

# ZIN

Studia informacyjne  
Information studies

VOL. 56 2018 NO. 2(112)

p-ISSN 0324-8194

e-ISSN 2392-2648



**WDiB UW**



STOWARZYSZENIE  
BIBLIOTEKARZY  
POLSKICH

## REDAKCJA | EDITORIALS

*Redaktor naczelny | Editor in Chief:* Barbara Sosińska-Kalata  
*Redaktor tematyczny, z-ca redaktora naczelnego | Co-editor in Chief:* Marcin Roszkowski  
*Sekretarz redakcji | Secretary:* Zuzanna Wiorogórska  
*Redaktor językowy – język polski | Philological editor (Polish):* Zuzanna Wiorogórska  
*Redaktor językowy – język angielski | Philological editor (English):* Agnieszka Kasprzyk  
*Redaktor statystyczny | Statistical editor:* Anna Grzecznowska  
*Redaktor techniczny i korekta | Technical editor and proofreading:* Anna Lis

<b>RADA REDAKCYJNA</b>	<b>EDITORIAL BOARD</b>	<b>RADA KONSULTACYJNA</b>	<b>CONSULTING BOARD</b>
Wiesław Babik ( <i>Uniwersytet Jagielloński</i> )		Hanna Batorowska, UP	
Peter A. Bruck ( <i>Research Studio, Austria</i> )		Sabina Cisek, UJ	
Laurence Favier ( <i>Université de Lille, France</i> )		Piotr Gawrysiak, PW	
Birger Hjørland ( <i>University of Copenhagen, Denmark</i> )		Ewa Głowacka, UMK	
Michèle Hudon ( <i>University of Montreal, Canada</i> )		Małgorzata Górska, UW	
Bruno Jacobfeuerborn ( <i>Deutsche Telecom, Germany</i> )		Mirosław Górny, UAM	
Tibor Koltay ( <i>Eszterházy Károly Egyetem University, Hungary</i> )		Elżbieta Gondek, UŚ	
Kazimierz Krzysztofek ( <i>Wyższa Szkoła Psychologii Społecznej</i> )		Artur Jazdon, UAM	
Dariusz Kuźmina (Przewodniczący   Chair) ( <i>Uniwersytet Warszawski</i> )		Małgorzata Kisilowska, UW	
Elena Maceviciute ( <i>University of Borås, Sweden</i> )		Katarzyna Materska, UKSW	
Krystyna Matusiak ( <i>University of Denver, USA</i> )		Marek Nahotko, UJ	
Mieczysław Muraszek ( <i>Politechnika Warszawska</i> )		Piotr Nowak, UAM	
Widad Mustafa El Hadi ( <i>Université de Lille, France</i> )		Zbigniew Osiński, UMCS	
David Nicholas ( <i>CIBER, UK; Tomsk State University, Russia</i> )		Diana Pietruch-Reizes, UJ	
Henryk Rybiński ( <i>Politechnika Warszawska</i> )		Maria Próchnicka, UJ	
Barbara Stefaniak (em., <i>Uniwersytet Śląski</i> )		Arkadiusz Pulikowski, UŚ	
Elżbieta Stefańczyk ( <i>SBP</i> )		Remigiusz Sapa, UJ	
Tomasz Szapiro ( <i>Szkoła Główna Handlowa w Warszawie</i> )		Jadwiga Sadowska, UwB	
Joseph T. Tennis ( <i>University of Washington, USA</i> )		Marta Skalska-Zlat, UW	
Jadwiga Woźniak-Kasperek ( <i>Uniwersytet Warszawski</i> )		Stanisław Skórka, UP	
Elżbieta B. Zybert ( <i>Uniwersytet Warszawski</i> )		Marzena Świągół, UWM	
		Jacek Tomaszczyk, UŚ	

Wersja papierowa jest wersją pierwotną czasopisma.  
Zawartość czasopisma jest indeksowana w CEJSH, CSA, LISTA,  
Knowledge Organization Literature, PBN, PBB.

The paper version is the original version of the journal.  
The content of the journal is indexed in CEJSH, CSA, LISTA,  
Knowledge Organization Literature, PBN, PBB.

Pełne teksty roczników 1962–2017 dostępne są w Archiwum SBP: [www.sbp.pl/archiwumcyfrowe](http://www.sbp.pl/archiwumcyfrowe)  
oraz (do 2015 r.) w Medioteka WDIB UW: <http://medioteka.uw.edu.pl/dlibra/>

**ZAGADNIENIA  
INFORMACJI  
NAUKOWEJ**  
Studia informacyjne

**ISSUES IN  
INFORMATION  
SCIENCE**  
Information Studies

VOL. 56 2018 NO. 2(112)  
p-ISSN 0324-8194  
e-ISSN 2392-2648



**WDIiB UW**



STOWARZYSZENIE  
BIBLIOTEKARZY  
POLSKICH



Warszawa 2018

## ISSUES IN INFORMATION SCIENCE – INFORMATION STUDIES

The core purpose of this journal is to provide a forum for the dissemination of scientific papers and research results in the field of information science and other disciplines which analyze social and technological aspects of various information-related activities performed by contemporary communities. Moreover, the journal is to disseminate critical reviews and summaries of new publications in the field of information science and reports from important conferences discussing contemporary information problems.

In addition to that, after fifty years *Issues in Information Science* (ZIN) becomes bilingual to reach foreign readers – papers are available either in Polish or English.

The reconstructed journal is published with slightly extended title *Issues in Information Science – Information Studies* (ZIN – *Information Studies*) – the subtitle emphasizes the interdisciplinary nature of its subject profile covering a broad spectrum of issues studied by various academic disciplines and professional activity domains related to access to resources of recorded information and knowledge and the use of these resources by contemporary man and society. Other subjects to be covered by ZIN involve: 1) theoretical ponderings on the practice of information-related activities performed by various communities, 2) the results of research on the conditions influencing those activities and ways of improving methods and tools employed for the activities in question, 3) the methodology of information science research, information science history and education concerning the information science. The subject profile of semiannual ZIN – *Information Studies* covers, among else, the issues of:

- information science in relation to library science, archival science, museology and other disciplines researching preservation and access to scientific and cultural heritage
- information and knowledge management
- traditional and online scholarly communication
- information and knowledge organization
- metadata theory and practice
- Web 2.0
- Semantic Web
- information architecture
- information websites usability
- digital humanities
- human-computer interaction
- natural language processing
- information retrieval
- use of information and behavior of the information users
- social response to modern information technologies
- digital humanities
- information and digital skills
- information policy
- information ethics.

ZIN – *Information Studies* is addressed to: 1) information science teachers and lecturers, researchers and students, 2) practitioners of information-related activities who analyze methods and tools used to implement those activities in various domains and organizational environments, 3) politicians and donors related to information activities in various domains. The journal content may also be of some interest to teachers, students and researchers in other disciplines of science which deal with various aspects of information existence and use in the contemporary world.

ZIN – *Information Studies* is included in 'B' list of journals scored by Polish Ministry of Science and Higher Education and indexed by: Central European Journal in Social Sciences and Humanities (CEJSH), Cambridge Scientific Abstracts (CSA), Library and Information Science and Technology Abstracts (LISTA), Polish Bibliography of Book Studies (PBB), Knowledge Organization Literature and Polish Scholarly Bibliography (PBN).

## ZAGADNIENIA INFORMACJI NAUKOWEJ – STUDIA INFORMACYJNE

Głównym celem niniejszego czasopisma jest zapewnienie forum dla rozpowszechniania artykułów naukowych i wyników badań z zakresu nauki o informacji (informatologii) oraz innych dyscyplin, w których podejmowane są analizy społecznych i technologicznych aspektów działalności informacyjnej prowadzonej w różnych sferach współczesnego życia społecznego. Czasopismo służyć ma również rozpowszechnianiu krytycznych recenzji i omówień publikacji z tego zakresu oraz problemowych sprawozdań z ważnych konferencji poświęconych współczesnym problemom informacyjnym.

W minionym pięćdziesięcioleciu *Zagadnienia Informacji Naukowej* (ZIN) były czasopismem publikującym teksty wyłącznie po polsku, a zatem adresowanym tylko do czytelnika polskiego.

W nowej formie czasopismo adresowane jest zarówno do czytelnika polskiego jak i zagranicznego, publikujemy artykuły zarówno w języku polskim jak i angielskim. Obecnie czasopismo ukazuje się pod rozszerzonym tytułem: ZIN – *Studia Informacyjne*. Dodany podtytuł podkreśla interdyscyplinarny charakter jego profilu tematycznego, który obejmuje szeroki zakres problemów podejmowanych przez dyscypliny akademickie i dziedziny działalności zawodowej związane z zapewnianiem dostępu do utrwalonych zasobów informacji i wiedzy oraz ich wykorzystywaniem przez współczesnego człowieka i współczesne społeczeństwo. Czasopismo publikuje też artykuły prezentujące teoretyczną refleksję o praktycznej działalności informacyjnej prowadzonej w różnych dziedzinach i obszarach życia społecznego, a także wyniki badań służących poznaniu różnych uwarunkowań tej działalności oraz doskonaleniu jej metod i narzędzi. Na łamach ZIN publikowane są także artykuły poświęcone metodologii badań informatologicznych, historii nauki o informacji oraz edukacji w zakresie nauki o informacji. Profil tematyczny półrocznika ZIN – *Studia Informacyjne* obejmuje m.in. problematykę:

- nauki o informacji w powiązaniu z bibliotekoznawstwem, archiwistyką, muzeologią innymi dyscyplinami zajmującymi się problematyką zachowania i zapewnienia dostępu do dziedzictwa nauki i kultury
- zarządzania informacją i wiedzą
- komunikacji naukowej i cyfrowej komunikacji naukowej
- organizacji informacji i wiedzy
- teorii i praktyki metadanych
- zagadnień Web 2.0
- zagadnień Sieci Semantycznej
- architektury informacji
- projektowania użytecznych serwisów informacyjnych
- humanistyki cyfrowej
- interakcji człowiek – komputer
- przetwarzania języka naturalnego
- wyszukiwania informacji
- wykorzystywania informacji i zachowań informacyjnych użytkowników
- społecznej recepcji nowoczesnych technologii informacyjnych
- kompetencji informacyjnych i cyfrowych
- polityki informacyjnej
- etyki informacyjnej.

*Zagadnienia Informacji Naukowej – Studia Informacyjne* adresowane są do wykładowców, badaczy i studentów nauki o informacji, a także praktyków działalności informacyjnej, krytycznie analizujących metody i narzędzia jej realizacji w różnych środowiskach dziedzinowych i organizacyjnych oraz polityków i donatorów działalności informacyjnej w różnych dziedzinach. Lektura czasopisma może też zainteresować wykładowców, studentów i badaczy innych dyscyplin, które zajmują się równymi aspektami funkcjonowania informacji we współczesnym świecie.

*Zagadnienia Informacji Naukowej* znajdują się na liście B czasopism punktowanych Ministerstwa Nauki i Szkolnictwa Wyższego. Czasopismo jest indeksowane w bazach: Central European Journal in Social Sciences and Humanities (CEJSH), Cambridge Scientific Abstracts (CSA), Library and Information Science and Technology Abstracts (LISTA), Polska Bibliografia Bibliologiczna (PBB), Knowledge Organization Literature, Polska Bibliografia Naukowa (PBN).



# Spis treści | Contents

## ROZPRAWY. BADANIA. MATERIAŁY | THESES. RESEARCH. MATERIALS

Barbara Sosińska-Kalata <i>Big data (dane masowe) w nauce o informacji [Big Data (Massive Data) in Information Science]</i> .....	7
Kamil Stępień, Jadwiga Woźniak-Kasperek <i>Fotografia w bibliotece i w sieci: dwa podejścia do reprezentacji i wyszukiwania [Photography in the Library and on the Web: Two Perspectives on Representation and Retrieval]</i> .....	36
Paulina Krzanowska, Marek Nahotko <i>Gatunek information scraps w indywidualnym/grupowym zarządzaniu informacją [Information Scraps Genre in Individual/Group Information Management]</i> .....	52
Anna M. Kamińska <i>O rozwoju graficznych języków komunikacji. Przykład wykorzystania UML w obszarze bibliologii i informatologii [On the Development of Graphical Languages of Communication. An Example of the Use of UML in the Field of Book and Information Sciences]</i> .....	74
Anna Matysek <i>Architektura informacji w piśmiennictwie zagranicznym 1982–2018 [Information Architecture in Foreign Literature 1982–2018]</i> .....	88

## RECENZJE I OMÓWIENIA | REVIEWS

<i>Renata Frączek (2017). Upowszechnianie wyników badań naukowych w międzynarodowych bazach danych. Analiza bibliometryczna na przykładzie nauk technicznych ze szczególnym uwzględnieniem elektrotechniki [The Dissemination of Scientific Research Results in International Databases. A Bibliometric Analysis Based on the Example of Technical Sciences with Special Regard to Electrotechnics] (Katowice: Wydaw. Uniwersytetu Śląskiego, 2017) (Małgorzata Kowalska-Chrzanowska) .....</i>	112
<i>Wśród zagranicznych książek [Foreign Publications] (Jacek Wojciechowski) .....</i>	115
<i>Przegląd polskich nowości wydawniczych [New Polish Publications] (Anna Stanis) ..</i>	124

KRONIKA | CHRONICLE

*ECIL – „Europejska konferencja edukacji informacyjnej” (Oulu, Finlandia, 24–27 września 2018 r.) [ECIL – European Conference on Information Literacy (Oulu, Finland, September 24–27, 2018)]* (Zuzanna Wiorogórska) ..... 129

WSKAZÓWKI DLA AUTORÓW | GUIDELINES FOR AUTHORS ..... 133



# Big data (dane masowe) w nauce o informacji

Barbara Sosińska-Kalata

*Katedra Informatologii, Wydział Dziennikarstwa, Informacji i Bibliologii  
Uniwersytet Warszawski*

---

## Abstrakt

**Cel/Teza:** Celem artykułu jest omówienie głównych cech zjawiska określanego mianem big data, jego znaczenia dla problematyki badawczej nauki o informacji oraz próba wstępnej oceny stopnia zainteresowania nim badaczy tej dyscypliny.

**Koncepcja/Metody badań:** Krytyczna analiza piśmiennictwa przedmiotu wykorzystana została do omówienia istoty zjawiska big data oraz związanych z nim zmian w modelu badań naukowych, który w coraz większym zakresie znajduje zastosowanie w różnych dziedzinach współczesnej nauki. Rosnącą popularność badań big data w nauce zilustrowano wynikami analizy bibliometrycznej piśmiennictwa zarejestrowanego w interdyscyplinarnej bazie Scopus. Ocenę stopnia zainteresowania problematyką big data w nauce o informacji oparto na bibliometrycznej analizie piśmiennictwa indeksowanego w dziedzinowej bazie EBSCO – Library and Information Science and Technology Abstracts (LISTA).

