://expertus.bm.cm-uj.krakow.pl/Bibliografia/index.html">http://expertus.bm.cm-uj.krakow.pl/Bibliografia/index.html</a>
Uniwersytet Medyczny w Lublinie	Uniwersytet Medyczny w Lublinie Baza Danych Publikacji Naukowych Pracowników	<a href="http://bpyf.am.lublin.pl/public_multiseek.php">http://bpyf.am.lublin.pl/public_multiseek.php</a>
Politechnika Lubelska	Publikacje Pracowników Politechniki Lubelskiej	<a href="http://pub.pollub.pl/">http://pub.pollub.pl/</a>
Uniwersytet Marii Curie-Skłodowskiej	Bibliografia Publikacji Pracowników UMCS	<a href="http://libases.umcs.lublin.pl/expertus/pub/">http://libases.umcs.lublin.pl/expertus/pub/</a>
Uniwersytet Przyrodniczy w Lublinie	Baza Publikacji Pracowników Uniwersytetu Przyrodniczego	<a href="http://syzyf.up.lublin.pl/~prace/public/">http://syzyf.up.lublin.pl/~prace/public/</a>
Akademia Sztuk Pięknych im. Wł. Strzemińskiego w Łodzi	Bibliografia Akademii Sztuk Pięknych im. Wł. Strzemińskiego w Łodzi	<a href="http://expertus.asp.lodz.pl/">http://expertus.asp.lodz.pl/</a>
Politechnika Łódzka	BIBLIO Bibliografia Dorobku Piśmienniczego Pracowników Politechniki Łódzkiej od 2004	<a href="http://212.51.210.184/~expertus/c/">http://212.51.210.184/~expertus/c/</a>

<b>Nazwa uczelni</b>	<b>Nazwa bibliografii</b>	<b>Adres URL</b>
Uniwersytet Łódzki	Bibliografia Publikacji Pracowników Uniwersytetu Łódzkiego	<a href="http://212.191.71.9/">http://212.191.71.9/</a> <a href="http://uni-lodz2.expertus.com.pl/cgi-bin/expertus.cgi">http://uni-lodz2.expertus.com.pl/cgi-bin/expertus.cgi</a>
Uniwersytet Medyczny w Łodzi	Bibliografia Publikacji Pracowników Uniwersytetu Medycznego w Łodzi	<a href="http://www.bg.umed.lodz.pl/pl/um/">http://www.bg.umed.lodz.pl/pl/um/</a>
Uniwersytet Ekonomiczny w Poznaniu	Bibliografia Publikacji Pracowników Uniwersytetu Ekonomicznego w Poznaniu	<a href="http://bg.ue.poznan.pl/all/">http://bg.ue.poznan.pl/all/</a>
Uniwersytet Medyczny w Poznaniu	Bibliografia Pracowników UM w Poznaniu	<a href="http://150.254.179.40/bazy/publikacje/">http://150.254.179.40/bazy/publikacje/</a>
Akademia Wychowania Fizycznego w Poznaniu	Bibliografia Publikacji Pracowników Akademii Wychowania Fizycznego w Poznaniu	<a href="http://klio.awf.poznan.pl/bp/index.php">http://klio.awf.poznan.pl/bp/index.php</a>
Politechnika Poznańska	Bibliografia Publikacji Pracowników Politechniki Poznańskiej BIBLIO	<a href="http://library.put.poznan.pl/bib/">http://library.put.poznan.pl/bib/</a>
Uniwersytet Rzeszowski	Bibliografia Publikacji Pracowników Uniwersytetu Rzeszowskiego (2001-...)	<a href="http://bibliografia.univ.rzeszow.pl/01/">http://bibliografia.univ.rzeszow.pl/01/</a>
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie	PUBLI Baza Publikacji Naukowych Pracowników ZUT	<a href="http://www.aleph.zut.edu.pl/F/?func=find-b-0&amp;local_base=tus03">http://www.aleph.zut.edu.pl/F/?func=find-b-0&amp;local_base=tus03</a>
Pomorski Uniwersytet Medyczny w Szczecinie	Bibliografia Dorobku Naukowego Pomorskiego Uniwersytetu Medycznego w Szczecinie	<a href="http://bibliografia.pum.edu.pl/">http://bibliografia.pum.edu.pl/</a>
Uniwersytet Mikołaja Kopernika w Toruniu	Bibliografia Publikacji Pracowników i Doktorantów Uniwersytetu Mikołaja Kopernika w Toruniu	<a href="http://212.122.203.4/splendor/umk/">http://212.122.203.4/splendor/umk/</a>
Warszawski Uniwersytet Medyczny	Bibliografia Publikacji Pracowników Warszawskiego Uniwersytetu Medycznego	<a href="http://bibl.amwaw.edu.pl/expertus/p/index.html">http://bibl.amwaw.edu.pl/expertus/p/index.html</a>
Politechnika Warszawska	BIBLIOgrafia Pracowników Politechniki Warszawskiej	<a href="http://gate.bg.pw.edu.pl/F/?func=find-b-0&amp;local_base=wtu03">http://gate.bg.pw.edu.pl/F/?func=find-b-0&amp;local_base=wtu03</a>
Uniwersytet Warszawski	Bibliografia Uniwersytetu Warszawskiego	<a href="http://bibliografia.icm.edu.pl/g2/">http://bibliografia.icm.edu.pl/g2/</a>
Uniwersytet Medyczny we Wrocławiu	Bibliografia Publikacji Pracowników Uniwersytetu Medycznego we Wrocławiu	<a href="http://156.17.60.59/expertus/w/">http://156.17.60.59/expertus/w/</a>
Akademia Wychowania Fizycznego we Wrocławiu	Bibliografia Publikacji Pracowników Akademii Wychowania Fizycznego we Wrocławiu	<a href="http://aleph.awf.wroc.pl/expertus/www_expertus/p/">http://aleph.awf.wroc.pl/expertus/www_expertus/p/</a>
Politechnika Wrocławska	Centrum Wiedzy i Informacji Naukowo-Technicznej Politechniki Wrocławskiej Baza DONA	<a href="http://apin2.bg.pwr.wroc.pl/Aleph/wysz_aut.htm">http://apin2.bg.pwr.wroc.pl/Aleph/wysz_aut.htm</a>
Uniwersytet Przyrodniczy we Wrocławiu	Publikacje Pracowników UP	<a href="http://aleph18.bibl.up.wroc.pl/F/?func=find-b-0&amp;local_base=arw03">http://aleph18.bibl.up.wroc.pl/F/?func=find-b-0&amp;local_base=arw03</a>
Uniwersytet Zielonogórski	System Komputerowej Ewidencji Publikacji	<a href="http://publikacje.uz.zgora.pl:7777/skep/toolbox.show_main_page_pl">http://publikacje.uz.zgora.pl:7777/skep/toolbox.show_main_page_pl</a>

Nazwa uczelni	Nazwa bibliografii	Adres URL
Uniwersytet Technologiczno-Przyrodniczy im. J.J. Śniadeckich w Bydgoszczy	Bibliografia Publikacji Pracowników Uniwersytetu Technologiczno-Przyrodniczego im. J.J. Śniadeckich w Bydgoszczy	<a href="http://www.bg.utp.edu.pl/expertus/">http://www.bg.utp.edu.pl/expertus/</a>
Wyższa Szkoła Oficerska Sił Powietrznych w Dęblinie	Bibliografia Publikacji Pracowników Wyższej Szkoły Oficerskiej Sił Powietrznych w Dęblinie	<a href="http://expertus.wsosp.pl/">http://expertus.wsosp.pl/</a>
Wyższa Szkoła Oficerska Wojsk Lądowych we Wrocławiu	Baza ZMECH	<a href="http://www.wso.wroc.pl/bg/expertus/zmech/">http://www.wso.wroc.pl/bg/expertus/zmech/</a>

---

# Baza e-PKA – rewitalizacja Starego Katalogu Biblioteki Jagiellońskiej

Krystyna Sanetra

Biblioteka Jagiellońska

---

## 1. Skanowanie jako alternatywa dla retrospektywnego opracowania zbiorów

W początkach lat 90. XX w. biblioteki w Polsce zaczęły wdrażać zautomatyzowane systemy biblioteczne, inicjując proces przechodzenia od tradycyjnych katalogów kartkowych (odrębnych dla poszczególnych kategorii dokumentów) do zintegrowanych katalogów komputerowych typu OPAC (ang. *online public access catalog*). Obok sporządzania opisów bibliograficznych dla bieżącego wpływu niemal od początku rozpoczęto też retrospektywne opracowanie zbiorów, tj. tworzenie rekordów bibliograficznych dla dokumentów opisanych wcześniej na kartach katalogowych. Opracowanie retrospektywne, a właściwie rekatalogowanie, to proces kosztowny, pracochłonny i czasochłonny, a przeniesienie metadanych z kart katalogowych do katalogu komputerowego było (i wciąż jest) jednym z najważniejszych wyzwań dla bibliotek. Potrzeba dostarczenia informacji w formie cyfrowej o zasobach tradycyjnych była i jest tak duża, że aby jej sprostać, dość powszechną praktyką stało się skanowanie kart katalogowych i tworzenie cyfrowych kopii katalogów kartkowych. Katalogi te nazywa się zwykle elektronicznymi, zeskanowanymi, zdigitalizowanymi katalogami kartkowymi lub po prostu katalogami kartkowymi albo katalogami obrazów kart. W odniesieniu do tego ostatniego terminu i przez analogię do nazwy OPAC pojawiła się też nazwa CIPAC (ang. *card-image public access catalog*). CIPAC zdefiniować można jako katalog biblioteczny dostępny *online*, oparty na bazie danych, w której zgromadzono zdigitalizowane karty katalogowe, oferujący bardziej lub mniej wyrafinowane mechanizmy przeglądania i wyszukiwania [Oberhauser 2003]. W 2001 r. Oberhauser utworzył istniejącą do dziś stronę internetową, zawierającą zestawienie katalogów typu CIPAC – The International CIPAC List [Oberhauser 2014,

dok. elektr.]. Jest ona nadal rozbudowywana; obecnie zawiera alfabetyczną listę 28 państw, a w ich obrębie listę miast i instytucji z linkami do 157 katalogów CIPAC (najwięcej, bo 48 katalogów z Niemiec, 25 – z Czech, 24 – z Austrii). Lista 7 polskich CIPAC jest z pewnością niekompletna i nie wszystkie linki są poprawne.

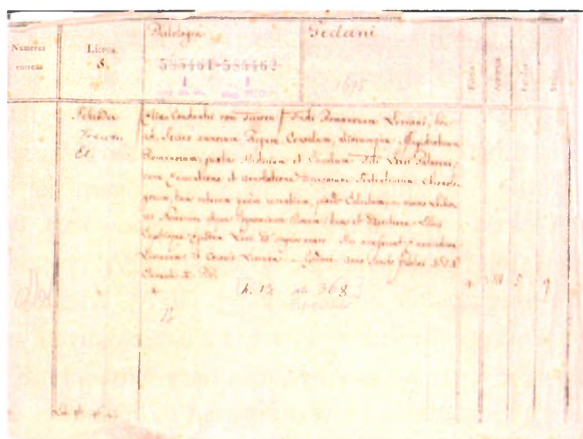
## 2. Elektroniczna wersja Podstawowego Katalogu Alfabetycznego, czyli CIPAC BI

CIPAC Biblioteki Jagiellońskiej, czyli baza e-PKA, powstał w latach 1998-2003, w którym to okresie skanowano i stopniowo udostępniano w internecie obrazy kart katalogowych Podstawowego Katalogu Alfabetycznego druków wydanych od czasów Gutenberga do 1949 r.

### 2.1. Podstawowy Katalog Alfabetyczny, zwany Starym Katalogiem

Podstawowy Katalog Alfabetyczny (PKA), zwany Starym Katalogiem, powstał w wyniku połączenia katalogu polskiego, katalogu obcego i katalogu cimeliów. Był pierwszym kartkowym katalogiem Biblioteki Jagiellońskiej. Najstarsze karty pochodzą z początku XIX w., a zaczęto je tworzyć z inicjatywy kierującego biblioteką w latach 1802-1807 Anzelma Speisera (ilustracja 1). Były to karty o wymiarach 19x25 cm, zwane cedułkami.

W prawym górnym rogu karty katalogowej znajdowały się rubryki do oznaczenia formatu i sygnatury topograficznej: oznaczenie szafy (łac. *armarium*),



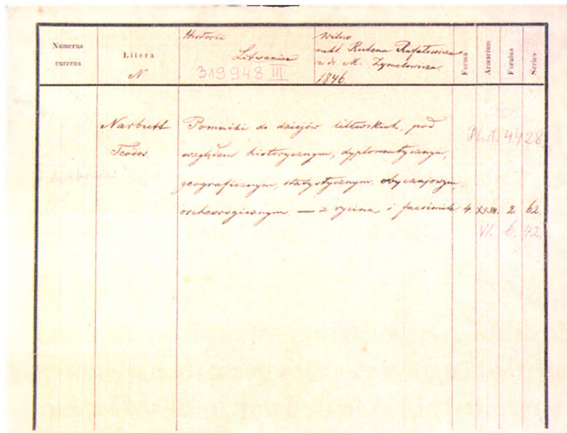
Ilustracja 1. Karta katalogowa sporządzona przez Anzelma Speisera  
Źródło: Biblioteka Jagiellońska



oznaczenie półki (łac. *forulus*) i kolejnego numeru na półce (łac. *series*). Zamiast sygnatury wpisywano zwykle tylko nazwę działu (tu: *Philologia*) i pierwszą literę nazwiska autora. Przyjęty format i wielkość kart pozostały właściwie takie same, później zmieniło się tylko rubrykowanie. Następcy Anzelma Speisera, Jerzy Samuel Bandtkie, Józef Muczkowski i Karol Estreicher, kontynuowali rozpoczęte prace katalogowe (ilustracje 2-4), wprowadzając sporo zmian i uzupełnień. Ostateczny kształt i układ kart został ustalony w latach 30. XX w., kiedy funkcję dyrektora pełnił Edward Kuntze (1926-1947).



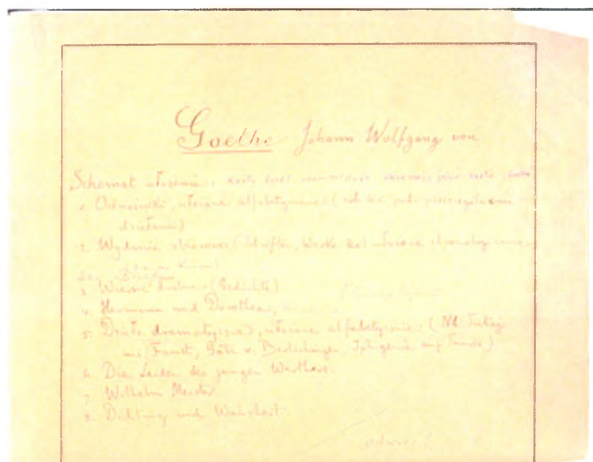
Ilustracja 2. Karta katalogowa sporządzona przez J. S. Bandtkiego, dyrektora Biblioteki Jagiellońskiej w latach 1811-1835  
Źródło: Biblioteka Jagiellońska



Ilustracja 3. Karta katalogowa sporządzona przez Józefa Muczkowskiego, dyrektora Biblioteki Jagiellońskiej w latach 1837-1858  
Źródło: Biblioteka Jagiellońska

Forma	WYMIAR (cm)	Tytuł	Liczba		
			W katalogu	W zbiorze	W sprzedaży
		głównie, alfabetycznie, rok druku, nakład, język, kraj, data wydania, w formie, okładka, twarda, miękka, papier, kolor, plastyk — Etykieta, Okładka, Miękka, zgrzewana			
		<i>Goethe, Johann Wolfgang von</i>			
		<i>Schemat utworów, które miał w swoim posiadaniu, jak następuje:</i>			
		<i>1. Odmianki utworów alfabetycznie (zob. listę w poprzednim katalogu)</i>			
		<i>2. Wykaz utworów (tytuły, rok, nakład, język, kraj, data wydania, w formie, okładka, twarda, miękka, papier, kolor, plastyk — Etykieta, Okładka, Miękka, zgrzewana)</i>			
		<i>3. Wykaz utworów (tytuły, rok, nakład, język, kraj, data wydania, w formie, okładka, twarda, miękka, papier, kolor, plastyk — Etykieta, Okładka, Miękka, zgrzewana)</i>			
		<i>4. Wykaz utworów (tytuły, rok, nakład, język, kraj, data wydania, w formie, okładka, twarda, miękka, papier, kolor, plastyk — Etykieta, Okładka, Miękka, zgrzewana)</i>			
		<i>5. Wykaz utworów (tytuły, rok, nakład, język, kraj, data wydania, w formie, okładka, twarda, miękka, papier, kolor, plastyk — Etykieta, Okładka, Miękka, zgrzewana)</i>			
		<i>6. Wykaz utworów (tytuły, rok, nakład, język, kraj, data wydania, w formie, okładka, twarda, miękka, papier, kolor, plastyk — Etykieta, Okładka, Miękka, zgrzewana)</i>			
		<i>7. Wykaz utworów (tytuły, rok, nakład, język, kraj, data wydania, w formie, okładka, twarda, miękka, papier, kolor, plastyk — Etykieta, Okładka, Miękka, zgrzewana)</i>			
		<i>8. Wykaz utworów (tytuły, rok, nakład, język, kraj, data wydania, w formie, okładka, twarda, miękka, papier, kolor, plastyk — Etykieta, Okładka, Miękka, zgrzewana)</i>			
		<i>9. Wykaz utworów (tytuły, rok, nakład, język, kraj, data wydania, w formie, okładka, twarda, miękka, papier, kolor, plastyk — Etykieta, Okładka, Miękka, zgrzewana)</i>			
		<i>10. Wykaz utworów (tytuły, rok, nakład, język, kraj, data wydania, w formie, okładka, twarda, miękka, papier, kolor, plastyk — Etykieta, Okładka, Miękka, zgrzewana)</i>			

Ilustracja 4. Karta katalogowa sporządzona przez Karola Estreichera, dyrektora Biblioteki Jagiellońskiej w latach 1868-1905  
Źródło: Biblioteka Jagiellońska



Ilustracja 5. Karta przewodnia sporządzona przez Aleksandra Birkenmajera, dyrektora Biblioteki Jagiellońskiej w latach 1947-1951  
Źródło: Biblioteka Jagiellońska

PKA rejestruje druki zwarte i serie wydawnicze, ale znajdują się w nim również sygnalne karty katalogowe dla czasopism, bez wyszczególnienia roczników i numerów. Najstarsze karty PKA mają przeszło 200 lat i stanowią dzisiaj unikalny w skali europejskiej zbiór autografów wybitnych polskich bibliotekarzy (Bandtkiego, Estreichera, Kuntzego, Piekarskiego i in.). Katalog zawiera ok. miliona wielkoformatowych kart katalogowych, przechowywanych luźno w 3120 dREW-

nianych skrzynkach, opatrzonych kolejnym numerem porządkowym i określeniem zakresu indeksu. Pewna liczba kart, szacunkowo ok. 5%, to karty trwale połączone w grupy, tzw. karty zszywane.



Ilustracja 6. Skrzynka katalogowa nr 26, zakres indeksu Albertus Magnus – Albi  
Źródło: Biblioteka Jagiellońska

Podstawą układu kart jest alfabet łaciński z uwzględnieniem liter k i w, obrazuje to poniższa tablica. Inne alfabety są poddane transliteracji.

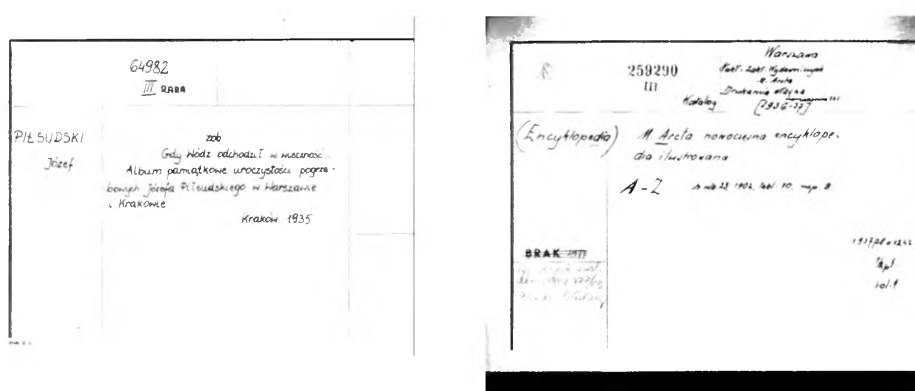
Tablica ilustrująca zasadę szeregowania kart w PKA i e-PKA

A = ā ä (ae)	I = j ĭ	S = ś š
B	K	T
C = ċ ċ	L = ĺ	U = ü (ue)
D	M	V
E = ē	N = ñ	W
F	O = ó ö (oe)	X
G	P, Q	Y
H	R	Z = ź ż ž

Źródło: Biblioteka Jagiellońska

PKA jest częściowo katalogiem krzyżowym, tzn. w jednym ciągu alfabetycznym ułożone są karty z hasłami: autorskimi, tytułowymi i przedmiotowymi. Karty opatrzone są hasłami autorskimi, jeżeli książka ma jednego, dwóch lub trzech autorów. Nazwiska świętych, władców, papieży i autorów starożytnych występują w formie łacińskiej (np. Casimirus, Joannes, Vergilius), a autorzy średniowieczni rejestrowani są pod imieniem w formie łacińskiej lub w formie narodowej.

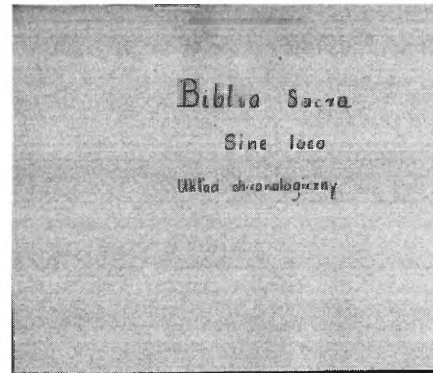
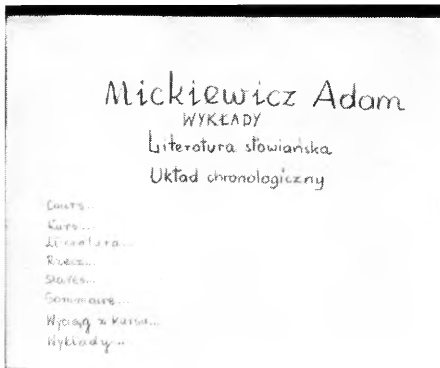
Dla autorów późniejszych zawsze stosuje się oryginalną pisownię, np. Shakespeare, Descartes, Voltaire. Jeżeli dzieło ma więcej niż trzech autorów lub jest anonimowe, hasłem jest pierwszy rzeczownik z tytułu stojący w mianowniku, np. dla *Wielkiej Ilustrowanej Encyklopedii Powszechnej* hasłem na karcie jest słowo „encyklopedia” (ilustracja 7). Hasła rzeczowe tworzone były dla dzieł dotyczących instytucji, uniwersytetów, towarzystw, bibliotek, muzeów, partii, organizacji społecznych, kulturalnych i oświatowych, miast i państw (w pisowni oryginalnej), wielkich i znanych wydarzeń historycznych i politycznych (np. I i II wojna światowa, powstanie listopadowe etc.) oraz nazw znanych osób (np. Piłsudski, ilustracja 7).



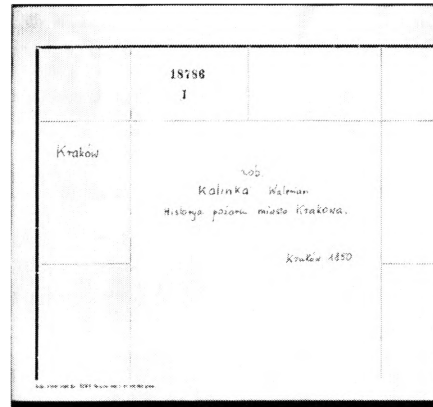
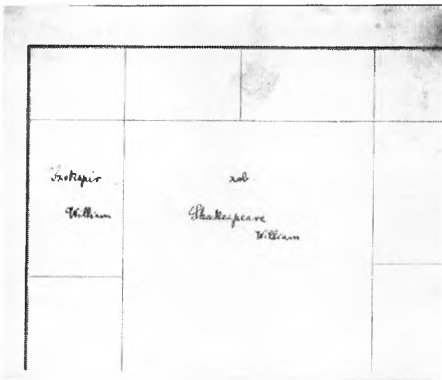
Ilustracja 7. Karty opatrzone hasłem rzeczowym i hasłem tytułowym  
Źródło: Biblioteka Jagiellońska

Do specyfiki PKA należy też uwzględnienie hasel takich, jak: antologie, druki ulotne (w obrębie hasła występuje układ przedmiotowy), sprawozdania, statuty (dla instytucji, towarzystw i organizacji – topograficzny układ kart według miast, a w obrębie miast – alfabetycznie według nazw instytucji), programy szkolne (spisy wykładów, sprawozdania dyrektorów, składy osobowe etc.; karty są ułożone topograficznie, a w obrębie miast – alfabetycznie według nazw instytucji). Hasło „Biblia” obejmuje zarówno całe Pismo Święte, jak i tytuły poszczególnych ksiąg (np. Stary Testament, Ewangelie, Księgę Psalmów), ułożone według miejsc wydania, a w ich obrębie – chronologicznie. Pomocą w wyszukiwaniu są karty przewodnie (ilustracje 5 i 8), pomagające uzyskać orientację w zastosowanym układzie kart w obrębie równobrzmiącego hasła, oraz karty odsyłaczowe (ilustracja 9).

Na kartach katalogowych występują różne sygnatury: działowe dla starych zbiorów (np. Prawo 141 II, Medic. 28137), *numerus currens* – numery inwentarowe (np. 567276 III), *incunabula* – dla druków z XV w. (np. Inc. 62), cimelia



Ilustracja 8. Karty przewodnie  
Źródło: Biblioteka Jagiellońska



Ilustracja 9. Karty odsyłaczowe  
Źródło: Biblioteka Jagiellońska

– dla najcenniejszych dzieł Biblioteki (np. Cim. Q4000), albumy – dla książek ilustrowanych grafiką oryginalną (np. 1200 II Albumy).

Od 1950 r. rozpoczęto rejestrowanie bieżących nabytków na kartach formatu międzynarodowego (7,5x12,5 cm) w nowym katalogu, dostępnym dla publiczności. Nadal jednak uzupełniano Stary Katalog o karty z opisami nabytków wydanych przed 1950 r., nawet jeżeli od 1994 r. rekordy bibliograficzne dla tych nabytków sporządzane były również w Komputerowym Katalogu Zbiorów Bibliotek UJ (KKZBUJ). Dlatego w PKA znajdują się nie tylko karty pisane ręcznie (ilustracja 10), ale również pisane na maszynie (ilustracja 11) oraz karty z naklejonymi wydrukami opisów z bazy KKZBUJ (ilustracja 12). PKA uzupełniano o karty rejestrujące nabytki wydane przed 1950 r., także gdy powstała baza e-PKA.

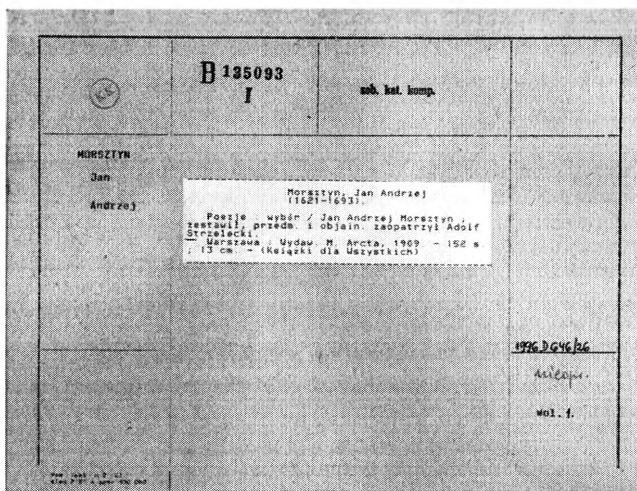
Wtedy karty skanowano i obrazami kart uzupełniano elektroniczną wersję Staroego Katalogu. Z włączania do PKA i do e-PKA kart / skanów kart dla druków wydanych do 1949 r., a nabytych w okresie późniejszym, zrezygnowano dopiero na początku 2008 r., na mocy zarządzenia nr 3/2008 dyrektora Biblioteki Jagiellońskiej Zdzisława Pietrzyka.

119	<b>370344</b> II	Warszawa S. Lewental 1883	
Morsztyn Jan Andrzej	Dozycie. Oryginały i fotokopie.  4 wyd. zbiorowe, opatrzone rękopiśmianym archiwiz- mów i imion własnych. s. XXVI, 29-465, 7 nb.  Biblioteka najcenniejszych utworów literatury europejskiej. - Literatura polska.		1937. 12.10. pt. vol. 1.

Ilustracja 10. Karta napisana ręcznie w latach 50. XX w.  
Źródło: Biblioteka Jagiellońska

	585356 I Woj. Sl. Dz.	Kraków Fr. Cesary 1676	
<i>Morsztyn</i> Antypasty	Antypasty maślane. 4-o. Knlb. 72 Sygn.: A-5 <sup>4</sup> Defekt: brak składek A, k. tyt. zastąpione prze- rysami. NB: Książeczka podaje mylnie ilość kart na 64.		1967 k 144 Sl. Dz.  Opr. 11/2 p.  vol. 1.
W. B., Rudnicka 20. Pod sygn. 801 I odbiła typogr. nieco odmienne.			

Ilustracja 11. Karta napisana na maszynie w latach 60. XX w.  
Źródło: Biblioteka Jagiellońska



Ilustracja 12. Karta z wklejonym wydrukiem opisu z KKZBUJ – lata 90. XX w.  
 Źródło: Biblioteka Jagiellońska

PKA jest bezcennym zabytkiem kultury narodowej, a jednocześnie – do czasu digitalizacji – był jedynym źródłem informacji o unikalnych zbiorach Biblioteki Jagiellońskiej, nie posiadano bowiem żadnej pełnej i trwałej jego kopii. Konieczność udostępniania czytelnikom narażała go na fizyczne uszkodzenie, a z biegiem czasu na niebezpieczeństwo całkowitego zniszczenia. Oznaczałoby to bezpowrotną utratę informacji o podstawowym księgozborze BJ, jednym z najstarszych, unikalnym w skali światowej i istotnym dla europejskiego dziedzictwa kultury.

## 2.2. Digitalizacja Podstawowego Katalogu Alfabetycznego

Projekt digitalizacji PKA był wielkim przedsięwzięciem zarówno finansowym, jak i organizacyjnym, ale przyniósł wymierne korzyści polskiemu i międzynarodowemu środowisku naukowemu, a także samej Bibliotece Jagiellońskiej – gwarantując zachowanie informacji o cennym księgozborze oraz podniesienie jakości świadczonych przez nią usług informacyjnych.

Podstawowym założeniem digitalizacji PKA było wierne odwzorowanie obrazu wszystkich kart katalogowych w procesie skanowania oraz odtworzenie fizycznej struktury katalogu kartkowego w logicznej strukturze bazy danych graficznych. W procesie skanowania zachowywano następujące parametry plików graficznych: format pliku – TIFF 5.0, średni rozmiar pliku – 750 KB, odwzorowanie rozmiaru – 1023x806 punktów (100 punktów/cal, skala 1:1), odwzorowanie barw – 256 odcieni szarości, wyostrenie – 3%. Uzyskane pliki graficzne

archiwizowano na nośnikach CD-ROM. Z kolei parametry plików graficznych udostępnianych w internecie były następujące: format pliku graficznego – JPEG, stopień kompresji – 75%, rozmiar obrazu – średnio 768x600 punktów, a średni rozmiar pliku – 75 KB. Udostępniono w trybie *online* dużą bazę danych graficznych (ok. 100 GB).

Projekt realizowany był w dwóch etapach. W latach 1998–1999 przeprowadzono pierwszy etap digitalizacji, obejmujący 550 000 kart katalogowych (1715 skrzynek). Prace zostały zrealizowane dzięki dotacji KBN (Komitetu Badań Naukowych), przy wkładzie własnym w wysokości 133 000 zł, co stanowiło 50% ogólnych nakładów wynoszących 266 000 zł (środki finansowe przyznano decyzją nr 1217/W/DOT/DI/98). Skanowanie kart powierzono firmie DjaF, kontrolę merytoryczną – bibliotekarzom BJ, natomiast informatycy BJ stworzyli oprogramowanie, skonfigurowali sprzęt i zajęli się obróbką informatyczną plików oraz ich archiwizowaniem. Skanowano dziennie ok. 13 skrzynek katalogu (ok. 300–330 obrazów w jednej skrzynce), prowadząc kontrolę poprawności odwzorowania dla wszystkich zeskanowanych obiektów. Nazwa pliku zawierającego graficzny obraz karty miała postać „skrzynka\_karta\_sufiks”, przy czym „skrzynka” to cztery cyfry numeru skrzynki PKA (dopełnione wiodącymi zerami), „karta” to cztery cyfry kolejnego numeru awersu karty w obrębie skrzynki (dopełnione wiodącymi zerami), a „sufiks” to kolejne małe litery alfabetu łacińskiego, określające karty<sup>1</sup> zapisane dwustronnie lub zszywane. Na kartach jednostronnych sufiks nie występował. Przykładowy opis wyglądał następująco:

0025\_0258 = 258 karta z 25 skrzynki; 0025\_0258\_b = rewers 258 karty z 25 skrzynki

Pliki umieszczano w katalogach o nazwach określających przynależność do skrzynki (tj. cztery cyfry określające numer skrzynki, np. 0025). Proces wprowadzania obrazów graficznych był zorganizowany w taki sposób, aby wykluczyć przemieszczenie kart pomiędzy skrzynkami, jak również zmianę kolejności kart w skrzynce. Dziennie opracowywano średnio 4 000 plików graficznych, o średnim rozmiarze w formacie wyjściowym TIFF – 2,7 GB, archiwizowanych na 5 płytach CD-ROM. Sumarycznie w pierwszym etapie realizacji procesu digitalizacji (ok. 140 dni) wykonano 550 000 obrazów kart, zarchiwizowanych na ok. 700 płytach CD-ROM, o objętości ok. 376 GB, w formacie wyjściowym plików TIFF, a także udostępniono, w formacie JPEG, ok. 46 GB danych. Początkowo wykorzystywano serwer o parametrach: procesor Pentium – 450 MHz, pamięć operacyjna – 256 MB, pamięć dyskowa – 72 GB (8 x HDD 9 GB Ultra Wide

---

<sup>1</sup> Karty PKA dzielą się na trzy podstawowe rodzaje: pojedyncze zapisane jednostronnie, pojedyncze zapisane dwustronnie, nierozłączne grupy kart (zsyte), zawierające od kilku do kilkunastu kart (jednostronnych i/lub dwustronnych), stanowiące szacunkowo ok. 1% wszystkich kart.



SCSI), OSLinux – Debian, a następnie serwer o konfiguracji: procesor Athlon – 1.3 GHz, pamięć RAM – 512 MB, pamięć dyskowa – 144 GB (2 x 72GB Ultra Wide SCSI LVD), OS Linux RedHat – 7.1. Podczas digitalizacji wykorzystywano oprogramowanie autorskie BJ: w procesie kontroli skanowania oprogramowanie własne z wykorzystaniem C, Perl, Java Script, w procesie konwersji formatu i obróbki grafiki – Image Magick, przy udostępnianiu w internecie – PostgreSQL 7, Apache 1.3.

W drugim etapie digitalizacji, w latach 2000–2003, prace kontynuowane były przez pracowników BJ w ramach obowiązków służbowych. Posuwały się one znacznie wolniej, digitalizowano ok. 2 skrzynki, tj. ok. 660 kart dziennie, co stanowiło ok. 40 skrzynek miesięcznie. Skanowano na skanerach płaskich HP model 6300c i 4570c, kontrolowano poprawność i kompletność obrazów graficznych kart, obrabiano pliki i wykonywano kopie na dyskach CD-ROM. Bazę e-PKA udostępniono w internecie 1 stycznia 1999 r. (ilustracja 13), a następnie

Biblioteka Jagiellońska

*Podstawowy Katalog Alfabetyczny Druków Zwartych wydanych do roku 1949*

**BIBLIOTECZNA GRAFICZNA BAZA DANYCH**

Realizacja I etapu projektu digitalizacji Katalogu.

Dotacja ze środków KBN, umowa nr 1217/W/DOT/DI/98

[O Katalogu](#) [Informacje o bazie danych](#) [Informacje o projekcie](#) [Biblioteka Jagiellońska](#)

**Przeszukiwanie Indeksu Katalogu**

Q R S T U V W X Y Z

**Wpisz hasło**  
**wyszukiwawcze:**  
( nazwisko autora lub rzeczownik w mianowniku )

srch

**Szukaj ?**

**Nasi sponsorzy**

Zalecana rozdzielczość 1024x768, minimum 800x600, High Color

Kodowanie polskich znaków: ISO-8859-2

Opracowanie WWW: Janusz Chwastek

e-mail: admwww@hpm.bj.uj.edu.pl

Ostatnia aktualizacja: 09/15/2004 11:25:19

Ilustracja 13. Pierwszy interfejs umożliwiający dostęp do e-PKA  
Źródło: Biblioteka Jagiellońska

sukcesywnie uzupełniano obrazy graficzne kolejnych kart z kolejnych skrzynek. Skanowanie ostatniej, 3120 skrzynki, zakończono w 2003 r. W sumie zdigitalizowano 970 346 kart. Realizacja tego przedsięwzięcia była ogromnym wyzwaniem ze względu na dużą liczbę wielkoformatowych kart katalogowych. Spora część kart była napisana ręcznie (atrament, tusz, ołówek, długopis), co miało wpływ na wybór technologii skanowania. Ze względu na unikalność katalogu i konieczność zapewnienia dostępu dla publiczności proces digitalizacji musiał być wykonany na terenie Biblioteki Jagiellońskiej. Nie było też możliwości zastosowania automatów skanujących z powodu odchyień w wymiarach kart, ich unikalności i różnorodnej gramatury papieru. Bezwzględny wymóg wiernego odwzorowania struktury katalogu i informacji zawartych na poszczególnych kartach wymuszał skomplikowane i czasochłonne operacje kontroli poprawności wszystkich procesów odwzorowywania.

Baza zawiera rekordy typu tekstowego, stanowiące podstawę indeksu alfabetycznego, oraz rekordy w postaci obrazów graficznych kart katalogowych Podstawowego Katalogu Alfabetycznego.

### **3. Rewitalizacja Starego Katalogu**

Podstawowy Katalog Alfabetyczny nie wymagał „przywrócenia do życia”, ponieważ, stanowiąc bezcenny zabytek kultury narodowej, był jednocześnie udostępniany czytelnikom jako jedyne źródło informacji o księgozbiorze w nim opisanym. Udostępnianie katalogu czytelnikom było konieczne. Roczna liczba kwerend i udzielonych informacji związanych z jego przeszukiwaniem szacowana była w 1996 r. na ok. 10 tys. i stanowiła 55% łącznej liczby 18 tys. kwerend i informacji związanych z wszystkimi zbiorami BJ [Elektroniczny... 1997]. Narażało to jednak katalog na fizyczne uszkodzenia, a z biegiem czasu na całkowite zniszczenie, ponieważ karty leżą luźno w skrzynkach, a papier znacznej ich części jest z powodu wieku słaby i kruchy. Nie było żadnej pełnej kopii katalogu, więc zastosowanie technologii komputerowych do wykonania i udostępnienia jego wersji cyfrowej w internecie, wzbogaconej o nowe narzędzia wyszukiwawcze i zintegrowanej z istniejącym systemem zdalnego zamawiania egzemplarzy w Bibliotece Jagiellońskiej, było rozwiązaniem służącym nie tylko udostępnieniu, ale i nadaniu mu funkcjonalności na miarę potrzeb współczesnego użytkownika. Udostępnienie bazy e-PKA umożliwiło też zamknięcie PKA dla publiczności, co zapewniło konserwację i ochronę tego katalogu (ilustracja 14).



Ilustracja 14. Zabytkowy Podstawowy Katalog Alfabetyczny druków wydanych do 1949 r. przechowywany w Bibliotece Jagiellońskiej  
Źródło: Biblioteka Jagiellońska

### 3.1. Narzędzia służące rewitalizacji

Baza e-PKA stanowi typowy CIPAC, zawiera cyfrowe duplikaty kart katalogowych PKA, które włożono w wirtualne skrzynki. Jest dostępna w internecie pod adresem <http://pka.bj.uj.edu.pl/PKA/>. Przez blisko 10 lat (do 2008 r.) uzupełniano skany, gdy nowa karta dla dokumentu wydanego do 1949 r. pojawiała się w zasobie PKA. Na początku 2008 r. zrezygnowano z uzupełniania PKA i tym samym zaprzestano uzupełniania e-PKA o nowe skany. Tym niemniej nie przewiduje się wycofania bazy e-PKA z użytkowania. Stanowi ona cyfrową kopię zabytkowego katalogu i jako katalog jest wykorzystywana. W 2012 r., dzięki modernizacji obsługującego ją oprogramowania, została wzbogacona o użyteczne funkcje. Zaprojektowano i wykonano nowy interfejs graficznie spójny z serwisem www Uniwersytetu Jagiellońskiego. Zastosowano także nową koncepcję nawigacji pomiędzy kartami katalogu, tj. podział strony na część elementów nawigacyjnych: listę z wykazem skrzynek lub listę z wykazem haseł oraz część z obrazem karty

katalogu wraz z opisem (hasła, adnotacje). Dotychczasowa wersja prezentowała na pojedynczym ekranie tylko listę skrzynek lub tylko obraz skanu karty, co znacznie ograniczało możliwość swobodnego poruszania się po katalogu.

### 3.2. Nawigacja za pomocą indeksów

Mechanizm wyszukiwania w indeksach e-PKA to funkcja wykorzystująca bazę danych zawierającą hasła przypisane do skanów kart katalogowych. Wynikiem wyszukiwania jest lista haseł pasujących do wyszukiwanego terminu oraz podgląd skanów kart odpowiadających hasłom z listy, z możliwością nawigacji pomiędzy kartami skanów. Funkcja ta w znaczący sposób rozszerzyła możliwości przeszukiwania bazy. Dotarcie do potrzebnej karty w PKA i w e-PKA przed wprowadzeniem indeksów wymagało ręcznego przewracania / przewijania kolejnych kart.

**BIBLIOTEKA JAGIELLOŃSKA**  
Katalog Podstawowy (zw. Stary) druków wydanych do 1949 roku włącznie

Wyszukiwanie skrzynek

A B C D E  
F G H I K  
L M N O P  
Q R S T U  
V W X Y Z

Wpisz poszukiwany termin ?

**Szukaj**

UWAGA!

Jeśli nad obrazem karty pojawi się pole **WYKAZ SKRZYNEK** to informacje w nim zawarte są bardziej aktualne niż zapisy na samej karcie.

wracaj

Klikaj znaki by zobaczyć hasła indeksu do skrzynek

O katalogu  
Zasady wkładki kart  
Instrukcja Korzystania  
Dostęp do druków online

Uwaga! Nowe nabytki, które wpływają do Biblioteki od r. 2008 nie są już rejestrowane w tym katalogu. Poszukiwane i zamawianie zbiorów BI należy rozpoczynać zawsze od Kompostowego Katalogu Zbiorów Biblioteki UJ, który notuje również wiele druków wydanych przed r.1950

<< Poprzedni	< Wstecz	WYKAZ SKRZYNEK	Dalej >	Następ >>
		WZ	XEM	# 3019
		XEN	XENOPHON ODSYŁACZE, DZIEŁA ZBIOROWE	# 3020
		XEHOPHON DZIEŁA FOSZCZ (A-Z)	XZ	# 3021
		Y	YOR	# 3022
		YOS	YZ	# 3023
		Z A	Z I	# 3024
		...	...	...
		Z K	Z L	# 3025
		...	...	...
		Z A	Z ABII	# 3026
		Z ABII	Z ABOK	# 3027
		Z ABOK	Z ACG	# 3028
		Z ACG	Z ACHEN	# 3029
		Z ACHEN	Z ADK	# 3030
		Z ADK	Z AGAM	# 3031
		Z AGAM	Z AGZ	# 3032
		Z AGZ	Z AHZ	# 3033

Ilustracja 15. Kolejny interfejs e-PKA (2004 r.)

Źródło: Biblioteka Jagiellońska

Początkowo bazę e-PKA można było zaliczyć do CIPAC-ów opartych na hasłach z wirtualnych skrzynek, odzwierciedlających hasła na skrzynkach fizycznych (naklejki z zakresem liter; ilustracja 15). Do wyszukiwania służył prosty indeks, który stanowił wykaz skrzynek, zawierający jedno- lub kilkukliterowe hasło początkowe i hasło końcowe (zakres literowy kart zebranych w skrzynce)

oraz numer skrzynki. Zainicjowanie wyszukiwania skrzynki z potrzebnym zakresem literowym umożliwiały litery A-Z, zgrupowane w prostokącie w lewym górnym rogu ekranu, opatrzonym napisem „wyszukiwanie skrzynki”. Poruszanie się po „wykazie skrzynek” zapewniały przyciski „dalej” i „wstecz” (wyświetlało się 20 następnych lub poprzednich skrzynek, aż do ostatniej skrzynki nr 3120). Aby wejść do skrzynki, należało kliknąć na hasło początkowe (wyświetlała się pierwsza karta) lub końcowe (wyświetlała się karta ostatnia). W obrębie skrzynki karty przewijało się ręcznie po jednej lub – oceniając jak daleko jest obraz karty z potrzebnym hasłem – w okienku „idź do karty” należało podać jej przewidywany numer (ilustracja 16). Jeżeli to nie dało odpowiedniego rezultatu, trzeba było przesuwać się do przodu lub do tyłu, aż do odnalezienia potrzebnego skanu. W trakcie wyszukiwania można było zmienić numer skrzynki lub wrócić do wykazu skrzynek, klikając na „powrót do wyszukiwania”.

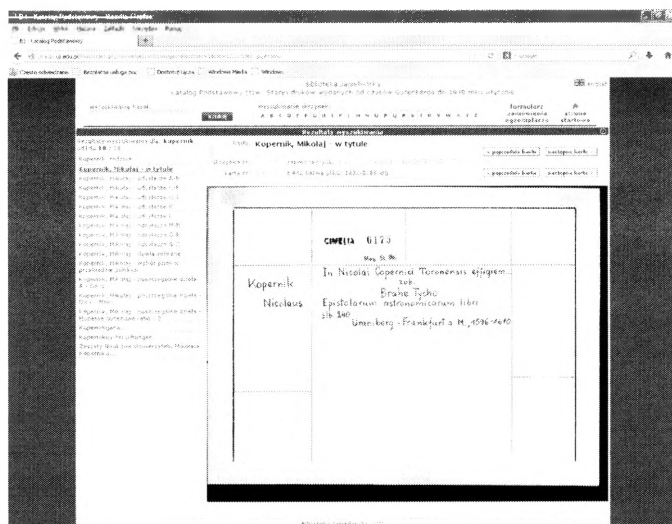


Ilustracja 16. Wyszukiwanie kolejnych kart w skrzynce  
 Źródło: Biblioteka Jagiellońska

W latach 2004-2012 sporządzono szczegółowy indeks haseł do kart w poszczególnych skrzynkach, przy zachowaniu struktury „wykazu skrzynek”. Po kliknięciu na znak „+” rozwija się wykaz haseł z danej skrzynki, który zwiija się po kliknięciu na znak „-”. Hasła indeksu ułożone są w takiej samej kolejności jak karty w skrzynce. Kliknięcie na hasło prowadzi do obrazu pierwszej karty z nim

związanej, do kolejnych kart przechodzi się, klikając przycisk „dalej” lub „następna karta”. W przypadku autorów hasła indeksu umożliwiają dostęp do pierwszej karty odsyłaczowej dotyczącej danej osoby, a jeżeli karty odsyłaczowej brak – do pierwszej karty głównej, dotyczącej dzieła jej autorstwa. W przypadku dzieł nieautorskich – umożliwiają dostęp do pierwszej karty danego tytułu, a dla haseł rzeczowych – dostęp do pierwszej karty z nimi związanej. Specjalny interfejs umożliwiał wprowadzanie haseł. Ich opracowanie wymagało niejednokrotnie sprawdzania nazw w bibliografiach, słownikach pseudonimów, katalogach *online* bibliotek i innych źródłach informacji oraz dokonywania melioracji. W sumie dla 3 120 skrzynek i 970 346 kart sporządzono i umieszczono w indeksach 321 394 haseł, czyli średnio trzy karty objęto jednym hasłem [Bułat 2012].

Dzięki indeksom możliwe jest wyszukiwanie za pomocą haseł (ilustracja 17). Poszukiwany termin należy wpisać w okienko „wyszukiwanie haseł” i potwierdzić klawiszem „enter” lub kliknięciem przycisku „szukaj”. Wielkość liter i znaki diakrytyczne nie są brane pod uwagę. Wyszukiwane są wszystkie hasła, które zawierają podany ciąg znaków (nie tylko te, które się od niego zaczynają). Klikając na wybrane hasło, przechodzi się do obrazu pierwszej związanej z nim karty. Do obrazów następnych kart, odnoszących się do danego hasła (jeżeli występują), przechodzi się, klikając odnośnik „następna karta”. Można też przejść do następnego hasła.



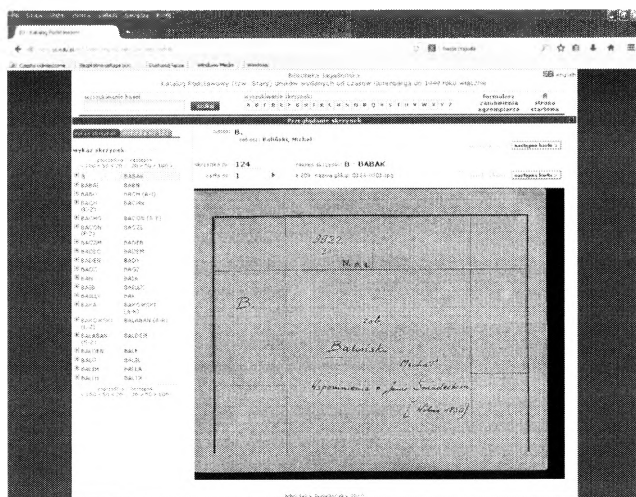
Ilustracja 17. Wyszukiwanie hasła „kopernik”

Z listy otrzymanych haseł wybrano drugie: „Kopernik, Mikołaj – w tytule”, wyświetlona została pierwsza karta, na której hasło „kopernik” występuje w tytule dzieła.

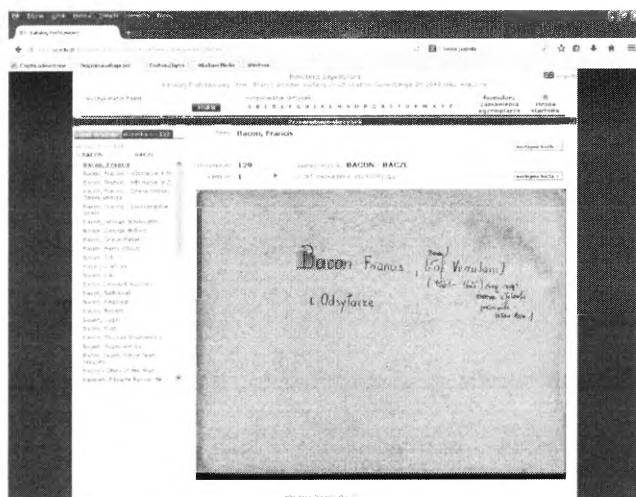
Źródło: Biblioteka Jagiellońska

### 3.3. Nawigacja za pomocą wirtualnych skrzynek

Drugi sposób wyszukiwania to wyszukiwanie skrzynek i przeglądanie ich zawartości, inicjowane wybraniem litery identyfikującej skrzynkę (ilustracja 18). W następnych krokach można przeszukiwać zawartości skrzynki, korzystając z przycisku „następna karta”, wpisując w okienko przewidywany numer potrzebnej karty lub skorzystać z indeksu haseł ukrytych pod przyciskiem „+” (ilustracja 19).



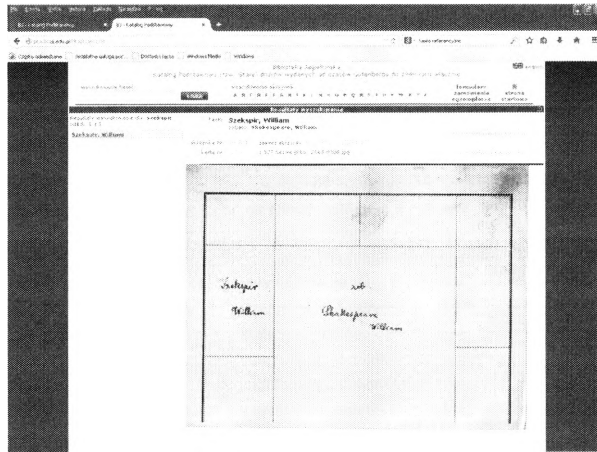
Ilustracja 18. Wykaz wirtualnych skrzynek na literę B  
Źródło: Biblioteka Jagiellońska



Ilustracja 19. Wykaz haseł związanych ze skrzynką o zakresie literowym BACON – BACZL  
Źródło: Biblioteka Jagiellońska

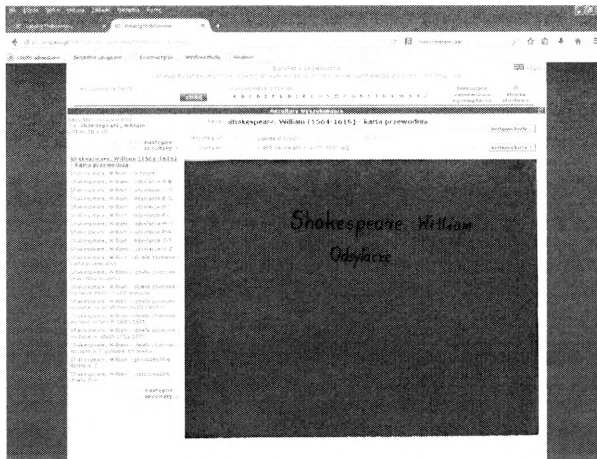
### 3.4. Nawigacja za pomocą mechanizmu linkowania haseł

Funkcjonalność e-PKA obejmuje mechanizm linkowania haseł odsyłaczy do odpowiednich kart (ilustracja 20). W przypadku kart, do których przypisane są odsyłacze całkowite („zob.”), hasło odrzucone (nieużywane w katalogu) linkowane jest z hasłem przyjętym (używany). Hasło przyjęte, prezentowane w opisie karty, jest jednocześnie linkiem do nowego okna, zawierającego wykaz haseł obejmujący hasło przyjęte oraz skan / skany kart katalogowych, których ono dotyczy (ilustracja 21).



Ilustracja 20. Karta odsyłacza całkowitego uzyskana po wyszukaniu hasła odrzuconego „Szekspir”. Hasło przyjęte („Shakespeare”) stanowi link do wykazu haseł zawierających to hasło.

Źródło: Biblioteka Jagiellońska

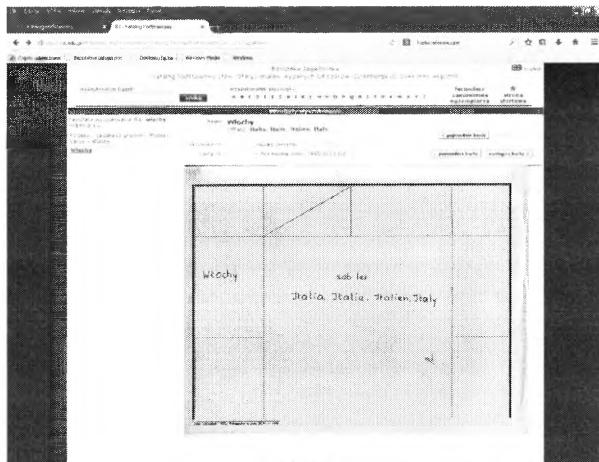


Ilustracja 21. Wykaz haseł związanych z hasłem przyjętym w odsyłaczu typu „zob.”

Źródło: Biblioteka Jagiellońska

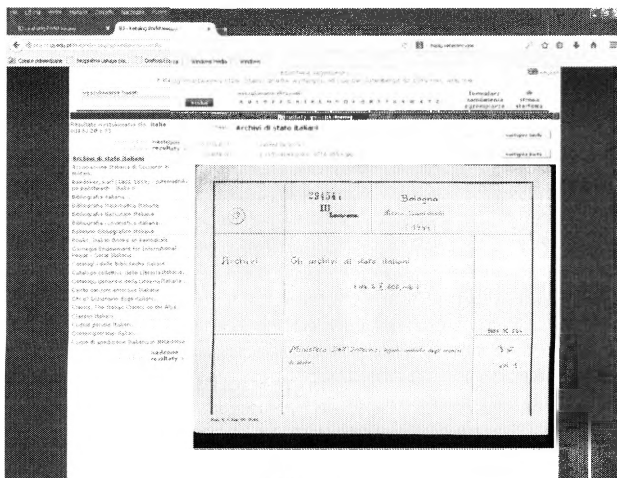


Dla kart, do których przypisane są odsyłacze uzupełniające („zob. też”), hasło będące nazwą stosowaną w e-PKA linkowane jest z równorzędnymi wariantami tego hasła, także używanymi w katalogu (ilustracja 22). Poszczególne warianty są jednocześnie odnośnikami – po kliknięciu na każdy z nich automatycznie wykonywane jest związanie z nim, nowe wyszukiwanie. Jako rezultat wyświetla się lista rankingowa zawierająca karty opisane tą formą hasła (ilustracja 23).



Ilustracja 22. Karta z odsyłaczem uzupełniającym „zob. też” i odnośnikami do hasła, z którymi związane są obrazy kart.

Źródło: Biblioteka Jagiellońska



Ilustracja 23. Lista haseł będąca rezultatem kliknięcia na pierwszy odnośnik („Italia”) w odsyłaczu uzupełniającym „zob. też”, z którym związane są obrazy kart.

Źródło: Biblioteka Jagiellońska

### 3.5. Adnotacje – aktualizacja danych

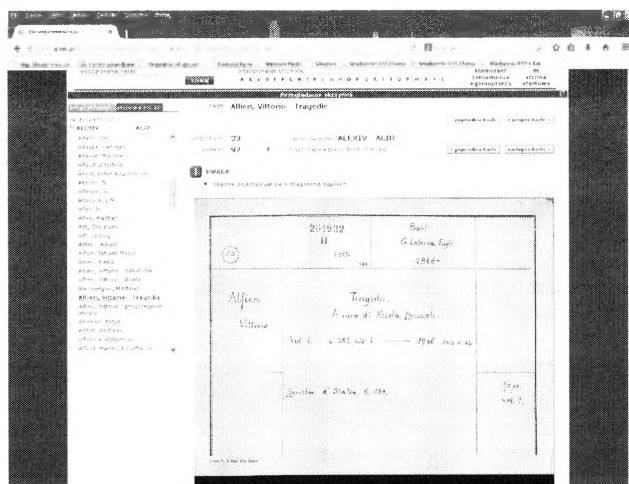
Aby zaktualizować istotne dane zawarte na obrazach kart, zastosowano mechanizm tworzenia odpowiednich adnotacji (ilustracja 24), czyli uwag pojawiających się nad niektórymi kartami.

**EDYCJA ADNOTACJI**

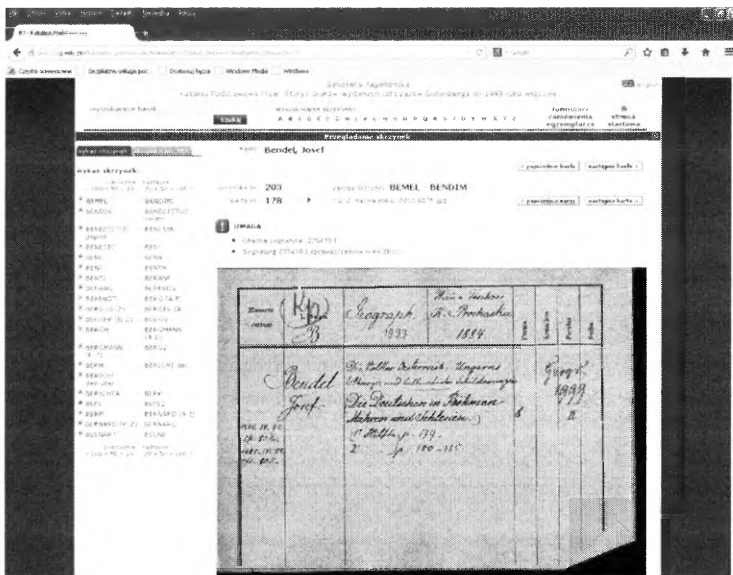
1	<input type="button" value="Dodaj"/>	Obecna sygnatura: <input type="text"/>
2	<input type="button" value="Dodaj"/>	<input type="text"/> obecnie przechowywane w: <input type="text" value="magazynie ogólnym"/>
3	<input type="button" value="Dodaj"/>	Udostępnia się także na mikrofilmie <input type="text"/>
4	<input type="button" value="Dodaj"/>	Stwierdzono brak <input type="text"/>
5	<input type="button" value="Dodaj"/>	Brak zastąpiono nowym/udostępnionym egzemplarzem <input type="text"/>
6	<input type="button" value="Dodaj"/>	<input type="text"/>
7	<input type="button" value="Dodaj"/>	Sygnatura <input type="text"/> <input type="button" value="sprawdź/zamów w KKZBUJ"/>
8	<input type="button" value="Dodaj"/>	Pełny tekst tej pozycji Biblioteka Jagiellońska udostępnia przez Polską Bibliotekę Internetową. <a href="http://">http://</a> <input type="text"/> <input type="button" value="pokaż ..."/>
9	<input type="button" value="Dodaj"/>	Dostęp online <input type="text"/> <small>[nałamy wpisać pełny URL: <a href="http://">http://</a>....]</small>

Ilustracja 24. Formularz służący do nanoszenia adnotacji  
Źródło: Biblioteka Jagiellońska

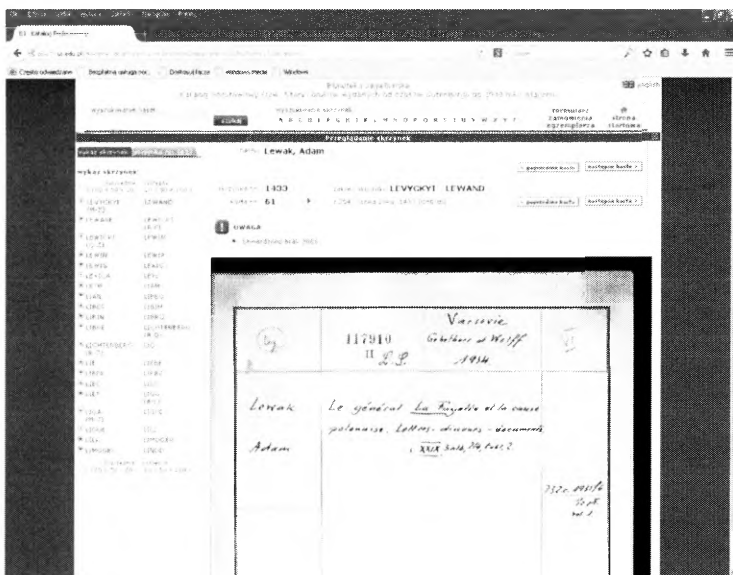
Są to uwagi informujące np. o aktualnym miejscu przechowywania egzemplarza (ilustracja 25), o zmianie sygnatury (ilustracja 26), stwierdzeniu braku (ilustracja 27), o dostępności kopii mikrofilmowej dokumentu (ilustracja 28).



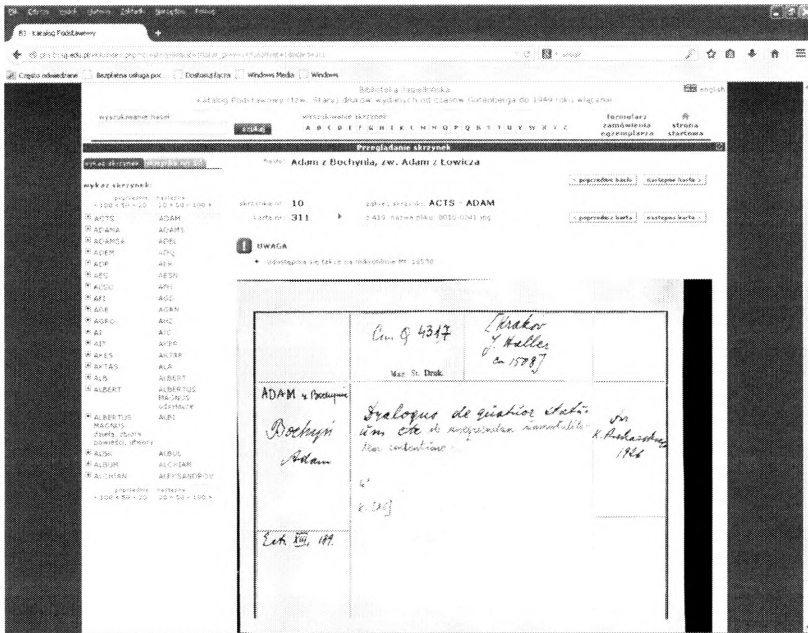
Ilustracja 25. Adnotacja o miejscu przechowywania dokumentu  
Źródło: Biblioteka Jagiellońska



Ilustracja 26. Adnotacja podająca aktualną sygnaturę  
 Źródło: Biblioteka Jagiellońska

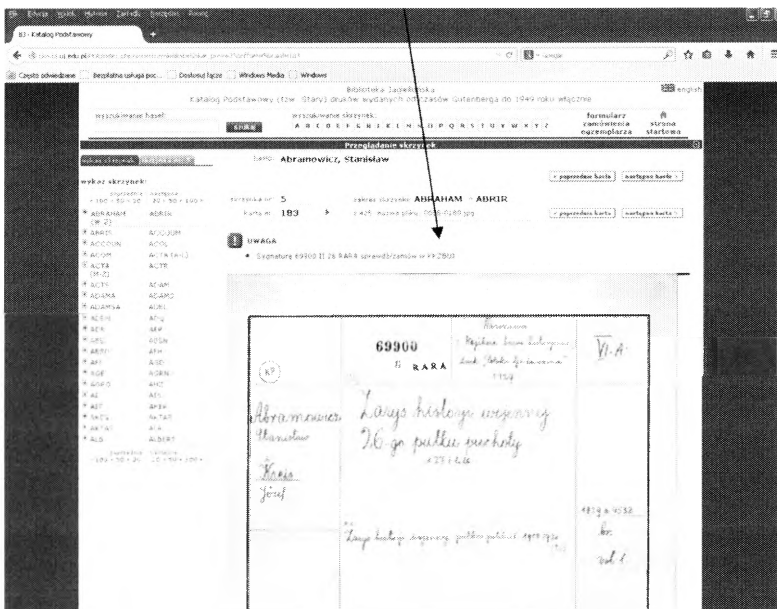


Ilustracja 27. Adnotacja informująca o stwierdzeniu braku  
 Źródło: Biblioteka Jagiellońska



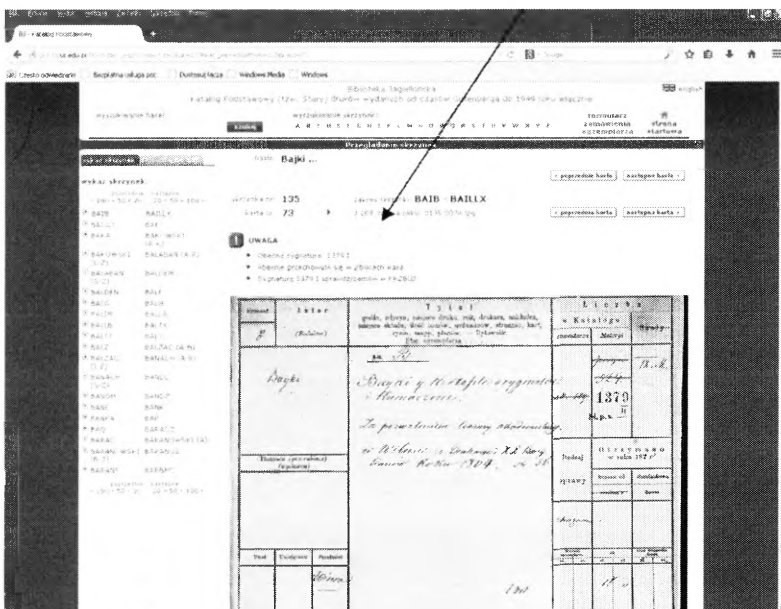
Ilustracja 28. Adnotacja o dostępności mikrofilmu dokumentu  
 Źródło: Biblioteka Jagiellońska

Wprowadzane są także adnotacje zawierające odnośniki / linki wiążące dane w bazie e-PKA z danymi w innych bazach. Funkcja adnotacji odsyłających do Komputerowego Katalogu Zbiorów Bibliotek UJ (adnotacja o treści: „Sygn.: sprawdź / zamów w KKZBUJ”, ilustracja 29) polega na podaniu aktywnego, bezpośredniego linku / odesłania do listy sygnatur, zawierającej sygnaturę wymienioną w adnotacji. Zwiększa to znacząco komfort pracy użytkownika, który egzemplarze odnotowane w obu bazach, e-PKA i KKZBUJ musi zamówić w bazie KKZBUJ. Adnotacje o treści „dostęp online” (link pod słowem *online*) umożliwiają połączenie z pełnym tekstem dokumentu opisanego na obrazie karty katalogowej (ilustracja 31), dostępnym w Jagiellońskiej Bibliotece Cyfrowej (ilustracja 32) lub w Federacji Bibliotek Cyfrowych (ilustracja 33).



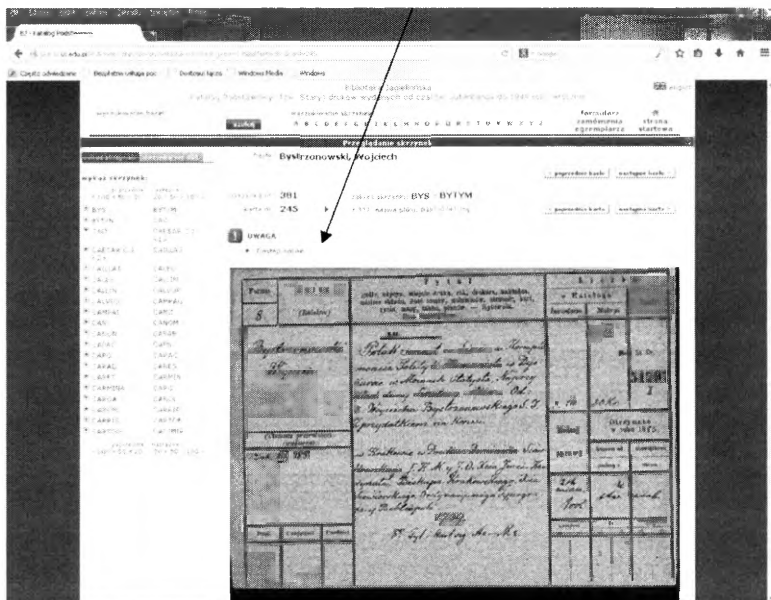
Ilustracja 29. Adnotacja przenoszona do bazy KKZBU, gdzie dokument można zamówić (link pod KKZBU).

Źródło: Biblioteka Jagiellońska



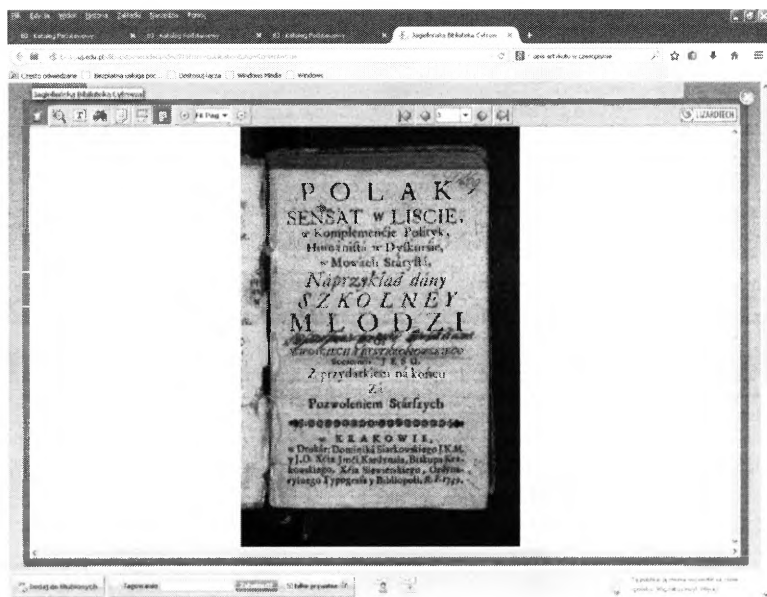
Ilustracja 30. Adnotacji może być kilka.

Źródło: Biblioteka Jagiellońska



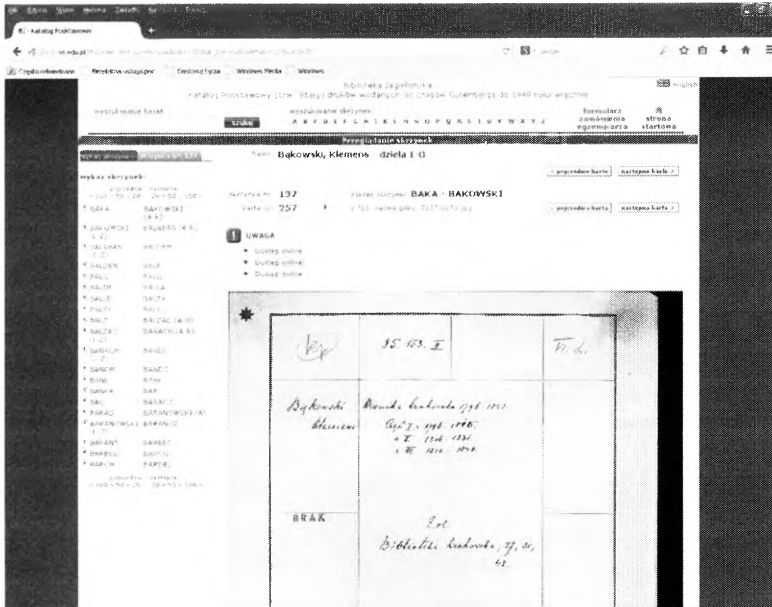
Ilustracja 31. Adnotacja umożliwiająca połączenie z pełnym tekstem dokumentu (link pod słowem *online*)

Źródło: Biblioteka Jagiellońska



Ilustracja 32. Pełny tekst w Jagiellońskiej Bibliotece Cyfrowej, uzyskany po kliknięciu na odnośnik *online* w adnotacji do karty, widocznej na poprzedniej ilustracji (31)

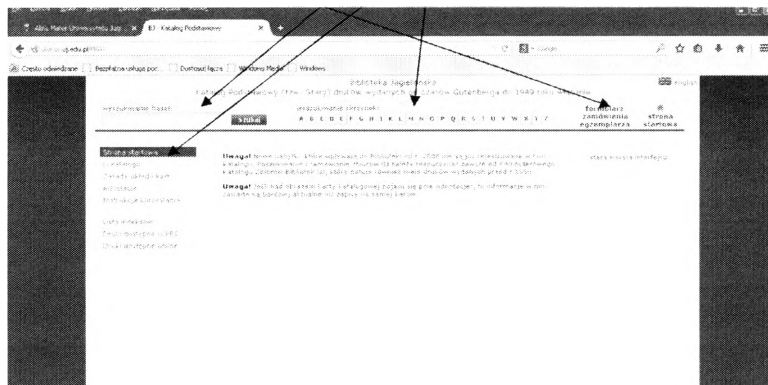
Źródło: Biblioteka Jagiellońska



Ilustracja 33. Adnotacja umożliwiająca dostęp *online* do pełnego tekstu (tu dostępnego w Małopolskiej Bibliotece Cyfrowej)  
Źródło: Biblioteka Jagiellońska

### 3.6. Nowoczesny interfejs e-PKA

Nowoczesny interfejs bazy e-PKA zaprojektowany został w stylu Portalu BJ (ilustracja 34). Na stronie startowej widnieje nazwa własna katalogu (nie w każdym CIPAC-u występuje), dostępne są narzędzia do zainicjowania wyszukiwania i przeglądania zawartości bazy oraz formularz zamówienia egzemplarza. Dla wygody czytelnika zamieszczono również szczegółową instrukcję korzystania z bazy, informacje o układzie kart w katalogu, o adnotacjach i o samym katalogu. Udostępniono listę indeksów wyszukiwawczych i listy druków dostępnych *online*. Dostępna jest także angielskojęzyczna wersja interfejsu bazy e-PKA.



Ilustracja 34. Strona startowa bazy e-PAK – narzędzia do zainicjowania wyszukiwania, formularz zamówienia, informacje o katalogu  
 Źródło: Biblioteka Jagiellońska

## 4. Podsumowanie

Rewitalizacja Podstawowego Katalogu Alfabetycznego polegała na utworzeniu CIPAC-u, bazy obrazów kart katalogowych, odwzorowującej fizyczny układ Starego Katalogu. Umożliwiło to powszechny, nieograniczony w czasie i przestrzeni dostęp do informacji o cennych zbiorach Biblioteki Jagiellońskiej i zdalne ich zamawianie. Przeglądanie wirtualnych kart odbywa się analogicznie do przeglądania katalogu kartkowego. Wykorzystane narzędzia informatyczne znacznie usprawniły dostęp i korzystanie z katalogu, umożliwiając też utworzenie powiązań / linków z innymi bazami dostępnymi w trybie *online*. Rewitalizacja umożliwiła więc z jednej strony zachowanie cech tradycyjnego katalogu kartkowego, a z drugiej – niezwykle rozszerzyła dostęp i usprawniła korzystanie z tego źródła informacji. Ważnym aspektem projektu była też ochrona samego zabytkowego katalogu. Liczący 200 lat Stary Katalog, stanowiący zabytek bibliotekarstwa, w swej fizycznej formie został wyłączony z użytkowania. Ocalony przed zniszczeniem, zaczął pełnić wyłącznie funkcje muzealne.

Obecnie często spotyka się opinię, że jeżeli coś nie jest dostępne w internecie, to nie istnieje. Zrewitalizowany Podstawowy Katalog Alfabetyczny druków wydanych od czasów Gutenberga do 1949 roku został udostępniony w internecie, spełnia swoją rolę i będzie ją spełniał zapewne jeszcze przez dłuższy czas. Wiadomo, że najbardziej korzystnym i pożądanym rozwiązaniem byłoby wprowadzenie pełnej informacji o wszystkich zbiorach Biblioteki Jagiellońskiej do OPAC-u (Komputerowego Katalogu Zbiorów Bibliotek UJ). Dlatego opracowanie retro-



spektywne (rekatalogowanie) jest priorytetowym zadaniem biblioteki, a prace w tym zakresie prowadzone są intensywnie. Jest to jednak długi i kosztowny proces – po 20 latach tworzenia OPAC-u odnotowano w nim ok. 34 % ogólnej liczby zbiorów Biblioteki Jagiellońskiej. Stosuje się więc też inne rozwiązania, jak to opisane w artykule.

## PRZYKŁADY KATALOGÓW TYPU CIPAC W BIBLIOTEKACH POLSKICH

1. Biblioteka Główna Uniwersytetu Ekonomicznego w Krakowie  
Katalog Klamrowy do 1954, <http://ikar1.uek.krakow.pl>  
Katalog Kartkowy 1955-1993, <http://ikar2.uek.krakow.pl>
2. Biblioteka Akademii Muzycznej w Krakowie  
Elektroniczny Katalog Kartkowy Biblioteki Akademii Muzycznej w Krakowie,  
<http://149.156.54.39/prezentacja/Default.aspx>
3. Biblioteka Główna Akademii Górniczo-Hutniczej w Krakowie  
Katalog Kartkowy BG AGH, <http://wbg.bg.agh.edu.pl/katalog/prezentacja/Default.aspx>
4. Biblioteka Uniwersytecka w Warszawie  
Cyfrowe Katalogi Kartkowe, BUW, <http://ckk.buw.uw.edu.pl>
5. Biblioteka Uniwersytecka w Poznaniu  
Cyfrowy Katalog Biblioteki, <http://buuam.digital-center.pl/#>
6. Biblioteka Uniwersytecka we Wrocławiu  
Katalogi zdigitalizowane, <http://www.bu.uni.wroc.pl/katalogi/zdigitalizowane>
7. Zakład Narodowy im. Ossolińskich  
Katalog Kartkowy Druków Zwartych i Ciągłych XIX i XX w., <http://www2.oss.wroc.pl/kat>
8. Biblioteka Główna Politechniki Wrocławskiej  
Katalogi Kartkowe BG PWR, <http://karty.bg.pwr.wroc.pl>
9. Biblioteka Główna – Politechnika Gdańska  
Katalog Kartkowy, <http://bg.pg.edu.pl/kartkowy-online>
10. Główna Biblioteka Lekarska  
Katalog Kartkowy 1972-1999  
<http://195.187.98.8/index.php/pl/przegląd/katalog-ksiazek>
11. Książnica Pomorska im. Stanisława Staszica w Szczecinie  
Zeskanowany Katalog Czasopism  
<http://www.ksiaznica.szczecin.pl/www/katalog-czasopism>
12. Biblioteka Publiczna m. st. Warszawy  
Katalog Alfabetyczny Książek, <http://ikar1.koszykowa.pl>  
Sprawozdania i czasopisma zagraniczne, <http://ikar2.koszykowa.pl>
13. Biblioteka Uniwersytecka w Toruniu  
Katalogi Kartkowe *online*, <http://www.bu.umk.pl/katalogi>

14. Komputerowy Katalog Kartkowy Łódzkiej Akademickiej Sieci Bibliotecznej  
<http://biblioteka.asp.lodz.pl/katalogi/kkk>
15. Biblioteka Główna Gdańskiego Uniwersytetu Medycznego  
Zdigitalizowane Katalogi Kartkowe  
<http://biblioteka1.gumed.edu.pl/katalogzeskanowany/index.php?kat=-1>

## BIBLIOGRAFIA

- Bułat, Barbara (2012). *Indeks do elektronicznej wersji starego katalogu – krótka historia* [prezentacja]. Materiały własne BJ.
- Elektroniczny Katalog Podstawowy Druków Zwartych Biblioteki Jagiellońskiej wydanych do 1949 r.* (1997). Wniosek do KBN o dofinansowanie rozbudowy i udostępniania baz danych o charakterze regionalnym lub ogólnokrajowym. Maszynopis, 15 s.
- Oberhauser, Otto Carl (2014). *The International CIPAC List* [dok. elektr.]. <http://cipacs.vfi-online.org/index.html> [odczyt: 15.05.2014].
- Oberhauser, Otto Carl (2003). *Card-image public access catalogues (CIPACs): an international survey*. Program: Electronic Library and Information Systems Vol. 37, No. 2, p. 73-84.

---

## KKZBUJ – OPAC nowej generacji?

Krystyna Sanetra

Biblioteka Jagiellońska

---

### KKZBUJ – charakterystyka bazy

Komputerowy Katalog Zbiorów Bibliotek UJ (KKZBUJ), dostępny pod adresem <https://chamo.bj.uj.edu.pl/uj/>, to centralny katalog Uniwersytetu Jagiellońskiego. Jest to baza typu OPAC (Online Public Access Catalog) umożliwiająca dostęp do zasobów wszystkich bibliotek tworzących jednolity system biblioteczno-informacyjny Uniwersytetu, tj. Biblioteki Jagiellońskiej, Biblioteki Medycznej, bibliotek wydziałowych i instytutowych UJ. Baza KKZBUJ utrzymywana i obsługiwana jest przez oprogramowanie VTLS/Virtua (od 2003 r., poprzednio VTLS).

Biblioteka Jagiellońska rozpoczęła wprowadzanie danych do zintegrowanego systemu bibliotecznego VTLS w 1994 r., równocześnie udostępniając komputerowy katalog *online*, początkowo przez telnet wyłącznie w Bibliotece Jagiellońskiej na 10 stanowiskach komputerowych, a od 1996 r. za pomocą przeglądarki internetowej – VTLS Gateway. Od 2003 r. dostęp dla czytelników przez www zapewniany był przez interfejs VTLS/Virtua Chameleon/iPortal, następnie od 2006 r. przez interfejs VTLS/Virtua Vectors iPortal, a od grudnia 2014 r. przez VTLS/Virtua CHAMO.

W 1994 r. współkatalogowanie w ramach KKZBUJ rozpoczęła Biblioteka Instytutu Informatyki UJ, w 1995 r. Biblioteka Instytutu Fizyki oraz Biblioteka Instytutu Matematyki. W następnych latach do współpracy przy tworzeniu wspólnej bazy katalogowej przystąpiły kolejne biblioteki UJ. W 1998 r. uruchomiono użytkowanie modułu rejestracji wypożyczeń i zwrotów w Wypożyczalni Miejscowej BJ, a następnie stopniowo w kolejnych bibliotekach UJ. Wkład bibliotek UJ we współtworzenie KKZBUJ przedstawia ilustracja 3. Od 2002 r. opracowanie zbiorów odbywa się we współpracy z katalogiem centralnym polskich bibliotek

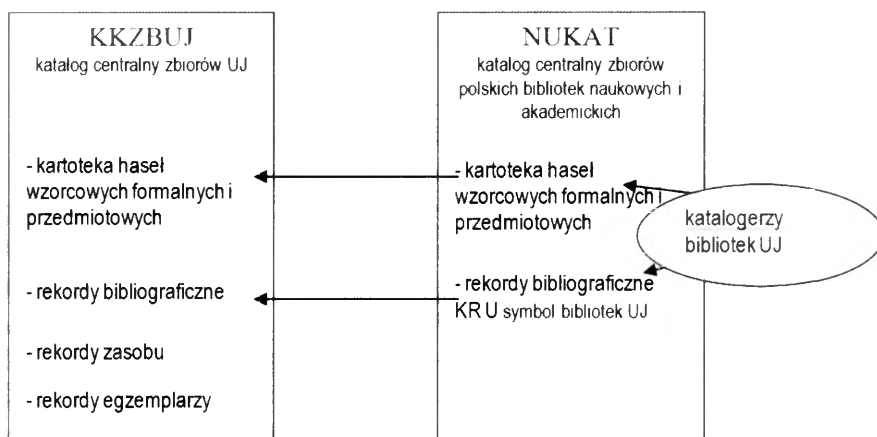
naukowych i akademickich NUKAT. Wkład bibliotek UJ we współtworzenie NUKAT przedstawia ilustracja 2.

Fakt, że baza KKZBUJ od początku jej istnienia podłączona była do globalnej sieci i była zdalnie dostępna, umożliwił przełamanie wyizolowania katalogów kartkowych poszczególnych bibliotek sieci UJ, zastosowanie tych samych standardów i zasad katalogowania oraz tworzenie katalogu centralnego, prezentującego zasoby bibliotek całej sieci UJ. Wiadomo bowiem, że „Sharing of data among libraries has created great efficiencies in catalog production, and it has also expanded the available universe of resources for library users” [Coyle 2010, p. 10]. KKZBUJ, autonomiczna baza katalogu centralnego UJ, jest bazą lokalną w stosunku do katalogu centralnego NUKAT. Metadane zawarte w obu bazach są ze sobą zharmonizowane i oddziałują na siebie. Metadane w NUKAT komunikują się z metadanymi w KKZBUJ za pomocą hiperlinków i internetu.

Podstawowym elementem danych w katalogu komputerowym, a więc i w bazie KKZBUJ, jest rekord bibliograficzny, na który składa się opis bibliograficzny dokumentu, dla którego został on utworzony, hasła ujednolicone i inne punkty dostępu oraz numery i kody niezbędne do prawidłowego przetwarzania danych. Rekord bibliograficzny, a w szczególności zawarty w nim opis bibliograficzny, identyfikuje wydanie dokumentu (wszystkie egzemplarze danego wydania). Informacje o indywidualnych cechach egzemplarza dokumentu (np. sygnatura, lokalizacja) zapisywane są w rekordach egzemplarzy. Rekordy bibliograficzne wchodzą w relacje z innymi rekordami bibliograficznymi, np. rekord bibliograficzny dla tytułu wydawnictwa ciąglego związany jest (link) z tytułem poprzednim i / lub następnym tego wydawnictwa, z rekordami khw dla haseł użytych w danym rekordzie bibliograficznym, z rekordami egzemplarzy (z tyłoma, ile egzemplarzy dzieła posiada biblioteka) i z rekordami zasobu. Wspomniane rodzaje rekordów wykorzystywane są dla potrzeb wyszukiwania informacji o zbiorach bibliotek sieci UJ, a także dla potrzeb katalogowania oraz rejestracji wypożyczeń i zwrotów.