**Wyniki i wnioski:** Zagadnienie big data można traktować jako kolejną fazę rozwoju technologii komputerowej i jej zastosowań w różnych dziedzinach nauki i praktyki. W środowisku wielkich zasobów danych zapisanych w cyfrowym formacie, technologie big data zapewniają wgląd w wiedzę, której nie można byłoby wydobyć tradycyjnymi metodami wyszukiwania informacji. W tym sensie technologie te wspierają procesy transferu wiedzy między ludźmi, które stanowią główny przedmiot zainteresowań nauki o informacji. Analiza piśmiennictwa indeksowanego w bazie LISTA potwierdziła, że rozwój technologii big data i jej zastosowań stanowi istotne wyzwanie dla nauki o informacji, którym zainteresowanie badaczy systematycznie rośnie, jakkolwiek nie jest ono jeszcze w tej dyscyplinie bardzo duże. Analiza tematyki tego piśmiennictwa potwierdziła też, że problematyka big data łączy się z kluczowymi obszarami badań nauki o informacji. Badania dotyczące big data najczęściej prezentowane są na łamach czasopism specjalizujących się w ilościowych badaniach informacji (bibliometrii, naukometrii, altmetrii), informatyce medycznej, problematyce systemów informacyjnych i wyszukiwania informacji oraz w zarządzaniu informacją. W czasopismach o szerokim profilu tematycznym obejmującym całe pole badawcze nauki o informacji publikacje na temat big data dotychczas ukazywały się rzadko. Autorami największej liczby artykułów dotyczących tej problematyki są badacze związani z ośrodkami naukowymi w Stanach Zjednoczonych, w Wielkiej Brytanii i w Chinach. Piśmiennictwo dotyczące badań big data w nauce o informacji charakteryzuje duża różnorodność podejmowanej tematyki szczegółowej. Dominuje tematyka należąca do obszaru nauk komputerowych oraz mediów społecznych, ale do zagadnień często omawianych należą też metadane, zarządzanie i dzielenie się wiedzą, biblioteki cyfrowe, bibliometria oraz kwestie związane z informatyką medyczną i ochroną zdrowia.

**Ograniczenia badań:** Omówione badanie ma charakter sondażowy i przeprowadzone zostało na indeksowanym w bazie LISTA piśmiennictwie, w którego opisie tematycznym użyty został termin „big data”. Piśmiennictwo prezentujące problematykę związaną z badaniem wielkich zbiorów danych, w którego indeksowaniu nie użyto tego terminu, nie zostało zatem uwzględnione w badaniu. Ponadto polityka indeksowania bazy LISTA, w szczególności względnie mała reprezentacja czasopism wydawanych w innych językach niż angielski wśród indeksowanych w niej źródeł, może ograniczać reprezentatywność uzyskanych wyników dla badań dotyczących big data, związanych z problematyką nauki o informacji, w skali globalnej.

**Oryginalność/Wartość poznawcza:** Zgodnie z wiedzą autorki, artykuł jest pierwszą próbą oceny stopnia zainteresowania problematyką big data w nauce o informacji.

**Słowa kluczowe**

Badanie bibliometryczne. Big data. Dane masowe. Nauka o informacji. Problematyka badawcza.

Otrzymany: 11 lutego 2018. Zrecenzowany: 23 lutego 2018. Zaakceptowany: 5 marca 2018.

## 1. Wprowadzenie

Dane masowe, obecnie najczęściej określane angielskim terminem „big data”, oznaczają zbiory danych, które przyrastają w sposób nieograniczony, dla których pamięć musi być rezerwowana dynamicznie i których nie da się przetwarzać metodami tradycyjnymi. Dane masowe występują zatem w nadzwyczaj dużych ilościach, liczonych w petabajtach (PB), zettabajtach (ZB), czy nawet jottabajtach<sup>1</sup>, a to implikuje konieczność nowego podejścia do ich gromadzenia, magazynowania, przetwarzania i transmisji. Według raportu firmy Intel, o produkcji big data można mówić wtedy, gdy organizacja generuje medianę 300 terabajtów danych tygodniowo (Intel, 2012). Należy dodać, że szacunki te odnoszą się przede wszystkim do wielkich zasobów danych wykorzystywanych w analizach wspierających procesy decyzyjne we współczesnym marketingu i zarządzaniu biznesem. Wielkie ilości danych generuje też nowoczesna aparatura naukowa i technologicznie zaawansowane narzędzia, np. jedno doświadczenie przeprowadzone w CERN-ie przy użyciu Wielkiego Zderzacza Hadronów generuje około 40 terabajtów (TB) danych w ciągu 30 minut, a jeden przelot odrzutowca dostarcza około 10 TB danych (Jacobfeuerborn, 2013). Wielkie zasoby danych tworzą ludzie, publikując w przestrzeni cyfrowej różnego typu komunikaty i pozostawiając ślady swojej aktywności w sieciach komputerowych, np. w postaci logów do różnych stron internetowych czy też kwerend kierowanych do wyszukiwarek. W 2012 r. na świecie opublikowanych zostało ponad 1,57 mln artykułów naukowych, co oznacza, że na każdą minutę tego roku przypadały 3 nowe artykuły (Ferstein, 2012). Już w 2009 r. Facebook informował, że spółka dysponuje 1 PB danych, natomiast w 2016 r. w posiadaniu Google było 15 eksabajtów<sup>2</sup> danych (Patgiri & Ahmed, 2016). Obecnie Google przetwarza 40 tys. pytań w każdej sekundzie, tj. 3,5 miliarda pytań dziennie<sup>3</sup>. Codziennie ogromne zasoby danych generuje sieć powiązanych z sobą milionów inteligentnych urządzeń współtworzących dynamicznie rozwijający się Internet Rzeczy (ang. *Internet of Things*, IoT).

Przytoczone przykłady dają wyobrażenie o wielkości zasobów określanych mianem „big data”. W różnych dziedzinach skala wielkości zasobów przetwarzanych metodami big data może być inna, zawsze jednak mówić będziemy o wielkościach znacząco większych niż te, które stanowiły podstawę analiz prowadzonych metodami tradycyjnymi. Zasoby big data mogą być zarówno ustrukturyzowane (jak np. dane o transakcjach biznesowych, przechowywane w relacyjnych bazach danych, czy też dane gromadzone w wielkich bazach bibliograficznych, patentowych, rejestrów medycznych itp.), częściowo ustrukturyzowane

<sup>1</sup> Petabajt (PB) jest jednostką używaną do oznaczania biliarde bajtów, w których mierzona jest pojemność największych pamięci masowych; 1 PB jest równy 1015 bajtów, tj. ok. 1020 TB, zettabajt (ZB) oznacza tryliard bajtów, tj. 1021 bajtów, a jottabajt (YT) to kwadrylion bajtów, czyli 1024 bajtów.

<sup>2</sup> Eksabajt (EB) jest równy 1018 bajtów, tj. ok. 1 040 816 TB.

<sup>3</sup> Dane za Google Search Statistics (<http://www.internetlivestats.com/google-search-statistics/>).

(np. pełne teksty opatrzone tagami), jak i nieustrukturyzowane (np. wiadomości e-mail czy komentarze generowane w mediach społecznościowych).

Dla big data wielkość zasobów jest kluczowa, jednak wielu badaczy przekonuje, że nie tylko o ich wielkość chodzi. Na przykład, za cechę równie istotną dla zasobów big data uznaje się ich heterogeniczność, wielką różnorodność formatów i sposobów reprezentacji danych. Pogląd ten znalazł odzwierciedlenie w jednej z najpopularniejszych do dziś definicji „big data”, którą w 2001 r. sformułował Doug Laney, odwołując się do trzech atrybutów, które uznał za konstytutywne dla tego typu zasobów i których nazwy zaczynają się na literę „v”, tj. do tzw. formuły „3Vs”: *volume* (wielkość), *velocity* (szybkość), *variety* (różnorodność); (Laney, 2001). Jednak dwanaście lat później, jak twierdzi Alon Friedman, w 2013 r. Laney zrewidował swoją głośną definicję, stwierdzając, iż:

Big Data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization (Friedman, 2017, 135).

Zatem big data charakteryzuje przede wszystkim wielkość zasobów, której towarzyszy albo szybki przyrost, albo duża różnorodność, albo obie te własności równocześnie. Nadzwyczajna wielkość zasobów big data w powiązaniu z szybkością ich narastania i różnorodnością formatów i reprezentacji danych czynią te zasoby zbyt złożonymi, aby można było je magazynować i przetwarzać tradycyjnymi metodami.

W miarę wzrostu zainteresowania problematyką big data kolejni badacze proponowali doprecyzowanie definicji „3Vs”, wskazując inne atrybuty szybko rosnących zasobów danych, które uznawano za specyficzne dla nich i których nazwy, wzorem definicji Laneya, zaczynają się od litery „v”: *variability* (zmiennosc), *veracity* (wiarygodność), *value* (wartość), *validity* (ważność), *volatility* (ulotność), *virtual* (faktyczność), *visualization/visibility* (wizualizacja/widoczność), *vitality* (witalność), *vaccum* (próżnia) (por. Patgiri & Ahmed, 2016). Dobieranie kolejnych v-atributów charakteryzujących big data stało się popularne w piśmiennictwie poświęconym temu zjawisku, jednak za faktycznie najistotniejsze jego cechy uznać należy przede wszystkim *volume* (wielkość) i *complexity* (złożoność).

Przetwarzanie zasobów big data wymaga stosowania nowej technologii i nowych metod analitycznych. Podstawą nowego podejścia do przetwarzania wielkich zasobów danych jest założenie, że jego głównym celem jest wydobywanie z tych zasobów ukrytej w nich nowej wiedzy przez stosowanie metod i technik określanych ogólnie analityką danych (ang. *data analytics*). Według Google Trends, *data analytics* jest terminem wyszukiwawczym najsilniej powiązanim z tematem big data. Przeprowadzona przez Jonathana Stuarta Warda i Adama Barkera analiza różnych definicji big data wykazała, że intensywnie rozwijane narzędzia i metody analizy danych masowych stanowią trzeci, konstytutywny składnik zjawiska big data (Ward & Barker, 2013). Do podobnych wniosków doszli też Andrea De Mauro, Marco Greco i Michele Grimaldi, analizując tematykę piśmiennictwa dotyczącego big data, które zarejestrowano w interdyscyplinarnej bazie Scopus (De Mauro et al., 2016). Metody eksploracji i analizy danych masowych ukierunkowane są na szukanie relacji i wzorów powiązań między danymi oraz szacowanie prawdopodobieństwa ich występowania, co pozwala następnie przewidywać trendy oraz rekomendować działania, decyzje i innowacje optymalne w określonych sytuacjach, w odniesieniu do potrzeb określonej grupy klientów i określonych celów. Technologie big data otworzyły nowe możliwości

zdobywania wiedzy niezwykle użytecznej dla nowoczesnego biznesu, gdzie od kilku lat są intensywnie wykorzystywane. Technologie te dostarczyły także nauce nowych, potężnych narzędzi badawczych, zapewniających znacznie bardziej niż dotąd szczegółowy wgląd w rozmaite zjawiska i procesy naturalne, techniczne i społeczne, a także w słabo dotąd poznane własności ludzkiej twórczości.

## 2. Big data i analityka danych masowych w nauce, w naukach społecznych i humanistyce

Informatyzacja aparatury badawczej, cyfryzacja informacji i wiedzy o człowieku i otaczającym go świecie, która obejmuje coraz większe obszary naszego życia, oraz wielki wzrost mocy obliczeniowej komputerów zmieniają sposób uprawiania nauki w niemal wszystkich już dziedzinach, oferując nowe podejście, nowe narzędzia i nowe metody poznawania świata i rozwiązywania problemów. Zjawisko to, zapoczątkowane w latach 90. XX w. intensywnym rozwojem technologii sieciowych, technik *data-mining* i *cloud computing* oraz ich wykorzystaniem w genetyce i astronomii, zostało określone mianem Czwartego Paradygmatu w ewolucji nauki (Hey et al., 2009; Jacobfeuerborn, 2013). Według Jima Graya, badacza z laboratoriów Microsoftu, który dziesięć lat temu określenie to zaproponował, trzy pierwsze podstawowe paradygmaty w rozwoju nauki stanowiły najdawniejszy paradygmat empiryczny, oparty na opisie zjawisk naturalnych, zapoczątkowany przez prace Galileusza, Johanna Keplera czy Tycho de Brahe'a paradygmat teoretyczny, oparty na modelowaniu zjawisk teoretycznej generalizacji oraz rozwijany w ostatnich kilkudziesięciu latach paradygmat komputacyjny, oparty na komputerowej symulacji złożonych zjawisk. Czwarty paradygmat, charakteryzujący e-naukę, oparty jest na intensywnym wykorzystywaniu danych cyfrowych w badaniach naukowych. Dane pozyskiwane są przez aparaturę badawczą lub generowane przez symulatory, a następnie przetwarzane przez oprogramowanie komputerowe; informacja i wiedza przechowywane są w pamięciach komputerowych; badacz analizuje zawartość baz danych czy plików komputerowych, korzystając z metod statystycznych i narzędzi zarządzania danymi. W ten sposób cały cykl badawczy oparty zostaje na cyfrowych danych reprezentujących badany świat oraz procesach ich komputerowego przetwarzania.

Viktor Mayer-Schönberger i Kenneth Cukier piszą, że technologie big data zmieniają nasze myślenie, pracę i życie (Mayer-Schönberger & Cukier, 2014). Wykorzystując metody eksploracji danych i wyodrębniania wzorów powiązań w wielkiej skali, zapewnianej przez dane masowe, przed nauką otworzyły one nowy sposób poznawania świata, który opiera się na zastąpieniu modelu wyjaśniania badanych zjawisk i procesów przez ustalanie ich przyczyn, modelem ustalania korelacji między szczegółowymi danymi opisującymi (reprezentującymi) te zjawiska i procesy oraz ich kontekst (środowisko, sytuacje) w przestrzeni cyfrowej. Przyjmuje się założenie, że ustalenie korelacji między elementami wielkich zbiorów danych wystarcza do uzyskania nowej wiedzy, poznania nieznanych dotąd własności, procesów i aspektów naszej rzeczywistości. Korelacje te nie muszą wyjaśniać, dlaczego coś się dzieje, ale informują, że to się dzieje, pozwalając przewidywać kierunki i sposoby rozwoju badanych zjawisk. Jak piszą Mayer-Schönberger i Cukier:

Nie zawsze musimy znać przyczyny jakiegos zjawiska, możemy po prostu pozwolić danym przemawiać za siebie (Mayer-Schönberger & Cukier, 2014, 30).

Zwiększenie ilości danych, na których przeprowadzane są analizy, umożliwia odkrywanie ukrytych powiązań i modelowanie ich powtarzalnych schematów, których dostrzeżenie było niemożliwe przy mniejszej ilości informacji. Technologie big data w niektórych przypadkach umożliwiają poddanie analizie wszystkich danych, które dotyczą badanych zjawisk, a nie tylko np. ich próby losowej, którą operuje się w tradycyjnych badaniach ilościowych, zakładając (starając się zapewnić) jej reprezentatywność. Znacznie większa szczegółowość i kompletność danych stanowiących podstawę poznania naukowego równocześnie umożliwia jego wielką skalowalność – od poznawania ogólnego kierunku do poznawania najbardziej szczegółowych detali. Charakterystyczną cechą nowego podejścia do poznawania świata jest też rezygnacja z zachowania dużej dokładności dokonywanych pomiarów, którą w wielu przypadkach może zastąpić lepsze zrozumienie badanych zjawisk dzięki wykorzystaniu wielkiej liczby różnorodnych danych opisujących te zjawiska.

Niektórzy badacze uważają, że model badania oparty na wykorzystaniu big data zapewnia też większy obiektywizm poznania niż badania oparte na formułowaniu założeń i hipotez, budowaniu teorii oraz ich weryfikacji na podstawie relatywnie małych prób badawczych. Wielką dyskusję w nauce wywołał Chris Anderson, redaktor naczelny magazynu *Wired*, ogłaszając nawet „koniec teorii” i twierdząc, że w epoce big data formułowanie spekulatywnych teorii wyjaśniających, czym są analizowane dane, jest niepotrzebne, wystarczy bowiem poznanie korelacji, w które te dane wchodzi (Anderson, 2008). Nie ulega wątpliwości, że identyfikacja korelacji zachodzących w wielkich zbiorach danych pozwala na stawianie nowych pytań, otwierając nowe możliwości poznawcze. Niemniej jednak, zarówno teza o obiektywizmie badań opartych na analizie wielkich zasobów danych, jak i teza o końcu teorii nie są przekonujące. W szczególności trzeba podkreślić, że chociaż nadzwyczajna wielkość analizowanych zbiorów danych zmniejsza ryzyko błędu związanego z pominięciem danych istotnych, to jednak wiarygodność wyników analiz metodami big data zawsze zależy od jakości danych poddawanych analizie, a ta z kolei – od metod pozyskiwania danych, wykorzystanych źródeł, a także przygotowania danych do analizy. Poglądowi o obiektywizmie badań opartych na zasobach big data i ilościowych metodach ich analizy przeczy też fakt, iż zakres gromadzenia danych (np. przez określone urządzenia pomiarowe) i interpretacja wyników statystycznych analiz wielkich zbiorów danych zależne są od instrumentarium i celu prowadzonych badań.

Niezależnie od kontrowersji dotyczących oceny stopnia wpływu zjawiska big data na transformację badań naukowych nie podlega dyskusji to, że wpływ ten jest coraz silniejszy i dotyczy coraz większej liczby dyscyplin. Kluczową rolę odgrywa tu zmiana nastawienia do tego, jak dane mogą być wykorzystywane, która nastąpiła w ostatnich kilkunastu latach. Po pierwsze, wraz z przekonaniem, że przydatność danych nie kończy się z chwilą osiągnięcia celu, dla którego były one gromadzone (np. do wykonania pewnego działania), upowszechniła się archiwizacja danych; wtórne wykorzystanie danych staje się źródłem inspiracji i innowacji. Po drugie, dostępność technologii cyfrowych i coraz szerszy zakres ich zastosowania implikują tzw. danetyzację rzeczywistości, tj. zbieranie danych o wszystkim, w tym o kwestiach, o których dotąd nie myślano jako o źródłach danych (np. naprężenia w konstrukcji mostu, wibracje silnika, miejsce przebywania konkretnej osoby, logi do serwisów internetowych, słowa wpisywane przez użytkowników do wyszukiwarek internetowych) i przetwarzanie ich w kwantyfikowalny format.

Wielkie możliwości wykorzystania nowego modelu badań naukowych przed naukami społecznymi otworzyły w szczególności media społecznościowe oraz masowe generowanie,

gromadzenie i przetwarzanie danych o ludzkich zachowaniach zbiorowych i indywidualnych. Big data staje się coraz powszechniejszym mechanizmem, którym ludzie posługują się, nadając sens otaczającej ich rzeczywistości i równocześnie dostarczając niezmiernie bogaty materiał badawczy (Klous & Wielaard, 2016). Z kolei masowa digitalizacja dziedzictwa kulturowego i nowe instrumentarium badań ludzkiej twórczości dały podstawę do coraz szerszego zastosowania metod big data w humanistyce.

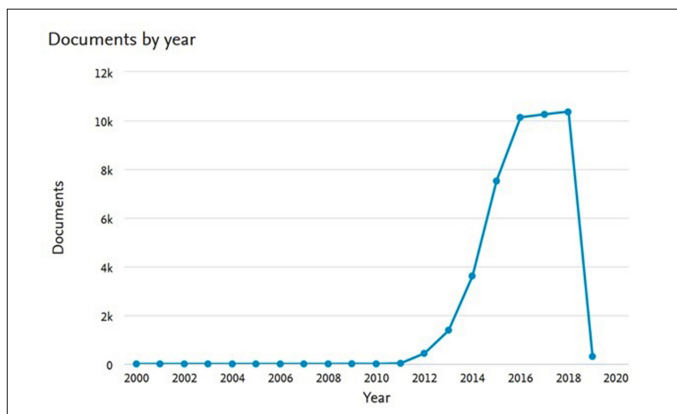
Szybko rosnącą popularność wykorzystywania danych masowych w badaniach naukowych oraz dystrybucję tego typu badań w różnych dziedzinach nauki ilustrują wyniki prostego badania bibliometrycznego, przeprowadzonego na interdyscyplinarnej bazie Scopus. Baza ta, utworzona przez wydawnictwo Elsevier w 2004 r., rejestruje piśmiennictwo ze wszystkich obszarów nauki od 1970 r., w tym artykuły opublikowane w ponad 22,8 tys. czasopismach naukowych i 8,3 mln artykułów opublikowanych w materiałach konferencyjnych. Obecnie baza Scopus zawiera ponad 71 mln rekordów (Scopus, 2018).

Na podstawie wyszukiwania za pomocą terminu „big data” przeprowadzonego 18 listopada 2018 r. w polu słów kluczowych, w którym umieszczane są terminy wyodrębnione jako charakterystyczne dla tematyki omawianej w indeksowanym dokumencie, z zasobów bazy Scopus wyodrębniono zbiór piśmiennictwa dotyczącego zjawiska big data. Ponieważ celem wyszukiwania było wyodrębnienie wszystkich publikacji na ten temat, niezależnie od ich formy i przynależności dziedzinowej, nie zastosowano ograniczeń formalnych ani dziedzinowych. W rezultacie otrzymano 44 052 rekordów publikacji wydanych w okresie od 2000 do 2019 r.<sup>4</sup>. Analiza rozkładu chronologicznego otrzymanego zbioru rekordów wykazała, że w Scopus zarejestrowano jedną publikację wydaną w 2000 r., a w latach 2001–2010 – od jednej do siedmiu publikacji. Wyraźny wzrost publikacji dotyczących big data następuje w 2011 r., w którym zarejestrowano już 25 dokumentów o tej tematyce, z czego 21 opublikowanych zostało w czasopismach i materiałach konferencyjnych z zakresu informatyki. W kolejnych latach wzrostowa tendencja utrzymuje się, a tempo wzrostu szybko rośnie: w 2012 r. zarejestrowano 435 publikacji o big data, w 2013 – 1374, w 2014 – 3618, w 2015 – 7536 publikacji. W 2016 r. liczba publikacji dotyczących big data przekroczyła 10 tys., a tempo jej wzrostu w kolejnych latach znacznie zmniejszyło się, co świadczyć może o pewnej stabilizacji aktywności środowiska badawczego zajmującego się tym zagadnieniem. W 2018 r. do połowy listopada zarejestrowano 10 369 publikacji oraz 306, które ukazały się już z datą wydania 2019 (Rys. 1).

Pierwszą publikacją dotyczącą problematyki big data, którą zarejestrowano w Scopus jest artykuł zakwalifikowany zarówno do nauk komputerowych, jak i nauk społecznych, dotyczący metod kompresji obrazów, wydany w 2000 r. w czasopiśmie *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives* (Rys. 2).

Rozkład geograficzny zbioru publikacji o big data wyodrębnionego z bazy Scopus prezentuje wyraźną dominację w badaniu tego zagadnienia dwóch krajów: Chin i Stanów Zjednoczonych (Rys. 3). Na chińskie ośrodki badawcze przypada 31% badań omówionych w tym piśmiennictwie, a na USA 24%. Dwoma kolejnymi państwami, w których prowadzona jest największa liczba badań o tej tematyce, są Indie (8%) i Wielka Brytania (ok. 6%).

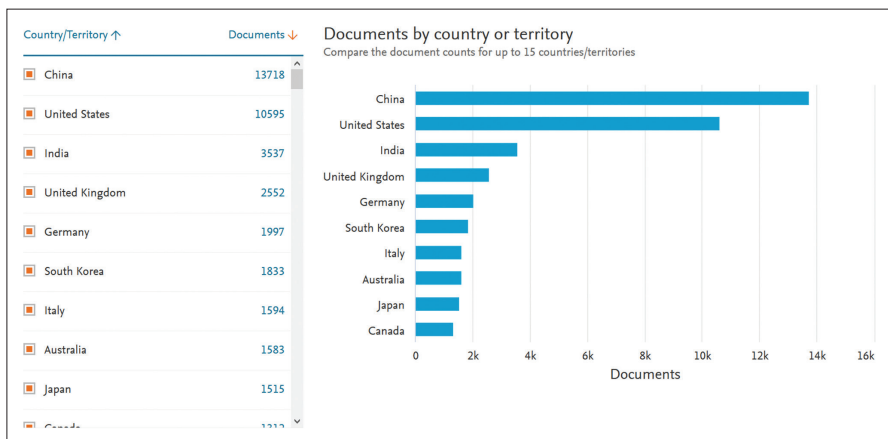
<sup>4</sup> Rejestracja artykułów z rokiem wydania 2019 wynika z coraz powszechniejszej tendencji przyspieszania publikacji online tekstów zaakceptowanych i przygotowanych do druku w numerach, które formalnie często ukazać się mają dopiero za kilka miesięcy.



Rys. 1. Chronologiczny rozkład publikacji dotyczących big data, zarejestrowanych w bazie Scopus [wyszukiwanie: 18.11.2018]

Document title	Authors	Year	Source	Cited by
1 Image compression versus matching accuracy	Kiefner, M., Hahn, M.	2000	International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives 33, pp. 316-323	7

Rys. 2. Rekord pierwszego artykułu dotyczącego problematyki big data, zarejestrowanego w Scopus [18.11.2018]

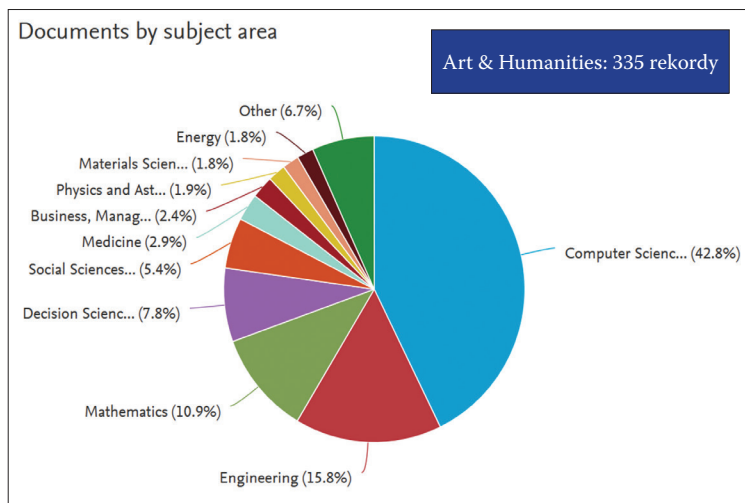


Rys. 3. Rozkład geograficzny publikacji dotyczących big data zarejestrowanych w bazie Scopus [18.11.2018]

Rozkład piśmiennictwa według dyscyplin naukowych uwidacznia jednoznaczną dominację informatyki w badaniach dotyczących big data: przypada na nią blisko 43% wszystkich publikacji (Rys. 4). Big data to przede wszystkim zagadnienie informatyczne, ale też związane z badaniami zarówno podstawowymi, jak i aplikacyjnymi w innych dziedzinach. W świetle danych otrzymanych w wyniku przeprowadzonego wyszukiwania dziedzinami wyodrębnionymi według kategoryzacji stosowanej w Scopus, w których zagadnienie big data zyskało znaczącą popularność, są: nauki techniczne (15,8%), matematyka (10,9%) i tzw. *decision sciences* – interdyscyplinarne pole badań zajmujące się wykorzystywaniem technik ilościowych w podejmowaniu decyzji w zarządzaniu i biznesie (7,8%). W dwóch pierwszych przypadkach mamy do czynienia z technicznymi aspektami przetwarzania danych masowych, organizacji i realizacji tego procesu, w drugim natomiast z matematycznymi aspektami tworzenia algorytmów wykorzystywanych w przetwarzaniu big data. *Decision sciences* zajmują się działalnością, w której technologie big data znajdują najszersze zastosowanie. Czwartą kategorię dziedzinową pod względem liczby publikacji o big data stanowią nauki społeczne, na które przypada 5,4% publikacji (4319 rekordów). Trzeba jednak zaznaczyć, że z kategorii tej w Scopus wyłączone zostały zarówno wspomniane *decision sciences*, jak i nauki o zarządzaniu i biznesie (1917 rekordów, 2,4%).

Analizując rozkład piśmiennictwa o big data według kryterium dziedzinowego, warto zwrócić uwagę na nieobecność humanistyki i nauk o sztuce wśród kategorii wyróżnionych przez narzędzia analityczne Scopus w uzyskanym wyniku wyszukiwania. Spośród publikacji z tego zakresu w Scopus zarejestrowano dotąd zaledwie 335 publikacji, w których podejmowany był temat big data. Wielkość taka stanowi zaledwie 0,76% wszystkich publikacji o tej tematyce, toteż publikacje te włączone zostały do grupy „Other”.

Podsumowując, można zatem stwierdzić, że wykorzystanie technologii big data w naukach społecznych stanowi nurt badań już wyraźnie wyodrębniony i ilościowo znaczący, natomiast w humanistyce zainteresowanie tym modelem badań jest jeszcze niewielkie.



Rys. 4. Rozkład piśmiennictwa dotyczącego big data według dyscyplin naukowych [wyszukiwanie: 18.11.2018]



W kategoryzacji dyscyplin naukowych stosowanej w bazie Scopus nauka o informacji ulokowana jest w podkategorii Library and Information Sciences, należącej do nauk społecznych. Narzędzia analityczne bazy nie umożliwiają jednak wyodrębnienia podkategorii dziedzinowych w wynikach wyszukiwania, dlatego też bardziej szczegółowe analizy dotyczące problematyki danych masowych w nauce o informacji przeprowadzone zostały na podstawie piśmiennictwa zarejestrowanego w dziedzinowej bazie EBSCO – Library and Information Science and Technology Abstracts (LISTA). Wyniki tych badań omówione są w następnej części artykułu.

### 3. Big data i analityka danych masowych a problematyka badawcza nauki o informacji

Zagadnienie big data można traktować jako kolejną fazę rozwoju technologii komputerowej i jej zastosowań w różnych dziedzinach nauki i praktyki. Wokół problematyki wykorzystania technologii komputerowej do zapewnienia sprawnego dostępu do utrwalonej informacji i wiedzy w połowie XX w. ukształtowała się współczesna nauka o informacji jako interdyscyplinarny obszar badań o własnym, jednoznacznie wyodrębnionym repertuarze problemów badawczych. Tefko Saracevic, analizując specyfikę tych problemów, opisał naukę o informacji następująco:

Information science is the science and practice dealing with the effective collection, storage, retrieval, and use of information. It is concerned with recordable information and knowledge, and the technologies and related services that facilitate their management and use. More specifically, information science is the field of professional practice and scientific inquiry addressing the effective communication of information and information objects, particularly knowledge records, among humans in the context of social, organizational, and individual need for and use of information (Saracevic, 2010, 2570).

Z kolei według popularnego *Online Dictionary of Library and Information Science* nauka o informacji to:

The systematic study and analysis of the sources, development, collection, organization, dissemination, evaluation, use, and management of information in all its forms, including the channels (formal and informal) and technology used in its communication (Reitz, 2014).

Warto tu jeszcze przytoczyć wyróżnienie przez Bruno Jacobfeuerborna dwóch wymiarów nauki o informacji (Jacobfeuerborn, 2013). W pierwszym, wykorzystuje ona interdyscyplinarne podejście do rozwijania podstaw teoretycznych i metodologicznych własnych problemów badawczych, skupionych na społecznym transferze informacji i wiedzy oraz szerokim kontekście jego uwarunkowań. W wymiarze drugim, celem nauki o informacji jest pomaganie badaczom, uczonym, inżynierom, wynalazcom i innym tzw. pracownikom wiedzy w lokalizacji i pozyskaniu informacji i wiedzy niezbędnych w ich pracy. W obu tych wymiarach zjawisko big data powinno być istotnym elementem problematyki badawczej dyscypliny. W pierwszym, może pomóc w znalezieniu rozwiązania problemów związanych z zarządzaniem informacją i wiedzą w środowisku ogromnych i szybko rosnących zbiorów danych, w drugim – może wskazać skuteczne metody obsługi użytkowników w tym nowym środowisku informacyjnym. Odpowiedzią na potrzeby związane z rozwiązaniem tych problemów jest kształtowanie się w ostatnich latach tzw.