Rekordy bibliograficzne nie są bezpośrednio wprowadzane do KKZBUJ – sporządzane są przez katalogerów w NUKAT z zaznaczeniem w polu 962 symbolu BJ (KR U). Pliki z rekordami nowymi, które weszły do bazy NUKAT, wystawiane są codziennie do pobrania i automatycznie pobierane przez partycypujące biblioteki. Do bazy KKZBUJ wgrywane są tylko te oznaczone symbolem KR U. Podobnie pobierane są z NUKAT pliki z rekordami bibliograficznymi zmodyfikowanymi. W bazie KKZBUJ dotychczasowe wersje rekordów bibliograficznych są nadpisywane przez nowe / zmodyfikowane wersje, pod warunkiem że obie wersje mają identyczny numer kontrolny w polu 035. Również rekordy kartoteki haseł wzorcowych, nowe i zmodyfikowane, wprowadzane są do KKZBUJ w ana-

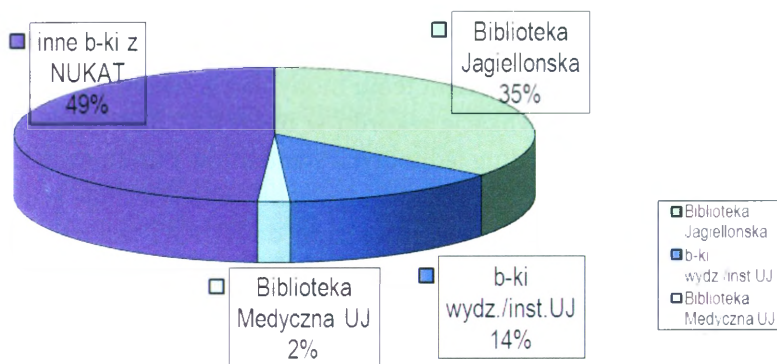
logiczny sposób. Kluczową rolę przy takim przepływie rekordów (ilustracja 1) ma identyczność numeru kontrolnego w obu bazach (numer systemowy rekordu jest właściwy dla danej bazy). Dzięki temu modyfikacja rekordu w NUKAT powoduje automatyczną modyfikację rekordu w bazie KKZBUJ. Rekordy bibliograficzne sporządzane są według jednolitych standardów i ustalonych zasad, tak aby opis wystarczająco identyfikował katalogowane dzieło. Cechuje je uniwersalność, która sprawia, że rekord raz wprowadzony do NUKAT jest pobierany wielokrotnie, do różnych baz, gdzie wciąż identyfikuje ten sam dokument. Podobna idea przyświeca sporządzaniu haseł wzorcowych, z tym że identyfikowane w tym przypadku jest hasło: formalne (dla nazwy osobowej, nazwy instytucji, imprezy, tytułu ujednoliconego, tytułu serii) lub przedmiotowe.



Ilustracja 1. Przepływ danych pomiędzy NUKAT i KKZBUJ



Ilustracja 2. Udział KKZBUJ we współtworzeniu NUKAT – rekordy bibliograficzne, stan na 31.12.2013 r.

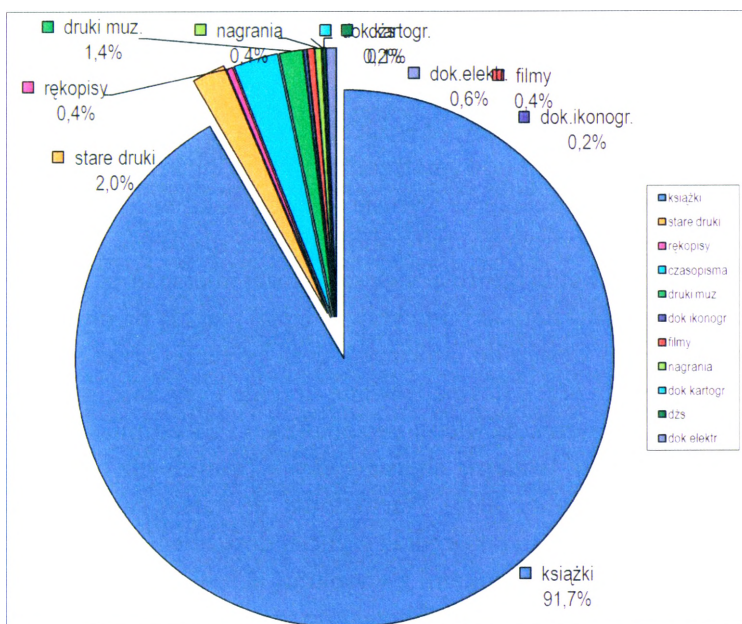


Ilustracja 3. Udział bibliotek UJ we współtworzeniu KKZBUJ – rekordy bibliograficzne, stan na 31.12.2013 r.

W KKZBUJ rekordy bibliograficzne uzupełniane są o pewne pola lokalne, zachowywane w tej bazie także przy wgrywaniu kolejnej wersji rekordu pochodzącej z NUKAT. W KKZBUJ do rekordów bibliograficznych dopinane są rekordy zasobu (dla wydawnictw ciągłych, wydawnictw wielotomowych) oraz rekordy egzemplarza. Rekordy zasobu i rekordy bibliograficzne nie są przedmiotem wymiany między bazami. W rekordach egzemplarza niezbędne jest wczytanie kodu paskowego z dokumentu, który jest konieczny do wykonywania transakcji wypożyczeń i zwrotów. Obok sygnatury bardzo istotną daną jest także oznaczenie lokalizacji egzemplarza, ponieważ do jednego uniwersalnego rekordu bibliograficznego podpinane są rekordy egzemplarzy należących do różnych bibliotek sieci UJ. Rekordy bibliograficzne są wspólne dla zasobów bibliotek UJ, zaś rekordy egzemplarzy – indywidualne dla każdego egzemplarza każdej biblioteki. Każda z bibliotek oznacza swoje egzemplarze własnymi sygnaturami i przypisuje je do swojej lokalizacji.

W KKZBUJ opracowywane i notowane są wszystkie typy dokumentów, zarówno analogowe, jak i elektroniczne (ilustracja 4). Dla dokumentów zdigitalizowanych, posiadających reprodukcje cyfrowe w Jagiellońskiej Bibliotece Cyfrowej, sporządzane są rekordy bibliograficzne dla dokumentu tradycyjnego – podstawy reprodukcji. To, że dokument posiada kopię cyfrową, zaznaczane jest w polu uwag, a w polu 856 dodawany jest adres elektroniczny reprodukcji cyfrowej, który w KKZBUJ stanowi dynamiczny link do kopii cyfrowej w JBC.

Dane (rekordy bibliograficzne, zasobu, rekordy khw, rekordy czytelnika) do KKZBUJ wprowadzane są w formacie MARC 21. Dostęp do danych zapewniały głównie indeksy: autorów, tytułów, tytułów czasopism, haseł przedmiotowych, haseł przedmiotowych w języku KABA, haseł przedmiotowych w języku MeSH,



Ilustracja 4. KKZBUJ – różne typy dokumentów – stan na 17.06.2014 r.

hasel przedmiotowych w języku angielskim i francuskim, indeksy sygnatur. Wyszukiwanie proste przez słowa obejmowało pola związane z autorami, tytułami, hasłami przedmiotowymi, ISBN, ISSN lub jednocześnie wszystkie wymienione kategorie. Możliwe było także wyszukiwanie zaawansowane oraz wyszukiwanie przez słowa w hasłach: tytułowych, autorskich i przedmiotowych.

## Next Generation Catalog a baza KKZBUJ dostępna za pomocą VTLS/Virtua CHAMO

Termin „katalog nowej generacji” (KNG, Next Generation Catalog = NGC) zaczęto stosować wraz z założeniem przez Erica Lease Morgana listy dyskusyjnej NGC4Lib mailing list, mającej umożliwić dyskusję na temat następnej generacji katalogów bibliotecznych typu OPAC [Nagy 2011]. W dyskusjach katalog taki określany był czasem jako katalog biblioteki 2.0 lub katalog 3. generacji [Yang, Hoffmann 2010]. Morgan przedstawił w swoim blogu ideę katalogu nowej generacji, poczynając od ogólnego stwierdzenia, że taki katalog nie jest rzeczywistym katalogiem, a raczej narzędziem tak zaprojektowanym, aby studentom łatwiej było się uczyć, nauczycielom nauczać, a naukowcom prowadzić badania. Zapewnia

ono użytkownikom narzędzia / sposoby wyszukania i wykorzystania danych i informacji. W trakcie dyskusji w literaturze pojawiło się wiele głosów rozważających problem z różnych punktów widzenia [Morgan 2006]. Merčun i Žumer w swoim eksperckim studium podkreślały, że wizja i potrzeba KNG wynika przede wszystkim z pojawienia się, wraz z nowymi trendami Webu, nowej generacji użytkowników. Chodzi głównie o pokolenie wyrosłe w środowisku technologii komputerowej, nazywane często cyfrowymi tubylcami, o *native speakerów* języka cyfrowego. Ich potrzeby i oczekiwania informacyjne są w dużej mierze inne. Jako użytkownicy oczekują interfejsów bardzo intuicyjnych, lepiej reagują na obraz niż na tekst, preferują nieliniowy dostęp do informacji, uczą się raczej przez testowanie i popełnianie błędów niż przez sformalizowane uczenie i czytanie, są współpracujący i wielozadaniowi, oczekują więcej personalizacji i natychmiastowej satysfakcji oraz mają świadomość istnienia alternatywnych (innych niż biblioteka) źródeł informacji [Merčun, Žumer 2008]. Warto dodać, że starsza generacja użytkowników, tzw. cyfrowi imigranci, także coraz szybciej dostosowują się do zmian, stając się profesjonalnymi użytkownikami z wysokimi oczekiwaniami. Tak więc nowa generacja użytkowników zmusza biblioteki do zanalizowania, jak ich serwisy nadążają za potrzebami użytkowników i za zmianami technologicznymi.

Charakterystyczne cechy KNG zestawiał i opisał szeroko Marshall Breeding [Breeding 2007]. Morgan wymienił 4 główne zasady KNG [Morgan 2006]. Yang i Hofmann w studium porównującym katalogi *online* w oprogramowaniach Koha, Evergreen i Voyager wymieniają 10 cech nowoczesnego katalogu, odzwierciedlających punkt widzenia bibliotekarzy, które miałyby raczej uzupełniać niż zastępować cechy tradycyjnego katalogu (OPAC-u). Ci sami autorzy w studium dotyczącym 260 bibliotek akademickich USA i Kanady wzięli pod uwagę 12 funkcjonalności KNG, z czego 10 jest tożsamyh z wyżej wspomnianymi [Yang, Hofmann 2011].

VTLS/Virtua CHAMO (CHAMO) to oparty na technologii webowej OPAC 2.0 umożliwiający użytkownikom dostęp do zasobów bibliotecznych UJ, wdrożony w grudniu 2014 r., zastępujący wykorzystywany dotąd Vectors iPortal. Interfejs CHAMO obsługuje nowe funkcje katalogu, ma bardziej nowoczesny wygląd oraz cechy społecznego OPAC-u. Wdrożono podstawowy samoistny moduł interfejsu. Dodatkowe, opcjonalne, wymagające dalszych nakładów finansowych funkcje VTLS-CHAMO Discovery to: Vitiation – do tworzenia cytowań bibliograficznych, Discovery Extension – pozwalający podłączyć także zewnętrzne zasoby biblioteczne (bazy danych) i stworzyć multiwyszukiwarkę, Drupal API – moduł umożliwiający zintegrowanie katalogu z innymi serwisami (np. z portalem biblioteki).

Uznając wspomniane wyżej 12 cech i funkcjonalności współczesnego katalogu OPAC nowej generacji za pewien konsensus, w niniejszym artykule przedsta-



wiono te cechy i porównano je z funkcjami i cechami KKZBUJ, dostępnego dla użytkowników za pomocą interfejsu VTLS/Virtua CHAMO.

## 1. Zintegrowane wyszukiwanie wszystkich zasobów biblioteki, dostępnych w bibliotece i poprzez bibliotekę

KNG powinien obejmować wszystkie zasoby biblioteki, nie tylko informacje bibliograficzne dotyczące książek, czasopism, zbiorów specjalnych, dokumentów elektronicznych na nośnikach fizycznych, ale także prowadzić do pełnych tekstów książek i artykułów z czasopism zawartych w subskrybowanych bazach elektronicznych oraz do biblioteki cyfrowej (ilustracja 5). Jedno zapytanie powinno umożliwić uzyskanie sfederowanego rezultatu, jednej listy wyników z jednoczesnego (sfederowanego) przeszukania pełnotekstowych materiałów w bazach, bibliotece cyfrowej i metadanych dotyczących materiałów bibliotecznych (wyszukiwanie sfederowane [ang. *federated search*], metawyszukiwanie, pozwala na przygotowanie odpowiedzi na pojedyncze zapytanie użytkownika na podstawie zawartości wielu baz danych [Nahotko 2010, s. 165]).



Ilustracja 5. Wyszukiwanie sfederowane – schemat

## KKZBUJ przez CHAMO

Jak wspomniano powyżej, KKZBUJ notuje wszystkie kategorie zbiorów bibliotek UJ, od rękopisów, poprzez druki stare, nowe, muzyczne, zwarte, ciągłe, dokumenty ikonograficzne i kartograficzne, dokumenty audiowizualne i elektroniczne na nośnikach fizycznych, aż do dokumentów oryginalnie elektronicznych (ang. *born digital*) i kopii cyfrowych w Jagiellońskiej Bibliotece Cyfrowej (JBC). W przypadku tych ostatnich w rekordach bibliograficznych w KKZBUJ umieszczone są permanentne linki, które dynamicznie przenoszą użytkownika

do pełnych tekstów w JBC, a ponadto metadane w formacie MARC 21 zawarte w rekordzie bibliograficznym są konwertowane w KKZBUJ do formatu Dublin Core i stanowią podstawę opisu obiektów cyfrowych w JBC. Metadane dotyczące dokumentu w KKZBUJ i obiektu cyfrowego w JBC są linkowane wzajemnie między obiema bazami. Co więcej, w KKZBUJ znajdują się rekordy bibliograficzne dotyczące zarówno subskrybowanych, jak i ogólnodostępnych baz danych, z linkami wiodącymi do baz, ale nie sięgającymi ich pełnej zawartości. W KKZBUJ nie ma opisów artykułów z czasopism i książek z baz, a pełnotekstowe wyszukiwanie możliwe jest dopiero po wejściu do danej bazy (w tym do JBC).

Reasumując, można powiedzieć, że KKZBUJ jest katalogiem hybrydowym, pozwala dotrzeć do zasobów tradycyjnych (analogowych), jak i cyfrowych, w tym będących rezultatem digitalizacji tych pierwszych. Natomiast zawartość subskrybowanych baz, powstających poza działalnością Biblioteki, udostępniania jest za pomocą związanych z nimi wyszukiwarek. Za takim rozwiązaniem przemawiają względy ekonomiczne i organizacyjne. Zresztą, mimo że niektórzy autorzy twierdzą, że sfederowane wyszukiwanie jest świętym Graalem w tzw. *discovery layers*<sup>1</sup> [Yang, Wagner 2010, p. 698], to inni zauważają, że mimo wielu korzyści, wyszukiwanie to ma pewne „wrodzone” ograniczenia, jak: 1) różnorodność zasobów utrudniająca formułowanie zapytań, 2) metodologiczna i techniczna trudność scalania w jeden zbiór wyników z baz pełnotekstowych, OPAC-u i biblioteki cyfrowej, 3) zróżnicowane struktury danych wykorzystywane w tak różnych bazach utrudniające tworzenie rankingu wyników podobnego do tego, jaki generowany jest w wyszukiwarkach webowych [Nahotko 2010, s. 166]. Tym niemniej scalanie danych bibliotecznych z danymi z innych źródeł powinno być wyznacznikiem sposobu działania katalogów bibliotecznych kolejnych generacji. Pełna integracja zasobów, ułatwiająca maksymalnie dostęp do informacji, jest celem niezwykle istotnym, stanowiącym wciąż zadanie do wykonania dla większości bibliotek. Ale, jak twierdzi Marek Nahotko, nie jest celem ostatecznym, bo obecnie dla zwiększenia efektywności pracy uczonych, dydaktyków i ich studentów niezbędne jest dostarczanie usług działających na dostępnych treściach. Oprócz dostarczania treści przyszłością bibliotek są usługi i narzędzia pozwalające na *mashup* treści [Nahotko 2011, s. 208].

**2. KNG powinien mieć przyjazny, nowoczesny, intuicyjny i wizualnie atrakcyjny interfejs,** podobny do Google’a, Amazon i innych popularnych stron internetowych.

---

<sup>1</sup> *Discovery layer* to relatywnie nowy komponent oprogramowania dla bibliotek zawierający interfejs wyszukiwawczy umożliwiający odnalezienie informacji zawartej w katalogu bibliotecznym i poza nim [What is... dok. elektr.].

## KKZBUJ przez CHAMO

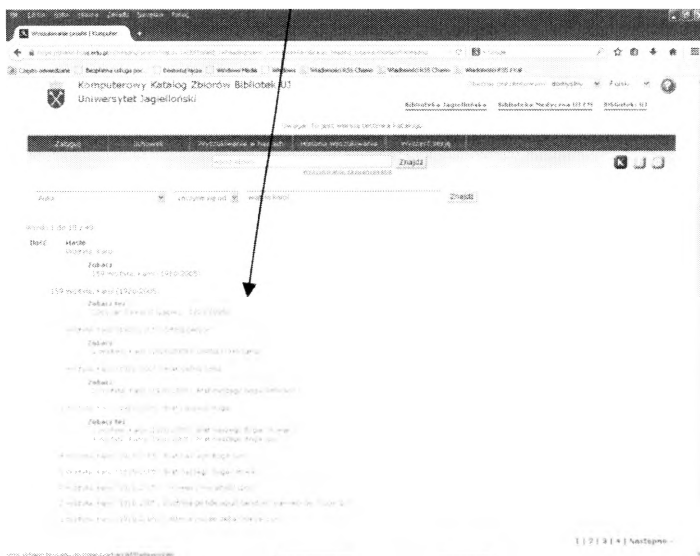
Ocena interfejsu, jego atrakcyjności i intuicyjności jest w dużej mierze subiektywna. Tym niemniej można powiedzieć, że interfejs CHAMO zapewnia łatwiejsze, bardziej intuicyjne wyszukiwanie niż poprzednie interfejsy. Terminem wyszukiwawczym w wyszukiwaniu prostym może być wybrane dowolnie słowo lub jego fragment (pozostała część jest maskowana gwiazdką). Wyniki wyszukiwania mogą być zawężane za pomocą skorelowanych z rezultatem wyszukiwania i dostępnych na ekranie faset. Wyszukiwanie przez słowa może być prowadzone również jako zaawansowane. W tej opcji można sprecyzować składnię wyszukiwania oraz skorzystać z możliwości zawężenia rezultatów wyszukiwania przez zastosowanie kilku filtrów równocześnie. Bez problemu można również usunąć z ekranu *Aktualne wyszukiwanie* dodatkowe terminy i filtry (kliknąć w x przy terminie), rozszerzając w ten sposób rezultaty wyszukiwania, jak również określić liczbę wyników na ekranie, wykorzystać lub wyczyścić historię wyszukiwania. Rezultaty wyszukiwania można też sortować według wybranego kryterium, a także personalizować. Na każdym ekranie dostępny jest schowek (ilustracja 8) przeznaczony do gromadzenia wyników wyszukiwania (egzemplarze), które można zapisać jako zestawienie bibliograficzne (po zalogowaniu się na swoje konto) i to zestawienie wysłać na wskazany adres e-mail bądź wyeksportować jako plik tekstowy w różnych formatach (.txt, .csv, .rec).

W KKZBUJ możliwe jest również wyszukiwanie przez słowa tylko w hasłach (zakładka w menu głównym), według kategorii takich jak: autor, tytuł, hasło przedmiotowe, hasło przedmiotowe w języku francuskim i angielskim, tytuł czasopisma. Rezultaty tego wyszukiwania wyświetlane są jako lista hiperlinków wyszukiwawczych, poprzedzonych cyframi określającymi liczbę dokumentów związanych z danym hasłem, prowadzących do ekranu *Wyniki wyszukiwania*. Na liście hiperlinków występują również odsyłacze typu *Zobacz* i *Zobacz też* (ilustracja 6).

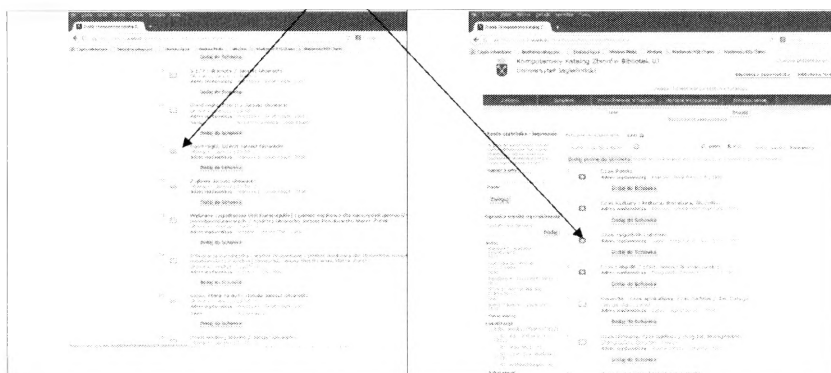
KKZBUJ może być dostępny w więcej niż jednym widoku, do wyboru przez użytkownika (np. całość katalogu lub tylko księgozbiór *Lectorium*), a także w wersji na komórkę. Przy poszczególnych pozycjach w KKZBUJ widnieją ikonki (ilustracja 7) określające graficznie poszczególne rodzaje dokumentów.

Interfejs CHAMO umożliwia czytelnikowi (po zalogowaniu się) sprawdzenie stanu swojego konta, jego ważności, samodzielnie aktualizację danych do kontaktu (e-mail, numer telefonu), zmianę hasła dostępu do konta, zgłoszenie zagubienia karty i otrzymywanie wiadomości od bibliotekarza bezpośrednio na konto. Może on także oczywiście złożyć zamówienie na książkę i prolongować jej zwrot, zarezerwować dla siebie książkę wypożyczoną i anulować niepotrzebne rezerwacje ze swego konta.

Inne przyjazne funkcje interfejsu CHAMO opisano przy omawianiu poszczególnych cech KNG.



Ilustracja 6. Rezultat wyszukiwania w hasłach



Ilustracja 7. Ikony przy poszczególnych rodzajach dokumentów

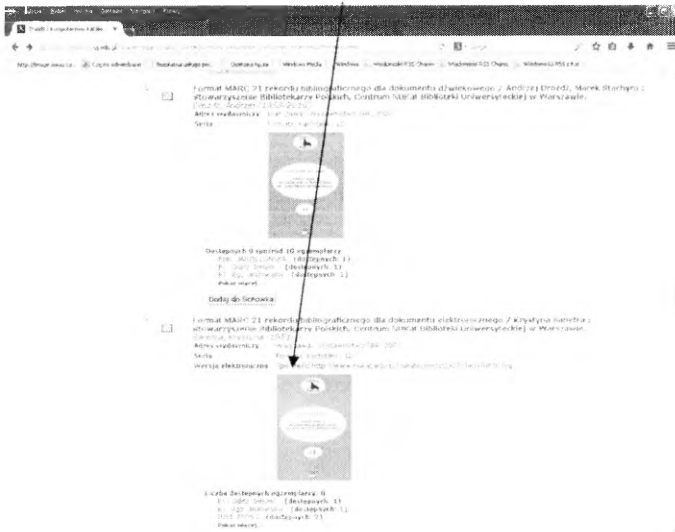


Ilustracja 8. Schowek i możliwości jego wykorzystania

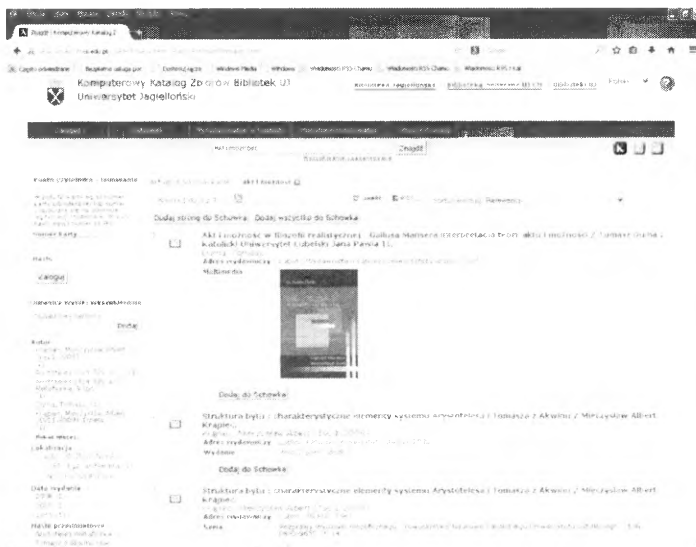
**3. KNG powinien zawierać wzbogaconą zawartość** z różnych źródeł w celu wzmocnienia wizualnej atrakcyjności i zwiększenia ilości informacji prezentowanych użytkownikowi. Niektóre kategorie tej zawartości mogą być połączone z podstawową informacją bibliograficzną katalogu. Te kategorie to np. skany okładek książek, pojemników na filmy i dokumenty elektroniczne lub inne wizualne przedstawienie graficzne, miniaturki przywołujące dokument. Mogą to być także skany spisów treści, streszczenia przejęte z dokumentu lub innych źródeł, np. promocyjnych opisów wydawcy lub opisy, oceny / recenzje książek, zwłaszcza dostarczone przez ekspertów.

### **KKZBUJ przez CHAMO**

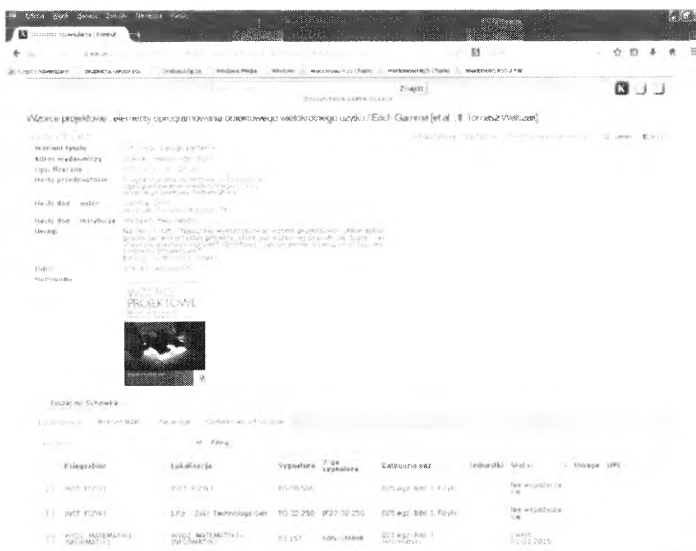
W KKZBUJ widoczne są skany okładek książek i innych typów dokumentów, spisy treści w nich zawarte (ilustracje 9, 10, 11), jak również, gdy metadane dotyczą zarówno tradycyjnej formy dokumentu, jak i jego kopii cyfrowej, miniaturki obiektów cyfrowych oraz miniaturki dokumentów natywnie elektronicznych udostępnianych w JBC. Adresy elektroniczne skanów okładek, spisów treści i miniaturek utworzonych w BJ wprowadzane są do rekordów bibliograficznych na poziomie NUKAT, do wykorzystania przez inne współpracujące biblioteki. Podobnie do KKZBUJ przejmowane są rekordy bibliograficzne z linkami do skanów / spisów treści rezydujących na zewnętrznych serwerach, stanowiąc dodatkowo zysk ze współpracy z NUKAT. Na podobnej zasadzie możliwe jest również dodawanie linków do opisów, recenzji i innych uzupełniających materiałów.



Ilustracja 9. KKZBUJ – skany okładek oraz spis treści na ekranie rezultatów wyszukiwawczych



Ilustracja 10. KKZBUJ – skany okładek na ekranie rezultatów wyszukiwawczych



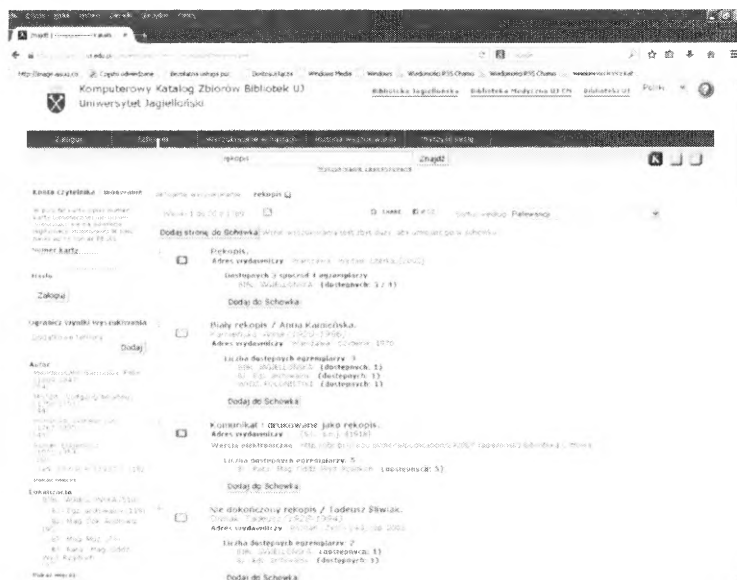
Ilustracja 11. KKZBUJ – skany okładek na ekranie wykazu egzemplarzy

**4. KNG powinien umożliwiać nawigację fasetową**, czyli pozwalać na wyświetlenie wyników wyszukiwania jako zestawu kategorii. Rekordy bibliograficzne w formacie MARC 21 dostarczają bogatego materiału do kreowania faset. Można łatwo wyekstrahować z nich kategorie faset: nazwy osobowe, nazwy geograficzne,

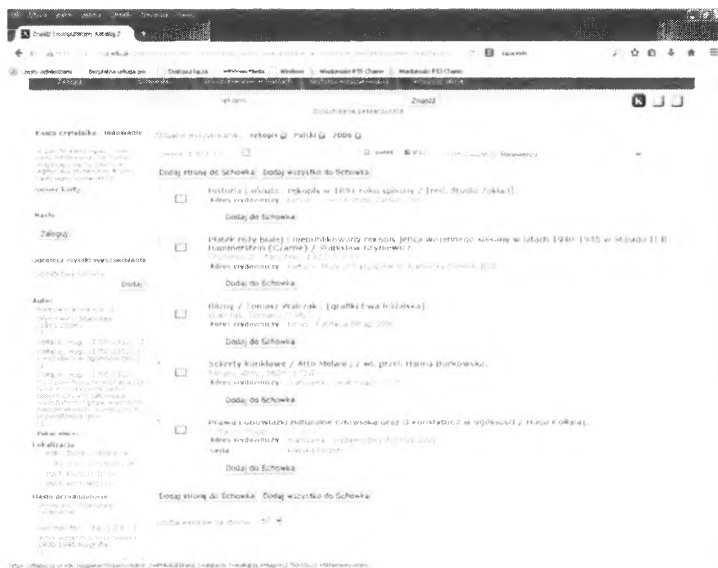
tematy / hasła przedmiotowe, zakresy dat, typy dokumentów, języki, lokalizacje itp. Nawigacja fasetowa umożliwia użytkownikom zawężenie rezultatów wyszukiwania. Fasety pojawiają się jako linki korespondujące ze słowami albo frazami w obrębie wyników wyszukiwania i zwykle towarzyszy im ujęta w nawiasy kwadratowe liczba dokumentów, do których można dotrzeć po wybraniu danej fasety. Jeżeli wyszukiwana informacja jest złożona, fasety mogą być grupowane w kategorie. Wyszukiwanie fasetowe pozwala użytkownikowi na interakcje ze źródłem przez odkrywanie informacji w nim zawartych, a nie zgadywanie ich z góry. Takie podejście stoi w fundamentalnej opozycji do tradycyjnego wyszukiwania, gdzie interfejs dostarcza stronę z możliwością zaawansowanego wyszukiwania, na której użytkownik może / musi konstruować złożony zestaw warunków / kwalifikatorów na początku procesu wyszukiwania.

## KKZBUJ przez CHAMO

Przy wyszukiwaniu w CHAMO użytkownicy mogą zawęzić rezultat wyszukiwawczy za pomocą faset (ilustracje 12, 13) znajdujących się z lewej strony ekranu w ramce pod nazwą *Ogranicz wyniki wyszukiwania*. Fasety są hiperlinkami, które odnoszą się do rekordów aktualnie wyświetlonych na ekranie. Wyświetlone fasety zmieniają się wraz ze zmianą parametrów wyszukiwania / kryterium wyszukiwawczego. Kliknięcie na fasetę powoduje dodanie tego terminu do wyrażenia wyszukiwawczego w menu nawigacyjnym i generuje jako rezultat krótszą,



Ilustracja 12. Rezultat wyszukiwania (1389 pozycji) – po wprowadzeniu terminu *rękopis*



Ilustracja 13. Zawężenie rezultatów poprzedniego wyszukiwania *rękopis* do 5 pozycji po użyciu fasyty *język (polski)* oraz fasyty *data wydania (2006)*

bardziej precyzyjną listę pozycji. Fasyty umożliwiają zawężenie lub przefiltrowanie rezultatów wyszukiwania według różnych kategorii informacji. Liczby ujęte w nawiasy okrągłe przy poszczególnych fasetach wskazują liczbę dokumentów związanych z daną fasetą w danym wyszukiwaniu. Wyświetlone fasyty są pogrupowane i opatrzone nagłówkami takimi jak: *Autor*, *Lokalizacja*, *Daty wydania*, *Hasło przedmiotowe*, *Typ rekordu*, *Forma literacka dokumentu* itp. Aby wyświetlić wszystkie fasyty pod pokreślonym nagłówkiem należy kliknąć *Pokaż więcej*.

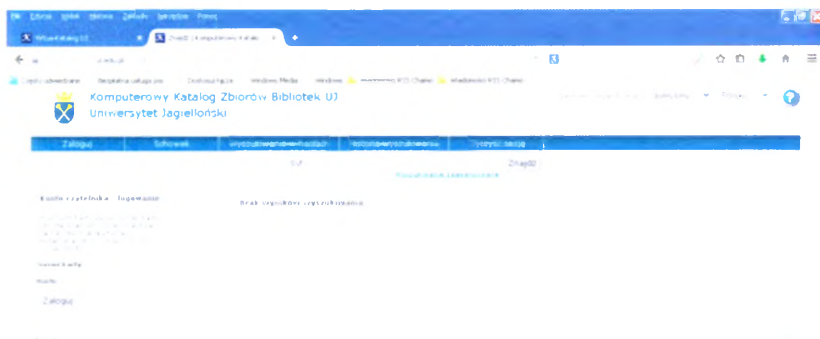
**5. W KNG okienko do wyszukiwania prostego z linkiem do wyszukiwania zaawansowanego powinno widnieć na każdej stronie.** Ponieważ wyszukiwanie przez słowa kluczowe stało się powszechną techniką w internecie, KNG winien proponować rozpoczynanie od wyszukiwania prostego. Może ono być niewystarczające dla wielu użytkowników, więc na każdej stronie winien być link do wyszukiwania zaawansowanego, które oferuje użytkownikowi bardziej precyzyjne i złożone opcje.

### **KKZBUJ przez CHAMO**

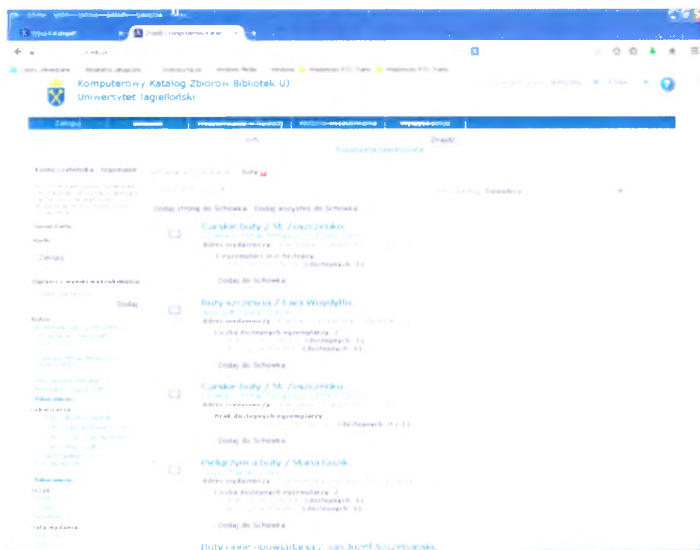
Standardowo w KKZBUJ wyszukiwanie rozpoczyna się od wyszukiwania przez słowa. Na każdym ekranie widnieje okienko do wyszukania prostego / szybkiego, do którego można wpisać dowolne słowo, aby zainicjować wyszukiwanie. Wy-



szukiwanie to obejmuje wszystkie indeksowane pola rekordu bibliograficznego, a trzeba mieć świadomość, że nie wszystkie pola są indeksowane. Jeżeli pierwsze wpisane słowo nie przyniesie żadnych rezultatów wyszukiwania (ilustracja 14), to należy wpisać inne (ilustracja 15).



Ilustracja 14. Brak rezultatów wyszukiwania



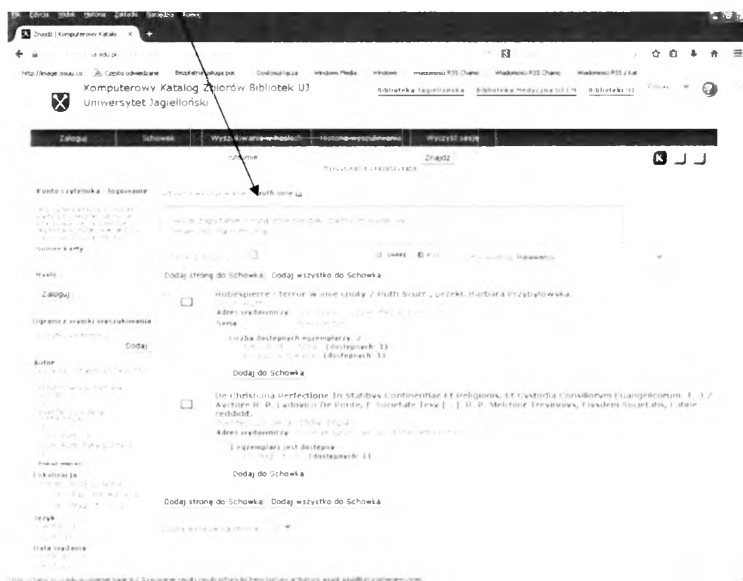
Ilustracja 15. Zmiana terminu wyszukiwawczego na liczbę mnogą (*buty*) spowodowała, że uzyskano rezultat wyszukiwawczy (34 dokumenty).

Poprawić i / lub zawęzić wyszukiwanie można za pomocą faset, a także wpisując dodatkowe słowa w okienku pod nazwą *Dodatkowe terminy*. Dodatkowe terminy pojawiają się, podobnie jak w przypadku faset, w menu nawigacyjnym *Aktualne wyszukiwanie* (ilustracja 16).

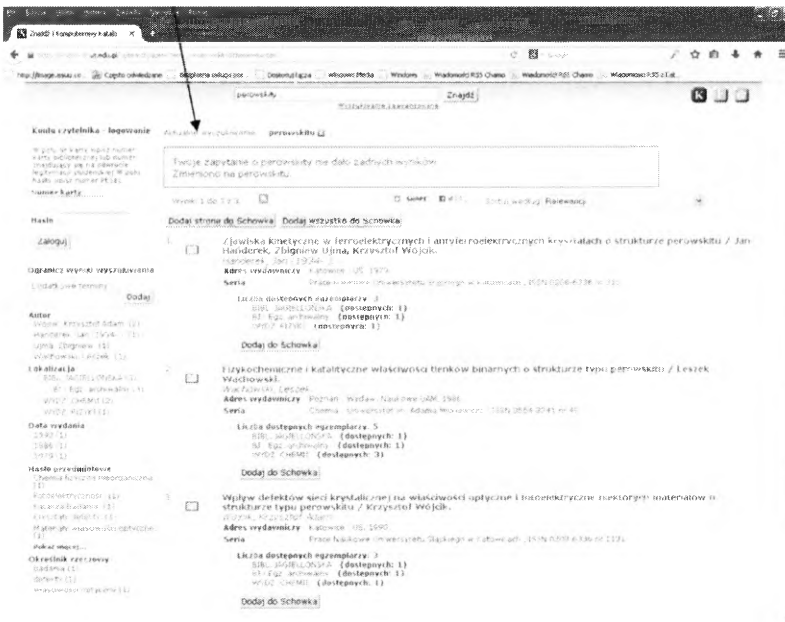


Ilustracja 16. Po dołożeniu dodatkowego terminu *magiczne* rezultat wyszukiwawczy został ograniczony do jednego rekordu / dokumentu.

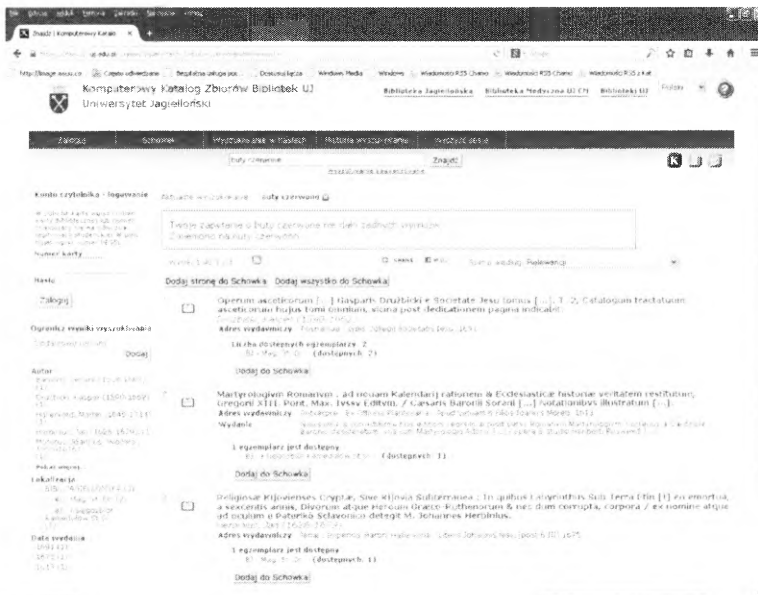
Czasem, gdy dodatkowe słowo nie daje żadnych rezultatów, system CHAMO zmienia termin wyszukiwawczy na podobny, co może być bardzo korzystne (ilustracje 17, 18) lub zupełnie nieużyteczne (ilustracje 19, 20).



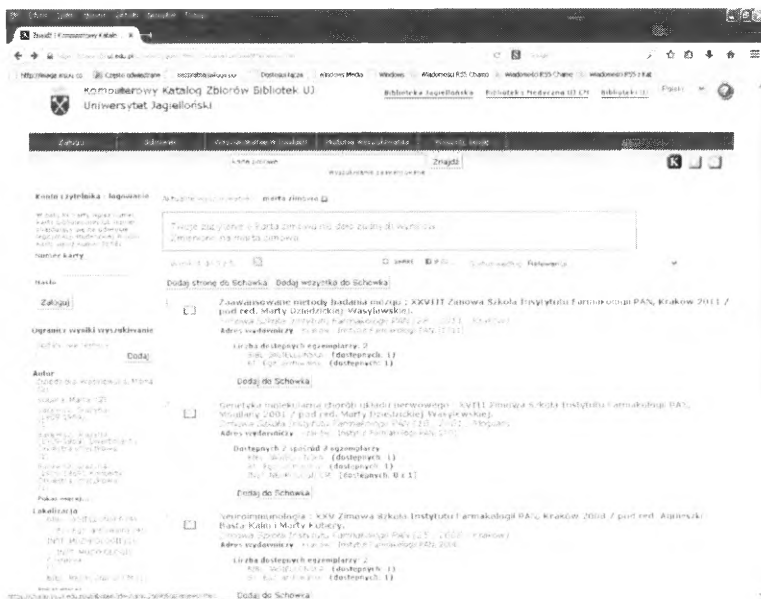
Ilustracja 17. Zmiana przez interfejs CHAMO terminu wyszukiwawczego *ruta imię* na *ruth imię* (angielska wersja imienia) dała pozytywny wynik wyszukiwawczy (2 pozycje).



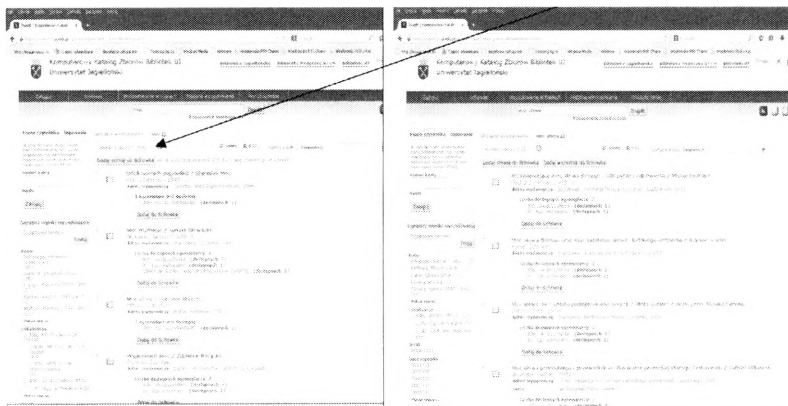
Ilustracja 18. Zmiana przez interfejs CHAMO terminu wyszukiwawczego *perowskity* na *perowskitu* dała pozytywny wynik wyszukiwawczy (3 pozycje).



Ilustracja 19. Zmiana przez interfejs CHAMO terminu wyszukiwawczego z *buty czerwone* na *nuty czerwone* przyniosła rezultat wyszukiwawczy niezwiązany z początkowo szukanym terminem.

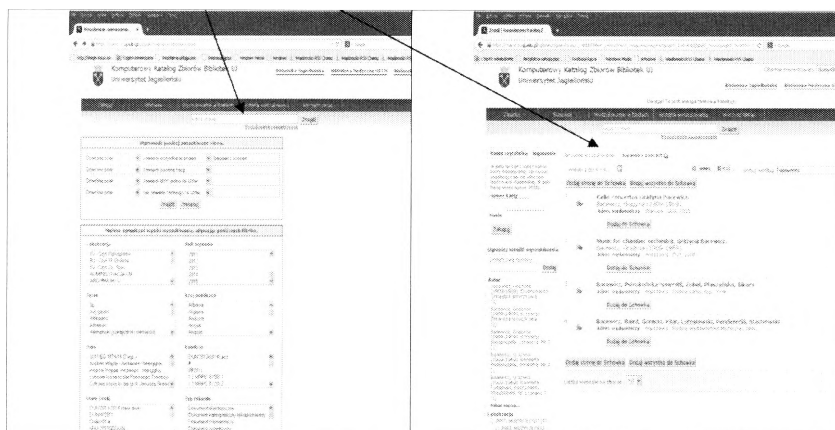


Ilustracja 20. Zmiana przez interfejs CHAMO terminu wyszukiwawczego z *karta zimowa* na *marta zimowa* przyniosła rezultat wyszukiwawczy niezwiązany z szukanym terminem.

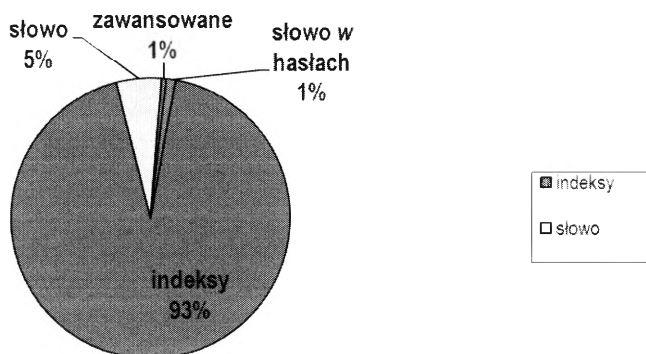


Ilustracja 21. Rezultaty wyszukiwania prostego – doprecyzowanie terminu wyszukiwawczego z *moc* na *moc słowa* zmniejszyło rezultat wyszukiwawczy do 12 pozycji (poprzednio ponad 1500).

Okienko do wyszukiwania prostego (ilustracja 21) przez słowo z linkiem do wyszukiwania zaawansowanego (ilustracja 22), dostępne na każdej stronie OPAC-u, jest czynioną przez biblioteki próbą naśladowania Google'a i innych popularnych wyszukiwarek. Celem jest zagwarantowanie, że użytkownik ma okienko wyszukiwawcze zawsze pod ręką, gdziekolwiek znajdzie się, korzystając z interfejsu [Yang,



Ilustracja 22. Link, opcje i rezultaty wyszukiwania zaawansowanego



Ilustracja 23. Średnia miesięczna statystyka wyszukiwań w KKZBU dostępnym przez Vectors iPortal (luty-lipiec 2014 r.)

Wagner 2010, p. 703]. Zakłada się, że wyszukiwanie proste przez słowo, uważane za jedną z najważniejszych cech KNG, jest cechą pożądaną przez użytkowników. Będzie interesujące sprawdzić, jak zostanie ono przyjęte przez użytkowników KKZBU obsługiwane przez CHAMO. Tegoroczna analiza statystyk wyszukiwania w KKZBU dostępnego przez Vectors iPortal (ilustracja 23) dowodzi bowiem, że absolutny priorytet jako kryterium wyszukiwania stanowił tytuł dzieła. Tytuł wykorzystywany był miesięcznie średnio 65 razy częściej niż również dostępne na pierwszej stronie interfejsu wyszukiwanie przez słowo za pomocą kategorii *Wszystko*. Wyszukiwanie proste przez słowo stanowiło średnio miesięcznie tylko 5%, wyszukiwanie zaawansowane i przez słowa w hasłach po 1%, dominowało natomiast wyszukiwanie przez indeksy, stanowiące 93% wszystkich wyszukiwań.

## 6. Relewantność

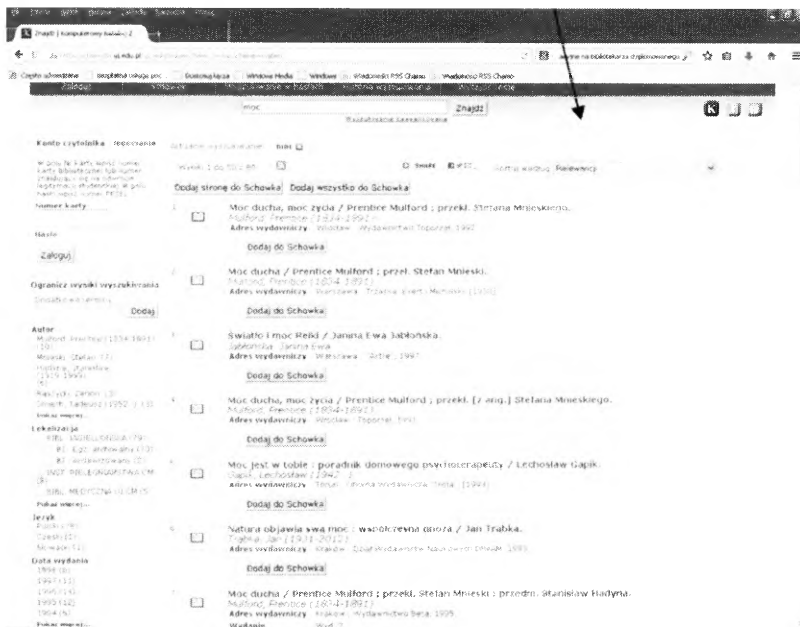
Ranking wyszukanych dokumentów według ich relewantności w KNG jest procesem sortowania wyników wyszukiwania w ten sposób, aby na początku listy pojawiły się te dokumenty, które wydają się najbardziej relewantne, najbardziej ważne / istotne, pozostające w logicznym związku z wprowadzonym zapytaniem. Tradycyjny ranking wyników wyszukiwania oparty jest na częstotliwości i pozycji wprowadzonych terminów wyszukiwawczych w rekordach bibliograficznych. Wielokrotne pojawienie się terminów wyszukiwawczych zwiększa relewantność (trafność, zgodność odpowiedzi z postawionym pytaniem). Istotne jest również, w którym miejscu rekordu bibliograficznego pojawiają się terminy wyszukiwawcze. Terminy wyszukiwawcze pojawiające się w tytule dokumentu lub w strefie tytułu mają większą wagę niż te w obrębie tekstu. Relewantność techniczna niekoniecznie zawsze odzwierciedla ważność dokumentu wśród wyników wyszukiwania. Innymi ważnymi czynnikami mogą być np. liczba i jakość cytowań dokumentu, a w kontekście bibliotek – wyżej w rankingu mogą być dokumenty o większej liczbie wypożyczeń, wskazującej na popularność i użyteczność, jak też dokumenty w większej liczbie egzemplarzy, ponieważ wydają się ważne na tyle, aby przechowywać je w więcej niż jednym egzemplarzu lub w więcej niż jednej bibliotece.

## KKZBUJ przez CHAMO

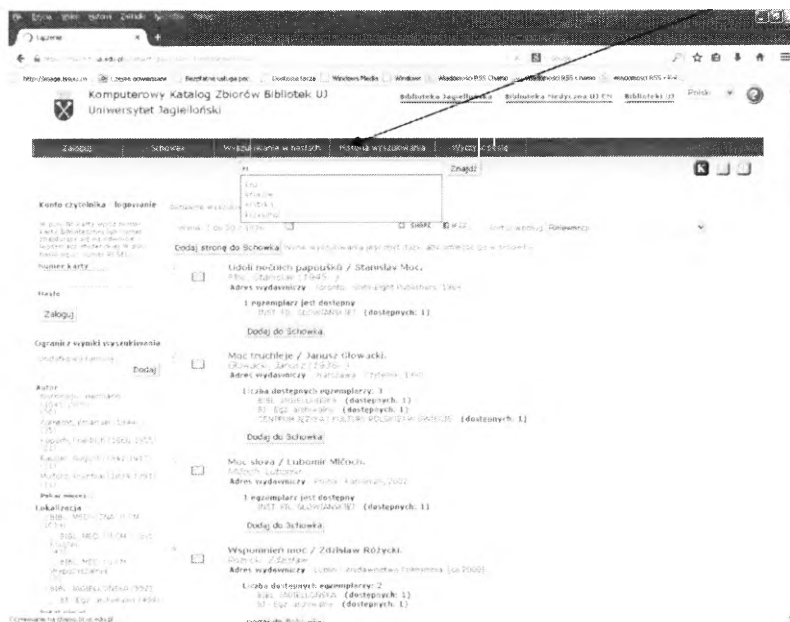
Rezultaty wyszukiwania prostego / szybkiego sortowane są zawsze według relewancji (ilustracja 24), nawet wówczas gdy sortowanie zostanie ustawione według innej kategorii. Relewancja w tym przypadku polega na tym, że jako pierwsze wyświetlają się rekordy bibliograficzne, w których szukane słowo występuje w polach: tytuł, autor, temat hasła przedmiotowego. To domyślne sortowanie można potem zmienić na inne i uzyskać posortowanie uzyskanych wyników wyszukiwania według innej kategorii. Przy przedstawianiu rezultatów wyszukiwania nie są brane pod uwagę takie czynniki wpływające na relewancję, jak transakcje wypożyczeń i rezerwacji oraz liczba egzemplarzy danego dzieła.

## 7. Korekta błędnych zapytań („Czy miałeś na myśli...?”)

KNG powinien sugerować korektę wyrażenia wyszukiwawczego, gdy pojawi się w nim błąd w pisowni oraz proponować listę poprawnych terminów do wykorzystania przez użytkownika. Taka cecha interfejsu jest powszechnie oczekiwana. Podpowiadanie użytkownikowi jest dobrą opcją, bo ogranicza liczbę nieudanych zapytań kierowanych do bazy i częstotliwość pojawiania się komunikatu, że nie znaleziono żadnych dokumentów.



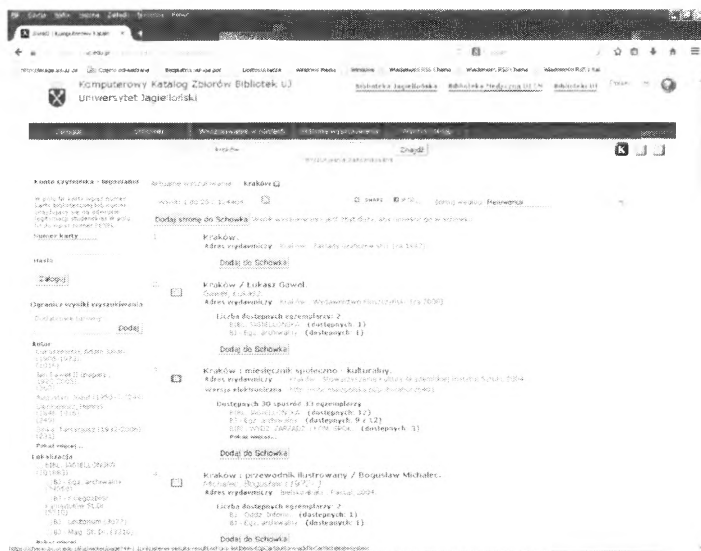
Ilustracja 24. Rezultaty wyszukiwania posortowane według relewancji – szukane słowo widoczne w tytułach



Ilustracja 25. Podpowiedzi interfejsu CHAMO po wpisaniu początku terminu wyszukiwawczego kr...

## KKZBUJ przez CHAMO

Interfejs CHAMO nie koryguje błędnie zapisanych wyrażen wyszukiwawczych przez interaktywne pytanie zawierające odpowiedź poprawnie napisanego słowa („Czy miałeś na myśli...?”). Jednakże podczas wpisywania słowa w okienko wyszukiwawcze pojawia się okienko z odpowiedziami słów (ilustracja 25) zawierających wpisane litery / sylaby, w miarę wpisywania kolejnych liter odpowiedzi są coraz bardziej adekwatne i można wybrać to lub podobne słowo z listy (ilustracja 26).



Ilustracja 26. Z odpowiedzi interfejsu CHAMO wybrano słowo *Kraków*

Ponadto przy wprowadzaniu wyrażenia wyszukiwawczego nie ma potrzeby wpisywania polskich znaków diakrytycznych, są one automatycznie dokładane i widoczne w rezultatach wyszukiwania.

## 8. Rekomendowanie pokrewnych / podobnych materiałów

Zakłada się, że KNG powinien rekomendować czytelnikom książki i inne materiały na podstawie logów transakcji, sugerując że czytelnik, który pożyczył daną książkę, pożyczyłby również następujące... lub zestawiać wykaz materiałów pod linkiem *Rekomendowana lektura*.

## KKZBUJ przez CHAMO

Interfejs CHAMO nie posiada funkcji rekomendowania użytkownikowi materiałów na podstawie logów transakcji wypożyczeń.



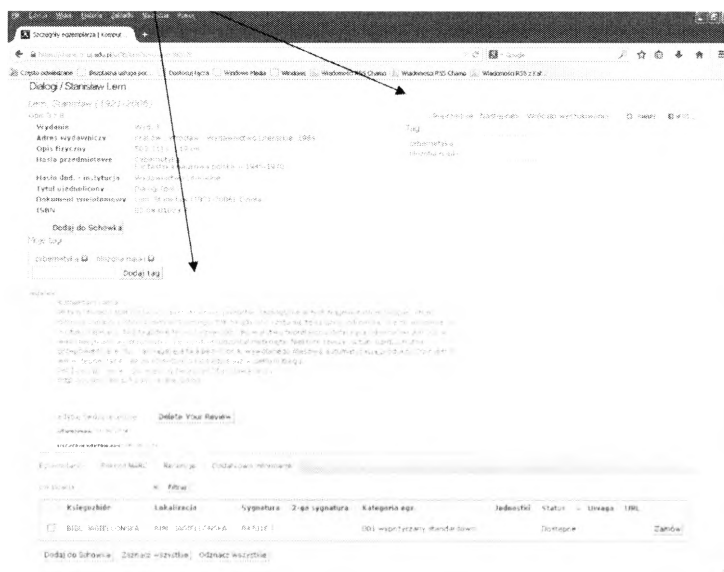
## 9. Wkład użytkownika – w duchu Web 2.0

Katalog nie powinien być jednostronną prezentacją informacji, ale umożliwiać współudział i zaangażowanie użytkowników. W KNG użytkownicy powinni być aktywnymi współautorami treści zawartych w OPACu – mogą oceniać, pisać recenzje, przeglądy, dopisywać tagi i komentarze dotyczące dokumentów. Wkład użytkowników może być ważnym wskaźnikiem w rankingu relewantności.

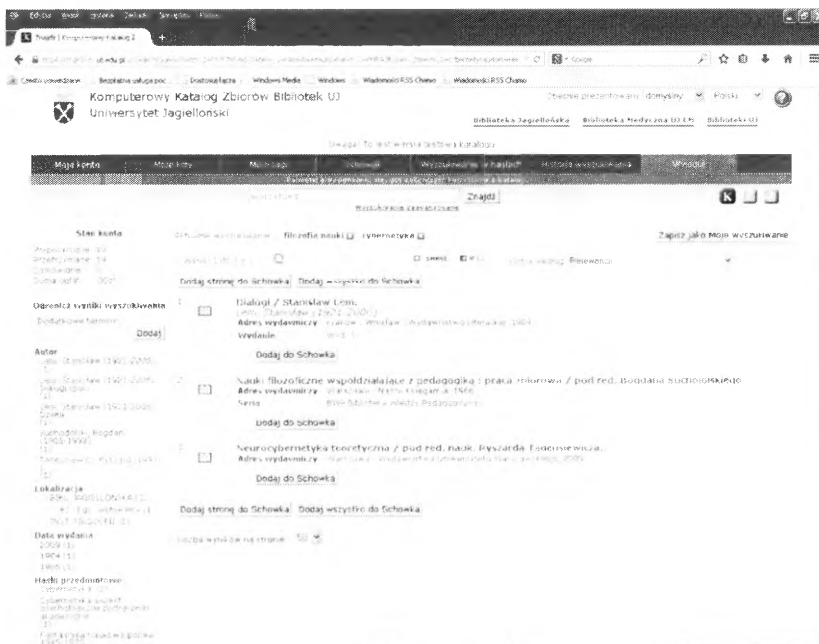
### KKZBUJ przez CHAMO

Interfejs CHAMO pozwala na włączenie opcji tagowania, czyli kategoryzacji treści przez użytkowników – dodawania tagów (ilustracja 27) przy przeglądanych pozycjach, najczęściej za pomocą słów kluczowych (słów, fraz) [Roszkowski 2008]. Pozwala również na sporządzenie, załadowanie i edytowanie przez administratora listy tagów zakazanych. Dodanie przez użytkownika wyrażenia, które znajduje się na liście tagów zakazanych, nie jest możliwe, o czym użytkownik zostanie poinformowany automatycznym komunikatem systemu. Tagi wyświetlają się na ekranie wykazu egzemplarzy, są widoczne dla wszystkich użytkowników i stają się dodatkowym kryterium wyszukiwania.

Użytkownik może dopisywać również komentarze, opinie, oceny danej pozycji, które są moderowane / zatwierdzane przez administratora i widoczne dla wszystkich – po kliknięciu przycisku *Recenzje* (ilustracja 27) – przy czym ostatnio dodane zawsze znajdują się na początku listy.



Ilustracja 27. Tagi i recenzja na ekranie wykazu egzemplarzy



Ilustracja 28. Tagi (ilustracja 27) jako kryterium wyszukiwania

**10. Wykorzystanie kanałów RSS** do informowania użytkowników o częstym uaktualnianiu zawartości katalogu. Może on być skonfigurowany do przesyłania informacji, takich jak lista nowości (według typów materiałów lub dziedzin wiedzy), lista najczęściej wypożyczanych / rezerwowanych materiałów i innych informacji dla użytkowników / subskrybentów.

### **KKZBUJ przez CHAMO**

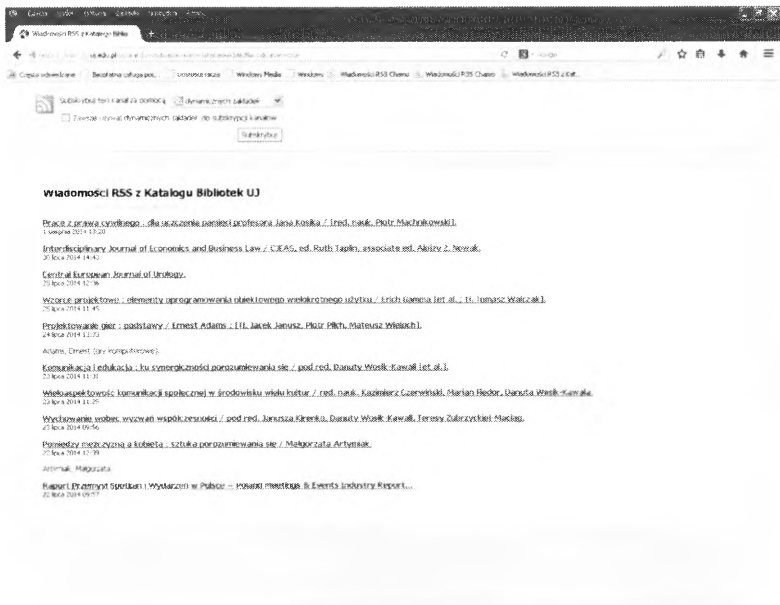
Interfejs umożliwia włączenie i wykorzystywanie takiej funkcji. Jest ona dostępna po kliknięciu na ikonkę (ilustracja 29) widoczną na ekranie z wynikami wyszukiwania.

### **11. Zintegrowanie z portalami społecznościowymi**

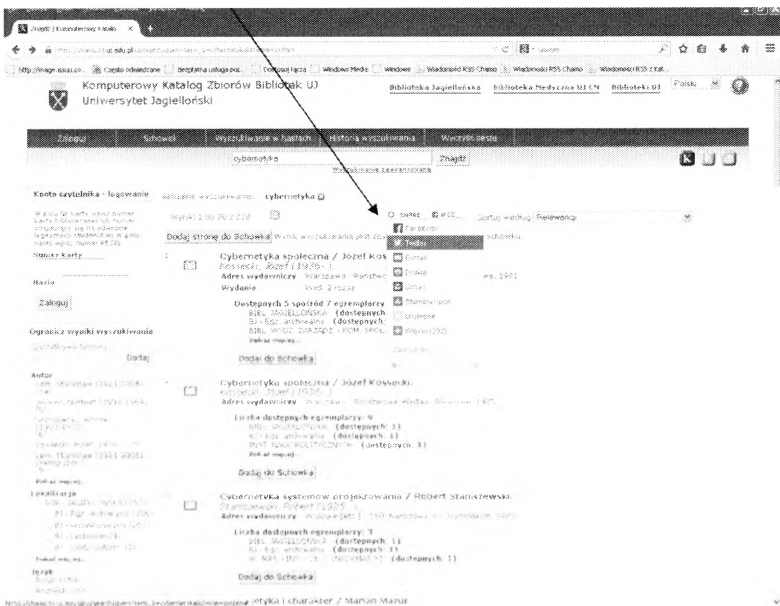
Użytkownicy mogą dzielić się linkami do egzemplarzy bibliotecznych z przyjaciółmi na Facebooku, Twitterze itp.

### **KKZBUJ przez CHAMO**

Interfejs CHAMO umożliwia bezpośrednie łączenie się użytkownika z portalami społecznościowymi, wysłanie e-maili itp. (ilustracja 30).



Ilustracja 29. Ekran z wiadomościami dostępnymi po kliknięciu na ikonkę



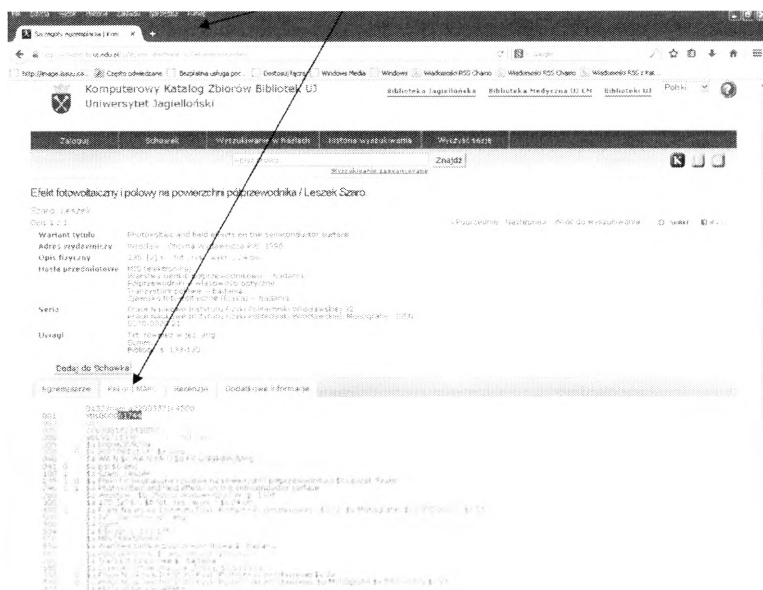
Ilustracja 30. Menu dostępnych portali społecznościowych i poczty elektronicznej

## 12. Stałe /stabilne linki

W KNG rekord bibliograficzny powinien posiadać stały URL, który można skopiować i który stanowi permanentny link do tego rekordu.

### KKZBUJ przez CHAMO

Każdy rekord bibliograficzny posiada stały adres elektroniczny, którego istotną częścią, poza adresem bazy katalogowej, jest numer systemowy rekordu w tej bazie. Skopiowanie tego adresu i wprowadzenie go do przeglądarki internetowej umożliwi bezpośrednie wyszukanie tego rekordu (adres elektroniczny poniższego rekordu bibliograficznego: <https://chamo.bj.uj.edu.pl/uj/lib/item?id=chamo:61744&theme=system>, ilustracja 31).



Ilustracja 31. Stały URL rekordu bibliograficznego, którego częścią jest numer systemowy rekordu (61744).

## Podsumowanie

„Katalogi lub narzędzia wyszukiwawcze nowej generacji (KNG) wykorzystują istniejące dane bibliograficzne z tradycyjnych, zintegrowanych systemów bibliotecznych, prezentując je w nowej, przyjaznej użytkownikowi formie” [Waleszko 2011]. Aby to było możliwe, dane w bazie katalogowej muszą być standardowe i spójne. „Anna Paluszkiwicz powtarzała nam zawsze – w procesie komputery-

zacji katalogu najważniejsze i najbardziej kosztowne są dane. Trzeba je tak budować, aby były odporne na zmiany technologiczne, zmiany zasad katalogowania, zmiany systemów i formatów” [Burchard 2014]. Tak zbudowane dane zawiera baza KKZBUJ. Z porównania funkcji i cech KKZBUJ dostępnego przez interfejs CHAMO wynika, że z 12 analizowanych cech KNG, CHAMO całkowicie nie realizuje jednej funkcji, a dwie inne realizuje częściowo. CHAMO nie sięga do logów transakcji wypożyczeń użytkowników i nie rekomenduje na tej podstawie pokrewnych materiałów do czytania. Wdrożone CHAMO nie zapewnia sfederowanego wyszukiwania w całym środowisku informacyjnym bibliotek UJ, chociaż obejmuje wszystkie rodzaje tradycyjnych zasobów bibliotecznych, a także obiekty cyfrowe powstałe w wyniku digitalizacji. CHAMO posiada również mechanizmy częściowego korygowania błędnych zapytań kierowanych do bazy.

Można zatem stwierdzić, że KKZBUJ dostępny przez interfejs CHAMO w dużej mierze odpowiada charakterystyce Katalogu Nowej Generacji. Dalsze zmiany wymuszać będzie fakt, że KKZBUJ powinno zapewniać dostęp do wszystkich zasobów biblioteki nowej generacji. Biblioteka Jagiellońska stała się już biblioteką hybrydową, w której współistnieją zasoby tradycyjne / analogowe i różnej proveniencji zasoby cyfrowe oraz ich metadane. Ekspansja mediów cyfrowych i internetu spowodowała, że choć cyfrowe środowisko informacyjne Biblioteki Jagiellońskiej stanowi tylko 7,8% ogólnej liczby 5 279 566 woluminów i jednostek zbiorów tradycyjnych, gromadzonych od XIV w. (stan na 31.12.2013 r.), to już roczny przyrost zbiorów cyfrowych jest jednak znaczny. W 2013 r. roczny przyrost zbiorów tradycyjnych wyniósł 149 090 wol. i jedn., a przyrost zbiorów cyfrowych był mniejszy o ok. połowę i wyniósł 72 627 jedn. (49% przyrostu zbiorów tradycyjnych). Cyfrowe środowisko informacyjne według stanu na 1.01.2014 r. wynosiło 410 848 jedn. udostępnionych zasobów cyfrowych, z czego 4,2% to zinventoryzowane dokumenty elektroniczne (7 345 jedn.), 42,2% to licencjonowane zbiory elektroniczne, do których BJ opłacała dostęp (173 994 jedn.), a 53,4% to materiały zdigitalizowane i udostępnione w JBC (219 509 jedn.). Biblioteka obsługuje symultanicznie wiele generacji użytkowników, z różnymi oczekiwaniami, doświadczeniami i różną wiedzą na temat strategii wyszukiwawczych. A trzeba dodać, że współczesny użytkownik biblioteki i internetu przyzwyczajony jest do otrzymywania żądanej informacji natychmiast [Nahotko 2011, s. 195]. Wpływ takiego środowiska informacyjnego wymaga wdrażania kolejnych nowych technologii, nowych narzędzi wyszukiwawczych. Bibliotekarzom zapewne wystarczy determinacji, ale niezbędne również są środki finansowe.

## BIBLIOGRAFIA

- Breeding, Marshall (2007). *Introduction to „Next Generation Library Catalogs”*. Library Technology Report Vol. 43, No. 4, p. 5-14.
- Burchard, Maria (2014). *Wspomnienie o Annie Paluszkiewicz*. Tytuł Ujednolicony nr 13, s. 29-29.
- Coyle, Karen (2010). *Library data in a modern context*. Library Technology Reports Vol. 46, Issue 1, p. 5-13.
- Morgan, Eric Lease (2006). A “next generation” library catalog – executive summary (Part #1-5, post) [dok. elektr.]. LITABlog. <http://litablog.org/2006/07/07/> [odczyt: 25.06.2014].
- Merčun, Tanja; Žumer, Maja (2008). *New generation of catalogues for the new generation of users: a comparison of six library catalogues*. Program: electronic library and information systems Vol. 42, No. 3, p. 243-261.
- Nagy, Andrew (2011). *Defining the Next Generation Catalog*. Library Technology Reports Vol. 47, No. 7, p. 11-16.
- Nahotko, Marek (2011). *Integracja wyszukiwania w zasobach informacyjnych*. Przegląd Biblioteczny R. 79, z. 2, s. 192-210.
- Nahotko, Marek (2010). *Komunikacja naukowa w środowisku cyfrowym: globalna biblioteka cyfrowa w informatycznej infrastrukturze nauki*. Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, 283 s.
- Roszkowski, Marcin (2008). *Czym jest tagowanie?* [dok. elektr.]. Warsztaty Bibliotekarskie nr 3-4. <http://www.pedagogiczna.edu.pl/warsztat/2008/3-4/080306.htm> [odczyt 22.09.2014].
- Waleszko, Małgorzata (2011). *Wpływ katalogów 3. generacji na katalogowanie w bibliotekach akademickich* [dok. elektr.]. BABIN 2.0. <http://babin.bn.org.pl/?p=625> [odczyt: 25.06.2014].
- What is a discovery layer?* [dok. elektr.]. <https://foss4lib.org/decision-support/discovery-layer-ssm/definition> [odczyt: 20.10.2014].
- Yang, Sharon Q.; Hofmann, Melissa A. (2010). *The Next Generation Library Catalog: a comparative study of the OPACs of Koha, Evedgreen, and Voyager*. Information Technology and Libraries Vol. 29, No. 3, p. 141-150.
- Yang, Sharon Q.; Hofmann, Melissa A. (2011). *Next generation or current generation? A study of the OPACs of 260 academic libraries in the USA and Canada*. Library Hi Tech Vol. 29, Issue 2, p. 266-300.
- Yang, Sharon Q.; Wagner, Kurt (2010). *Evaluating and comparing discovery tools: how close are we towards next generation catalog?* Library Hi Tech Vol. 28, No. 4, p. 690-709.

---

Część IV

# **Użytkownicy, źródła, narzędzia, systemy**

---

---

# Użytkownicy bibliotek cyfrowych w piśmiennictwie anglojęzycznym: system *versus* użytkownik

Małgorzata Janiak

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

## Wprowadzenie

Artykuł zawiera analizę przedmiotową anglojęzycznych publikacji naukowych o zasięgu międzynarodowym z lat 1996-2014, dotyczących użytkowników bibliotek cyfrowych. Zastosowano metodę analizy i krytyki piśmiennictwa, jako powszechnie wykorzystywaną w przygotowywaniu analityczno-syntetycznych relacji z dotychczasowego stanu wiedzy [Cisek 2010, s. 274]. Aby dotrzeć do relewantnej literatury fachowej i naukowej, wykorzystano dziedzinową bazę danych LISTA – Library, Information Science and Technology Abstracts, inne bazy ogólnie dostępne w internecie, np. DBLP (<http://www.dblp.org/>) i ProQuest (<http://www.proquest.com>), oraz wyszukiwarkę Google Scholar. Zaprezentowano wyniki analizy najważniejszych artykułów, w których poruszana była tematyka użytkowników informacji przeszukujących, oceniających i współtworzących biblioteki cyfrowe. Nie jest to pełna bibliografia wszystkich prac naukowych, w których pojawił się termin „użytkownik biblioteki cyfrowej”, a jedynie chronologiczny wykaz najciekawszych tematów badawczych, które dla ułatwienia zostały przedstawione w grupach wyodrębnionych po przeanalizowaniu całego materiału badawczego. W niniejszym artykule nie uwzględniono bowiem publikacji, w których użytkownik był wymieniany jedynie jako ostateczny odbiorca, a sama publikacja poświęcona była systemom lub zawartości bibliotek cyfrowych. Temat oceny jakości także został pominięty (poza nielicznymi przykładami), gdyż omawiało go już kilku polskich autorów, m.in. Małgorzata Janiak i Monika Krakowska [Janiak, Krakowska 2012]. Z tej racji nie przytaczano więc powtórnie np. analiz publikacji Hong Iris Xie z lat 2002-2008 [Xie 2006; 2008].



Biblioteka cyfrowa (dalej w skrócie: BC) rozumiana jest przez autorów analizowanych prac bardzo różnorodnie. Jest to przede wszystkim biblioteka, której zbiory – gromadzone lokalnie lub zdalnie – dostępne są w formatach cyfrowych za pośrednictwem komputerów podłączonych do sieci, właściwie z każdego miejsca świata, jak głosi definicja zawarta np. w Wikipedii [Biblioteka cyfrowa, dok. elektr.]. W konsekwencji w badanych opracowaniach terminu „biblioteki cyfrowe” używano m.in. na określenie bardzo szeroko rozumianych systemów informacyjnych skupiających wokół siebie społeczność, zgodnie z definicją sformułowaną w manifeście z 2011 r. „The Digital Library Reference Model – DL.org: Coordination Action on Digital Library Interoperability, Best Practices and Modelling Foundations”, przygotowanym w ramach Network of Excellence on Digital Libraries. Biblioteka cyfrowa jest w nim przedstawiona jako „potencjalna wirtualna organizacja, która w sposób kompleksowy gromadzi, zarządza i zachowuje na długi czas bogate treści cyfrowe oraz oferuje swoim docelowym społecznościom użytkowników wyspecjalizowane funkcje (umożliwiające także personalizację systemu), określoną jakość oraz zgodność z polityką tworzenia całego systemu. Lokuje się ona w centrum działalności intelektualnej człowieka, spełniając różne jego potrzeby: poza intelektualnymi i edukacyjnymi także kulturowe, estetyczne czy rozrywkowe”. Jednocześnie w innych pracach *digital library* utożsamiana jest z biblioteką wirtualną – ogromną, idealną, uniwersalną kolekcją informacji, do której istnieje stały dostęp, bez względu na to, gdzie fizycznie się ona znajduje. W trzecim wariantcie biblioteka cyfrowa jest rozumiana jako kontynuacja działalności rzeczywistej biblioteki, zwłaszcza akademickiej, z poszerzeniem jej funkcjonalności o dodatkowe elementy sfery elektronicznej: portale biblioteczne (z katalogami i systemami linkującymi do innych zasobów w sieci) oraz różnorodne bazy pełnotekstowe [Janiak 2012, s. 15-65].

Natomiast podstawowy dla tego artykułu termin „użytkownicy bibliotek cyfrowych” został w analizie potraktowany bardzo szeroko. Podczas wyszukiwania materiału badawczego nie myślano wyłącznie o użytkownikach ostatecznych / potencjalnych / końcowych, lecz o wszystkich osobach związanych z bibliotekami cyfrowymi. Posiłkowano się kategoryzacją autorstwa Ying Zhang (2010), która podzieliła interesariuszy bibliotek cyfrowych, czyli wszystkich odbiorców i osoby zaangażowane w ich tworzenie, na grupy: użytkowników końcowych (ang. *users*), badaczy (ang. *researchers*), bibliotekarzy (ang. *librarians*), zarządzających biblioteką cyfrową (ang. *developers*) oraz administratorów (ang. *administrators*) [Zhang 2010]. Drugim istotnym podziałem uwzględnionym w trakcie analiz była kategoryzacja scalająca wszystkich pośredników, tj. bibliotekarzy, administratorów, w pewnym sensie także deweloperów, w jedną grupę, zaproponowana przez Gobindę G. Chowdhury’ego, Davida McMenemy’ego i Alana Poultera w pracy

„MEDLIS: Model for Evaluation of Digital Libraries and Information Services” z 2008 r. Przedstawiona w tym artykule trójdzielna typologia użytkowników obejmuje: 1) zespół zarządzający wraz ze wsparciem technicznym, odpowiedzialny za przygotowanie kolekcji, opracowanie metod i technik docierania do nich, a także za tworzenie pomocy dla użytkowników końcowych, 2) pośredników (intermediatorów) używających systemu w imieniu użytkowników końcowych, znających ich potrzeby i umiejętności, 3) użytkowników końcowych [Chowdhury, McMenemy, Poulter 2008]. Wszystkie te osoby były istotne dla tej analizy. Jednakże w wyniku przeglądu relewantnych publikacji ustalono, że przeważają w nich prace dotyczące użytkowników końcowych, zaś te poświęcone projektantom lub administratorom pozostają w mniejszości.

Z punktu widzenia tematyki wiele publikacji odnosiło się do różnorodnych aspektów działania systemu biblioteki cyfrowej, oceny jego jakości, funkcjonalności lub użyteczności. Dlatego też rozpatrywane teksty zostały podzielone na dwa zespoły. Pierwszy z nich skupia publikacje dotyczące systemu biblioteki cyfrowej oraz dokonywanej przez użytkowników oceny jego funkcjonalności, której wyniki inspirują działania projektantów, a w efekcie wpływają na jakość systemu. W drugim natomiast ujęto relacje z badań zmierzających do ustalenia, kim są użytkownicy bibliotek cyfrowych, jakie mają potrzeby informacyjne, z jakimi problemami borykają się podczas wyszukiwania informacji. W zespole pierwszym, liczniejszym, wyodrębniono dodatkowo: 1) publikacje, których dominującym tematem jest system biblioteki cyfrowej i jego zawartość, oraz 2) prace poświęcone analizie możliwości wykorzystania bibliotek cyfrowych w edukacji. Trzeba ponadto zaznaczyć, że ze względu na szeroki zakres tematyczny wiele artykułów można było zaliczyć do kilku podrozdziałów jednocześnie. Ostateczne przyporządkowanie konkretnego tekstu zależało od najważniejszego prezentowanego w nim zagadnienia i od tego, czy główny nacisk padał np. na osobowość i potrzeby użytkowników, czy też na sam system itp.

## **1. System jako najważniejszy element, użytkownik jako „źródło” danych**

### **1.1. Wczesne prace, wydane do roku 2000**

Użytkownicy bibliotek cyfrowych pojawili się już w tytule artykułu z 1993 r. „Users, User Interfaces, and Objects: Envision, a Digital Library” Edwarda A. Foxa i jego współpracowników, opisującego projekt Envision, którego celem było stworzenie „skoncentrowanej na użytkownikach” (ang. *user-centered*) bazy

danych literatury informatycznej, opartej początkowo na publikacjach Association for Computing Machinery (ACM). Użytkownicy zostali określani przez autorów jako eksperci, którzy mogą wspomóc twórców biblioteki cyfrowej w wypracowywaniu rozwiązań w zakresie reprezentacji treści, architektury systemu oraz komunikacji człowiek – komputer (interfejsu) [Fox et al. 2003].

W 1996 r. autorzy dwóch publikacji: 1) „A Browsing Tool of Multi-lingual Documents for Users without Multi-lingual Fonts” [Sakaguchi et al. 1996] oraz 2) „Interactive Term Suggestion for Users of Digital Libraries: Using Subject Thesauri and Co-occurrence Lists for Information Retrieval” [Schatz et al. 1996] wskazywali kolejne problemy informatyczne ważne dla budowniczych bibliotek cyfrowych, związane z terminami służącymi wyszukiwaniu informacji, także wielojęzycznej. Użytkownicy zostali w nich potraktowani jako odbiorcy konkretnych produktów. Rok później ogłoszono podobny tematycznie artykuł „Serving users in many languages: cross-language information retrieval for digital libraries” autorstwa Douglasa Oarda [Oard 1997, dok. elektr.].

Chronologicznie następny cykl artykułów, z lat 1997 i 1998, opublikowany w czasopiśmie „d-Lib Magazine”, ukazał się pod wspólnym tytułem „Safeguarding digital library contents and users”. Został on przygotowany przez liczny zespół współautorów (Henry Gladney, Ulrich Kohl, Jeffrey Lotspiech, Marc Kaplan, Fred Mintzer, Norishige Morimoto i inni), dla których użytkownik nie był jednak głównym podmiotem badań. W artykułach tych rozważano rozmaite problemy informatyczne i prawne związane z udostępnianiem i wykorzystywaniem zawartości kolekcji, tzn. zapisywanie, wysyłanie, kontrolowanie, zabezpieczanie praw autorskich (np. przez znaki wodne), przestrzeganie warunków użytkowania, ustalanie jednoznacznych identyfikatorów itp. Odbiorca został w nich potraktowany jako ostatni element całego układu, ważny, lecz nie kluczowy dla poruszanego tematu [Gladney et al. 1997, dok. elektr.].

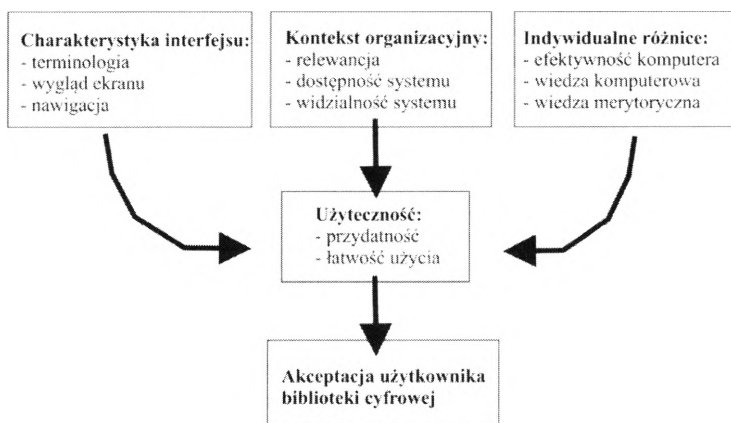
Wskazanie na klienta bibliotek cyfrowych jako na istotny element całego systemu biblioteki cyfrowej pojawiło się w 1997 r. w pracy „The user community as responsibility and resource: building a sustainable digital library” Davida Seamana, w tytule akapitu: „Użytkownicy jako źródło informacji dla przyszłego planowania” (*Users as a resource for future planning*). Jest ona przykładem tego, że w początkowych badaniach nad użytkownikami skupiano się na systemie i jego zawartości, gdyż poprzez badania użytkowników badano bazę. Jeszcze nie myślano o tym, kim jest użytkownik biblioteki cyfrowej, jakie są jego specyficzne potrzeby, zachowania informacyjne itp. Patrzono na bibliotekę cyfrową wyłącznie jako na pełnotekstową bazę danych, która ma być systemem doskonałym, ale jednak podobnym do innych baz danych. Użytkownik miał natomiast za zadanie pomagać w doskonaleniu systemu [Seaman 1997, dok. elektr.].