*data science* – nowego nurtu badań w szeroko rozumianej nauce o informacji skupionego na poszukiwaniu i praktycznym zastosowaniu metod derywowania wiedzy (znaczenia i wartości) z wielkich zbiorów danych.

W środowisku wielkich zasobów danych zapisanych w cyfrowym formacie, technologie big data zapewniają wgląd w wiedzę, której nie można byłoby wydobyć tradycyjnymi metodami wyszukiwania informacji. W tym sensie można powiedzieć, że technologie te wspierają procesy transferu wiedzy między ludźmi (jakkolwiek nie tylko między ludźmi<sup>5</sup>), które stanowią główny przedmiot zainteresowań nauki o informacji. A zatem zasadna jest teza, że rozwój technologii big data i jej zastosowań stanowi nowe i niezwykle ważne wyzwanie dla nauki o informacji, pozostając w ścisłej korelacji z jej kluczową problematyką badawczą. Teza ta powinna znaleźć potwierdzenie w rosnącej liczbie badań dotyczących big data w nauce o informacji. Omówiona poniżej analiza piśmiennictwa zarejestrowanego w bazie LISTA jest próbą weryfikacji tej tezy.

### **3.1. Metoda i próba badawcza**

Jak wspomniano wcześniej, zainteresowanie problematyką big data wśród badaczy nauki o informacji zostało zbadane na podstawie piśmiennictwa indeksowanego w międzynarodowej, dziedzinowej bazie LISTA (EBSCO), która należy do najbardziej wyczerpujących źródeł informacji o piśmiennictwie naukowym tej dyscypliny. W bazie tej indeksowanych jest ponad 560 czasopism z zakresu nauki o informacji i bibliotekoznawstwa (NIB) oraz ich dyscyplin pokrewnych, a także wybrane książki i materiały konferencyjne. W LISTA indeksowane jest piśmiennictwo wydawane w ponad 20 językach, jakkolwiek zdecydowaną większość stanowią publikacje w języku angielskim. Zasięg chronologiczny bazy obejmuje okres od połowy lat 60. XX w. Trzeba zaznaczyć, że – ponieważ zakres tematyczny bazy obejmuje nie tylko piśmiennictwo nauki o informacji, a baza nie zapewnia możliwości automatycznego wyodrębniania publikacji reprezentujących poszczególne subdyscypliny objęte indeksowaniem – trudno uzyskane wyniki jednoznacznie interpretować jako dotyczące wąsko rozumianej nauki o informacji. Notabene, na ogół niemożliwe jest wyznaczenie granic pola badawczego nauki o informacji jednoznacznie oddzielających je od problematyki bibliotekoznawstwa, tzw. informatyki stosowanej i wielu innych dziedzin, w których podejmowane są badania nad zjawiskami i procesami informacyjnymi. Stąd wyniki przeprowadzonego wyszukiwania trzeba interpretować w odniesieniu do całego obszaru tematycznego, objętego indeksowaniem w bazie LISTA. Aby jednak ocenić rolę badań dotyczących big data w nauce o informacji w sensie omówionym w poprzedniej

---

<sup>5</sup> Tradycyjnie przyjmuje się, że jednym z trzech głównych obszarów badań, składających się na intelektualną strukturę nauki o informacji, obok problematyki źródeł informacji i problematyki technologii informacyjnej, jest problematyka użytkowników informacji i użytkowania informacji, która dotychczas była łączona przede wszystkim z badaniem potrzeb i zachowań informacyjnych ludzi (por. Sosińska-Kalata, 2017). Rozwój inteligentnych technologii informacyjnych zmusza jednak do weryfikacji koncepcji użytkownika informacji, włączając do niej również problematykę użytkowania informacji przez np. inteligentne urządzenia, wspierające, a coraz częściej zastępujące człowieka w różnych działaniach związanych z poszukiwaniem i pozyskiwaniem potrzebnej informacji. W polskim piśmiennictwie kwestię konieczności redefinicji pojęcia użytkownika w nauce o informacji omawiał ostatnio Remigiusz Sapa (2018).

części artykułu, przyjrano się też obecności tej problematyki na łamach czasopism, które dotychczas uznawane były za najbardziej reprezentatywne dla tej dyscypliny<sup>6</sup>.

Wyszukiwanie przeprowadzone zostało 15 listopada 2018 r. za pomocą trzech kwerend, które miały zapewnić:

- (a) wyodrębnienie najwcześniejszych wystąpień określenia „big data” w piśmiennictwie zarejestrowanym w LISTA;
- (b) wyodrębnienie najwcześniejszych wystąpień określenia „big data” w czasopismach naukowych (recenzowanych) indeksowanych w LISTA;
- (c) wyodrębnienie zarejestrowanego w LISTA piśmiennictwa naukowego, w którym podejmowano problematykę danych masowych (big data).

Analiza ilościowa piśmiennictwa naukowego, w którym podejmowano problematykę big data, została przeprowadzona według pięciu kryteriów:

- (a) data publikacji (rozkład chronologiczny badań);
- (b) język publikacji;
- (c) czasopisma (koncentracja i rozproszenie publikacji o big data);
- (d) afiliacje autorów (rozkład geograficzny badań);
- (e) struktura tematyczna.

Trzeba zaznaczyć, że omówione poniżej badanie ma charakter sondażowy i przeprowadzone zostało na indeksowanym w bazie LISTA piśmiennictwie, które wyodrębniono za pomocą prostej kwerendy wymagającej jedynie użycia w opisie tematycznym terminu „big data”. Piśmiennictwo prezentujące problematykę związaną z badaniem wielkich zbiorów danych, w którego indeksowaniu nie użyto tego terminu, nie zostało zatem uwzględnione w badaniu.

### 3.2. Najwcześniejsze wystąpienia określenia „big data” w piśmiennictwie zarejestrowanym w bazie LISTA

Ogólny sondaż najwcześniejszych wystąpień problematyki big data w piśmiennictwie zarejestrowanym w bazie LISTA został przeprowadzony na podstawie kwerendy: „big data” (wszystkie pola). Poszukiwano więc wystąpień wyrażenia „big data” w całej zawartości rekordów. W rezultacie uzyskano 1288 rekordów publikacji z okresu 1974–2018.

<sup>6</sup> Za czasopisma najbardziej reprezentatywne („główne”, „kanoniczne”) dla nauki o informacji uważa się takie czasopisma, których profil tematyczny obejmuje szerokie spektrum problematyki badawczej tej dyscypliny, wokół których skupiają się jej uznani badacze i które należą do najczęściej cytowanych, co znajduje odzworowanie w wysokim wskaźniku wpływu (IF, SNIP). Na podstawie takich kryteriów do grupy najbardziej reprezentatywnych czasopism nauki o informacji obecnie należałoby zaliczyć: *International Journal of Information Management*, *Information Processing & Management*, *Journal of the Association for Information Science and Technology*, *Journal of Information Science*, *Aslib Journal of Information Management*, *Journal of Documentation*, *Information Research*. Wielu badaczy do tej grupy dodaje też *Scientometrics*, *Journal of Informetrics* oraz *Library Hi Tech* i *Library and Information Science Research* jako czasopisma o wysokim IF, które specjalizują się w subdyscyplinach nauki o informacji należących do jej kluczowych nurtów badawczych. Ze względu na fakt, że piśmiennictwo nauki o informacji jest rejestrowane w bazach zwykle obejmujących zakres szerszy niż pole badawcze tej dyscypliny, tego rodzaju podejście bywa stosowane przez badaczy, którzy podejmują próby oceny zjawisk i własności specyficznych dla samej nauki o informacji, np. stan rozwoju dyscypliny, trendy badawcze, front badań, współpraca międzynarodowa, poziom interdyscyplinarności itp. (zob. np. White & McCain, 1998; Zhao & Strotmann, 2008; Chang & Huang, 2012; Sosińska-Kalata, 2013).

Najwcześniejsze publikacje, wydane w latach 1974, 1977 i 1981, dotyczą przetwarzania dużych baz danych, a wyszukanie rekordów tych publikacji wynika z wystąpienia w ich abstraktach frazy „big data” w wyrażeniu „big data bases”. Tematem pierwszej z tych publikacji – artykułu, który ukazał się w periodyku *Naučno-techničeskaja informacija* – były algorytmy wyszukiwania informacji w dużych bazach dokumentacyjnych (Rys. 5). Jakkolwiek artykuł ten nie dotyczy technologii big data w sensie współczesnym, to warto zwrócić uwagę na to, iż omawiano w nim problemy związane z przetwarzaniem wielkich zbiorów danych oraz na to, iż ukazał się w jednym z głównych czasopism fachowych zajmujących się problematyką informacji naukowej, wydawanych w tamtym czasie w ZSRR.

<b>Algorithm, realizuiuschchii poisk v dokumental'noi ips.</b> (retrieval algorithm in a documentary irs.)	
<b>Język:</b>	Russian
<b>Autorzy:</b>	Avdeev, B A Borodin, V V
<b>Źródło:</b>	Nauchno Tehnicheskaya Informatsiya Series 2. 1974, Vol. 2 Issue 8, p30-32. 3p.
<b>Typ dokumentu:</b>	Article
<b>Abstrakt:</b>	Algorithms implementing retrieval operations in documentary irs are depicted. The number of external memory calls is evaluated, which is an important characteristic of algorithm performance in handling big data bases.
<b>Uwagi:</b>	Update Code: 1000
<b>Numer akcesji:</b>	ISTA1000442

Rys. 5. Pierwsze zarejestrowane w bazie LISTA użycie wyrażenia „big data” w abstrakcie artykułu

W 1983 r. w *Lecture Notes on Computer Science* ukazał się pierwszy artykuł, spośród zarejestrowanych w bazie LISTA, w którym użyte zostało określenie „big data” na oznaczenie wielkich zbiorów danych. Również ten artykuł dotyczył tworzenia nowych algorytmów przeszukiwania tego typu zbiorów danych (Rys. 6).

<b>Treatment of big values in an applicative language HFP.</b> <b>Translation from by-value access to by-update access</b>	
<b>Autorzy:</b>	Katayama, T <sup>1</sup>
<b>Źródło:</b>	Lecture Notes on Computer Science, Vol. 147. 1983.
<b>Typ dokumentu:</b>	Book Chapter
<b>Pojęcia tematu:</b>	COMPUTERS ENGINEERING LANGUAGE & languages
<b>Słowa kluczowe podane przez autora:</b>	Access methods
<b>Abstrakt:</b>	This paper proposes a method of treating big data by converting by-value access to by-update access, which is used in the implementation of an applicative language HFP. HFP is an applicative language which admits hierarchical and applicative programming and is based on the attribute grammar of Knuth. It also has a close relationship to Prolog. The author first introduces HFP and discusses its implementation which solves the big data problem by using a simple file processing program Book Published by Springer-Verlag, Germany, 1983
<b>Uwagi:</b>	Place of Publication: Germany Publisher: Springer-Verlag Update Code: 1900
<b>Afilacje autora:</b>	<sup>1</sup> Tokyo Inst. of Technology Tokyo, Japan
<b>Numer akcesji:</b>	ISTA1904357

Rys. 6. Pierwszy artykuł, zarejestrowany w bazie LISTA, w którym wyrażenie „big data” zostało użyte na oznaczenie wielkich zbiorów danych

<b>BIG DATA.</b>	
<b>Autorzy:</b>	Arnold, Stephen E. sa@arnolditf.com
<b>Źródło:</b>	Online. Mar/Apr2011, Vol. 35 Issue 2, p27-26. 3p.
<b>Typ dokumentu:</b>	Article
<b>Pojęcia tematu:</b>	*Information resources management *Information retrieval *Internet *Electronic publications *Access to information *Information overload Business
<b>Abstrakt:</b>	The article discusses the problem posed by big data to information service providers. The bigger problem of complying with the data required for a legal electronic discovery request is noted. The amount of data being handled by Amazon.com, eBay, Facebook, Google and telecommunications companies is explained. The Hadoop Wiki also presents a challenge to established software vendors such as Aster Data Systems Inc.
<b>Zliczanie słów pełnego tekstu:</b>	1415
<b>ISSN:</b>	0146-5422
<b>Numer akcesji:</b>	59290536

Rys. 7. Pierwszy artykuł, zarejestrowany w bazie LISTA, dotyczący problematyki big data w kontekście wykorzystywania danych generowanych przez użytkowników Internetu

Pierwsze publikacje, które dotyczą problemu danych masowych, zostały wydane w 2010 r. Są to krótkie komunikaty z konferencji poświęconych nowym metodom informatycznym oraz recenzja książki Davida Bolliera *The Promise and Peril of Big Data*.

W 2011 r. ukazał się natomiast pierwszy artykuł o problematyce big data w kontekście wykorzystywania nieustrukturyzowanych i różnorodnych danych generowanych przez użytkowników Internetu (serwisów e-commerce i mediów społecznościowych). Artykuł ten został wydany w nierecenzowanym magazynie *Online* (Rys. 7).

### 3.3. Najwcześniejsze wystąpienia określenia „big data” w czasopismach naukowych indeksowanych w bazie LISTA

Do śledzenia najwcześniejszych wystąpień określenia „big data” w czasopismach naukowych zostało wykorzystane ograniczenie wyników poprzedniej kwerendy za pomocą kryterium „czasopisma naukowe (recenzowane)”. Wyodrębniono w ten sposób zbiór 808 rekordów, w tym 741 rekordów artykułów naukowych i 67 rekordów artykułów recenzyjnych. Artykuły te ukazały się w latach 2008–2018.

Pierwszym recenzowanym artykułem naukowym, zarejestrowanym w bazie LISTA, w którym stwierdzono wystąpienie określenia „big data” w abstrakcie, jest artykuł kubańskich badaczy wydany w języku hiszpańskim w 2008 r. w czasopiśmie *Ciencias de la Información*. Tematem tego artykułu jest badanie z zakresu patentometrii, tj. analiza danych o kubańskich patentach zarejestrowanych w amerykańskich wielkich bazach patentowych, której celem jest wskazanie najbardziej innowacyjnych kubańskich technologii, ośrodków badawczych i badaczy. A zatem, podobnie jak w przypadku piśmiennictwa nierecenzowanego, również pierwsze użycia określenia „big data” w artykułach recenzowanych wiążą się z analizami wielkich baz danych.

Pierwszym zarejestrowanym w LISTA artykułem badawczym, w którym termin „big data” pojawia się wśród słów kluczowych, identyfikujących główne pojęcia omawianego tematu, jest artykuł wydany w 2011 r. w czasopiśmie *Journal of the American Medical Informatics Association*, dotyczący wykorzystania metod data-mining w symulacjach komputerowych (Rys.8).



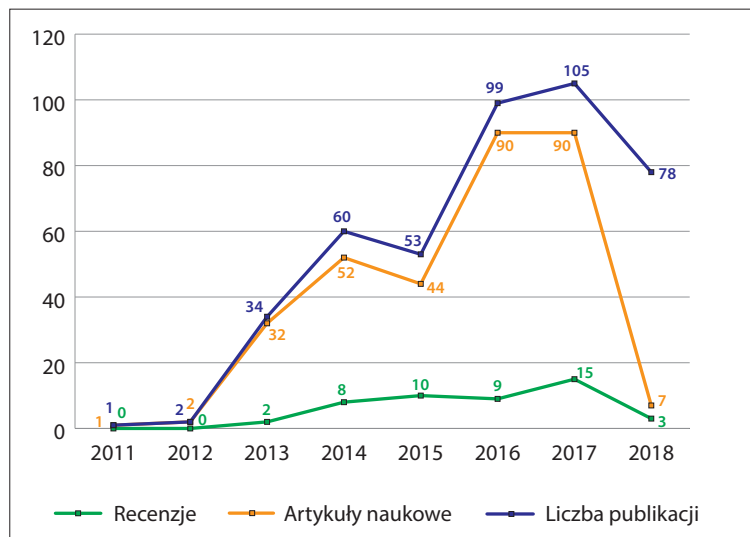
Rys. 8. Pierwszy zarejestrowany w bazie LISTA artykuł badawczy, w którego rekordzie termin „big data” występuje w polu pojęć tematu

### 3.4. Piśmiennictwo nauki o informacji, bibliotekoznawstwa i nauk pokrewnych, w którym podejmowano problematykę big data

W celu wyodrębnienia z bazy LISTA piśmiennictwa naukowego, którego przedmiotem są badania dotyczące big data, została użyta kwerenda, w której poszukiwane wystąpienia terminu „big data” ograniczono do pola pojęć tematu (SU) oraz typ dokumentu ograniczono do kategorii „czasopisma naukowe (recenzowane)”. Na podstawie tej kwerendy otrzymano zbiór 427 rekordów, w tym 381 rekordów artykułów naukowych i 47 rekordów recenzji w czasopismach naukowych. Publikacje te ukazały się w latach 2011–2018, a więc w ciągu ostatnich ośmiu lat. Najwcześniejszym artykułem zaindeksowanym terminem „big data” jest wspomniany w poprzedniej części artykuł z zakresu informatyki medycznej o zastosowaniu technik data-mining w symulacjach komputerowych.