Także w 1997 roku na konferencji 60th Annual Meeting of the American Society for Information Science poczyniono kolejne kroki w kierunku rozwijania badań *user-centered*. Zauważono wtedy, iż oprócz zagadnień informatycznych, takich jak bazy danych, sieci, pamięci, skalowalność i przetwarzanie obrazu, biblioteki cyfrowe należy badać również w aspekcie ekonomicznym i bezpieczeństwa oraz za pomocą metod etnograficznych i socjologicznych, skupiając się przede wszystkim na użyteczności (ang. *usability*). Podkreślono, iż system jest skuteczny tylko wtedy, gdy zdecydowana większość docelowych użytkowników jest w stanie wykorzystać zaprojektowaną dla nich funkcjonalność. Zaznaczano przy tym, iż opinie użytkowników The Alexandria Digital Library (ADL), pozyskane różnorodnymi technikami (ankiety, analizy informatycznych rejestrów użytkowników, sesje fokusowe, analiza wprowadzanych komentarzy i opinii), oprócz tego, że pozwoliły ustalić, kim są użytkownicy, posłużyły przede wszystkim analizie interfejsu, funkcjonalności i treści ADL oraz stworzeniu planów dalszych działań [Hill et al. 1997, dok. elektr.].

Podobna tematycznie publikacja, „Supporting scholarly inquiry. Incorporating users in the design of the digital library”, w której postulowano włączenie użytkowników w proces projektowania bibliotek cyfrowych, ukazała się rok później w 1998 r., a jej autorami byli Sandra D. Payette i Oya Y. Rieger. Po raz kolejny zaapelowali oni o włączanie użytkowników końcowych do badań nad systemem biblioteki cyfrowej już w fazie jego projektowania [Payette, Rieger 1998].

## **1.2. Publikacje po roku 2000 – system i jego zawartość**

Po roku 2000 ukazało się więcej publikacji, w których zajęto się problematyką użytkowników bibliotek cyfrowych. Tylko nieliczne prace dotyczyły zarówno konstrukcji systemu, jak i jego zawartości. Do tej wąskiej grupy tekstów prezentujących biblioteki cyfrowe w ujęciu całościowym zalicza się m.in. artykuł „Determinants of User Acceptance of Digital Libraries: An Empirical Examination of Individual Differences and System Characteristics” z 2001 r., w którym autorzy postawili tezę, iż pomimo wydawania milionów dolarów na biblioteki cyfrowe, nadal wielu użytkowników je ignoruje, głównie ze względu na brak wiedzy o nich [Hong et al. 2001, dok. elektr.]. Prawie ten sam zespół autorów powtórzył tę tezę także w artykule z 2004 r. „What leads to user acceptance of digital libraries?”. Ciekawym fragmentem tej publikacji było wyszczególnienie czynników wpływających na akceptację biblioteki cyfrowej przez odbiorców i na stopień zadowolenia osiąganego przez nich w trakcie pracy z systemem [Thong, Hong, Tam 2004, dok. elektr.]. Czynniki te przedstawiono na rysunku 1.



Rysunek 1. Model akceptacji przez użytkownika biblioteki cyfrowej  
 Źródło: [Thong, Hong, Tam 2004]. Rysunek przetworzony przez autorkę

Omówienie wyników badań użytkowników było przedmiotem także publikacji z 2002 r. „Evaluation of digital library impact and user communities by analysis of usage patterns”, poświęconej kwestii oceny jakości bibliotek cyfrowych. Przedstawiono w niej system służący do analizowania logów użytkowników, generowanych na poziomie poszczególnych serwerów bibliotek cyfrowych. Zaproponowano podejście ilościowe do oceny biblioteki cyfrowej, oparte na analizie wzorów zachowań informacyjnych i określeniu struktury danej społeczności użytkowników, a w szczególności ich przyzwyczajień i skłonności w zakresie wyboru i pozytywnej oceny zbiorów dokumentów tworzących dane kolekcje [Bollen, Luce 2002, dok. elektr.].

W kolejnym artykule, „Digital libraries: to meet or manage user expectations”, David Bawden i Polona Vilar stwierdzili natomiast, iż oczekiwania użytkowników bibliotek cyfrowych są często nierealistyczne, nadmiernie wysokie, głównie z powodu wszechobecności wyszukiwarek internetowych funkcjonujących jako podstawowe środowisko informacyjne. Typowe wymagania stawiane bibliotece cyfrowej są następujące: 1) biblioteka ma być kompleksowa – mieścić w sobie wszystko („panbiblioteka”), 2) cała jej zawartość ma być dostępna, 3) reakcje systemu mają być błyskawiczne, aby zapewnić użytkownikom natychmiastowe zadowolenie, 4) dane mają być jednolite, 5) system powinien być wyposażony w jeden, łatwy w obsłudze interfejs, 6) treści mają być zapisane w wielu formatach – tekst, obrazy, dźwięk. Chcąc uprzystępnąć użytkownikom zasady działania bibliotek cyfrowych, autorzy wyjaśnili, dlaczego niektóre z oczekiwań zgłaszanych wobec bibliotek cyfrowych są niemożliwe do spełnienia. W tym celu:

1) wyszczególnili, jakie zasoby są niedostępne w bibliotekach cyfrowych, a zarazem zaznaczyli, że 2) tylko niewiele materiałów prymarnych jest dostępnych, 3) istnieje wiele poziomów wyszukiwania, które należy wykorzystać w trakcie pracy oraz 4) istnieje wiele punktów dostępu do danych, ale nie są one jednolite. Wnioski z tych badań stały się też podstawą dla stworzenia narodowego programu Library of Congress's National Digital Library Program. Podsumowując, można stwierdzić, iż zasada „nie każ mi myśleć” jest niemożliwa do zastosowania w tak rozbudowanych systemach informacyjnych, jakimi są biblioteki cyfrowe [Bawden, Vilar 2006, dok. elektr.].

Następnym tekstem z grupy prac przedstawiających biblioteki cyfrowe w sposób globalny, zarówno w aspekcie systemowym, jak i z punktu widzenia zasobów, jest ogłoszony w 2009 r. artykuł „A Practical Guide for Building a User-Focused Digital Library Collection”, w którym autorzy znów wzywają do badania użytkowników, a zarazem apelują o dostarczanie użytkownikom pracującym w dowolnym miejscu i w dowolnym czasie interesujących ich treści cyfrowych. Znajomość potrzeb informacyjnych odbiorców umożliwi bowiem tworzenie i wdrażanie narzędzi pomocowych, które zwiększają widoczność zbiorów bibliotek cyfrowych, a co za tym idzie – ułatwiają klientom dotarcie do tych zasobów [Zhu, Guevara 2009, dok. elektr.].

W artykule z tego samego roku pt. „Characterizing a Digital Library's Users. Steps towards a nuanced view of the user” Lorraine F. Normore omówiła rolę badania użytkowników w doskonaleniu metod i efektów projektowania systemów informacyjno-wyszukiwawczych bibliotek cyfrowych. Autorka wyróżniła: 1) badania prowadzone w początkowej fazie tworzenia biblioteki cyfrowej lub innego projektu, mające służyć poznaniu potrzeb, umiejętności, zachowań informacyjnych odbiorców w celu stworzenia jak najlepszej kolekcji oraz 2) badania użyteczności systemu w późniejszych stadiach zaawansowania projektu, które jednocześnie podsumowują działania jego twórców [Normore 2009, dok. elektr.].

Najbardziej interesująca z tej grupy publikacji jest książka z 2012 r. przygotowana pod redakcją Mileny Dobrewej, Andy'ego O'Dwyera i Pierluigiego Feliciti pt. „User Studies for Digital Library Development”, przynosząca wieloaspektowe omówienie problematyki badań użytkowników. Po pierwsze, uzasadniono w niej celowość takich badań, stwierdzając że ich rezultaty pozwalają: 1) zapewniać lepsze usługi, 2) radzić sobie z nadmiarem informacji, 3) eliminować problemy związane z wartością i przepływem informacji w społeczeństwie informacyjnym i w gospodarce opartej na wiedzy, 4) elastycznie określać modele kosztów i pobierania opłat (koszt uzyskania dostępu, usługi abonamentowe) za korzystanie z bibliotek cyfrowych. Po drugie, zidentyfikowano powody, dla których rezygnuje się z badania odbiorców, a były to: 1) brak czasu, 2) brak pieniędzy, 3) brak wie-

dzy specjalistycznej, 4) brak odpowiedniego warsztatu do badania użyteczności, 5) brak wiedzy o tym, jak interpretować rezultaty badawcze. Po trzecie, zdefiniowano, czym jest jakość bibliotek cyfrowych, stwierdzając że dobra biblioteka to taki produkt programistyczny, który – w określonych okolicznościach użycia – umożliwia użytkownikom osiągnięcie założonych celów w sposób skuteczny, produktywny, bezpieczny i dający zadowolenie. Opisano też mity na temat użytkowników, które funkcjonują w środowisku projektantów, administratorów i deweloperów bibliotek cyfrowych. Najważniejszy z nich wyraża się w przeświadczeniu: „wiemy, jacy są nasi użytkownicy, znamy ich”. Tymczasem w większości przypadków twórcy bazują na własnej wiedzy niepopartej badaniami. A przecież należy zbadać: 1) co interesuje użytkowników bibliotek cyfrowych oprócz nauk humanistycznych i sztuki (które stanowią największy procent zbiorów na świecie), 2) jak ważna jest wielojęzyczność i międzykulturowe wykorzystanie informacji, 3) jak wyrafinowane są potrzeby i oczekiwania użytkowników, 4) jak wzbogacić i ponownie wykorzystać kolekcje cyfrowe, 5) jak zmienić przyzwyczajenia użytkowników, skłaniając ich do tego, aby nie tylko konsumowali informację, ale też włączali ją w „przeżywanie” bogatszych doświadczeń. Drugi mit symbolizuje hasło: „gdy zbudujemy bibliotekę cyfrową, użytkownicy sami przyjdą”. Jest to przekonanie błędne, gdyż „popyt” musi być napędzany przez wiedzę o potrzebach informacyjnych użytkowników, które zmieniają się w czasie, a prócz tego należy dopracowywać mechanizmy personalizacji: nie można skupiać się tylko na grupie, ważny jest każdy odbiorca. Trzecim mitem jest stwierdzenie: „im więcej zaoferujemy, tym bardziej wszyscy będą zadowoleni”. Trzeba dokładniej przemyśleć kwestie personalizacji systemu, a zwłaszcza mechanizmów rekomendujących wspomagających użytkowników w ich codziennej pracy, tak aby nie poczuli się oni przytłoczeni nadmiarem informacji. Kolejny mit głosi, że badania użytkowników są nierozdzielnie związane z badaniami jakości oraz użyteczności systemu, podczas gdy w istocie służą również lepszemu zrozumieniu klienta. Ostatni już, piąty mit mówi, iż jakość oznacza innowację, wobec czego należy zakładać, uznając to za pewnik, iż wszyscy użytkownicy chcą najnowszych technologicznych gadżetów i całej palety możliwych usług. Odbiorcy często chcą zaś przede wszystkim szybko dotrzeć do potrzebnych informacji i „narzucające” im się w trakcie wyszukiwania usprawnienia nie zawsze są przez nich pozytywnie przyjmowane.

W poszczególnych rozdziałach omawianej książki zaprezentowano pokrótce: modele bibliotek cyfrowych, zachowania użytkowników, problemy projektowania, potrzeby użytkowników oraz ocenę bibliotek cyfrowych. Przedstawiono także różne techniki i metody badań użytkowników, w tym sondaże i wywiady, analizy logów, śledzenie ruchu oczu (ang. *eye tracking*), analizy eksperckie. Poza tym omówiono wielojęzyczny dostęp do zasobów, biblioteki cyfrowe dla dzieci,

media społecznościowe, ochronę zasobów cyfrowych, rozwój urządzeń mobilnych, punkty dostępowe, metadane w Europeanie. Zilustrowano również możliwości wykorzystania list dystrybucyjnych (w szczególności w kontekście szkolnictwa wyższego) oraz kursów projektowania, e-learningu i zasobów bibliotek cyfrowych w działalności instytucji kultury, takich jak muzea i archiwa [Dobrevá, O'Dwyer, Feliciati eds. 2012].

Ostatnio, już w 2014 r., ukazała się praca „A Survey of User Perceptions of Digital Library E-Quality and Affinity” [Zha, Xiao, Zhang 2014]. Głównym jej tematem są systemy informacyjne (IS), lecz autorzy piszą również o modelu odnoszenia sukcesu przez klientów. Przedstawiają mianowicie wyniki badania percepcji użytkowników bibliotek cyfrowych (357 przebadanych osób) w odniesieniu do: 1) jakości informacji (ang. *information quality*), która wpływa na sukces semantyczny: użytkownicy rozumieją odbierane treści, 2) jakości technicznej systemu (ang. *system quality*), oddziałującej na sukces techniczny: użytkownicy potrafią poruszać się w obrębie systemu, 3) jakości usług (ang. *service quality*), która przyczynia się do sukcesu aplikacyjnego: użytkownicy sprawnie korzystają z poszczególnych funkcji serwisu. Autorzy dodali aplikację powinowactwa (ang. *affinity*), czyli poczucia wspólnoty, przynależności do społeczności biblioteki cyfrowej (sukces skuteczności), zaznaczając – w podsumowaniu wyników badań – że e-jakość wpływa na tworzenie się e-powinowactwa. Wzorcowy relacyjny model odnoszenia korzyści, opracowany przez W.H. DeLonea i E.R. McLena, przedstawiony został na rysunku 2.



Rysunek 2. Satisfakcja użytkownika

Źródło: [Thong, Hong, Tam 2004], wzorowany na [De Loeane, McLean 2003].

Rysunek przetworzony przez autorkę

### 1.3. Publikacje po roku 2000 – system: struktury, funkcjonalność, interfejs

Wśród publikacji, w których znalazły się nawiązania do użytkowników bibliotek cyfrowych, większość stanowią jednak te poświęcone w pierwszym rzędzie systemowi, jego funkcjonalnościom oraz możliwościom interfejsu. I tak autorom dwóch prac z 2000 r. przyświecała maksyma: „jeśli chce się stworzyć system, któ-

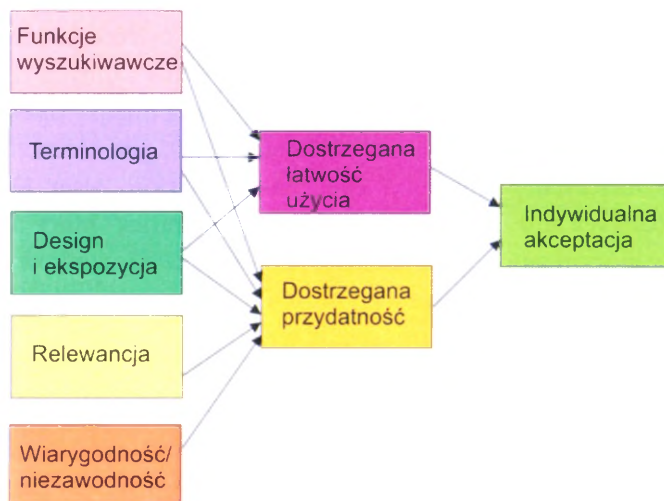


ry ma satysfakcjonować użytkowników, to należy zadać sobie pytanie, czy klienci końcowi są zadowoleni z usług oferowanych przez biblioteki cyfrowe i czy ich opinie na ten temat były badane”. Chodzi o artykuły Mounira A. Khalila „Are End-Users Satisfied by Using Digital Libraries?” oraz Tiziany Catarci „Digital Libraries: Did You Ask the Users?” [Khalil 2000, dok. elektr.; Catarci 2000, dok. elektr.]. W tym samym roku w publikacji „A user-centered interface for information exploration in a heterogeneous digital library” Michelle Q. Wang Baldonado, uzasadniając konieczność projektowania nowych interfejsów użytkownika dla bibliotek cyfrowych, zaprezentowała SenseMaker – interfejs do wyszukiwania informacji w heterogenicznym środowisku biblioteki cyfrowej, który umożliwia lepsze przedstawianie wyników, uwidocznienie cytatów i innych relacji między artykułami oraz zwiększenie płynności przechodzenia między operacjami przeglądania, wyszukiwania i filtrowania [Baldonado 2000, dok. elektr.].

Do najważniejszych publikacji z 2002 należy artykuł „User Assessment of Electronic Reserves and Implications for Digital Libraries” Mary Sellen i Brendy Hazard, w którym autorki podkreślają, iż reakcje użytkowników w zetknięciu z zasobami elektronicznymi mogą wyrażać ich aprobujący stosunek do tych zasobów, choć bywa też, że wskazują na aspekty wymagające udoskonalenia, a tym samym stawiają wyzwania techniczne przed twórcami bibliotek cyfrowych [Sellen, Hazard 2002, dok. elektr.]. W drugim artykule, „Understanding user acceptance of digital libraries: what are the roles of interface characteristics, organizational context, and individual differences?” James Y.L. Thong, Weiyin Hong i Kar-Yan Tam przeanalizowali informacje otrzymane od 397 użytkowników bibliotek cyfrowych, wykorzystując w tym celu model akceptacji technologii (TAM). Zidentyfikowali różne cechy interfejsu oraz zmienne kontekstowe i zewnętrzne, wpływające na ocenę użyteczności, a także indywidualne różnice w wykorzystywaniu zasobów. Wyniki pokazały, że użyteczność, przyjazność i funkcjonalność są podstawowymi czynnikami warunkującymi poziom akceptacji systemu przez użytkowników, a co za tym idzie – wysoką ocenę jakości bibliotek cyfrowych [Thong, Hong, Tam 2002, dok. elektr.]. W kolejnej publikacji z 2002 r., „Thesauri and ontologies in digital libraries”, Dagobert Soergel wspomniał o użytkownikach jako o odbiorcach słowników i ontologii, przygotowywanych na potrzeby bibliotek cyfrowych. Najważniejszym tematem była jednak ponownie tylko wybrana część systemu informacyjno-wyszukiwawczego BC, bez zamierzenia, by ująć go w sposób całościowy [Soergel 2002, dok. elektr.].

Kilka lat później, w 2005 r. Kirsten R. Butcher i Sonal Bhushan w artykule „Using strand maps to engage digital library users with science content” także skupili się na jednym z elementów systemu informacyjnego biblioteki cyfrowej i ustalili, że mapy stron, zestawy kolekcji tworzone w portalach bibliotek cyfro-

wych są oceniane przez użytkowników jako pomocne [Butcher, Bhushan 2005, dok. elektr.]. Dość ogólna była natomiast publikacja „User Acceptance of Digital Library: an empirical exploration of individual and system components”, której autorzy określili podstawowe elementy systemu wpływające na akceptowanie biblioteki cyfrowej przez użytkowników [Vaidyanathan, Sabbaghi, Bargellini 2005, dok. elektr.]. Wizualizacja uzyskanych przez nich wyników badawczych została przedstawiona na rysunku 3.



Rysunek 3. Czynniki wpływające na indywidualną akceptację biblioteki cyfrowej przez użytkownika  
Źródło: [Vaidyanathan, Sabbaghi, Bargellini 2005]. Rysunek przetworzony przez autorkę

Następnym składnikiem (komponentem) systemu biblioteki cyfrowej, opisywanym w kilku artykułach, są mechanizmy personalizacji. Problem ten poruszyli m.in. Enrique Frías-Martínez, George Magoulas, Sherry Chen i Robert Macredie w artykule „Automated user modeling for personalized digital libraries”, który stanowi głos w dyskusji nad koniecznością stałego doskonalenia jakości usług cyfrowych. Autorzy opisali mechanizm konstruowania list dystrybucyjnych, które odpowiadają potrzebom informacyjnym użytkownika, oraz mechanizm adaptacyjny, który automatycznie „uczy się”, jakie są preferencje i cele użytkownika, a następnie dostosowuje do nich procedury interakcji, np. przez podsuwanie klientowi różnego rodzaju podpowiedzi, sugestii wykorzystania określonych, potencjalnie interesujących dokumentów. W konkluzjach przedstawili dwa modele: generycznej zindywidualizowanej architektury adaptacyjnej i adaptacyjnej biblioteki cyfrowej (rysunki 4 i 5).

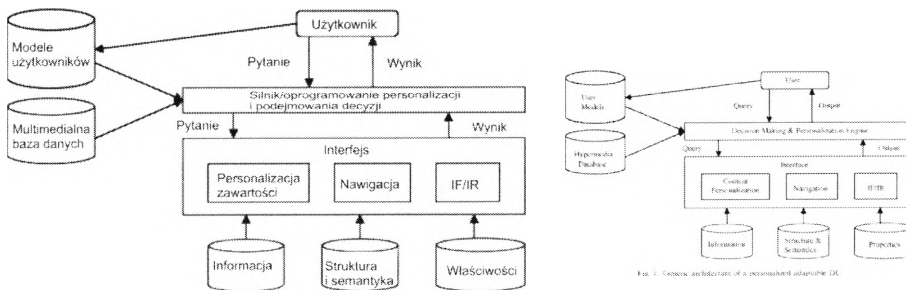
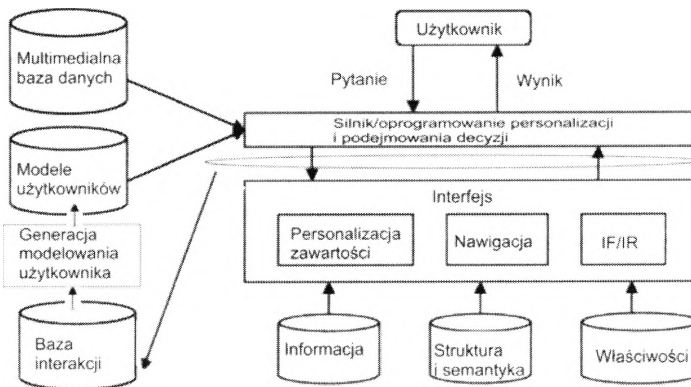


Fig. 3. Generic architecture of a personalized adaptive DL.

Rysunek 4. Generyczna zindywidualizowana architektura adaptacyjna

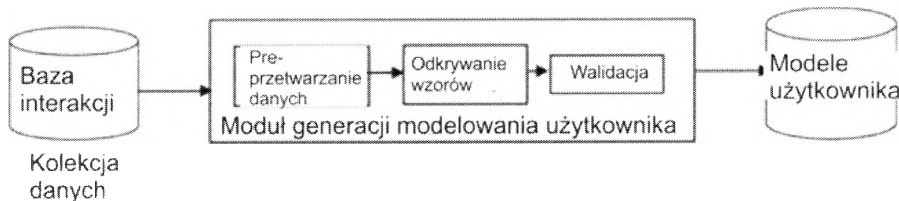
Źródło: [Frias-Martínez, Magoulas, Macredie 2006]. Rysunek przetworzony przez autorkę



Rysunek 5. Generyczna architektura adaptacyjna

Źródło: [Frias-Martínez, Magoulas, Macredie 2006]. Rysunek przetworzony przez autorkę

Autorzy zwiualizowali także proces tworzenia modeli użytkowników, który przedstawia się następująco:



Rysunek 6. Proces tworzenia modeli użytkowników

Źródło: [Frias-Martínez, Magoulas, Macredie 2006]. Rysunek przetworzony przez autorkę

Podjęcie klastrowe, które umożliwia zapewnienie bardziej efektywnej personalizacji biblioteki cyfrowej, zaprezentowali Enrique Frías-Martínez, Sherry Y. Chen, Robert D. Macredie i Xiaohui Liu w pracy z 2007 r. „The role of human factors in stereotyping behavior and perception of digital library users: a robust clustering approach”. Zidentyfikowali w niej główne „czynniki ludzkie” (ang. *human factors*) wyznaczające zachowania użytkowników. Są nimi: style poznawcze, poziom wiedzy oraz różnice między płciami. Podkreślili, że dzięki uprzedniemu skategoryzowaniu użytkowników (właśnie według wymienionych czynników) można lepiej zaprojektować elementy systemu, zwłaszcza te wspomagające wyszukiwanie [Frías-Martínez, Chen, Macredie, Liu 2007, dok. elektr.].

Ci sami autorzy, w nieco tylko uszczuplonym składzie, w innym artykule – „Automatic cognitive style identification of digital library users for personalization” – dobitnie zaznaczyli, iż nie ma jednego interfejsu bibliotek cyfrowych, który odpowiadałby wszystkim użytkownikom. Dlatego też jednym z kluczowych elementów, które mogą wpłynąć na poprawę satysfakcji odczuwanej przez użytkowników, jest właśnie personalizacja i dostosowywanie funkcjonalności systemu do konkretnych potrzeb. Niestety, zdobycie wiedzy o sposobach, w jaki klienci personalizują bibliotekę cyfrową, wymaga przeprowadzenia czasochłonnych testów określających style poznawcze. W swoim studium autorzy postanowili zatem zbadać, jak różne systemy klasyfikacji mogą być wykorzystywane do automatycznego identyfikowania stylów poznawczych użytkowników, które można zaobserwować w trakcie ich interakcji z bibliotekami cyfrowymi, a następnie opracowywać na tej podstawie automatyczne zestawy do personalizacji systemu [Frías-Martínez, Chen, Liu 2007, dok. elektr.]. Kontynuację tego wątku badawczego stanowi artykuł z następnego roku – „Investigation of behavior and perception of digital library users: A cognitive style perspective” [Frías-Martínez, Chen, Liu 2008, dok. elektr.].

Rok 2008 przyniósł ponownie publikacje o interfejsach. Jako pierwszą przedstawić należy pracę „Evaluation & Comparison of Free and Commercial User Interfaces of Digital Libraries” [Alijani, Nikkar 2008], w której porównano cztery biblioteki cyfrowe: Gutenberg project, 24x7, Netlibrary i Ebrary. Trybom wyszukiwania w różnych interfejsach poświęcona jest także publikacja „Effects of interaction design in digital libraries on user interactions”. Zrelacjonowano w niej wyniki badań podjętych w celu ustalenia, jak różne funkcje wyszukiwania i przeglądania w bibliotekach cyfrowych, oparte na trzech systemach operacyjnych: ACM, IEEE CS, IEEE Xplore, wpływają na poziom interakcji mało doświadczonego użytkownika z interfejsem. Przeprowadzono w tym celu eksperymenty z 35 klientami, które wykazały istnienie różnic w poziomie wykorzystania funkcjonalności bibliotek cyfrowych obsługiwanych przez różne systemy operacyjne [Zhang et al. 2008, dok. elektr.].

Rok później (2009) ukazał się artykuł „A Linguistic Recommender System For University Digital Libraries To Help Users In Their Research Resources Accesses”, w którym zaprezentowano systemy Recommender – narzędzia oparte są na językowych modelach zbiorów rozmytych, umożliwiające przefiltrowanie ogromnych ilości informacji dostępnych w internecie i ich ocenę. Badania z użyciem narzędzi Recommender prowadzone były w środowisku cyfrowej biblioteki uniwersyteckiej [Porcel, Lizarte, Herrera-Viedma 2009, dok. elektr.].

W kolejnej publikacji, „Understanding user requirements and preferences for a digital library Web portal” [Agosti et al. 2010], odnoszącej się do konkretnego produktu informacyjnego, w tym przypadku do portalu biblioteki cyfrowej, przytoczono wyniki badań przeprowadzonych w ramach projektu TELplus. Miały one dać wgląd w potrzeby użytkowników i ich preferencje co do usług bibliotek cyfrowych oferowanych przez portal Biblioteki Europejskiej. Celem przeprowadzonych analiz było zdobycie wiedzy o tym, w jaki sposób poprawić jakość obecnych usług portalu internetowego, aby przygotować go dla przyszłych użytkowników.

Do konkretnego projektu zawężone były też badania omówione w publikacji „A New Focus on End Users: Eye-Tracking Analysis for Digital Libraries” [Sykes et al. 2010]. Prototyp v1.0 Europeana był analizowany przez pryzmat danych dotyczących użytkowników i sposobów ich funkcjonowania, pozyskanych techniką *eye-tracking*. Określono w ten sposób obszary ekranu najintensywniej przyciągające wzrok odbiorcy, a zarazem zidentyfikowano te elementy interfejsu, które wprawdzie skupiają na sobie uwagę użytkowników, lecz nie są efektywnie wykorzystywane w wyszukiwaniach. Wnioski z badań posłużyły do stworzenia zaleceń w sprawie przyszłego rozwoju biblioteki cyfrowej Europeana, a zwłaszcza do przygotowania bardziej intuicyjnego panelu wyszukiwania zaawansowanego.

#### **1.4. Publikacje po roku 2000 – zawartość, kolekcje, różne typy dokumentów**

Kolejną grupę rozpatrywanych publikacji tworzą prace przedstawiające wyniki analiz zawartości bibliotek cyfrowych i dotyczące różnorodnych aspektów tworzenia kolekcji, w tym treści specjalistycznych. I tak w roku 2000 ukazał się tekst poświęcony wykorzystaniu źródeł medycznych przez wykładowców akademickich, lekarzy, pacjentów i bibliotekarzy: „Digital libraries’ support for the user’s Information Journey” [Adams, Blandford 2000]. Wyrażono w nim tezę, że to głównie bibliotekarze są świadomi, jak wielkie są zasoby bibliotek cyfrowych i jak w pełni wyzyskać ich potencjał. Oni też byli chętni do korzystania z usługi alertów informujących o zmianach w zawartości kolekcji.

W pracy z tego samego roku pt. „Usability, User Preferences, Effectiveness, and User Behaviors When Searching Individual and Integrated Full-Text Databases: Implications for Digital Libraries” Soyeon Park zwrócił uwagę na problem wspierania wyborów użytkownika przez integrację charakterystyk poszczególnych baz danych dostępnych w ramach biblioteki cyfrowej. Widać tu rozumienie „biblioteki cyfrowej” jako biblioteki wirtualnej zbudowanej z rozproszonych baz pełnotekstowych [Park 2000, dok. elektr.].

Do tej grupy zaliczyć można też prace, których tematyka skupia się wokół różnych typów zasobów gromadzonych w BC: filmów, nagrań muzycznych lub grafiki. Na przykład w artykule „Cumulating and sharing end users knowledge indexing in a video digital library” z 2001 r. Marc Nanard i Jocelyne Nanard opisali środowisko OPALES, stworzone w Institut National de l’Audiovisuel (INA) w Paryżu. Pozwalało ono użytkownikom na korzystanie z mechanizmu dzielenia się wiedzą i z narzędzi ułatwiających docieranie do zarchiwizowanych filmów. W portalu przygotowano obszary robocze dla użytkowników, wyposażone w funkcje nieingerujące w struktury opisów, ale pomagające opisywać i interpretować zasoby (zgodnie z wiedzą całej społeczności) [Nanard, Nanard 2001, dok. elektr.].

Filmów wideo dotyczyła także publikacja wieloautorska z 2002 r. pt. „Alternative Surrogates for Video Objects in a Digital Library: Users’ Perspectives on Their Relative Usability”, w której przedstawiono wyniki badania przydatności w realizacji konkretnych zadań surogatów filmów (scenopisów obrazkowych ze słowami kluczowymi, pokazów slajdów z tekstem oraz dźwiękiem, z usługą szybkiego przesuwania ich do przodu). Dziesięciu uczestników, z których wszyscy oglądali filmy i przynajmniej od czasu do czasu poszukiwali nagrań, przedstawiło komentarze na temat mocnych i słabych stron każdego z substytutów. Ponadto przeprowadzono szereg testów sprawdzających efektywność odbioru (określenie istoty filmu – ang. *gist*, rozpoznawanie obiektów, rozpoznawanie działań). Uczestnicy podkreślili, iż zrozumienie sedna przekazu audiowizualnego (zwłaszcza po obejrzeniu zamiennika / surogatu, będącego w dużym stopniu streszczeniem) zależy od trzech elementów: aktualności / tematyczności, fabuły i wizualnego streszczenia filmu [Wildemuth et al. 2002].

Jednym z problemów badawczych odnoszących się do zawartości bibliotek cyfrowych jest też recepcja udostępnianych w nich treści przez użytkowników. Gemma Madle, Patty Kostkova, Jane Mani-Saada i Julius R. Weinberg (2003) zajęli się tą kwestią m.in. w publikacji „Evaluating the Changes in Knowledge and Attitudes of Digital Library Users”, która dotyczyła informacji medycznej. W ramach eksperymentu z zastosowaniem pre- i post-kwestionariuszy ocenili zmiany zachodzące w stanie wiedzy i postawach użytkowników the National Electronic

Library for Communicable Disease, biblioteki cyfrowej gromadzącej piśmiennictwo z zakresu chorób zakaźnych [Madle et al. 2003, dok. elektr.].

Autorem, który poruszył z kolei problem dostępu do kolekcji wideo w interaktywnych bibliotekach cyfrowych, był Dan E. Albertson. W artykule „Influences of users' familiarity with visual search topics on interactive video digital libraries” z 2010 r. przedstawił on wyniki badań, które potwierdziły tezę, iż zrozumienie działań potencjalnych użytkowników pomaga naukowcom dopasować do nich prototypowe projekty systemowe. Albertson badał, czy i jak znajomość wizualnych tematów wyszukiwania jest skorelowana z innymi czynnikami w kontekście interaktywnego pobierania danych (tu: wideo). Przeprowadził w tym celu eksperyment, w którym 28 nauczycieli przedmiotów ścisłych miało wykonać sześć tematycznych wyszukiwań za pomocą prototypowego systemu do pobierania klipów wideo z kolekcji NASA Science Education Programs. Analiza wyników wykazała, że dostęp do informacji jest najbardziej efektywny, gdy merytoryczna wiedza użytkownika jest skorelowana z innymi czynnikami wpływającymi na jego zachowania informacyjne, a zwłaszcza z umiejętnością korzystania z interaktywnych możliwości systemu informacyjno-wyszukiwawczego. Wtedy też ocena prototypu systemu biblioteki cyfrowej jest najwyższa [Albertson 2010].

W publikacjach z 2010 r. omawiana była także problematyka wykorzystania zasobów specjalistycznych. W artykule „A User-Centered Design of a Personal Digital Library for Music Exploration” David Bainbridge, Brook J. Novak i Sally Jo Cunningham przedstawili wyniki swoich badań, podczas których sondowali główne oczekiwania dziewięciu muzyków względem zawartości i funkcjonalności biblioteki cyfrowej gromadzącej muzykalia. Uczestnicy badania domagali się: możliwości nagrywania, korzystania z dubbingu, dobrego opracowania materiału, możliwości przechowywania wybranych dzieł i organizowania zasobów stosownie do swoich potrzeb. Podobne badanie, tyle że dotyczące korzystania z kolekcji grafiki, a konkretnie z serwisu MediaGLOW, zostało podsumowane w artykule „Flexible Access to Photo Libraries via Time, Place, Tags, and Visual Features” [Girgensohn et al. 2010, dok. elektr.]. Dla klientów cyfrowej biblioteki zdjęć ważne były: estetyka, przypisanie lokalizacji GPS do sfotografowanych obiektów, nadane przez użytkownika znaczniki, możliwości filtrowania według różnych kryteriów. Bardzo zadowoleni byli z algorytmu, który automatycznie umieszczał fotografie z podobnymi atrybutami blisko siebie.

Dwa lata później (2012) opublikowano artykuł dotyczący wielojęzycznej zawartości bibliotek cyfrowych pt. „Multilingual Needs and Expectations in Digital Libraries: A Survey of Academic Users with Different Languages”, zawierający omówienie badania ankietowego użytkowników bibliotek akademickich, prze-

prowadzonego w celu określenia ich potrzeb i oczekiwań w zakresie przetwarzania informacji wielojęzycznej podczas interakcji z bibliotekami cyfrowymi. Autorzy przeanalizowali różne aspekty wielojęzycznych potrzeb i oczekiwań uczestników, ich wielojęzycznego zachowania, często używanych wielojęzycznych zasobów informacyjnych i żądanych funkcji dla usług pobierania treści, a także wielojęzyczny interfejs bibliotek cyfrowych. Podkreślili, iż dostępne narzędzia tłumaczeniowe nie są jeszcze zadowalające [Wu, He, Luo 2012, dok. elektr.].

Ważnym zagadnieniem związanym z kolekcjami bibliotek cyfrowych jest również ochrona ich zawartości. Temat ten został podjęty m.in. w publikacji „From digital libraries to digital preservation research: the importance of users and context”, wydanej w roku 2010, w której Gobinda Chowdhury przedstawił różne aspekty badań nad metodami i rozwiązaniami stosowanymi w ochronie zasobów cyfrowych, także tych elementów, które wykorzystywane są w obsłudze użytkowników do zabezpieczania ich praw, anonimowości itp. Podkreślił, że w obszarze tym wiele pozostaje jeszcze do zrobienia zarówno od strony prawnej, jak i systemowej [Chowdhury 2010, dok. elektr.].

Ostatnim z omawianych tu zagadnień związanych z zawartością bibliotek cyfrowych jest uzupełnianie istniejących kolekcji o nowe treści. W artykule z 2013 r. pt. „Social tagging in a scholarly digital library environment: users' perspectives” autorzy opisali tagowanie jako jeden ze sposobów ilościowego rozbudowywania treści w bibliotece cyfrowej [Noorhidawati, Fariza Hanum, Zohoorian-Fooladi 2013, dok. elektr.]. Tagowanie wydawało się autorom ważnym elementem, który, jak zakładali w hipotezie wyjściowej, jest chętnie wykorzystywany przez użytkowników. Tymczasem w wyniku badań na grupie użytkowników, przeprowadzonych techniką ankiety i wywiadu, okazało się że większość z nich nie używa tagowania społecznego, głównie dlatego, że nie wie o istnieniu takiej możliwości w systemie BC. We wnioskach swojego artykułu autorzy postulowali więc, aby podejmować akcje informowania odbiorców o najnowszych możliwościach systemów, starając się wspomóc przyszły rozwój narzędzi znakowania społecznego w bibliotekach cyfrowych.

Ostatni artykuł z tej grupy, „Understanding the Information Needs of Large-Scale Digital Library Users: Comparative Analysis of User Searching”, ogłoszony już w 2014 r., dotyczy dwóch amerykańskich bibliotek cyfrowych: National Science Digital Library i Opening History. Autorki Oksana Zavalina i Elena V. Vassilieva przeanalizowały w nim możliwości wyszukiwawcze obu systemów oraz zachowania informacyjne użytkowników, wskazując iż najczęstszym sposobem docierania do zasobów jest szukanie przez: słowo kluczowe / przedmiot (ang. *object*) lub pojęcie, potem dopiero przez miejsca, osoby, hasła korporatywne, wydarzenia itp. [Zavalina, Vassilieva 2013, dok. elektr.].



## 1.5. Publikacje po roku 2000 – edukacyjne biblioteki cyfrowe jako przedmiot badań

Warto omówić jeszcze serię artykułów dotyczących badań użytkowników bibliotek cyfrowych tworzonych dla celów edukacyjnych. Jedną z ciekawszych publikacji w tej grupie jest artykuł z roku 2001, w którym przytoczono wyniki dokonanej przez użytkowników oceny działania bibliotek cyfrowych, w tym oczywiście interfejsów: „Use of a digital work environment prototype to create a user-centered university digital library”. Autorzy zaproponowali w nim model architektury prototypowego cyfrowego środowiska pracy (DWE), które miało ułatwiać użytkownikom ze społeczności akademickiej docieranie do zasobów informacyjnych. Elementy DWE, które pomagają tworzyć bibliotekę wirtualną, pozwalają identyfikować naukowe zasoby informacyjne, różne kategorie użytkowników oraz związane z tymi kategoriami zadania. Cyfrowe zasoby informacyjne obejmują artykuły pełnotekstowe, bazy danych, prace doktorskie, e-czasopisma, e-książki, bazy danych multimedialnych, ale też kalendarze studiów uniwersyteckich, statuty uczelni, informacje o rejestracji na zajęcia, wytyczne dla tworzenia prac, poradniki itp. Ponieważ informacje te mają charakter rozproszony (znajdują się na różnych serwerach i na różnych stronach domowych), użytkownikom trudno jest dotrzeć do nich w pożądanym przez nich formacie i we właściwym czasie. Prototyp DWE został zaprojektowany w Szkole Informatyki Nanyang Technological University, a jego ważną cechą jest prywatny obszar roboczy systemu, służący do przechowywania publikacji użytkownika, zaznaczania ważnych hiperłączy i tworzenia notatek. Ponownie widać tutaj rozszerzenie definicji biblioteki cyfrowej i rozumienie jej jako systemu informacyjnego, który umożliwia dostęp do danych o toku studiów, proponowanych kierunkach studiów, zajęciach itp. [Meyyappan, Chowdhury, Foo 2001, dok. elektr.].

Temat edukacji pojawił się też w pracy o cztery lata późniejszej: „Influencing User Behavior through Digital Library Design” [Manduca et al. 2005]. Autorzy skupili się w niej na bibliotekach wirtualnych wspomagających nauczanie, w tym przypadku nauczanie w ramach kursów z zakresu geologii. Wskazali, iż najbardziej pomocne w projektowaniu i doskonaleniu portali bibliotecznych są: 1) techniki *user-centered design*, które angażują użytkowników do projektowania i testowania BC, 2) badania sprawdzające, jak wykorzystywane są biblioteki cyfrowe na potrzeby pracy / nauki i jakie nawyki można zaobserwować w docelowej grupie użytkowników oraz 3) techniki oceny webometrycznej, które pozwalają zrozumieć wzorce wyszukiwań realizowanych przez użytkowników. W badaniu przeprowadzonym techniką wywiadu sprawdzano sposoby wykorzystania źródeł w trakcie zajęć dydaktycznych oraz poza nimi.

Kolejne studium przypadku zaprezentował Pertti Vakkari w artykule z 2006 r. pt. „Trends in the use of digital libraries by scientists in 2000-2005”, w którym przytoczył, w układzie według tematów, statystyki wykorzystania dokumentów z biblioteki FinELib, Finnish Electronic Library przez trzy grupy użytkowników końcowych: profesorów, wykładowców, asystentów w latach 2000-2005 [Vakkari 2006, dok. elektr.].

Kolejna ważna publikacja z tego roku to „Digital libraries: what do users want?”, w której Elahe Kani-Zabihi, Gheorghita Ghinea i Sherry Y. Chen wysunęli tezę, iż bez względu na to, na jakim oprogramowaniu została zbudowana biblioteka cyfrowa, oczekiwania w stosunku do jej funkcjonalności są podobne. Teza ta została sformułowana po przebadaniu przedstawicieli społeczności akademickiej (studentów, pracowników administracji i wykładowców), którzy w następujący sposób określili swoje podstawowe preferencje: 1) informacje powinny być łatwo i szybko odnajdywalne, 2) biblioteka cyfrowa powinna dostarczać wskazówek, pomagać użytkownikom w docieraniu do pożądaných informacji, 3) użytkownicy powinni mieć możliwość łatwego zapoznania się / oswojenia się (ang. *to familiarize with*) z biblioteką cyfrową, 4) interakcja z biblioteką cyfrową powinna mieć pozytywny wpływ na metody pracy stosowane przez poszczególnych użytkowników, 5) biblioteka cyfrowa powinna wspierać wspólną pracę opartą na wymianie wiedzy. Badani użytkownicy zgłaszali też postulaty dotyczące funkcjonalności systemu informacyjnego biblioteki cyfrowej, stwierdzając że: 1) cała zawartość książek i artykułów musi być przeszukiwalna, 2) biblioteka cyfrowa powinna mieć cechy biblioteki wirtualnej, tzn. odsyłać przez linki do innych zasobów informacyjnych, 3) w wynikach wyszukiwania powinny się pojawiać wszystkie relewantne materiały, 4) powinny wyświetlać się wszystkie dostępne szczegółowe informacje na temat poszukiwanej pozycji, 5) użytkownicy powinni być w stanie zapisać swoje poprzednie wyszukiwania, 6) użytkownicy powinni mieć dostęp do opinii innych klientów o dziełach udostępnianych w konkretnej kolekcji, 7) wyszukiwarka powinna kierować użytkowników prosto do szukanej publikacji, 8) zasoby należy klasyfikować nie tylko według tematu, ale także według typu dokumentów, 9) mechanizmy systemu powinny wysłać użytkownikom e-mail lub powiadomienia SMS, kiedy zamówione / zapowiadane materiały są już dostępne, 10) wszystkie usługi oferowane przez biblioteki „rzeczywiste” powinny być oferowane także z poziomu biblioteki cyfrowej, 11) w BC powinny być osadzone zakładki „pomoc” na każdym etapie współpracy użytkownika z systemem informacyjno-wyszukiwawczym. Postulaty dotyczące interfejsu / użyteczności brzmiały: 1) BC powinna mieć wyraźnie oznaczone filtry służące do wyszukiwania różnych typów dokumentów: książek, płyt CD/DVD, wideo i czasopism, 2) BC powinna mieć interfejs oparty na komunikacji wizualnej, 3) interfejs BC powinien za-

wierać animacje, które ożywiałyby go, a nawet wprowadzały element humoru, 4) interfejs BC powinien być wyposażony w większą liczbę menu rozwijalnych, 5) interfejs BC powinien mieć więcej ikon i przycisków, 6) wyniki wyszukiwania powinny mieścić się na ekranowej długości strony, bez konieczności przewijania w dół. W odniesieniu do zawartości kolekcji wyrażano następujące oczekiwania: 1) BC powinna umożliwiać użytkownikom czytanie gazet i dostęp do e-książek (także najnowszych), 2) BC powinna zawierać zdjęcia, pliki muzyczne i filmy wideo. Jak wynika z przytoczonych opinii użytkowników, zarówno zawartość, jak i funkcjonalność biblioteki cyfrowej powinny spełniać bardzo wiele, często sprzecznych oczekiwań i potrzeb [Kani-Zabihi, Ghinea, Chen 2006, dok. elektr.].

W tym samym 2006 r. opublikowany został tekst Richarda Butterwortha „Designing Digital Library Resources for Users in Sparse, Unbounded Social Networks”, w którym autor przedstawił po raz kolejny bibliotekę cyfrową przeznaczoną dla środowiska akademickiego określonego jako: 1) dobrze zidentyfikowana grupa użytkowników (można stwierdzić, kto jest, a kto nie jest potencjalnym użytkownikiem systemu), 2) mające ściśle określone potrzeby i zadania, 3) łatwo dostępne – także w celu przeprowadzenia badań na temat jego potrzeb, 4) jednorodne, o bardzo podobnych wymaganiach (można łatwo dobierać próby do badań), 5) dzielące wspólne metody i tryby pracy, studiowania i prowadzenia badań. Cechy te mogą wspomóc twórców systemów w projektowaniu jak najlepszych bibliotek cyfrowych przeznaczonych dla tej wyodrębnionej grupy [Butterworth 2006, dok. elektr.].

Podobny tematycznie jest artykuł Mimi Recker „Perspectives on Teachers as Digital Library Users: Consumers, Contributors, and Designers”. Jego autorka propaguje włączanie do procesu projektowania i wdrażania systemów informacyjnych zarówno uczniów, jak i nauczycieli, wyrażając przekonanie, że przysłuży się to doskonaleniu jakości cyberinfrastruktury edukacji [Recker 2006, dok. elektr.].

Bardziej ogólny jest tekst z tego samego roku pt. „Print vs. electronic resources: A study of user perceptions, preferences, and use”, w którym Ziming Liu przedstawia wyniki badań wskazujące, iż dla studentów i absolwentów studiów najkorzystniejsze jest rozwiązanie hybrydowe: szybki dostęp do źródeł drukowanych i elektronicznych [Liu 2006, dok. elektr.].

Do studentów jako użytkowników odniesiono się także w pracy „Collaborative digital library of historical resources: Evaluation of first users” [Abdullah, Zaindab 2007], w której udokumentowano wyniki badań nad praktykami informacyjnymi studentów historii, wykonujących projekty o tematyce zbieżnej z ich kierunkiem studiów. Zaznaczono, iż biblioteki cyfrowe mogą być integralną częścią działań edukacyjnych oraz podkreślono, że są one uznawane przez użytkowników za funkcjonalne i użyteczne.

Dwa lata później ukazał się inny artykuł o podobnej tematyce, a mianowicie „Managing Digital Libraries in the Light of Staff and Users: An Approach” [Moghaddam 2009, dok. elektr.], w którym położono nacisk na funkcję edukacyjną biblioteki cyfrowej, przyjmując „szkoleniowe” podejście do zarządzania systemem.

W tym samym roku w artykule „A user-centered and evidence-based approach for digital library projects” przedstawiono wyniki testów funkcjonalnych wykonanych w ramach badania metodą *case study* w latach 2003-2006 przez profesorów i studentów California Polytechnic State University. Zastosowana technika badawcza umożliwiła iteracyjność ocen oraz ciągłe doskonalenie koncepcji projektowych i funkcjonalności produktu. Wskazano także na tzw. uczenie się przez działanie, a więc na wartość dydaktyczną biblioteki cyfrowej [Somerville, Brar 2009, dok. elektr.].

Wymieniony wcześniej Ziming Liu wraz z Lili Luo w publikacji „A Comparative Study of Digital Library Use: Factors, Perceived Influences, and Satisfaction” z 2011 r. powtórnie powrócili do tematu wykorzystania bibliotek cyfrowych przez członków społeczności akademickiej, przedstawiając tym razem wyniki swoich badań na grupie studentów z Chin, w której mierzyli poziom zadowolenia oraz częstotliwość korzystania z bibliotek cyfrowych (najwyższą satysfakcją odnotowali u użytkowników odwiedzających BC co najmniej raz w tygodniu) [Liu, Luo 2011, dok. elektr.].

Kolejny konkretny przykład biblioteki edukacyjnej opisany został w artykule „A simulation study for a number of simultaneous users of an institutional knowledge repository system” [Doctor 2011, dok. elektr.]. Jest to tekst o charakterze technicznym, opisujący pilotażowy projekt instytucjonalnego repozytorium w szkole biznesu w Indiach, a w szczególności wyniki symulacji przeprowadzonych w celu określenia liczby jednoczesnych użytkowników, którzy mogą korzystać z zawartości repozytorium instytucjonalnego, z założeniem że dane te posłużą do ustalenia minimalnej konfiguracji serwera.

Tematycznie podobna publikacja Krystyny K. Matusiak z 2012 r. pt. „Perceptions of usability and usefulness of digital libraries” odnosi się do badań przeprowadzonych wśród studentów i wykładowców, którzy za podstawowy składnik kolekcji bibliotek cyfrowych uznają dokumenty tekstowe. Źródła multimedialne, wizualne, dźwiękowe nie są przez nich często kojarzone z bibliotekami cyfrowymi. Takie ograniczone wykorzystanie zasobów BC ma, zdaniem autorki, dwie przyczyny: 1) systemy biblioteczne nie są postrzegane jako przyjazne, co z kolei zniechęca potencjalnych użytkowników, 2) biblioteki akademickie są postrzegane jako miejsca przechowywania przede wszystkim zasobów tekstowych. Ciekawa jest autorska propozycja modelu biblioteki cyfrowej, stanowiąca kolejny dowód na to, jak bardzo różnorodne mogą być conceptualizacje systemu BC spo-

tykane w piśmiennictwie naukowym. Według modelu Matusiak, odtworzonego na rysunku 7, BC jest „emanacją” biblioteki tradycyjnej, rozszerzeniem jej funkcjonalności i przeniesieniem ich w sferę wirtualną.



Rysunek 7. Budowa cyfrowej biblioteki akademickiej  
Źródło: [Matusiak 2012]. Rysunek przetworzony przez autorkę

W publikacji „Finding Problems: When Digital Library Users Act as Usability Evaluators” [Khoo, Kusunoki, MacDonald 2012] opisany został eksperyment edukacyjny przeprowadzony w Drexel University, w którym zespoły studentów uczestniczących w zajęciach z informatyki (na poziomie magisterskim – zajęcia z przedmiotu *HCI, human-computer interaction* / interakcja człowiek – komputer) otrzymały zadanie zaprojektowania i przeprowadzenia heurystycznej oceny interfejsu Internet Public Library. W wyniku tych działań okazało się, że z punktu widzenia badanych studentów najważniejszym elementem wpływającym na poziom oceny systemu były trudności z wyszukaniem konkretnych zasobów w bibliotece. Wskazania te porównano z wynikami równoległej oceny IPL dokonanej przez grupę doktorantów, którzy stwierdzili, iż najważniejszym elementem jest interfejs. Jak widać, przyjęcie – odpowiednio – roli ewaluatorów i twórców zmienia percepcję oraz wpływa na globalne rozumienie bibliotek cyfrowych i ocenę istotności ich poszczególnych elementów.

W ostatnim już omawianym tu tekście Yalan Yan i Xianjin Zha zajęli się problemem jakości bibliotek akademickich. W publikacji „Comparison between user affinity with digital libraries and virtual communities” przedstawili dyskusję na temat wartości i wiarygodności informacji tworzonych przez ekspertów m.in.

w zorganizowanych przez uniwersytety bibliotekach cyfrowych, w porównaniu do informacji wygenerowanych przez internautów skupionych w wirtualnych społecznościach. W konkluzjach podkreślili, że czynnikiem najsilniej motywującym użytkowników do zamieszczania danych w BC lub w wirtualnych społecznościach jest poczucie przynależności [Yan, Zha 2014].

## 2. Badania użytkowników

Do grupy publikacji dotyczących badań użytkowników zaliczone zostały teksty, w których autorzy opisali próby określenia profilu użytkowników bibliotek cyfrowych i cech formowanych przez nich zbiorowości.

Takie właśnie założenie zrealizowali Giuseppe Amato i Umberto Straccia w ogłoszonym w 1999 r. artykule „User Profile Modeling and Applications to Digital Libraries”, w którym przedstawili sposoby modelowania profili użytkowników na podstawie danych pozyskiwanych według schematu:

profil użytkownika:

kategoria danych osobowych

kategorie zbierania danych:

kategoria zawartość dokumentu

kategoria struktura dokumentu

kategoria źródła dokumentu

kategorie dostarczania danych:

kategoria dostarczenie znaczenia

kategoria czas realizacji

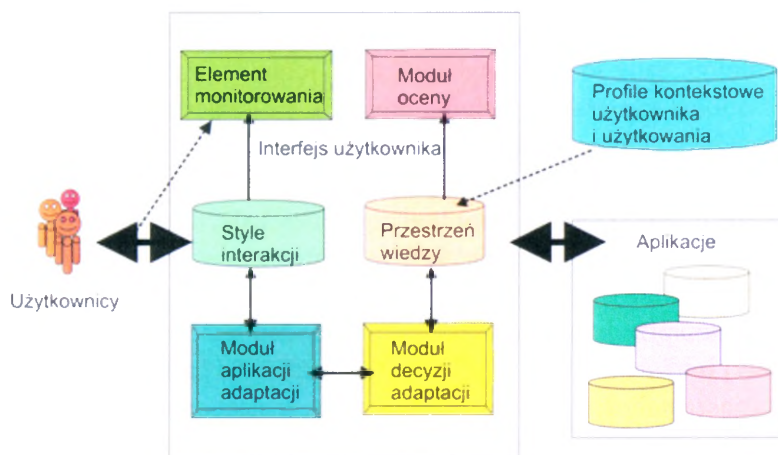
kategoria działania na danych

kategoria bezpieczeństwo danych.

Amato i Straccia stwierdzili, że opisując użytkowników przez pryzmat ich potrzeb skorelowanych z relewantnymi źródłami, twórcy bibliotek cyfrowych mogą dzielić odbiorców na kategorie, a następnie specjalnie dla nich przygotować odpowiednie kolekcje, opcje wyszukiwania itp. [Amato, Straccia 1999, dok. elektr.].

Specyficzny krąg odbiorców opisany został w pracy z 2000 r. „User interaction in digital libraries: coping with diversity through adaptation” [Stephanidis et al. 2000]. Omówiono w niej Projekt AVANTI, a zwłaszcza zmiany w interakcjach człowiek – komputer zachodzące w wyniku przekształceń oprawy estetycznej i architektury biblioteki cyfrowej. Czynniki poprawiające jakość tych interakcji badano na przykładzie trzech grup użytkowników końcowych: osób widzących i niewidzących oraz dzieci, dochodząc do wniosku, że w każdym przypadku klu-

czowe jest odpowiednie dobranie wizualnych cech interfejsu. Podkreślono, że adaptacje te muszą być jednocześnie zorientowane na kontekst. Odrębne rozważania poświęcono także dwóm innym grupom interesariuszy BC, a mianowicie deweloperom i twórcom. Przynależne im kompetencje zostały określone w następujący sposób: deweloperzy opracowują zunifikowaną specyfikację BC, a po stronie projektantów jest przewidywanie działań oparte na doświadczeniu oraz znajomości struktur, opracowywanie scenariuszy i przewodników dla użytkownika. Ciekawa jest w tej pracy architektura interfejsu (zobrazowana na rysunku 8), pokazująca że cały system BC traktowany jest jak system do zarządzania informacją i wspomagania. Składają się na niego: moduł podejmowania decyzji, serwer informacyjny użytkownika, serwer parametrów kontekstowych oraz komponenty interaktywne. Autorzy opisali też możliwości stworzenia profili użytkowników.



Rysunek 8. Architektura interfejsu

Źródło: [Stephanidis et al. 2000]. Rysunek przetworzony przez autorkę

Także w roku 2000 w artykule „Socially grounded user studies in digital library development” [Bishop et al. 2000] przedstawiono kolejny projekt: Afya Project, który miał przyczynić się do przezwyciężenia „przepaści cyfrowej” (ang. *digital gap*) istniejącej w lokalnych społecznościach w USA przez poszerzenie zakresu działań tamtejszych bibliotek publicznych. Zniwelowaniu barier w dostępie do informacji i usług zdrowotnych, których doświadczają afroamerykańskie kobiety, służyć miała społecznościowa sieć, w ramach której kobiety uczestniczyły w gromadzeniu danych (dla biblioteki wirtualnej) i tworzeniu analiz związanych z rozwijaniem, wdrażaniem i oceną zasobów internetowych. Jest to przykład za-

angażowania się użytkowników końcowych w tworzenie portalu, który pomaga im w codziennym życiu. Temat ten został też poruszony w 2002 r. w publikacji „Scenarios in the Afya project as a participatory action research (PAR) tool for studying information seeking and use across the *digital divide*” [Bharat et al. 2002], która również zawierała opis wyników projektu Afya. Jej autorzy podkreślili, że w ramach tego przedsięwzięcia stworzono różnorodne scenariusze promowania sprawiedliwości społecznej przez inicjowanie i podtrzymywanie partycypacyjnych działań równościowych, skupionych wokół rozwoju technologicznego i wykorzystania usług informacyjnych dotyczących zdrowia. W scenariuszach tych zapisane są konkretne strategie działań dla wzmocnienia pozycji Afroamerykanek i ich wkładu w budowanie społecznych technologii cyfrowych. Projekt przerodził się niemal w ruch społeczny, choć pierwotnie był próbą przezwyciężenia barier społecznych przez zgromadzenie danych i zbudowanie biblioteki wirtualnej.

Do problematyki specjalnych grup użytkowników powrócili też w 2002 r. Hayley Dunlop, Sally Jo Cunningham i Matt Jones w pracy „A digital library of conversational expressions: helping profoundly disabled users communicate”, w której omówili cyfrową bibliotekę wyrażen niezbędnych do prowadzenia rozmowy w języku naturalnym, wspomagającą odbiorców z głęboką niepełnosprawnością intelektualną [Dunlop, Cunningham, Jones 2002, dok. elektr.].

Bardziej ogólną klasyfikację klientów BC zaprezentowali rok później autorzy studium „Users and Uses of Online Digital Libraries in France” [Assadi et al. 2003], w którym przedstawione zostały wyniki badania zachowań użytkowników internetowej biblioteki cyfrowej, przeprowadzonych za pomocą trzech technik: kwestionariusza *online*, zliczania statystyk ruchu w sieci oraz wywiadu. W wyniku kwantyfikacji danych wyróżniono dwie główne grupy użytkowników: badaczy i bibliofilów. Wskazano, że różnią się oni zarówno od przeciętnych internautów, jak również od czytelników bibliotek klasycznych / tradycyjnych. Podkreślono, że biblioteki cyfrowe, dalekie od bycia tylko cyfrowymi wersjami zbiorów bibliotecznych, są atrakcyjne dla nowego typu publiczności, bowiem wnoszą nowy, unikalny i oryginalny sposób czytania i rozumienia tekstu, stanowią swoistą „nową arenę” czytania i komentowania prac.

W 2003 roku w tekście „Three looks at users: a comparison of methods for studying digital library use” Mark Notess odniósł się do konkretnego projektu BC, przedstawiając różne metody badań przydatne w pozyskiwaniu danych, na podstawie których można odtwarzać rzeczywiste praktyki informacyjne użytkowników bibliotek cyfrowych. Autor analizował zachowania użytkowników dwóch wersji muzycznej biblioteki *Variations* i *Variations 2*, przygotowanej przez Indiana University. Biblioteka ta (w obu swych wariantach, które były wynikiem



stopniowego doskonalenia systemu i jego zawartości) zawiera zarówno zapisy nutowe, jak i dźwiękowe utworów muzycznych. W obu fazach badania zostały wykorzystane kwestionariusze oceny biblioteki. Przeanalizowano także dzienniki sesji (metodami ilościowymi i jakościowymi). Trzecią metodą była obserwacja – obserwowano sposoby tworzenia zapytań kontekstowych, korzystania z informacji przez studentów i formułowania poleceń wyszukiwawczych wydawanych głosem. Najciekawszym aspektem tej pracy jest porównanie zastosowanych metod badawczych, które zostało zaprezentowane w poniższej tabeli, oraz podkreślenie, że najbardziej efektywne jest stosowanie metod mieszanych.

Tabela 1. Metody badania zachowań użytkowników biblioteki *Variations* i *Variations 2*

	<b>Kwestionariusz zadowolenia</b>	<b>Sesja logowania (aktywność)</b>	<b>Zapytania kontekstowe</b>
Ekspertyza	badanie projektu	techniczna (oprogramowanie, skrypty)	obserwacja, wywiady, modelowanie pracy
Czas utworzenia	umiarkowany	krótki	umiarkowany
Czas przeprowadzenia	żaden ( <i>online</i> ), do umiarkowanego (rekrutacja do próby i przeprowadzanie badania w wersji papierowej)	żaden	długi
Czas analizy	krótki	umiarkowany (metryki generacji) do bardzo długi (manualna analiza wszystkich plików)	długi
Korzyści	głównie podsumowujące; można odkryć tematy do dalszych badań	podsumowujące (metryki generacji) i kształtujące / twórcze (analiza manualna)	głównie kształtujące / twórcze

Źródło: [Notess 2003, dok. elektr.]

Kolejne analizy zostały przedstawione dwa lata później w pracy „Designing effective tasks for digital library user tests: lessons learned” [Notess, Kouper, Swan 2005], w której omówiono: iteracyjne automatyczne tworzenie zadań, projektowanie działań potencjalnych użytkowników, rekrutację autentycznych klientów oraz zaproponowano sposoby tworzenia realnych scenariuszy testowych.

W innej pracy z tego roku, „Users’ views about the usability of digital libraries” [Koohang, Ondracek 2005], zaprezentowano wyniki badań mających na celu określenie charakterystyki demograficznej oraz zidentyfikowanie tych cech użytkowników, które w największym stopniu wpływają na formułowane przez nich oceny użyteczności bibliotek cyfrowych – w chwili badania i w przyszłości, w kontekście długoterminowym. Badaniu poddano 107 studentów interdyscy-

plinarnych studiów licencjackich uniwersytetu stanowego Midwest, USA, a jako zmienne niezależne przyjęto: wiek, płeć, wcześniejsze doświadczenia z internetem oraz biegłość w korzystaniu z bibliotek cyfrowych. Natomiast poglądy użytkowników na temat przydatności bibliotek cyfrowych były traktowane jako zmienne zależne. W wynikach uzyskano istotne różnice dla zmiennych „wcześniejsze doświadczenia z internetem” oraz „poziom biegłości / znajomości biblioteki cyfrowej”, co oznacza, iż osoby z większą wiedzą technologiczną traktują biblioteki cyfrowe jako bardziej użyteczne, gdyż lepiej potrafią z nich korzystać. Dla pozostałych cech nie ujawniono statystycznie istotnych korelacji.

O specjalnym typie użytkownika – dzieciach oraz o tym, jak budować bibliotekę cyfrową wspólnie z dziećmi i z badaczami z różnych dyscyplin: informatologii, informatyki, edukacji, sztuki i psychologii – czyli o badaniach przeprowadzonych na University of Maryland, także w odniesieniu do International Children’s Digital Library – napisała Allison Druin w artykule „What Children Can Teach Us: Developing Digital Libraries for Children with Children” [Druin 2005]. Do biblioteki tej odniesiono się też w pracy „The International Children’s Digital Library: a Case Study in Designing for a Multilingual, Multicultural, Multigenerational Audience” [Hutchinson et al. 2005].

W kolejnej publikacji, „A library or just another information resource? A case study of users’ mental models of traditional and digital libraries” [Makri et al. 2007], przedstawiono studium dotyczące modeli mentalnych użytkowników bibliotek tradycyjnych i cyfrowych, które zostało opracowane na podstawie obserwacji i wywiadów przeprowadzonych z ośmioma osobami. Stwierdzono, że nikła znajomość ograniczeń dostępu prowadzi do zachowań: „awersja do podejmowania ryzyka”, natomiast nikła znajomość algorytmów wyszukiwania i znaczenia rangowania wyników do zachowań: „próby i błędy” („chybił-trafił”). Podkreślono przy tym, iż dla klientów ważna jest informacja zwrotna (ang. *feedback*) od twórców systemu, gdyż wspomaga ona konstruowanie przydatnych modeli mentalnych. Zaznaczono też, iż większość uczestników badania nie widziała wyraźnych różnic między poszczególnymi rodzajami narzędzi dostępu do zasobów cyfrowych i metodami korzystania z nich (przeglądanie katalogu biblioteki elektronicznej, usługi abstraktowania, wyszukiwanie w bibliotekach cyfrowych, budowanie kwerend w wyszukiwarkach internetowych).

Kolejna eksploracja empiryczna dotycząca użytkowników BC, omówiona w pracy „A study of the information needs of the users of a folk music library and the implications for the design of a digital library system” [Inskip, Butterworth, MacFarlane 2008], była jakościowym badaniem potrzeb informacyjnych użytkowników Vaughan Williams Memorial Library, małej specjalistycznej biblioteki muzyki ludowej z północnego Londynu. Podczas analizy jego wyników wyod-

rębniono cztery najważniejsze grupy potencjalnych i rzeczywistych użytkowników tej biblioteki, tzn. performerów, badaczy, zawodowców i entuzjastów. Na tej podstawie opracowano zbiór ogólnych wymagań projektowych dla muzycznej biblioteki cyfrowej.

Odniesienie do konkretnych bibliotek pojawiło się również w tym samym roku (2008) w artykule Anny Marii Tammaro „User perceptions of digital libraries: a case study in Italy”. Autorka przedstawiła w nim wyniki badań ankietowych i wywiadów przeprowadzonych wśród użytkowników włoskich bibliotek cyfrowych oraz kierowników instytucji kulturalnych uczestniczących w ich tworzeniu. Badania, przeprowadzone metodą *case study*, wykazały dość różnorodną (ale pozytywną) ocenę BC oraz tendencję do korzystania z usług więcej niż jednej instytucji. Tammaro zauważyła także, że użytkownicy często nie wiedzą, jak korzystać z biblioteki i nie zdają sobie sprawy ze wszystkich oferowanych przez nią usług. Nie skategoryzowała jednak klientów włoskich bibliotek cyfrowych [Tammaro 2008].

W następnym tekście, także z 2008 r., pt. „Users’ personality and perceived ease of use of digital libraries: the case for resistance to change” omówiono zależność między poziomem akceptowania systemu biblioteki cyfrowej (w tym przypadku: uniwersyteckiej biblioteki cyfrowej) przez użytkowników a sposobem, w jaki postrzegają oni BC i oceniają łatwość obsługi jej interfejsu. Postawy badanych oceniano za pomocą parametru osobowościowego określającego poziom akceptacji / oporu wobec zmian (ang. *resistance to change* – RTC). Po przebadaniu – techniką ankiety – 170 użytkowników wykazano, iż jest to jedna z istotnych cech wpływających na poziom oceny jakości BC [Nov, Ye 2008].

Dwa lata później ukazały się dwie publikacje, które zostały zaliczone do grupy tematycznej „rekonesans potrzeb i zachowań użytkowników”. Pierwszą z nich jest tekst Chiary Consonni (2010) „Non-users’ evaluation of digital libraries: A survey at the Università degli studi di Milano”, jeden z bardzo niewielu dotyczących nie-użytkowników. W artykule tym przedstawione zostało bardzo ciekawe, a mało zbadane zagadnienie: dlaczego nie korzysta się z bibliotek cyfrowych. Wyniki badań ankietowych nie-użytkowników usług bibliotek cyfrowych w Università degli Studi di Milano odnosiły się do różnych grup: studentów, nauczycieli oraz pracowników technicznych i administracyjnych uczelni, których łączyła jedna cecha: skorzystali z usług biblioteki, a następnie postanowili nie używać jej więcej. Wyniki wykazały, że nie-użytkownicy, pomimo wcześniejszego ograniczonego korzystania z biblioteki cyfrowej, uważali że byli zaznajomieni z serwisem oraz że był on zadowalający. Nie odwiedzali go więcej z różnych powodów, głównie dlatego, że „już go nie potrzebowali” lub też w danym czasie wykorzystywali inne serwisy. Pewnymi niedociągnięciami systemu biblioteki cy-

frowej, które zidentyfikowała Consonni, były niezbyt łatwa obsługa i ograniczona dostępność materiałów w niektórych dyscyplinach. Badania wykazały również, że serwis BC był wykorzystywany z różnych powodów, w tym: do przygotowania pracy dyplomowej, na potrzeby badań własnych, do zapoznawania się na bieżąco z wynikami najnowszych badań naukowych w jakiejś dziedzinie, ale także dla zabezpieczenia tzw. interesów osobistych [Consonni 2010, dok. elektr.].

Druga z publikacji to materiały z jednodniowej konferencji, która odbyła się w 2009 r. w Università degli Studi di Milano pod hasłem „Digital Library Futures: User Perspectives and Institutional Strategies” [Verheul, Tammaro, Witt eds. 2010]. Najciekawsze referaty związane z problematyką użytkowników zaprezentowali: David Nicholas, Daniel Teruggi, Elke Griefender, Einar Röttingen oraz Susan Hazan. Nicholas podsumował wyniki swoich siedmioletnich analiz logów z różnych platform naukowych, takich jak Elsevier ScienceDirect, na podstawie których rekonstruował nawyki użytkowników. Stwierdził, że typowi użytkownicy sieci „skaczą z miejsca na miejsce”, gromadzą duże ilości informacji, ale mało szczegółowych. Tak też zachowują się klienci bibliotek cyfrowych. „Wyszukiwanie informacji (ang. *information seeking*) jest frenetyczne, energiczne / skaczące (ang. *bouncing*), nawigacyjne, sprawdzające i przeglądające” [Nicholas 2009]. Teruggi omówił rezultaty oceny biblioteki cyfrowej Europeana, opartej na wykorzystaniu ankiet internetowych, analizie logów i treści opinii spontanicznie nadsyłanych przez użytkowników (ang. *feedback inbox*). Podkreślił, że klienci pragną uczestniczyć w tworzeniu treści Europeany i wchodzić z nią w żywe interakcje [Teruggi 2009]. Griefender, omawiając typowe trudności związane z badaniem użytkowników *online*, stwierdziła że naukowcy często skupiają się na użyciu metod ilościowych, których wyniki nie zawsze da się dobrze zinterpretować, zamiast zająć się metodami jakościowymi [Griefender 2009]. Röttingen przedstawił *case study* – wykorzystanie *Edvard Crieg Collection online* dla prześledzenia ewolucji zapisu muzycznego (Crieg Ballade op. 24) i podjęcia kluczowych decyzji dotyczących sposobu jego odtworzenia. Podkreślił wartość bibliotek muzycznych *online* dla artystów muzyków i naukowców oraz potrzebę współpracy z użytkownikami, aby poprawić możliwości wyszukiwawcze tych systemów [Röttingen 2009]. Hazan porównała biblioteki *online*, takie jak Project Gutenberg, World Digital Library i Internet Archive, zorientowane na dodawanie nowych treści, z serwisami, które zmieniają sposób interakcji użytkowników z ich kolekcjami: Google Books, Flickr i Facebook. Zasugerowała, że jeśli tradycyjne instytucje kultury, takie jak biblioteki, chcą utrzymać swoją publiczność, to muszą zacząć istnieć w mediach społecznościowych [Hazan 2009].

Rok 2012 przyniósł cykl artykułów dotyczących klientów bibliotek cyfrowych. Pierwszy z nich, „A review of quality evaluation of digital libraries based on

users' perceptions" [Heradio et al. 2012], stanowi przegląd piśmiennictwa naukowego dotyczącego oceny jakości bibliotek cyfrowych, dokonywanej na podstawie badania ich recepcji przez użytkowników. Główną wartością tego tekstu jest usystematyzowanie aktualnego stanu wiedzy o ewaluacji BC za pomocą metod jakościowych. W pracy „Discovering Digital Library User Behavior with Google Analytics” Kirk Hess opisał zaawansowane funkcje śledzenia zapytań Google Analytics, które można wykorzystywać do badań zachowania użytkowników [Hess 2012]. Natomiast w artykule „What Would ‚Google’ Do? Users’ Mental Models of a Digital Library Search Engine” Michael Khoo i Catherine Hall przedstawili modele mentalne użytkowników bibliotek cyfrowych i ich analizę w odniesieniu do użyteczności. Badanie treści tych modeli mentalnych wykazało istnienie generycznego modelu wyszukiwarki, będącego wynikiem kompilacji cech wielu wyszukiwarek internetowych traktowanych jako wzorce idealne, przy czym stwierdzono, że te wyobrażeniowe odniesienia mają charakter idealistyczny, a nie realistyczny, w związku z czym trudno byłoby na nich bazować przy tworzeniu interfejsów bibliotek cyfrowych [Khoo, Hall 2012].

Rok 2012 przyniósł także jedną z nielicznych prac dotyczących użytkowników polskich bibliotek cyfrowych. W tekście „Key users of Polish digital libraries” Mirosław Górny i Jolanta Mazurek omówili wyniki badań przeprowadzonych za pomocą kwestionariusza ankietowego, który zawierał 19 pytań dotyczących sposobów i powodów korzystania z biblioteki oraz preferencji w zakresie typów publikacji. W trakcie analiz autorzy skorzystali również z danych statystycznych zgromadzonych na serwerze biblioteki. Scharakteryzowali dominujące grupy użytkowników bibliotek cyfrowych, za główne kryterium typologiczne przyjmując ich zainteresowania, jakkolwiek opisali ich także pod względem wieku, płci, poziomu wykształcenia oraz miejsca zamieszkania (wielkość miejscowości). Ustalili, iż około 60% użytkowników to historycy amatorzy zainteresowani dziejami poszczególnych miejscowości lub genealogicy zainteresowani losami swojej rodziny. Pozostałe 40% to głównie pracownicy naukowci i studenci. Liczba stałych czytelników, tych, którzy korzystają z biblioteki co najmniej kilka razy w roku, wynosi około 30 000. Autorzy wskazali także czynniki, które mogą zwiększyć liczbę użytkowników odwiedzających biblioteki cyfrowe: rozbudowa infrastruktury informatycznej na obszarach wiejskich oraz szkolenia dla osób w wieku powyżej 70 lat. Grupą, która zdaje się rosnąć, są także studenci, zwłaszcza że są oni zainteresowani powszechnym dostępem do podręczników. Zdziwiałe było natomiast małe zainteresowanie bibliotekami cyfrowymi ze strony nauczycieli i uczniów, szczególnie jeśli zważyć, że zbiory BC zawierają większość lektur, bogaty zasób encyklopedii, dzieł historycznych itp. [Górny, Mazurek 2012].

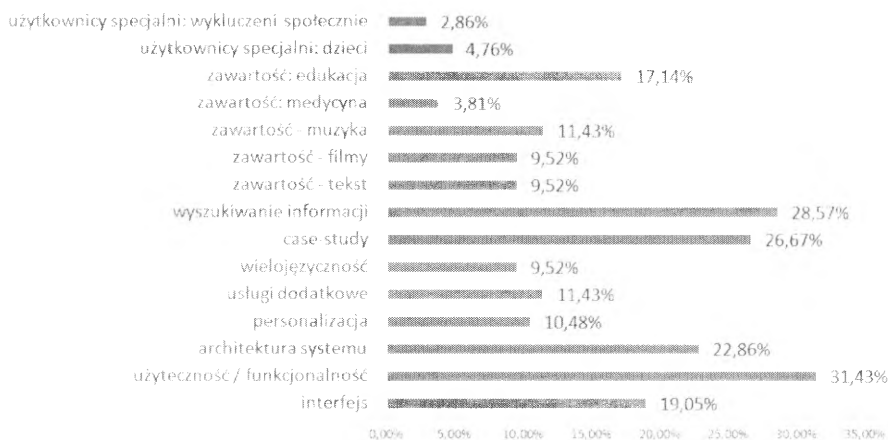
Następne dwa artykuły, już z 2013 r., dotyczą swoistych grup odbiorców. W pierwszym z nich, „Access to Digital Libraries for Disadvantaged Users” [Higgins, King 2013], omówiono znaczenie bibliotek cyfrowych jako spadkobierców misji bibliotek tradycyjnych, które pozwalają na eliminację fizycznych i czasowych ograniczeń w dostępie do informacji, a zarazem umożliwiają nawiązywanie kontaktów, ważnych zwłaszcza dla osób znajdujących się w niekorzystnej sytuacji życiowej (niezatrudnionych, starszych i chorych itp.).

Druga i ostatnia już z omawianych publikacji dotyczy bibliotek dla dzieci. Celem autorów tej pracy, pt. „Children Digital Libraries and User Interfaces: Proposing a Set of Criteria” [Sedighi, Gilvari, Nooshinfard 2013], było zbadanie interfejsów najważniejszych światowych bibliotek cyfrowych dla dzieci, porównanie ich oceny z opiniami ekspertów oraz zaproponowanie jednolitego autorytatywnego zestawu kryteriów oceny jakości dla irańskich bibliotek cyfrowych dla dzieci, wśród których wskazano: możliwość pobierania zasobów, możliwość personalizowania, opcje wyszukiwawcze dostosowane do wieku dziecka (wygląd okładki, treść) oraz jasny, „radosny” interfejs.

## Wnioski

W trakcie badań przeanalizowano w sumie 105 artykułów z lat 1993-2014. Większość z nich dotyczyła funkcjonalności i użyteczności (31,43%) systemu widzianego od strony użytkownika lub też zmian, jakie należy wprowadzić w architekturze i w działaniu biblioteki cyfrowej (22,86%), tak aby klienci byli bardziej zadowoleni z tych systemów informacyjno-wyszukiwawczych. Duża grupa tekstów dotyczyła konkretnych projektów wdrożeniowych o różnym charakterze oraz pojedynczych bibliotek cyfrowych, w odniesieniu do których omawiane były np. wybrane problemy programistyczne związane z doskonaleniem prezentacji danych, organizacją punktów dostępu do nich, możliwością pobierania informacji itp. (26,67%). Wiele publikacji odnosiło się wybiórczo do problemu wyszukiwania informacji (28,57%) i takiej budowy interfejsu, która w sposób optymalny i najbardziej intuicyjny wspomagałaby interakcje człowiek – komputer (19,05%). Autorzy licznych opracowań skupili się wyłącznie na konkretnych przykładach, prezentując swoiste *case-study* biblioteki cyfrowej lub wirtualnej. Jak zaznaczono bowiem we wstępie, wielu autorów pisząc o bibliotekach cyfrowych, miało na myśli także repozytoria materiałów np. dla studentów czy biblioteki wirtualne składające się głównie z odnośników do innych źródeł i systemów informacyjno-wyszukiwawczych dostępnych w sieci. W niektórych publikacjach podjęto temat dodatkowych aplikacji wspomagających działanie biblioteki cy-

frowej (11,43%) bądź jej personalizację (10,48%). W związku z szerokim dostępem do internetu ważkim tematem była także wielojęzyczność serwisu BC (9,52%). Dużą grupę stanowiły ponadto teksty poświęcone zawartości bibliotek cyfrowych: edukacyjnej (17,14%), muzycznej (11,43%), filmowej i tekstowej (po 9,52%) oraz serwisom specjalnym, w szczególności medycznemu (3,81%). Wyniki badań samych użytkowników, a zwłaszcza ich specjalnych grup, zawarte były w 4,76% artykułów; tematem 2,86% prac były potrzeby i zachowania informacyjne osób wykluczonych społecznie. Tylko jeden tekst był poświęcony użytkownikom polskim. Tylko raz w badanym zasobie piśmiennictwa zadano pytanie o to, dlaczego potencjalni użytkownicy, a w istocie nie-użytkownicy (ang. *non-users*), nie korzystają z bibliotek cyfrowych. W jednej książce i w jednym artykule omówiono szczegółowo metody badawcze, które można stosować w celu rozeznania zarówno docelowej grupy odbiorców, jak i twórców lub administratorów tych specyficznych baz danych. Dokładny procentowy rozkład tematów w rozpatrywanym zbiorze piśmiennictwa został przedstawiony na wykresie 1.



Wykres 1. Rozkład tematyczny artykułów  
Źródło: opracowanie własne

Trudno też byłoby wskazać jedną grupę autorów zajmujących się problematyką użytkowników bibliotek cyfrowych lub wybrać wiodące ośrodki badawcze. Wynika to z częstego publikowania rezultatów badań zogniskowanych na konkretnych projektach i problemach informatycznych. Wśród najczęściej pojawiających się nazwisk autorów wymienić należy: Gobindę Chowdhury'ego oraz współpracujące ze sobą zespoły: Milena Dobreva, Andy O'Dwyer i Perluigi Fe-

liciaty; Enrique Frías-Martínez, Sherry Y. Chen, Robert D. Macredie i Liu Xiaohui; James Y.L. Thong, Weiyin Hong i Kar Yan Tam; Xiangmin Li, Yuelin Zhang, Jingjing Liu i Ying Zhang.

Ostateczną konkluzją tego artykułu może być stwierdzenie, iż o ile na świecie często podejmuje się badania użytkowników ogólnych systemów informacyjnych, o tyle analizy użytkowników bibliotek cyfrowych pojawiają się w piśmiennictwie z zakresu informatologii zdecydowanie rzadziej. Prócz tego w przypadku części rozpatrywanych publikacji głównym celem była analiza systemu i jego wartości, a nie zbadanie samych klientów czy twórców bibliotek. Do badawczego zagospodarowania pozostaje wiele tematów: szczegółowe potrzeby informacyjne użytkowników, specyficzne wymagania i sposoby docierania do informacji np. multimedialnej, percepcja i akceptacja bibliotek cyfrowych, kształcenie użytkowników, personalizacja systemu, zaufanie i wiarygodność danych, prywatność, zabezpieczenie praw indywidualnych klientów. Na zbadanie czekają też ważne aspekty socjologiczne (np. tworzenie się społeczności rozbudowujących „swoje” biblioteki cyfrowe) czy kulturowe (przenikanie kultur a problem izolacji, włączanie klientów w tworzenie misji serwisów itp.). Przed informatologami stoi więc zadanie podjęcia tych zagadnień, aby dostęp do danych i informacji, które przez użytkowników zostaną przetworzone w wiedzę, był jakościowo jak najlepszy.

## BIBLIOGRAFIA

- Abdullah, Amirali; Zainab, A.N. (2007). *Collaborative digital library of historical resources. Evaluation of first users* [dok. elektr.]. *Malaysian Journal of Library & Information Science* Vol. 12, No. 2, p. 99-122. <http://arxiv.org/ftp/arxiv/papers/1301/1301.5398.pdf> [odczyt: 5.07.2014].
- Adams, Anne; Blandford, Ann (2000). *Digital libraries' support for the user's 'Information Journey'* [dok. elektr.]. [http://www.york.ac.uk/res/e-society/projects/8/aaab\]CDL-05preprint.pdf](http://www.york.ac.uk/res/e-society/projects/8/aaab]CDL-05preprint.pdf) [odczyt: 5.07.2014].
- Agosti, Maristella et al. (2010). *Understanding user requirements and preferences for a digital library Web portal*. *International Journal on Digital Libraries* Vol. 11, p. 225-238.
- Albertson, Dan E. (2010). *Influences of users' familiarity with visual search topics on interactive video digital libraries*. *Journal of the American Society for Information Science and Technology* Vol. 61, p. 2490-2502.
- Alijani, Rahim; Nikkar, Maliheh (2008). *Evaluation & comparison of free and commercial user interfaces of digital libraries* [dok. elektr.]. *International Journal of Information Science & Technology* Vol. 6. <http://ijism.ricest.ac.ir/ojs/index.php/ijism/article/view/83/68> [odczyt: 5.07.2014].



- Amato, Giuseppe; Straccia, Umberto (1999). *User profile modeling and applications to digital libraries* [dok. elektr.]. <https://who.rocq.inria.fr/Anne-Marie.Vercoustre/EuroDL99/Proceedings/Straccia/UserProfile.ps> [odczyt: 5.07.2014].
- Assadi, Houssein et al. (2003). *Users and uses of online digital libraries in France* [dok. elektr.]. [http://link.springer.com/chapter/10.1007/978-3-540-45175-4\\_1#page-1](http://link.springer.com/chapter/10.1007/978-3-540-45175-4_1#page-1) [odczyt: 5.07.2014].
- Bainbridge, David et al. (2010). *A user-centered design of a personal digital library for music exploration* [dok. elektr.]. <http://dl.acm.org/citation.cfm?doid=1816123.1816145> [odczyt: 5.07.2014].
- Balatsoukas, Panos (2012). *An eye-tracking approach to the evaluation of digital libraries*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 95-104.
- Baldonado, Michelle Q. Wang (2000). *A user-centered interface for information exploration in a heterogeneous digital library* [dok. elektr.]. *Journal of the American Society for Information Science* Vol. 51, p. 297-311. [https://intranet.library.gatech.edu/info/committees/gtel\\_content/docs/opac\\_user\\_centered.pdf](https://intranet.library.gatech.edu/info/committees/gtel_content/docs/opac_user_centered.pdf) [odczyt: 5.07.2014].
- Bangert-Drowns, Robert L. (2004). *Literature review* [dok. elektr.]. In: *Encyclopedia of evaluation* [dok. elektr.]. [http://www.sage-ereference.com/evaluation/Article\\_n319.html](http://www.sage-ereference.com/evaluation/Article_n319.html) [dostęp: 15.05.2010].
- Bawden, David; Vilar, Polona (2006). *Digital libraries. To meet or manage user expectations* [dok. elektr.]. *Aslib Proceedings* Vol. 58, p. 346-354. <http://www.emeraldinsight.com/journals.htm?articleid=1563468> [odczyt: 5.07.2014].
- Bharat, Mehra et al. (2002). *Scenarios in the Afya project as a participatory action research (PAR) tool for studying information seeking and use across the "digital divide"*. *Journal of the American Society for Information Science and Technology* Vol. 53, p. 1259-1266.
- Biblioteka cyfrowa* [dok. elektr.]. W: *Wikipedia*. [http://pl.wikipedia.org/wiki/Biblioteka\\_cyfrowa](http://pl.wikipedia.org/wiki/Biblioteka_cyfrowa) [dostęp: 15.05.2010].
- Bishop, Ann et al. (2000). *Socially grounded user studies in digital library development* [dok. elektr.]. *First Monday* Vol. 5. <http://firstmonday.org/ojs/index.php/fm/article/view/760/669> [odczyt: 5.07.2014].
- Bollen, Johan; Luce, Rick (2002). *Evaluation of digital library impact and user communities by analysis of usage patterns* [dok. elektr.]. *D-Lib Magazine*. <http://www.dlib.org/dlib/june02/bollen/06bollen.html> [odczyt: 5.07.2014].
- Butcher, Kirsten R.; Bhushan, Sonal (2005). *Using strand maps to engage digital library users with science content* [dok. elektr.]. [http://ieeexplore.ieee.org/xpl/login.jsp?tp=&number=4118574&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs\\_all.jsp%3Farnumber%3D4118574](http://ieeexplore.ieee.org/xpl/login.jsp?tp=&number=4118574&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D4118574) [odczyt: 5.07.2014].
- Butterworth, Richard (2006). *Designing digital library resources for users in sparse, unbounded social networks* [dok. elektr.]. [http://download.springer.com/static/pdf/973/chp%253A10.1007%252F11863878\\_16.pdf?auth66=1405074667\\_213ec7832eb2027c-6c71155f25ad1e9b&ext=.pdf](http://download.springer.com/static/pdf/973/chp%253A10.1007%252F11863878_16.pdf?auth66=1405074667_213ec7832eb2027c-6c71155f25ad1e9b&ext=.pdf) [odczyt: 5.07.2014].
- Catarci, Tiziana; (2000). *Digital libraries: did you ask the users?* [dok. elektr.]. *Kyoto International Conference on Digital Libraries*. [http://ieeexplore.ieee.org/xpl/login.jsp?tp=&number=942155&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs\\_all.jsp%3Farnumber%3D942155](http://ieeexplore.ieee.org/xpl/login.jsp?tp=&number=942155&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D942155) [odczyt: 5.07.2014].