#### 3.4.1. Rozkład chronologiczny

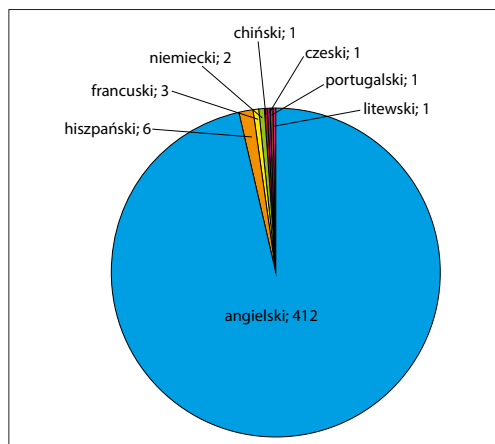
Chronologiczny rozkład publikacji na temat big data w czasopismach naukowych indeksowanych w bazie LISTA ukazuje systematyczny wzrost zainteresowania tą problematyką w ostatnich ośmiu latach (Rys. 9). Dane te pokazują też, iż, podobnie jak w innych dyscyplinach naukowych, również w nauce o informacji szybkie zwiększanie się liczby badań prowadzonych w tym zakresie następuje od 2011 r. W listopadzie 2018 r., kiedy przeprowadzane było wyszukiwanie, dane za rok 2018 były niepełne, stąd nie można brać ich pod uwagę w ocenie trendu.



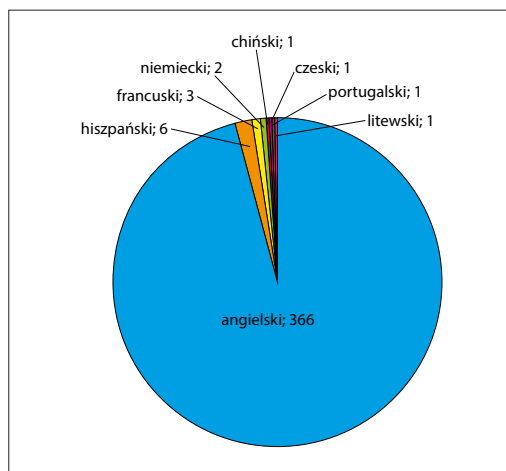
Rys. 9. Rozkład chronologiczny publikacji nt. big data w czasopiśmie naukowych indeksowanych w bazie LISTA

### 3.4.2. Języki publikacji

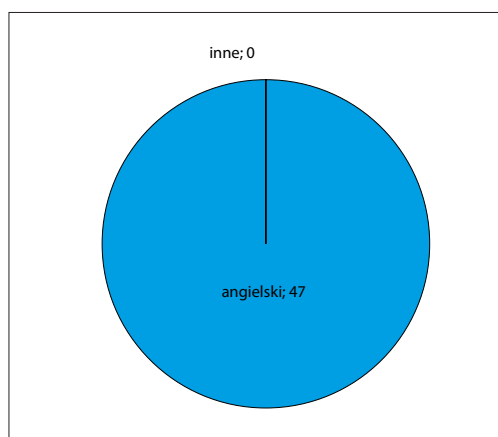
Analiza wyodrębnionego piśmiennictwa według języków publikacji (Rys. 10, 11 i 12) nie wnosi wiele do poznania specyfiki literatury dotyczącej tematyki big data w nauce o informacji i dziedzinach pokrewnych. Przeprowadzono ją, zakładając, że może ujawnić szczególne zainteresowanie tą problematyką w pewnych kręgach językowych, z wyłączeniem języka angielskiego, który jest podstawowym językiem komunikacji naukowej w obiegu międzynarodowym. Wyniki analizy nie pozwalają jednak na formułowanie takich wniosków.



Rys. 10. Rozkład wg języków publikacji wszystkich publikacji naukowych o big data, zarejestrowanych w bazie LISTA



Rys. 11. Rozkład wg języków publikacji artykułów naukowych o big data, zarejestrowanych w bazie LISTA



Rys. 12. Rozkład wg języków publikacji recenzji książek o big data, wydanych na łamach czasopism naukowych indeksowanych w bazie LISTA

Piśmiennictwo naukowe o big data, które zarejestrowano w bazie LISTA, wydano w ośmiu językach. Jak można było spodziewać się, przytłaczająca większość publikacji (97%) ukazała się w języku angielskim. W pozostałych językach wydano w ciągu ostatnich ośmiu lat od sześciu do jednego artykułu: sześć w języku hiszpańskim, trzy w języku francuskim, dwa w języku niemieckim oraz po jednym w językach: chińskim, czeskim, litewskim i portugalskim. Zwraca uwagę relatywnie większa liczba publikacji w języku hiszpańskim niż w językach pozostałych, jednak z reguły wielkości te są zbyt małe, aby można było mówić o szczególnym zainteresowaniu badaniami big data badaczy, którzy publikują w którymkolwiek języku innym niż język angielski. Także sam zestaw języków, w których opublikowano dotąd po jednym artykule, trzeba traktować jako dość przypadkowy. Uzyskane



dane demonstrują przede wszystkim dominację języka angielskiego zarówno w polityce indeksowania bazy LISTA, jak i ogólnie we współczesnej komunikacji naukowej.

### 3.4.3. Czasopisma nauki o informacji, w których ukazały się publikacje o big data

Więcej interesujących informacji dostarcza analiza czasopism, w których ukazały się artykuły i recenzje książek o tematyce big data. Publikacje te wydano na łamach 113 czasopism, czyli w około 20% źródeł indeksowanych w bazie LISTA. Piśmiennictwo to jest więc znacznie rozproszone. Ponad połowa (51,8%) artykułów i recenzji dotyczących zagadnień big data ukazała się jednak w zaledwie 13 czasopismach (Rys. 13). Można zatem stwierdzić, że wśród źródeł indeksowanych w bazie LISTA istnieje względnie nieduża grupa czasopism wyraźnie bardziej zainteresowana zagadnieniami big data niż pozostałe źródła. Biorąc pod uwagę liczbę opublikowanych artykułów, w grupie tej wyróżnić można trzy podgrupy:

- (1) czasopisma, w których ukazało się ponad 20 artykułów; podgrupę tę tworzą cztery periodyki (*Scientometrics*, *Information Systems*, *Journal of the American Medical Informatics Association*, *Choice: Current Reviews for Academic Libraries*);
- (2) czasopisma, w których ukazało się od 14 do 20 artykułów; do tej podgrupy należą trzy periodyki (*First Monday*, *Information Journal of Information Management*, *Information, Communication and Society*);
- (3) czasopisma, w których ukazało się od 8 do 10 artykułów; podgrupę tę tworzy sześć periodyków (*Information Services & Use*, *Library Hi Tech*, *El Profesional de la Informacion*, *Information Processing and Management*, *Journal of Medical Internet Research*, *Information Polity: The International Journal of Government & Democracy in the Information Age*).

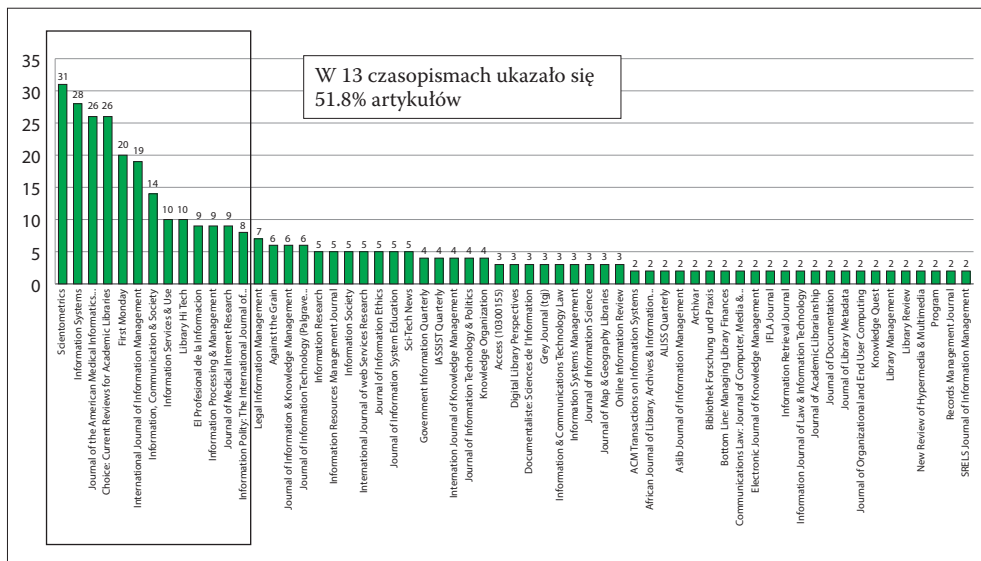
W pierwszej grupie znajdują się trzy periodyki, które zajmują się specjalistycznymi obszarami szeroko rozumianej nauki o informacji (naukometrią, zastosowaniami technologii inteligentnych, informatyką medyczną) oraz czasopismo przeglądowe, wydawane przez Association of College and Research Libraries, publikujące pisane przez naukowców recenzje książek naukowych oraz źródeł internetowych zawierających treści naukowe. Chociaż w bazie LISTA magazyn *Choice* został skategoryzowany jako recenzowane czasopismo naukowe, nie jest to czasopismo publikujące oryginalne artykuły badawcze.

Wśród trzech czasopism tworzących drugą grupę znajdują się czasopisma naukowe, zajmujące się problematyką szeroko rozumianej nauki o informacji, w tym dwa, które mają charakter interdyscyplinarnych czasopism zajmujących się problematyką komunikacji społecznej i współczesnych mediów (*Information, Communication and Society* oraz *First Monday*).

W trzeciej grupie znajdują się cztery czasopisma, których profil obejmuje szeroki zakres problematyki badawczej nauki o informacji (*Information Services & Use*, *Library Hi Tech*, *El Profesional de la Informacion*, *Information Processing and Management*) oraz dwa poświęcone specjalistycznym zagadnieniom nauki o informacji (*Journal of Medical Internet Research*, *Information Polity: The International Journal of Government & Democracy in the Information Age*).

Dane te pozwalają wysnuć wniosek, że w nauce o informacji badaniami big data interesują się w największym stopniu czasopisma, których profil ukierunkowany jest na problematykę zasobów i usług informacyjnych w obszarach specjalistycznych. Nieco rzadziej problematyka ta dotychczas pojawiała się na łamach czasopism nauki o informacji o profilu ogólnym, obejmującym szeroki repertuar problemów badawczych nauki

o informacji. Interesujące jest przy tym to, że wśród czasopism publikujących artykuły dotyczące badań big data jest niewiele takich, które należą do tzw. kanonicznych czy też głównych czasopism nauki o informacji – uznawanych za najważniejsze i najbardziej reprezentatywne dla tej dyscypliny.



Rys. 13. Ilościowy rozkład artykułów dotyczących big data w czasopismach naukowych z zakresu nauki o informacji i dyscyplin pokrewnych, w których ukazały się co najmniej dwa artykuły o tej tematyce

W tabeli 1 zostały zestawione tytuły czasopism, w których w latach 2011–2018 ukazało się co najmniej pięć artykułów poświęconych big data. Obok tytułów czasopism umieszczono ich aktualny<sup>7</sup> *impact factor* (IF), *Source Normalized Impact per Paper* (SNIP)<sup>8</sup> oraz liczbę artykułów zaindeksowanych w bazie LISTA do 18 listopada 2018 r. Zestawienie pokazuje, że pierwsze trzy czasopisma, w których w badanym okresie ukazało się najwięcej artykułów o analizowanej tu problematyce, należą do wysoko cytowanych periodyków naukowych, rejestrowanych zarówno w *Web of Science*, jak i w *Scopus*. Według danych *Journal Citation Reports*, w kategorii „Information Science & Library Science” *Scientometrics* zajmuje 25. pozycję w rankingu czasopism o największym oddziaływaniu, *Information Systems* pozycję szóstą, a *Journal of the American Medical Informatics Association* pozycję piątą. Czasopismo *International Journal of Information Management*, którego IF ma największą wartość wśród czasopism wyodrębnionych z bazy LISTA, w JCR w 2017 r. zostało sklasyfikowane na trzeciej pozycji wśród czasopism należących do kategorii „Information Science & Library Science”.

<sup>7</sup> Na podstawie *Journal Citation Reports* bazy *Web of Science* – raport za 2017 r.

<sup>8</sup> Na podstawie danych o indeksowanych źródłach bazy *Scopus* za 2017 r.

Tab. 1. Czasopisma indeksowane w bazie LISTA, w których ukazało się co najmniej pięć artykułów dotyczących problematyki big data

L.p.	Tytuł czasopisma	IF (2017)	SNIP (2017)	Liczba artykułów
1	<b>Scientometrics</b>	<b>2.173</b>	<b>1.378</b>	<b>31</b>
2	Information Systems	4.267	2.251	28
3	Journal of the American Medical Informatics Association	4.270	2.262	26
4	Choice: Current Reviews for Academic Libraries	–	–	26
5	First Monday	–	0.771	20
6	<b>International Journal of Information Management</b>	<b>4.516</b>	<b>2.824</b>	<b>19</b>
7	Information, Communication & Society	–	1.989	14
8	Information Services & Use	–	0.497	10
9	<b>Library Hi Tech</b>	<b>0.759</b>	<b>0.722</b>	<b>10</b>
10	El Profesional de la Informacion	–	1.130	9
11	<b>Information Processing &amp; Management</b>	<b>3.444</b>	<b>2.66</b>	<b>9</b>
12	Journal of Medical Internet Research	–	1.815	9
13	Information Polity: The International Journal of Government & Democracy in the Information Age	–	0.935	8
14	Legal Information Management	–	–	7
15	Against the Grain	–	–	6
16	Journal of Information & Knowledge Management	–	0.56	6
17	Journal of Information Technology (Palgrave Macmillan)	4.535	2.638	6
18	<b>Information Research</b>	<b>0.762</b>	<b>0.815</b>	<b>5</b>
19	Information Resources Management Journal	–	0.209	5
20	Information Society	1.889	1.225	5
21	International Journal of Web Services Research	–	0.387	5
22	Journal of Information Ethics	–	0.201	5
23	Journal of Information Systems Education	–	0.685	5
24	Sci-Tech News	–	–	5

Tabela 1 pokazuje też, iż więcej niż połowa czasopism, w których ukazały się artykuły o badaniach dotyczących big data w kontekście nauki o informacji i jej dziedzin pokrewnych, to periodyki, które nie są rejestrowane w Web of Science, a zatem o poziomie wpływu niższym niż wymagany od czasopism objętych tą bazą. W przypadku indeksowania w bazie Scopus, tylko cztery spośród tych czasopism nie są nim objęte.