- Chowdhury, Gobinda (2010). *From digital libraries to digital preservation research. The importance of users and context*. Journal of Documentation Vol. 66, p. 207-223.
- Chowdhury, Gobinda; McMenemy, David; Poulter, Alan (2008). *MEDLIS: Model for Evaluation of Digital Libraries and Information Services*. World Digital Libraries Vol. 1, p. 35-36.
- Chowdhury, Sudatta (2012). *User-centric studies*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 33-42.
- Cisek, Sabina (2010). *Metoda analizy i krytyki piśmiennictwa w nauce o informacji i bibliotekoznawstwie w XXI wieku*. Przegląd Biblioteczny t. 78 z. 3, s. 273-284.
- Clough, Paul (2012). *User-related issues in multilingual access to multimedia collections*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 117-126.
- Consonni, Chiara (2010). *Non-users' evaluation of digital libraries*. A survey at the Università degli studi di Milano. IFLA Journal Vol. 36, p. 325-331.
- DeLoene, William H.; McLean, Ephraim R. (1992). *Information systems success. The quest for the dependent variable*. Information Systems Research Vol. 3, p. 60-95.
- DeLoene, William H.; McLean, Ephraim R. (2002). *Information systems success Revisited*. In: *Proceedings of the 35<sup>th</sup> Hawaii International Conference on System Sciences (HICSS 02)*. Hawaii: IEEE Computer Society, p. 238-249.
- DeLoene, William H.; McLean, Ephraim R. (2003). *Information systems success. A ten-year update*. Journal of Management Information System Vol. 19, No 4, p. 9-30.
- Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi (2012). *And now ... to the brave real world*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 247-254.
- Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi (2012). *Introduction. User studies for digital library development*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 1-18.
- Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, 272 p.
- Doctor, Gayatri (2011). *A simulation study for a number of simultaneous users of an institutional knowledge repository system* [dok. elektr.]. World Digital Libraries Vol. 4. <http://iospress.metapress.com/content/4777v7140h587127/> [odczyt: 5.07.2014].
- Druin, Allison (2005). *What children can teach us. Developing Digital Libraries for Children with children* [dok. elektr.]. The Library Quarterly: Information Vol. 75, No. 1. <http://hcl2.cs.umd.edu/trs/2003-39/2003-39.pdf> [odczyt: 5.07.2014].
- Duff, Wendy M. (2012). *User studies in archives*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 199-206.
- Dunlop, Hayley; Cunningham, Sally Jo; Jones, Matt (2002). *A digital library of conversational expressions: helping profoundly disabled users communicate*. JCDL (Joint Conference on Digital Libraries), p. 273-274.
- Fox, Edward A. et al. (1993). *Users, user interfaces, and objects envision, a digital Library*. Journal of the American Society for Information Science and Technology Vol. 44, p. 480-491.
- Frías-Martínez, Enrique; Chen, Sherry Y.; Liu, Xiaohui (2007). *Automatic cognitive style identification of digital library users for personalization* [dok. elektr.]. Journal of the

- American Society for Information Science and Technology Vol. 58, p. 237-251. <http://onlinelibrary.wiley.com/doi/10.1002/asi.20477/pdf> [odczyt: 5.07.2014].
- Frías-Martínez, Enrique; Chen, Sherry Y.; Liu, Xiaohui (2008). *Investigation of behavior and perception of digital library users: A cognitive style perspective*. International Journal of Information Management Vol. 28, p. 355-365.
- Frías-Martínez, Enrique; Chen, Sherry Y.; Macredie, Robert D.; Liu, Xiaohui (2007). *The role of human factors in stereotyping behavior and perception of digital library users. A robust clustering approach*. User Model [dok. elektr.]. [http://download.springer.com/static/pdf/372/art%253A10.1007%252Fs11257-007-9028-7.pdf?auth66=1405075591\\_9f70d32c3256c5eec02f1439674f1ff9&ext=.pdf](http://download.springer.com/static/pdf/372/art%253A10.1007%252Fs11257-007-9028-7.pdf?auth66=1405075591_9f70d32c3256c5eec02f1439674f1ff9&ext=.pdf) [odczyt: 5.07.2014].
- Frías-Martínez, Enrique; Magoulas, George; Chen, Sherry; Macredie, Robert (2006). *Automated user modeling for personalized digital libraries* [dok. elektr.]. International Journal of Information Management Vol. 26, p. 234-248. <http://www.sciencedirect.com/science/article/pii/S0268401206000090> [odczyt: 5.07.2014].
- Girgensohn, Andreas et al. (2010). *Flexible access to photo libraries via time, place, tags, and visual features* [dok. elektr.]. JCDL (Joint Conference on Digital Libraries). <http://dl.acm.org/citation.cfm?id=1816151> [odczyt: 5.07.2014].
- Gladney, H.M.; Lotspiech, J.B. (1997). *Safeguarding digital library contents and users. Assuring convenient security and data quality* [dok. elektr.]. D-Lib Magazine. <http://www.dlib.org/dlib/may97/ibm/05gladney.html> [odczyt: 5.07.2014].
- Górny, Mirosław; Mazurek, Jolanta (2012). *Key users of Polish digital libraries*. The Electronic Library Vol. 30, p. 543-556.
- Griefender, Elke (2009). *A content analysis on the use of methods in online user research*. In: Verheul, Ingeborg; Tammara, Anna Maria; Witt, Steve eds. (2010). *Digital library futures. User perspectives and institutional strategies*. Berlin; München: De Gruyter Saur, p. 41-51.
- Griffiths, Jillian R. (2012). *Questionnaires, interviews and focus groups as means for user engagement with evaluation of digital libraries*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 65-74.
- Guin, Jeffery K. (2012). *User engagement and social media*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 137-148.
- Guldbæk Rasmussen, Katja; Petersen, Gitte (2012). *Personas*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 105-114.
- Hazan, Susan (2009). *What is a Library NOT a Library?* In: Verheul, Ingeborg; Tammara, Anna Maria; Witt, Steve eds. (2010). *Digital library futures. User perspectives and institutional strategies*. Berlin; München: De Gruyter Saur, p. 61-80.
- Hazan, Susan (2012). *User studies in museums. Holding the museum in the palm of your hand*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 207-216.
- Heradio, Ruben et al. (2012). *A review of quality evaluation of digital libraries based on users' perceptions*. Journal of Information Science Vol. 38, p. 269-283.
- Hess, Kirk (2012). *Discovering digital library user behavior with Google Analytics* [dok. elektr.]. Code4Lib Journal Issue 17, p. 1-10. <http://journal.code4lib.org/articles/6942> [odczyt: 5.07.2014].

- Higgins, Silke; King, Martin Luther (2013). *Access to digital libraries for disadvantaged users*. Library Philosophy & Practice p. 1-13.
- Hill, Linda L. et al. (1997). User evaluation. Summary of the methodologies and results for the Alexandria Digital Library, University of California at Santa Barbara [dok. elektr.]. <https://www.asis.org/annual-97/alexia.htm> [odczyt: 5.07.2014].
- Hong, Weiyin; Thong, James Y.L.; Wong, Wai-Man, Tam, Kar-Yan (2001). *Determinants of user acceptance of digital libraries. An empirical examination of individual differences and system characteristics*. Journal of Management Information Systems Vol. 18, No. 3, p. 97-124.
- Hutchinson, Hilary Browne et al. (2005). *The International Children's Digital Library: a case study in designing for a multilingual, multicultural, multigenerational audience*. Information Technology and Libraries Vol. 24, p. 4-12.
- Inskip, Charles; Butterworth, Richard; MacFarlane, Andrew (2008). *A study of the information needs of the users of a folk music library and the implications for the design of a digital library system*. Information Processing & Management Vol. 44, p. 647-662.
- Janiak, Małgorzata (2012). *Biblioteka cyfrowa, biblioteka elektroniczna, biblioteka wirtualna*. W: Janiak, Małgorzata; Krakowska, Monika; Próchnicka, Maria red. *Biblioteki cyfrowe*. Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, s. 15-65.
- Kani-Zabihi, Elahe; Ghinea, Gheorghita; Chen, Sherry Y. (2006). *Digital libraries: what do users want?* [dok. elektr.]. Online Information Review Vol. 30, p. 395-412 <http://www.emeraldinsight.com/journals.htm?articleid=1570032> [odczyt: 5.07.2014].
- Khalil, Mounir A. (2000). *Are end-users satisfied by using digital libraries?* [dok. elektr.]. [http://link.springer.com/chapter/10.1007%2F3-540-45268-0\\_40](http://link.springer.com/chapter/10.1007%2F3-540-45268-0_40) [odczyt: 5.07.2014].
- Khoo, Michael; Hall, Catherine (2012). *What would 'Google' do? Users' mental models of a digital library search engine* [dok. elektr.]. [http://link.springer.com/chapter/10.1007/978-3-642-33290-6\\_1](http://link.springer.com/chapter/10.1007/978-3-642-33290-6_1) [odczyt: 5.07.2014].
- Khoo, Michael; Kusunoki, Diana S.; MacDonald, Craig M. (2012). *Finding problems. When digital library users act as usability evaluators* [dok. elektr.]. [http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6149081&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs\\_all.jsp%3Farnumber%3D6149081](http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6149081&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D6149081) [odczyt: 5.07.2014].
- Klas, Claus-Peter (2012). *Expert evaluation methods*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 75-84.
- Kohl, Ulrich; Lotspiech, Jeffrey; Kaplan, Marc A. (1997). *Safeguarding digital library contents and users. Protecting documents rather than channels* [dok. elektr.]. D-Lib Magazine. <http://www.dlib.org/dlib/september97/ibm/09lotspiech.html> [odczyt: 5.07.2014].
- Konstantelos, Leo (2012). *Digital art online. Perspectives on user needs, access, documentation and retrieval*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 217-224.
- Koohang, Alex; Ondracek, James (2005). *Users' views about the usability of digital libraries*. British Journal of Educational Technology Vol. 36, p. 407-423.
- Law, Derek (2012). *User studies in libraries*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 191-198.
- Liu, Ziming (2006). *Print vs. electronic resources. A study of user perceptions, preferences, and use*. Information Processing & Management Vol. 42, p. 583-592.

- Liu, Ziming; Luo, Lili (2011). *A comparative study of digital library use. Factors, perceived influences, and satisfaction*. The Journal of Academic Librarianship Vol. 37, p. 230-236.
- Madle, Gemma et al. (2003). *Evaluating the changes in knowledge and attitudes of digital library users* [dok. elektr.]. [http://link.springer.com/chapter/10.1007/978-3-540-45175-4\\_4](http://link.springer.com/chapter/10.1007/978-3-540-45175-4_4) [odczyt: 5.07.2014].
- Makri, Stephann et al. (2007). *A library or just another information resource? A case study of users' mental models of traditional and digital libraries* [dok. elektr.]. Journal of the American Society for Information Science and Technology Vol. 58, p. 433-445. <http://onlinelibrary.wiley.com/doi/10.1002/asi.20510/full> [odczyt: 5.07.2014].
- Manduca, Cathy A. et al. (2005). *Influencing user behavior through digital library design* [dok. elektr.]. D-Lib Magazine. <http://www.dlib.org/dlib/may05/fox/05fox.html> [odczyt: 5.07.2014].
- Matusiak, Krystyna K. (2012). *Perceptions of usability and usefulness of digital libraries* [dok. elektr.]. International Journal of Humanities & Arts Computing Vol. 6, p. 133-147.
- Menzies, Kathleen; Birrell, Duncan (2012). *Significant others. User studies and digital preservation*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 149-158.
- Meyyappan, N.; Chowdhury, Gobinda; C.; Foo, Schubert (2001). *Use of a digital work environment prototype to create a user-centered university digital library*. Journal of Information Science Vol. 27, p. 249-265.
- Mihaylov, Petar (2012). *Design issues and user needs*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 43-50.
- Mintzer, Fred; Lotspiech, Jeffrey; Morimoto, Norishige (1997). *Safeguarding digital library contents and users. Digital watermarking* [dok. elektr.]. D-Lib Magazine. <http://www.dlib.org/dlib/december97/ibm/12lotspiech.html> [odczyt: 5.07.2014].
- Moghaddam, A. R. Isfandyari (2009). *Managing digital libraries in the light of staff and users. An approach*. International Journal of Information Science & Management Vol. 7, p. 31-40.
- Nanard, Marc; Nanard, Jocelyne (2001). *Cumulating and sharing end users knowledge to improve video indexing in a video digital library* [dok. elektr.]. <http://pagesperso.lina.univ-nantes.fr/~prie-y/ens/04-05/MasterRecherche/articles/1.pdf> [odczyt: 5.07.2014].
- Nicholas, David (2009). *The virtual scholar. The hard and evidential truth*. In: Verheul, Ingeborg; Tammaro, Anna Maria; Witt, Steve eds. (2010). *Digital library futures. User perspectives and institutional strategies*. Berlin; München: De Gruyter Saur, p. 23-32.
- Nicholas, David; Clark, David (2012). *Evidence of user behavior. Deep log analysis*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 85-94.
- Noorhidawati, Abdullah; Fariza Hanum, Nasaruddin, Zohoorian-Fooladi, N. (2013). *Social tagging in a scholarly digital library environment. Users' perspectives* [dok. elektr.]. Information Research Vol. 18. <http://www.informationr.net/ir/18-3/colis/paperC43.html#.VCw9hsscTVI> [odczyt: 5.07.2014].
- Normore, Lorraine F. (2009). *Characterizing a Digital Library's Users. Steps towards a nuanced view of the user*. Proceedings of the American Society for Information Science and Technology Vol. 45, p. 1-7.

- Notess, Mark (2003). *Three looks at users. A comparison of methods for studying digital library use* [dok. elektr.]. <http://variations2.indiana.edu/pdf/notess-3way-v2-prepub.pdf> [odczyt: 5.07.2014].
- Notess, Mark; Kouper, Inna; Swan, Maggie B. (2005). *Designing effective tasks for digital library user tests. Lessons learned*. OCLC Systems & Services Vol. 21, p. 300-310.
- Nov, Oded; Ye, Chen (2008). *Users' personality and perceived ease of use of digital libraries. The case for resistance to change*. Journal of the American Society for Information Science and Technology Vol. 59, p. 845-851.
- O'Dwyer, Andy (2012). *User studies for digital libraries' development: audiovisual collections*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 225-234.
- Oard, Douglas (1997). *Serving users in many languages: cross-language information retrieval for digital libraries* [dok. elektr.]. D-Lib Magazine. <http://www.dlib.org/dlib/december97/oard/12oard.html> [odczyt: 5.05.2014].
- Osborne, Nicola (2012). *Support for users within an educational or e-learning context*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 179-188.
- Park, Soyeon (2000). *Usability, user preferences, effectiveness, and user behaviors when searching individual and integrated full-text databases. Implications for digital libraries*. Journal of the American Society for Information Science Vol. 51, p. 456-468.
- Payette, Sandra D.; Rieger, Oya Y. (1998). *Supporting scholarly inquiry. Incorporating users in the design of the digital library*. Journal of Academic Librarianship Vol. 24, p. 121-130.
- Petrakieva, Lina (2012). *The shift to mobile devices*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 159-166.
- Porcel, Carlos; Lizarte, María Jesús; Herrera-Viedma, Enrique (2009). *A linguistic recommender system for university digital libraries to help users in their research resources accesses* [dok. elektr.]. <http://aisel.aisnet.org/mcis2009/133> [odczyt: 5.07.2014].
- Recker, Mimi (2006). *Perspectives on teachers as digital library users. Consumers, contributors, and designers* [dok. elektr.]. D-Lib Magazine. <http://www.dlib.org/dlib/september06/recker/09recker.html> [odczyt: 5.07.2014].
- Röttingen, Einar (2009). *A pianist's use of the digitised version of the Edvard Grieg Collection*. In: Verheul, Ingeborg; Tammaro, Anna Maria; Witt, Steve eds. (2010). *Digital library futures. User perspectives and institutional strategies*. Berlin; München: De Gruyter Saur, p. 53-60.
- Ruthven, Ian; Landoni, Monica; Lingnau, Andreas (2012). *Children and digital libraries*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 127-136.
- Sakaguchi, Tetsuo et al. (1996). *A browsing tool of multi-lingual documents for users without multi-lingual fonts*. Digital Libraries p. 63-71.
- Schatz, Bruce R. et al. (1996). *Interactive term suggestion for users of digital libraries. Using subject thesauri and co-occurrence lists for information retrieval*. Digital Libraries p. 126-133.
- Seaman, David (1997). *The user community as responsibility and resource: building a sustainable digital library* [dok. elektr.]. D-Lib Magazine. <http://www.dlib.org/dlib/july-97/07seaman.html> [odczyt: 5.07.2014].

- Sedighi, Zeinab; Gilvari, Abbas; Nooshinfard, Fatemeh (2013). *Children Digital Libraries and user interfaces. Proposing a set of criteria*. International Journal of Information Science & Management Vol. 11 Issue 2, p. 59-75.
- Sellen, Mary; Hazard, Brenda (2002). *User assessment of electronic reserves and implications for digital libraries*. Journal of Interlibrary Loan, Document Delivery & Information Supply Vol. 12, p. 73-83.
- Soergel, Dagobert (2002). *Thesauri and ontologies in digital libraries: Part 1. structure and use in knowledge-based assistance to users; Part 2: Design, evaluation, and development* [dok. elektr.]. <http://www.dsoergel.com/cv/B63.pdf> [odczyt: 5.07.2014].
- Somerville, Mary M.; Brar, Navjit (2009). *A user-centered and evidence-based approach for digital library projects*. The Electronic Library Vol. 27, No. 3, p. 409-425.
- Stephanidis, Constantine et al. (2000). *User interaction in digital libraries. Coping with diversity through adaptation*. International Journal on Digital Libraries Vol. 3, p. 185-206.
- Sykes, Jonathan et al. (2010). *A new focus on end users. Eye-tracking analysis for digital libraries* [dok. elektr.]. [http://link.springer.com/chapter/10.1007/978-3-642-15464-5\\_69](http://link.springer.com/chapter/10.1007/978-3-642-15464-5_69) [odczyt: 5.07.2014].
- Tammaro, Anna Maria (2008). *User perceptions of digital libraries. A case study in Italy*. Performance Measurement & Metrics Vol. 9, p. 130-137.
- Teruggi, Daniel (2009). *Who are the users of digital libraries. What do they expect and want? The Europeana experience*. In: Verheul, Ingeborg; Tammaro, Anna Maria; Witt, Steve eds. (2010). *Digital library futures. User perspectives and institutional strategies*. Berlin; München: De Gruyter Saur, p. 33-40.
- Thong, James Y.L.; Hong, Weiyin; Tam, Kar Yan (2004). *What leads to user acceptance of digital libraries?* Communications of the ACM – Bioinformatics Vol. 47, p. 79-83.
- Thong, James Y.L.; Hong, Weiyin; Tam, Kar-Yan (2002). *Understanding user acceptance of digital libraries: what are the roles of interface characteristics, organizational context, and individual differences?* International Journal of Human-Computer Studies Vol. 57, p. 215-243.
- Toms, Elaine G. (2012). *Models that inform digital library design*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 21-32.
- Tsakonas, Giannis (2012). *Users within the evaluation of digital libraries*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 51-62.
- Vaidyanathan, Ganesh; Sabbaghi, Asghar; Bargellini, Michael (2005). *User acceptance of digital library. An empirical exploration of individual and system components* [dok. elektr.]. [http://iacis.org/iis/2005/Vaidyanathan\\_Sabbaghi\\_Bargellini.pdf](http://iacis.org/iis/2005/Vaidyanathan_Sabbaghi_Bargellini.pdf) [odczyt: 5.07.2014].
- Vakkari, Pertti (2006). *Trends in the use of digital libraries by scientists in 2000-2005. A case study of FinELib* [dok. elektr.]. [http://eprints.rclis.org/8697/1/Vakkari\\_FinELib2000-05ASIST.pdf](http://eprints.rclis.org/8697/1/Vakkari_FinELib2000-05ASIST.pdf) [odczyt: 5.07.2014].
- Varga, Zsuzsanna (2012). *Resource discovery for research and course design*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 167-178.
- Verheul, Ingeborg; Tammaro, Anna Maria; Witt, Steve eds. (2010). *Digital library futures. User perspectives and institutional strategies*. Berlin; München: De Gruyter Saur, 175 p.

- Verwayen, Harry; Arnoldus, Martijn (2012). *A business-model perspective on end-users and open metadata*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. 235-244.
- Wildemuth, Barbara M. et al. (2002). *Alternative surrogates for video objects in a digital library. Users' perspectives on their relative usability* [dok. elektr.]. [http://link.springer.com/chapter/10.1007/3-540-45747-X\\_36](http://link.springer.com/chapter/10.1007/3-540-45747-X_36) [odczyt: 5.07.2014].
- Wilson, Tom (2012). *Preface*. In: Dobрева, Milena; O'Dwyer, Andy; Feliciati, Pierluigi eds. (2012). *User studies for digital library development*. London: Facet Publishing, p. IX- XII.
- Wolfram, Dietmar; Xie, Hong (2002). *Traditional IR for web users. A context for general audience digital libraries*. Information Processing & Management Vol. 38, p. 627-649.
- Wu, Dan; He, Daqing; Luo, Bo (2012). *Multilingual needs and expectations in digital libraries. A survey of academic users with different languages*. The Electronic Library Vol. 30, p. 182-197.
- Xie, Hong Iris (2006). *Evaluation of digital libraries. Criteria and problems from users' perspectives*. Library & Information Science Research Vol. 28, p. 433-452.
- Xie, Hong Iris (2008). *Users' evaluation of digital libraries (DLs). Their uses, their criteria, and their assessment*. Information Processing & Management Vol. 44, p. 1346-1373.
- Yan, Yalan; Zha, Xianjin (2014). *Comparison between user affinity with digital libraries and virtual communities*. Learned Publishing Vol. 27, p. 135-143.
- Zavalina, Oksana; Vassilieva, Elena V. (2014). *Understanding the information needs of large-scale digital library users. Comparative analysis of user searching* [dok. elektr.]. <http://alcts.metapress.com/content/h671q205qx565278/> [odczyt: 5.07.2014].
- Zha, Xianjin; Xiao, Zhiliang; Zhang, Jinchao (2014). *A survey of user perceptions of digital library e-quality and affinity*. Serials Review Vol. 40, p. 3-11.
- Zhang, Xiangmin; Li, Yuelin; Liu, Jingjing; Zhang, Ying (2008). *Effects of interaction design in digital libraries on user interactions*. Journal of Documentation Vol. 64, p. 438-463.
- Zhang, Ying (2010). *Developing a holistic model for digital library evaluation* [dok. elektr.]. Wiley Online Library. <http://onlinelibrary.wiley.com/doi/10.1002/asi.21220/pdf> [odczyt: 2.02.2011].
- Zhu, Qin; Guevara, Sophia (2009). *A practical guide for building a user-focused digital library collection*. Computers in Libraries Vol. 29, p. 6-10.



---

# **Intensity of information system use as a factor in evaluations made by information users. The example of Polish digital library users**

Mirosław Górny

Adam Mickiewicz University  
Institute of Linguistics, Department of Information Systems

Jolanta Mazurek

The Kórnik Library of the Polish Academy of Sciences

---

## **Introduction**

When investigating the effectiveness of information systems, we fairly frequently make use of evaluations made by the users of those systems. This is undoubtedly an entirely legitimate way of proceeding. After all, it is the user, or rather the user's evaluation, which would appear to be the most reliable source of information about the effects of the functioning of an information system.

In practice the question is not so clear-cut. A user making an evaluation of an information system is acting as a kind of measuring instrument. However, this instrument is very susceptible to the influence of a number of factors, and cannot be expected to provide exactly repeatable measurements.

It is sometimes claimed, of course, that much depends on the proper construction of such tools as questionnaires, and subsequently on the correct interpretation of the answers obtained. This is true; but nevertheless the results of surveys conducted among users are affected by very many factors – incomparably more than affect any measuring apparatus recording unambiguous measurable data such as size of sets, number and time of operations, etc. Indeed, even here certain difficulties arise, such as in relation to the comparability of units of measurement. This problem will be familiar to anyone who has tried, for example, to make a comparison of the infrastructures of several different libraries.

However, it is reliance on user evaluations that begins to cause real problems for those carrying out analysis of systems. How can any comparison be made when the measuring instrument is calibrated a little differently each time?

As a result, the results obtained through surveys of the opinions of users may show a relatively large degree of scatter, since on each occasion the users are a little different, or even completely different. Even when the composition of the respondent group remains fixed, the respondents themselves change. They have different knowledge, different experiences, a different attitude to the survey, or may simply be in a different psychological state. Users of information can be said to be a measuring instrument that has been to a certain degree “decalibrated”.

In that case, should user evaluations be used as a tool at all? The answer to this question is yes, but on condition that certain knowledge about the properties of this tool is taken into consideration.

## **Literature survey**

The literature concerning the evaluation of digital libraries is vast, and there is no space to discuss it in full detail here.

The greater part of the literature is concerned only with methods of library evaluation. This includes publications such as Frias-Marinez et al. (2008); Fuhr et al. (2007); Garibay et al., 2010; Gonçalves et al. (2007); Kimani et al. (2009); Pinto et al., (2007); Sfakakis et al., (2002), Tammaro (2008), Vinagre et al. (2011); Xie (2006) (2008).

Some publications contain evaluations collected for library design purposes. These include, for example, Roda et al. (2005), Ferreira and Pithan (2005) and Notess et al. (2005)

The influence of users' experience on evaluations of digital libraries has not yet been the subject of detailed research. Only a few publications, such as Kani-Zabihi et al. (2006), take account of this factor when reporting survey results.

The great diversity of literature on the subject of digital library evaluation does not make the researcher's task easier. There is a perception that almost every researcher in the field begins from scratch and plans to develop his or her own evaluation model. Another issue is that many of the indicators used are not specific to digital libraries, but relate, for example, to the problem of seeking information in general, or to general issues of the functionality of user interfaces of information systems. In evaluating digital libraries, it may be more productive to focus on questions that relate to such libraries specifically, omitting those that would more appropriately be the subject of research in more general fields.

A further issue is the effectiveness of the methodological approach taken. Most of the research involves surveying library users and giving a statistical presentation of the results. This may lead to the identification of relatively small differences that have little actual significance. Evaluations made using ten-point scales, for example, may be appropriate for medical, military and transport control systems, but less so in the case of libraries, which are used in different ways and for different purposes. It should be remembered that a user who is forced to evaluate something often does so without deeper consideration, and sometimes their evaluation reflects their personal condition emotional state or mental state rather than being a result of substantial analysis.

We are not questioning here the legitimacy of the research relating to evaluation of digital libraries. However we believe that in many cases, very complex evaluation models provide results that are not of great significance and have little importance for the practice of designing and building digital libraries.

## **Formulating the problem**

An information system, or rather its operation, can be approximately described by means of a set of suitable parameters. Each of the parameters takes values in some range. If we prepare a suitable set of parameters, we can ask users about their values. As a rule we do not do this directly – for each parameter we select suitable indicators whose value a user will be able to determine.

The responses of a user as evaluator depend on two groups of factors. One consists of factors affecting the state of the user, and the other of factors affecting the domain being evaluated by that user.

Aspects of the operation of a system can be classified depending on the difficulty of evaluating them and the ambiguity of the resulting evaluation. Where certain values can be measured, the evaluation is generally easier and unambiguous, but where it depends entirely on the user's subjective judgements, it clearly becomes harder to express and more ambiguous.

Mention should be made here of a very popular instrument for assessing the quality of libraries, namely LibQUAL. This includes evaluation in three dimensions, called Affect of Service, Information Control, and Library as Place.

Although the division into three dimensions here is not based on the degree of difficulty or ambiguity of evaluations, they are nonetheless linked to some extent. LibQUAL is a typical evaluation tool where the measuring instrument is the system user. It should be noted, however, that the results of the survey do not constitute a final evaluation of the system. They serve mainly to identify critical points in the functioning of the system.

As regards the goal of LibQUAL, its creators state the following:

“LibQUAL+® gives your library users a chance to tell you where your services need improvement so you can respond to and better manage their expectations. You can develop services that better meet your users’ expectations by comparing your library’s data with that of peer institutions and examining the practices of those libraries that are evaluated highly by their users.” [How Does LibQUAL+®... 2013, online doc.].

Hence the role of the user here comes down to pointing out, in various areas of the information system, clear deviations from the norm and places where the present level of service does not match expectations. At a later stage, however, researchers should focus on a precisely specified area of the information system’s operation, trying to determine why evaluations are either more than usually critical or more than usually positive. Each of these cases is interesting. The discovery of a configuration of factors which brings about exceptionally positive user evaluations is of equal importance as the reverse situation. Both determinations may equally well serve to improve the functioning of the information system.

The studies carried out by us had a similar goal. Digital libraries have existed in Poland for only 10 years, and most of them were created in the past two years. Two surveys were carried out, the first to identify the key users of digital libraries, and the second to obtain an evaluation of the libraries themselves.

The second survey was addressed to three identified groups of users (each group received a questionnaire prepared specially for that group, although part of the questionnaire was identical in each case). These three groups of key users were academic researchers (approximately 20% of all users), students (approximately 20%), and hobbyists, chiefly those interested in genealogy and local history (approximately 60%) [Górny, Mazurek 2012].

The questionnaires were sent individually to the e-mail addresses supplied in the first survey, and to the addresses of users who had written individually to the library administrators with various requests or comments. In total 9000 questionnaires were sent out, and 1349 completed forms were received back. This high return rate is a consequence of the enormous popularity, which digital libraries enjoy in Poland.

In analysing the results obtained, an attempt was made (although this was not the chief goal of the research) to identify groups of users based on the intensity of their use of digital libraries and to consider the differences in the evaluations made by them.

## Survey results

The respondents were divided into three basic groups – academics, students and hobbyists. Each of these groups was divided into three subgroups, based on intensity of library use. The first group, denoted HF, included users making frequent use of a digital library (more than once a week). The second group, denoted MF, contained users using the library at least once a month, but not more than once a week. The third group, denoted LF, included users who use the library less frequently than once a month.

The sizes of the groups were as follows:

Academics: (HF) N= 168; (MF) N=144; (LF) N=21

Students: (HF)=102; (MF) N=141; (LF) N=60

Hobbyists: (HF) N= 126; (MF) N=398; (LF) N=189

Users were asked to make evaluations using a system of scores, or grades, as used in Polish schools: excellent, very good, good, adequate, inadequate.

The results might have been analysed using the chi-squared test for fit, or by investigating the correlation between two sequences – the evaluation score and the stated frequency of library use. However the questionnaire results were not of a type that could be presented in the form of sequences of numbers. The frequency of use of a library was stated in various terms: sometimes as “rarely”, “from time to time” or “often”. Moreover it was necessary to decide whether “several times a month” means the same as “once a week”, etc. In the end a somewhat different method was chosen. For each group, a calculation was made of the percentage of respondents who had returned each of the scores (if a score had not been given at all, the percentage would obviously be zero). In this way a set of pairs was obtained: a score, together with the percentage of users who had indicated that score. Next the particular scores were assigned appropriate weights. The percentage of replies for a given score was multiplied by the weight. Namely, for a score of “excellent” the percentage of users in a given group was multiplied by 3. Multipliers for the other scores were 2 for “very good”, 1 for “good”, 0 for “adequate” and –1 for “inadequate”. To eliminate negative values from the graphs, each result was increased by 100. For clarity, a ten-point scale was adopted. That is, the results obtained were converted to a scale on which a result of 400 corresponded to 10 points, 200 to 5 points, etc.

Thus, for example, had all respondents in group H evaluated some parameter as “excellent”, it would have appeared on the graph with the value 10 (100% of responses  $\times 3 = 300$ ; adding the compensating value of 100 gives 400, which corresponds to the maximum of 10 points on the ten-point scale). Had all respondents indicated “good”, the graph would show 5 (100% of responses  $\times 1 = 100$ ; adding the compensating value gives 200; this corresponds to 5 points on the scale).

Table 1. Students' evaluations of digital library

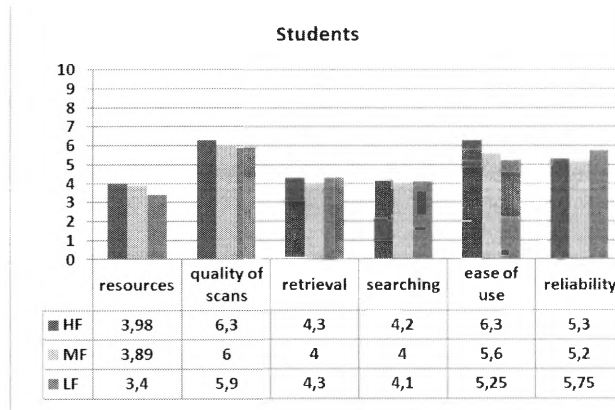


Table 2. Academics' evaluations of digital library

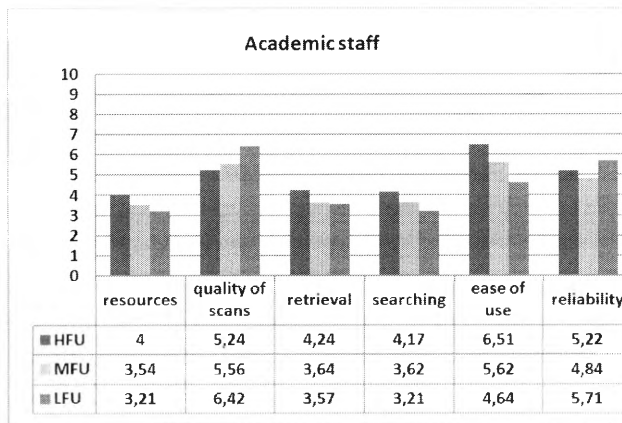
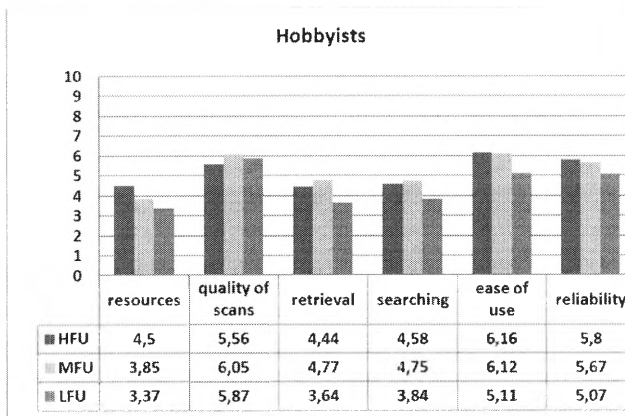


Table 3. Hobbyists' evaluations of digital library



Thus 10 points is equivalent to “excellent”, 7.5 to “very good”, 5 to “good”, 2.5 to “adequate”, and 0 to “inadequate”.

In relation to two features studied, a pattern was noticed whereby respondents in the group of most frequent users gave the highest scores, the MF group gave correspondingly lower scores, and the LF group gave the lowest scores. This applied to all of the studied classes of users – students, academics and hobbyists.

In the case of **evaluation of library resources** the results can be interpreted as follows:

a) someone who usually finds interesting publications in a digital library is likely to make frequent use of that library, thus frequent users can be expected to return higher scores;

b) frequent library users give higher scores because they know the resources better and can more easily find the publications they desire.

In the case of **evaluation of ease of use**, the correct explanation seems to be that those making frequent use of the library have simply come to know better the ways of using its interface, and thus give higher scores for ease of use than less frequent users.

Differences in scores between the basic types of users were also noteworthy. In the case of **evaluation of library resources** the highest scores were returned by hobbyists, and the lowest by professional academics. This can be interpreted as follows: hobbyists generally have more difficult access to academic libraries, and particularly to certain types of collections (such as publications from the 19<sup>th</sup> century or the interwar period), and hence the easy access to such publications that digital libraries provide gives many of them cause for enthusiasm. Academics, on the other hand, are more privileged users who have much easier access to collections that are not readily available to the general public.

## Conclusions

The results obtained from this survey certainly do not prove that there exists a relationship between the frequency of library use and the evaluations made concerning the library. However, neither do they exclude the existence of such a relationship.

We are fully aware that the analysis conducted has a number of weaknesses. One of them is that the result may be distorted by the proposed values for the criterion for classifying users according to frequency. A better solution would be to investigate the strength of the correlation between a sequence of values given for frequency of use and a sequence of score values. For this purpose a special

questionnaire would need to be prepared, eliminating the possibility of reporting frequency of use in a descriptive manner.

Our somewhat simplified method of visualization, based on the assignment of appropriate weights to the various possible scores and multiplication of those weights by the percentages of particular groups of respondents returning the respective scores, has not been analysed mathematically, and may therefore produce certain errors.

Our main aim, however, was to draw attention to a certain problem. The user learns the system, and the user's relationship with the system changes over time. It is therefore important to differentiate evaluations made by experienced users from those made by beginning users. Moreover it is undoubtedly necessary to distinguish users' evaluations depending on their past experience of information systems, on the length of time for which they have been using the system, and on the intensity of their use.

This applies, obviously, not only to the scores themselves. Certain assumptions can be made relating to the design and modernization of systems. A system designer may assume that a user will get to know the system over a certain amount of time, and only after that time will make full use of all of the possibilities provided by the system. This naturally affects the philosophy of design – for example, it may discourage efforts to achieve high system efficiency for novice users. By this token, a designer may avoid costly and complex solutions that can be expected to prove redundant to experienced users.

## REFERENCES

- Ferreira, Sueli Mara; Pithan, Denise Nunes (2005). *Usability of digital libraries: a study based on the areas of information science and human-computer-interaction. OCLC Systems & Services: International digital library perspectives*. Digital Library Usability Studies Vol. 21, p. 311-323.
- Frias-Marinez, Enrique; Chen, Sherry Y.; Liu, Xiaohui (2008). *Investigation of behavior and perception of digital library users: A cognitive style perspective*. International Journal of Information Management Vol. 28, p. 355-365.
- Fuhr, Norbert et al. (2007). *Evaluation of digital libraries*. International Journal on Digital Libraries Vol. 8, No. 1, p. 21-38.
- Garibay, Cecilia; Gutiérrez, Humberto; Figueroa, Arturo (2010). *Evaluation of a Digital Library by Means of Quality Function Deployment (QFD) and the Kano Model*. The Journal of Academic Librarianship Vol. 36, Issue 2, p. 125-132.
- Gonçalves, Marcos André et al. (2007). *What is a good digital library? A quality model for digital libraries*. Information Processing and Management Vol. 43, p. 1416-1437.



- Górny, Mirosław; Mazurek, Jolanta (2012). *Key users of Polish digital libraries*, The Electronic Library Vol. 30, No. 3, p. 543-556.
- How Does LibQUAL+® Benefit Your Library Users?* [online doc.]. [http://www.libqual.org/about/about\\_lq/general\\_info](http://www.libqual.org/about/about_lq/general_info) [accessed: 12.09.2014].
- Jeng, Judy (2005). *What Is Usability in the Context of the Digital Library and How Can It Be Measured?* Information Technology and Libraries Vol. 24, Issue 2, p. 47-56.
- Kimani, Stephen et al. (2009). *Digital Library Requirements: A Questionnaire-Based Study*. In: Theng, Yin-Leng et al. eds. *Handbook of Research on Digital Libraries: Design, Development, and Impact*. Information Science Reference. Hershey: IGI Global, p. 287-297.
- Kani-Zabihi Elahe; Ghinea, Gheorghita; Chen, Sherry Y. (2006). *Digital libraries: what do users want?* Online Information Review Vol. 30 Issue 4, p. 395-412.
- Notess, Mark; Kouper, Inna; Swan, Maggie B. (2005). *Designing effective tasks for digital library user tests: lessons learned*. OCLC Systems & Services: International digital library perspectives. Digital Library Usability Studies Vol. 21, p. 300-310.
- Pinto, Leonor Gaspar; Ochôa, Paula; Vinagre, Maria Helena (2008). *Integrated approach to the evaluation of digital libraries: an emerging strategy for managing resources, capabilities and results*. In: Heaney, Michael ed. *Library statistics for the twenty-first century world: proceedings of the conference held in Montréal on 18-19 August 2008 reporting on the global library statistics project*. München: K.G. Saur, p. 273-289.
- Roda, Claudia et al. (2005). *Digital image library development in academic environment: designing and testing usability*. OCLC Systems & Services: International digital library perspectives. Digital Library Usability Studies Vol. 21, p. 264-284.
- Sfakakis, Michalis; Kapidakis, Sarantos (2002). *User Behavior Tendencies on Data Collections in a Digital Library*. Lecture Notes in Computer Science Vol. 2458, p. 231-243.
- Tammaro, Anna Maria (2008). *User perceptions of digital libraries: a case study in Italy*. Performance Measurement and Metrics Vol. 9, No. 2, p. 130-137.
- Vinagre, Maria Helena; Pinto, Leonor Gaspar; Ochôa, Paula Paula Ochôa (2011). *Revisiting digital libraries quality: a multiple-item scale approach*. Performance Measurement and Metrics Vol. 12, Issue 3, p. 214-236.
- Xie, Hong (2006). *Evaluation of digital libraries: Criteria and problems from users' perspectives*. Library & Information Science Research Vol. 28, p. 433-452.
- Xie, Hong (2008). *Users' evaluation of digital libraries (DLs): Their uses, their criteria, and their assessment*. Information Processing and Management Vol. 44, p. 1346-1373.

---

# Digital libraries as a test bed for evaluating the effectiveness of information searching in OCR-processed texts

John Catlow, Mirosław Górny, Rafał Lewandowski

Adam Mickiewicz University

---

## 1. Introduction

The aim of this paper is to point out certain weaknesses of OCR with regard to the problem of information searching, and to describe the mechanisms involved. A methodology is presented for evaluating the effectiveness of information searching in OCR-processed texts. It is also shown to what extent relying exclusively on OCR techniques limits the possibilities of obtaining information from texts. It is indicated how and at what costs these limitations can be overcome by the use of keywords entered by a cataloguer. The research was conducted based on the resources and users of the Digital Library of Wielkopolska.

The resources of Polish digital libraries are largely of a historical nature. There are several reasons for this. Firstly, there is the desire to preserve the nation's heritage and enable access to it. We aim in particular to preserve for future generations those collections which are subject to gradual degradation. Also, many resources are not made available to readers physically because of the possibility that they might be further damaged. Converting them to digital form and making them available via the Internet enables them to be made accessible to a large number of readers. Another significant factor is undoubtedly the matter of copyright, which in the case of historical publications has generally expired. This means that such works can be converted to digital form and made available to readers much more easily, without the need for additional formalities and related costs.

A fundamental tool facilitating access to digital data is the index. We may distinguish two types of indexes used in digital libraries:

- indexes created manually by professional cataloguers;
- indexes created automatically from texts obtained via the process of optical character recognition (OCR).

## 2. Literature survey

Some literature already exists concerning the effectiveness of searching using indexes generated from OCR text. The authors of the paper “Evaluation of Model-Based Retrieval Effectiveness with OCR Text” [Taghva, Borsack, Condit 1996] carried out a series of experiments relating to information searching from data obtained from OCR-generated texts. They noted certain problems with the correct recognition of text, resulting from multi-column layouts, font variation, and graphical elements embedded in the text. Errors occurring during the OCR process, and characters generated as a result of misinterpretation of graphical elements as text, caused a significant increase in the size of the index. The authors reported that preliminary processing of the source materials and scanned images (increased clarity, scan alignment) proved more effective than attempts to improve the quality of documents after the OCR process.

The implementation of a search system based on OCR also carries increased disk space requirements. Even use of the manual zoning technique causes the size of the index to grow significantly.

The authors also describe problems with specific terms that cannot easily be corrected automatically – proper names and acronyms that are not commonly found in dictionaries, as well as numbers, dates and times of day. These data nonetheless carry high value for users of the information resources, and are important in a search system.

The authors report some negative results of using OCR technology, as well as certain methods for solving the problems. They also state that OCR can be useful in data search systems in which human intervention is minimal or absent.

The use of OCR with historical documents is addressed in the article “A Complete Optical Character Recognition Methodology for Historical Documents” [Vamvakas et al. 2008]. The authors present a methodology based on OCR technology for the indexing of printed and handwritten historical documents. The method involves three stages – in the first two a database is created on the basis of a set of documents, and in the third the database is used to index those documents.

The importance of the task of properly indexing historical documents is described in the article *Measuring Mass Text Digitization Quality and Usefulness* [Tanner, Muñoz, Ros 2009]. The authors researched the quality of OCR using two British Library databases: the 19th Century British Library Newspapers Database, a collection of 2 million pages of newspapers from England, Wales, Scotland and Ireland; and the Burney Collection Database, which covers the 17th and 18th centuries. The research covered approximately 1% of the 2 million pages contained in these resources. In the first database, “word accuracy” amounted to 78%, while

“number group accuracy” was only 64.1%. In the second database, which contains older publications, “significant word accuracy” fell to just 48.4%, which means that more than one half of significant words were incorrectly recognized. These results show that elapse of time has a significant effect on text indexing quality.

Another significant issue found in the literature is comparison of manual and automatic indexing. In the article *The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part I: Research, and the nature of human indexing* [Anderson, Perez-Carballo 2001] the authors draw attention to multiple variables of relevance to information retrieval systems:

- Document unit size – manual indexing concentrates on larger units, such as a complete article in a periodical, a complete chapter or a complete monograph. Automatic indexing, by contrast, is based on single paragraphs of text, not on whole documents.
- Scope of indexed material – automatic indexing covers the full content of a document, while manual indexing may be limited to the abstract or some other summary of the content.
- Exhaustivity – automatic indexing tends to be exhaustive, covering all terms contained in the text, while manual indexing is selective, taking in only those topics or aspects which appear to be most important for summarizing the content, significance or purpose of the text.
- Specificity – automatic indexing tends to use specific terminology (wide-ranging and differentiated vocabulary), as it uses the actual language of the text. Manual indexing tends to use more general terminology (and a numerically much smaller lexicon).

The authors also note that we know more about automatic than manual indexing, because machine methods need to be described rigorously so that they can be implemented by a computer algorithm. Understanding the way in which manual indexing is performed lies in the domain of such fields as cognitive psychology, cognitive science and neurology. Researchers are making constant progress in understanding how people think and perform mental tasks, but as the authors note, attempts to understand the steps that are taken in the mind of an indexer are still at the stage of speculative hypothesis.

### **3. Outline of the research problem**

Optical character recognition makes it possible to carry out automatic full-text indexing at relatively low cost. Because all phrases in the text are included, a level of completeness close to 100% can be obtained. In the case of historical publi-

cations, errors resulting from incorrect text recognition are a constant problem. A high level of completeness would appear to be beneficial to users, since their searches will return all documents that contain the search term. In the case of terms that appear rarely, this method is effective. However, in the case of more popular terms, the list of results may be extremely long – to such an extent that the user will not be able to look through all of them in a reasonable time. It is also to be expected that some of the documents returned will not be of interest, since their content will be of little or no significance to the user.

Manual indexing, on the other hand, ensures a high level of relevance – 100% in principle. The disadvantage of this method, however, is its high cost.

It can therefore be expected that automatic indexing based on OCR data will enable searching for all documents in a collection which contain the search term. Some of these documents, however, will not be relevant to the user. In order to find relevant documents, the user is forced to look through all of those returned by the search. It may even be necessary for the data browsing task to be divided between multiple users.

When the manual indexing method is used, all returned documents can be expected to have high relevance, and the probability of finding a significant document is close to 100%. On the other hand, this method has a low level of completeness (not all documents are indexed), and so the user cannot be sure that the list of results contains all relevant documents.

The creators and administrators of information systems often have to make a decision as to which indexing method to use. It can be expected that the main criteria to be considered will be indexing cost and user time.

The cost of automatic indexing is many times smaller than that of manual indexing. This saving is offset, however, by an increase in user time, as users have to perform the task of discarding irrelevant publications. If a resource is used intensively, this time is multiplied many times over – a large number of users will be spending time looking through irrelevant data. It can therefore be concluded that manual indexing becomes more beneficial when the collection is intensely used and when the level of relevance attained by automatic indexing is low. An increase in the costs of manual indexing (resulting from an increased number of indexed documents in the collection) is generally associated with a greater potential to achieve savings in user time.

Manual indexing should be treated as a supplementary indexing method, suitable for relatively large collections which have a lexical structure such that particular queries return results with a low degree of relevance. An additional factor is the intensity of use of the resource. In the longer term, the costs incurred will be paid back by savings in user time.

## 4. Optical character recognition

Optical character recognition (OCR) is a technology for recognizing characters and converting them from a raster image to the form of editable text. OCR maps the visual form of a particular symbol to its code in a character table – in the case of Polish texts, these usually correspond to the ISO-8859-2 or UTF-8 standards. The result may be recorded in a text file or merely displayed on a computer screen. The OCR process is often laborious and costly. Nonetheless, it is a task which many institutions choose to undertake. The main advantage of this technology is the fact that OCR data make it possible to generate indexes, which render the full texts of publications searchable.

In the case of historical documents, there are many technical difficulties involved in optical character recognition. These include:

- Paper degradation;
- Thinness of newspaper, which means that text on the other side of the page sometimes shows through during scanning;
- Imprecise printing;
- Incorrect storage;
- Creasing of the pages of newspaper and books;
- Fonts that are no longer in use;
- Damage that has occurred during use.

## 5. Models of publications contained in digital libraries

Every digital resource can be analysed as consisting of a certain number of objects. The analysis can be made at several levels of detail: we may consider the characters, the words, or the phrases which make up a publication.

In considering digital libraries, we may focus on three publication models:

- The newspaper – this has a small volume and covers varied topics. The articles are short, and their content chiefly informational. An important factor for the user is the possibility of looking through search results quickly and easily, due to the small volume. Newspapers were often printed on inferior paper using obsolescent printing machines. It can therefore be expected that the quality of OCR will be lowest in this case.
- The magazine – this is more extensive than a newspaper, containing long articles, often on varied topics. In the search results, the user is usually interested only in a selected article. In the case of publications of this type, it is helpful to use segmentation to make the data browsing process easier. OCR is more effective with magazines than with newspapers.

- The monograph – this is the most extensive out of the publication models being considered. Depending on the subject (whether it covers multiple topic areas or only one) it may be helpful to use segmentation to make browsing easier. Because of the print quality, this type of publication provides the highest OCR quality.

Apart from the publication model, the quality of OCR output depends also on the quality of scanning and the type of software used.

The OCR accuracy rate can be significantly improved using auxiliary software. This can provide such functionality as de-skewing, noise removal, document layout analysis, and binarization.

The OCR process may also include a correction stage. It may be automated – this solution has the advantage of speed, and enables large collections of data to be corrected in a short time. Automatic post-OCR correction involves three stages: tokenization, where the OCR system generates words; searching for words in a dictionary; and finally substitution of appropriate terms from the dictionary. This operation may also be performed in semi-automatic fashion, requiring the intervention of a human operator to select appropriate expressions.

Nonetheless, both of these methods are extremely costly and labour-intensive, and many libraries do not apply them, instead choosing to accept poorer-quality OCR results.

## **6. Tools to support data searching – indexes**

Indexes are a popular tool for users of information systems. For the purposes of our present analysis we will consider two types:

- indexes created based on terms entered manually by a cataloguer;
- indexes created automatically from OCR-processed text.

Indexes created for the same resource (one publication or a larger collection) may differ significantly, and thus may be received differently by users.

An automatically created index takes an unordered form. It includes terms which were not correctly recognized during the OCR process, and may therefore be incomprehensible to users. This applies in particular to terms that were not correctly recognized by the OCR software because they were split between lines. In languages such as Polish there are many words that have the same root, but appear in different inflected forms. Finally, in the case of historical literature and publications, there appear many terms which are no longer used and will be unknown to users.

The index will be very extensive and will include many terms that users will not in fact use during the data search process.

A manually created index is more friendly to the user and will be free of the defects listed above. It will nonetheless contain fewer terms, and it may well happen that a user fails to find a desired phrase in a manually created index, even though it would be present in an index produced automatically.

An automatically created index can be transformed into a more user-friendly form, although this requires time and costs. Incorrectly recognized terms can be corrected, lemmatization and stemming can be applied, and a dictionary of synonyms can be attached. Such an operation would need to be performed whenever a new publication is added. It would therefore require the allocation of certain hardware, software and human resources.

It is also possible to generate automatic indexes in the form of a tag cloud, which indicates not only the terms themselves, but also their popularity (frequency of occurrence) in a particular resource (publication or full collection).

The usual reason for using only automatic full-text indexing is the large volume of material involved. Many digital libraries lack the funds required to provide manual indexing of all of their resources.

A drawback of full-text indexing is that it normally leads to a very large and highly diversified dictionary, to which digital library users have to adapt. With manual indexing it is possible to use controlled dictionaries, which make the content of the descriptions uniform. Most digital libraries, including the Digital Library of Wielkopolska, in which the present study was carried out, contain material covering a wide variety of subjects. This means that searching would be made much easier if controlled dictionaries were used. Nonetheless the preparation of such dictionaries is very expensive, and even the implementation of already existing dictionaries requires the purchase of appropriate software.

## **7. Differences in searching resulting from the use of OCR or manual indexing**

Full-text searching usually generates a large number of results. Particularly in the case of general or popular terms, a marked difference can be seen in comparison with manual indexes. In the case of rarely encountered terms, it may be that they are omitted from a manually created index. In that case an index created automatically from OCR data offers a clear advantage, as it provides the only possibility of finding the desired term.

On the other hand, the large number of generated results makes searching through them difficult and laborious, and may discourage the user from investigating the resources further. There may also be a large number of irrelevant results, which further discourages the user.



The smaller number of results obtained in the case of manual indexing offers the possibility that all of the items found will be read.

On the other hand, the person making the catalogues and descriptions of the content may fail to mark all of the significant materials which would appear in full-text searching.

## **8. Costs of creating an index using OCR and manually**

The costs of creating indexes from OCR output can be significant. The text obtained directly from scanned materials may be highly imperfect. Its quality is affected by several factors:

- The quality of the source materials – in the case of Polish digital libraries these are often historical materials, with low paper and printing quality. The text obtained from such scanned materials is often very imperfectly recognized.
- OCR software – the performance of such programs can vary markedly. Also of importance may be the list of dictionaries used to support text recognition and the ability to add non-standard dictionaries. In the case of larger libraries, when extensive quantities of materials are processed using OCR, the speed of the process will also be significant.

The texts obtained will usually require correction – by automatic means, but also with the involvement of humans. There is a need for “manual” text correction, removal of flexional variants, and addition of synonyms. It is also necessary to remove “stop words” (words with little information content).

In the case of the resources from the Digital Library of Wielkopolska which were used in this study, no such operations are performed, in view of the high cost. The words generated by the OCR process are placed on the server and indexed without any additional correction. A similar procedure is applied by other Polish digital libraries.

## **9. Time required for searching and evaluation of information**

The list of publications found as a result of a search, particularly a full-text search, may be extremely long. There may be several thousand publications and hundreds of pages to look through. The time needed to familiarize oneself with the results of a search is often unacceptable to a user. The user has to look through all of the publications in order to find which of them are valuable (significant) for him or her, and which are not. The time required to investigate the results may differ greatly depending on the model of publication involved. Checking the results found in a newspaper, in view of its small volume, will take only a short time.

In the case of a monograph, when a term occurs in many places in the text, full investigation of the text may take longer.

## 10. Relevance

Relevance is a means of evaluating the value of information. Wiesław Babik, in his “Encyclopaedic dictionary of information, languages and search systems” [Bojar Bożenna oprac. 2002, p. 229], distinguishes four types of relevance:

- technical relevance;
- semantic relevance;
- pragmatic relevance;
- logical relevance.

In our study we take account of two of these types of relevance:

“Technical relevance is a relationship which, in a search set, e.g. in a set of characteristics of searched documents, for a given information search system, defines a subset of information elements, each of which satisfies the condition that its conformance (similarity) to the search instruction is not less than a level fixed for the system in question. Information elements which satisfy this condition are returned to the user in the response to a query for information, i.e. they are considered by the system to be documents that are relevant from the user’s point of view. (...)”

Pragmatic relevance is the relevance existing between a set of information found by the search system as a result of the application of a search instruction, and the user of the information, in particular the user’s information needs” [Bojar Bożenna oprac. 2002, s. 229].

## 11. Tests

A series of tests was carried out on the digital data stored in the Digital Library of Wielkopolska. The publications stored in the library are manually indexed, and selected publications are also indexed automatically using texts produced by OCR.

In both full-text and manual indexing, resources are searched at the level of complete publications. In the case of magazines, this means a particular issue of the magazine, and not a particular article within it. The results list directs the user to the issue of the magazine in which the sought term was found.

### Methodology

The search terms were found in publication descriptions, and the number of successful hits was recorded. After that the same search was carried out in full-text mode, based on data from the index produced by OCR processing.

The search terms were placed in quotes, so as to search for them in exactly the form in which they were entered in the search field.

The search terms used in the tests related to various subject areas.

Table 1. Numbers of results obtained by searching the Digital Library of Wielkopolska for terms being names of Polish cities, using automatically and manually generated indexes.

Term	Number of results from full-text searching	Number of results from searching publication descriptions	Publication description results as percentage of full-text results
“Poznań”	202,066	102,858	51.09%
“Kalisz”	7,388	648	8.78%
“Piła”	9,391	17	0.18%
“Konin”	3,062	142	4.64%
“Gniczno”	24,294	3,602	14.83%
“Leszno”	11,924	5,004	41.97%
“Bydgoszcz”	21,505	2,831	13.16%
“Toruń”	14,991	708	4.72%
“Gdańsk”	19,887	697	3.50%

The results of the tests in the “cities” category show that the number of hits in the case of full-text searching is significantly greater. In view of the origin of the collections (the Wielkopolska region) the numbers of hits for cities outside that region are noticeably lower.

In certain cases, a word being the name of a city also has other meanings. Some of the search results may thus be misleading. This applies to terms such as *Piła* (which also means “saw” in Polish), *Łódź* (“boat”), *Brzeg* (“edge”), *Łuków* (genitive form of “bows” or “arcs”), *Krosno* (“loom”), *Koło* (“wheel”), *Turek* (“Turk”), etc. This may be one of the reasons for the much greater number of results returned by full-text searching than by searching of manual descriptions (as in the case of “Piła”). It should be remembered that information about these cities may also be available under their German or even Latin names – such results may appear when a dictionary of synonyms is used.

Table 2. Numbers of results obtained by searching the Digital Library of Wielkopolska for terms being names of smaller towns in Poland, using automatically and manually generated indexes.

Term	Number of results from full-text searching	Number of results from searching descriptions of publications
“Wągrowiec”	15,313	3,891
“Lubon”	1,676	17
“Szamotuły”	13,250	3,002
“Rogoźno”	5,057	41
“Kościan”	8,226	510
“Gostyń”	6,331	331
“Jarocin”	7,925	4,644
“Pleszew”	6,639	3,372
“Krotoszyn”	9,415	1,353
“Wolsztyn”	6,374	2,996
“Tuchola”	1,513	7
“Tczew”	3,001	8
“Włocławek”	3,689	122
“Świebodzin”	114	20

Table 3. Numbers of results obtained by searching the Digital Library of Wielkopolska for terms being names of Polish villages, using automatically and manually generated indexes.

Term	Number of results from full-text searching	Number of results from searching descriptions of publications
“Kąkolewo”	459	5
“Dopiewo”	867	2
“Tarnowo Podgórne”	390	6
“Krzemieniewo”	263	6
“Damaśławek”	1,144	19
“Mieścisko”	1,293	7
“Ryczywół”	902	7
“Budziszewo”	119	0
“Pałczewo”	9	0

Table 4. Numbers of results obtained by searching the Digital Library of Wielkopolska for terms being surnames or full personal names, using automatically and manually generated indexes.

Term	Number of results from full-text searching	Number of results from searching descriptions of publications
“Pietrzak”	1,732	6
“Gajewski”	2,002	16
“Wierzbicki”	2,725	8
“Milewski”	2,456	24
“Śliwiński”	2,892	29
“Paderewski”	2,600	41
“Chłapowski”	5,314	114
“Ratajski”	2,472	15
“Bojanowski”	1,017	17
“Cegielski”	7,311	692
“Hipolit Cegielski”	292	3
“Dowbor-Muśnicki”	532	17
“Tytus Działyński”	266	83
“Arkady Fiedler”	226	3
“Iłłakowiczówna”	378	25
“Feliks Nowowiejski”	417	4
“Piotr Wawrzyniak”	265	14

Table 5. Numbers of results obtained by searching the Digital Library of Wielkopolska for terms having inflectional variants, using automatically and manually generated indexes.

Term	Number of results from full-text searching	Number of results from searching descriptions of publications
“Teatr Polski w Poznaniu” (Polish Theatre in Poznań)	1,229	86
“filatelistyka” (philately)	239	386
“filatelistów” (philatelists, genitive case)	319	0
“filatelistami” (philatelists, instrumental case)	12	0
“filatelistyczny” (philatelic)	127	0
“filatelista” (philatelist)	157	0
“filatelistka” (philatelist, fem.)	1	0
“Powstanie wielkopolskie” (Wielkopolska Uprising)	797	1,755

There is a certain group of terms which were not included in any manually created descriptions of publications. This brings to light an advantage of automatic full-text indexing – without such an index, the terms in question would not have been found at all.

Table 6. Numbers of results obtained by searching the Digital Library of Wielkopolska for terms referring to institutions and events, using automatically and manually generated indexes.

Term	Number of results from full-text searching	Number of results from searching descriptions of publications
“Gimnazjum w Wolsztynie” (Wolsztyn Middle School)	31	0
“Gimnazjum w Kaliszu” (Kalisz Middle School)	26	0
“Biblioteka w Lesznie” (Leszno Library)	6	0
“Dom Kultury w Kościanie” (Kościan Culture Centre)	4	0
“Kino w Lesznie” (cinema in Leszno)	6	0
“kapela dudziarska” (bagpipe band)	24	0
“turniej recytatorski” (public recitation competition)	38	0
“wystawa akwarystyczna” (aquarium exhibition)	0	0
“mecze piłkarski Polska Niemcy” (Poland-Germany football match)	14	0
“Unia Leszno” (name of a speedway club)	15	0
“Ostrovia Ostrów” (name of a speedway club)	43	0
“przegląd filmów amatorskich” (amateur film review)	4	0
“Parafia w Pleszewie” (Pleszew Parish)	2	0
“zawody wędkarskie” (angling competition)	42	0
“koncert muzyczny” (musical concert)	21	0
“koncert muzyki” (music concert)	1418	0

## Relevance of search results

A query was constructed to return a list of publications which were indexed automatically under a particular term, but had not been indexed under that term by human indexers.

Example query:

Text of publication “Leszno”

This produces a list of documents whose description does not contain the search term, although the OCR-processed text is indexed under that term. Only newspapers were included in this test.

The results indicate the number of publications containing the search term. The user does not know how many times a term appears in each publication. Every publication must be opened and searched. In the test, information was recorded concerning the source of the publication, the number of occurrences of the search term in the publication, and the context in which the word appeared.

Table 7. Numbers of newspapers in the Digital Library of Wielkopolska in which given search terms were found using automatically generated indexes, but not when using manual indexes.

Term	Number of occurrences
“Leszno”	1,319
“Poznań”	6,146
“Gniezno”	1,303
“Szamotuły”	547
“Chłapowski”	464
“filatelista”	4
“Cegielski”	367
“Powstanie wielkopolskie”	48

Analysis was also carried out using the program AntConc 3.2.4w. This enables the rapid display of information concerning the number of occurrences of a defined term in text publications. It is then possible to calculate the frequency with which terms occur. An additional feature of the program is that it displays results in a way that makes it possible to determine easily the context of every occurrence of the term in the publication.

Table 8. Selected test results (in English translation) containing information about documents in which the term “Leszno” was found using automatically generated indexes, but not when using manually generated indexes.

No.	Magazine title	Issue	No. of occurrences	Context
1	<i>Żołnierz Wielkopolski</i> : illustrated weekly magazine devoted to physical, military and general education. Published by the Poznań Provincial Committee for Military and Physical Education	1928.06.24 R.9 No. 26	2	Names of two delegates to the Congress of the Union of Reserve Officers, with the name of their town in brackets.
2	<i>Żołnierz Wielkopolski</i> : illustrated weekly magazine devoted to physical, military and general education. Published by the Poznań Provincial Committee for Military and Physical Education	1928.07.12 R.9 No. 28	1	Report on relay race for the <i>Głos Leszczyński</i> Cup – 10 lines.
3	<i>Żołnierz Wielkopolski</i> : illustrated weekly magazine devoted to physical, military and general education. Published by the Poznań Provincial Committee for Military and Physical Education	1928.01.01 R.9 No. 1 (275)	2	Multi-page article on the Wielkopolska Uprising. The name of Leszno appears among many towns which took part in the Uprising.
4	<i>Żołnierz Wielkopolski</i> : illustrated weekly magazine devoted to physical, military and general education. Published by the Poznań Provincial Committee for Military and Physical Education	1922.10.11 R.3 No. 24	1	Report on a football match between Polonia Leszno and a team from the 3 <sup>rd</sup> Air Regiment (1/4 page).
5	<i>Żołnierz Wielkopolski</i> : illustrated weekly magazine devoted to physical, military and general education. Published by the Poznań Provincial Committee for Military and Physical Education	1928.06.10 R.9 No. 24	1	A speech on the occasion of the blessing of the flag of the Association of Uprising Veterans and Fighters in Krzywiń includes information on the actions of the Leszno group.
6	<i>Żołnierz Wielkopolski</i> : with illustrations	1924.09.01 R.5 No. 24/25	1	List of names of the Executive of the Union of Reserve Junior Officers in Western Poland. Among more than 20 names is the name of a deputy councillor from Leszno.
7	<i>Żołnierz Wielkopolski</i> : illustrated weekly magazine devoted to the military and military education. Published by the Union of Reserve Officers and Junior Officers of Western Poland and the Union of Uprising Veterans and Fighters O.K. VII	1927.06.19 R.8 No. 33 (247)	2	Report on a journey around Poland by professional junior officers of the Poznań Corps. In Zakopane they visited the sick former commander of the 55 <sup>th</sup> regiment from Leszno. Report on summer camps of military and physical education. Leszno appears among several dozen names of towns.



No.	Magazine title	Issue	No. of occurrences	Context
8	<i>Żołnierz Wielkopolski</i> : illustrated weekly magazine devoted to the military and military education. Published by the Union of Reserve Officers and Junior Officers of Western Poland and the Union of Uprising Veterans and Fighters O.K. VII	1927.08.14 R.8 No. 41 (255)	3	List of Regional Commands of the Union of Uprising Veterans and Fighters. The Leszno region is listed among a total of 16 regions.
9	<i>Żołnierz Wielkopolski</i> : illustrated weekly magazine devoted to the military and military education. Published by the Union of Reserve Officers and Junior Officers of Western Poland and the Union of Uprising Veterans and Fighters O.K. VII	1927.08.07 R.8 No. 39/4 (253/254)	1	Report on military sporting competitions. A brief mention of a football match between military teams.
10	<i>Żołnierz Wielkopolski</i> : with illustrations	1923.03.15 R.4 No. 3	1	Report on a corporal from Leszno, who had won a lottery prize
11	<i>Orlątko</i> : magazine for Polish boys and girls	1934 April R.2 No. 7	1	Replies to letters to the editors. One of the letters was written by brothers from Leszno.
12	<i>Orlątko</i> : magazine for Polish boys and girls	1934 June R.2 No. 9/10	1	A scout from Leszno correctly identified other scouts in fancy dress. Report on 8 lines.
13	<i>Kino Teatr</i> : illustrated fortnightly magazine on cinema and theatre	1928.11.15 R.1 No. 4	1	Photographic plates were produced by the Cynkograf firm, from the village of Leszno near Warsaw
14	<i>Kino Teatr</i> : illustrated fortnightly magazine on cinema and theatre	1928.12.01 R.1 No. 5	1	Photographic plates were produced by the Cynkograf firm, from the village of Leszno near Warsaw
15	<i>Mucha</i>	1911 R.43 No. 32	2	Two advertisements for firms from the village of Leszno near Warsaw (a corn cure manufacturer and a tar board factory).

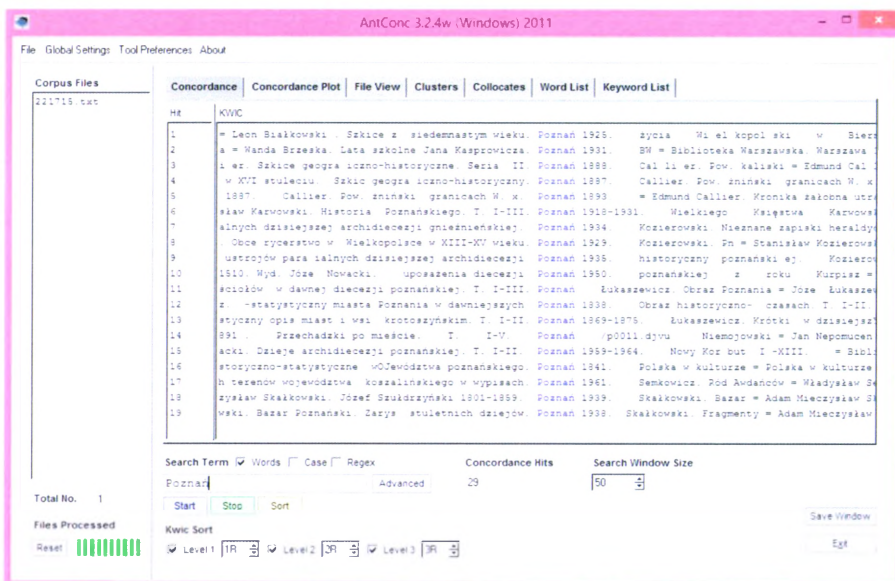


Figure 1. A screenshot from the AntConc 3.2.4w program.

## 12. Conclusions

A digitized text is a file – sometimes large, sometimes small – containing a certain number of strings of characters. These strings form words which have some significance for the user. If we map them to the file addresses at which they occur, we can obtain a list of files in which those words appear. This is how searching in digital texts works. The user provides a specific word in a query, and obtains as a result a set of texts which contain that word. The problem is that words usually appear in many different texts, and the user is often required to look through large numbers of texts in order to find one of interest.

Of course, computer tools provide certain means of narrowing down one's search results. A user may add other words to the query, or may define the position of the words in the text – for example by requiring that certain words appear within a defined distance of each other – or may also require that certain words appear with a certain frequency, and that certain others must not occur. This, however, is all that computerization can offer. The presence or absence of a given word in a text, its position, the co-occurrence of other words and the frequency of occurrence of a given word are effectively all that an automated system can allow one to specify. (It may also take account of the different ways in which particular words are distinguished, but the user is not usually able to use such features in

a query – they are more likely to be used by the system itself, which may apply various default procedures to try to determine the relevance of a result.)

Use of these techniques requires a significant amount of effort and experience on the user's part, as well as knowledge about the collections being searched. The effect may still be unsatisfactory (with an excessively large number of results being returned). It should also be borne in mind that the creation of such possibilities entails the high costs of building and operating a computer system, not to mention the problem of erroneous reading of certain words in the OCR process.

An automated solution may be complemented by a manual indexing system. The indexer uses his or her knowledge of users' needs and applies an appropriate strategy. Because this work is expensive, it is desirable to limit its application to a selected group of sources.

Lists of addresses and other types of list do not require manual indexing, as they contain sets of personal or other names that are relatively well ordered and well suited to being looked through manually by a user (in the case of printed works), and are relatively easy to process with OCR. For similar reasons, manual indexing is not applied as a rule to encyclopaedias, dictionaries and other such works of reference. Indexing of monographs is useful to a limited extent. This is because they usually concern a single subject, and contain extensive material that needs to be read as a whole. A description of their content that is adequate for a reader's needs is usually contained in the bibliographic description and in the table of contents, if this is included as a subcollection with purely informational functions.

An enormous problem is posed, however, by press publications (particularly daily newspapers). It may be assumed, with a certain degree of simplification, that manual indexing is a procedure that should be applied above all to newspapers and magazines.

Press publications have a different structure to that of other types of source. They consist of relatively short reports, covering varied topics and often having significant informational significance. The indexer must therefore act in a manner appropriate to the expected needs of users. We attempt to describe an appropriate procedure here. The indexer must pay attention to:

- Events – the indexer should assess whether a given event may be of significance to a potential reader. It is not possible to lay down any hard guidelines for this, as much depends on the indexer's knowledge of the subject matter and of readers' information needs. Nonetheless, in regional newspapers for example, the indexer will normally omit events of international or nationwide importance, on the assumption that such articles are merely reprinted from national newspapers, instead focusing attention on local events, which may not be recorded in any other sources.

- Absent key words – sometimes the indexer may decide that a word might be used by a reader even though it is absent from the text. For example, if the text contains extensive descriptions of military activity in some area, but the name of that area is not given explicitly, the indexer will often deduce what area is being described, and include its name as a key word.
- Deformed words – this may be an error resulting from font smudging, for example.
- Proper names which deserve to be exposed for some reason.
- Volume of text on a given topic – the indexer will pay attention to relatively extensive texts devoted to a particularly subject area, assuming that long texts by their nature may contain significant information, opinions, etc.
- Exceptional information – sometimes a relatively minor reference, which to contemporaries may not have seemed important, may now be of much greater significance, for instance as a step in the development of technology – an example of this was a report of an offer by German firms to install video telephones between Łódź and Warsaw in 1936.

In order for manual indexing to be of help to a reader, the words entered by indexers must be given greater weight during searching. Indeed, this is a feature provided by the software used by many systems.

It is difficult to assess the economic effectiveness of manual indexing. This is not due to any difficulty in calculating the costs of the process (if an indexer can index approximately 100 pages of text daily, then he or she will complete approximately 2000 pages per month, and approximately 20,000 pages per year – taking the annual cost of employing such a person in Poland to be around 40,000 zloty, the cost of indexing one page comes out to approximately 2 zloty or 0.5 euro). The obstacle faced is that it is not possible to place a value on the results of this indexing work.

The effectiveness of searching should not be assessed merely in terms of a percentage of appropriate results. In the case of searches performed by users of digital libraries, a single good result is sometimes of enormous value. Hence even if the large number of words inserted by indexers seems not to play any role during searching, it is enough that they may cause just one extremely important document to be returned.

Statistically, it is theoretically possible for a comparison of effectiveness to be made between searching using automatically constructed indexes and searching using manual indexes. Nonetheless this requires analysis to be performed in the course of real searching, since otherwise there is no possibility of determining the relevance of the results – only the person performing the search is able to make such a determination.

A solution, up to a point, is to prepare tests according to appropriate scenarios. Then, however, the results may be unconsciously distorted by the person arranging the test. We do not have access to material that has been manually indexed in full, and a degree of error is likely to be introduced when samples of library content are prepared for testing purposes.

As a consequence, the tests are reduced to identifying a certain set of sources being found on the basis of given words from automatically constructed indexes. That set is then analysed to identify in it the words which ought to be exposed. The size of samples depends on the size of the basic collection and on the query made. At present the size of the basic collections of some digital libraries (including the Digital Library of Wielkopolska) would appear to be entirely adequate.

It must nonetheless be remembered that as the size of the basic collection increases, the effectiveness of searching of full-text resources is reduced, because of the increase in the size of the result sets. In that case the importance of manual indexing increases.

## REFERENCES

- Anderson, James D.; Perez-Carballo, Jose (2001). *The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part I: Research, and the nature of human indexing*. Information Processing and Management Vol. 37, p. 231-254.
- Bojar, Bożenna oprac. (2002). *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych* [Encyclopaedic dictionary of information, languages and search systems] (in Polish). Warszawa: Wydaw. Stowarzyszenie Bibliotekarzy Polskich, 2002, 363 s.
- Bieniecki, Wojciech (2005). *Analiza wymagań dla metod przetwarzania wstępnego obrazów w automatycznym rozpoznawaniu tekstu* [Analysis of requirements for methods of pre-processing of images in automatic text recognition] (in Polish). [http://wbieniec.kis.p.lodz.pl/research/files/05\\_Bronislawow\\_OCR.pdf](http://wbieniec.kis.p.lodz.pl/research/files/05_Bronislawow_OCR.pdf) [accessed: 22/06/2014].
- Droettboom, Michael (2003). *Correcting broken characters in the recognition of historical printed documents*. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.7866&rep=rep1&type=pdf> [accessed: 22/06/2014].
- EDL Report on Digitisation in European National Libraries 2006-2012 (2008). [http://www.cenl.org/docs/Report\\_digitisation\\_NLs.pdf](http://www.cenl.org/docs/Report_digitisation_NLs.pdf) [accessed: 22/06/2014].
- Eikvil, Line (1993). *OCR – Optical Character Recognition*. Report No. 876, 1993. <http://www.nr.no/~eikvil/OCR.pdf> [accessed: 22/06/2014].
- Hauser, W. Andreas (2007). *OCR Postcorrection of Historical Texts*. <http://www.cip.ifi.lmu.de/~hauser/papers/histOCRNachkorrektur.pdf> [accessed: 22/06/2014].
- Holley, Rose (2009). *Many Hands Make Light Work: Public Collaborative OCR Text Correction in Australian Historic Newspapers*. [http://www.nla.gov.au/ndp/project\\_details/documents/ANDP\\_ManyHands.pdf](http://www.nla.gov.au/ndp/project_details/documents/ANDP_ManyHands.pdf) [accessed: 22/06/2014].

- Impact: Improving access to text: Concept*, online document. <http://www.impact-project.eu/about-the-project/concept/> [accessed: 22/06/2014]
- Taghva, Kazem; Borsack, Julie; Condit, Allen (1996). *Evaluation of model-based retrieval effectiveness with OCR text*. ACM Transactions on Information Systems Vol. 14, Issue 1, p. 64-93.
- Lin, Leo (2009). *Improving Digital Library Support for Historic Newspaper Collections*. <http://researchcommons.waikato.ac.nz/bitstream/10289/3262/1/thesis.pdf> [accessed: 22/06/2014].
- Tanner, Simon; Muñoz Trevor; Ros, Pich Hemy (2009). *Measuring Mass Text Digitization Quality and Usefulness*. D-Lib Magazine Vol. 15, No. 7/8. <http://www.dlib.org/dlib/july09/munoz/07munoz.html> [accessed: 26/06/2014].
- Vamvakas, Goergios et al. (2008). *A Complete Optical Character Recognition Methodology for Historical Documents*. DAS, 08. The Eighth IAPR International Workshop on, Document Analysis Systems, p. 525-532.

---

# Wizualizacje modeli mentalnych bibliotek cyfrowych

Małgorzata Janiak, Monika Krakowska

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

W artykule omówiono wyniki badań dotyczących wizualizacji modeli bibliotek cyfrowych, stworzonych w latach 2011-2014 przez studentów Instytutu Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego studiów 1. (st1st) oraz 2. stopnia (st2st). Celem analiz było zdefiniowanie wzorców biblioteki cyfrowej (dalej w skrócie: BC) na podstawie prac rysunkowych studentów oraz wyodrębnienie najważniejszych elementów tego systemu informacyjno-wyszukiwawczego, które pojawiały się w modelach kognitywnych badanych użytkowników. Wyniki oceny skonfrontowano z dwoma modelami bibliotek cyfrowych: referencyjnym schematem DELOS, przedstawionym w manifestie *The Digital Library Reference Model* opracowanym dla Sieci Doskonałości dla Bibliotek Cyfrowych – *The DELOS Network of Excellence for Digital Libraries* [wersja z 2011 r.: Candela et al., dok. elektr.; Casarosa 2007], stanowiącym przykład modelu systemu informacyjnego, oraz modelem 5S – *Stream, Structures, Spaces, Scenarios, Societies* [Gonçalves et al. 2002, 2004], obrazującym wzorzec systemu informatycznego. Zastosowano metodologię badań jakościowych zapożyczoną z nauk społecznych, polegającą na analizie i interpretacji wykonanych rysunków i innych reprezentacji graficznych [Theron, Mitchell, Smith 2011].

## MODEL REFERENCYJNY BIBLIOTEKI CYFROWEJ DELOS

Międzynarodowa Sieć Doskonałości dla Bibliotek Cyfrowych DELOS w ogłoszonym manifestie *The Digital Library Manifesto* z 2006 r. [Candela et al. 2011, dok. elektr.] przedstawia przestrzeń biblioteki cyfrowej jako wielopłaszczyzno-

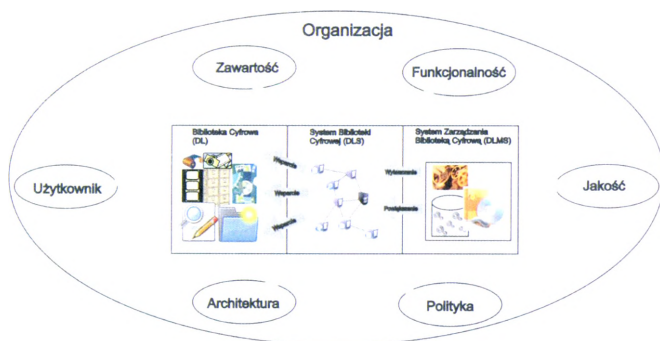
wą i złożoną strukturę, a sam termin „biblioteka cyfrowa” rozumie jako heterogeniczny, różnorodnie przedstawiany przez badaczy. W deklaracji DELOS biblioteka cyfrowa w najbardziej ogólnym rozumieniu charakteryzowana jest jako „połączenie usług i treści, system zapewniający funkcjonalność przekazywania informacji między obiektami i użytkownikami w kontekście rozproszonych zbiorów różnorodnych danych. Ta funkcjonalność obejmuje dostęp, publikowanie, dostarczanie, przechowywanie, personalizację itp., określoną jakość oraz zgodność z polityką tworzenia całego systemu” [Candela et al. 2011, dok. elektr.]. Należy zauważyć, że nie istnieje ujednolicona i powszechnie akceptowana definicja biblioteki cyfrowej [Janiak 2012, s. 15-52] i że BC bywa określana również jako narzędzie lokujące się w centrum działalności intelektualnej człowieka, które nie powinno mieć logicznych, koncepcyjnych, fizycznych lub czasowych granic, a które dodatkowo przełamuje osobiste i kulturowe bariery w dostępie do informacji [Candela et al. 2007, dok. elektr.]. Celem istnienia bibliotek cyfrowych jest tworzenie powszechnie dostępnego kanału komunikacyjnego, z założeniem że każdy może też uczestniczyć w dyskusjach nad jego doskonaleniem, w ewaluacji jego zasobów oraz w podnoszeniu jakości informowania o tych zasobach.

Główna wizualizacja zaprezentowana w manifeście, czyli „przedstawienie dowolnego obiektu za pomocą programu graficznego” (termin komputerowy) lub „wynik pracy polegającej na wzmożonym wytwarzaniu wyobrażeń wytwórczych” (termin psychologiczny) [Wizualizacja, dok. elektr.; Visualization in Scientific Computing 1987], sprowadziła całą przestrzeń biblioteki cyfrowej do trzech podstawowych konfiguracji (systemów): biblioteki cyfrowej (zasobu, z którym podejmują interakcje użytkownicy – DL), systemu biblioteki cyfrowej (aplikacji, które są zainstalowane i skonfigurowane dla prowadzenia biblioteki cyfrowej – DLS) oraz systemu zarządzania biblioteką (systemu oprogramowania, które wspiera tworzenie i zarządzanie biblioteką cyfrową – DLMS) [Candela et al. 2011, dok. elektr.].

Systemy te tworzą warstwy uniwersum biblioteki cyfrowej, gdzie każde pole dotyczy różnych grup użytkowników, którzy są zaangażowani np. w implementację biblioteki cyfrowej lub podejmują różne zachowania informacyjne w obrębie jej systemu. Wzajemne relacje łączące wybrane typy użytkowników (użytkownik końcowy, twórca-projektant, administrator oraz zarządzający) zostały zilustrowane na rysunkach 1 i 2.

Według przyjętych w artykule definicji i charakterystyki biblioteki cyfrowej rozumianej jako system można wyodrębnić pięć przenikających się wzajemnie oraz uzupełniających się zakresów, stanowiących fundament systemu informacyjnego: zawartość (ang. *content*), użytkownik (ang. *user*), funkcjonalność (ang.



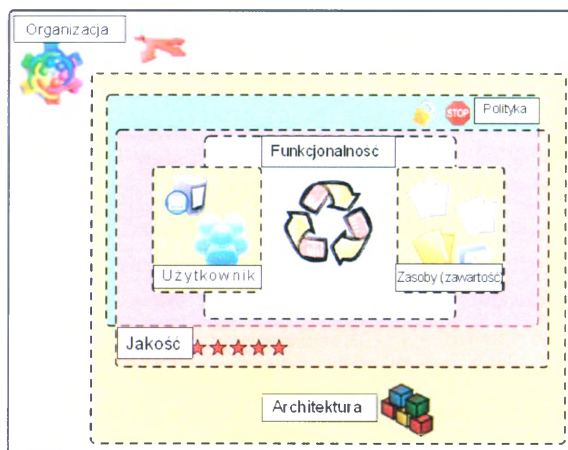


Rysunki 1 i 2. Koncepcja biblioteki cyfrowej i jej użytkowników.  
 Obydwa rysunki przetworzone za: [Candela et al. 2011, dok. elektr.]

*functionality*), polityka (ang. *policy*), jakość (ang. *quality*), a dodatkowym, szóstym elementem jest architektura, struktura systemu (ang. *architecture*).

Na rysunku 3 przedstawiono graficznie wszystkie obszary składające się na system biblioteki cyfrowej. Należy zauważyć, że pod terminem „użytkownicy” kryją się zarówno twórcy, jak i odbiorcy, zaangażowani we wszelkie procesy i działania związane m.in. z tworzeniem, implementowaniem, kreowaniem, korzystaniem lub promowaniem BC.

Obszary, zwłaszcza te dotyczące zawartości, użytkowników, funkcjonalności i polityki, mają wiele podobnych cech. Są to płaszczyzny odnoszące się do wewnętrznych jednostek, komponentów biblioteki cyfrowej, które mogą być rozpoznawane przez zewnętrzne podmioty (np. użytkowników różnego typu). Wprowadzenie, w kontekście systemu biblioteki cyfrowej (DL), pojęcia zasobów, które wiąże się z wymienionymi obszarami, ułatwia spójne rozumienie systemu biblioteki cyfrowej.



Rysunek 3. Elementy biblioteki cyfrowej oraz użytkownicy.  
Rysunek przetworzony za: [Candela et al. 2011, dok. elektr.]

Model referencyjny obejmuje wszystkie etapy tworzenia bibliotek cyfrowych, od projektowania aż po implementację systemu informacyjnego. Stanowi ramy pojęciowe, skonkretyzowaną strukturę, w której ujęte są różne kategorie użytkowników oraz ich relacje z systemem biblioteki cyfrowej. Model składa się ze zbioru ujednoczonych pojęć, aksjomatów i relacji, odnoszących się do każdego z wyodrębnionych obszarów, i jest niezależny od konkretnych standardów i technologii lub indywidualnych wdrożeń.

Manifest przedstawiony przez DELOS stwarza przede wszystkim podstawy do opracowania i implementowania modelu referencyjnego biblioteki cyfrowej, którego jednym z nadrzędnych celów jest odwzorowanie mentalnych reprezentacji biblioteki powstających w umysłach interesariuszy – odbiorców, bibliotekarzy, twórców, administratorów, menadżerów – oraz dostosowanie systemu do ich specyficznych potrzeb [Casarosa 2007, dok. elektr.].

## MODEL BIBLIOTEKI CYFROWEJ 5S

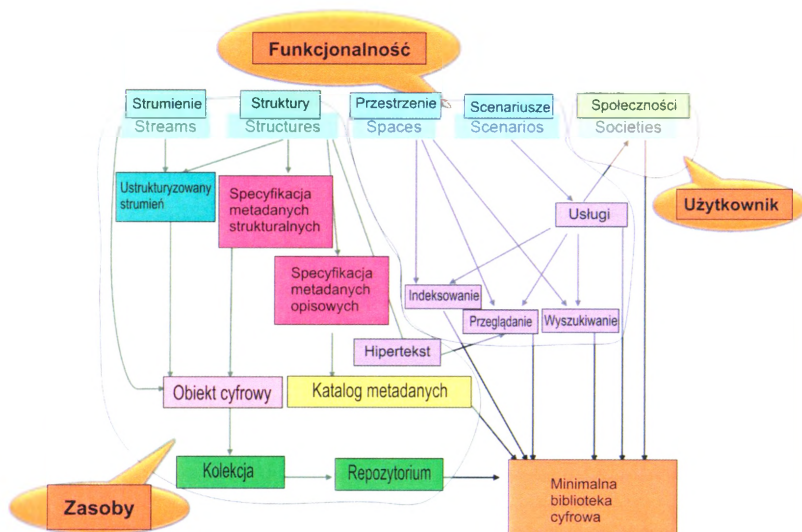
Abstrakcyjnym modelem bibliotek cyfrowych, bazującym także na wizualizacji mentalnej interesariuszy, jest model 5S: *Streams, Structures, Spaces, Scenarios, Societies* (określający elementy: strumieniowy, strukturalny, przestrzenny, sytuacyjny oraz społecznościowy) [Gonçalves et al. 2004, p. 271].

Schemat ten został oparty na wzajemnych relacjach zachodzących między obszarami: 1) strumieni (ang. *streams*), czyli różnych rodzajów informacji multime-

dialnej kreowanej i zarządzanej przez biblioteki cyfrowe, 2) struktur (ang. *structures*), czyli informacji zorganizowanej w bibliotece cyfrowej (metadane, formaty itp.), obiektów multimedialnych zebranych w kolekcje, 3) przestrzeni (ang. *spaces*), czyli odwzorowanych w systemie operacji, działań i usług, 4) wszelkich interakcji pomiędzy użytkownikiem a systemem oraz różnorodnych usług oferowanych przez bibliotekę cyfrową (ang. *scenarios*) oraz 5) współpracujących użytkowników – twórców, administratorów, badaczy, bibliotekarzy cyfrowych, odbiorców końcowych, jak i wszystkich możliwości tego systemu, którego funkcjonalność ma zaspokajać potrzeby informacyjne (ang. *societies*) [Gonçalves et al. 2004, p. 272-276].

Podobieństwo modelu 5S do referencyjnego schematu DELOS dotyczy przede wszystkim czterech obszarów: funkcjonalności (ang. *functionality*), zawartości (ang. *content*), struktur (ang. *structures*) i użytkowników (ang. *users*). Pierwsza z domen schematu DELOS obejmuje procesy indeksowania, wyszukiwania i przeglądania zasobów biblioteki cyfrowej oraz aspekt sytuacyjny korzystania z serwisów i usług oferowanych w ramach systemu. Obszar zawartości, obejmujący repozytorium cyfrowe, można uznać za odpowiadający obszarowi strumieni w modelu 5S, jako że mieszczą się w nim elementy takie, jak wybór odpowiedniego kodu językowego oraz opis i forma multimediiów – zasobów cyfrowych i kolekcji. Płaszczyzna *content* w DELOS koresponduje także z wymiarem *structures* w modelu 5S, ponieważ zawiera specyfikację organizacyjną struktury metadanych, katalog metadanych, organizację wiedzy w bibliotece cyfrowej. Obszar użytkowników w DELOS jest zbieżny z dymensją *societies* w 5S i odnosi się do wszystkich interesariuszy (administratorów, bibliotekarzy, twórców i odbiorców końcowych) oraz do wzajemnych relacji pomiędzy nimi. 5S formalizuje i objaśnia strukturę systemu informacyjnego przez ukazanie wielostronnych interakcji między jego komponentami [Gonçalves et al. 2004, dok. elektr., p. 276-279].

Model schematu biblioteki cyfrowej 5S, w którym nacisk pada na funkcjonalność systemu, uwzględnia najistotniejsze usługi, czynności i działania realizowane w jej strukturze, odnoszące się do poszczególnych procesów informacyjnych, np. wyszukiwania, przeglądania, odnajdywania zasobów (ang. *discovery*), a także do wizualizacji zasobów [Fox, Gonçalves, Shen 2012, p. 44]. Może on być skutecznie opisywany w swoistym języku (5S Language, 5SL) złożonym z podstawowych lub rozbudowanych obiektów matematycznych reprezentujących obiekty cyfrowe, metadane, kolekcje, usługi [Gonçalves, Fox 2002, dok. elektr.]. Język 5S, opracowany na potrzeby modelu 5S, umożliwia przygotowanie specyfikacji systemu informatycznego w tych właśnie pięciu wymiarach, które jednocześnie dotyczą także jakości informacji multimedialnej oraz zachowań informacyjnych użytkowników, chociaż nie zostało to *explicite* uwidocznione w samej wizualizacji modelu [Gonçalves, Fox 2002, dok. elektr.].



Rysunek 4. Obszary relatywne do referencyjnego modelu DELOS w modelu abstrakcyjnym biblioteki cyfrowej 5S. Rysunek przetworzony za: [Gonçalves, Fox, Watson, Kipp 2004]

W badaniach nad modelami mentalnymi użytkowników bibliotek cyfrowych, prowadzonych w pierwszym semestrze każdego roku akademickiego w latach 2011-2014, starano się powiązać rysunki studentów z poszczególnymi domenami zarówno modelu 5S, jak i referencyjnego schematu DELOS, a także wskazać te elementy, które były najczęściej typowane jako najważniejsze komponenty systemu informacyjno-wyszukiwawczego. Obszary stanowiące fundamenty bibliotek cyfrowych niekoniecznie muszą być bowiem tożsame z elementami, które zostały zawarte w najczęściej implementowanych modelach. Każdy z użytkowników, bazując na swoich indywidualnych obrazach mentalnych, stworzył wyobrażenie określonego systemu informacyjnego, zbudował go z różnych modułów, które następnie rozpoznał w rzeczywistości, np. w konkretnej bibliotece cyfrowej.

## Modele mentalne

W informatologii coraz częściej podejmowane są badania nad modelami mentalnymi, które stanowią podstawę rozumienia rzeczywistości, obiektów, systemów oraz inicjowania procesów informacyjnych przez użytkowników. Reprezentacje tworzone w ludzkim umyśle określane są, w terminologii różnych dyscyplin naukowych, jako struktury wiedzy, stany wiedzy, struktury kognitywne lub stany kognitywne [Chen, Ke 2014, p. 1676]. Według Vicki Ahlstrom i Kennetha

Allendoerfera wiedza na temat modeli mentalnych użytkowników przyczynia się do zrozumienia i wyjaśnienia procesów poznawczych człowieka, a w szczególności tego, jak informacja strukturyzowana jest w umyśle i w jaki sposób podmiot nią zarządza [Chen, Ke 2014, p. 1675]. Możliwe jest zatem zastosowanie wiedzy o mentalnych reprezentacjach rzeczywistości do procesów tworzenia systemów informacyjno-wyszukiwawczych, w celu lepszego zrozumienia interakcji człowiek – komputer.

Do modeli mentalnych odniesiono się także w psychologii kognitywnej. Na uwagę zasługuje definicja Kennetha Craika, który określił je jako imitacje świata fizycznego konstruowane w umyśle człowieka, na których ludzie opierają swoje działania i zachowania oraz przewidywania zdarzeń będących następstwem aktualnie podejmowanych działań. Stanowią one, zdaniem Philipa Johnsona-Lairda, indywidualne struktury wiedzy, wewnętrzne reprezentacje, które są podstawą rozumienia świata, nadawania mu sensu, a także wzajemnego oddziaływania człowieka z zewnętrznym uniwersum [Krakowska 2012, s. 313]. Wydaje się więc, że modele takie wnoszą istotny wkład do rozważań nad budową i funkcjonowaniem bibliotek cyfrowych, zwłaszcza przy analizowaniu odbioru i oceny jakości BC przez użytkowników. W przypadku badań nad sztuczną inteligencją w naukach informatycznych, a zwłaszcza w projektowaniu systemów informacyjno-wyszukiwawczych, modele mentalne stanowią realne odzwierciedlenie i kognitywną rekonstrukcję fizycznego systemu. Co istotne, człowiek tworzy i uaktualnia mentalny konstrukt na podstawie interakcji np. z nowymi technologiami, narzędziami komunikacji i informacji, opierając inicjowane procesy i działania także na wcześniejszych doświadczeniach oraz na posiadanej wiedzy [Chen, Ke 2014, p. 1676].

Biblioteka cyfrowa, rozumiana jako system i centrum działań intelektualnych oraz interakcji odbiorców, twórców, bibliotekarzy, administratorów, może również wspomagać umysłową aktywność człowieka [Chowdhury 2010, p. 209]. Istotne jest jednak, aby specyfikacja funkcjonalna systemu uwzględniała różnego typu modele mentalne wszystkich jej użytkowników [Roda 2005, p. 268-269]. Kompleksowość i wielorakość możliwych sposobów postrzegania biblioteki cyfrowej sprawia, że wyobrażenia użytkowników mogą odznaczać się dużym zróżnicowaniem. W psychologii poznawczej przyjmuje się, że każdy podmiot tworzy subiektywne poznawcze reprezentacje otaczającego świata, pewne wzorce porównawcze. Abstrakcyjna struktura umysłowa to zwykle „nietrwała reprezentacja słowno-obrazowa o dość dużym poziomie złożoności, powstająca w toku procesu myślenia lub rozumowania, jako sposób na przedstawienie sobie w umyśle istoty problemu i składających się nań zależności” [Nęcka, Orzechowski, Szymura 2008, s. 640]. Według Johnsona-Lairda model mentalny powstaje w wyniku

procesów tworzenia nowej i przekształcenia już nabytej wiedzy oraz podejmowania relewantnych zachowań informacyjnych, jest nietrwały, indywidualny, subiektywny, zależny od konkretnej sytuacji i posiadanej wiedzy [Westbrook 2006, p. 564; Nęcka, Orzechowski, Szymura 2008, s. 63].

Modele mentalne mogą być uznane za pojęciowe reprezentacje systemu informacyjno-wyszukiwawczego, zaś przez odkrywanie lub wywoływanie konstrukcji poznawczej świata rzeczywistego użytkowników można przekonać się, czy odbiorca potrafi ocenić efektywność działania systemu, a także świadomie kierować własnymi zachowaniami informacyjnymi [Rieh et al. 2010, p. 165]. Analiza rozumienia systemu przez użytkowników pozwala rozwijać wiedzę na temat sposobu korzystania z niego oraz ulepszać procesy informacyjne [Norman 1986, p. 7-8].

W badaniach nad interakcjami człowiek – komputer zauważono, że modele mentalne użytkowników odnoszą się do ich wewnętrznej wiedzy o systemach informacyjno-wyszukiwawczych (np. o bibliotekach cyfrowych, katalogach bibliotecznych, wyszukiwarkach internetowych), relacjach pomiędzy elementami architektury i wewnętrznym funkcjonowaniu tych systemów. Yang Zhang uważa, że model powstający w umyśle użytkownika umożliwi mu rozumienie całego systemu, jego struktury, zasad i celowości działania [Zhang, 2010, p. 2210]. Stwierdzono, iż akceptacja i rozumienie funkcjonowania systemów informacyjno-wyszukiwawczych determinowane są przez dominujące struktury poznawcze [Rieh et al. 2010, p. 166]. Rezultaty analiz tych reprezentacji pozwalają przede wszystkim na ocenę jakości istniejących systemów, ich udoskonalanie, a także eliminację ludzkich błędów pojawiających się podczas interakcji z systemem. Na struktury poznawcze tworzone w umyśle użytkownika, według Tefko Saracevica, mają wpływ zmienne, które mieszczą się w obszarze: 1) poznawczym, i są to wcześniejsze doświadczenia, wiedza dziedzinowa, percepcja użyteczności oraz łatwość użytkowania systemu, 2) społecznym, do którego należą motywacje, zainteresowania oraz potrzeby informacyjne, a także 3) demograficznym, określonym przez takie cechy, jak wiek, płeć, kontekst kulturowy i poziom wykształcenia [Pan et al. 2006, dok. elektr.].

Użytkownicy tworzą zbiory synchronicznych i diachronicznych zasad wyznaczających interpretowanie i rozumienie rzeczywistości oraz podejmowanie działań (w tym mieszczą się np. zachowania informacyjne). Niechętnie także odrzucają pierwotnie zbudowane modele mentalne, nawet wówczas, gdy są one niekompletne i mogą wzajemnie się na siebie nakładać podczas różnorodnych interakcji z systemem [Zhang 2010, p. 2207]. Mimo możliwości doskonalenia modeli w trakcie ich wielokrotnego stosowania w procesach informacyjnych, użytkownicy częstokroć posługują się znanymi odwzorowaniami, np. przenoszą model jednej biblioteki cyfrowej na drugą (na zasadzie: „to co jest znane, jest

bezpieczniejsze, szybsze i prostsze w obsłudze”) [Nęcka, Orzechowski, Szymura 2008, s. 60].

Każdy podmiot zaangażowany w procesy projektowania, zarządzania i użytkowania bibliotek cyfrowych ma własne reprezentacje mentalne. Twórcy budują modele pojęciowe (ang. *conceptual models*) systemu oraz modele użytkownika (ang. *user models*). Modele mentalne systemów bibliotek cyfrowych powstają z kolei w umysłach użytkowników końcowych. O ile model użytkownika odzwierciedla zestaw cech, jakie twórca systemu przypisuje domyślnie klientowi końcowemu tego systemu, o tyle pojęciowe wyobrażenie systemu powstające w umyśle użytkownika (wizja systemu) jest dla niego swoistą metaforą wyrażającą strukturę i działanie systemu. Zarówno twórcy systemu (projektanci oraz osoby zarządzające i administrujące systemem), jak i użytkownicy końcowi kreują własne wyobrażenia o sobie nawzajem oraz o obiektach i narzędziach, z którymi się stykają. Obrazy te są oparte na wiedzy o obiektach, spełnianych przez nie funkcjach, zachodzących między nimi interakcjach, zdarzeniach występujących podczas pracy z systemem [Westbrook 2006, p. 566]. Na model mentalny stworzony przez użytkownika ma wpływ zarówno wiedza i doświadczenie, jak i wewnętrzne modyfikacje reprezentacji wynikające z odwzorowywania rzeczywistości. Konieczne jest, aby podczas konsytuacji modelu projektu (ang. *design model*), odpowiadającej wizji twórcy, uwzględniana była właśnie informacja o wiedzy i potrzebach użytkownika końcowego, łącząca reprezentacje umysłowe systemu i użytkowników [Makri et al. 2007, p. 434].

Modele mentalne mogą być rozpatrywane z różnych perspektyw. Badania nad obrazowaniem systemów i sieci informacyjnych pozwoliły na ocenę reprezentacji umysłowych z: 1) perspektywy technicznej (ang. *technical view*), gdzie najważniejsze elementy to system komputerów, serwerów, narzędzi ICT oraz ich adekwatność do oczekiwań użytkowników, 2) perspektywy funkcjonalnej (ang. *functional view*), zakładającej analizowanie różnego rodzaju usług, w tym wyszukiwania informacji, 3) perspektywy procesu (ang. *processed view*), gdzie system lub serwis traktowane są jako centrum wszelkich procesów, interakcji, 4) perspektywy zależności (ang. *connection view*), która uwzględnia charakter powiązań pomiędzy poszczególnymi elementami systemu lub sieci, warunkujących przebieg poszczególnych procesów, interakcji, działań [Zhang 2008, p. 1331].

Efektywne funkcjonowanie systemów informacyjno-wyszukiwawczych oraz skuteczne korzystanie z nich wymaga uwzględnienia czynników afektywnych, emocjonalnych i motywacyjnych, zatem podczas wizualizowania struktur kognitywnych należy brać pod uwagę również aspekty psychologiczne. W nawiązaniu do jednej z podanych wcześniej definicji bibliotek cyfrowych, zgodnie z którą jest ona centrum nieograniczonych procesów poznawczych podmiotu realizowa-

nych podczas interakcji człowiek – komputer – system, w Instytucie Informacji Naukowej i Bibliotekoznawstwa UJ podjęto badania nad modelami mentalnymi użytkowników bibliotek cyfrowych, których celem było uchwycenie dominujących sposobów postrzegania i rozumienia tych systemów informacyjno-wyszukiwawczych przez studentów informatologii i bibliologii. Celem przedstawionych analiz było dokonanie opisu i charakterystyki graficznych reprezentacji bibliotek cyfrowych, stworzonych przez studentów IINiB UJ, a następnie zbadanie stopnia ich przystawalności, odpowiednio, do modelu referencyjnego biblioteki cyfrowej zaproponowanego przez sieć DELOS oraz do modelu 5S.

## Badania

W badaniach przeprowadzonych w latach 2011-2014 uczestniczyło: 132 studentów studiów 1. stopnia (st1st) oraz 54 studentów studiów stacjonarnych i niestacjonarnych 2. stopnia (st2st), realizujących specjalizację *Zasoby cyfrowe* (dalej: ZC). Zadaniem uczestników było przygotowanie rysunków, na których przedstawiali własne modele mentalne bibliotek cyfrowych. Każdy student otrzymał bardzo ogólną instrukcję, aby wykonać tylko jeden rysunek, który miał przedstawić mentalny, skojarzeniowy obraz biblioteki cyfrowej pojawiający się w jego umyśle (bez szczegółowych poleceń typu: tylko szkic, diagram, schemat, rysunki bez napisów itp.). Na wizualizacjach studenci na podstawie subiektywnej wiedzy ukrytej (ang. *implicite knowledge*) oraz doświadczenia w korzystaniu z bibliotek cyfrowych uwidocznili najistotniejsze komponenty systemu informacyjno-wyszukiwawczego, jego funkcjonowanie oraz interakcje zachodzące między biblioteką a użytkownikiem. Przygotowywane rysunki nie miały być wizualizacjami artystycznymi<sup>1</sup> i oddawać jedynie model strukturalny i / lub funkcjonalny biblioteki cyfrowej. Sposób prezentacji modeli mentalnych był bardzo różnorodny, a wyniki analiz przedstawione zostały w dalszej części artykułu.

Celem badań było między innymi uzyskanie odpowiedzi na pytanie, czy studenci studiów pierwszego i drugiego cyklu wizualizują biblioteki cyfrowe w różny sposób, zależny przede wszystkim od ich wiedzy o tych obiektach. W programie studiów 1. stopnia w Instytucie Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego kształcenie w zakresie bibliotek cyfrowych prowadzone było, do roku akademickiego 2013/2014, w ramach przedmiotów „Źródła informacji” oraz „Podstawy informacji naukowej i bibliotekarstwa”. Należy zatem

---

<sup>1</sup> Studenci w trakcie badania zgłaszali problem: „nie umiemy rysować”, dlatego zamiast, a czasem obok elementów graficznych, na szkicach pojawiły się napisy.



przyjąć, że studenci pierwszego roku studiów 1. stopnia dysponują ogólną wiedzą o bibliotekach cyfrowych jako o obiektach informacyjnych, często traktując je jako element biblioteki hybrydowej lub też jako „przeniesienie” realnej biblioteki do cyberprzestrzeni. Nie do końca jednak potrafią rozpoznać cechy konstytutywne biblioteki cyfrowej odróżniające ją od innych systemów informacyjnych, nie posiadają bowiem praktyki w zakresie tworzenia oraz ewaluowania bibliotek cyfrowych. Nie posiadają także odpowiedniej wiedzy teoretycznej dotyczącej użytkowników informacji oraz roli modeli mentalnych w opisie i wyjaśnianiu zachowań informacyjnych użytkowników.

Studenci wyższych lat poznają problematykę zachowań informacyjnych, potrzeb informacyjnych, funkcjonowania systemów informacyjnych, obiektów cyfrowych, stron www, interfejsów w trakcie realizowanego programu studiów, a więc mają określoną wiedzę i doświadczenie, które należy brać pod uwagę jako istotny czynnik podczas analizowania indywidualnych struktur mentalnych, których materializacją były stworzone przez nich rysunki. Należy przypuszczać, że studenci studiów 2. stopnia, którzy realizują specjalizację ZC, zgromadzili już obszerną wiedzę o bibliotekach cyfrowych, ich historii, różnorodnych formach, zasobach, działaniu, procesach informacyjnych oraz użytkownikach, aspektach prawnych, ocenie jakości, a dodatkowo posiadają duże doświadczenie w korzystaniu z tych systemów. Oprócz wiedzy teoretycznej studenci nabywają również umiejętności praktyczne w wyniku realizacji własnych projektów kolekcji cyfrowych, uczestniczenia w procesie digitalizowania obiektów, opisywania zasobów, a także zarządzania BC. Trzeba także domniemywać, że studenci studiów 2. stopnia, którzy wybrali specjalizację ZC, potrafią określić cechy użytkowników (obecnych i przyszłych) oraz przewidzieć pewne zachowania informacyjne, zarówno własne, jak i klientów biblioteki cyfrowej. Na własnych wyobrażeniach o BC, doświadczeniu i znajomości podstaw funkcjonowania każdego logicznego systemu informacyjnego opierać się będzie również ich ewentualna współpraca z bibliotekami różnego typu podczas tworzenia kolekcji cyfrowych. Ważne zatem jest określenie, jaki indywidualny obraz (reprezentację kognitywną) nie tylko samego systemu biblioteki cyfrowej, ale i jego użytkowników posiadają przyszli twórcy bibliotek cyfrowych.

## **Metodologia badań**

Modele mentalne, jako abstrakcyjne struktury umysłu, są trudne do zdefiniowania, rozpoznania oraz precyzyjnego scharakteryzowania. Badając je, naukowcy wykorzystywali przede wszystkim wywiad i protokoły myślenia na głos (ang. *think-aloud protocol*) [Brandt, Uden 2003; Katzeff 1990; Borgman 1986; Efthemia-

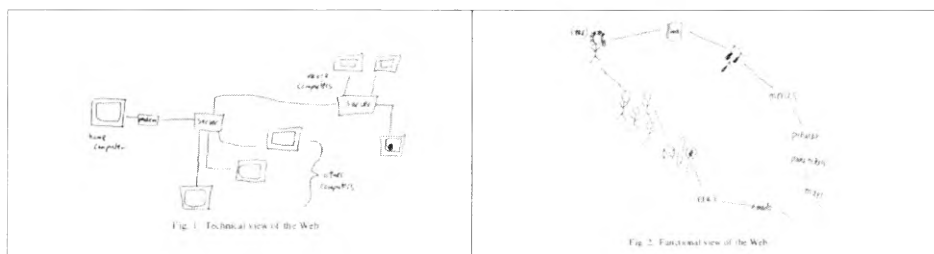
dis, Hendry 2005: za Zhang 2008, p. 1331]. W związku z tym, iż wyniki tych badań w dużym stopniu zależą od zasobu leksykalnego respondentów (znanych im terminów), w badaniach nad modelami mentalnymi zaczęto stosować rysunki dla zobrazowania struktury wyobrażeń tworzonych przez użytkowników, traktując rysunek jako jedną z pierwotnych form komunikacji [Gray 1990 za: Zhang 2008, p. 1331]. W niektórych ujęciach teoretycznych (Roland Barthes, W.J.T. Mitchell) obraz traktowany jest jako *quasi*-tekst, przekaz mogący podlegać analizie treści na podobieństwo przekazów wyrażonych w języku naturalnym. Nie można też zapomnieć o tym, że ludzie uczą się tworzyć i interpretować narracje wizualne w trakcie zdobywania zarówno wiedzy, jak i doświadczenia (Gunther Kress, Theo Van Leeuwen 2006)<sup>2</sup>. Opierając się właśnie na tych założeniach, podczas realizowania badań autorki postanowiły nie tylko odtworzyć modele mentalne studentów, lecz także porównać je zależnie od tego, czy ich autorami byli studenci studiów 1., czy 2. stopnia, co zostało już wspomniane na początku artykułu.

Metoda rysunków (szkiców) odręcznych była stosowana w badaniach informatologicznych m.in. przez Stephena T. Kerra (1990) oraz Sue Denham (1993), która użyła jej do zbadania dziecięcych wyobrażeń o tym, czym jest komputer. Z kolei Annelise Mark Pejtersen oraz Raya Fidel (1998) analizowały, przez ocenę graficznych wizualizacji, możliwość zastosowania schematów kognitywnych do tworzenia ram dla projektowania i oceny systemów informacyjno-wyszukiwawczych. Andrew Thatcher i Mike Greyling (1998) wykorzystali metodę rysunków do zbadania konceptualizacji internetu, zaś Judy Sheard i Jason Ceddia (2004) badali dezorientację użytkowników sieci wyrażoną w języku obrazów. Efthimis N. Efthimiadis i David G. Hendry (2005) próbowali określić, w jaki sposób studenci rozumieją funkcjonowanie internetu oraz wyszukiwarek. Metodę rysunku odręcznego zastosowali również Marina Papastergiou (2005), która analizowała modele mentalne internetu tworzone przez uczniów szkół średnich z Grecji, i Yang Zhang (2008), która na podstawie badań nad zachowaniami informacyjnymi studentów opracowała różne modele systemów informacyjno-wyszukiwawczych: procesowy, ogólny (globalny), związany z interfejsem oraz interaktywny. W swoim artykule Yang Zhang zaprezentowała wybrane rysunki, które okazały się stosunkowo podobne do tych wykonanych przez studentów polskich, mimo że nie dotyczyły tego samego, gdyż respondenci amerykańscy przedsta-

---

<sup>2</sup> Z tego powodu w artykule zaproponowano podzielnie badanych użytkowników na dwie grupy: studentów studiów pierwszego stopnia, którzy dopiero zapoznają się z tematyką bibliotek cyfrowych, oraz studentów studiów drugiego stopnia, którzy specjalizują się w niej na studiach magisterskich. Ze względów stylistycznych stosowano wymiennie określenia „studia 1. stopnia” / „studia licencjackie” oraz „studia 2. stopnia” / „studia magisterskie”, choć według obowiązującej nomenklatury mówi się wyłącznie o stopniach studiów.

wiali model internetu, a nie biblioteki cyfrowej [Zhang 2008, p. 1331]. Uczestnicy badania najczęściej ujmowali internet w perspektywie technologicznej, jako najważniejsze jego elementy wskazując wówczas serwery, komputery i połączenia między nimi (czyli sieć), oraz w perspektywie funkcjonalnej, podkreślając istotność takich składników, jak sieć, użytkownicy, dane oraz linki [Rieh 2010, p. 168; Zhang 2008, p. 1334-1336]. Przykłady wizualnych reprezentacji tych modeli mentalnych w ujęciu technologicznym i funkcjonalnym przedstawiono na rysunkach 5 i 6. Prace te podobne są do rysunków studentów polskich, o czym będzie jeszcze mowa, choć respondenci amerykańscy przedstawiali model internetu, a nie biblioteki cyfrowej. Dla autora rysunku odzwierciedlającego techniczną wizję internetu najważniejsze są serwery, komputery i połączenia między nimi (sieć), natomiast w wizji funkcjonalnej na pierwszy plan wysuwają się sieć, użytkownicy, dane, linki.



Rysunki 5 i 6. Techniczna i funkcjonalna wizualizacja internetu [Zhang 2008, dok. elektr., p. 1335]

## Wizualizacje

Rysunek jako jedna z technik wizualizacyjnych znany jest w nauce od dawna, a w niektórych dyscyplinach, np. geometrii, stosuje się go od zarania. Sama wizualizacja rozumiana jest jako proces przekształcania informacji w formę wizualną umożliwiającą obserwację, nadawanie sensu i rozumienie informacji [What is Visualization?, dok. elektr.], jako metoda obliczeń (ang. *computing*) transformująca symboliczne w geometryczne, metoda zobaczenia niewidzialnego (Visualization in Scientific Computing 1987, dok. elektr.). Nie na darmo mówi się: „muszę to zobaczyć” czy chociaż „wyobrazić sobie”. Łączy się to z ogólną kondycją człowieka i dominacją zmysłu wzroku, na który przypada 70% wszystkich receptorów, jakimi dysponują poszczególne organy zmysłowe. Właśnie dlatego wizualizacja pozwala poznać sens oraz ułatwia znalezienie rozwiązania problemu, gdyż odwzorowuje rzeczywistość, zmusza do stworzenia modelu, abstrakcyjnej wizualnej metafory, a więc wymaga analizy treści. Stworzenie wizualizacji

poprzedzone jest procesem poznawczym, najczęściej połączonym ze zrozumieniem. Efektem tych działań jest także obraz tworzony w umyśle człowieka (ang. *mental image*), który jest reprezentacją fizycznego zewnętrznego świata. Schemat ten zwany jest w przypadku tworzenia i funkcjonowania baz danych modelem zewnętrznym i stanowi wyobrażenie użytkownika o tym, jak działa, a czasem także o tym, jak zbudowana jest baza<sup>3</sup>. Model zewnętrzny jest najczęściej nie w pełni spójny z modelami twórców: konceptualnym i implementacyjnym. Na różnice w modelu twórcy systemu w stosunku do wyobrażeń użytkownika zwrócił uwagę m.in. Norman, co zaprezentowano na rysunku 7.

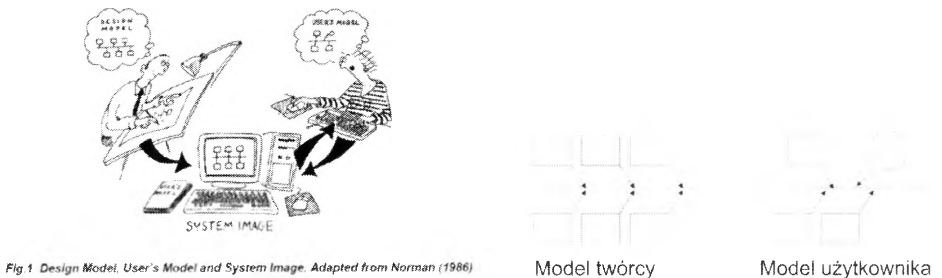


Fig. 1. Design Model, User's Model and System Image. Adapted from Norman (1986)

Rysunek 7. Skomplikowany model twórcy i prostszy użytkownika.  
Rysunek przetworzony za: [Makri et al 2007, p. 434; Norman 1986, p. 46]

Cel badań podjętych w Instytucie INIB UJ, polegający na analizie tego, czy i pod jakim względem modele interesariuszy są analogiczne do modeli specjalistów (model DELOS oraz 5S), można porównać do różnorodnych analiz prowadzonych przez badaczy zajmujących się wizualizacjami, a zwłaszcza wizualizacją informacji (ang. *infovis*), która pozwala na wykorzystanie interaktywnych reprezentacji obrazowych poszczególnych danych i obiektów w celu wzmocnienia ludzkich procesów poznania.

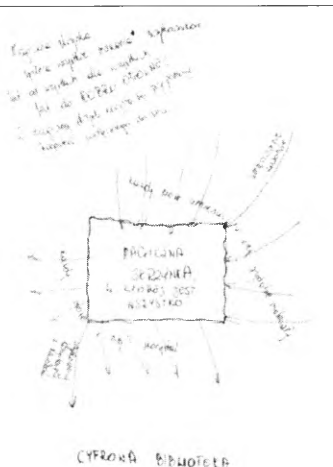

## Rodzaje rysunków

W wyniku badań nad wizualizacjami modeli mentalnych bibliotek cyfrowych otrzymano 186 rysunków<sup>4</sup> różnego rodzaju, które były dość abstrakcyjnymi porównaniami systemu do innych znanych obiektów, np. do przestrzeni cyfrowej,

<sup>3</sup> Należy przy tym zaznaczyć, że nie wszyscy klienci analizują struktury danych: niewielu użytkowników internetu rozmyśla nad tym, jak jest zbudowana baza wyszukiwarki Google.

<sup>4</sup> Autorki musiały niestety dokonać wyboru ilustracji, choć wszystkie prace były bardzo interesujące.

portretowały bibliotekarzy upodobnionych do cyborgów lub postaci z filmów SF bądź, zgodnie z wyobrażeniami autorów, przedstawiały BC np. jako wynik dematerializacji biblioteki tradycyjnej przez przeniesienie jej w całości (w tym budynku) do przestrzeni internetu. Zastosowany podział nie jest oczywiście podziałem rozłącznym – wiele rysunków można zaliczyć do różnych zarysowanych tu grup. Pierwsza typologia dzieli prace na: artystyczne, ilustracyjne i techniczne, druga odwołuje się do obowiązującej normy PN-ISO 10209-1:1994 „Dokumentacja techniczna wyrobu – Terminologia – Terminy dotyczące rysunków technicznych: ogólne i rodzaje rysunków”. Rysunkami ilustracyjnymi są wszystkie prace, gdyż każda z nich miała przedstawić lub objaśnić model mentalny – widać to na wybranych przykładach, będących bardzo ogólnymi, wręcz abstrakcyjnymi obrazami „panbiblioteki”, „wszechbiblioteki”, dobra ogólnego, które zawiera wszystkie informacje.

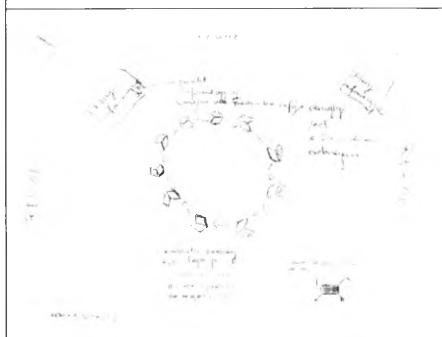
Rysunki ilustracyjne	
<i>Studia 1. stopnia</i>	<i>Studia 2. stopnia</i>
 <p>Rysunek 8. Biblioteka cyfrowa jako dobro ogólne, magiczna skrzynka, w której jest wszystko</p>	 <p>Rysunek 9. Biblioteka cyfrowa „wsysająca” wszystko: książki, Ziemię, kosmos = panbiblioteka</p>

Niektóre z prac są artystyczne<sup>5</sup>, czasem wręcz humorystyczne, jak te ukazane przykładowo na rysunkach 10-15. Wśród ilustracji artystycznych (z narysowanym komputerem, osobą, książką itp.) znalazły się prace dwojakiego rodzaju:

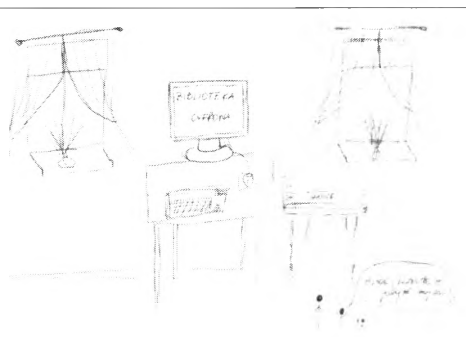
<sup>5</sup> Stworzone pomimo obaw badanych.

1) bardzo proste, szkicujące główną myśl, jaka nasunęła się autorowi w trakcie rysowania – np. rysunki studentów 1. stopnia, na których widać komputery wyglądające jak wianek z kwiatów, ustawione wokół okrągłego stołu (zgodnie z przetworzoną ideą arturiańską: wszyscy mają równe prawa i dostęp do informacji) lub „udomowioną” bibliotekę cyfrową, „urządzoną” w czyimś salonie oraz 2) bardziej rozbudowane pod względem liczby uwzględnionych obiektów / elementów – np. rysunki studentów 2. stopnia przedstawiające różnorodne dokumenty, które mogą być widoczne na ekranie komputera (minimalizm i symbolizm) lub czytelną jako sześciąt symbolizujący Matrix (z filmu braci Wachowskich).

*Prace studentów studiów 1. stopnia*

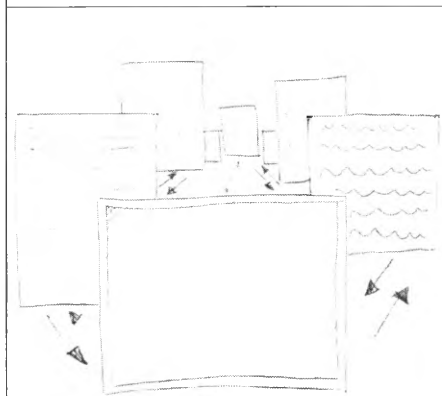


Rysunek 10. BC w przestrzeni pokoju;  
„okrągły stół”



Rysunek 11. „Udomowiona biblioteka” (zadziwiająca jest rozbieżność proporcji: komputer – użytkownik)

*Prace studentów studiów 2. stopnia*



Rysunek 12. Artystyczna wizja:  
„co stoi za ekranem”



Rysunek 13. Sześciąt Matriksa

Na rysunkach humorystycznych studenci ukazywali m.in. swoje wyobrażenia o rodzaju relacji łączącej wysoko wykwalifikowanego pracownika biblioteki i odbiorcę, nazwanego tu „uży-szkodnikiem”, samozadowolenie użytkownika (studenci st1st), a także akcentowali fakt, iż to twórcy wpadają na pomysły, które są akceptowane przez użytkowników pomimo ich skomplikowania, wreszcie przedstawiali bibliotekę cyfrową jako specyficzny bioniczny twór, który pożera, „zjada” wszystkie źródła informacji (st2st). Na jednej z prac umieszczono postać studenta, który po skorzystaniu z zasobów bibliotek cyfrowych z pewnością już zda egzamin<sup>6</sup>.

## Rysunki humorystyczne

### Studia 1. stopnia

### Studia 2. stopnia



Rysunek 14. Matrix z „uży-szkodnikiem”



Rysunek 15. Biblioteka cyfrowa „zjada” wszystkie zasoby informacyjne.

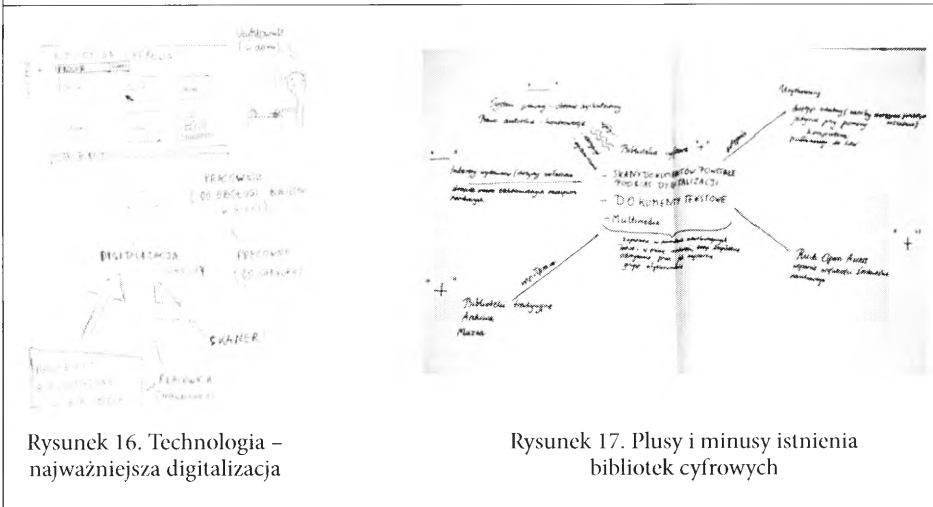
Część prac to również typowe diagramy, czyli uproszczone reprezentacje graficzne pewnych pomysłów, idei, wykorzystane do obrazowej reprezentacji wiedzy [Diagram, dok. elektr.], tak jak widać to na rysunkach 16 i 17, pokazujących najważniejsze elementy składające się na biblioteki cyfrowe.

<sup>6</sup> Należy zwrócić uwagę, że na rysunkach pojawiają się reprezentacje stanów afektywnych odnoszących się do samopoczucia autorów, postrzegania siebie i innych użytkowników albo jako nieakceptowanych odbiorców („uży-szkodników”, pasożytów), albo jako użytkowników bardzo zadowolonych, którzy po „wchłonięciu” wiedzy z biblioteki cyfrowej np. „na pewno zdadzą egzamin”. Na obecność takich elementów emocjonalnych w modelach kognitywnych zwracali uwagę inni badacze, m.in. Zhang, Saraccvic, Norman, Makri [Pan et al. 2006, dok. elektr.; Zhang 2008; Rieh et al. 2010].

## Diagramy

Studia 1. stopnia

Studia 2. stopnia



Rysunek 16. Technologia – najważniejsza digitalizacja

Rysunek 17. Plusy i minusy istnienia bibliotek cyfrowych

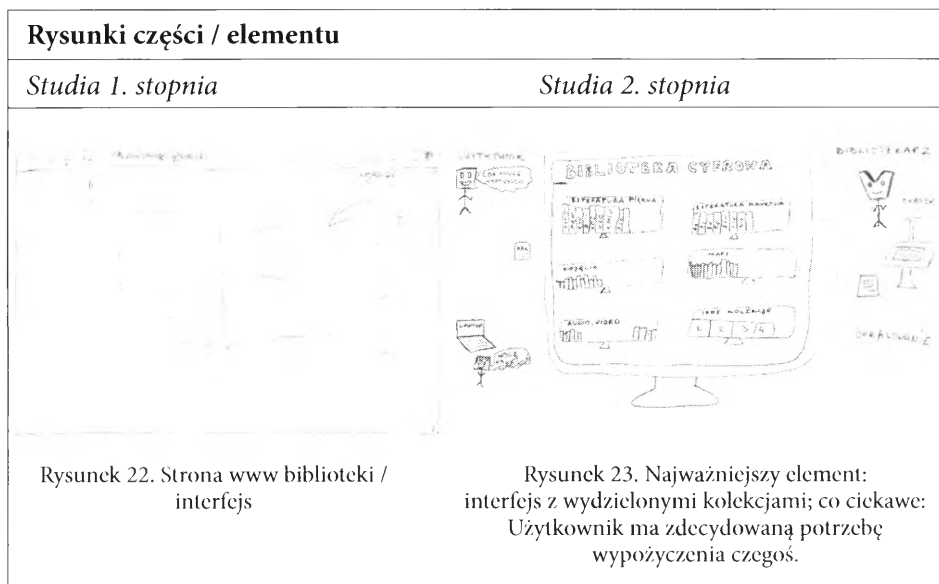
Z terminów określających typy rysunków technicznych, wymienionych we wspomnianej już normie PN-ISO 10209-1:1994, wybrano te, które można było zastosować do otrzymanych wizualizacji, a mianowicie: 1) układ: przedstawienie zarysów przedmiotów (jak na rysunkach artystycznych zaprezentowanych wcześniej), 2) przekrój (jak na rysunkach złożeniowych; oczywiście są to abstrakcyjne przekroje, czyli modele budowy i funkcjonalności BC), 3) szkic, 4) schemat – rysunek, w którym zastosowano symbole graficzne w celu pokazania funkcji części składowych obiektu i ich współzależności, 5) rysunek złożeniowy ogólny przedstawiający wszystkie zespoły i części całości, 6) rysunek części / elementu, 7) rysunek zestawu elementów, przedstawiający wymiary, sposób wyróżniania oraz zawierający dane dotyczące wykonania, 8) rysunek wykonawczy. Należy zauważyć, że nie są to ilustracje przygotowane na podziątkach, z dokładnym rozrysowaniem budowy, zastosowaniem skali itp.

Schematy przygotowane przez studentów to w przeważającej liczbie szkice organizacyjne prezentujące strukturę biblioteki cyfrowej, a w szczególności zależności pomiędzy jej komórkami funkcjonalnymi. W wyniku badań otrzymano głównie rysunki, na których uwzględniono elementy takie, jak: dane, sprzęt, oprogramowanie, odbiorcy, twórcy i sieci, a więc korelaty charakterystyczne dla każdego systemu informacyjno-wyszukiwawczego. Z drugiej strony pojawiły się też rysunki, na których studenci przedstawiali detale odnoszące się do procesu zarządzania biblioteką cyfrową oraz osoby odpowiedzialne za administrowanie systemem.





Rysunek części / elementu odnosi się, jak sama nazwa wskazuje, do poszczególnych składowych obiektu. Najczęściej pojawiającymi się komponentami BC były: komputery, interfejs, zasoby oraz użytkownicy określani jako „gracze”, co można skojarzyć z postrzeganiem sieci jako wirtualnej rzeczywistości lub gry komputerowej, w której jednostka aktywnie uczestniczy. Poszczególne elementy przedstawione na rysunkach wykonanych przez studentów 1. roku, odpowiednio, studiów licencjackich i studiów 2. stopnia, jak również przykładowe różnice w postrzeganiu biblioteki cyfrowej przez obie te grupy badanych zostały zobrażowane na rysunkach 22 i 23.



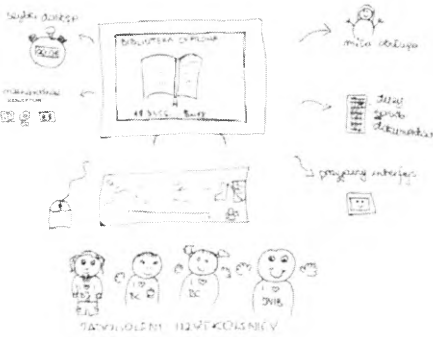
Na rysunkach mających postać zestawu elementów badani studenci pokazali zależności pomiędzy poszczególnymi składowymi BC, sygnalizując zarazem reakcje emocjonalne użytkownika, np. zadowolenie z korzystania z obiektów lub z miłej obsługi (st1st). Przedstawiali również ciąg zdarzeń od pojawienia się pomysłu stworzenia biblioteki cyfrowej do udostępnienia kolekcji wielu odbiorcom (z odzewem i interaktywną komunikacją, tzw. sprzężeniem zwrotnym) (st2st).

Rysunki wykonawcze odwzorowują sekwencje czynności towarzyszących tworzeniu biblioteki cyfrowej oraz niezbędne do tego sprzęty i aparaturę. Niektóre prace studentów st2st, w związku z posiadaną już przez nich wiedzą, a nawet doświadczeniem, dobrze oddały procesy związane z kreowaniem systemów, tzw. *workflow* biblioteki cyfrowej, przedstawiały pracownie digitalizacyjne lub profe-

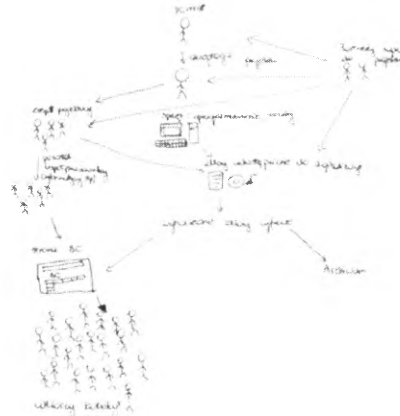
## Rysunki zestawu elementów

### Studia 1. stopnia

### Studia 2. stopnia



Rysunek 24. Interakcje i komponenty BC wzbudzające zadowolenie użytkowników: duży zasób dokumentów, różnorodność zbiorów, przyjazny interfejs, miła obsługa, szybki dostęp



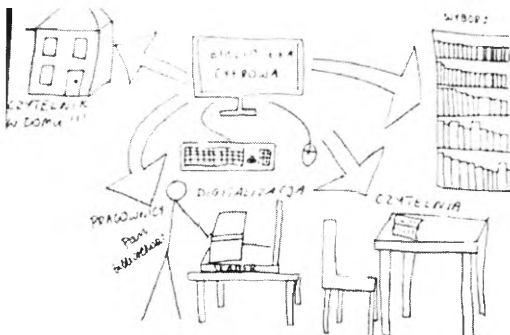
Rysunek 25. Aktorzy biblioteki cyfrowej: od pomysłu, przez partnerów projektu, do odbiorców kolekcji, przez sprzęt, oprogramowanie, wytworzone zbiory i stronę www

sjonalne skanery. Przykładowe elementy składające się na właściwie pojmowany system informacyjno-wyszukiwawczy, uwzględnione na szkicach przez badanych studentów st1st i st2st, odtworzono na rysunkach 26 i 27.

## Rysunki wykonawcze

### Studia 1. stopnia

### Studia 2. stopnia



Rysunek 26. Przestrzeń biblioteki z regałami, stolikami oraz pracownikami digitalizującymi zbiory



Rysunek 27. Od twórców, projektu, pracowników itp., do radości internautów (ang. workflow)

## Wyniki

W rezultacie przeprowadzonych badań otrzymano w sumie 186 odwzorowań indywidualnych modeli mentalnych, z czego 54 stanowiły rysunki studentów studiów 2. stopnia, specjalizujących się w zdobywaniu wiedzy na temat bibliotek cyfrowych w ramach specjalizacji ZC, a pozostałe 132 zostały wykonane przez studentów studiów 1. stopnia, którzy pozyskiwanie tej wiedzy dopiero rozpoczęli.

Analiza otrzymanych wyników badań, czyli wizualizacji modeli mentalnych studentów obu stopni studiów, prowadzi do następujących wniosków: wśród modeli mentalnych studentów 1. roku st1st, którzy nie posiadają jeszcze ugruntowanej wiedzy na temat bibliotek cyfrowych, przeważały dość abstrakcyjne porównania systemu BC np. do przestrzeni cyfrowej, źródła, które zbiera wszystko, czasem wręcz „zasysa” cały świat, lub do elementu świata gier czy wirtualnej rzeczywistości. Studenci st2st stworzyli bardziej złożone modele, często w formie diagramów lub nawet modeli wykonawczych. Schematy te odnosiły się głównie do aspektu funkcjonalnego oraz usługowego (ang. *processed view*) związanego z serwisami i usługami w bibliotece cyfrowej, jak również do aspektu technologicznego. Jest to zbieżne z wynikami uzyskanymi przez Yang Zhang (funkcjonalny i technologiczny ogląd internetu) [Zhang 2008].

Najważniejszym elementem strukturalnym figurującym w modelach referencyjnym i 5S, który badani przedstawili na swoich rysunkach, były zasoby (st1st – 83,33%; st2st – 98,15%). Wskazuje to, iż w ogólnym rozumieniu systemu zasoby traktowane są jako cała biblioteka cyfrowa, kompleksowa całość, a nie wyłącznie jako jeden z komponentów, czyli zawartość tego systemu. Często ta zawartość obrazowana była jako: teksty drukowane, filmy, grafiki, zdjęcia, płyty CD-ROM, serwisy, bazy danych, linki, zdigitalizowana treść (czasem zawartość = wszystko, co jest dostępne, jak na rysunkach 28 i 29), co można odnieść do wymiaru strumieniowego biblioteki cyfrowej w modelu 5S. Biblioteka cyfrowa pojmowana była także jako wynik całkowitej wirtualizacji biblioteki tradycyjnej (w tym budynku i jego zawartości, np. z szafami lub półkami bibliotecznymi – co można zauważyć na rysunku 30) przez przeniesienie jej przestrzeni internetu (wręcz tak jak w tezie: „wirtualne jest realnym”, jak można zauważyć na rysunku 31). To wyobrażenie pojawiło się w 26,52% prac studentów st1st oraz 14,81% prac studentów st2st. Ta dysproporcja w wynikach dla obu cykli studiów wskazuje, iż studenci specjalizacji ZC wiedzą więcej o mechanizmach funkcjonowania biblioteki cyfrowej, zaś definicje biblioteki cyfrowej / wirtualnej / hybrydowej nie mylą im się i nie nakładają znaczeniowo, a w związku z tym nie stanowią bariery w korzystaniu z systemów i w pojmowaniu istniejących między nimi różnic. Co ciekawe, tylko w pracach studentów studiów licencjackich zobrazowana została przestrzeń użytkownika:

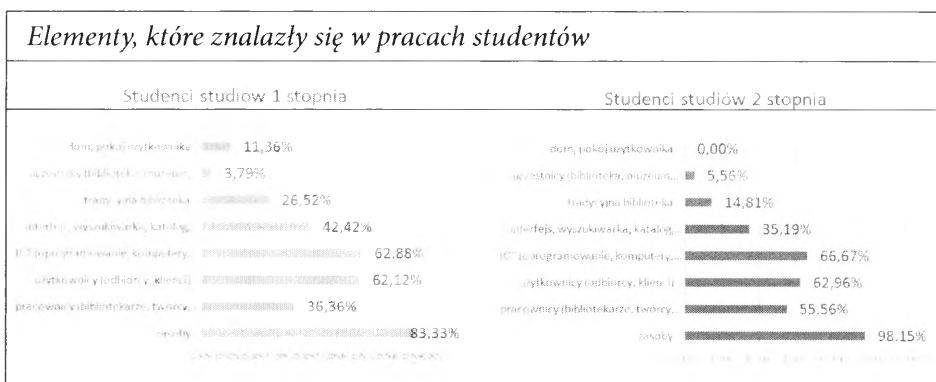
pokój, dom, wygodne krzesło i biurko (rysunek 32 – bardzo zadowolony użytkownik z kubkiem kawy), a raz słoneczna łąka z kocem (rysunek 33). Można na tej podstawie przyjąć, że dla członków badanej grupy bardzo ważnym elementem w procesach informacyjnych jest komponent afektywny, związany z emocjami towarzyszącymi zachowaniom informacyjnym i podejmowanym interakcjom człowiek – komputer.

Bardzo liczne modele mentalne odzwierciedlały postrzeganie biblioteki cyfrowej w perspektywie technologicznej, jako że komputery, sieć i oprogramowanie znalazły się na wizualizacjach 62,88% badanych z 1. roku studiów stacjonarnych oraz w pracach 66,67% studentów specjalizacji ZC. Podobnie interfejs / katalog / wyszukiwarkę ukazało na swoich rysunkach 42,42% studentów st1st i 35,19% studentów st2st. Na niektórych rysunkach w badanej zbiorowości wskazywano poszczególne komponenty bibliotek cyfrowych (np. serwery, bazy danych, sieć), – bez szczegółowego zwracania uwagi na funkcjonalność systemu, różnorodne procesy i interakcje w nim zachodzące – lub przyglądano się bibliotekom cyfrowym z perspektywy globalnego aktu komunikacji, łączącego poszczególne obszary biblioteki cyfrowej. Niekiedy można było także zaobserwować próby zaznaczenia relacji w ogólnym modelu (jak uwidoczniono to na rysunku 34). Więcej wzajemnych powiązań między komponentami systemu informacyjno-wyszukiwawczego przedstawili studenci studiów 2. stopnia. Dotyczy to zwłaszcza diagramów i rysunków wykonawczych, na których znalazły się terminy dotyczące funkcjonowania biblioteki cyfrowej, udostępniania serwisów. Terminy te można odnieść – w modelu referencyjnym – do wymiarów funkcjonalnego i aktywności użytkowników, zaś w modelu 5S – do obszarów przestrzennego, scenariuszowego i społecznościowego. Wśród zapisanych na ilustracjach sformułowań najczęściej występowały: digitalizowanie i implementowanie zasobów tradycyjnych, udostępnianie, wyszukiwanie, obsługa użytkowników, pozyskiwanie, wymiana międzybiblioteczna, rejestrowanie użytkownika, gromadzenie, przechowywanie, opracowanie, ochrona, katalogowanie, tagowanie, funkcja bibliograficzna, informacyjna i transfer danych (na jednym diagramie pojawił się wręcz schemat elektroniczny – rysunek 35). Wśród określeń dotyczących użytkownika pojawiały się natomiast: wyszukiwanie, poszukiwanie informacji niekoniecznie skonkretyzowanej, otrzymywanie informacji, czytanie, możliwość robienia notatek, drukowanie, ściąganie na dysk, zadowolenie oraz zdobywanie wiedzy.

Powtarzającym się elementem badanych szkiców były reprezentacje twórców, menedżerów i użytkowników bibliotek cyfrowych, przy czym ilościowy rozkład wskazań na każdą z tych grup był różny u studentów st1st i st2st. Spośród wszystkich interesariuszy bibliotek cyfrowych najczęściej wymieniani byli użytkownicy, zapewne rozumiani jako odbiorcy końcowi (klienci, konsumenci): kategoria ta

pojawiła się w 62,12% prac z grupy badawczej studentów studiów 1. stopnia oraz na 62,96% rysunków studentów studiów magisterskich. Natomiast pracownicy / twórcy / administratorzy / informatycy zostali wymienieni lub narysowani tylko na 36,36% prac studentów st1st oraz na 55,56% szkiców studentów realizujących specjalizację ZC. Według wyobrażeń utrwalonych na szkicach studentów użytkownicy zdobywają wiedzę, szukają, obserwują, są zadowoleni z korzystania z bibliotek cyfrowych, a bibliotekarze już posiadają określoną wiedzę, dzielą się nią, mają większe kwalifikacje i umiejętności (co można zauważyć na rysunku 36). Jeden raz pojawiło się porównanie opracowania zbiorów do nowego procesu twórczego, zaś na innej pracy studenta st2st rozdzielone zostały kompetencje twarde i miękkie, jako element doskonalenia zawodowego (co przedstawiono na rysunku 37). Wśród twórców biblioteki cyfrowej trzy razy zostali wymienieni także autorzy książek, zaś jeden raz pojawiły się nawet postaci z bajek (zaprezentowano to na rysunku 38).

Najbardziej obrazowane były instytucje współtworzące systemy: inne biblioteki, muzea, archiwa. Ich ważność uznało tylko 3,79% studentów st1st oraz 5,56% st2st. Biblioteka cyfrowa raczej kojarzona jest z biblioteką tradycyjną<sup>7</sup>, sporadycznie z siecią instytucji, a także ze współpracą z innymi instytucjami kultury. Badani studenci częściej traktowali biblioteki cyfrowe jako biblioteki hybrydowe lub biblioteki wirtualne, istniejące tylko w rzeczywistości sieciowej, cyberprzestrzeni, rzadziej natomiast przedstawiali wizję tworzenia bibliotek federacyjnych lub regionalnych.



Wykres 1 i 2. Komponenty bibliotek cyfrowych pojawiające się w modelach mentalnych studentów studiów licencjackich, 1. stopnia oraz magisterskich, 2. stopnia (realizujących specjalizację *Zasoby cyfrowe*)  
Źródło: badania własne, 2011-2014

<sup>7</sup> Podobne wnioski uzyskała Stephann Makri w swoich badaniach modeli mentalnych użytkowników bibliotek cyfrowych [Makri et al. 2008].

Dokładny rozkład elementów bibliotek cyfrowych zaobserwowany w modelach mentalnych badanych studentów przedstawiono na wykresach 1 i 2.

Przykładowe elementy bibliotek cyfrowych, które studenci tworzący modele mentalne uznali za na tyle istotne, by ująć je w swoich szkicach w postaci reprezentacji graficznych lub terminologicznych, zostały także przedstawione na rysunkach 28-39.

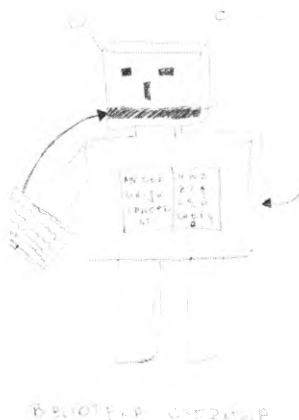
## Rysunki – analiza treściowa

### Studia 1. stopnia

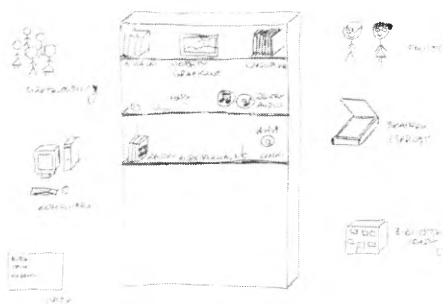


Rysunek 28. Biblioteka wirtualna zbierająca wszystko; st1st

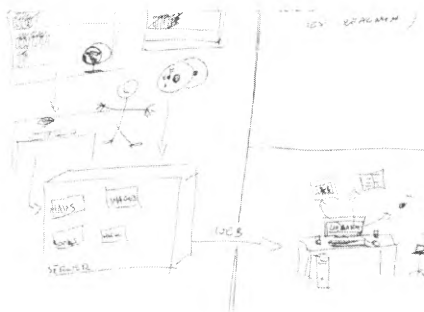
### Studia 2. stopnia



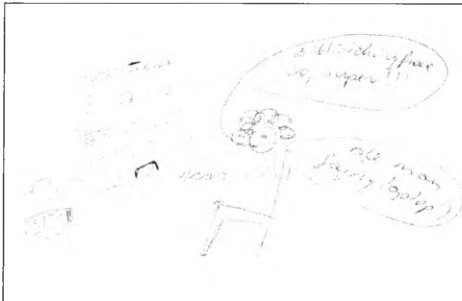
Rysunek 29. Abstrakcyjny model; st2st



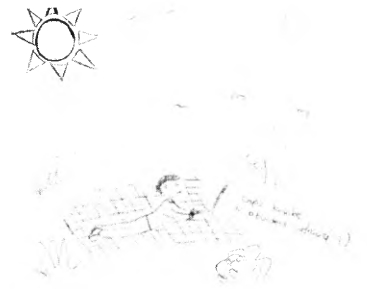
Rysunek 30. Biblioteka cyfrowa jako „szafa” z różnymi typami dokumentów oraz jej ważne elementy: użytkownicy, komputery, opisy oraz bibliotekarze, sprzęt i biblioteka tradycyjna; st2st



Rysunek 31. Wirtualne jest realnym; st1st



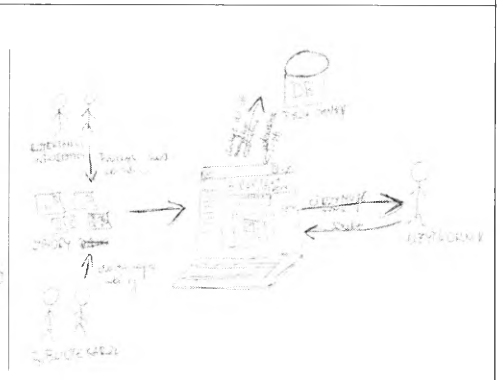
Rysunek 32. Samozadowolenie użytkownika; st1st



Rysunek 33. Wszędostęp; st1st



Rysunek 34. Komputeryzacja; st1st



Rysunek 35. Funkcjonalność; st1st

SCHEMAT BIBLIOTEKI CYFROWEJ



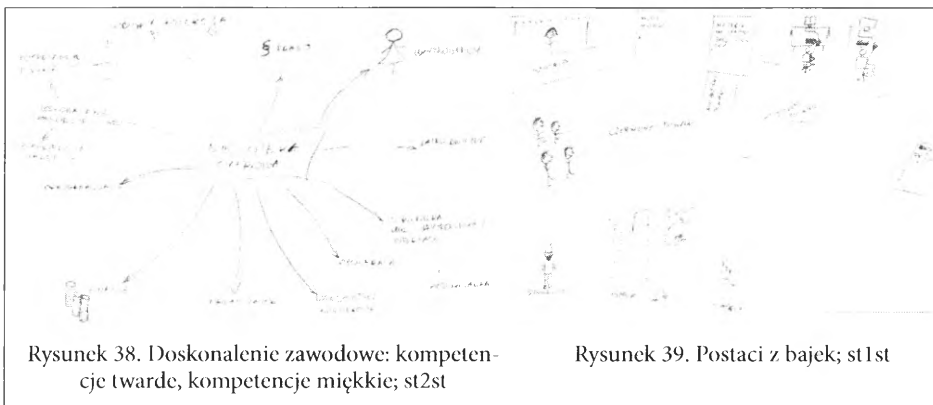
- +/- użytkownik/początek wyszukiwania
- ☑ przełącznik - wybór funkcji wysz.
- |—| Indeks tytułowy
- ☐ Indeks autorów
- ☒ Indeks nazwy
- ⊗ wyzniki wyszukiwania

Rysunek 36. Schemat elektroniczny; st1st



Rysunek 37. Model i jego elementy; Bardziej obeznany bibliotekarz; st1st





Rysunek 38. Doskonalenie zawodowe: kompetencje twarde, kompetencje miękkie; st2st

Rysunek 39. Postaci z bajek; st1st

## KORELACJA Z MODELEM DELOS

W odniesieniu do modelu referencyjnego sieci DELOS można zauważyć, iż najlepiej identyfikowanym jego elementem jest ogólnie pojmowana biblioteka cyfrowa (DL), która jako termin pojawiła się na większości rysunków. Trudno jednak twierdzić, iż jest to jeden z trzech elementów tworzących model DELOS, gdyż napis „biblioteka cyfrowa” zobaczyć można na narysowanych budynkach, w interfejsach lub np. w nagłówkach prac. Rzadziej pojawiał się termin „system biblioteki cyfrowej” (DLS), natomiast wyodrębniane były jego elementy: katalog, wyszukiwarka, interfejs, strona www, baza danych, sieć (jak opisano to we wcześniejszym podrozdziale). System zarządzania (DLMS), jako element najsilniej kojarzony z możliwościami samego oprogramowania, dzięki któremu zbudowano bibliotekę cyfrową, i z osobami nią zarządzającymi, został zwizualizowany przez wymienienie elementów dotyczących administrowania zasobami. Na żadnej z prac rysunkowych nie został odwzorowany idealny podział na trzy współistniejące części systemu, czyli bibliotekę cyfrową jako obiekt (DL), system (DLS) oraz system zarządzania (DLMS). Wydaje się więc, iż studenci patrzą na ten model jako na całość, współzależną od wymienionych komponentów.

Pozostałe składniki modelu referencyjnego są już lepiej kojarzone przez studentów i obrazowane na wizualizacjach. Jak już wspomniano, najważniejszą częścią biblioteki cyfrowej jest jej zawartość, którą na swoich rysunkach uwzględniło 83,33% studentów st1st i 98,15% studentów st2st (jak na rysunku 39, gdzie zasoby to cała armia). Reprezentacje różnych grup interesariuszy BC były obecne na 62,12% wszystkich rysunków studentów st1st i w 74,07% prac studentów st2st. Należy przy tym zaznaczyć, iż wielu autorów szkiców rozdzielało osoby współtworzące bibliotekę cyfrową na dwie grupy: twórców (administratorów, infor-

matyków) oraz bibliotekarzy, przygotowujących zasób (rysunek 40). Nie zawsze wszystkie zespoły pojawiały się w jednej pracy i ta uwaga dotyczy głównie rysunków studentów studiów 2. stopnia, którzy czasem przedstawiali administratorów, bibliotekarzy i informatyków, a pomijali użytkowników. Wyraźna specjalizacja programu kształcenia narzucająca spojrzenie z punktu widzenia twórców bibliotek cyfrowych oraz realizacja takich przedmiotów, jak: Tworzenie bibliotek cyfrowych, Ewaluacja bibliotek cyfrowych, Digitalizacja, obróbka i opis zdigitalizowanego zasobu, Projekt biblioteki cyfrowej, mogły spowodować wytworzenie specyficznego modelu mentalnego, charakterystycznego dla tej grupy użytkowników systemów (informatyków, administratorów, bibliotekarzy cyfrowych współtworzących bibliotekę), a zarazem nieuwzględnienie w modelu mentalnym użytkownika końcowego, dla którego powstaje taki system informacyjno-wyszukiwawczy. Twórcy rysunków, czyli studenci, bardziej wczuli się w rolę przyszłych twórców bibliotek cyfrowych.

Funkcjonalność została sprowadzona głównie do procesu wyszukiwania informacji i udostępniania danych (skanowania, pobierania zasobów), także przez ich dobre przygotowanie (organizowanie, kategoryzowanie) (rysunek 41). Cecha ta okazała się ważna dla 54,55% studentów st1st i 53,70% studentów st2st. Należy jednak zwrócić uwagę, iż pomimo mniejszej liczby prac studentów studiów 2. stopnia, w których znaleziono odniesienie do aspektu praktyczności, funkcjonalności, to właśnie ta grupa badawcza przygotowała znacznie bardziej pogłębione charakterystyki BC, wymieniając cechy, które są istotne dla użytkowników końcowych bibliotek cyfrowych. Cechy te to: powszechny dostęp, dostępność wielodzielnicowych źródeł zgromadzonych w jednym miejscu, szybkie dotarcie do wybranej publikacji o każdej porze, możliwość archiwizacji cennych zbiorów dziedzictwa narodowego, digitalizowanie zasobów tradycyjnych i implementowanie elektronicznych, mechanizm „zbierający” informację o użytkowniku w celu najlepszego dostosowania interfejsu do jego potrzeb oraz organizacja społecznego tagowania, rejestrowanie i kategoryzowanie użytkowników, a także ochrona, transfer danych itp.

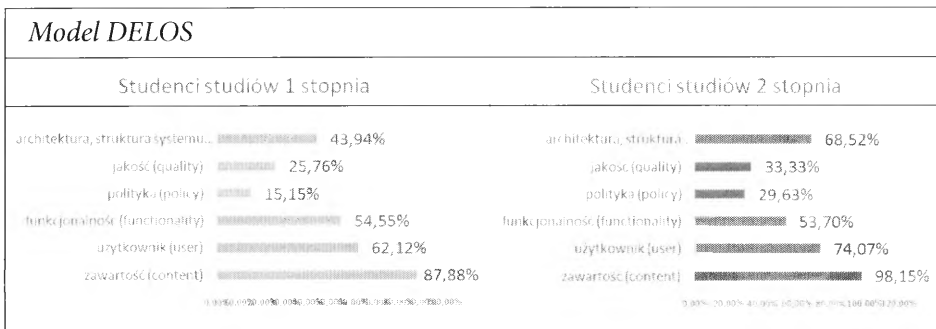
Znacznie lepiej została zidentyfikowana struktura systemu – połączona z funkcjonalnością – której zasadnicze znaczenie przypisało 43,94% studentów studiów licencjackich oraz 68,52% studentów studiów magisterskich. Różnica ta wynika najprawdopodobniej także z dysproporcji w posiadanej wiedzy, ukrytej i jawnej. Studenci studiów magisterskich specjalizujący się w tworzeniu bibliotek cyfrowych mają większe kompetencje i wiedzę więcej o procesie tworzenia zdigitalizowanych lub elektronicznych zbiorów, ich opracowywania, kategoryzowania i podziału kolekcji, tworzenia relacji odsyłaczowych itp. W modelach mentalnych studentów studiów 1. stopnia zarówno katalog, jak i kolekcje biblioteki cyfrowej

utożsamiane są z możliwościami zaawansowanych lub specjalistycznych wyszukiwarek (w taki sposób rysowane są np. interfejsy). Na rysunkach uzyskanych od tej grupy badanych właściwie funkcjonujący system informacyjno-wyszukiwawczy był określany dość enigmatycznie i ogólnie jako „dobre zorganizowanie zasobów” lub „spójna architektura serwisu” (rysunek 42).

Kolejną cechą wyróżnioną w trakcie analizy modeli mentalnych zaprezentowanych w postaci rysunków, którą można połączyć z wcześniej wymienionymi korelatami bibliotek cyfrowych, jest jakość, uwzględniona przez 25,76% badanych studentów st1st i przez 33,33% studentów st2st. Niezwykle ciekawe są użyte przymiotniki, które mają charakter opisowy i pojawiają się w formie słów kluczowych na grafikach prezentujących strukturę systemu. Studenci 1. roku studiów licencjackich charakteryzowali bibliotekę cyfrową za pomocą następujących określeń: intuicyjność serwisu, wygoda, szybkość, łatwy dostęp, różnorodność. Z kolei studenci 1. roku studiów magisterskich opisywali poszczególne elementy systemu, używając sformułowań: jakość zasobów, dostosowanie kolekcji do potrzeb i oczekiwań użytkowników (rysunek 43), aktualizacja kolekcji. Pojawiły się również abstrakcyjne określenia, jak choćby: „centrum wszechstronnej wiedzy” oraz „pozytywna emocja, radości internautów z możliwości korzystania z zasobów cyfrowych”. Wśród prac studentów studiów 1. stopnia jedna przedstawiała także problemy i ograniczenia, które negatywnie wpływają na jakość całego systemu. Prace takie wskazują na emocjonalny stosunek studentów do biblioteki cyfrowej (paradygmat afektywny badań użytkowników).

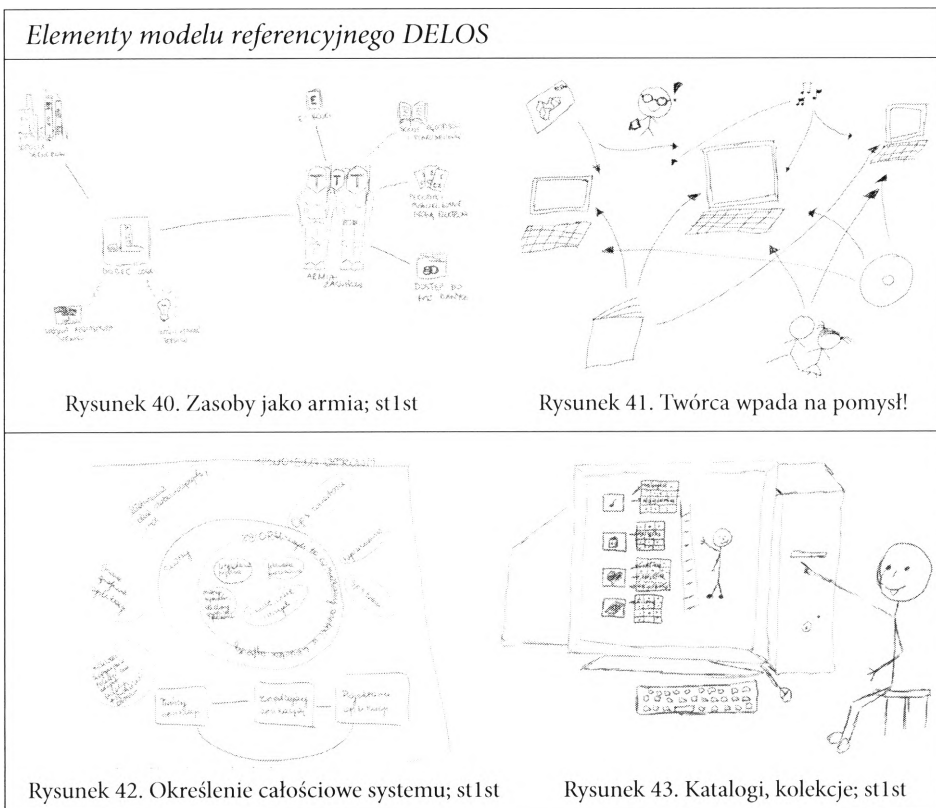
Polityka tworzenia bibliotek cyfrowych, która została wyeksponowana w modelu referencyjnym DELOS, miała zasadnicze znaczenie dla 15,15% studentów st1st i dla 29,63% studentów st2st. Należy wziąć pod uwagę ponownie różnice wynikające z posiadanej wiedzy i umiejętności, wyraźnej specjalizacji studentów, znajomości praktyki związanej ze sposobami tworzenia kolekcji (rysunek 44), przygotowywania projektów, świadomości specyfiki *workflow* itp. Studenci st2st wymienili w obszarze polityki biblioteki cyfrowej terminy takie, jak: pomysł, finanse, prawo, partnerzy, myślenie o przyszłości. Na rysunkach uwidoczniono także problemy związane z brakiem polityki Public Relations („jak dotąd za mało się o bibliotece cyfrowej mówi”), a także domniemany (przez autora szkicu) niski poziom wiedzy na temat korzystania z repozytoriów i bibliotek cyfrowych wśród osób zamieszkujących obszary wiejskie.

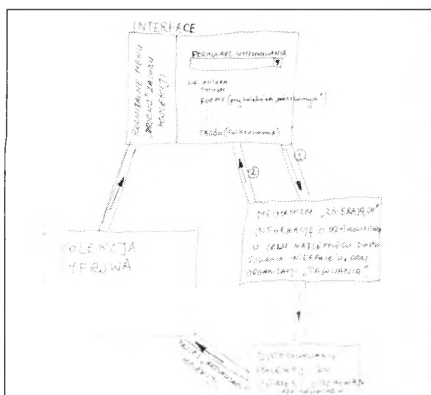
Poszczególne komponenty referencyjnego modelu biblioteki cyfrowej DELOS, które mają swoje odzwierciedlenie w strukturze mentalnej badanych użytkowników tych systemów informacyjno-wyszukiwawczych, zostały przedstawione na wykresach 3 i 4.



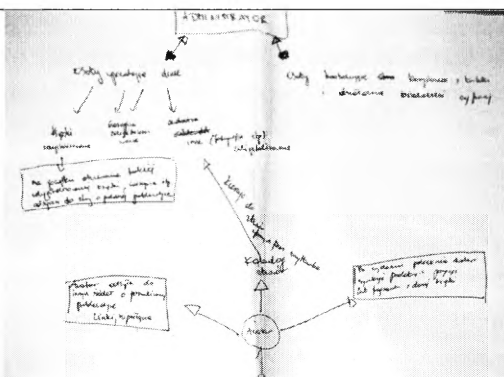
Wykres 3 i 4. Komponenty referencyjnego modelu bibliotek cyfrowych DELOS pojawiające się w modelach mentalnych studentów studiów licencjackich, 1. stopnia oraz magisterskich, 2. stopnia (realizujących specjalizację Zasoby cyfrowe)  
 Źródło: badania własne, 2011-2014

Przykładowe reprezentacje wizualne i terminologiczne korespondujące z referencyjnym modelem DELOS, które znalazły się na szkicach uczestników badania, zostały także przedstawione na rysunkach 39-44.





Rysunek 44. Mechanizmy: interfejs, mechanizm zbierający dane o użytkowniku ze sprzężeniem zwrotnym (feedback), dostosowanie kolekcji; st2st



Rysunek 45. W centrum awatar; nadzoruje administrator; st1st

## KORELACJA Z MODELEM 5S

Analizując wyniki przez pryzmat modelu 5S, można zauważyć, że studenci najczęściej rysowali komponenty, które można powiązać zarówno z obszarem strumieniowym, w tym zasoby i sieci, jak i z obszarem społecznościowym. W ich wizualizacjach pojawiły się należące do tych obszarów narzędzia informacyjno-komunikacyjne i elementy strukturalne systemów informacyjno-wyszukiwawczych. Wszystkie te elementy odnosiły się do użytkowników i do różnorodnych składowych systemu umożliwiających procesy komunikacji.

Na graficznych reprezentacjach modeli mentalnych, które można uznać za przybliżone odbicie modelu 5S, został uwzględniony komponent strumieniowy, który wskazało 86,36% studentów st1st i 94,4% studentów st2st. Obszar ten został ukazany zarówno przez zobrazowanie zawartości (rysunek 45), struktury, katalogów, jak i poszczególnych typów dokumentów w bibliotece cyfrowej. Wagę wewnętrznej i zewnętrznej organizacji systemu informacyjno-wyszukiwawczego BC, rozumianego jako zwarty kompleks powyższych elementów, dostrzegło i wyraziło na swoich rysunkach 68,52% studentów st2st i znacznie mniej, bo 43,94%, studentów st1st. Studenci studiów magisterskich zwrócili uwagę w swoich pracach na kategoryzowanie wiedzy, organizowanie zasobów, metadane itp. Studenci studiów 1. stopnia częściej wskazywali natomiast na aspekty przestrzenne i scenariuszowe, które zauważali także uczestnicy specjalizacji ZC, jednak w mniejszym stopniu.

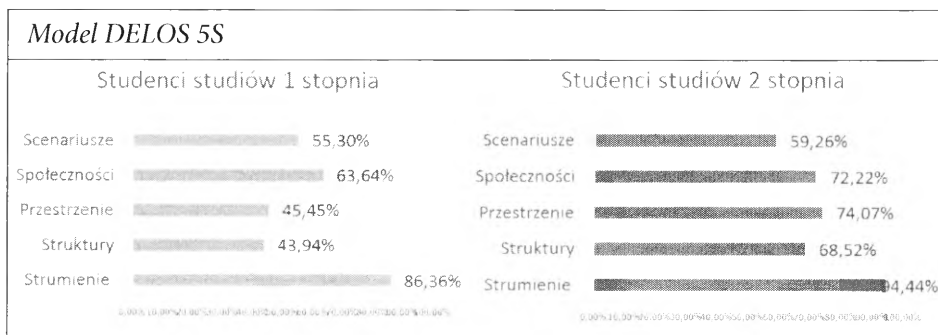
Różnego rodzaju procesy: indeksowania, przeglądania, wyszukiwania, możliwe do realizowania dzięki różnym strukturom danych (rysunek 46), wskazano na 45,45% rysunków studentów st1st i 74,07% prac studentów st2st. Ponownie zauważalna znacząca różnica w treści wizualizacji wynika najprawdopodobniej ze specjalistycznej wiedzy studentów studiów magisterskich, którzy już znają różne sposoby przygotowywania i opracowywania danych oraz organizacji informacji i zarządzania nią. Modele mentalne studentów studiów 1. stopnia okazały się natomiast zdominowane przez kojarzenie biblioteki cyfrowej z wyszukiwarką: autorzy szkiców identyfikowali procesy informacyjne zachodzące w systemie biblioteki cyfrowej z prostym wyszukiwaniem tekstowym za pomocą haseł i słów kluczowych. W taki właśnie dość nieskomplikowany sposób zobrazowane zostały: interfejs (rysunek 47), wyszukiwarka i katalog biblioteki cyfrowej. Część ilustracji zawierała elementy odnoszące się do domeny usługowej systemu, czyli obszaru scenariuszy w modelu 5S. Bibliotekę cyfrową studenci rozumieją jako serwis umożliwiający różnego rodzaju procesy biblioteczne (przeniesienie do przestrzeni wirtualnej tradycyjnych działań wymiany międzybibliotecznej, katalogowania, wypożyczania – rysunek 48). Na wizualizacjach pojawiły się także wyrażenia określające bibliotekę cyfrową w terminach abstrakcyjnych, jako przestrzeń sieciową, cyfrową, wirtualną, która umożliwia dotarcie do informacji w sposób nieokreślony. Co ważne, studenci studiów 2. stopnia lepiej niż ich młodsi koledzy rozumieli zachowania informacyjne, złożoność procesu docierania do informacji (rysunek 49) oraz ważność pertynencji informacji.

Wśród istotnych komponentów systemu BC badani studenci wyszczególnili aspekt społeczności (model 5S: *societies*), który został zwizualizowany przez 63,64% studentów st1st i przez 72,22% studentów st2st. Jak zauważono wcześniej, na rysunkach studentów pojawili się wszyscy interesariusze bibliotek cyfrowych: administratorzy, twórcy, bibliotekarze cyfrowi. Najczęściej wymieniano użytkowników końcowych, twórców, zaś najrzadziej – deweloperów i menadżerów (rysunek 50). Wśród prac magistrantów wyróżnić należy jeden rysunek, na którym zaznaczono aspekt badania potrzeb użytkowników jako niezbędną część implementacji bibliotek cyfrowych, a także kilka prac podkreślających aspekt afektywny – satysfakcję użytkownika.

Poszczególne komponenty oraz obszary charakterystyczne dla matematycznego modelu biblioteki cyfrowej 5S, które znalazły odzwierciedlenie w strukturze mentalnej badanych użytkowników tych systemów informacyjno-wyszukiwawczych, zostały przedstawione na wykresach 5 i 6.

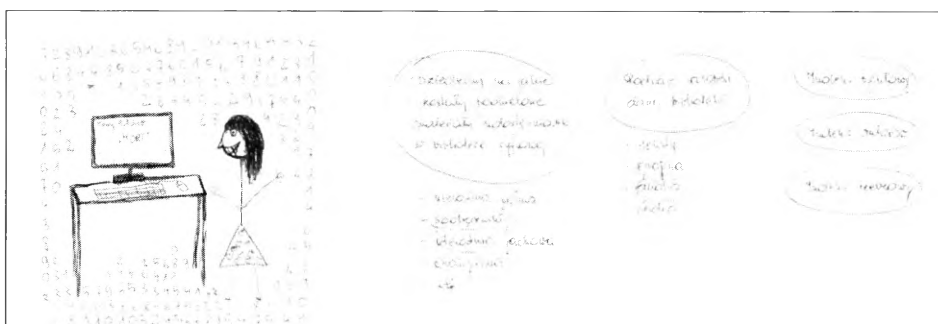
Przykładowe reprezentacje wizualne i terminologiczne zbieżne z modelem 5S, które znalazły się w pracach badanych uczestników (z podziałem na studentów studiów 1. i 2. stopnia), zostały także przedstawione na rysunkach 45-50.

## Model DELOS 5S

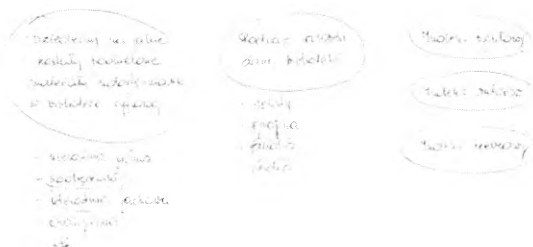


Wykres 5 i 6. Komponenty modelu bibliotek cyfrowych 5S, pojawiające się w modelach mentalnych studentów studiów licencjackich, 1. stopnia oraz magisterskich, 2. stopnia (realizujących specjalizację *Zasoby cyfrowe*).

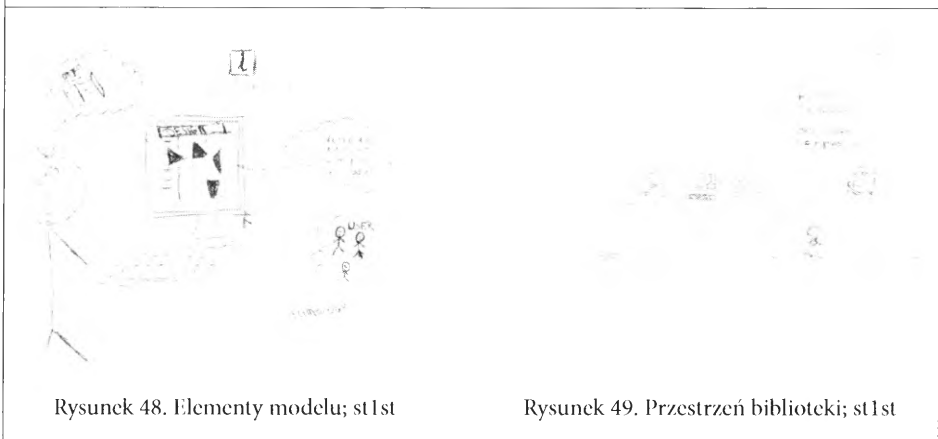
Źródło: badania własne 2011-2014



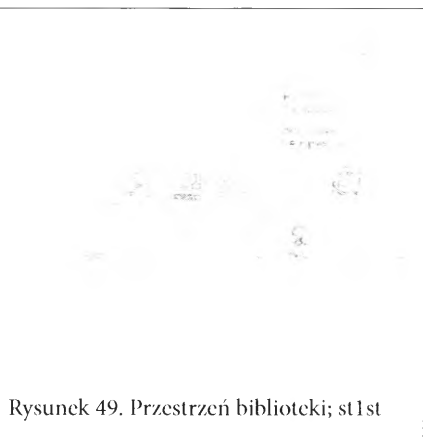
Rysunek 46. Zasoby, czasem niebezpieczne (*à propos* treści książki, która mówi o Śmierci potrzebującej wakacji, wybierającej chłopca na zastępstwo, Morta, niezwykle pragnącego zdobyć wiedzę, której stanowczo nie powinien posiadać); st1st



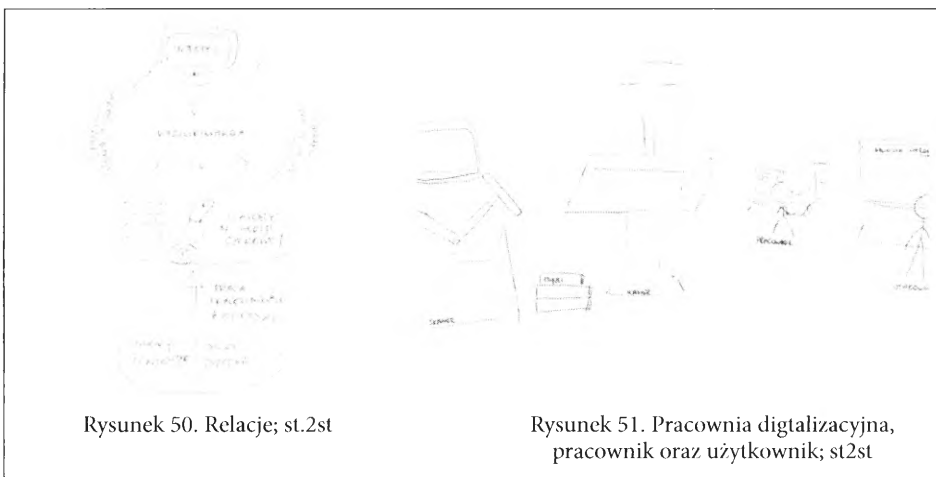
Rysunek 47. Katalog; st2st



Rysunek 48. Elementy modelu; st1st



Rysunek 49. Przestrzeń biblioteki; st1st



Rysunek 50. Relacje; st.2st

Rysunek 51. Pracownia digitalizacyjna, pracownik oraz użytkownik; st2st

## WNIOSKI

Badanie wizualizacji i modelowania mentalnego bibliotek cyfrowych jest niezwykle ważne w przypadku tworzenia ogólnej struktury dla tych systemów. Zarówno twórcy systemu (projektanci oraz osoby zarządzające i administrujące systemem, bibliotekarze, informatycy), jak i jego użytkownicy końcowi tworzą własne wyobrażenia nie tylko o sobie nawzajem, ale także o wszelkiego rodzaju komponentach budujących bibliotekę cyfrową – o obiektach i narzędziach, z którymi się stykają. Te umysłowe reprezentacje są zwykle oparte na wiedzy o elementach systemu i ich wzajemnych interakcjach, o realizowanych przez nie funkcjach, a także o zdarzeniach występujących podczas pracy z serwisem. Przebieg relacji człowiek – komputer jest również wypadkową zdobytych w trakcie pracy doświadczeń oraz intuicyjnych wyobrażeń o architekturze systemu i procesach informacyjnych podejmowanych w trakcie oddziaływania użytkownika z biblioteką cyfrową.