W tabeli 1 czcionką półgrubą zostały zaznaczone czasopisma, które należą do tzw. kanonicznych czasopism nauki o informacji. Wśród periodyków, w których ukazało się co najmniej pięć artykułów o big data jest ich zaledwie pięć. Pełniejszy obraz obecności tej problematyki w czasopismach, które można uznać za najważniejsze we współczesnej nauce o informacji, prezentuje tabela 2. W okresie objętym badaniem w czasopismach tych opublikowano łącznie 82 artykuły dotyczące problematyki big data, co stanowi nieco więcej niż jedną piątą wszystkich publikacji zaindeksowanych terminem „big data” w bazie LISTA.

W tej grupie najwięcej, bo blisko 38% publikacji, ukazało się w *Scientometrics*, podobnie zresztą jak w całej badanej próbie (ponad 8%). Ilościowe badania informacji o nauce i piśmiennictwie naukowym wyraźnie stanowią nurt badań, w którym technologie i metody big data znajdują najczęstsze zastosowanie.

Tab. 2. „Kanoniczne” czasopisma nauki o informacji, w których ukazały się artykuły dotyczące problematyki big data

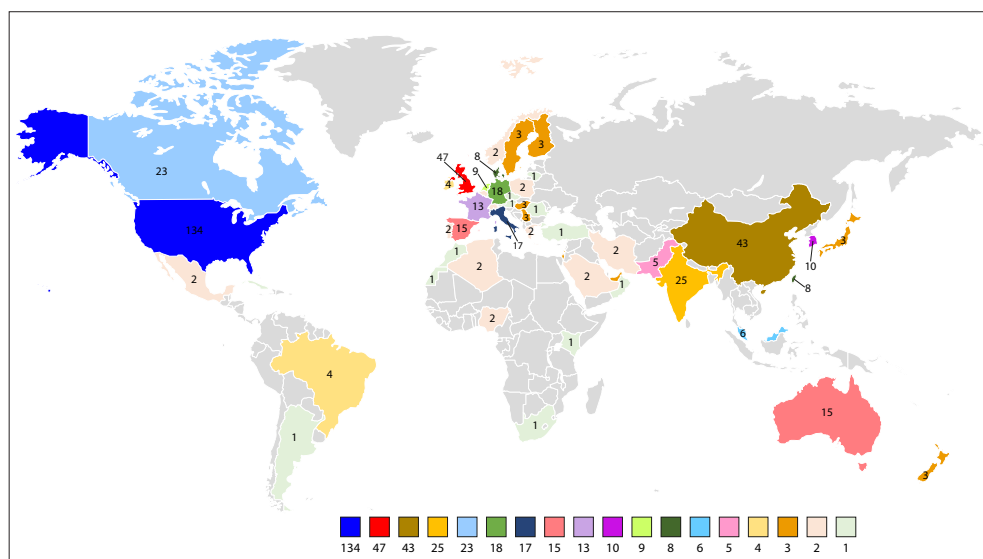
Lp.	Tytuł czasopisma	IF (2017)	SNIP (2017)	Liczba artykułów
1	<i>Scientometrics</i>	2.173	1.378	31
2	<i>International Journal of Information Management</i>	4.516	2.824	19
3	<i>Library Hi Tech</i>	0.759	0.722	10
4	<i>Information Processing &amp; Management</i>	3.444	2.66	9
5	<i>Information Research</i>	0.762	0.815	5
6	<i>Journal of Information Science</i>	1.939		3
7	<i>Aslib Journal of Information Management</i>	1.461		2
8	<i>Journal of Documentation</i>	1.157		2
9	<i>Journal of the Association for Information Science &amp; Technology</i>	2.835		1
<b>Razem artykułów</b>				82 (21.85%)

Zaskakujące może wydawać się to, że w ciągu ostatnich ośmiu lat zaledwie jeden artykuł o tematyce big data ukazał się w JASIST (*Journal of the Association for Information Science and Technology*, do 2013 r. *Journal of the American Society for Information Science and Technology*) – czasopiśmie często uznawanym za najważniejsze źródło dla nauki o informacji i najbardziej reprezentatywne dla jej pola badawczego. Równie zastanawiające jest to, że w *Journal of Documentation*, najstarszym europejskim czasopiśmie poświęconym tej dyscyplinie i także uznawanym za należące do najbardziej dla niej reprezentatywnych, dotychczas ukazały się tylko dwa artykuły dotyczące big data. Trzeba jednak przypomnieć, że wyniki przeprowadzonej analizy piśmiennictwa indeksowanego w bazie LISTA wskazały, iż liczba publikacji o tematyce big data w czasopismach o szerokim profilu, obejmującym różne obszary badań nauki o informacji, jest wyraźnie mniejsza niż w czasopismach o profilu bardziej specjalistycznym. Wyjątek stanowi *International Journal of Information Management*, którego profil określony jest względnie szeroko, ale w którym dominuje jednak problematyka zarządzania informacją w organizacjach, a więc dotycząca dziedziny, w której zastosowania technologii big data należą do najczęstszych.

#### 3.4.4. Geograficzny rozkład publikacji naukowych na temat big data w nauce o informacji

Na podstawie analizy afiliacji autorów publikacji naukowych wyodrębnionych z bazy LISTA został ustalony rozkład geograficzny tego piśmiennictwa, który interpretować można też jako geograficzny rozkład ośrodków badawczych, zajmujących się zagadnieniami big data w obszarze problemowym nauki o informacji (Rys. 14 i 15). Rozkład ten jednoznacznie wskazuje na zdecydowaną dominację amerykańskich ośrodków badawczych: w Stanach Zjednoczonych znajdują się ośrodki badawcze, przy których afiliowana jest blisko jedna

trzecia (29%) autorów artykułów o tej tematyce. Do grupy krajów, w których znajduje się najwięcej ośrodków afiliujących badaczy zajmujących się big data w nauce o informacji należą: Wielka Brytania (10%) i Chiny (9%). Na Stany Zjednoczone, Wielką Brytanię i Chiny przypada łącznie prawie połowa badań o tej tematyce, które omówione zostały w piśmiennictwie indeksowanym w LISTA. W Indiach i w Kanadzie zlokalizowanych jest po 5% ośrodków, przy których afiliowani są autorzy badanego piśmiennictwa, w Niemczech i we Włoszech po 4%, a na Australię, Hiszpanię i Francję przypada po 3% ośrodków.

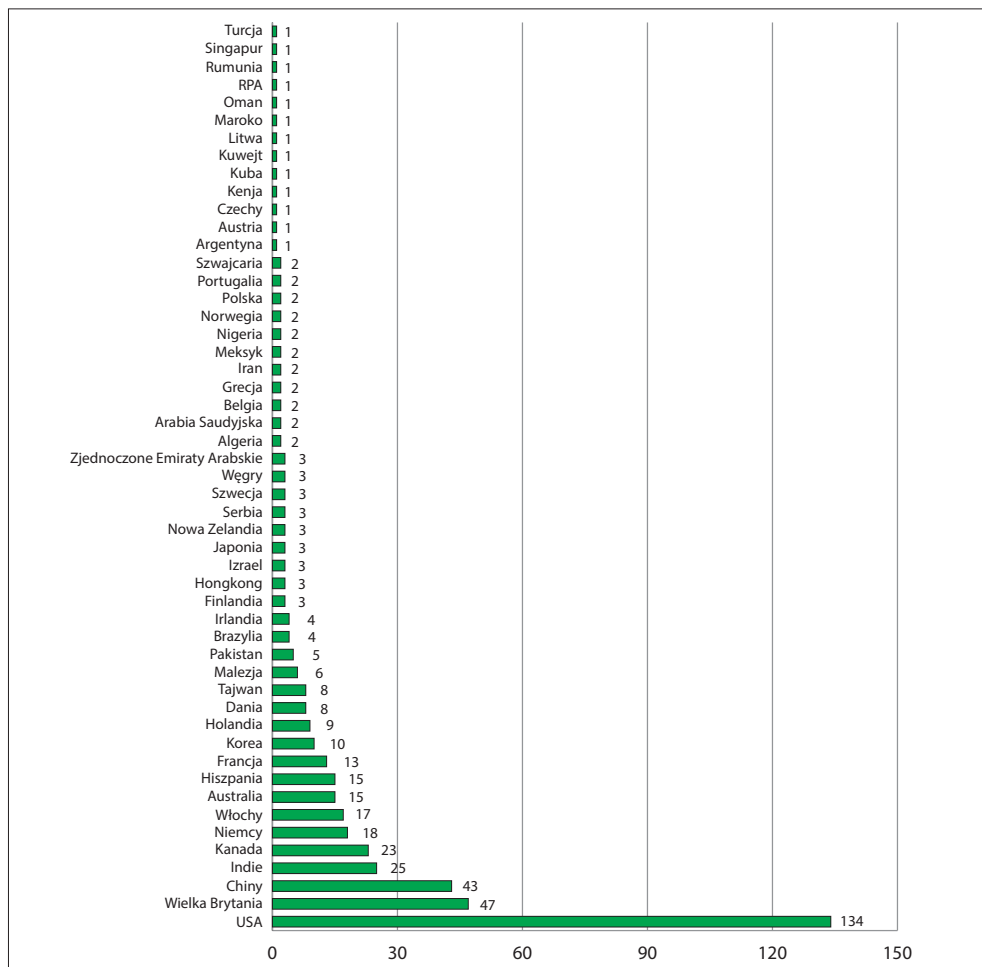


Rys. 14. Geograficzny rozkład publikacji naukowych dotyczących big data, zarejestrowanych w bazie LISTA

Nie jest zaskoczeniem fakt, że najbardziej aktywni w problematyce dotyczącej big data i zastosowań tej technologii w obszarze problemowym nauki o informacji są badacze związani z ośrodkami naukowymi w Stanach Zjednoczonych. Powszechnie znana jest supremacja tego kraju na polu badań nowych technologii i ich zastosowań, i tam nauka o informacji niemal od początku rozwija się najbardziej prężnie. Nie można zapominać też, że to właśnie amerykańskie firmy (Microsoft, Google, Amazon, Apple) mają najlepszy dostęp do globalnych big data. Również wysoka intensywność badań dotyczących problematyki informacyjnej zarówno w wymiarze technologicznym, jak i społecznym w Wielkiej Brytanii znajduje potwierdzenie w znaczącym zainteresowaniu tamtejszych ośrodków badawczych tą nową technologią i jej aplikacjami. Ciekawym zjawiskiem jest duża liczba chińskich ośrodków naukowych zaangażowanych w nurt badań dotyczących big data w nauce o informacji. Coraz bardziej intensywna eksploracja zaawansowanych technologii przez chińskich badaczy jest widoczna także w wielu innych dziedzinach. Warto zauważyć, że szybki rozwój takich firm jak Alibaba, Baidu czy Hauwei zapewnia Chinom coraz lepszy dostęp do big data.

Porównanie rozkładu geograficznego afiliacji autorów publikacji dotyczących big data w nauce o informacji z omówionym wcześniej analogicznym rozkładem piśmiennictwa

z różnych dziedzin zarejestrowanego w bazie Scopus (Rys. 3) pozwala stwierdzić, że chińska nauka jest obecnie wyraźnym liderem w badaniach nad big data na świecie, jednak w nauce o informacji Stany Zjednoczone i Wielka Brytania nadal utrzymują nad nią przewagę. Poza tym oba rozkłady są dość podobne. Interesującą kwestią jest też zaobserwowana, w piśmiennictwie o tematyce big data zarejestrowanym w bazie LISTA, stosunkowo duża liczba afiliowanych w ośrodkach amerykańskich autorów, których nazwiska sugerują ich chińskie pochodzenie. Wyciągnięcie wniosków z tej obserwacji wymaga jednak dokładniejszych badań.



Rys. 15. Ilościowy rozkład publikacji naukowych dotyczących zagadnień big data według państw, w których znajdują się ośrodki badawcze afiliujące ich autorów

Uwagę bez wątpienia zwraca nieobecność Rosji w geograficznym rozkładzie piśmiennictwa dotyczącego big data w nauce o informacji. Można przypuszczać, że pewien wpływ na taki stan rzeczy ma polityka indeksowania bazy LISTA, preferująca publikacje w języku

angielskim, jednakże w bazie tej zaindeksowano łącznie 756 publikacji wydanych w języku rosyjskim, z czego 684 to teksty opublikowane w okresie 2011–2018. Z dużym prawdopodobieństwem problem dotyczy więc zainteresowania zjawiskiem big data rosyjskich badaczy zajmujących się nauką o informacji. W interdyscyplinarnej bazie Scopus wśród 44 052 publikacji zaindeksowanych słowem kluczowym „big data” 590 ma autorów afiliowanych w rosyjskich ośrodkach badawczych, z czego 539 opublikowano w języku angielskim, a tylko 51 w języku rosyjskim. W bazie Scopus, w zestawieniu piśmiennictwa o big data według krajów afiliacji autorów, Federacja Rosyjska zajmuje 14 pozycję.

Wśród piśmiennictwa o big data zaindeksowanego w LISTA znalazły się dwie publikacje autorów z polskich ośrodków badawczych, obie wydane w języku angielskim. W bazie Scopus zarejestrowano 358 publikacji o tej tematyce autorów afiliowanych w polskich instytucjach naukowych, co w zestawieniu piśmiennictwa według krajów afiliacji lokuje Polskę na 27. miejscu.

### 3.4.5. Tematyka publikacji dotyczących problematyki big data w nauce o informacji

Próbę ustalenia struktury tematycznej piśmiennictwa o zagadnieniach big data zarejestrowanego w bazie LISTA podjęto na podstawie analizy słów kluczowych występujących w polu tematu. Analiza ta została przeprowadzona w dwóch wariantach: na podstawie danych o całym zbiorze 381 artykułów naukowych dotyczących problematyki big data, które zarejestrowano w bazie LISTA oraz na podstawie danych o zbiorze zarejestrowanych w LISTA 82 artykułów na ten temat, opublikowanych w dziewięciu czasopismach uznawanych za najbardziej reprezentatywne dla badań prowadzonych w nauce o informacji.

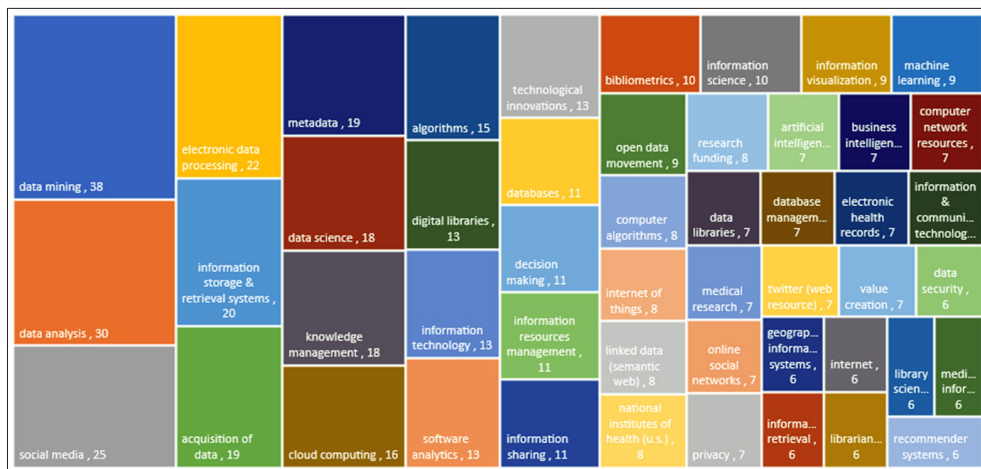
Strukturę tematyczną całego zbioru artykułów naukowych dotyczących big data, które zarejestrowano w bazie LISTA, prezentuje rysunek 16. Ze względu na zapewnienie czytelności wykresu, prezentacja ilościowego rozkładu słów kluczowych została ograniczona do terminów użytych w indeksowaniu co najmniej sześciu publikacji. W analizie nie uwzględniono wystąpień terminu *big data*, którym zaindeksowane były wszystkie artykuły w badanym zbiorze.