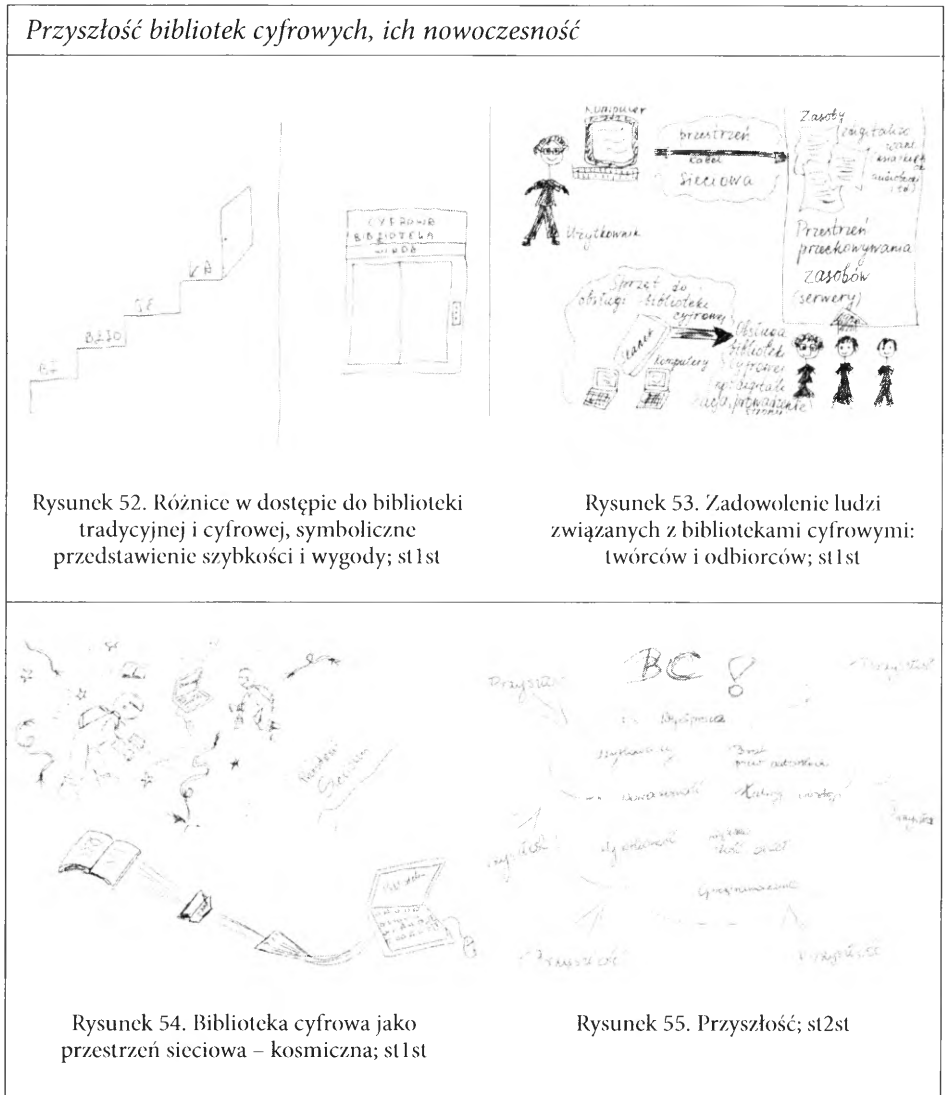
Analiza i interpretacja danych uzyskanych za pomocą zastosowanej metody rysunków odręcznych pozwoliła na określenie i dokonanie oceny modeli mentalnych otrzymanych w trakcie badań. Rysunki obrazowały przede wszystkim perspektywę techniczną (zasoby, system komputerowy, sieci itp. oraz ich użytkownicy) oraz funkcjonalną (akcent na wyszukiwanie informacji). Tylko nieliczne prace demonstrowały dokładne rozumienie struktury biblioteki cyfrowej, z uwzględnieniem perspektywy procesu (biblioteka cyfrowa jako centrum, złożona z poszczególnych elementów), a wśród nich zaledwie pojedyncze prezentowały szczegółowe zależności między komponentami budującymi cały system informacyjno-wyszukiwawczy BC.



Wszystkie wizualizacje przeanalizowano pod kątem ich zgodności z dwoma uprzednio wytypowanymi modelami odniesienia, czyli DELOS i 5S. Obydwa, pomimo immanentnej różnicy (informacyjny model DELOS *versus* informatyczny 5S), pozwoliły na pogrupowanie treści prac oraz określenie ogólnego modelu mentalnego osób z mniejszą lub większą wiedzą merytoryczną na temat bibliotek cyfrowych: studentów studiów 1. stopnia oraz 2. stopnia (specjalizujących się podczas studiów magisterskich w dziedzinie bibliotek cyfrowych, przyszłych bibliotekarzy cyfrowych, twórców i administratorów tych systemów informacyjno-wyszukiwawczych). Dla rozrysowania podstawowej struktury systemu biblioteki cyfrowej, w postaci choćby bardzo ogólnej abstrakcji wizualnej, potrzeba, rzecz jasna, elementarnej wiedzy o tych systemach, jak dowiodły tego np. prace studentów studiów 1. stopnia. W wyniku badań nie otrzymano schematu zawierającego wszystkie elementy zidentyfikowane przez specjalistów, profesjonalistów i praktyków. Główne różnice w stosunku do modelu referencyjnego DELOS i 5S, dostrzeżone w pracach studentów specjalizacji *Zasoby cyfrowe*, dotyczyły kompletności oraz szczegółowości graficznego reprezentowania struktur bibliotek cyfrowych – ich autorzy stosowali bardziej fachowe pojęcia w podpisach swoich rysunków. Studenci studiów 2. stopnia w szczególności zwracali uwagę na takie elementy modelu biblioteki cyfrowej, jak: struktury danych, np. metadane, układ zasobów oraz relacje między poszczególnymi informacjami, jakość, a także polityka tworzenia kolekcji. Uzyskany w wyniku badań obraz preferencji i sposobów rozumienia bibliotek cyfrowych każe także zwrócić uwagę na odmienność w interpretowaniu biblioteki cyfrowej przez studentów. Studenci st1st częściej łączyli biblioteki cyfrowe z rzeczywistym, tradycyjnym wizerunkiem biblioteki, co może wskazywać na skojarzenia z pojęciem biblioteki hybrydowej. Studenci tej grupy badawczej nie posiadali i nie zaprezentowali na wizualizacjach dokładnej wizji biblioteki cyfrowej, a w ich wyobrażeniach nakładały się na siebie różne serwisy internetowe, wyszukiwarki, bazy danych lub wręcz pojawiły się w nich obrazy internetu jako „panbiblioteki”.

Wszyscy badani studenci jako najważniejszy element całego uniwersum biblioteki cyfrowej – według modelu referencyjnego – wskazali jej zasoby, postrzegając pozostałe systemy (DLS i DLMS) jako jeden element (oprogramowanie, sprzęt, sieć), będący koniecznym dodatkiem do zasobów, czasem traktowany w domyśle jako jeden z komponentów biblioteki cyfrowej. Obserwując graficzne i terminologiczne wyobrażenia interesariuszy BC, można z łatwością dostrzec, że studenci redukują ich do użytkowników końcowych, a zapominają o innych osobach zaangażowanych w tworzenie biblioteki i zarządzanie nią. W odniesieniu do modelu 5S badana grupa studentów analogicznie wskazała na najważniejsze komponenty: strumieniowy i społecznościowy.

Wiedza uzyskana w wyniku badań użytkowników bibliotek cyfrowych może być w sposób wielokierunkowy i kompleksowy wykorzystana w procesie projektowania i doskonalenia tych systemów, szczególnie w zakresie tworzenia modelu wdrożeniowego, odpowiadającego oczekiwaniom i potrzebom interesariuszy. Studenci, jako obecni użytkownicy i przyszli twórcy, wyobrazili sobie bowiem przyszłość bibliotek, ekspansję środowiska cyfrowego, elementy i uwarunkowania ich rozwoju, co można zauważyć na ostatnich już zaprezentowanych ilustracjach (rysunki 51-54).



## BIBLIOGRAFIA

- Boyack, Kevin W.; Wylie, Brian N.; Davidson, George S. (2002). *Information Visualization, Human-Computer Interaction, and Cognitive Psychology: Domain Visualizations*. Visual Interfaces to Digital Libraries Lecture Notes in Computer Science Vol. 2539, p. 145-158.
- Brandt, D. Scott (2001). *Reference, Mental Models and Teaching Technology*. The Reference Librarian Vol. 74, p. 37-48.
- Brandt, D. Scott; Uden, Lora (2003). *Insight into mental models of novice Internet searchers*. Communications of the ACM Vol. 46, Issue 7, p. 133-136.
- Candela, Leonardo et al. (2007). *Setting the foundations of Digital Libraries* [dok. elektr.]. D-Lib Magazine Vol. 13, No. 3-4. <http://www.dlib.org/dlib/march07/castelli/03castelli.html> [odczyt: 19.08.2014].
- Candela, Leonardo et al. (2011). *The digital library reference model* [dok. elektr.]. <https://www.coar-repositories.org/files/D3-2b-Digital-Library-Reference-Model.pdf> [odczyt: 2.09.2014].
- Casarosa, Vittore (2007). *DELOS reference model for digital libraries* [dok. elektr.]. <http://elag2007.upf.edu/papers/casarosa.pdf> [odczyt: 2.09.2014].
- Chen, Ya-Ning; Ke, Hao-Ren (2014). *A Study on Mental Models of Taggers and Experts for Article Indexing Based on Analysis of Keyword Usage*. Journal of the Association for Information Science and Technology Vol. 65, No. 8, p. 1675-1694.
- Chowdhury, Gobinda (2010). *From digital libraries to digital preservation research: the importance of users and context*. Journal of Documentation Vol. 66, No. 2, p. 207-223.
- Cole, Charles et al. (2007). *A classification of mental models of undergraduates seeking information for a course essay in history and psychology: preliminary investigations into aligning their mental models with online thesauri*. Journal of the American Society for Information Science and Technology Vol. 58, No.13, p. 2092-2104.
- Denham, Pearl (1993). *Nine- to fourteen-year-old children's conception of computers using drawings*. Behavior and Information Technology Vol. 12, Issue 6, p. 346-358.
- Efthimiadis, Efthimis N.; Hendry, David G. (2005). *Search engines and how students think they work*. Proceedings of ACM SIGIR'05. ACM: New York, p. 595-596.
- Efthimiadis, Efthimis N.; Hendry, David G. (2008). *Conceptual Models for Search Engines*. Information Science and Knowledge Management Vol. 14, p. 277-307.
- Fox, Edward A.; Gonçalves, Marcos A.; Shen, Rao (2012). *Theoretical foundations for digital libraries: the 5S (societies, scenarios, spaces, structures, streams) approach*. California, San Raphael: Morgan & Claypool Publishers, 180 p.
- Fuhr, Norbert (2010). *Evaluation of digital libraries*. International Journal of Digital Libraries [dok. elektr.]. <http://www.scribd.com/doc/185523/Evaluation-of-Digital-Libraries-Fulltext> [odczyt: 2.07.2014].
- Gombrich, Ernst (1990). *Obraz wizualny*. W: Głowinski, Michał red. *Symbole i symbolika*. Warszawa: Czytelnik, s. 312-338.
- Gonçalves, Marcos André; Fox, Edward A. (2002). *5SL: a language for declarative specification and generation of digital libraries*. Proceedings of the 2<sup>nd</sup> ACM/IEEE-CS Joint Conference on Digital Libraries (Portland, Oregon, USA, July 14-18, 2002). New York: ACM, p. 263-272.

- Gonçalves, Marcos André Et al. (2004). *Streams, Structures, Spaces, Scenarios, Societies (5S): A formal model for digital libraries* [dok. elektr.]. <http://www.dlib.vt.edu/projects/5S-Model/5s6.pdf> [odczyt: 20.08.2014].
- Gonçalves, Marcos André; Fox, Edward A.; Watson, Layne T. (2008). *Towards a digital library theory: a formal digital library ontology*. *International Journal on Digital Libraries* Vol. 8, No. 2, s. 91-114.
- Gray, Susan H. (1990). *Using Protocol Analyses and Drawings to Study Mental Model Construction during Hypertext Navigation*. *International Journal of Human-Computer Interaction* Vol. 2, p. 359-377.
- Green, Thomas R. G. (1990). *Limited theories as a framework for human-computer interaction*. In: Ackermann, David; Tauber, Michael J. eds. *Mental models and human-computer interaction*. Amsterdam: North-Holland Publishing Co., p. 3-39.
- Gwizdka, Jacek (2010). *Distribution of Cognitive Load in Web Search*. *Journal of the American Society for Information Science and Technology* Vol. 61, Issue 11, p. 2167-2187.
- Infovis.org = Information Visualisation Resources (1995-). *What is Visualization?* [dok. elektr.]. <http://www.infovis.org/> [odczyt: 19.03.2014].
- Janiak, Małgorzata (2012). *Biblioteka cyfrowa, biblioteka elektroniczna, biblioteka wirtualna*. W: Janiak, Małgorzata; Krakowska, Monika; Próchnicka, Maria red. *Biblioteki cyfrowe*. Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, s. 15-65.
- Johnson-Laird, Philip N. (1983). *Mental models: Toward a cognitive science of language, influence, and consciousness*. Cambridge, England: Cambridge University Press, 513 p.
- Katzeff, Cecilia (1990). *System demands on mental models for a fulltext database*. *International Journal of Man-Machine Studies* Vol. 32, Issue 5, p. 483-509.
- Kerr, Stephen T. (1990). *Wayfinding in an electronic database: The relative importance of navigational cues vs. mental models*. *Information Processing & Management* Vol. 26, Issue 4, p. 511-523.
- Khoo, Michael; Hall, Catherine (2012). *What Would 'Google' Do? Users' Mental Models of a Digital Library Search Engine*. In: Zaphiris, Panayiotis et al. eds. *TPDL 2012, LNCS 7489*. Berlin: Springer-Verlag, p. 1-12.
- Krakowska, Monika (2012). *Modele mentalne użytkowników bibliotek cyfrowych*. W: Janiak, Małgorzata; Krakowska, Monika; Próchnicka, Maria red.. *Biblioteki cyfrowe*. Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, s. 309-329.
- Kress, Gunther; Van Leeuwen, Theo (2006). *Reading Images. The Grammar of Visual Design*. London: Routledge, 312 p.
- Latour, Bruno (1983). *Visualisation and Cognition: Drawing Things Together*. In: Kuklick, Henrika ed. *Knowledge and Society Studies in the Sociology of Culture Past and Present* [dok. elektr.]. Greenwich, Connecticut: JAI Press Inc. Vol. 6, p. 1-40. <http://worldream.com/refs/Latour%20-%20Visualisation%20and%20Cognition.pdf> [odczyt: 19.08.2014].
- Lee, Jiunde; Boling, Elizabeth (2008). *Information-Conveying Approaches and Cognitive Styles of Mental Modeling in a Hypermedia-Based Learning Environment*. *Journal of the American Society for Information Science and Technology* Vol. 59, Issue 4, p. 644-661.
- Literat, Ioana (2013). *A Pencil for Your Thoughts: Participatory Drawing as a Visual Research Method with Children and Youth*. *International Journal of Qualitative Methods* Vol. 12, p. 84-98.

- Makri, Stephann et al. (2007). *A library or Just another information resource? A case study of users' mental models of traditional and Digital libraries*. Journal of the American Society for Information Science and Technology Vol. 58, Issue 3, p. 433-445.
- Marcin Drabek (2010). *Dominacja wzroku. Kryzys pisma po zwrocie wizualnym* [dok. elektr.]. Kultura-Historia-Globalizacja Nr 8, s. 37-46. <http://www.khg.uni.wroc.pl/files/5khg8drabek.pdf> [odczyt: 12.08.2014].
- Murthy, Uma et al. (2007). *Extending the 5S Digital Library (DL) Framework: From a Minimal DL towards a DL Reference Model* [dok. elektr.]. In: *1<sup>st</sup> Workshop on Digital Library Foundations*, ACM IEEE Joint Conference on Digital Libraries. June 18-23, 2007, Vancouver, British Columbia, Canada. <http://www.dlib.vt.edu/projects/5S-Model/extending5S.pdf> [odczyt: 10.08.2014].
- Nęcka, Edward; Orzechowski, Jarosław; Szymura, Błażej (2008). *Psychologia poznawcza*. Warszawa: Academica Wydawnictwa SWPS; Wydawnictwo Naukowe PWN, 740 s.
- Norman, Donald A. (1986). *Cognitive Engineering*. In: Norman, Donald; Draper, Stephen W. eds. *User Centered System Design*. Hillsdale NJ: Lawrence Erlbaum, p. 31-62.
- Pan, Bing et al. (2006). *Bridging the gap: A conceptual model of the access of digital libraries* [dok. elektr.]. Journal of Digital Information Vol. 7, Issue 2. <http://www.panb.people.cofc.edu/pan/Pan-missingLinkPaper.pdf> [odczyt: 19.08.2014].
- Papastergiou, Marina (2005). *Students' mental models of the Internet and their didactical exploitation in informatics education*. Education and Information Technology Vol. 10, Issue 4, p. 341-360.
- Parsons, Paul; Sedig, Kamran (2014). *Adjustable Properties of Visual Representations: Improving the Quality of Human-Information Interaction*. Journal of the Association for Information Science and Technology Vol. 65, Issue 3, p. 455-482.
- Pejtersen, Annelise Mark; Fidel, Raya (1998). *A framework for work centered evaluation and design: A case study of information retrieval on the Web* [dok. elektr.]. Working Paper for MIRA workshop, Grenoble. <http://www.dcs.gla.ac.uk/mira/workshops/grenoble/fp.pdf> [odczyt: 19.08.2014].
- Rieh, Soo Young et al. (2010). *Conceptualizing institutional repositories work: Using co-discovery to uncover mental models*. In: *Proceedings of the third symposium on Information interaction in context*. New York: ACM, p. 165-174.
- Roda, Claudia et al. (2005). *Digital image library development in academic environment: designing and testing usability*. OCLC Systems & Services: International digital library perspectives Vol. 21, Issue 4, p. 264-284.
- Roth, Sandra P. (2010). *Mental models for web objects: where do users expect to find the most frequent objects in online shops, news portals, and company web pages*. Interacting with Computers Vol. 22, p. 140-152.
- Sheard, Judy I.; Ceddia, Jason D. (2004). *Conceptualisation of the web and disorientation, AusWeb X: The Challenge of Integration as the Web Moves Into Its Second Decade* [dok. elektr.]. Proceedings of the Tenth Australian World Wide Web Conference, 03 July 2004 to 07 July 2004, Southern Cross University, Lismore NSW Australia, p. 362-367. <http://ausweb.scu.edu.au/aw04/papers/refereed/ceddia/paper.html> [odczyt: 19.08.2014].
- Thatcher, Andrew; Greyling, Mike (1998). *Mental models of the internet*. International Journal of Industrial Ergonomics Vol. 22, No. 4-5, p. 299-305.

- Theron, Linda; Mitchell, Claudia; Smith, Ann; Stuart, Jean eds. (2011). *Picturing Research. Drawing as Visual Methodology* [dok. elektr.]. Rotterdam: Sense Publishers, 260 p. <https://www.sensepublishers.com/media/757-picturing-research.pdf> [odczyt: 19.08.2014].
- Visualization in Scientific Computing* (1987) [dok. elektr.]. <http://www.sci.utah.edu/vrc2005/McCormick-1987-VSC.pdf> [odczyt: 19.08.2014].
- Ware, Colin (2004). *Information Visualization: Perception for Design* [dok. elektr.]. <http://www.trilemon.com/wp-content/uploads/2012/08/Information-Visualization-Perception-for-Design.pdf> [odczyt: 12.08.2014].
- Westbrook, Lynn (2006). *Mental models: a theoretical overview and preliminary study*. Journal of Information Science Vol. 32, No. 6, p. 563-579.
- Wizualizacja (2014) [dok. elektr.]. W: *Wikipedia*. [http://pl.wikipedia.org/wiki/Wizualizacja\\_\(informatyka\)](http://pl.wikipedia.org/wiki/Wizualizacja_(informatyka)) oraz [http://pl.wikipedia.org/wiki/Wizualizacja\\_\(psychologia\)](http://pl.wikipedia.org/wiki/Wizualizacja_(psychologia)) [odczyt: 19.08.2014].
- Zhang, Yang (2008). *The influence of mental models on undergraduate students' searching behavior on the Web*. Information Processing and Management Vol. 44, p. 1330-1345.
- Zhang, Yang (2010). *Dimensions and elements of people's mental models of an information-rich web space*. Journal of the American Society for Information Science and Technology Vol. 61, Issue 11, p. 2206-2218.

---

# Źródła internetowe z zakresu nauk biologicznych i ich ocena w kontekście kształtowania kompetencji informacyjnych studentów

Monika Krakowska

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

## Wprowadzenie

Przemiany ekonomiczne, społeczne i technologiczne uzależnione są od kompetencji jednostek tworzących społeczeństwo oparte na wiedzy, zaangażowanych w pogłębianie swoich kwalifikacji oraz intelektualny rozwój. Odpowiednie wykorzystanie posiadanych umiejętności – obywatelskich, ekonomicznych, matematycznych, informacyjnych, naukowych czy kulturalnych – w sposób istotny wpływa na rozwój całej zbiorowości, przejawiający się między innymi w podejmowanych przez jej członków formach aktywności i w rozkwicie rozmaitych dyscyplin. Ekspansja nowych form i narzędzi informacyjno-komunikacyjnych zwiększa dostępność wyników badań naukowych publikowanych przez uczonych oraz przez instytucje i organizacje szkolnictwa wyższego. Aktywne uczestnictwo jednostek w kształtowaniu społeczeństwa wiedzy wymaga nie tylko znajomości różnorodnych źródeł, w tym naukowych, i umiejętności wyszukiwania w nich relewantnych informacji, ale też ciągłego doskonalenia w procesie uczenia się oraz właściwego rozumienia kultury informacyjnej (ang. *information literacy*).

Celem artykułu jest zdefiniowanie założeń i koncepcji kształtowania kultury informacyjnej, zwłaszcza w środowisku naukowym, akademickim. Zwrócono także uwagę na wybrane standardy kształcenia na studiach wyższych w zakresie szeroko rozumianych kompetencji informacyjnych oraz na ich korelację z macierzami celów i efektów kształcenia w poszczególnych dyscyplinach nauk bio-

logicznych. Zaprezentowano najważniejsze wytyczne w zakresie kształtowania kultury informacyjnej, w szczególności rozwijania kluczowych kompetencji informacyjnych wspomagających kreowanie warsztatu badawczego.

Właściwie rozwinięte umiejętności informacyjne przyczyniają się przede wszystkim do podejmowania adekwatnych działań w cyfrowym środowisku informacyjnym. Wskazano zatem na konieczność wspierania inicjatyw mających na celu podnoszenie kwalifikacji informacyjnych użytkowników pochodzących ze środowiska akademickiego. Omówiono także najważniejsze działania związane ze współtworzeniem platformy naukowej SYNAT oraz jej poszczególne elementy, w tym bazę naukowych zasobów elektronicznych, w kontekście oceny wartościującej źródeł *online*, ich gromadzenia i tworzenia charakterystyk wyszukiwawczych.

## **Implikacje kształtowania krytycznych i dziedzinowych kompetencji informacyjnych w środowisku szkolnictwa wyższego i ich standaryzacja**

Wielopłaszczyznowe spojrzenie na kulturę informacyjną, zwłaszcza w aspekcie jej oddziaływania na rozwój społeczeństwa wiedzy, stanowi podstawę rozważań nad rolą umiejętności informacyjnych w warunkowaniu zachowań użytkowników. Wzrasta zainteresowanie zrozumieniem procesów i zachowań informacyjnych, szczególnie w perspektywie kształcenia i uczenia się przez całe życie (ang. *lifelong learning*). Kultura informacyjna to szereg specyficznych kompetencji, które odnoszą się do funkcjonowania użytkownika w środowisku informacyjnym. Do najistotniejszych umiejętności cechujących kompetentnego użytkownika, który bierze udział w kreowaniu struktur społeczeństwa opartego na wiedzy, należą m.in. kompetencje informacyjne, cyfrowe, medialne oraz te dotyczące krytycznej oceny informacji. Umiejętności te, wpływające na inicjowanie skutecznych zachowań informacyjnych, odnoszą się również do podejmowanych przez użytkownika procesów edukacji na różnych poziomach. Kompetencje informacyjne, jako korelat właściwego zastosowania potencjału intelektualnego oraz wiedzy, stały się niezwykle istotne dla funkcjonowania we współczesnym społeczeństwie.

Do procesów kształtowania kwalifikacji i predyspozycji niezbędnych w środowisku informacyjnym odnoszą się różnorodne wzorce paradygmatyczne, związane z edukowaniem i formowaniem użytkowników społeczeństwa opartego na wiedzy. Wzorce te są wypracowywane przez liczne organizacje bibliotekarskie, np. International Federation of Library Associations (IFLA), American Library



Association (ALA), Association for College and Research Libraries (ACRL) oraz Society of College, National and University Libraries (SCONUL), jak również przez instytucje szkolnictwa wyższego, w tym ośrodki akademickie. Organizacje te definiują i wdrażają do programów kształcenia najistotniejsze koncepcje kultury informacyjnej oraz standardy i modele przekazywania wiedzy o kompetencjach informacyjnych. Wskazują także na najważniejsze czynniki decydujące o właściwych efektach procesu uczenia się i kształcenia permanentnego. Do tych determinantów należą między innymi: znajomość potrzeb informacyjnych – własnych i odbiorców, orientacja w zasobach informacyjnych, umiejętność tworzenia strategii wyszukiwawczych, dobór i krytyczna ocena jakości informacji [ACRL, dok. elektr.].

Wiele modeli oraz standardów określających poszczególne probieże umiejętności stanowiących fundament *information literacy* jest z powodzeniem wykorzystywanych i wprowadzanych do systemów edukacji. Normy te mają również jednorodną konfigurację elementów składających się na paradygmat rozwijania kompetencji informacyjnych.

Wyróżnia się wśród nich nie tylko znanstwo w zakresie źródeł informacji, umiejętność ich identyfikowania, określania ich dostępności, efektywnego wyszukiwania, syntetyzowania oraz krytycznej ewaluacji. Do kluczowych umiejętności zaliczono też kompetencje związane z definiowaniem i wyrażaniem potrzeb informacyjnych, poprawnym wnioskowaniem oraz formułowaniem stanowisk na podstawie udostępnianej wiedzy, jak również rozpowszechnianiem i transmisją informacji oraz kreowaniem nowej wiedzy i poszerzaniem systemu wartości. Istotna jest także wykazywana przez użytkownika znajomość zróżnicowanego kontekstu ekonomicznego, prawnego i społecznego wykorzystywania informacji. Wszystkie te kwalifikacje stanowią fundament nowoczesnego, globalnego społeczeństwa informacyjnego [ACRL, dok. elektr.].

W przestrzeni szkolnictwa wyższego wprowadza się różnorodne normatywy w zakresie kształtowania odpowiednich umiejętności informacyjnych, które wspomagają realizowanie programów studiów i osiągnięcie należytych efektów kształcenia. Modele takie, np. opracowany przez SCONUL *Seven pillars of Information Literacy*, są udoskonalane, zwłaszcza w perspektywie współczesnych zmian ekonomiczno-społecznych. Wzorzec SCONUL uwidacznia przede wszystkim konsyтуacyjność i dynamiczność procesu kształcenia oraz zdobywania kompetencji informacyjnych. Odnosi się do określonej społeczności, środowiska, w którym funkcjonuje użytkownik, ale także wskazuje na indywidualizację całego procesu edukacji [SCONUL 2011, p. 4]. W uzupełnionym i zoptymalizowanym przez stowarzyszenie schemacie należy zwrócić uwagę na wykształcenie wśród użytkowników pochodzących z różnorodnych środowisk, szczególnie academic-

kich, umiejętności uświadamiania sobie i zapełniania tzw. luk informacyjnych (ang. *information gaps*), co pozwala na kontynuowanie metapoznawczego procesu edukacji, w którym istotne staje się konstruowanie strategii wyszukiwania informacji (w dostępnych źródłach), identyfikowanie kontrolowanych słowników, stosowanie taksonomii oraz właściwych słów kluczowych. W procesie tym ważną rolę odgrywa rozwijanie świadomości użytkownika w zakresie funkcjonowania instytucji udostępniających informację, rozpowszechniających wiedzę w społeczności lokalnej i globalnej oraz konieczności ciągłego dostosowywania się do dynamicznie rozwijającego się środowiska informacyjnego [SCONUL 2011, p. 5-11]. Zarówno standardy kompetencji informacyjnych, jak i modele ich implementacji oraz poszczególne wytyczne i wskazówki odnośnie do formowania umiejętności istotnych w danej kulturze informacyjnej pełnią kluczową rolę w rozwoju globalnej przestrzeni szkolnictwa wyższego.

Opracowywane w formie standardów wytyczne, dotyczące rezultatów kształcenia w określonej dyscyplinie, są przygotowywane w większości państw. Zawierają one matrycę czynników kształtujących kulturę informacyjną, wypracowaną przez środowiska bibliotekarskie, mogącą stanowić podstawę do tworzenia optymalnych programów kształcenia. Nadrzędnym celem ich wdrażania jest przede wszystkim harmonizacja, przejrzystość i normalizacja procesu edukacji na poziomie wyższym. Takie wzorce zostały wypracowane na potrzeby stworzenia Europejskiego Obszaru Szkolnictwa Wyższego m.in. w Wielkiej Brytanii przez Quality Assurance Agency for Higher Education, organizację zajmującą się oceną jakości kształcenia [QAAHE 2007, dok. elektr.]. Zestawienie kryteriów dziedzinowych (ang. *subject benchmark statement*), utworzone na potrzeby kształcenia akademickiego w ramach różnych dyscyplin, w tym także nauk biologicznych, ma określać charakterystykę studiów wyższych w danej dziedzinie oraz specyfikę programów kształcenia. Poszczególne wskaźniki obrazują oczekiwania względem efektów kształcenia, zdobytej wiedzy i umiejętności, a ponadto odnoszą się do narzędzi ewaluacji odpowiednich kwalifikacji [QAAHE 2007, dok. elektr.]. Kryteria dziedzinowe stanowią podstawę pomiaru jakości edukacji przez analizę wyszczególnionych, wyspecjalizowanych rezultatów kształcenia (ang. *learning outcomes*) dla elastycznie i progresywnie konstruowanych programów nauczania.

W dokumencie *Subject benchmark statement. Biosciences*, prezentującym podstawy kształcenia w naukach biologicznych, zaproponowano zestaw metod i obszarów badawczych skupionych wokół eksploracji procesów życiowych i wzajemnych relacji pomiędzy organizmami żywymi [QAAHE 2007, p. 2, dok. elektr.]. Wskazano tym samym na specyfikę programów edukacji, standardy nauczania oraz właściwości i umiejętności, jakie studenci kierunków biologicz-

nych, na poziomie studiów pierwszego i drugiego stopnia, powinni zrealizować i zdobyć podczas całego okresu kształcenia. W naukach biologicznych, stanowiących obszar badań praktycznych i eksperymentalnych, wymaga się od badaczy i studentów odpowiedniej wiedzy w zakresie tworzenia warsztatu badawczego, pracy indywidualnej i współpracy w grupie, a także umiejętności informacyjno-medialnych: korzystania z różnorodnych źródeł informacji, stosowania różnych form komunikacji, w tym wykorzystujących zasoby multimedialne i cyfrowe, oraz korzystania z narzędzi technologii informacyjno-komunikacyjnych. We wspomnianym dokumencie wytypowano także najistotniejsze kompetencje, które student jest zobowiązany zdobyć i rozwijać w toku procesu kształcenia oraz w trakcie edukacji przez całe życie (ang. *lifelong learning*), zarówno profesjonalnej i zawodowej, jak i pozaformalnej. Zaliczono do nich:

- a) wiedzę dziedzinową i jej rozumienie,
- b) umiejętności o charakterze dziedzinowym (ang. *subject-specific skills*), kompetencje naukowe,
- c) umiejętności akademickie i tzw. przenośne (ang. *graduate and transferable skills*), wśród których wyróżniono kompetencje intelektualne, praktyczne, matematyczne, komunikacyjne, związane z prezentacją i wykorzystaniem technologii informacyjnych, interpersonalne i grupowe oraz kompetencje samodzielnego zarządzania i rozwoju zawodowego [QAAHE 2007, p. 2-4, dok. elektr.].

Wśród kompetencji intelektualnych zwrócono uwagę na kwalifikacje związane z gromadzeniem, krytycznym analizowaniem i wnioskowaniem na podstawie zebranych informacji, klasyfikowaniem zasobów tradycyjnych i cyfrowych oraz krytycznym ewaluowaniem ich jakości. Zaliczono do nich także umiejętności matematyczne (ang. *numeracy skills*) oraz zdolność rozumienia i interpretowania źródeł tekstowych, statystycznych i graficznych, zarówno *online*, jak i tradycyjnych [QAAHE 2007, p. 6-7, dok. elektr.]. Wyróżniono również generyczne standardy kształcenia, nieprzyporządkowane do specyficznego obszaru badań w naukach biologicznych, w tym dotyczące tzw. umiejętności progowych (ang. *threshold standard*). Do tych standardów włączono zdolność precyzyjnego zapisywania i kodowania danych, umiejętność przeprowadzania analiz danych jakościowych i statystycznych, jak również rozwijanie sposobów uaktualniania wiedzy oraz predyspozycji do organizowania i zarządzania informacją biologiczną [QAAHE 2007, p. 10, dok. elektr.]. Wytyczne dotyczą także procesów nauczania oraz strategii i metod kształcenia stosowanych przez nauczycieli akademickich. W tym przypadku zwrócono uwagę na konieczność inicjowania edukacji na odległość, tworzenia i korzystania z zasobów *online* oraz informowania o różnorodnych źródłach informacji elektronicznej i udostępniania ich.

Wypracowane standardy efektywnego kształcenia, zwłaszcza obejmujące nauczanie specyficznych kwalifikacji dziedzinowych oraz umiejętności informacyjnych, są implementowane w uczelniach brytyjskich realizujących programy studiów w naukach biologicznych. Praktyka ta związana jest najczęściej z rozwijaniem umiejętności w zakresie krytycznej analizy, interpretacji i oceny jakości informacji, selekcji i syntezy relewantnej informacji, sprawnego komunikowania, biegłości w stosowaniu systemów informacyjno-wyszukiwawczych, jak i z podwyższaniem kompetencji naukowych, cyfrowych, medialnych etc. [University of Derby, dok. elektr.; Oxford Brooks University, dok. elektr.; Queen Mary University of London, dok. elektr.].

Rozpatrując wytyczne dotyczące efektywnego kształcenia w poszczególnych dziedzinach, można odwołać się do założeń Deklaracji Bolońskiej. Określają one ideę tworzenia Europejskiego Obszaru Szkolnictwa Wyższego – sformalizowanego i transparentnego – w którym możliwe jest konfrontowanie ze sobą nie tylko programów studiów wyższych i efektywności procesu kształcenia, ale także różnorodnych kompetencji i umiejętności, do których zdobywania i doskonalenia są zobligowani studenci [Bologna Process, dok. elektr.]. W Polsce zagadnienie to zostało uregulowane w Rozporządzeniu Ministra Nauki i Szkolnictwa Wyższego z dnia 2 listopada 2011 r. w sprawie Krajowych Ram Kwalifikacji dla Szkolnictwa Wyższego. W dokumencie tym – wzorując się na postanowieniach innych europejskich rządów i instytucji, a także na wytycznych Unii Europejskiej [Bologna Process, dok. elektr.] – określono efekty kształcenia dla profili zarówno ogólnoakademickich, jak i dziedzinowych w obszarze nauk humanistycznych, społecznych, technicznych, ścisłych, medycznych i przyrodniczych, w tym biologii, nauk rolniczych, weterynaryjnych i leśnych [Rozporządzenie... 2011]. Rozporządzenie stanowi fundament planowania procesu kształcenia, określania efektów uczenia się, definiowania zasobu wiedzy, umiejętności i kompetencji społecznych, które student jest zobowiązany uzyskać podczas realizowania programów kształcenia na danym kierunku studiów pierwszego i drugiego stopnia. Odwołano się w nim również do standardów i modeli kształcenia kultury informacyjnej [Ministerstwo..., dok. elektr.].

W niniejszym artykule odniesiono się wyłącznie do tych elementów procesu kształcenia, które precyzują i uświadamiają konieczność budowania warsztatu bibliograficzno-informacyjnego, rozwijania umiejętności krytycznej ewaluacji informacji, gromadzenia i stosowania różnorodnych zasobów tradycyjnych i cyfrowych oraz tworzenia i stosowania wiedzy dziedzinowej o kontrolowanej jakości. W aspekcie pogłębiania wiedzy zasygnalizowano konieczność posiadania orientacji w zakresie podstawowych kategorii pojęciowych i terminologii przyrodniczej oraz znajomości kierunków rozwoju nauk biologicznych, pokrewnych

dyscyplin i specjalizacji. Zwrócono szczególną uwagę na te fragmenty w dokumencie rządowym, które odnoszą się do: 1) rozwijania kompetencji związanych z krytyczną oceną aktualnie podejmowanych w środowisku naukowym dyskusji oraz wyników badań prowadzonych na szczeblu międzynarodowym, publikowanych w literaturze fachowej, 2) stosowania odpowiednich metod, technik i narzędzi badawczych, 3) pozyskiwania funduszy na realizację projektów naukowych, a także 3) możliwości wdrażania wyników badań i osiągnięć naukowych w życiu społeczno-gospodarczym [Rozporządzenie..., s. 14743-14744]. W zakresie kreowania kultury informacyjnej podkreślono natomiast, iż student jest zobowiązany zdobyć kwalifikacje umożliwiające mu korzystanie z dostępnych źródeł tradycyjnych i elektronicznych, ich krytyczną analizę i selekcję informacji oraz optymalne wnioskowanie i formułowanie uzasadnionych tez na podstawie informacji pozyskanych z różnorodnych zasobów. W obszarze kompetencji społecznych Krajowe Ramy Kwalifikacji wskazują przede wszystkim na konieczność zrozumienia potrzeby uczenia się przez całe życie, podnoszenia kompetencji zawodowych i personalnych, permanentnego zdobywania wiedzy przy zastosowaniu odpowiednich źródeł fachowych i naukowych oraz systematycznego aktualizowania wiedzy z zakresu uprawianej dziedziny naukowej i jej praktycznego wykorzystywania [Rozporządzenie..., s. 14748].

Jednym z aspektów kultury informacyjnej, bazującej na dorobku naukowym bibliotekoznawstwa, jest kształtowanie odpowiednich kwalifikacji w zakresie zdobywania i wykorzystywania wiedzy o różnorodnych źródłach informacji, krytycznej ewaluacji tych źródeł oraz ich typologii [Batorowska 2013, s. 61; s. 116]. W odniesieniu do zachowań informacyjnych kompetencje użytkowników związane są z umiejętnością samodzielnego identyfikowania własnych potrzeb informacyjnych, lokalizowania i eksplorowania źródeł w celu pozyskania z nich relewantnej informacji, jak również z zarządzaniem, organizacją i zastosowaniem informacji, poddanej wcześniej ocenie jakościowej [Lau 2011, p. 16]. Zarówno modele kształcenia kompetencji informacyjnych, jak i macierze celów i efektów kształcenia zaproponowane w *Subject Benchmark Statements* wskazują na istotność wypracowywania skutecznego sposobu inicjowania procesu edukacji w zakresie kultury informacyjnej oraz rozwijania właściwych umiejętności i kwalifikacji w środowisku akademickim.

Na szczególną uwagę zasługują umiejętności naukowe (ang. *scientific literacy*) oraz kluczowe kompetencje informacyjne (ang. *critical information literacy*), za pomocą których student ma możliwość budowania dziedzinowego warsztatu informacyjnego i podejmowania zachowań informacyjnych w kontekście akademickim [Nitz et al. 2014, p. 13-14; Dold 2014, p. 180]. Wykorzystywanie umiejętności informacyjnych, interpretowanie, konstruowanie, przetwarzanie komu-

ników, jak również ewaluacja naukowych źródeł informacji stanowią kluczowe kwalifikacje studentów, umożliwiające zrozumienie samej nauki lub określonej dyscypliny [Nitz et al. 2014, p. 13]. Kluczowe kompetencje informacyjne umożliwiają identyfikowanie specyficznych zasobów i form komunikacji stosowanych w różnych naukach, np. fizyce, biologii, socjologii etc. [Dold 2014, p. 180]. Wspomagają one również kształtowanie środowiska informacyjnego w danej grupie społecznej, w tym także w społeczności akademickiej, przez tworzenie reguł korzystania z określonego języka profesjonalnego, zwyczajów oraz praktyk wykształconych w obrębie poszczególnych dyscyplin. James Elmborg definiuje te umiejętności nie tylko jako zespół kwalifikacji, ale też jako zdolność pojmowania całego systemu nauki i obowiązujących w niej zasad tworzenia wiedzy. Co istotne, kompetencje te mieszczą w sobie umiejętności krytycznej oceny samego systemu, traktowanego jako całość [Elmborg 2006, p. 196]. Wspomaga to nie tylko indywidualne analityczne podejście do naukowych hipotez, ale także dynamiczny rozwój dyscypliny przez integrowanie i konfrontowanie dominujących doktryn, a zarazem uwzględnianie interdyscyplinarnego kontekstu, w którym funkcjonuje dana dziedzina [Dold 2014, p. 181]. Konieczność kształcenia kompetencji informacyjnych adekwatnych do indywidualnych potrzeb użytkownika i cech środowiska informacyjnego, w którym ten użytkownik funkcjonuje, została wielokrotnie potwierdzona w badaniach informatologicznych [m.in. Brem, Russell, Weems 2011; Lorenzen 2001; Barranoik 2001; Urquhart, Rowley 2007], przy czym wykazano też, że w przypadku osób preferujących elektroniczne zasoby informacji *online* kluczowe znaczenie ma umiejętność wyszukiwania relewantnej informacji, jej oceny jakościowej oraz późniejszego nią zarządzania [Julien, Barker 2009, s. 13].

Przedstawione w artykule badania nad ewaluacją i tworzeniem zasobów z dziedziny nauk biologicznych w ramach realizacji projektu SYNAT miały za zadanie wspomóc proces kształtowania kompetencji informacyjnych przedstawicieli środowiska akademickiego i naukowego. Rezultatem badań było stworzenie wytycznych na potrzeby budowania warsztatu informacyjnego studentów biologii, pracujących w tej dyscyplinie badaczy oraz nauczycieli akademickich realizujących programy studiów. W tworzeniu tego warsztatu powinno się również uwzględniać macierze celów i efektywności edukacji w zakresie kultury informacyjnej na poziomie studiów wyższych.

## **Źródła informacji z dziedziny biologii i krytyczna ocena ich jakości**

Wiedza na temat różnorodnych źródeł informacji dziedzinowej jest warunkiem stosowania adekwatnych metod nauczania i uczenia się oraz umiejętnego wykorzystania kompetencji informacyjnych zarówno przez specjalistów z wszystkich dziedzin, jak i przez studentów. Z racji ekspansywnego rozwoju wielu obszarów wiedzy i technologii oraz postulatów otwarcia dostępu do zasobów nauki warto zwrócić uwagę na rozwój systemów informacyjno-wyszukiwawczych oferujących wgląd do wyników badań i opracowań teoretycznych z danej dziedziny, umożliwiających budowanie wiedzy.

Przykładem wdrażania inicjatyw zmierzających do zaspokojenia potrzeb informacyjnych naukowców i ogółu społeczności akademickiej oraz kształtowania kultury informacyjnej w tym środowisku był projekt SYNAT realizowany w latach 2010-2013 przez konsorcjum szesnastu jednostek naukowych polskich uczelni wyższych. Celem tego zadania badawczego było stworzenie infrastruktury otwartej platformy hostingowej i systemu baz wiedzy rejestrujących zasoby naukowe wykorzystywane na potrzeby edukacji. Wśród etapów realizacji projektu można wyróżnić m.in. identyfikację, opracowanie, organizację i koncentrację merytorycznie wartościowych, naukowych zasobów elektronicznych, które składają się na zawartość platformy.

Proces identyfikacji, ewidencjonowania, zbierania oraz ewaluowania zróżnicowanych źródeł internetowych rozpoczęto w styczniu 2011 r. Jednym z zespołów realizujących tę część projektu SYNAT była interdyscyplinarna grupa badaczy z Uniwersytetu Jagiellońskiego, w tym pracowników naukowych Instytutu Informacji Naukowej i Bibliotekoznawstwa UJ. Wśród analizowanych naukowych zasobów cyfrowych z obszaru matematyczno-przyrodniczego weryfikacja i wstępna organizacja repozytorium wiedzy dotyczyła takich dyscyplin nauki, jak biologia, chemia, fizyka, geografia, medycyna i matematyka. W artykule przedstawiono wyłącznie źródła z zakresu biologii, gdyż autorka uczestniczyła w pracach zespołu, poddając analizie zasoby z tej właśnie dziedziny.

## **Tworzenie warsztatu badawczego w zakresie gromadzenia zasobów elektronicznych**

Wyszukiwanie źródeł informacji w dziedzinie biologii oparto na trzech systemach klasyfikacyjnych: Uniwersalnej Klasyfikacji Dziesiętnej (UKD), Polskiej Klasyfikacji Tematycznej Komitetu Badań Naukowych (PKT KBN) oraz klasy-

fikacji systematycznej Jagiellońskiej Biblioteki Cyfrowej (JBC). Zastosowanie systemu podziału zasobów i hierarchizacji dziedzinowej w obszarze nauk biologicznych pozwoliło sprecyzować kryteria lokalizowania źródeł *online*. Zgodnie z Uniwersalną Klasyfikacją Dziesiątą, według której – w obrębie głównego działu Matematyka. Nauki Przyrodnicze – wyodrębniono poddziały związane z biologią, poszukiwano zasobów informacyjnych za pomocą następujących słów kluczowych (nazw poddziałów):

- biologia – podręczniki szkolne i akademickie,
- biologia – materiały do ćwiczeń praktycznych (zadania, pytania, testy),
- antropologia – pochodzenie człowieka – poszczególne rasy,
- hydrobiologia,
- genetyka ogólna – dziedziczność,
- ewolucja – powstawanie gatunków – darwinizm,
- biologia komórki – cytologia,
- materiałowe podstawy życia – biochemia – biologia molekularna – biofizyka,
- wirusologia,
- mikrobiologia,
- botanika (w tym podręczniki akademickie oraz materiały do ćwiczeń praktycznych),
- zachowanie i ochrona roślin w stanie dzikim,
- ogrody botaniczne – rezerваты roślin,
- klucze do oznaczania roślin,
- zoologia (w tym podręczniki akademickie oraz materiały do ćwiczeń praktycznych),
- zachowanie i ochrona zwierząt w stanie dzikim,
- ogrody zoologiczne – rezerваты zwierząt,
- sposób życia zwierząt – etologia – psychologia zwierząt,
- zoologia systematyczna,
- bezkęgowce ogólnie,
- pierwotniaki – jamochłony – koralowce i inne,
- mięczaki (małże, ślimaki, ośmiornice, inne),
- pierścieniowce (robaki, skorupiaki, pajęczaki i inne),
- owady – entomologia,
- strunowce,
- kręgowce,
- ryby – ichtiologia,
- płazy (kumaki, ropucha, salamandra, żaby),
- gady (jaszczurki, węże, żółwie, krokodyły),
- ptaki – ornitologia,



ssaki,

naczelne – człekokształtne,

*Homo sapiens* – człowiek [Biblioteka w szkole, dok. elektr.].

W przypadku dostosowywania haseł wyszukiwawczych do standardu tematycznej klasyfikacji KBN wyszukiwanie prowadzono w zasobach elektronicznych dotyczących:

biologii teoretycznej i doświadczalnej,

systematyki biologicznej – taksonomii,

biologii komórkowej i subkomórkowej,

biologii opisowej i makrobiologii,

biologii eksperymentalnej,

biometrii (biomatematyki i biostatyki),

bioinżynierii,

metod i technik badań biologicznych,

autekologii,

populacjologii,

biocenologii,

fitosocjologii,

biochemii,

biochemii porównawczej,

biochemii drobnoustrojów – biochemii roślin – biochemii zwierząt,

biochemii molekularnej,

wirusologii klinicznej,

bakteriologii – mykologii mikrobiologicznej,

bakteriologii lekarskiej,

systematyki bakterii,

parazytologii,

antropologii – antropometrii,

antropogenezy,

genetyki osobniczej,

genetyki zwierząt,

embriologii zwierząt – teratologii,

anatomii porównawczej – morfologii porównawczej,

fizjologii – funkcji i układów organizmów zwierzęcych,

zoogeografii (faunistyki),

ochrony gatunkowej zwierząt,

genetyki roślin,

embriologii roślin,

anatomii roślin – morfologii roślin,

patologii roślin,  
radiobiologii<sup>1</sup> [KBN, dok. elektr.]

Wykorzystując klasyfikację systematyczną JBC, która dzieli biologię na kilka poddziałów, uzupełniono proces wyszukiwawczy słowami kluczowymi z zakresu oceanologii (zwłaszcza w przypadku pozyskiwania źródeł z zakresu biologii oceanicznej i morskiej) oraz paleontologii. Dodatkowo skorzystano z podziału nauk biologicznych zaproponowanego przez INTUTE. Obok nazw dyscyplin, które już zostały wymienione bądź pojawiły się w zastosowanych klasyfikacjach, użyto dodatkowych słów kluczowych do poszukiwania źródeł mieszczących się w problematyce:

bioróżnorodności,  
historii naturalnej,  
paleobotaniki,  
paleozoologii,  
protistologii,  
paleontologii stratygraficznej,  
mykologii,  
biologii słodkowodnej [INTUTE, dok. elektr.].

W przypadku typologii źródeł zastosowano model podziału różnorodnych zasobów opracowany przez INTUTE, który rozpoczął swoją działalność projektową w 1996 r. w ramach programu Electronic Libraries (eLib) Programme. Inicjatywa, opierająca się na współdziałaniu bibliotekarzy i naukowców z poszczególnych dyscyplin, miała na celu utworzenie serwisu o kontrolowanej jakości Resource Discovery Network (RDN). Konsorcjum składające się z przedstawicieli University of Birmingham, University of Bristol, Heriot-Watt University, University of Manchester, Manchester Metropolitan University, University of Nottingham, University of Oxford zakończyło działalność w 2011 r. Serwis udostępnia jednak nadal zweryfikowane i poddane krytycznej ewaluacji wielod dziedzinowe zasoby elektroniczne [INTUTE, dok. elektr.]. Zaindeksowano w nim m.in.: archiwa, projekty artystyczne (ang. *arts projects*), strony www stowarzyszeń, bibliograficzne bazy danych, niebibliograficzne bazy danych, materiały bibliograficzne, materiały biograficzne, blogi, studia przypadku (ang. *case studies*), kolekcje, serwisy internetowe przedsiębiorstw, zbiory danych (ang. *datasets*), e-booki, wydarzenia, wy-

---

<sup>1</sup> W prezentowanej liście słów kluczowych, opartych na klasyfikacji KBN i UKD, pominięto część poddziałów, nie zapisano dublujących się terminów oraz nie uwzględniono haseł odnoszących się do weterynarii, rolnictwa, ekologii, technologii, medycyny, gdyż procesy wyszukiwania w tym obszarze prowadzone były przez innych badaczy. W przypadku źródeł m.in. biomedycznych, bioinżynierskich, biochemicznych, biofizycznych starano się w wyniku ewaluacji treści odnaleźć zasoby, których autorstwo afiliowało do instytucji, organizacji prowadzących badania, szkolenia, kształcenie przede wszystkim w zakresie biologii.

stawy i galerie, przewodniki dziedzinowe (ang. *subject guides*), badania terenowe, organy rządowe, publikacje rządowe, obrazy, zasoby interaktywne, czasopisma – spisy treści i streszczenia, czasopisma pełnotekstowe, raporty prawne, materiały edukacyjne, konspekty wykładów, prawodawstwo, biblioteki, listy i grupy dyskusyjne, mapy, dokumenty dźwiękowe, dokumenty audiowizualne (ang. *moving images*), muzea, aktualności, organizacje *non-profit*, inne organizacje, dokumenty / raporty, artykuły / teksty naukowe, patenty / normy, ulotki informacyjne dla pacjentów, praktyki – wytyczne, źródła pierwotne, informacje produktowe, organizacje zawodowe, informatory bezpośrednie (encyklopedie, słowniki itp.), ośrodki badawcze i projekty, przewodniki i katalogi zasobów, recenzje, oprogramowanie, bazy danych obiektów muzealnych (ang. *specimen databases*), statystyki, przeglądy statystyczne, poradniki, niezależne komitety doradcze (ang. *think tank*) [INTUTE, *Resource type(s)*, dok. elektr.]. Kryterium rodzaju źródeł informacji było niezwykle istotne podczas realizacji projektu, zarówno w fazie eksploracji i oceny zasobów sieciowych, jak i przyporządkowywania ich do określonych kategorii. Typologia źródeł przyjęta w INTUTE stanowiła podstawę gromadzenia i selekcji różnorodnych zasobów z dziedziny biologii oraz ich wstępnej ewaluacji jakościowej, którą w trakcie badań znacznie pogłębiono. Na potrzeby projektu opracowano także nowy, bardziej szczegółowy format opisu bibliograficznego.

W projekcie SYNAT została zaproponowana typologia zasobów naukowych, którą należało posługiwać się podczas pozyskiwania dokumentów i tworzenia bazy danych. Podział ten, opracowany w wyniku ścisłej współpracy między Biblioteką Narodową, Biblioteką Główną Politechniki Warszawskiej i Biblioteką Jagiellońską, uwzględniał zasoby naukowe najczęściej występujące w środowisku komunikacji elektronicznej. Wyróżniono w nim następujące polskie i zagraniczne źródła informacji: archiwa, bazy danych faktograficzne, bazy danych pozostałe, bibliograficzne bazy danych, biblioteki, biblioteki cyfrowe, blogi, centra badawcze, czasopisma – spis treści i streszczenia, czasopisma pełnotekstowe, dane statystyczne, dokumenty graficzne, dokumenty rządowe, firmy / spółki, informacje biograficzne, inne organizacje, inne teksty, zbiory dokumentów, instytucje i organizacje edukacyjne, instytucje i organizacje naukowe, instytucje rządowe, książki elektroniczne, księgarnie, listy dyskusyjne / fora, mapy, atlasy, materiały branżowe, materiały konferencyjne, multimedia, muzea, organizacje rządowe, organizacje użytku publicznego, patenty, katalogi norm, pełnotekstowe bazy danych, portale, programy komputerowe, projekty badawcze, raporty, rozprawy doktorskie i habilitacyjne, serwisy wiadomości, słowniki, encyklopedie, informatory, spisy bibliograficzne, stowarzyszenia, strony domowe pracowników naukowych, strony konferencji, wydawnictwa, wykłady, materiały szkoleniowe, kursy, wystawy / galerie, zbiory aktów prawnych.

Ewaluacja rozproszonych zasobów przeprowadzona została według określonych kryteriów oceny jakości, odnoszących się do trzech obszarów: 1) treści decydujących o przeznaczeniu źródła dla określonej grupy odbiorców – studentów, badaczy, kadry naukowej lub dydaktycznej, 2) przystępności i zakresu tematycznego oraz 3) formy, ze szczególnym uwzględnieniem szaty graficznej serwisu, struktury interfejsu oraz innych cech formalnych ułatwiających odbiór treści. Zastosowano metodę selekcji pozwalającą na zebranie jak największej liczby wiarygodnych zasobów internetowych użytecznych dla szerokiego grona odbiorców informacji naukowej, zarówno specjalistów, naukowców i badaczy, jak i studentów studiów pierwszego, drugiego i trzeciego stopnia.

W trakcie gromadzenia źródeł elektronicznych wstępna ocena stanowiła podstawę preselekcji zasobów naukowych, z uściśleniem że pod uwagę brane były tylko źródła przydatne do uprawiania nauki i edukacji w szkolnictwie wyższym. Na tym etapie zastosowano następujące mierniki jakości: aktualność informacji, dostępność (ze szczególnym uwzględnieniem oprogramowania i oferowanych w danym serwisie narzędzi wspomagających wgląd do zasobów elektronicznych), zgodność z obowiązującym prawem autorskim i innymi prawami własności intelektualnej, forma i architektura informacji (struktura, interfejs), jak również etyczność zasobów, czyli zgodność z dobrymi obyczajami przyjętymi w środowisku akademickim i w społeczeństwie w ogóle (eliminowano treści wulgarne i ogólnie nieaprobowane kulturowo) [Sapa, maszynopis].

W kolejnej fazie oceny jakości zasobów naukowych przeznaczonych do prezentacji i posadowienia w systemie baz wiedzy ewaluacja dotyczyła zasobów naukowych oraz wspierających naukę, a także narzędzi dostępu do nich. Źródła te oceniano według: a) poziomu naukowości (obecność recenzji, wyposażenie tekstu w aparat bibliograficzny, logiczność wywodu), b) wiarygodności określonej na podstawie autorstwa i weryfikowalnej za pomocą dodatkowych informacji o autorze lub instytucji sprawczej, c) istotności, oryginalności, pragmatyczności, rzetelności i aktualności informacji, d) funkcjonalności, identyfikacji obiektów składowych, integracji, dostępności i możliwości wykorzystania informacji, e) stabilności instytucji lub organizacji zarządzającej zasobem [Cornell University Library, dok. elektr.; Bednarek-Michalska, dok. elektr.; Smith, dok. elektr.]. Posłużono się także dokonaniem twórców INTUTE. W serwisie opisano różnorodne typy zasobów, których kategoryzacja została wykorzystana zarówno podczas wyszukiwania informacji w trakcie badań, jak i aktualizowania danych udostępnionych na stronach internetowych [INTUTE, dok. elektr.].

Podczas prowadzonej analizy i oceny źródeł elektronicznych wzięto pod uwagę zarówno zasoby naukowe zawierające wyniki badań, prezentowane w formie artykułów, monografii, prac zbiorowych, materiałów konferencyjnych, raportów,

zbiorów tekstów naukowych, repozytoriów cyfrowych, jak i źródła w postaci informatorów bezpośrednich, czasopism, baz danych, blogów, skryptów, map i atlasów lub też serwisów wiadomości. Do ewaluacji włączono również bibliograficzne bazy danych i spisy bibliograficzne, a także strony domowe pracowników naukowych, strony internetowe projektów badawczych, portale, serwisy *online* różnorodnych instytucji naukowo-badawczych, stowarzyszeń naukowych, instytucji edukacyjnych, akademickich lub organizacji rządowych i centrów badawczych oraz źródła naukowe wspierające kształcenie na poziomie akademickim.

W przypadku ewaluacji źródeł z zakresu biologii nie odnaleziono w okresie, w którym prowadzone były badania<sup>2</sup>, żadnych zasobów odpowiadających przyjętym kryteriom, obejmujących archiwa, faktograficzne bazy danych, biblioteki cyfrowe, dane statystyczne, dokumenty rządowe, informację biograficzną, księgarnie, materiały branżowe, patenty, katalogi norm, rozprawy doktorskie i habilitacyjne, wydawnictwa, zbiory aktów prawnych. Dogłębna analiza cyberprzestrzeni w poszukiwaniu tego typu źródeł wymagałaby przede wszystkim zwiększenia nakładów czasu oraz pogłębionej weryfikacji dostępnych informacji.

Oszacowanie jakości dotyczyło ponad 2500 źródeł informacji naukowej *online* z dziedziny biologii, w tym botaniki, zoologii, anatomii i fizjologii, genetyki, ewolucji i bioróżnorodności, biologii komórki, biotechnologii, homeostazy. Wyłączono z poszukiwań źródła z obszaru ekologii i ochrony środowiska, rolnictwa i weterynarii, gdyż ten zakres był analizowany przez innych uczestników projektu.

Po dokonaniu oceny opracowano opisy bibliograficzne analizowanych źródeł w programie EndNote, służącym do tworzenia bibliograficznych baz danych. Zasoby charakteryzowano, podając m.in. dane dotyczące typu dokumentu, autora (-ów), tytułu, roku publikacji / autoryzacji / aktualizacji, miejsca publikacji, instytucji sprawczej, afiliacji autora, daty dostępu do źródła, krótkiego tytułu, tłumaczenia tytułu, skrótu tytułu, klasyfikacji KBN, JBC i UKD, słów kluczowych, krótkiego abstraktu oraz uwag (związanych z dostępem do źródła, interfejsem, poruszaniem się po stronie, możliwości wyszukiwawczych etc.), jak również oryginalnego języka źródła. W rozpatrywanym zbiorze przeważały anglojęzyczne zasoby cyfrowe. Prawidłowość ta dotyczyła baz danych różnego typu, czasopism naukowych, stron organizacji lub ośrodków akademickich oraz archiwów OA. Zasoby polskie, pozostające w mniejszości, obejmowały głównie fachowe czasopisma z dziedziny nauk biologicznych, strony uniwersytetów i wydziałów realizujących programy kierunkowych studiów 1. i 2. stopnia. Zwracały

---

<sup>2</sup> Z tego względu badanie wymaga kontynuacji (dalszego wyszukiwania i ewaluowania zasobów elektronicznych w dziedzinie biologii).

uwagę rozproszone strony prywatne naukowców, prowadzone w formie blogów naukowych<sup>3</sup>.

W efekcie, po ocenie jakości i weryfikacji, uzyskano 680 źródeł informacji, do których zaliczono bazy danych (ogólne), bibliograficzne bazy danych, biblioteki, blogi, centra badawcze, czasopisma pełnotekstowe, czasopisma – spisy treści i streszczenia dokumenty graficzne, firmy / spółki, inne organizacje, źródła stanowiące inne teksty i zbiory dokumentów, instytucje i organizacje edukacyjne, instytucje i organizacje naukowe, instytucje rządowe, książki elektroniczne, listy dyskusyjne / fora, mapy i atlasy, materiały konferencyjne, multimedia, muzea, organizacje rządowe, organizacje użytku publicznego, pełnotekstowe bazy danych, portale, programy komputerowe, projekty badawcze i raporty, serwisy wiadomości, słowniki, encyklopedie, informatory, spis bibliograficzny, stowarzyszenia, strony domowe pracowników naukowych, strony konferencji, wykłady, materiały szkoleniowe, kursy, wystawy / galerie.

W poniższej tabeli zaprezentowano wykaz (wraz z wartościami liczbowymi i procentowymi) wszystkich rozpatrywanych źródeł polskich i zagranicznych, które poddano krytycznej ocenie jakości podczas przeprowadzonych badań.

Tabela 1. Rodzaje źródeł informacji w naukach biologicznych, poddanych ewaluacji i uwzględnionych w bazie projektu SYNAT

Rodzaje źródeł informacji	Liczba poszczególnych źródeł informacji w układzie rangowym	Procentowa liczba poszczególnych źródeł informacji
zbiory dokumentów, inne teksty	124	18,24%
czasopisma – spisy treści i streszczenia	71	10,44%
portale	63	9,26%
projekty badawcze, raporty	63	9,26%
stowarzyszenia	47	6,91%
centra badawcze	37	5,44%
instytucje i organizacje edukacyjne	34	5,00%
czasopisma pełnotekstowe	28	4,12%
słowniki, encyklopedie, informatory	21	3,09%
bazy danych pozostałe	20	2,94%
serwisy wiadomości	19	2,79%

<sup>3</sup> W przypadku polskich blogów naukowych nie ma platformy analogicznej do anglojęzycznych ScienceBlogs (<http://scienceblogs.com>) lub Scientific American (<http://blogs.scientificamerican.com>), co może stanowić barierę w dotarciu do tego typu źródeł informacji.

c.d. tabeli 1

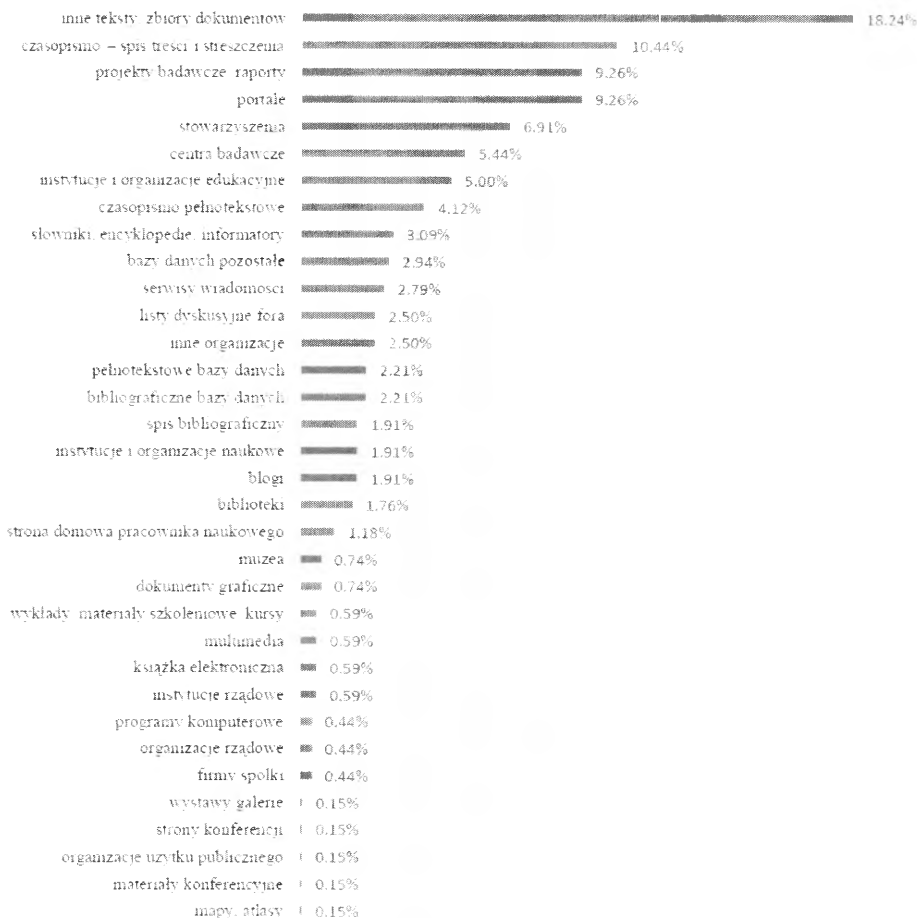
Rodzaje źródeł informacji	Liczba poszczególnych źródeł informacji w układzie rangowym	Procentowa liczba poszczególnych źródeł informacji
inne organizacje	17	2,50%
listy dyskusyjne / fora	17	2,50%
bibliograficzne bazy danych	15	2,21%
pełnotekstowe bazy danych	15	2,21%
blogi	13	1,91%
instytucje i organizacje naukowe	13	1,91%
spisy bibliograficzne	13	1,91%
biblioteki	12	1,76%
strony domowe pracownika naukowego	8	1,18%
dokumenty graficzne	5	0,74%
muzea	5	0,74%
instytucje rządowe	4	0,59%
książki elektroniczne	4	0,59%
multimedia	4	0,59%
wykłady, materiały szkoleniowe, kursy	4	0,59%
firmy / spółki	3	0,44%
organizacje rządowe	3	0,44%
programy komputerowe	3	0,44%
mapy, atlasy	1	0,15%
materiały konferencyjne	1	0,15%
organizacje użytku publicznego	1	0,15%
strony konferencji	1	0,15%
wystawy / galerie	1	0,15%
suma	680	100%

Źródło: badania własne

Na wykresie 1 przedstawiono, w porządku malejącym, procentowy udział poszczególnych typów zasobów sieciowych z zakresu nauk biologicznych, zweryfikowanych i zarejestrowanych w bazie SYNAT w postaci rozbudowanych opisów bibliograficznych zgodnych z formatem wprowadzania danych programu End-Note.

Najliczniejszą grupę wśród ewaluowanych zasobów *online* (124) tworzą tzw. inne teksty i zbiory dokumentów, obejmujące m.in.:

### Najczęściej ewaluowane zasoby elektroniczne z zakresu biologii



Wykres 1. Najliczniej reprezentowane naukowe zasoby internetowe z zakresu biologii, poddane krytycznej ocenie podczas procesu ewaluacji jakości do projektu SYNAT

Źródło: opracowanie własne

- opracowania naukowe z serii Mill Hills Essays amerykańskiego Narodowego Instytutu Badań Medycznych (National Institute for Medical Research);
- publikacje USGS Education Resources instytutu United States Geological Survey Education – materiały multimedialne, teksty, zasoby e-learningowe zarówno z biologii, jak i geografii czy geologii;
- teksty naukowe i edukacyjne Society for Experimental Biology (SEB), brytyjskiego Stowarzyszenia Biologii Eksperymentalnej;



- publikacje dotyczące różnorodnych zagadnień z zakresu biologii, umieszczone na stronach serwisu ScienceIQ, gdzie udostępniono zbiór krótkich artykułów naukowych z dziedziny m.in. biologii ogólnej, zoologii, botaniki, biologii człowieka, anatomii człowieka i historii naturalnej;
- artykuły naukowe dotyczące genetyki, genomów i biologii komórki, które pod nazwą Nature milestones: DNA Technologies zostały upublicznione na stronach serwisu Nature.com.

Jak wskazano na wykresie, drugą najliczniejszą grupą dziedzinowych źródeł elektronicznych z obszaru biologii, po zbiorach dokumentów i innych tekstów, są zasoby prezentujące spisy treści i streszczenia czasopism (71 pozycji, co stanowi 10,44% analizowanych źródeł). Wśród nich ewaluacji poddano m.in. czasopisma: „Journal of Theoretical Biology”, którego problematyka obejmuje zachowania zwierząt, biologię komórki, złożoność systemów biologicznych, ekologię, kinetykę enzymów, ewolucję, sieci genów, genetykę, biologię komórki, struktury molekularne, neurobiologię, pochodzenie życia, fizjologiczne mechanizmy biologiczne i biologię populacji; „Ethology: International Journal of Behavioural Biology” – periodyk poświęcony badaniom z zakresu biologii behawioralnej, etologii i zachowań zwierząt; recenzowane czasopismo „Plant Biotechnology Journal”, w którym publikowane są wyniki oryginalnych badań z zakresu biologii molekularnej roślin i raporty z zastosowania tych wyników w biotechnologii roślin. Należy również zaznaczyć, że czasopisma takie jak wspomniane „Journal of Theoretical Biology” udostępniają pełne teksty wyłącznie odpłatnie lub dla prenumeratorów.

W grupie 63 projektów badawczych i raportów (9,26% omawianych źródeł) uwzględniono zasoby związane z licznymi krajowymi i międzynarodowymi inicjatywami naukowymi oraz sprawozdania odnoszące się zarówno do zagadnień biologii ogólnej, jak i do specjalistycznych badań w dziedzinie genetyki, paleontologii, mikrobiologii.

Przykładem może być program badawczy TEStoplek „Rola transporterów oporności wielolekowej w farmakokinetyce i toksykologii – testy *in vitro* w praktyce farmaceutycznej i klinicznej” (rozpoczęty w 2009 r.), w którym głównym celem jest dostarczenie bazy merytorycznej dla opracowania nowych technik diagnostycznych, czynników i działań terapeutycznych oraz narzędzi dla przemysłu farmaceutycznego, umożliwiających podniesienie jakości życia przez kompleksowe przeciwdziałanie negatywnemu wpływowi zjawiska oporności wielolekowej nowotworów w onkologicznej praktyce klinicznej. Realizatorem projektu jest Uniwersytet Łódzki, a wykonawcą – Katedra Biofizyki Molekularnej (KBM) [UŁ, dok. elektr.].

Wyróżniono także Centrum doskonałości ANIMBIOGEN in EU (Centre of Excellence in Genomics, Biotechnology and Quality of Animal Products in Susta-

inable Production Systems with consideration of Animal Welfare), które zajmuje się badaniami w zakresie genetyki, hodowli zwierząt, genomiki funkcjonalnej, biotechnologii, biologii rozwoju i transkryptomiki. Celem działalności Centrum jest wzmocnienie jego potencjału badawczego i wzrost aktywności w prowadzeniu działalności naukowo-badawczej na skalę europejską. Kluczowym elementem projektu jest zacieśnianie współpracy z najlepszymi europejskimi ośrodkami naukowymi, zmierzające do utworzenia strategicznego partnerstwa w celu realizacji wspólnych interdyscyplinarnych projektów badawczych [IGHZ, dok. elektr.].

W trakcie analizy źródeł *online* dokonano oceny jakości i zweryfikowano również 63 biologiczne portale informacyjne (9,26% analizowanych źródeł), takie jak BioCentrum Edukacji Naukowej, TIMETREE. The Timescale of Life, BioAlive. The life science video share, Barcode of Life lub Young European Biotech Network. Kolejno uwzględniono polskie i zagraniczne stowarzyszenia naukowe (47 podmiotów), w tym m.in. Stowarzyszenie Studentów Nauk Przyrodniczych z Lublina, Polskie Towarzystwo Biologii Eksperymentalnej Roślin Uniwersytetu Wrocławskiego, Polskie Towarzystwo Mikrobiologów, Polskie Towarzystwo Przyrodników im. Kopernika, a także International Brain Research Organization, The Botanical Society of America, Federation of European Microbiological Societies, Federation of European Societies of Plant Biology (FESPB) oraz Society of Systematic Biology. Home of Systematic Biology, które zajmuje się problematyką biologii systematycznej, jej teorii, zasad, metodologii i praktyk, zarówno dla organizmów żyjących, jak i kopalnych. Serwis zawiera informacje o stowarzyszeniu, organizacji i składzie, misji, spotkaniach i publikacjach, w tym bezpłatny dostęp do treści i abstraktów artykułów z czasopisma „Systematic Biology”.

Wśród wprowadzonych do zasobów platformy SYNAT 37 centrów badawczych (5,44% źródeł) znalazły się m.in.: Krajowa Sieć Informacji o Bioróżnorodności oraz polski Instytut Biochemii i Biofizyki PAN z Warszawy. Ewaluacji poddano także 34 instytucje i organizacje edukacyjne (5% źródeł), w tym National Academies: Division on Earth and Life Studies, Fundację BioEdukacji, a także strony internetowe wydziałów i instytutów, np. amerykańskiego Duke University: Department of Biology, Department of Zoology; University of Otago z Nowej Zelandii, szwajcarskiego University of Zurich: Institute of Plant Biology oraz instytucji polskich: Zakładu Paleobiologii i Ewolucji Wydziału Biologii Uniwersytetu Warszawskiego, ośrodków Polskiej Akademii Nauk, np. Muzeum i Instytutu Zoologii Polskiej Akademii Nauk, Instytutu Fizjologii Roślin im. Franciszka Górskiego Polskiej Akademii Nauk.

Do grupy 28 czasopism pełnotekstowych zaliczono polskie i zagraniczne periodyki prezentujące ogólne lub szczegółowe wyniki analiz, badań, rezultaty

eksperymentów z biologii, genetyki, mikrobiologii, botanik i zoologii, ewolucji, etc., m.in. „New Scientist”, „Acta Mycologica”, „Molecular microbiology”, „PLOS Computational Biology”, „Experimental Biology Online”, „Folia Histochemica et Cytobiologica”, „Acta Biologica et Marinistica Universitatis Stetinensis”.

W bazie uwzględniono również 21 słowników, encyklopedii i informatorów (4,12% ogółu źródeł), w tym *Biology-Online Dictionary* oraz *EOL Encyclopedia of Life*, która sama w sobie stanowi również formę repozytorium, gdyż jej twórcy zmierzają do zgromadzenia rozproszonych informacji z książek, czasopism, baz danych, stron internetowych, projektów i wiedzy naukowców, umożliwiającego swobodny dostęp do zasobów cyfrowych poddanych weryfikacji i kontroli jakości.

Rezultatem pogłębionej ewaluacji było otrzymanie wiarygodnych i ujednoliconych opisów 20 baz danych ogólnych (co stanowi tylko 2,94% wszystkich analizowanych i zweryfikowanych źródeł informacji), np. NEMBASE, The Universal Database of the International Committee on Taxonomy of Viruses, Research in Entomology. Opisano także 19 serwisów wiadomości, m.in. GenomeWeb, Scientific American, Bioinformatyk.eu. Wyselekcjonowano 17 organizacji (np. International Society for Computational Biology, GenomeWeb, Society for Integrative and Comparative Biology) oraz tyle samo list i forów dyskusyjnych (po 2,50%). Przykładowo w bazie zasobów *online* z dziedziny biologii umieszczono bionet, neuroscience, bionet.molbio.emblatbank, Archives of iyor-uk@jiscmail.ac.uk.

Po dokonaniu oceny jakości wybrano 15 bibliograficznych baz danych, w tym np. NEIbank – dotyczącą problematyki ekspresji genów i białek w oku, przydatną dla badaczy genomiki i bioinformatyki, opracowaną przez National Eye Institute, Narodowy Instytut Zdrowia Amerykańskiego Departamentu Zdrowia i Opieki Społecznej oraz 15 pełnotekstowych baz danych (2,21%), m.in. Online mendelian inheritance in animals (OMIA), NeMys: Generic biological information system i Entrez cross-database search.

Wśród zasobów o charakterze społecznościowym wytypowano 13 blogów i 8 stron domowych pracowników naukowych (kolejno 1,91% i 1,18% zweryfikowanych opisów bibliograficznych), wśród których znalazły się blogi: „To jest eksperyment”, poświęcony zagadnieniom z zakresu biologii, medycyny, a także polityki społecznej, w którym umieszczono zapisy obserwacji i raporty z wyników badań, komentarze do artykułów, decyzji, zjawisk społecznych; „Botaniczna Galaktyka”, gdzie są prezentowane najnowsze wyniki badań i analiz z zakresu botaniki; ScienceBlogs, portal naukowy utworzony przez Seed Media Group, którego celem jest umożliwienie naukowcom o podobnych poglądach prowadzenia dyskusji w zakresie nauk biologicznych, medycznych, technologii, filozofii nauki oraz badań behawioralnych.

Po zweryfikowaniu uwzględniono także zawartość stron 13 instytucji i organizacji o charakterze naukowym, krajowych lub międzynarodowych, np. Institute of Arctic Biology, Instytut Biologii Doświadczalnej im. M. Nenckiego, Instytut Biologii Ssaków Polskiej Akademii Nauk w Białowieży, Biological Sciences Curriculum Study (BSCS).

Sprawdzono również 13 spisów bibliograficznych (1,91%), m.in. Biology bibliography, Internet Directory of Botany, Genomika oraz 12 bibliotek (1,76%), w tym np. Virtual Library on Genetics i Polską Bibliotekę Nauki: Kolekcję Nauk Przyrodniczych, kumulujących wybrane źródła naukowe z biologii, genetyki, biotechnologii, ewolucji, genetyki etc.

Wśród zasobów elektronicznych, które poddano ocenie jakości, znalazło się po 5 dokumentów graficznych i muzeów (np. „Zielnik Simona Syrenniususa” z 1613 r., „Krasiejów. Triasowe gady i płazy. Wykopaliska paleontologiczne Instytutu Paleobiologii PAN”, serwis Ogródu Botanicznego Uniwersytetu Jagiellońskiego, SYNTHESYS. Synthesis of Systematic Resources), 4 instytucje rządowe (np. National Science Foundation. Directorate for Biological Sciences – BIO), książki elektroniczne (np. „On-Line Biology Book”, „Płazy i gady Europy”), multimedia oraz materiały szkoleniowe, wykłady i kursy (np. „Marine Genomics Europe” (MGE), „Statystyka w biologii”, „Dr. Wayne’s Biology in 24 hours”, Mendel-Web). Wyróżniono firmy i spółki, organizacje rządowe, programy komputerowe (po 3) oraz jedno (1) materiały konferencyjne: „MMM 2010. Fifth International Conference Multiscale Materials Modeling. Conference Proceedings Biological & Soft Materials” i jedną organizację pożytku publicznego – NANONET Fundację Wspierania Nanonauk i Nanotechnologii.

Wybrane źródła informacji stanowią część dużej liczby zasobów *online* w obszarze szeroko rozumianej biologii. Jak wspomniano, tworzenie warsztatu informacyjnego w dziedzinie nauk biologicznych wymaga stałej i dynamicznej obserwacji środowiska cyfrowego, znacznych nakładów czasu i pracy w celu odszukania i zweryfikowania źródeł informacji, jak również specjalistycznej wiedzy biologicznej, a ponadto powinno być prowadzone w ścisłej współpracy z jednostkami badawczymi i akademickimi w celu uzyskania kompletnego zbioru o wysokiej jakości informacji dziedzinowej.

Wszystkie rodzaje źródeł informacji, które zostały wybrane i udostępnione w formie bazy zasobów cyfrowych, mogą stanowić nie tylko warsztat informacyjny dla studentów, naukowców i dydaktyków biologii, lecz także wspomagać kształtowanie pożądaných kompetencji informacyjnych, w tym umiejętności wartościowania informacji elektronicznej.

## Zakończenie

Działania zmierzające do wprowadzenia koncepcji *information literacy* do programów kształcenia oraz tworzenie dobrych jakościowo zasobów cyfrowych mają na celu rozwijanie kompetencji informacyjnych członków społeczności akademickiej, naukowców i studentów.

W toku przeprowadzonej analizy i oceny jakości źródeł informacji z zakresu biologii wytypowano różnorodne zasoby naukowe dostępne *online*, które mogą zostać potraktowane przez naukowców z tej dyscypliny jako potencjalne źródła informacji. Z uwagi na wielość typów i form tych zasobów, zarówno polskich, jak i zagranicznych, oraz czasochłonność oceny jakości postanowiono – z myślą przede wszystkim o naukowcach zajmujących się też dydaktyką akademicką i studentach – zbudować zestaw sprawdzonych źródeł informacji naukowej, które można zastosować w pracy dydaktycznej i badawczej, realizacji zadań ćwiczeniowych, przygotowaniu prac seminaryjnych, licencjackich lub magisterskich. Należy zakładać, że pogłębiony proces przeszukiwania zasobów internetowych doprowadziłby do odnalezienia jeszcze większej liczby interesujących materiałów naukowych, a więc także do rozbudowania bazy danych, która powstała w wyniku relacjonowanych tu badań. Baza ta zawiera ujednolicone opisy źródeł poddanych ocenie jakości za pomocą przyjętych kryteriów, przeznaczonych dla specyficznego użytkownika reprezentującego profesjonalne dziedzinowe środowisko akademickie. Konieczne wydaje się zatem kontynuowanie procesu ewaluacji jakości i weryfikacji elektronicznych zasobów informacji naukowej z obszaru nauk biologicznych. Powszechny dostęp do zweryfikowanych źródeł, starannie sklasyfikowanych i logicznie uporządkowanych, pozwoliłyby również na wzbogacenie wiedzy wszystkich potencjalnie zainteresowanych członków społeczeństwa informacyjnego.

## BIBLIOGRAFIA

- Association of College and Research Libraries (ACRL). *Information Literacy Competency Standards for Higher Education* [dok. elektr.]. <http://www.ala.org/acrl/standards/informationliteracycompetency> [odczyt: 10.09.2014].
- Barranoik, Lois (2001). *Research Success with Senior High School Students*. *School Libraries Worldwide* Vol. 7, No. 1, p. 28-45.
- Batorowska, Hanna (2013). *Od alfabetyzacji informacyjnej do kultury informacyjnej: rozważania o dojrzałości*. Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, 155 s.

- Bednarek-Michalska, Bożena (2007). *Ocena jakości serwisu internetowego* [dok. elektr.]. [http://moodle.weinoe.us.edu.pl/file.php?file=%2F1%2FTechnologia\\_informacyjna\\_w\\_educacji\\_wczesnoszkolnej%2FKryteria\\_Oceny\\_Stron\\_internetowych-1.doc](http://moodle.weinoe.us.edu.pl/file.php?file=%2F1%2FTechnologia_informacyjna_w_educacji_wczesnoszkolnej%2FKryteria_Oceny_Stron_internetowych-1.doc) [odczyt: 12.08.2014].
- Biblioteka w szkole. Schemat UKD* [dok. elektr.]. [http://www.bibliotekawshkole.pl/archiwum/2009/03/BwS\\_03\\_2009\\_UKD\\_WWW\\_schemat.pdf](http://www.bibliotekawshkole.pl/archiwum/2009/03/BwS_03_2009_UKD_WWW_schemat.pdf) [odczyt: 11.09.2014].
- Bologna Process. Towards the European Higher Education Area* [dok. elektr.]. <http://www.ond.vlaanderen.be/hogeronderwijs/bologna/ql/ql.asp> [odczyt: 19.09.2014].
- Bowden, Teresa S.; DiBenedetto, Angela (2001). *Information literacy in biology laboratory session. An example of librarian-faculty collaboration*. *Research Strategies* Vol. 18, p. 143-149.
- Brem, Sarah K.; Russel, Janet; Weems, Lisa (2001). *Science on the Web: Student evaluations of scientific arguments*. *Discourse Processes* Vol. 32 Issue2 / 3, p. 191-213.
- Callinan, Joanne E. (2004). *Information-seeking behaviour of undergraduate biology students*. *Library Review* Vol. 54, No. 2, p. 86-99.
- Ciołek, Matthew T. (1997). *Information Quality WWW Virtual Library: the Internet guide to construction of quality online resources* [dok. elektr.]. <http://www.ciolek.com/WWWVLIInfoQuality.html> [odczyt: 12.08.2014].
- Dold, Claudia J. (2014). *Critical Information Literacy: A Model for Transdisciplinary Research in Behavioral Sciences*. *The Journal of Academic Librarianship* Vol. 40, p. 179-184.
- Elmborg, James (2006). *Critical information literacy: Implications or instructional practice*. *The Journal of Academic Librarianship* Vol. 32, Issue 2, p. 192-199.
- Greenleaf, Cynthia L. et al. (2011). *Integrating literacy and science in biology: Teaching and Learning impacts of reading apprenticeship professional development*. *American Educational Research Journal* Vol. 48, p. 647-717.
- Instytut Genetyki i Hodowli Zwierząt PAN (IGHZ). *Centrum Doskonałości ANIMBIO-GEN* [dok. elektr.]. <http://www.ighz.edu.pl/?p0=3&p1=140> [odczyt: 15.08.2014].
- INTUTE* [dok. elektr.]. <http://www.intute.ac.uk/> [odczyt: 19.09.2014].
- Jagiellońska Biblioteka Cyfrowa (JBC). *Klasyfikacja systematyczna* [dok. elektr.]. <http://jbc.bj.uj.edu.pl/dlibra/collectiondescription?dirids=117> [odczyt: 11.09.2014].
- Julien, Heidi; Barker, Susan (2009). *How high-school students find and evaluate scientific information: a basis for information literacy skills development*. *Library & Information Science Research* Vol. 31, p. 12-17.
- Komitet Badań Naukowych (KBN). *Polska Klasyfikacja Tematyczna. 330 000 Nauki Biologiczne* [dok. elektr.]. <http://kbn.icm.edu.pl/pub/kbn/docs/pkt.html> [odczyt: 11.09.2014].
- Korobili, Stella et al. (2008). *Assessing information literacy skills in the Technological Education Institute of Thessaloniki, Greece*. *Reference Services Review* Vol. 37, No. 3, p. 340-354.
- Lau, Jesus (2011). *Kompetencje informacyjne w procesie uczenia się przez całe życie. Wytuczne* [dok. elektr.]. <http://www.ifla.org/files/assets/information-literacy/publications/ifla-guidelines-pl.pdf> [odczyt: 11.09.2014].
- Lee, Jennifer; MacMillan, Don (2004). *Evolving instruction in biology: using the web to improve the in-class instruction*. *Reference Services Review* Vol. 32, No. 4, p. 374-382.
- Lorenzen, Michael (2001). *The land of confusion? High School Students and their use of the World Wide Web for research*. *Research Strategies* Vol. 18, p. 151-163.
- Mansourian, Yazdan; Madden, Andrew D. (2007). *Perceptions of the web as a search tool amongst researchers in biological sciences*. *New Library World* Vol. 108, No. 9 / 10, p. 407-423.

- Milner-Bolotin, Marina; Nashon, Samson Madera (2012). *The essence of student visual-spatial literacy and higher order thinking skills in undergraduate biology*. Photoplasma Vol. 249 (Suppl 1), p. S25-S30.
- Ministerstwo Nauki i Szkolnictwa Wzwyższego (MNiSW). *Krajowe Ramy Kwalifikacji uczenia się przez całe życie* [dok. elektr.]. <http://www.nauka.gov.pl/krajowe-ramy-kwalifikacji-szkolnictwo> [odczyt: 12.08.2014].
- Nitz, Sandra et. al. (2014). *Do student perceptions of teaching predict the development of representational competence and biological knowledge?* Learning and Instruction Vol. 31, p. 13-22.
- Olin & Uris Libraries; Cornell University Library. *Evaluating Web Sites: Criteria and Tools* [dok. elektr.]. <http://olinuris.library.cornell.edu/ref/research/webeval.html> [odczyt: 12.08.2014].
- Oxford Brookes University (2013). *Programme Specification BSc (Hons) Biology (BG), Faculty of Health and Life Sciences* [dok. elektr.]. <https://www.brookes.ac.uk/about/directorates/asa/registry/apqo/programmes/mc10-bg.pdf> [odczyt: 10.09.2014].
- The Quality Assurance Agency for Higher Education (QAAHE) (2007). *Biosciences* [dok. elektr.]. <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/Biosciences07.pdf> [odczyt: 15.08.2014].
- Queen Mary University of London (2009). *Programme Specification BSc Biology, School of Biological & Chemical Sciences* [dok. elektr.]. [http://www.arcs.qmul.ac.uk/qa/programme\\_specifications/sbcs/September%202011/PS\\_C100\\_BScBiology\\_SBCS\\_23022011.pdf](http://www.arcs.qmul.ac.uk/qa/programme_specifications/sbcs/September%202011/PS_C100_BScBiology_SBCS_23022011.pdf) [odczyt: 10.09.2014].
- Rozporządzenie Ministra Nauki i Szkolnictwa Wzwyższego z dnia 2 listopada 2011 r. w sprawie Krajowych Ram Kwalifikacji dla Szkolnictwa Wzwyższego. Dz.U. 2011 nr 253 poz. 1520.
- Sapa, Remigiusz. *Kryteria oceny jakości zasobów informacji na potrzeby projektu SYNAT* [maszynopis].
- SCONUL (2011). *The SCONUL Seven Pillars of Information Literacy Core Model For Higher Education* [dok. elektr.]. <http://www.sconul.ac.uk/sites/default/files/documents/coremodel.pdf> [odczyt: 10.09.2014].
- Smith, Alistair (1996). *Criteria for Evaluation of Internet Information Resources* [dok. elektr.]. <http://www.vuw.ac.nz/~agsmith/evaln/index.htm> [odczyt: 12.08.2014].
- University of Derby (2012). *Programme specification BSc Biology* [dok. elektr.]. <http://www.derby.ac.uk/media/derbyacuk/contentassets/documents/programmespecifications/science/BSc-Biology-Programme-Specification.pdf> [odczyt: 10.09.2014].
- Urquhart, Christine; Rowley, Jennifer (2007). *Understanding student information behavior in relation to electronic information services: Lessons from longitudinal monitoring and evaluation, Part 2*. Journal of the American Society for Information Science and Technology Vol. 58, Issue 8, p. 1188-1197.
- Winterman, Brian (2009). *Building Belter biology undergraduates through information literacy integration* [dok. elektr.]. Issues in Science and Technology Librarianship No. 58, p. 1-9 <http://www.istl.org/09-summer/refereed1.html> [odczyt: 12.08.2014].

---

# Wybrane zagadnienia wizualizacji w nauce o informacji

Paulina Motylińska

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

Wizualizacja jest zagadnieniem interdyscyplinarnym, stanowiącym przedmiot zainteresowania specjalistów z wielu dyscyplin nauki i sztuki, m.in. nauki o informacji, bibliologii, informatyki, grafiki komputerowej. Każda z wymienionych dyscyplin bada proces wizualizacji z własnej perspektywy, zwracając uwagę na inne elementy wchodzące w jego skład. W artykule skupiono się na wybranych aspektach wizualizacji, które wchodzą w zakres zainteresowań nurtów badawczych informatologii. Ustalenie głównych zagadnień poprzedziła analiza i krytyka piśmiennictwa. Stwierdzono, że w polskiej literaturze przedmiotu, publikowanej w ramach nauki o informacji i bibliologii, brakuje przeglądowych prac dotyczących wizualizacji informacji, które mogłyby być przydatne dla początkujących badaczy tego zagadnienia. Jedną z niewielu publikacji, omawiającą ogólne zagadnienia wizualizacji w kontekście informatologii, jest książka Veslavy Osińskiej *Wizualizacja i wyszukiwanie dokumentów* [Osińska 2010b]. Dostępne są także podręczniki informatyczne, dotyczące wizualizacji informacji, pisane przez specjalistów w tej dziedzinie.

Znaczenie wizualizacji w nauce o informacji zwiększa się wraz ze zmianami zachodzącymi w samej dyscyplinie oraz w postrzeganiu roli pracownika informacji. Rozwój społeczeństwa informacyjnego, masowa produkcja zasobów informacyjnych i inne czynniki wpływające na rozbudowę informacyjnego sektora gospodarki sprawiły, że informacja traktowana jest jako towar, a zarządzanie informacją stało się niezbędnym elementem funkcjonowania każdej organizacji.



Uznaje się nawet, że posiadanie kompetencji wizualnych (ang. *visual literacy*) jest niezbędne, aby właściwie funkcjonować we współczesnym świecie [m.in. Dylak 2012].

W artykule zwrócono uwagę na różnice terminologiczne i problemy definicyjne w zakresie pojęć: „wizualizacja danych”, „wizualizacja informacji” i „wizualizacja naukowa”. Uwzględniono także podstawowe zagadnienia dotyczące wizualizacji w podziale na dwie główne grupy:

- zagadnienia związane ze wsparciem wizualizacyjnym oferowanym użytkownikowi systemów informacyjno-wyszukiwawczych;
- zagadnienia wizualizacji w nauce.

Wszystkie zaprezentowane w artykule problemy wymagają pogłębienia i dalszych badań. Celem niniejszego opracowania jest jedynie ogólna prezentacja zagadnień związanych z wizualizacją, które mogą stanowić obiekt zainteresowania specjalistów informacji. Artykuł stanowi wstęp i punkt odniesienia dla dalszej analizy.

## Problemy definicyjne

Próby jednoznacznego zdefiniowania terminów „dane”, „informacje” i „wiedza” były podejmowane w nauce o informacji i bibliotekoznawstwie od najwcześniejszych lat istnienia tych dyscyplin. Ten sam problem terminologiczny pojawia się w przypadku definiowania pojęć „wizualizacja danych” i „wizualizacja informacji”. Sama wizualizacja traktowana jest zarówno jako „proces reprezentacji wyrazów i pojęć w formie wizualnej”, jak i zastosowanie „narzędzi i metod interpretacji danych graficznych oraz generowania obrazów na podstawie danych wielowymiarowych” [Osińska 2010b, s. 15]. W powyższej definicji podkreślono, że obecnie wizualizacja obejmuje nie tylko proces tworzenia graficznych reprezentacji danych, ale również szereg specjalistycznych narzędzi i metod własnych, stosowanych w obrazowaniu danych złożonych. Jedna z często powtarzanych w literaturze przedmiotu definicji wizualizacji pochodzi z książki Stuarta Carda, Jocka Mackinlaya i Bena Shneidermana z 1999 r. i brzmi następująco: wizualizacja to „wizualna reprezentacja informacji służąca wzmocnieniu funkcji poznawczych użytkowników” [Card, Mackinlay, Shneiderman 1999, p. 7]. Osińska w swojej definicji wizualizacji zaakcentowała wykorzystanie specjalistycznych narzędzi w procesie wizualizowania, natomiast Card, Mackinlay i Shneiderman skupili się na podstawowym celu wizualizacji. Wśród definicji wizualizacji, dostępnych w literaturze przedmiotu, najczęściej spotykany podział na „wizualizację informacji” i „wizualizację danych” opiera się na różnicy

w poziomie abstrakcyjności wizualizowanych danych. W artykule postanowiono stosować termin „wizualizacja informacji” jako termin ogólny dla określenia całości zagadnienia.

Michael Friendly w publikacji z 2009 roku podaje, że termin „wizualizacja informacji” (ang. *information visualization*) jest terminem najszerszym. Wizualizacja informacji zawiera w sobie wizualizację danych (ang. *data visualization*) i wizualizację naukową (ang. *scientific visualization*). Friendly definiuje wizualizację informacji jako „wizualną reprezentację dużych zbiorów danych nie-numerycznych, takich jak pliki i kody źródłowe w systemach oprogramowania, biblioteczne i bibliograficzne bazy danych, sieci relacji w internecie itd.” [Friendly 2009, dok. elektr.]. W definicji „wizualizacji danych” Friendly określa dane jako „informacje, które zostały wyodrębnione w schematycznej formie, włączając atrybuty i zmienne pojedynczych jednostek informacji” [Friendly 2009, dok. elektr.]. Definicję wizualizacji danych Friendly’ego uzupełniono przykładami z zakresu grafiki statystycznej i kartografii tematycznej. Wizualizacją zjawisk i przedmiotów trójwymiarowych zajmuje się natomiast wizualizacja naukowa. Obejmuje ona m.in. dane architektoniczne, zjawiska meteorologiczne oraz dane biologiczne i medyczne. Wizualizacja naukowa kładzie nacisk na realistyczne przedstawienie symulacji m.in. objętości i powierzchni prezentowanych zjawisk, a także na pokazanie zmienności tych zjawisk, np. wskutek lub w czasie działania na nie czynników zewnętrznych, takich jak światło [Friendly 2009, dok. elektr.].

Melanie Tory i Torsten Möller stwierdzili, że tradycyjnym podziałem wizualizacji jest podział na wizualizację informacji (ang. *information visualization*) i wizualizację naukową (ang. *scientific visualization*), nieuwzględniający wyodrębnienia wizualizacji danych. Tory i Möller podali kilka propozycji rozróżnień wizualizacji informacji i wizualizacji naukowej, które opracowali na podstawie analizy literatury przedmiotu. Jedną z nich jest ustalenie kontekstu zastosowania danej wizualizacji; jeśli wizualizacja wykorzystywana jest w celach naukowych, to jest to wizualizacja naukowa. Wizualizację stosowaną w celach nienaukowych można skategoryzować jako wizualizację informacji. Innym sposobem odróżnienia wizualizacji naukowej od wizualizacji informacji jest określenie poziomu abstrakcyjności wizualizowanych danych. W zależności od tego, czy wizualizowane dane poparte są fizycznymi dowodami i dotyczą zjawisk fizycznych, czy są to dane abstrakcyjne, można mówić o wizualizacji naukowej lub wizualizacji informacji. W podsumowaniu autorzy stwierdzają, że wizualizacja naukowa najczęściej dotyczy danych naukowych, które umiejscowione są w konkretnej przestrzeni, nierozzerwalnie związanej z danymi, natomiast wizualizacja informacji obejmuje dane abstrakcyjne, nie umiejscowione w przestrzeni. Wśród przykła-

dów wizualizacji naukowej znalazły się trójwymiarowe obrazy medyczne, a jako przykład wizualizacji informacji zaproponowano wizualizacje danych finansowych lub kolekcji dokumentów [Tory, Möller 2004, dok. elektr.].

## Klasyfikacje wizualizacji

Trudności w jednoznacznym zdefiniowaniu danych i informacji w kontekście wizualizacji spowodowały podjęcie prób stworzenia klasyfikacji wizualizacji w oparciu o typ danych, które są poddawane temu procesowi. Jednym z częściej wspominanych w literaturze przedmiotu podziałów wizualizacji jest taksonomia Bena Shneidermana, przedstawiona w artykule z 1996 roku. Shneiderman zaproponował klasyfikację wizualizacji ze względu na typ wizualizowanych danych oraz „zadanie” wizualizacji. Dane zostały podzielone na siedem grup: dane tworzące jedno-, dwu- i trójwymiarowe obrazy, dane tymczasowe, dane wielowymiarowe, dane o strukturze hierarchicznej i dane o strukturze sieciowej [Shneiderman 1996, dok. elektr.]. Wśród zadań wizualizacji znalazły się: ukazywanie przekroju zjawiska (ang. *overview*), zbliżenie na określone elementy zjawiska (ang. *zoom*), filtrowanie informacji (ang. *filter*), ukazywanie szczegółów z całości wizualizacji na żądanie (ang. *details-on-demand*), obrazowanie powiązań między danymi (ang. *relate*), przedstawianie historii lub przebiegu zjawiska w czasie (ang. *history*) oraz wydobywanie interesujących elementów (ang. *extract*) [Shneiderman 1996, dok. elektr.].

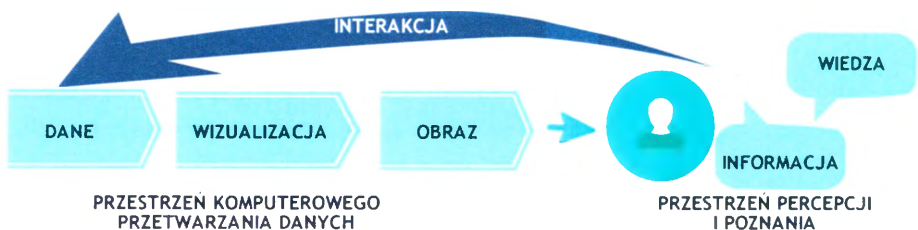
Inną propozycją podziału jest taksonomia technik wizualizacji zaproponowana przez Eda H. Chi. W oparciu o *Data State Model*, czyli model procesu wizualizacji uwzględniający transformacje danych podczas wizualizowania, Chi przeanalizował 36 technik wizualizacji, grupując je następująco:

- wizualizacje naukowe,
- wizualizacje oparte na danych geograficznych,
- wizualizacje dwuwymiarowe,
- wizualizacje wielowymiarowe,
- wizualizacje informacji krajobrazowych i przestrzennych,
- wizualizacje hierarchiczne,
- wizualizacje relacji,
- wizualizacje informacji tekstowych,
- wizualizacje sieci internetowej,
- wizualizacje na podstawie danych z arkusza kalkulacyjnego [Chi, dok. elektr.].

Chi zastosował różne kryteria podziału, co można uznać za jeden z podstawowych błędów podczas tworzenia klasyfikacji. Analizując powyższe przykłady typologii wizualizacji, należy zwrócić uwagę na płynność granic wszystkich kategorii; przyporządkowanie wizualizacji do określonej grupy może sprawiać trudności ze względu na niejednoznaczności w klasyfikowaniu wizualizowanych danych.

## Modele procesu wizualizacji

Min Chen, David Ebert, Hans Hagen i in. zwrócili uwagę na dowolne i zamienne stosowanie pojęć „dane”, „informacje” i „wiedza” w kontekście wizualizacji. W oparciu o tradycyjny sposób definiowania pojęć „dane”, „informacje” i „wiedza” stwierdzili, że należy je rozróżnić także w kontekście wizualizacji. Tradycyjny sposób definiowania określa dane jako symbole, które, po przetworzeniu i w określonym kontekście, stają się informacjami, natomiast wiedza to przetworzone w umyśle odbiorcy informacje. Dane poddawane są procesowi wizualizacji, w wyniku którego powstaje ich przedstawienie graficzne, odbierane przez użytkownika wizualizacji jako informacja. Informacja następnie zostaje przetworzona w wiedzę [Chen et al., 2009, s. 13]. Bazując na modelu łańcucha dane-informacje-wiedza Chen i in. przedstawili prosty model procesu wizualizacji (rysunek 1).



Rysunek 1. Model procesu wizualizacji

Źródło: Opracowanie na podstawie [Chen et al. 2009, p. 13]

Innym przykładem modelu procesu wizualizacji jest czteroetapowy model uwzględniony w książce Colina Ware'a z 2004 roku. Etap pierwszy procesu wizualizacji obejmuje gromadzenie i przechowywanie danych. Etap drugi to przetwarzanie zgromadzonych danych do postaci zrozumiałej dla odbiorcy wizualizacji. Podczas etapu trzeciego za pomocą odpowiednich narzędzi i oprogramowania

komputerowego następuje wyświetlenie wizualizacji na ekranie. Etap czwarty realizowany jest w umyśle odbiorcy wizualizacji, gdzie aktywowane są systemy percepcji i poznania [Ware 2004, p. 5]. Analizując model procesu wizualizacji Ware'a, warto zwrócić uwagę na występowanie pętli odpowiedzi zwrotnych na wszystkich etapach tworzenia wizualizacji. W trakcie projektowania wizualizacji projektant może powrócić do wszystkich poprzednich etapów w celu udoskonalenia efektu końcowego i dostosowania wizualizacji do potrzeb odbiorcy, np. powrócić do etapu pierwszego, by zebrać więcej potrzebnych danych.

## **Wizualizacja w nauce o informacji**

Powiązane z zagadnieniem wizualizacji obszary badawcze nauki o informacji można podzielić na dwie główne grupy. Pierwszą z nich stanowią zagadnienia dotyczące „wspomagania wizualizacyjnego” użytkownika systemów informacyjno-wyszukiwawczych. W tej grupie umieścić można problematykę udoskonalania możliwości wyszukiwawczych systemów informacyjnych, włączając w nią projektowanie interfejsów użytkownika z elementami wizualizacji informacji i rozwój wyszukiwarek graficznych, a także tworzenie multimedialnych baz danych. Druga grupa zagadnień dotyczy wizualizacji w nauce, a w szczególności przedstawiania wizualizacyjnego wyników badań naukowych. Szczególnie istotne są takie obszary badań, jak tworzenie nowych źródeł informacji, badania informetryczne i mapowanie nauk.

## **Interfejs użytkownika i użyteczność**

Chaomei Chen stwierdza, że wizualizacja informacji jest jednym z perspektywicznych nurtów we współczesnym projektowaniu interfejsów użytkownika [Chen 2006, p. 1]. Projektanci interfejsów dążą do zwiększenia możliwości bezpośredniego manipulowania obiektami składającymi się na interfejs, opierając się na zasadzie „wskaż i kliknij, przeciągnij i upuść” (ang. *point-and-click and drag-and-drop*) [Chen 2006, p. 1]. Budowane są interfejsy graficzne, w których m.in. wszelkie możliwe do zastąpienia elementy słowne zastępowane są odpowiednikami graficznymi i ikonami, a struktura serwisu lub informacji prezentowana jest w postaci drzewa hierarchicznego. Podczas projektowania takich rozwiązań często pomijane jest zagadnienie użyteczności (ang. *usability*) i dostępności (ang. *accessibility*) interfejsu użytkownika. Według Chen jednym z dziesięciu głównych nierozwiązanych problemów z wizualizacją informacji

jest projektowanie interfejsów w taki sposób, by wizualizacje spełniały kryteria użyteczności [Chen 2005, p. 12]. W ostatnim czasie problem ten zyskuje na znaczeniu, zwłaszcza w świetle zmian zachodzących w środowisku internetu, gdzie zaczęto zwracać większą uwagę na dostępność serwisów internetowych dla jak największej grupy użytkowników. Jedne z najpopularniejszych wytycznych dotyczących dostępności serwisów internetowych – Web Content Accessibility Guidelines 2.0 (WCAG 2.0) – zyskały status normy ISO (ISO/IEC 40500:2012). Dokument WCAG 2.0 zawiera podstawowe wytyczne dla projektowania serwisów internetowych, tak aby serwisy te były jak najbardziej dostępne, szczególnie dla osób niepełnosprawnych [WCAG 2.0, dok. elektr.]. Zmiany widać nawet w obowiązującym w Polsce prawie; strony internetowe prowadzone przez podmioty publiczne powinny zostać dostosowane do wymogów WCAG 2.0 do 2015 r. [Rozporządzenie... 2012]. Zagadnienia użyteczności i dostępności interfejsów są coraz częściej uwzględniane w badaniach informatologicznych z uwagi na zainteresowanie specjalistów z tej dziedziny tematyką interakcji użytkowników z informacją w środowisku internetowym.

Jedną z najprostszych i jednocześnie najczęściej powtarzanych w literaturze przedmiotu definicji użyteczności jest „łatwość, z jaką produkt może być używany” oraz „łatwość użycia” [Chowdhury, Chowdhury 2011, p. 14; Lauesen 2005, p. 6]. Podobnie definiuje użyteczność Jakob Nielsen, jeden z uznanych autorytetów w dziedzinie użyteczności – „użyteczność jest cechą jakości, przy pomocy której ocenia się, jak łatwy w użyciu jest interfejs użytkownika” [Nielsen 2012, dok. elektr.]. Dostępność, w wąskim ujęciu, definiowana jest natomiast jako dostosowanie interfejsu do potrzeb użytkowników niepełnosprawnych, a w szerszym ujęciu można mówić o tzw. uniwersalnej użyteczności (ang. *universal usability*), odnoszącej się do możliwości korzystania z interfejsu przez jak największą grupę użytkowników, niezależnie od ich wieku, doświadczenia czy fizycznych ograniczeń [Garlock, Piontek 1999, p. 61; Lynch, Horton 2008, p. 53; Stone et al. 2005, p. 177]. Ocena użyteczności interfejsu jest niezbędna do sprawdzenia, czy system spełnia potrzeby odbiorców. W jej ramach oceniane są także wizualizacje. Takie podejście do oceny użyteczności wizualizacji jest uzasadnione, gdy elementy wizualizacyjne stanowią integralną część interfejsu, Emile Morse, Michael Lewis i Kai Olsen zaproponowali jednak metodę oceny wizualizacji, które są częścią interfejsu, po ich uprzednim z niego wyodrębnieniu. Stwierdzili, że taki sposób oceny pozwala na lepsze rozpoznanie roli wizualizacji w całym systemie informacyjnym [Morse, Lewis, Olsen 2000, p. 637].

Problem dostępności wizualizacji informacji został zauważony także przez Catherine Plaisant. W artykule z 2005 r. Plaisant przedstawiła propozycje rozwiązań możliwych do zastosowania przy tworzeniu wizualizacji informacji geogra-

ficznych, tak by uwzględnione zostały zasady uniwersalnej użyteczności [Plaisant 2005, dok. elektr.]. Podczas projektowania pod kątem użyteczności należy zwrócić uwagę na trzy główne problematyczne aspekty:

- różnorodność użytkowników, m.in. ich wiek, niepełnosprawność, język,
- niejednorodność stosowanych przez użytkowników rozwiązań technologicznych, np. rozmiaru ekranu i szybkości łącza internetowego,
- różnice w wiedzy użytkowników, dotyczące poziomu wiedzy ogólnej, znajomości danego interfejsu czy systemu informacyjno-wyszukiwawczego.

Poszukiwanie publikacji zawierających wskazówki odnośnie do projektowania wizualizacji zgodnie z zasadami uniwersalnej użyteczności nie przyniosło zadowalających rezultatów. Przy projektowaniu wizualizacji stosuje się powszechnie przyjęte zasady projektowania interfejsów użytkownika, zgodne z zasadami użyteczności i dostępności, sformułowane m.in. przez Jakoba Nielsena, Bena Shneidermana czy Sorena Lauesena [Nielsen 2000; Shneiderman 1992; Lauesen 2005]. Istnieje niewiele wyspecjalizowanych opracowań dotyczących dostępności i użyteczności w kontekście wizualizacji, jednak w literaturze przedmiotu dostępne są raporty z badań dostępności i użyteczności konkretnych narzędzi do wizualizacji informacji. Wingyan Chung, Hsinchun Chen, Luis Chaboya i in. w artykule z 2005 r. zaprezentowali badanie użyteczności narzędzia COPLINK. COPLINK pozwala na tworzenie przestrzenno-czasowych wizualizacji zdarzeń, stosowanych w pracy policji podczas trwania śledztwa. W badaniu użyteczności porównywano możliwości przetwarzania informacji wizualnych w aplikacji COPLINK i w programie Microsoft Excel. Wizualizacje tworzone przez COPLINK uzyskały znacząco wyższe wyniki w ocenie poziomu użyteczności (łatwość użycia, prezentacja informacji, projekt interfejsu) niż tabelaryczne przedstawienie informacji przez program Excel [Chung et al. 2005, p. 152].

Oceną użyteczności interfejsów stron internetowych udostępniających mapy zajęli się Annu-Maaria Nivala, Stephen Brewster i L. Tiina Sarjakoski. Ocenie eksperckiej i badaniu użyteczności z udziałem użytkowników zostały poddane Google Maps, MSN Maps & Directions, MapQuest i Multimap. Badanie wykazało istnienie deficytów użyteczności zarówno w samym interfejsie użytkownika, jak i w wizualizacji map i narzędzi pozwalających na bezpośrednie manipulowanie elementami graficznymi [Nivala, Brewster, Sarjakoski 2008, p. 129]. Tematyką projektowania interfejsów użytkownika do serwisów udostępniających geowizualizacje oraz oceną ich użyteczności zajmowali się także Etien Koua, Alan MacEachren i Menno-Jan Kraak [Koua, MacEachren, Kraak 2006], Catherine Emma Jones, Mordechai Haklay, Sam Griffiths i Laura Vaughan [Jones et al. 2009], Olivier Swienty, Tumasch Reichenbacher, Simone Reppermund i Josef Zihl [Swienty et al. 2008] oraz Petr Kubíček i Čeněk Šašinka [Kubíček, Šašinka

2011]. Większość odnalezionych publikacji dotyczących zagadnień użyteczności wizualizacji dotyczyła właśnie wizualizacji w środowisku informacji geograficznej. Zainteresowanie badaniami użyteczności geowizualizacji może być związane z dużą dostępnością narzędzi prezentujących mapy oraz ich popularnością wśród osób korzystających z internetu (m.in. serwis Google Maps). Użytkownicy ogólnodostępnych geowizualizacji nie muszą posiadać wyspecjalizowanej wiedzy na temat działania tego typu narzędzi, dlatego też łatwiej jest przeprowadzić testy użyteczności z udziałem rzeczywistych odbiorców tych usług.

## Wizualizacja w wyszukiwaniu informacji

Wyszukiwanie informacji, jako jeden z klasycznych obszarów badawczych w nauce o informacji, jest także jednym z najczęściej analizowanych tematów w powiązaniu z wizualizacją. Wyszukiwanie wizualizacyjne (ang. *visual search*) uznawane jest za jeden z wyspecjalizowanych rodzajów wyszukiwania informacji (ang. *information retrieval*), podczas którego użytkownik otrzymuje dodatkowe „wsparcie wizualizacyjne” [Pajić 2014]. Pomoc wizualizacyjna może przyjąć różne formy, np. graficznej reprezentacji wyszukanych dokumentów i stron internetowych, graficznego przedstawienia wykorzystanych słów kluczowych lub wizualizacji powiązań między dokumentami. Dejan Pajić podaje przykłady terminów stosowanych dla określenia wyszukiwania wizualizacyjnego, m.in. *visual information retrieval*, *visualization of information retrieval results*, *visualization-based information retrieval* [Pajić 2014, s. 146]. W niniejszym artykule skupiono się głównie na wyszukiwarkach internetowych, jednak elementy wizualizacyjne występują także w innych systemach informacyjno-wyszukiwawczych, m.in. bazach danych (np. EBSCO) i bibliotekach cyfrowych.

Literaturę przedmiotu dotyczącą wizualizacji w kontekście wyszukiwania informacji można podzielić na grupy tematyczne, które nie są rozłączne i w ogólny sposób pokazują analizowane zagadnienia:

- opracowania ogólne, poruszające tematykę m.in. wyszukiwania informacji w internecie i zachowań informacyjnych w kontekście wizualizacji [m.in. Alhenshiri, Blustein 2011; Hoeber, Yang 2010; Koshman 2006; Newby 2002];
- opisy prototypów i projektów systemów informacyjno-wyszukiwawczych z elementami wizualizacji [m.in. Ahn, Brusilovsky 2013; Chu et al. 2007; Fluit, Sabou, Harmelen 2006; Osińska, Bala, Gawarkiewicz 2012; Pajić 2014; Wiza 2012];
- porównanie systemów z elementami wizualizacji i systemów tradycyjnych, tekstowych [m.in. Koshman 2004; Rivadeneira, Bederson 2003];



- oceny systemów informacyjno-wyszukiwawczych z elementami wizualizacji [m.in. Ridsen et al. 2000; Shiri 2008; Sutcliffe, Ennis, Hu 2000].

Natalia Pamuła-Cieślak w następujący sposób grupuje rozwiązania wizualizacyjne stosowane w wyszukiwarkach internetowych:

- elementy graficzne w logo wyszukiwarki, interfejsie wyszukiwawczym i na całej stronie;
- elementy graficzne pomagające użytkownikowi w wyszukiwaniu;
- graficzne porządkowanie wyników wyszukiwania;
- graficzna prezentacja wyników wyszukiwania [Pamuła-Cieślak 2014, s. 259].

Elementy wizualne wspomagające użytkownika podczas wyszukiwania to m.in. filtry pozwalające na zawężenie wyników wyszukiwania lub zmianę kwerendy wyszukiwawczej oraz prezentacja najczęściej wyszukiwanych, powiązanych ze sobą zapytań. Anwar Alhenshiri i James Blustein podzielili elementy wizualizacyjne opcji pomocy dla użytkownika w formułowaniu kwerendy wyszukiwawczej na wizualną pomoc w rozszerzeniu kwerendy i wizualne określenie wagi relewancji danego dokumentu w wynikach kwerendy [Alhenshiri, Blustein 2011, p. 322-323]. Wśród przykładów projektów systemów, które pomagają użytkownikowi w rozbudowaniu lub zawężeniu formułowanych przez niego pytań, można wskazać projekt Orlanda Hoebera i Xue Dong Yang. Hoeber i Yang zaproponowali zastosowanie kolorowego histogramu, prezentującego alternatywne terminy wyszukiwawcze. Zostały one zaczerpnięte z wyszukanych dokumentów, umiejscowionych na kilku najwyższych miejscach w rankingu wyszukiwarki, a użytkownik mógł bezpośrednio wykorzystać je do zmiany kwerendy [Hoeber, Yang 2008]. Podobny system pomocy w modyfikowaniu terminów wyszukiwawczych został zastosowany w przeznaczony dla dzieci wyszukiwarce Quintura. Po wyświetleniu wyników wyszukiwania na ekranie użytkownika pojawia się chmura zawierająca powiązane terminy wyszukiwawcze, które interaktywnie można dodawać do kwerendy (rysunek 2).

Narzędzie wizualizacyjne do obrazowania poziomu relewancji wyszukanego dokumentu zostało zaprojektowane m.in. przez Jesusa Vegasa, Pabla de la Fuente i Fabia Crestanego. Vegas i in. zaproponowali zastosowanie narzędzia *Webdocball*, pozwalającego na graficzne przedstawienie oszacowanej zgodności poszczególnych podstron serwisu internetowego z hasłem wyszukiwawczym. Wizualizacja tworzona przez *Webdocball* przybiera formę okręgów, których środek reprezentuje cały dokument, a poziom zgodności oznaczany jest kolorystycznie [Vegas, Fuente, Crestani 2003, p. 354].

Podczas korzystania z tradycyjnych wyszukiwarek użytkownik otrzymuje listę wyników wyszukiwania, zawierającą tytuły dokumentów wraz z odnośnikami. Alternatywę dla tradycyjnego sposobu prezentacji listy rankingowej oferują wy-



Rysunek 2. Interfejs użytkownika do wyszukiwarki Quintura Kids  
 Źródło: Quintura Kids [dok. elektr.], <http://quinturakids.com/> [odczyt: 6.09.2014]

szukiwarki klastrowe, które dzielą otrzymane wyniki wyszukiwania na klastry. Mogą być one przedstawione tradycyjnie, jako spis nazw poszczególnych grup, lub w sposób graficzny. Jedną z wyszukiwarek, która prezentuje klastry w postaci graficznej, jest Carrot Search. Użytkownik ma możliwość wybrania sposobu prezentacji wyników wyszukiwania; wyniki mogą być przedstawione w formie tradycyjnych klastrów, okręgów (rysunek 3) i „Foam Tree” (rysunek 4). Inne wyszukiwarki, które oferują różnorodne niestandardowe formy wsparcia wizualizacyjnego dla użytkownika, to m.in. Wolfram Alpha [Wolfram Alpha, dok. elektr.] i VisGets [VisGets, dok. elektr.]. Szeroki opis różnorodnych narzędzi wizualizacyjnych można znaleźć w artykule Orlanda Hoeblera i Xue Dong Yang z 2010 r. [Hoerber, Yang 2010], a dodatkowe przykłady w artykule Billa Kulesa, Maxa Wilsona i Bena Shneidermana z 2008 r. [Kules, Wilson, Shneiderman 2008, dok. elektr.]. W artykule Hoeblera i Yang narzędzia te zostały podzielone na trzy główne grupy: narzędzia do wizualizacji kwerendy wyszukiwawczej, wizualizacji dokumentu i wizualizacji dokumentu zastępczego, obejmującego metadane dokumentu i fragment zawartości dokumentu [Hoerber, Yang 2010, p. 209].

Największy rozkwit funkcjonowania wyszukiwarek umożliwiających graficzne przedstawienie wyników wyszukiwania przypada na lata 2008-2010 [Pamuła-Cieślak 2014, s. 263], jednak pierwsza aplikacja tego typu, KartOO, została zaprojektowana już w 2001 r. [Dürsteler 2002, dok. elektr.]. Większość narzędzi działających w tamtym okresie, wśród których wymienić można np. Kool Torch, Ujiko czy Search Me, nie jest już dostępna. Część z nich udostępniona jest tyl-



Top 26 results of about 5,150,000 for information literacy

[Information Literacy Competency Standards for Higher Education](#)  
NOTE: While the **Information Literacy Competency Standards** for higher education during 2013-14, an ACRL task force is extensively...

[Introduction to Information Literacy - American Library Association](#)  
**Information Literacy** is the set of skills needed to find, evaluate, analyze, and begining of the 21st century has been called the **Information...**

[Information Literacy - Wikipedia, The Free Encyclopedia](#)  
The United States National Forum on **Information Literacy** defines one able to identify, locate, evaluate, and effectively use that...

[Information Literacy - Home - University of Guelph](#)  
**Information Literacy** is the ability to identify what **information** is, how **information** is organized, identify the... **Information Literacy** Portal

[Project Information Literacy](#)  
Project **Information Literacy** (PIL) is a large-scale, national study about habits, conducted in partnership with the University of...

[Information Literacy - April 2, 2014](#)  
Apr 2, 2014... The primary purpose of the **Information Literacy** Seattle cooperation in the development of **information literacy**...

[Information Literacy - Skills & Tools of Knowledge - University of...](#)  
**Information Literacy** shares a fundamental set of core thinking and p...

Rysunek 3. Prezentacja wyników wyszukiwania w wyszukiwarce Carrot Search: okrąg  
Źródło: Carrot Search [dok. elektr.], <http://search.carrotsearch.com/carrot2-webapp/search> [odczyt: 6.09.2014]



Top 26 results of about 5,150,000 for information literacy

[Information Literacy Competency Standards for Higher Education](#)  
NOTE: While the **Information Literacy Competency Standards** for higher education during 2013-14, an ACRL task force is extensively...

[Introduction to Information Literacy - American Library Association](#)  
**Information Literacy** is the set of skills needed to find, evaluate, analyze, and begining of the 21st century has been called the **Information...**

[Information literacy - Wikipedia, The Free Encyclopedia](#)  
The United States National Forum on **Information Literacy** defines one able to identify, locate, evaluate, and effectively use that...

[Information Literacy - Home - University of Guelph](#)  
**Information Literacy** is the ability to identify what **information** is, how **information** is organized, identify the... **Information Literacy** P...

[Project Information Literacy](#)  
Project **Information Literacy** (PIL) is a large-scale, national study about habits, conducted in partnership with the University of...

[Information Literacy - April 2, 2014](#)  
Apr 2, 2014... The primary purpose of the **Information Literacy** Seattle cooperation in the development of **information literacy**...

[Information Literacy - Skills & Tools of Knowledge - University of...](#)  
**Information Literacy** shares a fundamental set of core thinking and p...

Rysunek 4. Prezentacja wyników wyszukiwania w wyszukiwarce Carrot Search: Foam Tree  
Źródło: Carrot Search [dok. elektr.], <http://search.carrotsearch.com/carrot2-webapp/search> [odczyt: 6.09.2014]

ko w formie testowej lub wymagającej uiszczenia opłaty. Pajic przytacza wyniki badań, z których wynika, że zastosowanie wizualizacji w wyszukiwarkach internetowych zwiększa skuteczność i precyzję wyszukiwania informacji, szczególnie gdy użytkownik stosuje ogólne, nieuszczerżowane kwerendy wyszukiwawcze [Pajic 2014, p. 146]. Pomimo że wyniki badań wskazują na pozytywny wpływ wizualizacji na proces wyszukiwania informacji, wizualizacyjne elementy pomocy dla użytkownika są usuwane z opcji wyszukiwarek. Przykładem może być opcja Wonder Wheel, udostępniona przez Google w 2009 r., która umożliwiała wizualizację powiązań między słowami kluczowymi z dokumentów a kwerendami wyszukiwania. Narzędzie to zostało usunięte w 2011 r. [Price 2011, dok. elektr.]. Rozbieżności pomiędzy wynikami badań, wskazującymi na wysoką skuteczność tego typu narzędzi, a ich rzeczywistym udostępnianiem są zastanawiające. Konieczne byłoby przeprowadzenie badań pozwalających ustalić, z jakich powodów wprowadzanie narzędzi wizualizacyjnych na rynek nie kończy się sukcesem. Potencjalne przyczyny mogą dotyczyć np. strat finansowych, konieczności stosowania skomplikowanych rozwiązań technologicznych lub braku zainteresowania ze strony użytkowników.

## Wizualizacja w nauce

Drugim obszarem badań nad wizualizacją w dziedzinie informatologii jest wizualizacja w nauce. Obszar ten obejmuje takie zagadnienia, jak wizualizacja wyników badań naukowych, uwzględniająca tworzenie wartości naddanej i wydobycie nowej wiedzy ze zbiorów danych, a także tworzenie nowych źródeł informacji. Informatolodzy zajmują się także badaniami bibliometrycznymi i mapowaniem nauk.

Stuart Card, Jock Mackinlay i Ben Shneiderman stwierdzili, że wizualizacja wspomaga użytkownika m.in. poprzez rozwinięcie możliwości pamięci i przetwarzania informacji, skrócenie czasu wyszukiwania informacji, poprawienie rozpoznawania wzorców w wizualizowanych danych i usprawnienie wnioskowania [Card, Mackinlay, Shneiderman 1999]. Wizualizacja ma także służyć lepszemu zrozumieniu informacji oraz wydobyciu faktów i zależności, które nie byłyby bez niej widoczne; powinna wspomagać proces tworzenia nowej wiedzy. Jednym z zadań wizualizacji jest również ukazanie przekroju danego zjawiska czy przestrzeni informacyjnej, czyli przedstawienie ogólnego spojrzenia na całość prezentowanych danych. W anglojęzycznej literaturze przedmiotu takie zadanie wizualizacji nazywane jest *overview*. Nie odnaleziono dobrego polskiego odpowiednika terminu *overview*, dlatego w artykule postanowiono posługiwać

się nazwą angielską. Termin ten stosowany jest zarówno na określenie ogólnego celu wizualizacji, jak i konkretnej cechy interfejsu użytkownika *overview-plus-detail*, pozwalającej na jednoczesne wyświetlanie zarówno całości, jak i szczegółów określonej przestrzeni informacyjnej. Na podstawie analizy literatury przedmiotu Kasper Hornbaek i Morten Hertzum zaproponowali model *overview* w wizualizacji informacji [Hornbaek, Hertzum 2011, p. 511].

W modelu Hornbaeka i Hertzuma cele wizualizacji, takie jak odkrywanie, rozumienie i monitorowanie, osiągnane są poprzez *overview*. Mają one szczególne znaczenie w badaniach naukowych. Jean-Daniel Fekete, Jarke J. van Wijk, John T. Stasko i in. podjęli próbę określenia, jaka jest rzeczywista wartość wizualizacji, szczególnie dla badaczy i naukowców. Według Fekete i in. obiektywne opisanie wartości wizualizacji jest problematyczne, na podstawie dobrze zaprojektowanych przykładów autorzy stwierdzili jednak, że wizualizacja stanowi skuteczne narzędzie analityczne, jej rzeczywistą wartość można natomiast poznać jedynie w praktyce, podczas tworzenia i wykorzystywania w konkretnych projektach [Fekete et al. 2008]. Oprócz niekwestionowanych zalet w literaturze przedmiotu często przedstawia się trudności związane z całym procesem wizualizacji. Greg McInerney, Min Chen, Robin Freeman i in. zauważają, że wielu naukowców nie potrafi zaprojektować dobrych wizualizacji. Wskazują także na brak profesjonalnych szkoleń w tym zakresie. Nieumiejętne tworzenie wizualizacji może prowadzić do pominięcia istotnych odkryć naukowych, nieporozumień, a także niepokojącej tendencji do podejmowania tylko tych badań, których wyniki da się bez trudności zwizualizować [McInerney et al. 2014, p. 148]. McInerney i współpracownicy zwrócili także uwagę na relacje pomiędzy atrakcyjnym projektem graficznym wizualizacji a poziomem zaufania w stosunku do prezentowanych danych. Zdarza się, że niska jakość projektu graficznego utożsamiana jest z niską jakością wyników badań. Autorzy wspominają także o trudnościach w interpretacji wizualizacji przez nieprofesjonalistów [McInerney et al. 2014]. Wizualizacje stanowią reprezentację danych obrazujących rzeczywistość, dlatego też pomijanie w nich danych niepewnych, powodujących niejednoznaczność interpretacji, przyczynia się do stwarzania pozornej precyzji wizualizowanych danych lub modeli.

Pojęcie „niepewności” w wizualizacji może być rozumiane w dwojaki sposób: jako wizualizacja niepewności, czyli wizualizacja danych różnej jakości, oraz niepewność wizualizacji, rozumiana jako poziom niedokładności, który pojawia się podczas przetwarzania danych. Ukazywaniem niepewności w wizualizacji zajęli się także m.in. Meredith Skeels, Bongshin Lee i in. Na podstawie analizy istniejących wcześniej klasyfikacji niepewności w wizualizacji zidentyfikowali pięć głównych typów niepewności [Skeels et al. 2009]. Zaliczyli do nich: szacowanie (ang. *approximation*), prognozowanie (ang. *prediction*), nieporozumienie

lub niespójność (ang. *disagreement or inconsistency*), kompletność (ang. *completeness*) oraz wiarygodność (ang. *credibility*) danych [Skeels et al. 2009, p. 3-4]. W swojej klasyfikacji Skeels i in. połączyli oba sposoby rozumienia niepewności w wizualizacji. Wszystkie typy niepewności w wizualizacji mogą wpłynąć na prawdziwość wniosków wyciąganych na podstawie wizualizowanych surowych danych, będących wynikiem badań naukowych. Aby wizualizacje rzeczywiście stanowiły skuteczne i pomocne narzędzie przy prowadzeniu badań naukowych, konieczne jest opracowanie zasad właściwego projektowania wizualizacji. Christa Kelleher i Thorsten Wagener zaproponowali dziesięć zasad tworzenia wizualizacji na potrzeby publikacji naukowych [Kelleher, Wagener 2011]. Znalazły się wśród nich m.in. staranne dobieranie kolorów do typów wizualizowanych danych oraz tworzenie najprostszych prezentacji graficznych, jakie są możliwe dla danego typu danych. Chaomei Chen postuluje natomiast konieczność wyboru wskaźników jakości wizualizacji, które są niezbędne do oceny i selekcji wizualizacji [Chen 2005, s. 14]. Wskazówki dotyczące projektowania takich wizualizacji, które zwiększałyby możliwości odkrywania nowej wiedzy, można znaleźć także w publikacjach Roberta Pattersona i in. [Patterson et al. 2014] oraz Tanji Keller i in. [Keller et al. 2006]. Keller i in. skupili się na znaczeniu wielowymiarowości oraz kolorystyki w wizualizacji, natomiast Patterson i in. sformułowali wskazówki dla projektowania, które mogą korzystnie wpłynąć na procesy poznawcze odbiorcy.

## **Wizualizacja w badaniach informetrycznych i mapowaniu nauk**

Zmiany w systemie komunikacji naukowej, obejmujące przeniesienie zasobów piśmiennictwa naukowego do środowiska elektronicznego i sieciowego, spowodowały wzrost zainteresowania informetrią. W artykule postanowiono traktować termin informetria jako termin najszerzy, za definicją Leo Egghe: „informetria jest terminem najszerzym, obejmującym wszystkie badania metryczne związane z nauką o informacji, włączając bibliometrię (...), naukometrię (...) i webometrię (...)” [Egghe 2005, p. 1311]. Odnalezione piśmiennictwo dotyczące wizualizacji w informetrii można podzielić na dwie główne grupy:

- publikacje przeglądowe, opisujące m.in. historię informetrii,
- artykuły badawcze, zawierające wyniki badań informetrycznych.

Bazy danych piśmiennictwa naukowego już w latach 50. XX wieku zaczęły wprowadzać opcję indeksu cytowań [Borner, Chen, Boyack 2003]. Koncepcja analizy cytowań stała się podstawą dla wszelkich badań informetrycznych. Jedną ze stosowanych w nich metod jest wizualizacja dyscyplin wiedzy lub nauki

(ang. *Knowledge Domain Visualization*), nazywana także mapowaniem nauk (ang. *Mapping Science*) lub naukoografią (ang. *scientography*). Termin *Knowledge Domain Visualization* został wprowadzony przez Chaomei Chen w 2001 r., natomiast termin *scientography* przez Eugene'a Garfielda w 1960 r. [Osińska 2010a]. Szczegółowy opis historii wizualizacji dyscyplin nauki można znaleźć w publikacji Katy Borner, Chaomei Chen i Kevina Boyack [Borner, Chen, Boyack 2003]. Wizualizacja dyscyplin nauki służy analizie struktury i zmian zachodzących w poszczególnych działach nauki. Badania informetryczne pozwalają na odkrycie m.in. luk w wiedzy lub tematów nadmiernie eksploatowanych. Rozwój technik wizualizacji dużych zbiorów danych przyczynił się do udoskonalenia i usprawnienia badań informetrycznych. Przykładem dobrze obrazującym ilość koniecznych do przetworzenia danych w badaniach informetrycznych jest praca Kevina Boyack, Richarda Klavansa i Katy Borner, którzy opracowali mapę struktury wszystkich nauk. Do badań wykorzystali 7121 czasopism naukowych, zawierających ponad milion artykułów [Boyack, Klavans, Borner 2005]. Do przetworzenia i analizy tak dużych zbiorów danych informetrycznych, np. zbiorów danych bibliograficznych, wykorzystywane są specjalistyczne narzędzia wizualizacyjne. Osińska dzieli narzędzia stosowane w badaniach informetrycznych na narzędzia do obróbki danych i narzędzia do mapowania i wizualizacji [Osińska 2010a]. Wśród narzędzi do obróbki danych wymieniono Bibexcel i CiteSpace. CiteSpace został umieszczony także wśród narzędzi do mapowania i wizualizacji. Dodatkowo, jako narzędzia do mapowania i wizualizacji wymieniono m.in. program Pajek, UCINET, Permap oraz programy do analizy statystycznej. Opisy działania poszczególnych narzędzi można znaleźć np. w artykułach dotyczących bibliometrii, publikowanych w czasopismach branżowych [Kowalska, Radzicka 2012].

Borner, Chen i Boyack opisali podstawowe typy map w wizualizacji domen naukowych [Borner, Chen, Boyack 2003]. Wyróżnione zostały mapy dokumentów (ang. *documents maps*), mapy naukowców (ang. *author-based maps*) i mapy semantyczne (ang. *semantic maps*). Podstawową jednostką wykorzystywaną w wizualizacji domen naukowych są dokumenty (np. artykuły naukowe lub patenty). Mapy dokumentów wykorzystywane są m.in. do wyszukiwania dokumentów, analizy domen, oceny dorobku naukowego i zarządzania nauką. Mapy naukowców przyjmują dwie główne formy: mapy współcytowania (ang. *co-citation maps*) i współautorstwa (ang. *co-authorship maps*). Mapy współcytowania pokazują strukturę danej dyscypliny naukowej, natomiast mapy współautorstwa odkrywają sieci społeczne w danej dyscyplinie lub społeczności badaczy. Mapy semantyczne, nazywane także analizami współwystępowania słów (ang. *co-word analyses*), pomagają w zrozumieniu struktury kognitywnej pola badawczego.

Badania wizualizacji domen naukowych podejmowane są zarówno w naukach ścisłych i medycznych, jak i społecznych i humanistycznych. Osińska stwierdza, że ze względu na swoją interdyscyplinarność informatologia i bibliologia należą do dyscyplin najczęściej wybieranych do wizualizacji [Osińska 2010a, s. 47]. Postanowiono przytoczyć kilka przykładów interesujących badań domen naukowych. Abebe Rorissa i Xiaojun Yuan przeprowadzili mapowanie struktury domeny wyszukiwania informacji (ang. *information retrieval*). Badanie objęło literaturę przedmiotu z lat 2000-2009. Do odnalezienia relewantnych dokumentów wykorzystano Web of Science, a badanie pozwoliło m.in. na odkrycie najczęściej cytowanych czasopism i artykułów naukowych, aktywnych instytucji badawczych oraz sieci współpracujących ze sobą naukowców [Rorissa, Yuan 2012]. Mehri Sedighi i Ammar Jalalimanesh przeprowadzili analizę naukometryczną obszaru zarządzania wiedzą. Wizualizacja słów kluczowych z publikacji z tego zakresu pozwoliła stwierdzić, że przedmiot badań w zarządzaniu wiedzą znacznie się zmienił i rozwinął w latach 2004-2010 [Sedighi, Jalalimanesh 2014]. W literaturze przedmiotu można wyróżnić osobną podgrupę publikacji dotyczących analizy informacji patentowej i badań nad patentami. Informacja patentowa jest popularnym przedmiotem badań informetrycznych. Liu Gui-Feng, Sun Hua-Ping i Song Xin-Ping zwizualizowali strukturę badań nad patentami w naukach o informacji i zarządzaniu. Korzystając z narzędzia CiteSpace II, opracowali mapy m.in. sieci współcytowania autorów i czasopism, powiązań między słowami kluczowymi oraz najbardziej aktywnych ośrodków badań nad patentami [Gui-Feng, Hua-Ping, Xin-Ping 2014]. Wśród przeglądowych prac dotyczących wizualizacji w informacji patentowej można wymienić np. artykuł YunYun Yang i in. [Yang et al. 2008].

Obszerne analizy literatury przedmiotu dotyczącej badań informetrycznych i mapowania nauk można znaleźć m.in. w publikacjach Judit Bar-Ilan [Bar-Ilan 2008] oraz Katy Borner, Chaomei Chen i Kevina Boyacka [Borner, Chen, Boyack 2003].

## Podsumowanie

Zmiany zachodzące w sposobie postrzegania informacji miały wpływ na włączenie wizualizacji w proces zarządzania informacją. Zagadnienie wizualizacji, tak jak pokazano na podstawie analizy literatury przedmiotu, jest uwzględniane w zainteresowaniach badawczych informatologii i bibliologii. Badania wizualizacji w kontekście nauki o informacji i bibliologii podzielono na dwie główne grupy: badania nad wsparciem wizualizacyjnym dla użytkownika systemów informacyjno-wyszukiwawczych oraz badania dotyczące wizualizacji w nauce.



W literaturze przedmiotu dostępnych jest wiele definicji „wizualizacji danych i informacji” oraz „wizualizacji naukowej”; uznano, że można stosować termin „wizualizacja informacji” do określenia ogółu zagadnień związanych z wizualizacją. W pracy pominięto rozważania nad definicją wizualizacji wiedzy, skupiono się głównie na wizualizacji danych i informacji. Część badaczy postuluje, by wyodrębnić wizualizację wiedzy z wizualizacji informacji [m.in. Burkhard 2004], jednak rozstrzygnięcie tej kwestii zależy od sposobu, w jaki definiowane są pojęcia „dane”, „informacje” i „wiedza”. Analiza konkretnych przykładów wizualizacji, które nazywane są wizualizacjami danych, informacji lub wiedzy, mogłaby pomóc w odnalezieniu rzeczywistych różnic pomiędzy nimi. W systemach informacyjno-wyszukiwawczych stosowane są obecnie różnorodne formy wsparcia dla użytkownika, szczególnie widoczne w wyszukiwarkach internetowych. Osobną kategorię badań nad wizualizacją w interfejsach użytkownika stanowią badania nad wizualizacją w bibliotekach cyfrowych. Zagadnienie to nie zostało uwzględnione w artykule, jednak planuje się sporządzenie odrębnego opracowania dotyczącego zastosowania wizualizacji w repozytoriach cyfrowych. Analiza literatury przedmiotu pokazała, że w obrębie nauk informatycznych publikowane są prace poruszające tematykę projektowania wizualizacji zgodnie z zasadami użyteczności i dostępności, jednak w informatologii takie prace są nieliczne. Ze względu na własną specyfikę wizualizacji, które nie mogą być tożsame z interfejsem użytkownika, konieczne jest podjęcie badań nad użytecznością wizualizacji, ze szczególnym naciskiem na projektowanie pod kątem uniwersalnej użyteczności. Badania wskazują, że wykorzystanie wizualizacji w systemach informacyjno-wyszukiwawczych może zwiększać efektywność procesu wyszukiwania informacji, jednak wyniki badań nie przekładają się na rzeczywiste działanie takich narzędzi. Proponuje się przeprowadzenie badań nad faktycznym zainteresowaniem użytkowników narzędziami wizualizacyjnymi, wspomagającymi wyszukiwanie informacji. W badaniach naukowych wizualizacja wykorzystywana jest zarówno jako narzędzie do opracowywania wyników badań, jak i bezpośrednia metoda badawcza. W procesie mapowania domen naukowych wizualizacja stanowi nieodłączny element badania. W artykule skupiono się tylko na mapowaniu domen naukowych w zakresie informatologii i bibliologii, nie analizowano publikacji zawierających opracowania struktur nauk ścisłych. Ukierunkowanie tematyczne baz danych, które indeksują tylko niewielką część zasobów z nauk humanistycznych, sprzyja jednak wizualizacji domen nauk ścisłych [Iribarren-Maestro, Lascrain-Sanchez, Sanz-Casado 2009].

Wizualizacja informacji może być atrakcyjnym problemem badawczym dla informatologów i bibliologów zainteresowanych badaniami o charakterze interdyscyplinarnym.

## BIBLIOGRAFIA

- Ahn, Jae-wook; Brusilovsky, Peter (2013). *Adaptive visualization for exploratory information retrieval*. Information Processing and Management Vol. 49, p. 1139-1164.
- Alhenshiri, Anwar; Blustein, James (2011). *Exploring visualization in web information retrieval*. International Journal of Internet Technology and Secured Transactions Vol. 3, No. 3, p. 320-330.
- Bar-Ilan, Judit (2008). *Informetrics at the beginning of the 21<sup>st</sup> century – a review*. Journal of Informetrics Vol. 2, p. 1-52.
- Borner, Katy; Chen, Chaomei; Boyack, Kevin (2003). *Visualizing Knowledge Domains*. Annual Review of Information Science & Technology Vol. 37, p. 179-255.
- Boyack, Kevin; Klavans, Richard; Borner, Katy (2005). *Mapping the backbone of science*. Scientometrics Vol. 64, No. 3, p. 351-374.
- Burkhard, Remo (2004). *Learning from Architects: The Difference between Knowledge Visualization and Information Visualization*. W: *Proceedings of the Information Visualization, Eighth International Conference*, p. 519-524.
- Card, Stuart; Mackinlay, Jock; Shneiderman, Ben (1999). *Readings in Information Visualization: Using Vision to Think*. San Francisco: Morgan-Kaufmann, 686 p.
- Chen, Chaomei (2005). *Top 10 Unsolved Information Visualization Problems*. IEEE Computer Graphics and Applications July/August 2005, p. 12-16.
- Chen, Chaomei (2006). *Information Visualization: Beyond the Horizon*. 2<sup>nd</sup> edition. Miejsce wydania: Springer Science & Business Media, 332 p.
- Chen, Min et al. (2009). *Data, Information, and Knowledge in Visualization*. Computer Graphics and Applications Vol. 29, Issue 1, p. 12-19.
- Chi, Ed. H. (2000). *A Taxonomy of Visualization Techniques using the Data State Model* [dok. elektr.]. <http://www-users.cs.umn.edu/~echi/papers/infovis00/Chi-Taxonomy-Visualization.pdf> [odczyt: 5.08.2014].
- Chowdhury, Gobinda; Chowdhury, Sudatta (2011). *Information users and usability in the digital age*. London: Facet, 208 p.
- Chu, Shixian et al. (2007). *A Treemap-based Result Interface for Search Engine Users*. In: Smith, M.J.; Salvendy, G. eds. *Human Interface*. Berlin: Springer-Verlag, p. 401-410.
- Chung, Wingyan et al. (2005). *Evaluating event visualization: a usability study of CO-PLINK spatio-temporal visualizer*. International Journal of Human-Computer Studies Vol. 62, p. 127-157.
- Dürsteler, Juan (2002). *KartOO* [dok. elektr.]. <http://www.infovis.net/printMag.php?num=97&lang=2> [odczyt: 21.09.2014].
- Dylak, Stanisław (2012). *Alfabetyzacja wizualna jako kompetencja współczesnego człowieka*. W: Skrzydlewski, Wojciech; Dylak, Stanisław red. *Media – Edukacja – Kultura*. Poznań; Rzeszów: Polskie Towarzystwo Technologii i Mediów Edukacyjnych, s. 119-132.
- Egghe, Leo (2005). *Expansion of the field of informetrics: Origins and consequences*. Information Processing and Management Vol. 41, No. 6, p. 1311-1316.
- Fekete, Jean-Daniel et al. (2008). *The Value of Information Visualization*. In: Kerren, Andreas et al. eds. *Information Visualization: Human-Centered Issues and Perspectives*. Berlin: Springer-Verlag, p. 1-18.

- Fluit, Christiaan; Sabou, Marta; Harmelen, Frank van (2006). *Ontology-based Information Visualization: Toward Semantic Web Applications*. In: Geroimenko, Vladimir; Chen, Chaomei eds. *Visualizing the Semantic Web*. London: Springer, p. 45-58.
- Friendly, Michael (2009). *Milestones in the history of thematic cartography, statistical graphics, and data visualization* [dok. elektr.]. <http://www.math.yorku.ca/SCS/Gallery/milestone/milestone.pdf> [odczyt: 5.08.2014].
- Garlock, Kristen; Piontek, Sherry (1999). *Designing Web interfaces to library services and resources*. Chicago: American Library Association, 128 p.
- Gui-Feng, Liu; Hua-Ping, Sun; Xin-Ping, Song (2014). *Visualizing and mapping the research on patents in information science and management science*. *Malaysian Journal of Library & Information Science* Vol. 19, No. 1, p. 87-103.
- Hoerber, Orland; Yang, Xue Dong (2008). *HotMap: Supporting visual exploration of Web search results*. *Journal of the American Society for Information Science and Technology* Vol. 60, Issue 1, p. 90-110.
- Hoerber, Orland; Yang, Xue Dong (2010). *Supporting Web Search with Visualization*. In: Yao, J.T. ed. *Web-Based Support Systems: Advanced Information and Knowledge Processing*. London: Springer-Verlag, p. 183-214.
- Hornbaek, Kasper; Hertzum, Morten (2011). *The notion of overview in information visualization*. *International Journal of Human-Computer Studies* Vol. 69, p. 509-525.
- Iribarren-Maestro, Isabel; Lascurain-Sanchez, Maria Luisa; Sanz-Casado, Elias (2009). *The use of bibliometric techniques in evaluating social sciences and humanities. Celebrating scholarly communication studies: a festschrift for Olle Persson* Vol. 05-S, p. 25-38.
- Jones, Catherine Emma et al. (2009). *A less-is-more approach to geovisualization – enhancing knowledge construction across multidisciplinary teams*. *International Journal of Geographical Information Science* Vol. 23, No. 8, p. 1077-1093.
- Kelleher, Christa; Wagener, Thorsten (2011). *Ten guidelines for effective data visualization in scientific publications*. *Environmental Modelling & Software* Vol. 26, p. 822-827.
- Keller, Tanja et al. (2006). *Information visualizations for knowledge acquisition. The impact of dimensionality and color coding*. *Computer in Human Behavior* Vol. 22, p. 43-65.
- Koshman, Sherry (2004). *Comparing usability between a visualization and text-based system for information retrieval*. *Journal of Documentation* Vol. 60, Issue 5, p. 565-580.
- Koshman, Sherry (2006). *Visualization-based information retrieval on the Web*. *Library & Information Science Research* Vol. 28, p. 192-207.
- Koua, Etien; Maceachren, Alan; Kraak, Menno-Jan (2006). *Evaluating the usability of visualization methods in an exploratory geovisualization environment*. *International Journal of Geographical Information Science* Vol. 20, No. 4, p. 425-448.
- Kowalska, Aneta; Radzicka, Joanna (2012). *Programy BibExcel i Pajek w analizach bibliometrycznych* [dok. elektr.]. *Biuletyn EBIB nr 3 (120)*. [http://www.ebib.pl/images/stories/numery/130/130\\_kowalska\\_radzicka.pdf](http://www.ebib.pl/images/stories/numery/130/130_kowalska_radzicka.pdf) [odczyt: 9.09.2014].
- Kubíček, Petr; Šašinka, Čeněk (2011). *Thematic uncertainty visualization usability – comparison of basic methods*. *Annals of GIS* Vol. 17, Issue 4, p. 253-263.
- Kules, Bill; Wilson, Max; Shneiderman, Ben (2008). *From Keyword Search to Exploration: How Results Visualization Aids Discovery on the Web* [dok. elektr.]. <http://eprints.soton.ac.uk/265169/1/VSRWeb-TR.pdf> [odczyt: 6.09.2014].

- Lauesen, Soren (2005). *User interface design*. Harlow: Addison-Wesley, 640 p.
- Lynch, Patrick; Horton, Sarah (2008). *Web Style Guide: Basic principles for creating Web sites*, 3<sup>rd</sup> edition. New Haven: Yale University Press, 338 p.
- McInerny, Greg et al. (2014). *Information visualisation for science and policy: engaging users and avoiding bias*. Trends in Ecology and Evolution Vol. 29, No. 3, p. 148-157.
- Morse, Emile; Lewis, Michael; Olsen, Kai (2000). *Evaluating visualizations: using a taxonomy guide*. International Journal of Human-Computer Studies Vol. 53, p. 637-662.
- Newby, Gregory (2002). *Empirical Study of a 3D Visualization for Information Retrieval Task*. Journal of Intelligent Information Systems Vol. 18, p. 31-53.
- Nielsen, Jakob (2000). *Designing web usability*. Indianapolis: New Rider Publishing, 420 p.
- Nielsen, Jakob (2012). *Usability 101: Introduction to usability* [dok. elektr.]. <http://www.useit.com/alertbox/20030825.html> [odczyt: 20.10.2014].
- Nivala, Annu-Maaria; Brewster, Stephen; Sarjakoski, L. Tiina (2008). *Usability Evaluation of Web Mapping Sites*. The Cartographic Journal Vol. 45, No. 2, p. 129-138.
- Osińska, Veslava (2010a). *Rozwój metod mapowania domen naukowych i potencjał analityczny w nim zawarty*. Zagadnienia Informatyki nr 2(96), s. 41-51.
- Osińska, Veslava (2010b). *Wizualizacja i wyszukiwanie dokumentów*. Warszawa: Stowarzyszenie Bibliotekarzy Polskich, 183, [1] s.
- Osińska, Veslava; Bala, Piotr; Gawarkiewicz, Michał (2012). *Information Retrieval cross Information Visualization*. In: Ganzha, M.; Maciaszek, L.; Paprzycki, M. eds. *Proceedings of the Federated Conference on Computer Science and Information Systems 2012*. Warsaw: Polskie Towarzystwo Informatyczne, p. 233-239.
- Pajić, Dejan (2014). *Browse to search, visualize to explore: Who needs an alternative information retrieving model?* Computers in Human Behavior Vol. 39, p. 145-153.
- Pamuła-Cieślak, Natalia (2014). *Wizualizacja i prezentacja informacji w wyszukiwarkach internetowych*. W: Głowacka, Ewa; Kowalska, Małgorzata; Krysiński, Przemysław red. *Współczesne oblicza komunikacji i informacji: problemy, badania, hipotezy*. Toruń: Wydaw. Naukowe Uniwersytetu Mikołaja Kopernika, s. 257-268.
- Patterson, Robert et al. (2014). *A human cognition framework for information visualization*. Computers & Graphics Vol. 42, p. 42-58.
- Plaisant, Catherine (2005). *Information Visualization and the Challenge of Universal Usability* [dok. elektr.]. <http://hcil2.cs.umd.edu/trs/2004-36/2004-36.html> [odczyt: 6.08.2014].
- Price, Gary (2011). *Official: The Google Wonder Wheel is Gone* [dok. elektr.]. <http://searchengineland.com/official-the-google-wonder-wheel-is-gone-84105> [odczyt: 6.09.2014].
- Risden, Kirsten et al. (2000). *An initial examination of easy of use for 2D and 3D information visualizations of web content*. International Journal of Human-Computer Studies Vol. 53, p. 695-714.
- Rivadeneira, Walky; Bederson, Benjamin (2003). *A study of search results clustering interfaces: comparing textual and zoomable user interfaces* [dok. elektr.]. <ftp://ftp.cs.umd.edu/pub/hcil/Reports-Abstracts-Bibliography/2003-36html/2003-36.htm> [odczyt: 6.09.2014].
- Rorissa, Abebe; Yuan, Xiaojun (2012). *Visualizing and map ping the intellectual structure of information retrieval*. Information Processing and Management Vol. 48, p. 120-135.

- Rozporządzenie Rady Ministrów z dnia 12 kwietnia 2012 r. w sprawie Krajowych Ram Interoperacyjności, minimalnych wymagań dla rejestrów publicznych i wymiany informacji w postaci elektronicznej oraz minimalnych wymagań dla systemów teleinformatycznych. Dz.U. 2012 nr 0, poz. 526.
- Sedighi, Mehri; Jalalimanesh, Ammar (2014). *Mapping research trends in the field of knowledge management*. Malaysian Journal of Library & Information Science Vol. 19, No. 1, p. 71-85.
- Shiri, Ali (2008). *Metadata-enhanced visual interfaces to digital library*. Journal of Information Science Vol. 34, Issue 6, p. 763-775.
- Shneiderman, Ben (1992). *Designing the user interface: Strategies for effective human-computer interaction*. 2<sup>nd</sup> ed. Reading, MA: Addison-Wesley, 606 p.
- Shneiderman, Ben (1996). *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations* [dok. elektr.]. <http://www.cs.ubc.ca/~tmm/courses/old533/readings/shneiderman96eyes.pdf> [odczyt: 5.08.2014].
- Skeels, Meredith et al. (2009). *Revealing uncertainty for information visualization*. [dok. elektr.]. Information Visualization <http://research.microsoft.com/en-us/um/redmond/groups/cue/publications/ivs2009-uncertainty.pdf> [odczyt: 20.07.2014].
- Stone, Debbie et al. (2005). *User interface design and evaluation*. London: Morgan Kaufmann, 669 p.
- Sutcliffe, Alistair; Ennis, Mark; Hu, J (2000). *Evaluating the effectiveness of visual user interfaces for information retrieval*. International Journal of Human-Computer Studies Vol. 53, p. 741-763.
- Swienty, Olivier et al. (2008). *The Role of Relevance and Cognition in Attention-guiding Geovisualization*. Cartographic Journal Vol. 45, Issue 3, p. 227-238.
- Tory, Melanie; Möller, Torsten (2004). *Rethinking Visualization: A High-Level Taxonomy* [dok. elektr.]. <http://www.cs.sfu.ca/~torsten/Publications/Papers/infovis04.pdf> [odczyt: 5.08.2014].
- Vegas, Jesus; Fuente, Pablo de la; Crestani, Fabio (2003). *WebDocBall: A Graphical Visualization Tool for Web Search Results*. In: Seabstiani, F. *ECIR 2003*. Berlin: Springer-Verlag, p. 351-362.
- VisGets [dok. elektr.]. <http://mariandoerk.de/visgets/demo/> [odczyt: 6.09.2014].
- Ware, Colin (2004). *Information Visualization: perception for design*. 2<sup>nd</sup> edition. Miasto: Morgan Kaufmann, 486 p.
- WCAG 2.0 *Web Content Accessibility Guidelines*: autoryzowane tłumaczenie na język polski (2013) [dok. elektr.]. <http://fdc.org.pl/wcag2/> [odczyt: 8.10.2014].
- Wiza, Wojciech (2012). *Interactive 3D Visualization of Search Results*. In: Cellary, Wojciech; Walczak, Krzysztof eds. *Interactive 3D Multimedia Content*. London: Springer-Verlag, p. 253-291.
- Wolfram Alpha [dok. elektr.]. <http://www.wolframalpha.com/> [odczyt: 6.09.2014].
- Yang, YunYun et al. (2008). *Text mining and visualization tools – Impressions of emerging capabilities*. World Patent Information Vol. 30, p. 280-293.

---

# Benchmarking czasopism naukowych jako element strategii publikacyjnej naukowca. Przykład medycyny weterynaryjnej

Agnieszka Korycińska-Huras

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

Publikowanie w czasopismach naukowych jest podstawową formą uczestnictwa w formalnej komunikacji naukowej. W obszarze nauk doświadczalnych: chemii, biologii, fizyki, medycyny i nauk technicznych (ang. *Science, Technology & Medicine – STM*) zawartość periodyków naukowych stanowi ponad 80% całego dostępnego piśmiennictwa. Bezpośrednio warunkuje to możliwość udziału publikacji w globalnym obiegu naukowym, jako że wielkie bazy, będące głównymi węzłami na mapie powiązań komunikacyjnych w dziedzinie STM, np. Web of Science i Scopus, indeksują głównie zawartość czasopism. Ten, kto nie publikuje w czasopismach, staje się zatem *de facto* niewidzialny [Haustein 2012].

Na gruncie informatologii od co najmniej dwóch dekad prowadzone są badania zmierzające do ustalenia możliwie uniwersalnego zestawu parametrów oceny, który skutecznie wspomagałby publikujących badaczy w wyborze czasopism naukowych, do których decydują się kierować swoje teksty. Badania te realizowane są z perspektywy użytkownika (naukowca nieposiadającego wiedzy eksperckiej w zakresie naukometrii), za pomocą metod ilościowych i mieszanych, z założeniem cyklicznych powtórzeń i stopniowego ich rozciągania na kolejne dyscypliny wiedzy. Aktualny przegląd piśmiennictwa podsumowujący dorobek naukowy w tym zakresie można znaleźć w artykule Petera Jacso z 2013 r. [Jacso 2013].

W warunkach polskich opracowanie takiej siatki kryteriów, stanowiącej praktyczny przewodnik dla naukowców, jest szczególnie pożądane ze względu na niedawne reformy akademickiego trybu oceny pracy naukowej. Wdrażane w całym kraju systemy parametryzacji dorobku publikacyjnego jednoznacznie skłaniają

do preferowania wydawnictw zagranicznych, zwłaszcza czasopism o wysokim współczynniku *impact factor*. Mało opłacalne, z punktu widzenia wymiaru oceny pracowniczej i skorelowanych z nią szans awansu naukowego, stały się natomiast publikacje o wąskim zasięgu krajowym, m.in. referaty w materiałach konferencyjnych, którym przypisuje się obecnie niską wartość punktową w oficjalnych sprawozdaniach uczelnianych.

Nowe zasady sterowania aktywnością publikacyjną uczonych, wynikające z ambicji umiędzynarodowienia polskiej nauki, wymuszają zmiany w dotychczasowych praktykach uczestnictwa w transferze wiedzy naukowej. W interesie uczonych leży zatem znajomość wskaźników, które pozwolą im stworzyć wyważoną strategię publikowania, zapewniającą osiągnięcie jak największej liczby indywidualnych celów przy racjonalnym stosunku nakładów do korzyści.

Podstawowym celem badaczy jest eksternalizacja swojej wiedzy w taki sposób i za pośrednictwem takich kanałów komunikacji, aby dotarła ona w stosownym czasie do pożądaných adresatów, czyli, najczęściej, do lokalnego lub międzynarodowego środowiska badaczy reprezentujących określoną dziedzinę lub dyscyplinę nauki. Wśród naukowców niezmiennie przeważa pogląd, że – gdy chodzi o publiczność docelową – ważniejsze jest skuteczne zainteresowanie kompetentnych odbiorców, do których przekaz jest faktycznie kierowany, aniżeli przyciągnięcie czytelników liczniejszych, ale spoza branży. Deklaracje te zdają się jednak nieco na wyrost, skoro badania prowadzone od 2004 r. dowodzą, że na czoło listy kryteriów oceny wysuwają się reputacja i prestiż czasopisma, przy czym cechę tę utożsamia się, w obu przypadkach niesłusznie, już to z utrwaloną opinią środowiskową („to jest szacowny periodyk” / „to wydawnictwo ma markę”), już to z samym tylko indeksem *impact factor* [Swan, Brown 2004, dok. elektr.; Mierzejewska 2010].

Wielu informatologów przypomina wszakże, że współczynnik *impact factor* nie może być uznawany za pełne odzwierciedlenie jakości merytorycznej i środowiskowego statusu czasopisma [Rousseau 2002, dok. elektr.; Coleman 2007, p. 1148; Lozano, Larivière, Gingras 2012, dok. elektr.; Kamalski, Plume, Amin 2014, p. 390]. Jest to bowiem parametr ściśle ilościowy, odnoszący się do zaledwie jednego aspektu społecznego funkcjonowania czasopisma, tj. do liczby cytowań. Tymczasem wiarygodna i uczciwa ocena rangi periodyku musi mieć charakter wielopłaszczyznowy i obejmować wszystkie istotne atrybuty przesądzające o rezonansie publikowanych w nim treści. Stwierdzenie to wydaje się oczywiste dla każdego, kto zna choćby rudymenty metodologii naukowej, dlatego wybiórcze skoncentrowanie uczonych na wskaźniku IF może dziwić. Znaczący komunikacji naukowej widzą jednak proste wyjaśnienie tej tendencji. Otóż według Stefanie Hausteina w piśmiennictwie naukowym z zakresu informatologii wciąż jeszcze

brakuje przystępnych opracowań o charakterze poradnikowym, które pomogłyby badaczom, także tym mniej doświadczonym, odnaleźć się w gąszczu tajemniczych, często nieprzejrzyście wyłożonych metod i parametrów bibliometrycznych. Odstręczający hermetyczny metajęzyk analiz naukometrycznych sprawia, że dla bezpieczeństwa i zwykłej wygody uczeni wolą selektywnie zaufać temu, co dobrze znają, czyli wskaźnikowi *impact factor*, nie zważając na jego dowiedzione słabości i fakt, że z racji choćby zdroworozsądkowych nie może on uchodzić za jedyną i wystarczającą miarę wartości czasopisma. To, czy naukowiec szukający czasopisma adekwatnego do swoich potrzeb i zamierzeń wykorzysta w swoim wyborze również inne kryteria ocenne, zależy od tego, czy będzie je znał. Jak zauważają Craig, Ferguson i Finch, w praktyce przyjmują się tylko takie narzędzia i metody ewaluacji, które współbrzmia z potrzebami i nawykami intelektualnymi danego środowiska naukowego, a przede wszystkim spełniają zasadę facylitacji, czyli dają zadowalające efekty przy maksimum prostoty i poręczności [Craig, Ferguson, Finch 2010].

Zgodnie z przytoczoną dewizą w artykule przedstawiono wyniki analizy benchmarkingowej pięciu międzynarodowych anglojęzycznych czasopism naukowych z dziedziny weterynarii, przeprowadzonej za pomocą formularza stworzonego przez kontaminację dwóch matryc oceny, autorstwa Björka i Holmströma (2006) oraz Knight i Steinbach (2008), z empirycznie przetestowanym modelem fasetowej oceny jakości czasopism naukowych Stefanie Haustein (2012). Jako źródło danych będących podstawą porównania wykorzystano wyłącznie jawną zawartość serwisów internetowych badanych czasopism lub ich wydawców. Wyboru medycyny weterynaryjnej jako egzemplifikacji dokonano intuicyjnie, zakładając że polscy badacze z tej dziedziny rozwinęli charakterystyczne praktyki publikacyjne, jednak w roku 2014 zasadność tego wyboru została potwierdzona wynikami badań ICM UW, ogłoszonymi w publikacji „Otwarta nauka w Polsce 2014. Diagnoza”, w której stwierdzono, że „autorów publikacji z obszaru nauk rolniczych, leśnych i weterynaryjnych wyróżnia odmienne niż w naukach humanistycznych podejście do prestiżu, który jest dla nich najmniej istotny w porównaniu z pozostałymi grupami. Inaczej niż w wypadku humanistów, najbardziej liczy się dla nich punktacja, która jest najważniejszym kryterium wyboru miejsca publikacji” [Szprot red. 2014, dok. elektr., s. 85-86].

Praca ta, o charakterze przyczynkarskim, pogładowym i praktycznym, ma na celu po pierwsze, przypomnienie głównych wymienianych w literaturze informatologicznej wyznaczników rangi czasopism naukowych z punktu widzenia użytkownika (autora) oraz po drugie, ukazanie, jak ich wzajemne oddziaływania potęgują lub zmniejszają wagę *impact factor*. W Polsce ewidentnie pierwszoplanowym kryterium wyboru jest punktacja przyznana danemu periodykowi przez



Ministerstwo Nauki i Szkolnictwa Wyższego, w dużym stopniu proporcjonalna do IF, dlatego też do analizy kontrastywnej wybrano tylko czasopisma z tej listy. Do próby zakwalifikowano czasopisma o wartości oszacowanej przez MNiSW na 30 punktów w roku 2013. Nie ma wśród nich czasopism polskich, ponieważ jedyne dwa tytuły krajowe figurujące w Części A wykazu MNiSW na rok 2013, „Polish Journal of Veterinary Science” oraz „Bulletin of the Veterinary Institute in Pulawy”, zostały wycenione, odpowiednio, na 15 i 20 punktów.

Pod względem metodycznym przedstawione tu badanie spełnia kryteria benchmarkingu, rozumianego, zgodnie ze źródłową definicją przejętą z nauk o zarządzaniu, jako „proces oceny działalności przedsiębiorstw uznanych za liderów w danej branży, pozwalający na określenie najkorzystniejszych rozwiązań możliwych do zastosowania w firmie i ustalenie na ich podstawie strategicznych celów poprawy wyników działalności tej firmy” [Kaszubska 2012, dok. elektr.; Mikulski, Mikulska 2006, dok. elektr.]. W roli „firmy” występuje badacz (autor tekstów naukowych) poszukujący najbardziej rentownych kanałów publikacji z myślą o zapewnieniu sobie bezpieczeństwa zatrudnienia i przewagi konkurencyjnej w środowisku naukowym. Wymienione w definicji „przedsiębiorstwa” to anglojęzyczne czasopisma naukowe z dziedziny nauk weterynaryjnych, uznane za „liderów branży” na podstawie wysokiej (w danym przypadku maksymalnej) punktacji na liście rankingowej MNiSW na rok 2013. Termin „benchmark”, czyli cel wyrażalny za pomocą wskaźników liczbowych, odnosi się natomiast do czasopisma, które w wyniku wykonanej analizy okaże się najkorzystniejsze dla badacza w aspekcie rachunku zyski-nakłady. Należy uściślić, że w scenariuszu badawczym w pełni zgodnym z regułami sztuki w roli podmiotu typującego benchmarki powinno występować wydawnictwo, nie autor. Na potrzeby tej pracy przyjęto jednak, że autor może selekcjonować liderów w branży czasopism naukowych, ponieważ dzieli z nimi strategiczne cele, tj. poszerzenie zasięgu penetracji ogłaszanych treści (jak największa publiczność, jak najwięcej cytowań), a zasadniczo różnią go od nich tylko źródła osiągnięcia dochodów: dla wydawców czasopism jest to sprzedaż nakładów, a dla autora wynagrodzenie pobierane od pracodawcy lub wypłacane w ramach umów grantowych.

Porównanie czasopism o identycznej punktacji, a więc z perspektywy autora pozornie równowartościowych, ma tę zaletę, że bardzo czytelnie ujawnia wpływ poszczególnych cząstkowych składników oceny na rzeczywistą opłacalność publikacji. Okazuje się mianowicie, że różnice w często pomijanych aspektach, takich jak czas oczekiwania na publikację (ang. *publication delay* vel *lag*, *lapse* [Amat 2008, dok. elektr.]) bądź nadmierna heterogeniczność tematyczna, która utrudnia stosowanie jednolitych standardów recenzji, a zatem pośrednio podnosi odsetek materiałów odrzuconych, sprawiają zazwyczaj, że uzyskanie tej samej

puli punktów wymaga, w zależności od czasopisma, znacząco różnej inwestycji własnej i obarczone jest różnym stopniem niepewności. Innymi słowy, zdobyte przez autora punkty nie są zyskiem netto, lecz przychodem, od którego należy odliczyć „koszty uzyskania”. Przed podjęciem decyzji o zgłoszeniu rękopisu do danego czasopisma autor powinien zatem skalkulować obciążające ją ryzyko i szanse osiągnięcia satysfakcjonującej stopy zwrotu z inwestycji, próbując zabezpieczyć czasem niełatwe do pogodzenia interesy lub oczekiwania: własne (subiektywne, odnoszące się do percepcji osobistego prestiżu), macierzystej instytucji, organizacji finansującej badania, odbiorców docelowych, a bywa też, że i osób trzecich – swoich mistrzów, promotorów, współpracowników.

W artykule zestawiono ważniejsze „pozycje kosztowe” kształtujące końcowy bilans przedsięwzięcia publikacyjnego na przykładzie czasopism z dziedziny weterynarii. Ich znajomość może być szczególnie przydatna polskim uczonym z następujących powodów: 1) założony na szczeblu ministerialnym plan internacjonalizacji działalności naukowej implikuje zmianę obyczajów i praktyk uczestnictwa w formalnej komunikacji naukowej, wymuszając zwrot w stronę najwyżej punktowanych czasopism anglojęzycznych, 2) w następstwie zaostrzenia kryteriów oceny osiągnięć naukowych, powiązanego pośrednio z prognozowanym spadkiem liczby studentów i zapowiedziami redukcji kadry akademickiej, autorzy coraz silniej podlegają dyktatowi *publish or perish*, mogącemu skłaniać do pośpiesznych i powierzchownie uzasadnionych decyzji publikacyjnych, 3) wprowadzenie do nauki modelu biznesowego opartego na rygorystycznie rozliczanych grantach (krajowych, europejskich) prowadzi do sytuacji, w których błąd na etapie planowania strategii publikacyjnej, zwłaszcza jeżeli wymóg terminowego upubliczniania wyników badań zapisany jest w umowie grantowej, może oznaczać dla autora straty w wynagrodzeniu lub nawet konieczność zwrócenia części subwencji, 4) w polskiej kulturze naukowej silnie zakorzenione są honorowe rytuały społeczne i zachowania o charakterze kurtuazyjnym, budujące środowiskowy prestiż symboliczny wydawnictw lub pojedynczych czasopism, nie zawsze korespondujący z ich pozycją w rankingach oficjalnych; należy przypuszczać, że te lokalnie zmienne racje o rodowodzie kulturowym, zwykle nacechowane emocjonalnie, mogą utrudniać uczonym wdrożenie się do oceny czasopism w kategoriach rachunku ekonomicznego (nawet jeżeli, jak wskazują badania, badacze reprezentujący nauki ścisłe ujawniają mniejszy opór wobec zmian w modelu komunikacji naukowej aniżeli przedstawiciele humanistyki i nauk społecznych [Sapa 2007; Sapa 2008]), 5) do Polski dotrze niebawem podnosząca się właśnie w środowisku nauk biomedycznych fala niezgody na swoistą fetyszyzację wskaźnika *impact factor*, której wyrazem stała się oddolna międzynarodowa inicjatywa San Francisco Declaration on Research Assessment. Zawiera ona protest prze-

ciwko przekształcaniu IF w nadrzędne kryterium finansowania badań i oceny produktywności kadry naukowej we wszystkich dyscyplinach, połączony z apelem o poszukiwanie alternatywnych, wielopłaszczyznowych i czulszych instrumentów pomiaru wartości czasopism [Misteli 2013, dok. elektr.].

Przyjęty w analizie klucz oceny wartości czasopism obejmuje łącznie 5 kryteriów głównych i w tym aspekcie odwzorowuje model wypracowany przez Björka i Holmströma (2006). W poniższej tabeli przedstawiono wszystkie płaszczyzny oceny wskazane w oryginalnej propozycji szwedzkiej, także te, które sami autorzy określają jako parametry trudno mierzalne, a nawet niemierzalne z punktu wi-

Tabela 1. Kryteria oceny rentowności czasopisma z perspektywy badacza procesów komunikacji naukowej

Parametr	Dostępność danych
<b>Infrastruktura</b>	
Cena prenumeraty	DD lub DZ
Włączenie do międzynarodowych indeksów	DD
Przyjazność dla autorów	BE
Rozwiązania techniczne usprawniające proces wydawniczy	DD + BE
<b>Publiczność odbiorcza</b>	
Liczba prenumeratorów instytucjonalnych	DD lub DZ
Liczba prenumeratorów indywidualnych	DD lub DZ
Nakład wydania papierowego	OW
Liczba pobrań pełnego tekstu	OW + DD (np. w repozytoriach OA)
Struktura grupy docelowej (naukowcy / praktycy / przemysł)	OW + BE
Liczba abonentów kanałów alertowych	OW
Zasięg geograficzny czasopisma	OW
Cytowania (liczba cytowań na artykuł) w rozbiściu na cytowania w czasopismach naukowych i fachowych	BE
<b>Prestiż</b>	
Prestiż wydawcy	DD
Impact factor	DD
Miejsce czasopisma w rankingach dziedzinowych	DD
Renoma członków rady naukowej / kolegium recenzenckiego	BE
Subiektywne preferencje autora starającego się o publikację	DD
Regulacja współpracy między redakcją a autorami	OW
Odsetek rękopisów odrzuconych	DD lub OW
Jakość procesu recenzji	BE
Średni czas oczekiwania na publikację	DD lub DZ

Źródło: [Bjork, Holmström 2006, dok. elektr., p. 152]

Legenda: DD = dane dostępne, DZ = dane możliwe do zdobycia z pozycji autora, OW = dane od wydawcy, BE = konieczność przeprowadzenia badań empirycznych

dzenia indywidualnego badacza, z tej przyczyny, że albo wymagałyby odrębnych szeroko zakrojonych badań empirycznych, albo wyliczenie ich musiałyby opierać się na informacjach niejawnych, objętych tajemnicą biznesową lub strzeżonych przez wydawców jako potencjalnie zagrażające ich wizerunkowi. Parametry te przywołano tu tylko w celu ilustracyjnym, jako elementy pełnego, idealnego drzewa decyzyjnego, którego zastosowanie w praktyce przez indywidualnego naukowca jest w zasadzie niewykonalne.