Uzyskany rozkład słów kluczowych przede wszystkim uwidacznia dużą różnorodność omawianych zagadnień szczegółowych. Ze zrozumiałych względów wśród terminów użytych w indeksowaniu badanego piśmiennictwa najczęściej występują wyrażenia dotyczące zagadnień informatycznych (*data mining*, *data analysis*) oraz termin *social media*, odnoszący się do najczęściej wykorzystywanego źródła nich wykorzystywanych w analizach big data.

Jeśli do liczby wystąpień terminu *social media* (25) dodamy liczbę wystąpień nazwy *Twitter (web resources)* (7) oraz termin *online social networks* (7), to otrzymamy wartość 39, wyższą nawet od liczby wystąpień terminu *data mining* (38), jak i *data analysis* (30). Możemy zatem powiedzieć, że pojęcia eksploracji danych, analizy danych oraz sieci społecznych w przestrzeni cyfrowej wskazują podstawową ramę konceptualną tematyki badań dotyczących big data w nauce o informacji i jej dziedzinach pokrewnych.

Terminy o nieco mniejszej frekwencji ukazują specyficzne dla nauki o informacji zagadnienia badań szczegółowych, związanych ze zjawiskiem big data. W tej grupie także wskazać można terminy dotyczące zagadnień informatycznych i terminy odnoszące się do obszarów zastosowań analiz big data oraz ich społecznych aspektów. Wśród zagadnień informatycznych, które wydają się w największym stopniu, interesować badaczy szeroko rozumianych nauk informacyjnych występują: *electronic data processing*, *information*

*storage & retrieval systems, acquisition of data, metadata, data science, cloud computing, algorithms, data science, information technology, software analysis, technological innovations, information visualization, machine learning, artificial intelligence, database management, linked data (semantic web).* Warto zauważyć, że wśród tych zagadnień widoczne są tematy typowe dla badań nauki o informacji (np. gromadzenie, przechowywanie, wyszukiwanie informacji, problematyka metadanych, zarządzanie bazami danych) oraz tematy związane z zastosowaniem innowacyjnych technologii (sztuczna inteligencja, maszynowe uczenie się, przetwarzanie w chmurze, wizualizacja informacji). Z kolei wśród terminów odnoszących się do obszarów zastosowań analiz i technologii big data oraz społecznych aspektów tego typu badań względnie wysoką częstotliwość w badanym zbiorze mają wyrażenia *knowledge management, digital libraries, decision making, information resources management, information sharing, bibliometrics, information science, open data movement, national institutes of health (US), medical research i ecelctronic health records.* Warto zwrócić uwagę na dość znaczny udział wśród analizowanych słów kluczowych terminów związanych z ochroną prywatności i bezpieczeństwem danych.



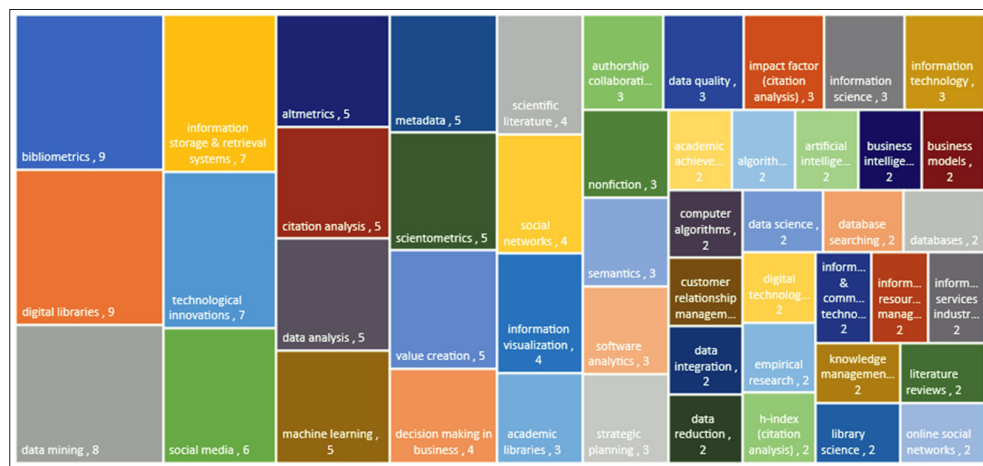
Rys. 16. Struktura tematyczna artykułów naukowych dotyczących zagadnienia big data, zarejestrowanych w bazie LISTA

Celem analizy słów kluczowych użytych do zaindeksowania 82 artykułów opublikowanych w dziewięciu „kanonicznych” czasopismach nauki o informacji (zob. Tab. 2) jest porównanie struktury tematycznej piśmiennictwa badawczego reprezentującego cały obszar nauki o informacji, bibliotekoznawstwa oraz dyscyplin pokrewnych ze strukturą tematyczną publikacji na ten temat, które uznać można za piśmiennictwo najbardziej reprezentatywne dla nauki o informacji. Jak widać na rysunku 17, rozkład słów kluczowych charakteryzujących tematykę artykułów z czasopism „kanonicznych” nieco różni się od rozkładu prezentującego strukturę tematyczną całego zbioru piśmiennictwa o big data wyodrębnionego z bazy LISTA. W głównych czasopismach nauki o informacji problematyka big data podejmowana była przede wszystkim w kontekście bibliometrii (*bibliometrics*) i bibliotek cyfrowych (*digital libraries*). Można stwierdzić, że temat big data



w tej grupie czasopism podejmowany był najczęściej w kontekście zagadnień ilościowych badań informacji (po pięć wystąpień terminów: *altmetrics*, *citation analysis*, *scientometrics*, trzy wystąpienia *impact factor* (*citation analysis*) oraz dwa wystąpienia *h-index* (*citation analysis*)). Niewątpliwie duży wpływ na to ma umieszczenie czasopisma *Scientometrics* w grupie kanonicznych periodyków nauki o informacji. Obecność terminów *data mining* i *social media* wśród wyrazów najczęściej występujących w polu tematu także w tej grupie artykułów uzasadnić należy analogicznie, jak w przypadku całego zbioru publikacji o tematyce big data wyodrębnionego z bazy LISTA. Terminy te wskazują zatem najczęstszy aspekt technologiczny badań big data i najczęstsze źródło danych poddawanych analizie. Warto natomiast zauważyć, że na czasopisma „kanoniczne” przypada większość spośród publikacji, w których problematykę big data podejmowano w kontekście bibliotek cyfrowych (9 spośród 13). W czasopismach „kanonicznych” relatywnie często omawiano też zagadnienia big data w kontekście systemów wyszukiwania informacji – na czasopisma te przypada 35% wszystkich publikacji na ten temat zarejestrowanych w bazie LISTA. Z kolei w kontekście metadanych o badaniach big data w „kanonicznych” czasopismach nauki o informacji pisano dotąd w pięciu artykułach, co stanowi 26% wszystkich publikacji na ten temat zarejestrowanych w bazie LISTA. Do ciekawych obserwacji można też zaliczyć to, że na czasopisma tradycyjnie uznawane za najbardziej reprezentatywne dla nauki o informacji przypada blisko 56% publikacji o problematyce big data dotyczących maszynowego uczenia się oraz 45% publikacji o big data dotyczących wizualizacji informacji.

Wnioski z przedstawionego porównania trzeba jednak traktować z dużą ostrożnością zarówno ze względu na umowność kategorii „kanonicznych” czasopism nauki o informacji, jak i ze względu na relatywnie małą liczbę analizowanych publikacji i duże rozproszenie ich tematyki.



Rys. 17. Struktura tematyczna artykułów naukowych dotyczących zagadnienia big data opublikowanych w „kanonicznych” czasopismach nauki o informacji

## 4. Podsumowanie

Badanie piśmiennictwa zarejestrowanego w bazie LISTA potwierdziło ogólną tezę, że rosnące w nauce zainteresowanie badaniami dotyczącymi analizy wielkich zbiorów danych i jej wykorzystania w rozmaitych zastosowaniach staje się widoczne również w nauce o informacji. Liczba 381 artykułów naukowych poświęconych tej problematyce, które zostały opublikowane w czasopismach naukowych indeksowanych w LISTA, wskazuje jednak, że intensywność tego zainteresowania nie jest jeszcze bardzo duża. Niemniej jednak analiza tematyki tych artykułów potwierdziła też, że problematyka big data łączy się z kluczowymi obszarami badań nauki o informacji.

W nauce o informacji badania dotyczące big data pojawiły się w 2011 r., a więc w tym samym czasie, gdy w całej nauce nastąpił wyraźny wzrost zainteresowania tą problematyką. Podobnie jak w innych dyscyplinach, również w nauce o informacji w ostatnich ośmiu latach następował systematyczny wzrost publikacji poświęconych tej problematyce.

Artykuły, w których podejmowano tematykę big data, ukazały się na łamach aż 113 czasopism należących do szeroko pojmowanych dyscyplin informacyjnych. Równocześnie jednak wyraźnie wyodrębnia się stosunkowo nieduża grupa kilkunastu czasopism, w których w ciągu ostatnich ośmiu lat do tej tematyki powracano wielokrotnie. Są to czasopisma z zakresu: naukometrii i bibliometrii, systemów informacyjnych, informatyki medycznej, zarządzania informacją oraz zaawansowanych technologii bibliotecznych. Dostrzec też można, że o badaniach big data w problemowym obszarze nauki o informacji najczęściej publikują czasopisma specjalizujące się w pewnych jej węższych subdyscyplinach, przede wszystkim ilościowych badaniach informacji (bibliometrii, naukometrii, altmetrii), informatyce medycznej, problematyce systemów informacyjnych i wyszukiwania informacji oraz w zarządzaniu informacją. W czasopismach o szerokim profilu tematycznym publikacje na temat big data ukazują się dość rzadko.

Około 22% artykułów o big data zarejestrowanych w LISTA ukazało się w czasopismach, które w wielu badaniach uznawano dotąd za najważniejsze czasopisma nauki o informacji, najbardziej reprezentatywne dla jej pola badawczego i o wysokim wpływie na rozwój jej badań. W tej grupie najwięcej artykułów o big data opublikowano w: *Scientometrics*, *International Journal of Information Management*, *Library Hi Tech*, *Information Processing & Management* i *Information Research*. Zaskakujący jest fakt, iż w takich czasopismach jak *Journal of the Association for Information Science and Technology* czy *Journal of Documentation*, uznawanych za główne czasopisma nauki o informacji, problematyka big data występowała dotąd sporadycznie.

Zagadnieniami big data w nauce o informacji zajmują się badacze z różnych krajów świata, są wśród nich również badacze polscy. Rozkład geograficzny afiliacji autorów artykułów o big data zarejestrowanych w bazie LISTA na ogół potwierdza z jednej strony, znaną od dawna największą aktywność ośrodków amerykańskich i brytyjskich w badaniach z zakresu nauki o informacji, w tym także związanych z tą nową technologią, a z drugiej – obserwowaną także w innych dziedzinach, gwałtownie rosnącą liczbę badań dotyczących big data prowadzonych w krajach azjatyckich, w szczególności w Chinach i w Indiach.

Piśmiennictwo dotyczące badań big data, które zarejestrowano z bazie LISTA, charakteryzuje duża różnorodność podejmowanej tematyki szczegółowej. Dominuje tematyka należąca do obszaru nauk komputerowych oraz mediów społecznych, ale do zagadnień

często omawianych należą też metadane, zarządzanie i dzielenie się wiedzą, biblioteki cyfrowe, bibliometria, w tym analiza cytowań oraz kwestie związane z informatyką medyczną i ochroną zdrowia. Z kolei w głównych czasopismach nauki o informacji na plan pierwszy wysuwają się badania bibliometryczne oraz dotyczące bibliotek cyfrowych.

Nie ulega wątpliwości, że badania dotyczące technologii big data, możliwości ich wykorzystania w usługach informacyjnych oraz rozmaitych ich aspektów społecznych będą przyciągać coraz większą uwagę badaczy nauki o informacji. Wielkość i złożoność zasobów informacji i wiedzy zgromadzonych w środowisku cyfrowym oraz stale rosnące tempo ich przyrostu wymuszają sięganie po tę nową technologię w działalności, która służyć ma skutecznemu transferowi informacji i wiedzy w społeczeństwie.

## Bibliografia

- Anderson, Ch. (2008). *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete* [online]. Wired, 6.23.08 [18.11.2018], <https://www.wired.com/2008/06/pb-theory/>
- Chang, Y.-W., Huang, M.-H. (2012). A Study of the Evolution of Interdisciplinarity in Library and Information Science: Using Three Bibliometric Methods. *Journal of the American Society for Information Science and Technology*, 63(1), 22–33.
- De Mauro, A., Greco, M., Grimaldi, M. (2016). A Formal Definition of Big Data Based on Its Essential Features. *Library Review*, 65(3), 122–135.
- Feinstein, S. (2012). *Ignorance: How It Drives Science*. New York: Oxford University Press.
- Friedman, A. (2018). Measuring the Promise of Big Data Syllabi. *Technology, Pedagogy and Education* vol. 27, nr 2, 135–148.
- Hey, T., Tansley, S., Tolle, K. (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Redmond, Wash.: Microsoft Research.
- Intel (2012). *Peer Research Big Data Analytics. Intel's IT Manager Survey on How Organizations Are Using Big Data* [online]. Intel IT Center [18.11.2018], <https://www.intel.com/content/dam/www/public/us/en/documents/reports/data-insights-peer-research-report.pdf>
- Jacobfeuerborn, B. (2013). Is Big Data a Paradigm Challenge to Information Science? *Zagadnienia Informatyki Naukowej*, 51 (2), 52–63.
- Klous, S., Wielaard, N. (2016). *We Are Big Data. The Future of the Information Society*. Amsterdam: Atlantis Press.
- Laney, D. (2001). *3D Data Management: Controlling Data Volume, Velocity, and Variety*. [online]. Gartner, file No.949. 6 Feb. 2001 [18.11.2018], [https://www.researchgate.net/publication/311642627\\_Big\\_Data\\_The\\_V%27s\\_of\\_the\\_Game\\_Changer\\_Paradigm](https://www.researchgate.net/publication/311642627_Big_Data_The_V%27s_of_the_Game_Changer_Paradigm)
- Mayer-Schönberger, V., Cukier, K. (2014). *Big data: rewolucja, która zmieni nasze myślenie, pracę i życie*. Warszawa: MT Biznes.
- Patgiri, R., Ahmed, A. (2016). *Big Data: The V's of the Game Changer Paradigm* [online]. IEEE 18th International Conference on High Performance Computing and Communications 2016 [18.11.2018]. [https://www.researchgate.net/publication/311642627\\_Big\\_Data\\_The\\_V%27s\\_of\\_the\\_Game\\_Changer\\_Paradigm](https://www.researchgate.net/publication/311642627_Big_Data_The_V%27s_of_the_Game_Changer_Paradigm)
- Reitz, J. M. (2014) *Online Dictionary for Library and Information Science* [online]. ABC-Clio [5.12.2018], [https://www.abc-clio.com/ODLIS/odlis\\_i.aspx](https://www.abc-clio.com/ODLIS/odlis_i.aspx)
- Sapa, R. (2018). Reinterpretacja koncepcji użytkownika usług informacyjnych. W: *Nauka o informacji w okresie zmian: innowacyjne usługi informacyjne*. Warszawa: Wydaw. SBP, 17–26.
- Saracevic, T. (2010). Information Science. In: *Encyclopedia of Library and Information Sciences. Third Edition*. Boca Raton, FL: CRC Press, vol. 4, 2570–2584, DOI: 10.108/E-ELIS3-120043704

- Scopus (2018). Scopus: An Eye on Global Research [online]. Elsevier [5.12.2018], [https://www.elsevier.com/\\_\\_data/assets/pdf\\_file/0008/208772/ACAD\\_R\\_SC\\_FS.pdf](https://www.elsevier.com/__data/assets/pdf_file/0008/208772/ACAD_R_SC_FS.pdf)
- Sosińska-Kalata, B. (2013). Obszary badań współczesnej informatologii (nauki o informacji). *Zagadnienia Informatologii Naukowej*, 51(2), 9–41.
- Sosińska-Kalata, B. (2017). Kierunki rozwoju współczesnej informatologii. *Forum Bibliotek Medycznych*, 10(2), 25–46.
- Ward, J.S., Barker, A. (2013). Undefined by Data: A Survey of Big Data Definitions [online]. Cornell University Library [18.11.2018], <https://arxiv.org/pdf/1309.5821.pdf>
- White, H.D., McCaine, K.W. (1998). Visualizing a Discipline: An Author Cocitation Analysis of Information Science, 1972–1995. *Journal of the American Society for Information Science*, 49(4), 327–355.
- Zhao, D., Strotmann, A. (2008). Information Science during the First Decade of the Web: An Enriched Author Cocitation Analysis. *Journal of the American Society for Information Science and Technology*, 59(6), 916–937.
- 

## Big Data (Massive Data) in Information Science

### Abstract

**Purpose/Thesis:** The aim of the paper is to discuss main features of the phenomenon known as big data, its importance for the research issues of information science and an attempt to pre-assess the researchers' level of interest in the topic in question.