## Dobór czasopism do analizy benchmarkingowej

W Wykazie czasopism naukowych MNiSW na rok 2013, Część A: „Czasopisma naukowe posiadające współczynnik wpływu Impact Factor (IF) znajdujące się w bazie Journal Citation Reports (JCR)” wykazano 142 wydawnictwa ciągłe z zakresu medycyny weterynaryjnej. Przyporządkowana im punktacja została przedstawiona w poniższej tabeli.

Tabela 2. Punktacja czasopism naukowych z listy Journal Citation Reports według MNiSW

Punktacja MNiSW		15	20	25	30	35	40	45	50
Liczba czasopism	N= 142	22	26	28	32	17	10	5	2
Udział procentowy		15,3%	18,0%	19,4%	22,2%	11,8%	8,3%	3,5%	1,4%

Źródło: [Komunikat... 2012, dok. elektr.]

Do porównania wybrano czasopisma z najliczniej reprezentowanej grupy, czyli dające autorom prawo do 30 punktów rozliczeniowych za każdy opublikowany artykuł [Komunikat... 2012, dok. elektr.]. Z puli 32 tytułów wyselekcjonowano pięć czasopism o najszerszym spektrum tematycznym, zaklasyfikowanych w trzeciej edycji wykazu Basic List of Veterinary Medical Serials (BLVMS) do kategorii „General veterinary medicine” [Ugaz et al. 2010, dok. elektr.]. Na potrzeby tego rekonesansu badawczego, mającego na celu jedynie przetestowanie matrycy oceny rentowności publikacji z punktu widzenia autora, uznano że właśnie takie periodyki z największym prawdopodobieństwem przyciągną uwagę polskiego naukowca szukającego łamów dla swojego tekstu. Liczba oferowanych punktów (30) jest wysoka, sytuuje się wyraźnie powyżej środka rozpiętości skali (15 – 50), a zarazem nie wkracza w rewiry najbardziej wyśrubowane. W świetle wiedzy o motywacjach decyzji inwestycyjnych pułap 30 może być określony jako bezpieczny, choć nieasekuracyjny, ambitny, obciążony znacznym ryzykiem,

a jednocześnie pozbawiony odcienia brawury, która samym zainteresowanym (inwestorom) kojarzy się zazwyczaj z wybieraniem celów skrajnie ambitnych. To jedna z domniemyanych zachęt. Drugą jest natomiast szeroki profil czasopisma, dopuszczający zasadniczo artykuły ze wszystkich specjalności medycyny weterynaryjnej. Niezależnie od tego, że audytowane periodyki przynależą do grupy ogólnomedycznych w wykazie BLVMS, informację tę zweryfikowano z zawartością ich serwisów www (zakładka *Scope & Aim*), potwierdzając w szczególności, że oficjalnie wyznaczony zakres obejmuje wszystkie subdyscypliny weterynarii w odniesieniu do wszystkich gromad kręgowców i bez względu na cechę domestykacji (zwierzęta domowe i udomowione oraz fauna dzika).

Obie te przesłanki są w pewnej mierze arbitralne i zdroworozsądkowe. Nie szkodzi to wprawdzie wynikom przeprowadzonego benchmarkingu, jako że jest on na razie ćwiczeniem demonstracyjnym, ale motywuje do zaplanowania regularnych, metodologicznie poprawnych badań pozwalających uchwycić zmienne, które rzeczywiście warunkują decyzje polskich uczonych z dziedziny weterynarii, począwszy od etapu najwcześniejszego, czyli wyboru profilu tematycznego i punktacji czasopisma. Z zastrzeżeniem umowności obu wspomnianych kryteriów inkluzji, do próby badawczej zaliczono następujące czasopisma:

1) Acta Veterinaria Scandinavica, skrót tytułu wg JCR: Acta Vet Scand  
instytucja sprawcza: Veterinary Associations of the Nordic Countries  
wydawca: BioMed Central, kraj: Wielka Brytania  
wydawane od: 1959, *online* od 2006  
częstotliwość: kwartalnik + suplement  
strona www: [www.actavetscand.com/](http://www.actavetscand.com/)

2) Veterinary Record, skrót tytułu wg JCR: Vet Rec  
instytucja sprawcza: British Veterinary Association  
wydawca: BMJ Group, kraj: Wielka Brytania  
wydawane od: 1888  
częstotliwość: tygodnik (wersja drukowana) + funkcja OnlineFirst (wersja elektroniczna)  
strona www: [veterinaryrecord.bmj.com/](http://veterinaryrecord.bmj.com/)

3) New Zealand Veterinary Journal, skrót tytułu wg JCR: NZ Vet J  
wydawca: Thomson&Reuters, kraj: Nowa Zelandia  
wydawane od: 1952  
częstotliwość: dwumiesięcznik  
strona www: [www.vetjournal.org.nz/contact](http://www.vetjournal.org.nz/contact)

4) Canadian Veterinary Journal (CVJ), skrót tytułu wg JCR: Can Vet J  
 instytucja sprawcza i wydawca: Canadian Veterinary Medical Association (CVMA)  
 kraj: Kanada  
 wydawane od: 1960  
 częstotliwość: miesięcznik  
 strona www: <https://www.canadianveterinarians.net/science-knowledge/cvj>

5) Journal of Veterinary Science, skrót tytułu wg JCR: J Vet Sci  
 instytucja sprawcza: Korean Society of Veterinary Science  
 wydawca: College of Veterinary Medicine, Seoul National University  
 wydawane od: 2000  
 kraj: Korea Południowa  
 częstotliwość: kwartalnik  
 strona www: <http://www.vetsci.org/paper/list.html>

Serwisy internetowe tych czasopism, a w razie potrzeby również ich wydawców, poddano audytowi według formularza będącego wypadkową propozycji Björka i Holmströma (2006), Knight i Steinbach (2008) oraz Haustein (2012). W wykazie badanych parametrów uwzględniono tylko takie, których ocena nie wymagała przeprowadzenia badań empirycznych z udziałem osób trzech (np. potencjalnych lub rzeczywistych użytkowników, w szczególności autorów mających już na swoim koncie publikacje w danym czasopiśmie). Zbiorcze zestawienie rozpatrywanych cech wraz z ich wartościami dla każdego z periodyków przedstawiono w tabeli 3.

Tabela 3. Wyniki audytu serwisów internetowych wybranych czasopism z zakresu medycyny weterynaryjnej

	Acta Vet Scand	Vet Rec	NZ Vet J	Can Vet J	J Vet Sci
<b>Parametr</b>					
Model biznesowy					
Czasopismo komercyjne	-	+	+	+	-
Open access	+	-	-	-	+
Stanowisko wydawcy wobec autoarchiwizacji	Creative Commons	żółty wydawca	niebieski wydawca	biały wydawca	Creative Commons
Model dystrybucji	E	D+E	D+E	D+E	E
<b>Infrastruktura</b>					
Cena prenumeraty					
Wersja <i>online</i>	OA	554 USD	b.d.	185 USD	OA
Wersja drukowana + <i>online</i>		681 USD	b.d.		

c.d. tabeli 3

	Acta Vet Scand	Vet Rec	NZ Vet J	Can Vet J	J Vet Sci
Średnia cena artykułu w bazach	OA	24 GBP	OA	32 USD	OA
Opłata za publikację (ang. <i>fee</i> ) <sup>1</sup>	2410 USD	-	-	160 CND	<300 USD
Opłata za publikację OA	nie dotyczy	1950 GB	tak - b.d.	-	-
Dopłata za druk kolorowy	+	-	-	150 CND	+
Opłata za reprint	+	-	-	-	+
Indeksowanie w bazach piśmiennictwa					
Index Veterinarius	-	+	+	-	-
Agricola	-	+	+	+	+
CAB Abstracts	+	+	+	+	+
Medline	-	-	+	-	+
Scopus	-	+	-	-	-
JRC	+	+	+	+	+
Scimago	+	+	+	+	+
Wskazanie innych baz	+	+	+	-	-
Obecność w rankingach dziedzinowych					
Red Jaspers	+	+	+	+	+
BLVMS	+	+	+	+	-
Kompletność i użyteczność serwisu www					
Instrukcje dla autorów	zdawkowe	zdawkowe	zdawkowe	zdawkowe	zdawkowe
Standardy międzynarodowe	+	+	+	+	+
Rekomendacje i wymogi bioetyczne	+	+	+	+	+
Wytyczne dotyczące konfliktu interesów	-	+	+	+	+
Instrukcje dla recenzentów	+	+	+	+	+
Pomoce dla recenzentów typu <i>checklist</i>	-	+	-	-	-
Odpłatna usługa <i>proof reading</i>	-	+	-	-	-
Aplikacje z rodziny JMS	+	+	+	+	-
Kanały alertowe (RSS...)	-	+	-	-	-
Wartość dodana publikacji elektronicznej					
<i>Rich Internet Publication</i>	-	-	-	+	-
Publikacje semantyczne	-	-	-	-	-

<sup>1</sup> Według stawek podanych na stronie www J Vet Sci cena 10-stronicowego artykułu z jedną ilustracją kolorową wynosi 750 USD, bez kosztów korekty językowej.

c.d. tabeli 3

	Acta Vet Scand	Vet Rec	NZ Vet J	Can Vet J	J Vet Sci
<b>Publiczność odbiorcza</b>					
Zasięg geograficzny przedmiotowy	cały świat				
Deklarowany zasięg geograficzny odbioru	cały świat				
Liczba prenumeratorów instytucjonalnych	b.d.	b.d.	2000	b.d.	b.d.
Liczba abonentów kanałów alertowych	b.d.	b.d.	b.d.	b.d.	b.d.
Parametry cytowań	zob. tabela 4				
Wskazanie tytułów cytujących	-	-	+	-	-
Monitoring pobrań pełnego tekstu	+	-	-	-	-
Opcja <i>Higly accessed</i>	+	-	-	-	-
Ścieżka wyszukiwania / przeglądania	+	+	+	+	+
Opcja <i>OnlineFirst</i>	-	+	-	-	-
Numery specjalne / monograficzne	+	+	-	-	-
<b>Prestiż</b>					
Indeksy wpływu	zob. tabela 4				
Żywotność tekstu w obiegu naukowym	zob. tabela 4				
Miejsce czasopisma w rankingach dziedzinowych	zob. tabela 4				
Subiektywne preferencje autora	zmienna delta liczona według skali Lickerta lub innej podobnej				
<b>Regulacja współpracy między autorami i redakcją</b>					
Tryb walidacji	niejawny				
Maksymalna liczba wezwań do korekty rękopisu (potem dyskwalifikacja)	b.d.	b.d.	b.d.	b.d.	brak wezwań
Odsetek rękopisów odrzuconych	b.d.	45%	56%	b.d.	62%
Skład komitetu redakcyjnego	międzynarodowy				
Liczba członków komitetu redakcyjnego	18	3	9	8	35
Skład kolegium recenzentów	niejawny				
Średni czas od złożenia rękopisu do otrzymania recenzji	6 mies.	4 mies.	8 mies.	3 mies.	8 tyg.
Czas na odpowiedź po recenzji	3 mies.	b.d.	b.d.	b.d.	8 tyg.
Średni czas oczekiwania na publikację	11 mies.	2 mies.	20 mies.	9 mies.	6 mies.
Średnia liczba współautorów artykułu	3,2	4,1	1,4	3,1	2,8
Liczba autorów z Europy Środkowej i Wschodniej w latach 2009-2013	18	34	14	4	6
Liczba autorów z polską afiliacją 2009-2013	2	9	4	1	2

Źródło: opracowanie własne na podstawie [Bjork, Holmström 2006, dok. elektr.; Knight, Steinbach 2008, dok. elektr.; Haustein 2012]



## Legenda i wnioski

Poniższe omówienie ogranicza się do tych rubryk tabeli 3, których zawartość można pożytecznie skomentować na podstawie aktualnej wiedzy informatologicznej. Przyjęto, że kryteria oceny wymienione w nagłówkach „Publiczność odbiorcza” i „Regulacja współpracy między autorami i redakcją”, wyrażające się głównie w wartościach liczbowych, nie wymagają dodatkowych objaśnień, a ich interpretacja przez autora uzależniona jest przede wszystkim od obiektywnych przymusów (zobowiązań, limitów czasowych, konieczności osiągnięcia określonego progu IF lub indeksu H itd.), którym podlega on w chwili dokonywania wyboru czasopisma.

**Model biznesowy.** Dwa z pięciu czasopism, *Acta Vet Scand* i *J Vet Sci*, wydawane są w formule Open Access. Pozostałe trzy to organy komercyjne, wydawane w systemie D+E (druk i równoległe wersja elektroniczna), lecz udostępniające bezpłatnie zawartości niektórych numerów. Archiwum na stronie domowej *Vet Rec* sięga retrospektywnie do roku 1996, artykuły *NZ Vet J* za lata 2009-2011 dostępne są *via* SciQuest, Australasian e-Library of Animal and Veterinary Science, zaś *Can Vet J*, korzystający z PubMed Central, odblokowuje pełne teksty po 6 miesiącach od publikacji.

Stanowisko wydawców w sprawie autoarchiwizacji zostało ustalone na podstawie instrukcji dla autorów udostępnionych w serwisach www, które dla pewności porównano z rekordami w bazie Romeo na platformie Sherpa. Żaden z trzech wydawców komercyjnych nie daje autorom pełnej swobody w dysponowaniu swoimi tekstami przed i po oficjalnej publikacji w czasopiśmie. Stowarzyszenia będące właścicielami *Can Vet J* wymagają od autorów bezwzględnej cesji praw autorskich i nie dopuszczają upubliczniania e-printów (biali wydawcy). Nieco mniej restrykcyjną politykę prowadzi BJM Group, wydawca *Vet Rec*, oddając autorom do zatwierdzenia rękopisy po recenzji (żółty wydawca), zaś względnie liberalny Thomson&Reuters, wydawca *NZ Vet J*, daje do dyspozycji tylko post-printy (w klasyfikacji Romeo: kolor niebieski). Uprawnienia w zakresie samoarchiwizacji są z perspektywy autora ważnym czynnikiem oceny czasopisma, gdyż rozpowszechnianie elektronicznych kopii swoich tekstów, zwłaszcza w starannie wyselekcjonowanych miejscach, np. w archiwach dziedzinowych i w naukowych portalach społecznościowych (*Academia.edu*, *ResearchGate* itd., *Mendeley*), daje realną szansę zwiększenia ich widoczności, a przy tym również obserwacji przyrostu i kierunku cytowań, reakcji odbiorców wyrażanych za pośrednictwem narzędzi Web 2.0 (komentarze, tagi, rankingi), a nawet imiennego ich identyfikowania [Haustein, Peters 2012, dok. elektr.; Haustein, Siebenlist 2011].

**Infrastruktura.** W tej grupie parametrów, prócz oczywistych kwestii finansowych dotyczących wysokości opłaty za publikację z ewentualnymi narzuta-

mi (wiele czasopism, a w badanej stawce Can Vet J, dolicza słoną należność za druk kolorowy) lub opłaty opcjonalnej za udostępnienie tekstu w trybie Open Access, na uwagę autora zasługuje w szczególności informacja o tym, w jakich licencjonowanych i otwartych bazach piśmiennictwa naukowego indeksowana jest zawartość danego czasopisma. W formularzu wymieniono wprost kilka baz, którym w literaturze bibliometrycznej przypisuje się wysoki potencjał w generowaniu naukowego wpływu publikacji, a są to, w dziedzinie medycyny weterynaryjnej, bazy Agricola, CAB Abstracts, Medline i Scopus. Rejestracja w tych bazach stanowi 10% oceny wartości czasopisma w autorytatywnym rankingu Basic List of Veterinary Medicine Serials, opracowywanym przez Veterinary Medical Libraries Section w ramach Medical Libraries Association [Ugaz et al. 2010, dok. elektr.]. Dla naukowców reprezentujących nauki biomedyczne ważne jest ponadto, czy inne bazy, o węższym zakresie, w których ewentualnie rozpisywane jest czasopismo, wchodzi w skład pakietów dystrybuowanych przez wielkich dostawców dostępu do piśmiennictwa naukowego, zwłaszcza tych współpracujących z uczelniami wyższymi, np. EBSCOHost. Nie należy też zaniedbywać rosnącej roli wyszukiwarki Google Scholar jako narzędzia istotnie wpływającego na widoczność publikacji, a zarazem umożliwiającego podgląd przyrostu i struktury jakościowej cytowań w czasie rzeczywistym. Wiedza o tym, w jakich bazach rejestrowana jest treść czasopisma, ma też wartość praktyczną, ponieważ autor, wykorzystując systemy organizacji informacji (SOI) charakterystyczne dla danej bazy, może retrospektywnie lub na bieżąco śledzić trasy cyrkulacji swoich artykułów w społeczności naukowej, o ile, rzecz jasna, baza ta została wyposażona w liczniki (statystyki) pobrań pełnego tekstu i odsłon abstraktu i / lub narzędzia automatycznego grupowania cytowań bądź wizualizacji kwerend i ścieżek przeglądania, realizowanych przez czytelników artykułu.

Kolejnym polem oceny z punktu widzenia autora jest kompletność i użyteczność strony domowej czasopisma. Podstawowe standardy jakości serwisów internetowych w komunikacji naukowej wymagają, aby witryna podmiotu świadczącego usługi wyczerpująco i przejrzysto powiadamiała potencjalnych klientów o zakresie oferty i o warunkach optymalnego z niej skorzystania. Należy więc zakładać, że zainteresowany autor znajdzie na stronie www czasopisma co najmniej informacje o minimalnych wymogach formalnych i treściowych, których spełnienie uchroni go przed ryzykiem odrzucenia rękopisu z powodów pozamerytorycznych. W tym kontekście duże znaczenie ma obszerność i szczegółowość instrukcji dla autorów, od których naukowiec ma prawo oczekiwać uregulowania kwestii kluczowych dla publikowania w naukach biomedycznych, takich jak zabezpieczenie przed konfliktem interesów (m.in. w przypadku badań sponsorowanych), atrybucja autorstwa w przypadku publikacji wieloautorskich, uzyski-

wanie akredytacji komisji bioetycznych lub wymóg uprzedniej rejestracji badań klinicznych. Warunków tych nie spełniają dwa z pięciu analizowanych czasopism. W pozostałych instrukcje dla autorów są rozbudowane i czytelnie posegmentowane. Zawierają zarówno precyzyjne wskazówki metodyczne i redakcyjno-edytorskie (NZ Vet J odsyła np. do międzynarodowego standardu „Uniform Requirements for Manuscript Submitted to Biomedical Journals”, wydanego przez International Committee for Medical Journal Editors, ICMJE, a J Vet Sci rekomenduje „Good Publication Practice Guidelines for Medical Journals”), jak i dokładne informacje o warunkach weryfikowania zgodności publikacji z obowiązującymi standardami bioetycznymi. Wydawcy Vet Rec, Acta Vet Scand i J Vet Sci odwołują się w tym względzie do powszechnie stosowanych rekomendacji międzynarodowych, tj. dokumentu Committee on Publication Ethics (COPE) oraz World’s Association’s Declaration of Helsinki. Dla autora polskiego kłopotliwy może być natomiast fakt, że w instrukcjach NZ Vet J wymienia się wyłącznie regulacje nowozelandzkie.

W czasopiśmie o szerokim profilu tematycznym nie bez znaczenia jest również udostępnianie na stronach www czasopiśm instrukcji dla recenzentów, a nawet szczegółowych kwestionariuszy oceny jakości typu *checklist*, w których określa się m.in. wyznaczniki kwalifikacji artykułu jako oryginalnej pracy badawczej (najcenniejszej w aspekcie liczby cytowań i indeksów wpływu) oraz mierniki poprawności metodologicznej. Za pomocą takiego narzędzia autor może dokonać samodzielnej oceny własnego tekstu *ex ante*, przed wysłaniem go do redakcji.

Jak postulują Björk i Holmström, Haustein oraz Cookson et al., przed podjęciem decyzji o wyborze czasopisma autor powinien wiedzieć jak najwięcej o jakości i wydajności procesu recenzowania artykułów [Björk, Holmström 2006, dok. elektr.; Haustein 2012; Cookson et al. 2012, dok. elektr.]. Badania dowodzą, że naukowcy są gotowi tolerować długi czas oczekiwania na publikację, tzw. *willingness-to-wait* (WTW), jeżeli negocjacje z recenzentami, jakkolwiek żmudne i wieloetapowe, przynoszą w efekcie wymierną poprawę w warstwie merytorycznej tekstu, a sam artykuł trafia w czasopiśmie do rubryki *high quality peer-reviewed research paper* [Rousseau, Rousseau 2012, p. 1213]. W praktyce autorowi, który nigdy nie przebył procedury walidacji w danym czasopiśmie, bardzo trudno jest zdobyć informacje o tym, jak ona faktycznie przebiega i jaka jest jej wartość dodana. W równaniu zysków i nakładów niewiadoma ta wydatnie podnosi poziom ryzyka, ponieważ bezzasadne i bezproduktywne przeciąganie procesu recenzji może spowodować stratę możliwości opublikowania badań pionierskich, przekroczenie wiążących autora terminów, a w następstwie nawet rotację ze stanowiska zatrudnienia, nie mówiąc już o szkodzie poznawczej w postaci przedawnienia samego artykułu [Luwel, van Wijk 2012, dok. elektr.].

Pośrednim prognostykiem jakości procesu wartościowania może być jednak obecność w serwisie www czasopisma intranetowych aplikacji z rodziny Journal Management Systems, służących do sformalizowanej obsługi *online* kontaktów między redakcją, recenzentami i autorami. Narzędzia te nie gwarantują wprawdzie bezkolizyjnej współpracy, niemniej w dużej mierze wykluczają wpływ rozmaitych czynników dezorganizujących, cechujących np. zwykłą korespondencję mailową [Hasan, Abuelrub 2013, dok. elektr.]. Rozwiązania tego rodzaju stosowane są w czterech audytowanych czasopismach. Dominuje aplikacja ScholarOne, przeznaczona zasadniczo do przesyłania odpowiednio sformatowanych rękopisów *online*, a w NZ Vet J funkcjonuje dodatkowo CATS, autorski system Thomson&Reuters, wspomagający wymianę informacji między autorem a recenzentami. Wobec faktu, że we wszystkich pięciu czasopismach ewaluacja ma w całości charakter niejawnny, co zastrzega się w instrukcjach dla autorów, dostępność narzędzi JSM powinna być postrzegana jako cenne udogodnienie, zwłaszcza wówczas, gdy – jak w przypadku CATS – pozwalają one na bieżąco obserwować status przedłożonego tekstu i monitorować kolejne kroki pracy recenzenta.

Dla autorów z dziedziny nauk biomedycznych inną atrakcyjną cechą czasopisma jest możliwość opublikowania swojego artykułu w technologii Rich Internet Publication (ang. *enhanced publication*), która polega na wmontowywaniu w dokument tekstowy artefaktów informacyjnych wyrażonych w kodach znaczących innych niż pismo, tj. multimediów (filmy, wideokasty, podcasty, animacje, nagrania dźwiękowe), pakietów zawierających tzw. *raw data*, czyli surowe dane, np. skompresowane, nieprzetworzone zbiory danych zarejestrowanych przez urządzenia pomiarowe w trakcie eksperymentów itp. [Breure, Voorbij, Hoogerwerf 2011, dok. elektr.].

**Prestiż czasopisma.** Wśród parametrów skupionych w tej klasie odrębnego wyjaśnienia wymaga ten określony jako „subiektywne preferencje autora”. Jest to w istocie wartość aleatoryczna, zmienna inter- i intrapersonalnie, wyznaczana przez wiązkę czynników określających sytuację autora w punkcie N na osi czasu, tj. w chwili przemyślenia nad wyborem czasopisma. Do czynników tych zalicza się z pewnością aktualne położenie zawodowe, ale też osobiste wyobrażenia i afekty. Rozsądnie jest zakładać, że wewnętrzne nastawienia autora decydują o tym, jaką wagę w ocenie rentowności publikacji przypisze on poszczególnym kryteriom obiektywnym. Proponowane tu podejście benchmarkingowe powinno umożliwić badaczowi przeprowadzenie oceny w taki sposób, by uwarunkowania osobiste nie zafałszowały jej końcowego rezultatu. W tym względzie pomocne będzie udzielenie sobie odpowiedzi na następujące pytania:

- Czy mam gotowy artykuł i szukam czasopisma, które go opublikuje, czy też zamierzam odpowiedzieć na zaproszenie (ang. *call for papers*) do zgłaszania tekstów mających wypełnić np. numer monograficzny?

- Jeśli mam gotowy artykuł, to czy w danym czasopiśmie w ostatnich kilku latach publikowane były teksty o podobnej tematyce?
- Jeśli nie, to czy ukazały się w nim teksty mogące stanowić podstawę teoretyczną lub wzorzec metodologiczny dla mojego artykułu?
- Czy mam świadomość, że czasopismo o zasięgu międzynarodowym oczekująco do zasady artykułów o użyteczności ponadlokalnej?
- Czy mogę dołączyć się do już uformowanej grupy badawczej i wraz z nią przeprowadzić badania (napisać artykuł), czy też muszę sam skompletować zespół?
- Jaką metodą i w jakim czasie jestem w stanie stworzyć zespół współautorski?
- Jak wiele czasu mogę poświęcić na stworzenie artykułu, uwzględniając np. konieczność uzyskania akredytacji krajowej lub zagranicznej komisji bietycznej (bądź załatwienia innych formalności)?
- Jak długo mogę czekać na wynik recenzji i decyzję o dopuszczeniu do publikacji?
- Jak długo mogę czekać na samą publikację?
- Czy naglą mnie terminy (np. habilitacja, rozliczenie grantu)?
- Jak ważny jest dla mnie środowiskowy prestiż danego czasopisma?
- Jak ważne są dla mnie ewentualne czynniki pozamerytoryczne, np. chęć zadowolenia oczekiwania mojego mistrza / przełożonego?
- Jak ważne są dla mnie pozytywne lub negatywne doświadczenia kolegów, którzy znajdują się lub znajdowali w sytuacji podobnej do mojej?
- Czy jestem skłonny brać te rekomendacje / porady za dobrą monetę?
- Jak wiele mogę stracić – w kontekście zawodowym – jeżeli dokonam błędnego wyboru?
- Czy potrafię określić przyczyny wyjaśniające ogólnie wysoki odsetek odrzuconych rękopisów w czasopismach mających wysoką punktację MNiSW?

Idąc za sugestią Knight i Steinbach, można wyobrazić sobie, że autor zobiektywizuje odpowiedzi na powyższe pytania za pomocą skali Lickerta lub dowolnej innej prostej skali pomiarowej, a następnie potraktuje uzyskane wyniki jako wagi, które przyłoży do poszczególnych kryteriów obiektywnych wskazanych w formularzu oceny (tabela 3) [Knight, Steinbach 2008, dok. elektr.]. W praktyce przykładowe postępowanie autora przedstawiałoby się następująco:

- 1) autor wytypował z listy MNiSW czasopisma o punktacji odpowiedniej do swoich potrzeb,
- 2) scharakteryzował je za pomocą relewantnych partii kwestionariusza (tabela 3),
- 3) spełnienie przez czasopismo danego kryterium zakodował, wpisując wartość „1” w odpowiednią rubrykę formularza (zgodnie z intuicyjną zasadą: „1” – cecha obecna, „0” – brak cechy),

- 4) przechodząc do oceny swoich motywacji subiektywnych, określił jako priorytetowy krótki czas oczekiwania na publikację (np. maksimum 13 miesięcy), tzn. przypisał mu wartość skrajną, „5”, w skali „zupełnie nieważne – bardzo ważne”,
- 5) jeżeli w analizowanej stawce czasopism znalazły się wydawnictwa oferujące czas oczekiwania krótszy niż 14 miesięcy, to przyporządkowane im wartości „1” w każdym przypadku pomnoży przez „5”, zwiększając wagę tego kryterium w kalkulacji sumarycznej,
- 6) zsumowuje wyniki dla każdego czasopisma, uzyskując wynik w postaci liczby naturalnej lub mieszanej,
- 7) czasopismo, które uzyska największą ocenę, będzie dla niego benchmarkiem, czyli wyrażonym liczbowo celem określającym najbardziej opłacalną w danych okolicznościach strategię publikacyjną.

Niezależnie jednak od tego, co powiedziano wcześniej, w samodzielnej ewaluacji nie można sztucznie umniejszać znaczenia obiektywnych wskaźników wpływu i pozycji czasopisma w międzynarodowych rankingach dziedzinowych. Dla badanych czasopism zobrazowano je zbiorczo w tabeli 4.

Tabela 4. Pozycja czasopism w indeksach cytowań, indeksach wpływu i rankingach dziedzinowych

	VET REC	ACTA VET SCAND	NZ VET J	CAN VET J	J VET SCI
<b>Journal Citation Reports 2013</b>					
Total cites	9859	544	1457	2315	805
IF 2013	1.633	1.382	1.219	0.474	1.142
5-years IF	1.579	1.719	1.322	0.982	1.317
Immediacy Index	0.552	0.216	0.443	>10.0	0.029
Articles	250	88	61	0.00288	70
Cited Half-life	>10.0	>10.0	>10.0	0.288	5.9
Eigenfactor® Score	0.00983	0.00261	0.00182	2315	0.00185
Article Influence® Score	0.411	0.481	0.385	0.474	0.358
<b>Scimago 2013</b>					
Kwartyl	Q2	Q2	Q1	Q2	Q2
Indeks H	74	31	32	40	26
SJR	0,466	0,479	0,612	0,377	0,484
Cytowania 3 lata	1393	317	230	366	255
Citable 2013	1291	186	166	526	187
Cites/Docs 4 lata	1,10	1,75	1,35	0,82	1,51
% IC	7,82	0,00	13,24	12,24	22,86

c.d. tabeli 4

	VET REC	ACTA VET SCAND	NZ VET J	CAN VET J	J VET SCI
<b>Red Jaspers 2010</b>					
Rank RJ	3740	5038	5496	5786	7433
JII	8,28	0,70	0,93	1,42	0,27
PII	32,33	18,21	15,02	13,02	4,60
ISI IF 2008-2010	1,99	1,39	2,19	1,01	–
ISI IF 2006-2010	2,43	2,04	1,68	1,31	–
ISI IF 2004-2010	2,46	2,36	1,55	1,27	–

## Legenda

Wskaźniki stosowane w Journal Citation Reports 2013

5-Year IF: średnia liczba cytowań artykułów opublikowanych w czasopiśmie w ciągu ostatnich pięciu lat, odnotowana w roku ogłoszenia danej edycji JCR; miernik szczególnie cenny w dyscyplinach, w których cytowania utrzymują się przez dłuższy czas po ukazaniu się publikacji; Immediacy Index: średni czas upływający od momentu publikacji artykułu do pierwszych jego cytowań; Cited Half-Life: miernik „żywołności” i „przeżywalności” artykułu w obiegu naukowym, pokazujący, jak długo po publikacji artykułu utrzymują się jego cytowania; Eigenfactor Score: indeks wpływu czasopisma obliczony na podstawie liczby cytowań, ale też z uwzględnieniem pochodzenia cytowań; ES przypisany dwóm różnym czasopismom będzie różnił się w zależności od tego, czy opublikowane w nich artykuły były cytowane w czasopismach, które same mają wysoki IF; Article Influence Score: odpowiada *grosso modo* wskaźnikowi 5-Year IF, ; najlepiej jeśli jego wartość dla czasopisma plasuje powyżej 1.00

Wskaźniki stosowane w rankingu Red Jaspers

JII: Journal Influence Index, współczynnik obliczony na podstawie własnego algorytmu Red Jaspers, im wyższy, tym lepiej dla oceny naukowej wpływowości czasopisma; PII: Paper Influence Index, iloraz JII i ogólnej liczby artykułów opublikowanych w czasopiśmie, pomnożony przez 1000; Rank: pozycja czasopisma w rankingu według wielkości wskaźnika JII.

Przedstawione w artykule robocze narzędzie do benchmarkingowej oceny rentowności czasopisma można by z pewnością rozbudowywać i niuansować, chociaż stosowanie go stałoby się wówczas niezwykle żmudne i czasochłonne. W zamyśle ma to być natomiast poręczny pomocnik dla autorów, stosunkowo łatwy w użyciu, w miarę możliwości intuicyjny i niewymagający rozległej wiedzy z dziedziny bibliometrii. Zgodnie z tendencją, która silnie zaznacza się w bada-

niach informatologicznych w ostatnim dziesięcioleciu, opisaną tu propozycję benchmarkingu należy koniecznie podbudować badaniem rzeczywistych zachowań autorów stojących przed wyborem strategii publikacji w czasopismach naukowych z dziedziny nauk weterynaryjnych. Pierwszy pilotaż tego narzędzia przeprowadzono z udziałem pięciorga pracowników uniwersyteckich wydziałów medycyny weterynaryjnej podczas warsztatów z cyklu „Akademia Medycyny Behavioralnej” zorganizowanych w czerwcu 2014 r. przez Katedrę Diagnostyki Klinicznej Wydział Medycyny Weterynaryjnej UWM w Olsztynie, Katedrę Chorób Wewnętrznych z Kliniką UWM w Olsztynie, Katedrę Chorób Wewnętrznych z Kliniką Koni, Psów i Kotów UP we Wrocławiu, Sekcję Patologii i Fizjologii Psów i Kotów PTNW oraz firmę Vetoquinol Biowet Sp. z o.o. Druga faza pilotażu, zakładająca ostateczne dopracowanie narzędzia, została przeprowadzona na większej próbie badawczej podczas XX Jubileuszowego Kongresu Polskiego Stowarzyszenia Lekarzy Weterynarii Małych Zwierząt w październiku 2014 r. W pierwszym kwartale 2015 r. planowane jest natomiast regularne badanie strategii publikacyjnych polskich uczonych z dziedziny medycyny weterynaryjnej za pomocą spersonalizowanych formularzy *online*, stworzonych na podstawie przetestowanego narzędzia benchmarkingowego.

## BIBLIOGRAFIA

- Amat, Carlos Benito (2008). *Editorial and Publication Delay of Papers Submitted to 14 Selected Food Research Journals. Influence on Online Posting* [dok. elektr.]. *Scientometrics* Vol. 74, No. 3, p. 379-389. [http://digital.csic.es/bitstream/10261/3640/3/Manuscrito\\_Scientometrics.pdf](http://digital.csic.es/bitstream/10261/3640/3/Manuscrito_Scientometrics.pdf) [odczyt: 01.07.2014].
- Björk, Bo-Christer; Holmström, Jonas (2006). *Benchmarking scientific journals from the submitting author's viewpoint* [dok. elektr.]. *Learned Publishing* Vol. 19, No. 2, p. 147-155. <http://dx.doi.org/10.1087/095315106776387002> [odczyt: 01.07.2014].
- Breure, Leen; Voorbij, Hans; Hoogerwerf, Maarten (2012). *Rich Internet Publications: „Show What You Tell”* [dok. elektr.]. *Journal of Digital Information* Vol. 13, No. 1. <http://journals.tdl.org/jodi/index.php/jodi/article/view/1606> [odczyt: 01.07.2014].
- Coleman, Anita (2007). *Assessing the value of a journal beyond the impact factor*. *Journal of the American Society for Information Science and Technology* Vol. 58, Issue 8, p. 1148-1161.
- Cookson, Rod et al. (2012). *What do Canadian Journal Editors Want from Peer Review?* [dok. elektr.]. *Thomson&Reuters Editors' Bulletin* Vol. 8, No. 1, p. <http://www.tandfonline.com/doi/full/10.1080/17521742.2012.713757> [odczyt: 01.07.2014].
- Craig, Iain; Ferguson, Liz; Finch, Adam T. (2010). *Journals ranking and impact factors: how the performance of journals is measured* In: Cope, Bill; Phillips, Angus eds. *The Future of the Academic Journal*. Waltham: Chandos Publishing, p. 159-193.



- Hasan, Layla; Abuelrub, Emad (2013). *Usability Testing for IAJIT OpenConf Journal Management System* [dok. elektr.]. *Journal of Software* Vol. 8, No. 2, p. 387-396. <http://ojs.academypublisher.com/index.php/jsw/article/view/7684> [odczyt: 01.07.2014].
- Haustein, Stefanie (2012). *Multidimensional Journal Evaluation: Analyzing Scientific Periodicals beyond the Impact Factor*. Berlin, Boston: De Gruyter/Saur, 408 p.
- Haustein, Stefanie; Peters, Isabella (2012). *Using social bookmarks and tags as alternative indicators of journal content description* [dok. elektr.]. *First Monday* Vol. 17, No. 11. <http://firstmonday.org/ojs/index.php/fm/article/view/4110/3357> [odczyt: 01.07.2014]. <http://www.firstmonday.org/ojs/index.php/fm/article/viewArticle/4110/3357> [odczyt: 27.06.2013].
- Haustein, Stefanie; Siebenlist, Tobias (2011). *Applying social bookmarking data to evaluate journal usage*. *Journal of Informetrics* Vol. 5, Issue 3, p. 446-457.
- Jacso, Peter (2013). *Multidimensional Journal Evaluation: Analyzing Scientific Periodicals beyond the Impact Factor*. *Online Information Review* Vol. 37, Issue 4, p. 662-663.
- Kamalski, Judith; Plume, Andrew; Amin, Mayur (2014). *A Publishing Perspective on Bibliometrics*. In: Cronin, Blaise; Sugimoto, Cassidy R. eds. *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*. Cambridge: The MIT Press, p. 385-396.
- Kaszubska, Elzbieta (2012). *Benchmarking – nowoczesna metoda planowania strategicznego* [dok. elektr.]. [archiwum.parp.gov.pl/partnerinfo/przewodnik/11dz02.rtf](http://archiwum.parp.gov.pl/partnerinfo/przewodnik/11dz02.rtf) [odczyt: 01.07.2014].
- Knight, Linda V.; Steinbach, Theresa A. (2008). *Selecting an Appropriate Publication Outlet: A Comprehensive Model of Journal Selection Criteria for Researchers in a Broad Range of Academic Disciplines* [dok. elektr.]. *International Journal of Doctoral Studies* Vol. 3, p. 59-79. <http://www.ijds.org/Volume3/IJDSv3p059-079Knight84.pdf> [odczyt: 01.07.2014].
- Komunikat Ministra Nauki i Szkolnictwa Wyższego w sprawie wykazu czasopism naukowych z 21 grudnia 2012 r.* [dok. elektr.]. [http://www.nauka.gov.pl/komunikaty/komunikat-ministra-nauki-i-szkolnictwa-wyzszego-w-sprawie-wykazu-czasopism-naukowych\\_31733.html](http://www.nauka.gov.pl/komunikaty/komunikat-ministra-nauki-i-szkolnictwa-wyzszego-w-sprawie-wykazu-czasopism-naukowych_31733.html) [odczyt: 01.07.2014].
- Lozano, George A.; Larivière, Vincent; Gingras, Yves (2012). *The weakening relationship between the impact factor and papers' citations in the digital age*. [dok. elektr.]. *Journal of the American Society for Information Science and Technology* Vol. 63, Issue 11, p. 2140-2145. <file:///Users/Aga/Downloads/Lozano+2012.pdf> [odczyt: 01.07.2014].
- Luwel, Marc; van Wijk, Eric (2012). *Publication Delays Revisited: 1998-2012* [dok. elektr.]. [http://sticonference.org/Proceedings/vol2/Luwel\\_Publication\\_569.pdf](http://sticonference.org/Proceedings/vol2/Luwel_Publication_569.pdf) [odczyt: 01.07.2014].
- Mierzejewska, Bożena I. (2010). *The Ecology of Academic Journals*. Waltham: Chandos Publishing, 200 p.
- Mikulski, Henryk; Mikulska, Agnieszka (2006). *Benchmarking jako nowoczesna metoda zarządzania przedsiębiorstwem* [dok. elektr.]. *Monitor Rachunkowości i Finansów* nr 6. [http://www.mrf.pl/index.php?mod=m\\_artykuly&cid=87&id=1](http://www.mrf.pl/index.php?mod=m_artykuly&cid=87&id=1) [odczyt: 01.07.2014].
- Misteli, Tom (2013). *Eliminating the impact of the Impact Factor* [dok. elektr.]. *Journal of Cell Biology* Vol. 201, No. 5, p. 651-652. <http://jcb.rupress.org/content/201/5/651.full> [odczyt: 01.07.2014].

- Rousseau, Ronald (2002). *Journal evaluation. Technical and practical issues* [dok. elektr.]. *Library Trends* Vol. 50, No. 3, p. 418-439. [https://www.ideals.illinois.edu/bitstream/handle/2142/8418/librarytrendsv50i3i\\_opt.pdf?sequence=1](https://www.ideals.illinois.edu/bitstream/handle/2142/8418/librarytrendsv50i3i_opt.pdf?sequence=1) [odczyt: 01.07.2014].
- Rousseau, Sandra; Rousseau, Ronald (2012). *Interactions between journal attributes and authors' willingness to wait for editorial decisions*. *Journal of the American Society for Information Science and Technology* Vol. 63, Issue 6, p. 1213-1225.
- Sapa, Remigiusz (2007). *Indywidualne strony www historyków i matematyków we współczesnej komunikacji naukowej*. W: Kocójowa, Maria red. *Między przeszłością a przyszłością: książka, biblioteka, informacja naukowa – funkcje społeczne na przestrzeni wieków*. Kraków: Wydaw. Uniwersytetu Jagiellońskiego, s. 272-278.
- Sapa, Remigiusz (2008). *Globalizacja komunikacji w naukach przyrodniczych w Polsce i jej niektóre konsekwencje*. W: Pietruch-Reizes, Diana red. *Zarządzanie informacją w nauce*. Katowice: Wydaw. UŚ, s. 90-102.
- Swan, Alma P.; Brown, Sheridan N. (2004). *JISC/OSI Journal Authors Survey Report. „JISC and OSI”* [dok. elektr.]. [www.jisc.ac.uk/uploaded\\_documents/JISCOAreport1.pdf](http://www.jisc.ac.uk/uploaded_documents/JISCOAreport1.pdf) [odczyt: 01.07.2014].
- Szprot, Jakub red. (2104). *Otwarta nauka w Polsce 2014. Diagnoza* (dok. elektr.). <http://pon.edu.pl/index.php/nasze-publicacje?pubid=13> [odczyt: 05.11.2014].
- Ugaz, Ana G. et al. (2010). *Basic list of veterinary medical serials, third edition: using a decision matrix to update the core list of veterinary journals* [dok. elektr.]. *Journal of Medical Libraries Association* Vol. 98, No. 4, p. 282-292. <http://europemc.org/articles/PMC2947140/> [odczyt: 01.07.2014].

---

# Systemy informacyjne wspomagające podejmowanie decyzji w zakresie edukacji na poziomie wyższym

Małgorzata Jaskowska, Magdalena Wójcik

Instytut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Jagielloński

---

## Przedmiot i cel badań

Przedmiot badania stanowią systemy eksperckie wspierające studentów i kandydatów na studia w podejmowaniu decyzji związanych z procesem edukacji w szkołach wyższych. Celem przeprowadzonej analizy było wskazanie głównych tendencji związanych z projektowaniem systemów doradczych (ang. *e-advising*), wytypowanie najlepszych praktyk w tym zakresie, a także sformułowanie – w oparciu o analizę literatury przedmiotu oraz studia przypadków – rekomendacji dla projektantów uczelnianych systemów informacji o kierunkach kształcenia, przez wyznaczenie rozwiązań modelowych.

## Terminologia i definicje

Uczelniane doradztwo polega na wspomaganii studentów przez pracowników uczelni w zakresie podejmowania słusznych decyzji związanych z całym tokiem studiów, począwszy od wyboru kierunku, na określeniu planowanej drogi zawodowej zakończywszy. Doradztwo może, jak wykazuje analiza piśmiennictwa, przybierać różne formy ze względu na rodzaj relacji między uczestnikami tego procesu. Doradztwo nakazowe (ang. *prescriptive advising model*) opiera się na wiodącej roli doradcy i biernej postawie studenta. Na drugim biegunie znajduje

się tzw. model rozwojowy (ang. *developmental advising model*), w którym doradca i student wspólnie podejmują decyzje i dzielą się odpowiedzialnością, przy czym student jest zachęcany do samodzielnej eksploracji źródeł informacji proponowanych przez doradcę. Perspektywa łącząca działania charakterystyczne dla obu omówionych typów nosi nazwę doradztwa zintegrowanego (ang. *integrated advising model*). Ciekawym modelem wydaje się doradztwo zaangażowane (ang. *engagement model*), oparte na budowaniu przez doradcę relacji ze studentem w celu wzmocnienia jego kompetencji personalnych i społecznych oraz budowania poczucia własnej wartości [Feghali, Zbib, Hallal 2011].

**E-doradztwo** (ang. *e-advising*) definiowane jest jako wykorzystanie narzędzi informatycznych dostępnych *online* w celu wspierania studentów w podejmowaniu decyzji (zwłaszcza zrutynizowanych), które przyczynia się do poprawy wydajności i dostępności doradztwa tradycyjnego, zmierzając tym samym do wspierania akademickiego sukcesu studentów [Zuhrieh, Shubair 2014].

Celem wprowadzania elektronicznych systemów doradczych na uczelniach, w których doradztwo ma dobrze ugruntowaną pozycję w systemie obsługi studentów, jest najczęściej nie tyle zastąpienie doradztwa tradycyjnego elektronicznym, ile raczej ułatwienie gromadzenia standardowych danych na temat studentów i ich postępów w nauce. Dzięki temu doradcy mogą skupić się na pogłębionej analizie danych i budowaniu w oparciu o nie rekomendacji dla studentów [Barker 2009; Feghali, Zbib, Hallal 2011]. Warto zaznaczyć, że choć systemy *e-advising* najczęściej funkcjonują jako wsparcie dla tradycyjnych procesów doradczych, to jednak mogą też być projektowane pod kątem obsługi studentów korzystających z nauczania na odległość (ang. *online students*), i wtedy całkowicie zastępować kontakt bezpośredni. Doradztwo jest niezbędnym komponentem w budowaniu pozytywnych doświadczeń akademickich, dlatego też tak istotny jest sposób prowadzenia działań doradczych, czy to w formie tradycyjnej, czy elektronicznej [Feghali, Zbib, Hallal 2011].

Na uczelniach polskich idea doradztwa w formie tradycyjnej jest słabo rozwinięta. Znajduje to odzwierciedlenie choćby w nazewnictwie funkcjonujących w szkołach wyższych systemów informacyjnych, opartych na technologii komputerowej: stosowane są raczej określenia związane z „obsługą” studentów lub „zarządzaniem” uczelnią niż z doradztwem bądź wspomaganiami decyzji. Jednak właśnie tym systemom chcemy w artykule poświęcić uwagę, analizując możliwości ich rozwoju w kierunku systemów doradczych, ponieważ, wobec braku doradztwa tradycyjnego, systemy elektroniczne mogłyby przejąć jego rolę. Wymagałoby to poszerzenia zakresu ich funkcji poza dotychczasowe: informacyjną i obsługi procesów. Aby móc sprawnie wypełniać zadania doradcze, nowe rozwiązania powinny być projektowane w oparciu o analizy funkcjonujących już

systemów, przede wszystkim badania ich jakości, a także wewnętrzne badania potrzeb wszystkich interesariuszy – użytkowników systemu. Powinny także być systematycznie udoskonalane przez poddawanie ich bieżącej ocenie.

## Podstawy analizy

Dokonałyśmy systematycznej analizy i krytyki piśmiennictwa naukowego z zakresu projektowania i oceny systemów wspierających podejmowanie decyzji w zakresie edukacji na poziomie wyższym. W celu zebrania relewantnego piśmiennictwa sięgnęłyśmy do następujących źródeł: Academic Search Complete (EBSCO), ERIC, Scopus, ISI Web of Science, Elsevier, Google Scholar. Artykuły relewantne do tematu były przez autorów określane za pomocą pojęć takich, jak: *e-advising*, *computer-based advising in educational systems*, *student information systems*, *technology-based advisory systems*, *academic advising*, *advising systems*, *recommendation systems*, *automated advising systems*, *student decision support system*. Szczególną wagę przywiązałyśmy do tekstów najnowszych (opublikowanych po 2008 roku), ponieważ zależało nam na wskazaniu aktualnej problematyki badawczej akademickich systemów rekomendacyjnych.

W celu wytypowania dobrych praktyk, stosowanych w projektowaniu systemów typu *e-advising* lub systemów zarządzania uczelnią z wbudowanymi funkcjami wspierania podejmowania decyzji przez studentów, dokonałyśmy systematycznego przeglądu sfery zasobów Web 2.0 w poszukiwaniu opisów wdrożeń i studiów przypadków. Przeszukałyśmy zasoby serwisów typu *video sharing* – YouTube i TED oraz *presentation sharing* – Slideshare i Slideboom, blogi i portale związane z wdrażaniem technologii informatycznych w szkolnictwie wyższym oraz mikroblog Twitter, używając takich samych haseł jak w przypadku wyszukiwania artykułów naukowych.

Zestawienie piśmiennictwa naukowego z wynikami wyszukiwania w sferze Web 2.0 pozwoliło na pełniejsze scharakteryzowanie przedmiotu badań, w tym wskazanie zarówno teoretycznych modeli prawidłowo zaprojektowanych systemów informacyjno-wyszukiwawczych, jak i przykładów konkretnych wdrożeń.

## Najlepsze praktyki, najpopularniejsze programy

Analiza piśmiennictwa wskazuje, że najpopularniejszym programem typu *student decision support system* na świecie jest zaprojektowane w Uniwersytecie Standforda oprogramowanie CourseRank, wdrożone od 2009 r. w ponad 500

uniwersytetach na całym świecie [CourseRank, dok. elektr.; Koutrika et al., 2009]. Artykuły naukowe relacjonujące badania prowadzone w oparciu o ten system świadczą, że jest on stale rozwijany i udoskonalany. Oprogramowanie należy do grupy aplikacji zaprojektowanych na potrzeby konkretnej uczelni i następnie wdrożonych w innych szkołach wyższych, podobnie jak polski system USOS.

Drugą grupę stanowią programy lokalne, zaprojektowane i wdrożone w jednej uczelni, jak Student Academic Advising System (SAAS) [Student Academic Advising System 2014, dok. elektr.] bądź eQuad [Waldner et al. 2012]. W polskich szkołach wyższych są to m.in.: Uczelnia.XP, HMS Solutions, Sokrates [Modele... 2011], EdukacjaCL [Sobecki 2012]. Takie rozwiązanie jest uzasadnione, dlatego że uwzględnia specyfikę uczelni [Shatnavi et al. 2014]. Analiza opisów wdrożeń publikowanych w sferze zasobów Web 2.0 pokazuje, że tendencja do tworzenia unikatowych, zaprojektowanych dla potrzeb konkretnej uczelni programów wspomagających realizację procesów edukacyjnych jest silna. Może o tym świadczyć liczba dostępnych w serwisie YouTube tutoriali obrazujących proces projektowania i tworzenia studenckich systemów informacyjnych (Student Information Systems – SIS) [Student..., dok. elektr. (a); Student..., dok. elektr. (b)].

Istnieje też grupa programów komercyjnych firm, takich jak: Red Lantern (DARS), Jenzabar, Datatel, Sungard (SCT Banner & CAPP), Conclusive Systems, PeopleSoft (Oracle) [Parameswaran et al. 2011], a w Polsce: Bazus [SIMPLE.EDU – zintegrowany system obsługi Uczelni 2014] i Invariant [System zarządzania dokumentacją dydaktyczną. S4. Sylabusy Siatki, Studia, Studenci, dok. elektr.].

## **Charakterystyka systemów akademickich**

Charakterystyka systemów informacyjnych została oparta na identyfikacji wzajemnie powiązanych w nich struktur: informacyjnej, funkcjonalnej, organizacyjno-przestrzennej i techniczno-technologicznej [Nowicki, Sitarska 2009]. Dodałyśmy do niej jeszcze jedną kategorię: rodzaje decyzji studenckich, jakie mogą być wspomagane przez system.

### **Struktura informacyjna**

Struktura informacyjna określa dostępne i przetwarzane zasoby informacyjne w systemie, tzw. węzły informacyjne. Akademickie systemy doradcze można scharakteryzować ze względu na zakres udostępnianych informacji, dzieląc je na systemy służące dokonywaniu wyborów w obszarach takich, jak: fakultety [Aslam, Nassimullah 2011]; specjalizacje [Grupe 2002]; kursy [Vialardi et al.

2009; Sobecki 2012; Shatnavi et al. 2014]; stypendia [Engin et al. 2014]. Zakres i szczegółowość informacji umieszczanych w konkretnych obszarach są niezwykle istotne w systemach eksperckich, o których mowa będzie dalej, ponieważ w głównej mierze na bazie wiedzy oparte są w nich reguły wspomagające podejmowanie decyzji. Im bardziej wyczerpująca jest charakterystyka obiektu (indywidualnego studenta, wykładowcy, kursu, nabytych kompetencji), tym większa szansa na skuteczny system podpowiedzi. Autorzy wskazują na dużą różnorodność źródeł informacji wykorzystywanych przez studentów, na przykład przy wyborze określonego kursu. Są to: opisy kursów, sylabusy, katalogi kursów, profile wykładowców, prywatne opinie absolwentów kursów wyrażane swobodnie na forach dyskusyjnych, rankingi i oceny kursów itd. [Kardan, Sadeghi 2013]. System powinien zatem uwzględniać tę szeroką paletę źródeł informacji i nie opierać możliwości wyboru jedynie na zawartości sylabusów.

W każdym przypadku definiowanie struktury informacyjnej systemu związane jest też z koniecznością przeprowadzenia badań potrzeb informacyjnych poszczególnych interesariuszy systemu.

## **Struktura organizacyjno-przestrzenna**

Struktura organizacyjno-przestrzenna określa miejsce poszczególnych węzłów informacyjnych w systemie oraz relacje pomiędzy nimi. Struktura ta jest często związana ze schematem organizacyjnym instytucji. W rozwiązaniach typu *e-advicing* każdy z wymienionych powyżej modułów (fakultety, specjalizacje, kursy, stypendia) może być odrębnym systemem informacyjnym bądź stanowić węzeł informacyjny w systemie zintegrowanym.

Analiza najlepszych praktyk pokazuje, że widoczna jest tendencja do tworzenia programów o elastycznej konstrukcji modułowej i otwartym kodzie, których funkcjonalności mogą być łatwo dostosowane do potrzeb konkretnej uczelni [Fedena, dok. elektr.]. Kolejną zaobserwowaną tendencją jest dążenie do integrowania rozproszonych systemów informacyjno-wyszukiwawczych w jedno uniwersalne narzędzie. Takie działanie odpowiada potrzebom studentów, którzy wyraźnie deklarują chęć korzystania z jednego systemu, skupiającego różne funkcjonalności [SIS, dok. elektr.].

Powstałe w ten sposób kompleksowe systemy typu *life map* pozwalają nie tylko efektywnie zarządzać bieżącymi procesami edukacyjnymi, ale także, a może przede wszystkim, planować przyszłe działania naukowe i zawodowe [Romano, dok. elektr.; Podcast..., dok. elektr.; Students..., dok. elektr. (c)]. Innym istotnym zagadnieniem jest ściśle integrowanie systemów informacyjno-wyszukiwawczych z platformami e-learningowymi w ramach jednego interfejsu [CobaltLMS, dok.

elektr.], które także znajduje odzwierciedlenie w piśmiennictwie naukowym, w pracach autorów omawiających problematykę projektowania systemów dostosowanych do potrzeb studentów uczących się zdalnie [Waldner, McDaniel, Widener 2011; Waldner et al. 2012]. Dla tej grupy odbiorców potrzeba stałego kontaktu z pracownikami uczelni i możliwość sprawnego organizowania procesu kształcenia za pomocą odpowiednich systemów jest szczególnie istotna. W literaturze przedmiotu podkreśla się potrzebę uwzględniania w procesie projektowania systemów przeznaczonych dla studentów, którzy nie odbywają zajęć na terenie uczelni, elementów sprzyjających budowaniu u nich poczucia więzi z uczelnią i innymi studentami [Waldner, McDaniel, Widener 2011; Waldner et al. 2012; Starks 2012].

Dążenie do integrowania rozproszonych aplikacji lub modułów w jeden spójny system sprzyja urzeczywistnieniu koncepcji wspomaganie przez systemy informacyjne realizacji typowych dla danej organizacji procesów. W procesowym (dynamicznym) podejściu do projektowania system rozpatrywany jest pod kątem kompleksowej obsługi procesów zachodzących w organizacji i w jej bezpośrednim otoczeniu. Podobnie jak w projektowaniu struktury funkcjonalnej i informacyjnej, celem jest usprawnienie realizacji działań organizacji. Dokonuje się to jednak nie przez odwzorowanie jej struktury i zadań realizowanych przez poszczególne jednostki organizacyjne lub przypisywanie funkcji poszczególnym węzłom informacyjnym. Istotny jest ciąg wzajemnie powiązanych działań, realizowanych w poprzek całej struktury organizacyjnej, w celu sprawnego przepływu ogółu zasobów. W systemach uczelnianych w ten sposób realizowane mogą być procesy: rejestracji na studia, specjalizacje i kursy, kształcenia e-learningowego, pozyskiwania stypendiów lub poszukiwania pracy tuż po studiach. Składać się one mogą *summa summarum* na jeden wspólny proces, rozpoczynający się od podjęcia edukacji, a zakończony znalezieniem pracy. Byłby to oczywiście model idealny.

## **Struktura funkcjonalna**

Struktura funkcjonalna systemu określa jego rolę we wspomaganie realizacji celów i zadań instytucji. W projektowaniu systemów pełni ona rolę nadrzędną w stosunku do pozostałych struktur. Każdy z wyróżnionych węzłów informacyjnych systemu ma wyznaczone indywidualne zadania: przetwarzanie danych oraz funkcję informacyjną.

W analizowanych na potrzeby tego tekstu badaniach podstawowe funkcje systemów zostały ujęte w trzy grupy: wspomaganie przebiegu studiów, wspomaganie kariery zawodowej i system doradztwa. Ten ostatni oparty jest na ułatwieniu komunikacji pomiędzy wszystkimi członkami społeczności uniwersyteckiej, przy wykorzystaniu dostępnych środków technicznych, w tym: tablic ogłoszeń, dys-



kusji *online*, FAQ, czatów, wideokonferencji [Ishak, Lehat 2012]. Analiza najlepszych praktyk wykazała, że obecnie podejmowane są próby udoskonalania systemów przez zapewnienie studentom możliwości formułowania pytań w języku naturalnym [Natural..., dok. elektr.] i dodawanie elementów multimedialnych, zwłaszcza tutoriali [CobaltLMS, dok. elektr.]. Widoczna jest także, choć na razie w stosunkowo niewielkim stopniu, tendencja do włączania w obręb systemów uczelnianych elementów funkcjonalności o charakterze społecznościowym, takich jak: rozbudowane profile użytkowników, możliwość zadawania pytań lub prowadzenia dyskusji *online* [Koutrika et al. 2009]. Widać to dobrze na przykładzie serwisu CobaltLMS [CobaltLMS, dok. elektr.].

## Struktura techniczno-technologiczna

W piśmiennictwie najczęściej wymienia się – ze względu na zasadę działania – dwa rodzaje systemów informacyjno-wyszukiwawczych funkcjonujących na uczelniach: systemy eksperckie i rozwiązania *data mining*. Systemy eksperckie, najczęściej przybierające formę tzw. *rule-based systems*, wykorzystują zdefiniowane reguły (klasyczne zasady logiki, regułę skojarzeń, zasady logiki rozmytej, zasadę podobieństw) do wspomagania i podejmowania trafnych decyzji [Duch dok. elektr.; Engin et al. 2014]. Z kolei *data mining* to systemy działające w oparciu o zaawansowane mechanizmy ekstrakcji danych z dużych zbiorów (reguła skojarzeniowa, klastry, drzewo decyzyjne, reguła najbliższego sąsiedztwa, analizy linków, sieci neuronowe, regresja, pozostałe, w tym ontologie) [Park et al. 2012; Vialardi et al. 2009; Sobecki 2012].

## Wspomaganie decyzji

W uczelnianym elektronicznym systemie informacyjnym funkcje doradcze realizowane są, jak wskazują autorzy analizowanego piśmiennictwa i wyniki obserwacji samych systemów, przy pomocy:

- wdrożonych, rozwijanych i udoskonalanych narzędzi komunikacyjnych, pozwalających studentom:
  - uzyskiwać odpowiedzi na pytania kierowane do pracowników uczelni;
  - korzystać z doradztwa społecznościowego (np. pytając, który z dostępnych kursów jest najciekawszy, zdaniem studentów, którzy już je ukończyli).
- systemów eksperckich, które, w oparciu o bazę wiedzy, wspomagają podejmowanie decyzji w zakresie pytań zrutyinizowanych (np. które kursy należy ukończyć, aby zrealizować wybrany moduł, ścieżkę, stopień studiów).

## Systemy eksperckie

Badania użyteczności serwisów uczelnianych przeprowadzane w grupie studentów i pracowników – zarówno opisywane w przytaczanych tu tekstach, jak też nasze (autorek) własne badania empiryczne – wskazują często na zgłaszany przez studentów postulat uzupełniania systemów informacyjnych o funkcje wspomaganie decyzji lub wyborów przez system podpowiedzi, symulowania możliwych rozwiązań i ich konsekwencji. Takich funkcji nie spełnia oczywiście zwykły informacyjny system bazodanowy, a tylko system ekspertowy. System ekspertowy na podstawie szczegółowej wiedzy i jedynie w jej granicach może wyciągać wnioski, podejmować decyzje, działać w sposób zbliżony do procesu rozumowania człowieka. Elementy składowe systemów ekspertowych to: 1) baza wiedzy zawierająca wiedzę dziedzinową, istotną dla podejmowanych decyzji, 2) maszyna wnioskująca, korzystająca z bazy wiedzy i 3) pamięć robocza, zawierająca wyniki pośrednie toku rozumowania. Baza wiedzy zawiera fakty, które są niezbędne do rozwiązania problemu w specyficznej dziedzinie i reguły, które kodyfikują wiedzę o obowiązujących w wybranej dziedzinie zależnościach (jeśli ..., to ...). Relacje odzwierciedlają zależności i asocjacje (skojarzenia) pomiędzy faktami w bazie wiedzy. Systemy ekspertowe służą do podejmowania rutynowych, powtarzalnych, łatwo dających się strukturalizować decyzji.

Obszary decyzyjne, w których uczelniane elektroniczne systemy doradcze są najczęściej wykorzystywane, to m.in.: dokonywanie wyborów fakultetów, specjalizacji i kursów, pozyskiwanie stypendiów, budowanie indywidualnych ścieżek edukacyjnych i podejmowanie decyzji dotyczących dalszego kształcenia [Biletskiy, Brown, Ranganathan 2009].

W analizowanym przez nas najnowszym piśmiennictwie omawiano różnorodnie zagadnienia związane z problematyką wykorzystania systemów eksperckich w szkolnictwie wyższym. W swoich dociekaniach skupiliśmy się na tematycznym obszarze decyzji wspomaganych przez systemy, pomijając teksty ściśle informatyczne (np. projektowanie reguł w systemach eksperckich [Shatnavi et al., 2014], eksperymenty z metodami filtrowania w systemach rekomendacyjnych [Sobecki 2012], ekstrakcja opisów kursów z języka html do xml w celu wykorzystania treści w akademickich systemach rekomendacji [Biletskiy, Brown, Ranganathan 2009]).

Jednym z ciekawszych tematów jest zaprojektowanie modelu systemu, w którym reguły dotyczą testowania i pomiaru zdolności ucznia w zakresie wyboru, a przede wszystkim – szans powodzenia i ukończenia określonego kierunku studiów [Aslam, Nassimullah 2011]. Badaniu wstępnemu podlegają tu: inteligencja, rozumienie, zdolność pojmowania, rozumienie pojęć matematycznych, przebieg

studiów, poziom inteligencji. Są to wymagania wyjściowe do studiowania danego kierunku. Jest to bardzo dobry przykład na to, jak można rozbudować bazę wiedzy systemu eksperckiego, aby wskazywane przez niego rekomendacje były bardziej wiarygodne i obiektywne. Podobna w tym zakresie jest propozycja rozszerzenia bazy wiedzy o kursach, zawartej w sylabusach, o informacje pochodzące od absolwentów tych kursów, wyrażane w bardziej (oceny kursów pochodzące z ankiet) bądź mniej (dyskusje *online*) sformalizowany sposób [Ishak, Lehat 2012]. Funkcje systemu pozwalające studentom nie tylko wybierać kursy z dostępnej oferty, ale też zawierające podpowiedzi z uwzględnieniem wymagań programu studiów i w ten sposób wspierające ich w takim wyborze, który pozwoli zaliczyć cały moduł i spełnić ostatecznie wszystkie warunki wymagane do tego, aby ukończyć dany stopień studiów, to kolejny przykład praktycznego wykorzystania systemów eksperckich [Parameswaran et al. 2011]. Inne to np. rozwiązania umożliwiające studentom wybór kursów na dany semestr i jednoczesne utworzenie indywidualnego harmonogramu zajęć [Engin et al. 2014].

## Postulaty użytkowników systemów akademickich

Analiza stanu badań pokazała, że problematyka oceny systemów informacyjno-wyszukiwawczych wspierających podejmowanie decyzji w zakresie edukacji na poziomie wyższym jest także przez autorów podejmowana, choć niestety w mniejszym stopniu. Poruszano następujące tematy: badania oczekiwań co do funkcji systemów [Ishak, Lehat 2012]; badania postrzegania przez studentów informacji zawartych na stronach akademickich (w szczególności postrzegania informacji dodatkowych – nierelevantnych do poszukiwań, ale potencjalnie przydatnych), służące lepszej organizacji treści w serwisach [Boatright-Horowitz, Langley, Gunnip 2009]; badania studentów dotyczące elementów formy i treści opisów kursów, jakie decydują o wyborze kursu *online* [Marshall, Greenberg, Machun 2012]; badania użyteczności serwisów [Zuhrieh, Shubair 2014; Feghali, Zbib, Hallal 2011]. Analiza powyższego piśmiennictwa pozwoliła na wyodrębnienie postulatów formułowanych przez studentów w odniesieniu do zmian, jakie powinny zostać wprowadzone w celu zwiększenia przyjazności i użyteczności systemów doradczych. Do najważniejszych rekomendacji można zaliczyć: uzupełnienie w opisach kursów informacji bieżących (np. o tym, że dany kurs nie odbędzie się w danym semestrze) [Zuhrieh, Shubair 2014]; uzupełnienie informacji o zgodnej z logiką i harmonogramem programu studiów wskazanej kolejności wyboru kursów [Zuhrieh, Shubair 2014]; obcojęzyczne wersje opisów kursów [Zuhrieh, Shubair 2014]; dodanie tutoriali wspomagających sprawną ob-

sługę systemu [Zuhrieh, Shubair, 2014]; w przypadku systemów przeznaczonych dla studentów studiujących zdalnie – włączenie dużej liczby elementów społecznościowych, sprzyjających budowaniu więzi z uczelnią i z innymi studentami [Waldner, McDaniel, Widener 2011; Waldner et al. 2012]. Rekomendacje dotyczące sposobów zwiększania funkcjonalności serwisów *e-advising*, opracowane na podstawie analizy literatury przedmiotu, uzupełniono o analogiczne refleksje, pojawiające się w zasobach Web 2.0. Do najważniejszych postulatów należy: integrowanie rozproszonych systemów w jedno narzędzie [SIS, dok. elektr.] i tworzenie kompleksowych systemów typu *life map* [Romano, dok. elektr.; Podcast, dok. elektr.; Students, dok. elektr. (c)]; wprowadzanie przyjaznych interfejsów, nawiązujących do popularnych wśród użytkowników narzędzi, takich jak Google czy Facebook [CobaltLMS, dok. elektr.]; zapewnienie możliwości formułowania zapytań w języku naturalnym [Natural..., dok. elektr.]; wprowadzanie multimedii oraz opcji o charakterze społecznościowym [CobaltLMS, dok. elektr.].

## **Modelowy zakres uczelnianego systemu wspomagającego decyzje studentów**

Modelowy system wspomagający podejmowanie decyzji powinien być spójnym i kompletnym narzędziem, umożliwiającym studentom całościowe planowanie kariery naukowej i zawodowej, adresowanym zarówno do studentów stacjonarnych, jak i tych kształcących się w trybie e-learningowym. Integracja wszystkich modułów informatycznych wspomagających obsługę poszczególnych procesów związanych ze studiowaniem na uczelni wymaga maksymalnego wykorzystania zgromadzonych zasobów, a z punktu widzenia użytkowników – tworzy jednorodne środowisko pracy. Idealny system doradczy powinien wspierać podejmowanie decyzji na różnych etapach kształcenia, pomagając w wyborze fakultetów, specjalizacji, kursów i stypendiów, a w końcu pracy zawodowej, najlepszych z punktu widzenia potrzeb konkretnych użytkowników. System taki powinien zapewniać możliwość przeprowadzania symulacji ścieżek kariery naukowej i zawodowej, tworzyć zindywidualizowane rankingi propozycji i oceniać szanse powodzenia wybranych przedsięwzięć, a także umożliwiać budowanie indywidualnych harmonogramów zajęć.

Rola systemu doradczego nie ogranicza się jedynie do gromadzenia i przetwarzania danych, ale opiera się przede wszystkim na zaawansowanych mechanizmach rekomendacyjnych systemów eksperckich, ułatwiających podejmowanie trafnych decyzji w trakcie trwania całego procesu kształcenia. Co ważne, podpowiedzi optymalnych rozwiązań, umożliwiające niezakłócony przebieg studiów,

powinny być uzupełnione o symulacje konsekwencji dokonywanych wyborów. Bazę wiedzy systemów eksperckich stanowić powinny nie tylko ustrukturalizowane opisy fakultetów, kursów, stypendiów czy ofert pracy, lecz także rekomendacje udzielane w ramach rozwiniętego systemu komunikacji wewnętrznej, na który składać się powinny sprawdzone narzędzia zapożyczone ze środowiska Web 2.0 i serwisów komercyjnych.

W warstwie informacyjnej kompleksowe systemy typu *life map* powinny gromadzić dane o studencie, w tym także wyniki testów indywidualnych predyspozycji oraz testów inteligencji, informacje o zawartości i warunkach realizacji poszczególnych kierunków, fakultetów i specjalizacji, informacje o zasadach i możliwościach optymalnego gospodarowania punktami ECTS, informacje o zasadach przyznawania różnych typów stypendiów wraz z oszacowaniem, na podstawie danych o studencie, szans na ich otrzymanie oraz – co bardzo istotne – informacje o rynku pracy i perspektywach zawodowych studenta wraz z konkretnymi propozycjami ofert zatrudnienia.

Szczególną wagę należy przywiązywać do stałego monitorowania funkcjonalności serwisów, zważywszy na ich rozmiary i liczbę zadań, których realizację mają wspomagać. Sprawnej obsłudze służyć powinny systemy podpowiedzi i tutoriali. Komunikacja z systemem doradczym powinna odbywać się w miarę możliwości w języku naturalnym, a samo oprogramowanie powinno być przyjazne i interaktywne. Powinno także udostępniać popularne wśród studentów opcje społecznościowe (opcje dodawania komentarzy, dyskusji, wymiany zasobów itp.). Istotne jest, aby system wspierał internacjonalizację studiów przez wdrażanie dobrze opracowanych wersji językowych (przynajmniej w języku angielskim).

Naszym zdaniem twórcy funkcjonujących obecnie na uczelniach polskich systemów informacyjno-wyszukiwawczych powinni w pierwszej kolejności:

- 1) w obrębie poszczególnych uczelni: dążyć do integracji rozproszonych systemów i platform w jedno kompleksowe narzędzie obsługujące wszystkie procesy akademickie – od rejestracji na studia, poprzez dokonywanie indywidualnych wyborów odnośnie do programu studiów, aż po wybór dalszej ścieżki edukacji lub kariery zawodowej,
- 2) zwiększać elastyczność systemów, tak by mogły one służyć w obrębie jednej uczelni celom i specyfice różnych dyscyplin,
- 3) zapewniać studentom spersonalizowaną pomoc i wsparcie w podejmowaniu decyzji dostosowanych do ich indywidualnych zainteresowań i predyspozycji.

W celu podsumowania rozważań na temat struktury i funkcji systemów eksperckich opracowałyśmy model prawidłowo zaprojektowanego uczelnianego systemu doradczego (tabela 1). Wyróżniłyśmy elementy struktury informacyjnej

i funkcjonalnej systemu, obszar wspomaganie decyzji oraz realizacji procesów, a następnie, w obrębie przyjętych kategorii, opisałyśmy komponenty niezbędne dla prawidłowego działania systemu eksperckiego.

Tabela 1. Modelowy zakres uczelnianego systemu doradczego

Oferta łączna dla studiów stacjonarnych i e-learningowych				
	fakultety	kursy ECTS	stypendia	praca
struktura informacyjna	<ul style="list-style-type: none"> <li>• podstawowe dane o studencie</li> <li>• test indywidualnych predyspozycji</li> <li>• test inteligencji</li> <li>• opis fakultetu</li> </ul>	<ul style="list-style-type: none"> <li>• plany studiów</li> <li>• informacja o kursach i punktach ECTS, sylabusy</li> <li>• profil studenta, wykładowcy</li> <li>• oceny, całościowe zestawienia osiągnięć edukacyjnych, także w formie graficznej</li> <li>• liczba punktów zdobytych oraz brakujących do ukończenia stopnia studiów</li> <li>• najczęściej zadawane pytania</li> <li>• aktualne informacje o dostępności kursów</li> </ul>	<ul style="list-style-type: none"> <li>• informacje o zrealizowanych kursach, umiejętnościach, dochodach i preferencjach studenta</li> <li>• szczegółowe informacje o wybieranym stypendium</li> <li>• odpowiedzi na pytania studentów</li> </ul>	<ul style="list-style-type: none"> <li>• informacje o kwalifikacjach studentów, zdobytych w czasie studiów</li> <li>• oferty pracy i wymagania pracodawców</li> </ul>
struktura funkcjonalna		<ul style="list-style-type: none"> <li>• narzędzie do tworzenia spersonalizowanych harmonogramów</li> </ul>		
	<ul style="list-style-type: none"> <li>• możliwość zadawania pytań w języku naturalnym</li> <li>• tutoriale multimedialne</li> <li>• funkcje społecznościowe</li> <li>• obce wersje językowe</li> <li>• przyjazny, nowoczesny interfejs, na wzór systemów komercyjnych, społecznościowych</li> <li>• fora dyskusyjne</li> <li>• aktualizacja informacji (zwłaszcza dotyczących oferty na dany rok akademicki)</li> </ul>			
wspomaganie decyzji	<ul style="list-style-type: none"> <li>• rekomendacje wyboru fakultetu</li> <li>• symulacja drogi zawodowej</li> </ul>	<ul style="list-style-type: none"> <li>• rekomendacja wyboru kursu akademickiego</li> <li>• komentarze do kursów, oceny</li> <li>• podgląd nazwisk innych studentów, którzy wybrali ten sam kurs</li> <li>• ranking propozycji</li> <li>• odpowiedzi kursów potrzebnych do zaliczenia danego stopnia studiów czy specjalizacji</li> </ul>	<ul style="list-style-type: none"> <li>• rekomendacja wyboru stypendium</li> <li>• ocena szans uzyskania</li> </ul>	
procesy	scenariusz rejestracji na fakultet (kierunek studiów)	scenariusze rejestracji na kursy w danym semestrze	scenariusz pozyskiwania stypendium	scenariusz ubiegania się o pierwszą pracę

Źródło: badanie własne, 2014

## Podsumowanie

Systemy informacyjno-wyszukiwawcze wspierające podejmowanie decyzji pełnią istotną rolę w kształtowaniu doświadczeń akademickich studentów, a tym samym wpływają na kształtowanie wizerunku uczelni. Prawidłowo zaprojektowany system powinien nie tylko ułatwiać automatyzację rutynowych czynności, ale przede wszystkim prowadzić, rekomendować i doradzać. Nowoczesne systemy doradcze typu *life map* powinny zapewniać wsparcie na różnych etapach kształcenia: ułatwiać wybór właściwego kierunku studiów, pomagać w bieżącym modyfikowaniu ścieżki edukacyjnej i podejmowaniu najkorzystniejszych decyzji w oparciu o indywidualne predyspozycje i zainteresowania studenta oraz wskazywać najciekawsze opcje dalszego rozwoju naukowego i zawodowego. Funkcjonującym obecnie na rynku systemom obsługi studentów najczęściej brakuje opcji umożliwiających prowadzenie procesu doradczego w sposób bardziej zindywidualizowany. Zwiększenie elastyczności systemów e-doradczych wydaje się być zatem jednym z największych wyzwań stojących przed projektantami nowoczesnych systemów informacyjno-wyszukiwawczych.

Tworzenie uczelnianych systemów eksperckich jest nie tylko praktycznym działaniem mającym na celu zwiększenie atrakcyjności oferty dydaktycznej i jej promocję w środowisku akademickim, ale wpisuje się również w szerszą koncepcję propagowania przez uczelnie postawy otwartości i przejrzystości. Prawidłowo zaprojektowane systemy informacyjno-wyszukiwawcze, zapewniające intensywny i ciągły przepływ informacji, mogą być postrzegane jako kolejny krok w kierunku zwiększania efektywności komunikacji uczelni z ich otoczeniem.

## BIBLIOGRAFIA

- Aslam, Muhammad Zaheer; Nasimullah, Abdur Rashid Khan (2011). *A Proposed Decision Support System/Expert System for Guiding Fresh Students in Selecting a Faculty in Gomal University, Pakistan*. Industrial Engineering Letters Vol. 1, No. 4, p. 33-40.
- Barker, Rachel (2009). *University considering advising system: Online advising program may simplify process* [dok. elektr.]. <http://www.gwhatchet.com/2009/02/23/university-considering-advising-system/> [odczyt: 22.08.2014].
- Biletskiy, Yevgen; Brown, Anthony J.; Ranganathan, Girish (2009). *Information Extraction From Syllabi for Academic e-Advising*. Expert Systems with Applications Vol. 36, No. 3, p. 4508-4516.
- Boatright-Horowitz, Su L.; Langley, Michelle; Gunnip, Matthew (2009). *Depth-of-Processing Effects as College Students Use Academic Advising Web Sites*. Cyberpsychology & Behavior Vol. 12, No. 3, p. 331-335.

- CobaltLMS – Student Information System built on Moodle (2014) [dok. elektr.]. <http://www.youtube.com/watch?v=IFgSSwpTWq8> [odczyt: 19.06.2014].
- CourseRank (2014) [dok. elektr.]. <https://www.courserank.com/w/home> [odczyt: 14.06.2014].
- Duch, Włodzisław (2013). *Rule discovery*. In: Dubitzky, W. et al. eds. *Encyclopedia of Systems Biology*. New York: Springer Science+Business Media [dok. elektr.]. <http://www.fizyka.umk.pl/publications/kmk/11-Rule%20Discovery.pdf> [odczyt: 19.06.2014].
- Engin, G Gökhan et al. (2014). *Rule-based Expert Systems for Supporting University Students*. *Procedia Computer Science* Vol. 31, p. 22-31.
- Fedena – Student Information System by Foradian Technologies (2013) [dok. elektr.]. [http://www.slideshare.net/ProductNation/fedena?qid=72780410-e1ed-41d2-a3ef-828ae4575aac&v=default&b=&from\\_search=4](http://www.slideshare.net/ProductNation/fedena?qid=72780410-e1ed-41d2-a3ef-828ae4575aac&v=default&b=&from_search=4) [odczyt: 19.06.2014].
- Feghali, Tony; Zbib, Imad; Hallal, Sophia (2011). *A Web-Based Decision Support Tool for Academic Advising*. *Educational Technology & Society* Vol. 14, Issue 1, p. 82-94.
- Grupe, Fritz H. (2002). *Student advisement: applying a web-based expert system to the selection of an academic major*. *College Student Journal* Vol. 36, No. 4, p. 573.
- Ishak, Isma B.; Lehat, Mohd Lezam B. (2012). *A conceptual framework of web-based academic advisory information system* [dok. elektr.]. SHUSER IEEE Symposium on Humanities, Science and Engineering Research. <http://www.scopus.com/inward/record.url?eid=2-s2.0-84867907916&partnerID=40&md5=f946afb585cdc1293bde4a07512d67bb> [odczyt: 19.06.2014].
- Kardan, Ahmad A.; Sadeghi, Hamid (2013). *A Decision Support System for Course Offering in Online Higher Education Institutes*. *International Journal of Computational Intelligence Systems* Vol. 6, No. 5, p. 928-942.
- Koutrika, Georgia et al. (2009). *Flexible recommendations for course planning*. In: Li, Jianzhong; Yu, Philip S. eds. *IEEE 25<sup>th</sup> International Conference on Data Engineering*. Los Alamitos, California, Washington, Tokyo: Institute of Electrical and Electronics Engineers, p. 1467-1470.
- Marshall, James; Greenberg, Heather; Machun, Patricia A. (2012). *How would they choose? Online student preferences for advance course information*. *Open Learning* Vol. 27, No. 3, p. 249-263.
- Modele zarządzania uczelniami w Polsce – raport końcowy (2011) [dok. elektr.]. [http://www.nauka.gov.pl/g2/oryginal/2013\\_05/bdfa51cb239812bbeac12dfa59be461c.pdf](http://www.nauka.gov.pl/g2/oryginal/2013_05/bdfa51cb239812bbeac12dfa59be461c.pdf) [odczyt: 29.06.2014].
- Natural Language Advising System – ECE – University of Florida (2014) [dok. elektr.]. <http://www.youtube.com/watch?v=pTu5oIbQKLc> [odczyt: 29.06.2014].
- Nowicki, Adam; Sitarska, Monika red. (2009). *Procesy informacyjne w zarządzaniu*. Wrocław: Wydaw. Uniwersytetu Ekonomicznego we Wrocławiu, 209 s.
- Parameswaran, Aditya; Venetis, Petros; Garcia-Molina, Hector (2011). *Recommendation Systems with Complex Constraints: A Course Recommendation Perspective*. *ACM Transactions on Information Systems* Vol. 29, No. 4, p. A:1-A:32.
- Park, Deuk Hee et al. (2012). *A literature review and classification of recommender systems research*. *Expert Systems with Applications* Vol. 39, No 11, p. 10059-10072.
- Podcast: *Using Degree Map in Student Advising at ACC* (2013) [dok. elektr.]. <http://www.youtube.com/watch?v=cO3v8mFAeXm> [odczyt: 28.06.2014].



- Romano, Joyce (2013). *LifeMap Developmental Advising System* [dok. elektr.]. [http://www.youtube.com/watch?v=gZEmeAdSG\\_c](http://www.youtube.com/watch?v=gZEmeAdSG_c) [odczyt: 29.06.2014].
- Shatnavi, Raed et al. (2014). *Building A Smart Academic Advising System Using Association Rule Mining. Presented at the The 4<sup>th</sup> International Conference on Information & Communication Systems*, At Jordan University of Science and Technology, Jordan [dok. elektr.]. <http://arxiv.org/ftp/arxiv/papers/1407/1407.1807.pdf> [odczyt: 27.06.2014].
- SIMPLE.EDU – zintegrowany system obsługi Uczelni. (2014) [dok. elektr.]. <http://bazus.pl/produkty/simpleedu/dla-kogo> [odczyt: 02.08.2014].
- SIS Project Kickoff (2013) [dok. elektr.]. <http://www.youtube.com/watch?v=RSfV9mD-2gwE> [odczyt: 26.06.2014].
- Sobecki, Janusz (2012). *Comparison of Nature Inspired Algorithms Applied in Student Courses Recommendation*. In: Pan, Jeng-Shyang; Chen, Shyi-Ming; Nguyen, Ngoc-Thanh eds. *Computational Collective Intelligence. Technologies and Applications*. Berlin Heidelberg: Springer, p. 278-287.
- Starks, Steven (2012). *Distance Advising: A Personalized Approach* [dok. elektr.]. <http://distanceadvising.com/2012/05/distance-advising-a-personalized-approach/#.U7QVREDn-I> [odczyt: 25.06.2014].
- Student Academic Advising System* (2014) [dok. elektr.]. <http://cnc.ucr.edu/sais/saas.html> [odczyt: 28.06.2014].
- Student Information System in java* (2014) [dok. elektr. (a)]. <https://www.youtube.com/watch?v=gZ6TVxPf69w> [odczyt: 28.06.2014].
- Student Information System in PHP with Encryption and Code* (2014) [dok. elektr. (b)]. <https://www.youtube.com/watch?v=EYHJGfJ6xfU> [odczyt: 28.06.2014].
- Student Success Plan (SSP) Overview of Version 1.2 and Preview* (2013) [dok. elektr. (c)]. <http://www.youtube.com/watch?v=qXeijr7jf8> [odczyt: 28.06.2014].
- System zarządzania dokumentacją dydaktyczną S4 Sylabusy Siatki Studia Studenti*. (2014) [dok. elektr.]. <http://www.invariant.pl/produkt/system-zarzadzania-dokumentacja-dydaktyczna-s4-sylabusy-siatki-studia-studenti/> [odczyt: 28.06.2014]
- Vialardi, Cesar et al. (2009). *Recommendation in Higher Education Using Data Mining Techniques*. In: Barnes, T. et al. eds. *Educational Data Mining*. Cordoba, Spain, July 1-3, 2009. Cordoba: International Working Group on Educational Data Mining. [dok. elektr.]. <http://www.educationaldatamining.org/EDM2009/uploads/proceedings/vialardi.pdf> [odczyt: 28.06.2014].
- Waldner, Leora et al. (2012). *The eQuad: A Next-Generation eAdvising Tool to Build Community and Retain Students* [dok. elektr.]. *The Mentor. Academic Advising Journal* Vol. 14. <http://dus.psu.edu/mentor/2012/10/equad-eadvising-tool-build-community-retain-students> [odczyt: 22.08.2014].
- Waldner, Leora; McDaniel, Dayna; Widener, Murray (2011). *E-Advising Excellence: The New Frontier in Faculty Advising* [dok. elektr.]. *Journal of Online Learning and Teaching* Vol. 7, No. 4. [http://jolt.merlot.org/vol7no4/waldner\\_1211.htm](http://jolt.merlot.org/vol7no4/waldner_1211.htm) [odczyt: 22.08.2014].
- Zuhrieh, Shana; Shubair, Abdul Karim Abdullah (2014). *SAAS: Creation of an e-Advising Tool to Augment Traditional Advising Methods* [dok. elektr.]. *Computer and Information Science* Vol. 7, No. 1. <http://ccsenet.org/journal/index.php/cis/article/view/32664> [odczyt: 22.08.2014].

Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich  
00-335 Warszawa, ul. Konopczyńskiego 5/7 tel. (22) 827-52-96  
www.sbp.pl; wydawnictwo@sbp.pl, biuro@sbp.pl  
Warszawa 2014. Wyd. I. Ark. wyd. 29,0. Ark. druk. 29,0  
Łamanie: Studio Kałamarz (http://www.studiokalamarz.pl)  
Druk i oprawa: Fabryka Druku sp. z o.o.  
ul. Zgrupowania AK „Kampinos” 6, 01-943 Warszawa  
fabrykadruku@fabrykadruku.pl



Monografia odzwierciedla zainteresowania naukowców teoretyków i praktyków pracujących w ośrodkach w całej Polsce, którzy w swojej działalności badawczej lub wdrożeniowej zajmują się różnymi aspektami komunikacji naukowej w środowisku cyfrowym. Większość tekstów powstała w Instytucie Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego. W zamyśle książka nie jest pełną panoramą wiedzy o wirtualnym obiegu przekazów naukowych, lecz uchwyconym w kadrze obrazem stanu badań i refleksji teoretycznej na ten temat, tak jak przedstawia się on w roku 2014. Problematyka tomu obejmuje aktualny kontekst prawny i modelowanie procesów komunikacji naukowej w sieci, tworzenie i zarządzanie kolekcjami dokumentów cyfrowych w ramach Zielonej Drogi dostępu do publikacji naukowych, metody optymalizacji katalogów bibliotecznych i rejestracji bibliograficznej dorobku naukowego pracowników uczelni oraz różnorodne zagadnienia związane z badaniem potrzeb użytkowników informacji cyfrowej.

Seria wydawana przez Wydawnictwo  
STOWARZYSZENIA BIBLIOTEKARZY POLSKICH  
we współpracy  
Z INSTYTUTEM INFORMACJI NAUKOWEJ  
I STUDIÓW BIBLIOLOGICZNYCH  
UNIwersytetu WARSZAWSKIEGO



Cena 45 zł