**Approach/Methods:** The critical analysis of literature has been used to discuss the essence of the big data phenomenon and related changes in the research model, increasingly applicable in various fields of modern science. The growing interest in big data in science is illustrated with the results of a bibliometric analysis of the literature indexed in the interdisciplinary Scopus database. The assessment of the level of interest in big data within the field of information science is based on a bibliometric analysis of the literature indexed in the domain-based EBSCO database – Library and Information Science and Technology Abstracts (LISTA).

**Results and conclusions:** Big data technologies can be treated as a next phase of development in computer technology and its applications in various fields of science and practice. In the environment of large data resources stored in a digital format, big data technologies provide an insight into knowledge that could not be extracted with traditional methods of information retrieval. In this sense, the afore-mentioned technologies support knowledge transfer processes occurring among people and those processes are the main focus of information science. The analysis of the literature indexed in the LISTA database confirmed that the development of big data technology and its applications is a significant challenge for information science and the interest in it is systematically growing, although it has not become very large so far. The analysis of topics of this literature also confirmed that the big data issues are related to the key areas of information science research. Most often big data research is presented in information science journals focused on quantitative information research (bibliometrics, scientometrics, altmetrics), medical computer science, information systems, information retrieval and information management. Journals with a broad thematic profile covering the whole field of information science research have been publishing papers rather rarely so far. The authors of the largest number of articles on big data in information science are affiliated to research centers in the United States, Great Britain and China. The literature on big data research in information science is distinguished with a large diversity of specific topics. Topics that belong to the area of computer science and social media dominate the field, but fairly often researchers also

discuss metadata, management and knowledge sharing, digital libraries, bibliometrics and issues related to medical computer science and health protection.

**Research limitations:** The research discussed in the paper is a preliminary recognition of interest in the big data phenomenon within the field of information science and it was built on the literature indexed in the LISTA database, which subject description includes the term "big data". Hence, the literature presenting the issues related to the study of large datasets where this index term was not used was not included in the study. In addition, the policy of indexing of the LISTA database, in particular the relatively small representation of journals published in languages other than English, may limit the representativeness of the results obtained for big data research related to information science issues on a global scale.

**Originality/Value:** To the best of the author's knowledge, the research presented in the paper is the first attempt to assess the level of interest in big data within the field of information science.

### Keywords

Bibliometric study. Big data. Information science. Massive data. Research problems.

---

*Prof. dr hab. BARBARA SOSIŃSKA-KALATA jest kierownikiem Katedry Informatologii na Wydziale Dziennikarstwa, Informacji i Bibliologii Uniwersytetu Warszawskiego oraz redaktor naczelną czasopisma Zagadnienia Informacji Naukowej – Studia Informacyjne i członkiem Komitetu Naukoznawstwa PAN. Specjalizuje się w problematyce nauki o informacji, w szczególności organizacji wiedzy, a także ilościowych badań informacji oraz historii, teorii i metodologii nauki o informacji. Opublikowała ponad 250 prac, w tym 12 książek i ponad 160 artykułów naukowych. Do jej najważniejszych publikacji należą monografie Modele organizacji wiedzy w systemach wyszukiwania informacji o dokumentach (1999) i Klasyfikacja. Struktury organizacji wiedzy, piśmiennictwa i zasobów informacyjnych (2002) oraz artykuły Évolution des Systèmes d'Organisation des Connaissances et établissement de critères pour leur évaluation praxeologique (2011), Obszary badań współczesnej informatologii (nauki o informacji) (2013), Nauka o informacji wśród nauk o kulturze (2017), Książka (dokument) w środowisku informacyjnym (2017), The impact of the works by Paul Otlet and Suzanne Briet on the development of the epistemology of documentation and information science in Poland (2018).*

*Kontakt z autorką:*

*b.sosinska@uw.edu.pl*

*Katedra Informatologii*

*Wydział Dziennikarstwa, Informacji i Bibliologii*

*Uniwersytet Warszawski*

*ul. Nowy Świat 69*

*00-046 Warszawa*

---

# Fotografia w bibliotece i w sieci: dwa podejścia do reprezentacji i wyszukiwania

Kamil Stępień

*Institut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Marii Curie-Skłodowskiej w Lublinie*

Jadwiga Woźniak-Kasperek

*Katedra Bibliografii i Dokumentacji, Wydział Dziennikarstwa, Informacji i Bibliologii  
Uniwersytet Warszawski*

---

## Abstrakt

**Cel/Teza:** Celem jest przedstawienie dwóch podejść do reprezentowania i wyszukiwania fotografii – tradycyjnego, metadaneowego, specyficznego dla bibliotek z automatycznym, sieciowym. Rozważania zogniskowane są na możliwości ich wzajemnego uzupełniania się w celu stworzenia nowej jakości informacyjno-wyszukiawczej w wielokanałowej komunikacji międzyludzkiej.

**Koncepcja/Metody badań:** Wykorzystano metodę analizy i krytyki piśmiennictwa.

**Wyniki i wnioski:** Optymalnym rozwiązaniem problemów reprezentacji i wyszukiwania fotografii byłby dziś system hybrydowy łączący opis tworzony przez profesjonalistę z treściami generowanymi oddolnie przez internautów oraz z rezultatami komputerowej analizy obrazu. Opis obiektu wizualnego tworzony równolegle na różnych płaszczyznach pozwoliłby na zmniejszenie ryzyka powstawania opisów błędnych, niezrozumiałych i niekompletnych oraz na wzrost możliwości wyszukiwania obiektów wizualnych z uwzględnieniem zindywidualizowanych potrzeb użytkowników.

**Oryginalność/Wartość poznawcza:** Artykuł uzupełnia niedostatecznie reprezentowaną w polskiej literaturze bibliologiczno-informatologicznej problematykę komputerowej analizy obrazu, zwłaszcza fotografii, i jej użyteczności informacyjno-wyszukiawczej.

## Słowa kluczowe

Biblioteka. Fotografia. Komputerowa analiza obrazu. Reprezentacja. Wyszukiwanie.

*Otrzymany: 14 lutego 2019. Zrecenzowany: 28 lutego 2019. Zaakceptowany: 14 marca 2019.*

---

## 1. Wstęp

Środowisko wyszukiwawcze Internetu w znacznej części stanowią obrazy, treści wizualne zestawione na ogół z mniejszą lub większą porcją tekstu. „Zestawienie tekstu i ilustracji składa się na przekaz o «podwójnej konceptualizacji»” (Tabakowska, 2006, 10). Pewną część ikonosfery sieciowej stanowią cyfrowe fotografie, w tym również cyfrowe kopie fotografii przechowywanych w instytucjach takich jak biblioteki, wraz z ich metadanymi. W artykule chcemy porównać na przykładzie fotografii dwa „paradygmaty” reprezentowania i wyszukiwania obrazów – biblioteczny, „tradycyjny”<sup>1</sup>, bazujący na metadanych z sieciowym,

---

<sup>1</sup> W artykule ograniczamy perspektywę do sfery bibliotek, zdając sobie sprawę, że nie tylko one dysponują zasobami fotografii oraz że podobnie postępuje się z fotografiami w innych instytucjach należących

możliwym do stosowania także w bibliotekach, choć na pewno nie bez problemów i łamania pewnych schematów. Zawarte w artykule treści dotyczące komputerowego rozumienia obrazu i wyszukiwania obrazem nie mają charakteru poradnika implementacyjnego, ale tworzą podstawową ramę konceptualną, z której można wywieść wnioski praktyczne.

Odwołując się do języka specjalistycznego bibliotekarzy i pracowników informacji, można powiedzieć, że w artykule będziemy skupiać uwagę na determinantach możliwości wyszukiwania (ang. *retrieval*) fotografii oznaczającego „kolejne, indywidualne realizacje wyszukiwania (...), obejmujące ciąg interakcji jednostki poszukującej informacji z systemem informacyjnym, zarówno na poziomie manipulowania systemem, jak i na poziomie intelektualnym, tj. tworzenia instrukcji wyszukiwawczej, ustalania kryteriów selekcji wyszukanej informacji oraz dotyczącym oceny trafności wyszukanej informacji” (Wilson, 2000, 49)<sup>2</sup>, a także „odnajdywania” (ang. *seeking*) będącego procesem mniej sformalizowanym i ustrukturyzowanym, ale coraz bardziej spersonalizowanym i złożonym, angażującym technikę komputerową. W praktyce sieciowej szukający płynnie przechodzi od pytania przez przeglądanie, filtrowanie, po wyszukiwanie, by wrócić do pytania itd. Wyszukiwanie stwarza bowiem szansę na lepsze zrozumienie problemu i ewentualne powtórzenie czynności. Można zaobserwować dynamiczny rozwój łączenia sposobów wizualizacji treści z samym wyszukiwaniem, co daje pożądane rezultaty nie tylko w postaci bogatego zestawu hiperłączy, ale także połączeń z innego rodzaju treściami, takimi jak: multimedia, obrazy, tłumaczenia czy dane geograficzne. Jesteśmy przekonani, że metadane słowne (opisy, reprezentacje przedmiotowe, rzeczowe) są nadal niezbędnym elementem środowiska wyszukiwawczego, ale im poświęcono relatywnie dużo publikacji, na które zwracamy uwagę zainteresowanych czytelników (Baca 2002; Baca 2006; Beghtol 1986; Clough, Ireson, Marlow, 2009; Jörgensen, 2003; Jörgensen et al., 2001; Jörgensen, 1999; Jörgensen, 1998; Rygiel, 2012; Shatford-Layne, 1994; Shatford, 1986; Svenonius, 1994; Włodarczyk & Woźniak-Kasperek, 2017).

## 2. Fotografia w bibliotece – poza tradycyjne kategorie

Dokumenty ikonograficzne, a wśród nich fotografie, zgromadzone w bibliotekach, w muzeach, w archiwach i innych instytucjach kultury i informacji są opracowywane i przygotowywane do wyszukiwania zazwyczaj w podobny, „konwencjonalny” sposób, właściwy dla danego środowiska. Zasoby te stanowią nadal nie w pełni doceniane źródło (Talbierska 2009), mimo że przekaz za pomocą kodu wizualnego pozwala na transmisję tych treści, które są niemożliwe lub trudne do wyrażenia w inny sposób, co ma określoną wartość poznawczą, a mieści się w sferze zainteresowania różnych dziedzin wiedzy i dyscyplin naukowych, m.in. historii sztuki, socjologii, etnologii, architektury czy urbanistyki. Specyficzne dla bibliotek podejście do opracowania fotografii sprawia, że ich wyszukiwanie może być realizowane poprzez odwołanie jedynie do kategorii przyjętych na etapie tworzenia reprezentacji.

---

do klasy GLAM (Galleries, Libraries, Archives, Museums). Przyjeliśmy również założenie, że sama zmiana kodu z analogowego na cyfrowy nie zmienia statusu fotografii, inaczej mówiąc, zdigitalizowanie fotografii i umieszczenie jej w bibliotece cyfrowej nie czyni z niej automatycznie fotografii sieciowej.

<sup>2</sup> Wszystkie cytaty z prac nieopublikowanych dotychczas w języku polskim są tłumaczeniami autorów niniejszego artykułu [red.].

Analizę i opracowanie rzeczowe fotografii w bibliotece można zrealizować na kilka sposobów. Szczegółowy opis procedury analizy i tworzenia opisu zawiera m.in. publikacja B. Włodarczyka i J. Woźniak-Kasperek (Włodarczyk & Woźniak-Kasperek, 2017). W rezultacie powstają opisy zawierające odpowiedzi na pytania o: przedmiot główny widoczny na fotografii, przedmioty równorzędne i poboczne, własności przedmiotów, w tym lokalizację przestrzenną (geograficzną) i czasową, ich ujęcie (punkt widzenia), adresatów, formę (np. fotografia czarno-biała polska), czy postać fizyczną (np. odbitka srebrowo-żelatynowa). W przypadku zasobów o dużym wolumenie podejście takie może powodować generowanie tak licznych odpowiedzi, że wyszukiwanie właściwie należy uznać za nieudane. Użytkownik nie jest bowiem w stanie zapoznać się z całością otrzymanych wyników, a nie ma do dyspozycji narzędzi umożliwiających ich ograniczenie. Z różnych powodów (brak możliwości głębszej, poprawnej i kompletnej interpretacji poszczególnych warstw treści fotografii, wiedzy o nośniku i technikach utrwalenia, niedostatków stosowanych słowników kontrolowanych lub niekontrolowanych itd.) nie jest bowiem możliwa rafinacja odpowiedzi. Z pomocą mogłyby tu przyjść techniki automatycznego rozpoznawania i wyszukiwania obrazu obrazem (wyszukiwanie odwrotne), o czym będzie mowa w dalszych częściach artykułu.

Digitalizacja i prezentowanie cyfrowych kopii i cyfrowych oryginałów kolekcji muzealnych, bibliotecznych czy archiwalnych stały się w ostatnich latach powszechną praktyką. Obecność dokumentów ikonograficznych w środowisku sieciowym sprawia, że z jednej strony, uaktualniane są w duchu nowych możliwości i potrzeb odbiorców praktyki transformacji danych i prezentowania informacji o nich oraz sposób ich utrwalenia, z drugiej zaś, wprawdzie jeszcze nieśmiało, ale instytucje te zaczynają czerpać z oferty zaawansowanych technologii cyfrowych. Cyfryzacja zasobów oraz umieszczenie ich w internecie miały wpływ na zmiany w strukturze i prezentacji danych opisu. Nadal jednym z nierozwiązanych problemów jest jednak problem takiego języka opisu rzeczowego, systemu organizacji wiedzy, ontologii itp., które odpowiadałyby potrzebom informacyjnym różnych odbiorców (indywidualnych, ale też instytucjonalnych) oraz specyfice obrazu. Brakuje systemowego podejścia do opisu przedstawień wizualnych, w tym tu nas interesujących fotografii, które sprzyjałoby konsekwencji w tworzeniu poprawnych i kompletnych reprezentacji fotografii, a co za tym idzie, eliminowaniu błędów i trudności wyszukiwawczych. Niezbędna jest rewizja dotychczasowych rozwiązań, wsparta pogłębioną refleksją nad sposobami tworzenia i zarządzania metadanymi opisowymi oraz wzbogacenie repertuaru tradycyjnych metod i środków o te, które są dostępne dzięki technologiom przetwarzania obrazów cyfrowych.

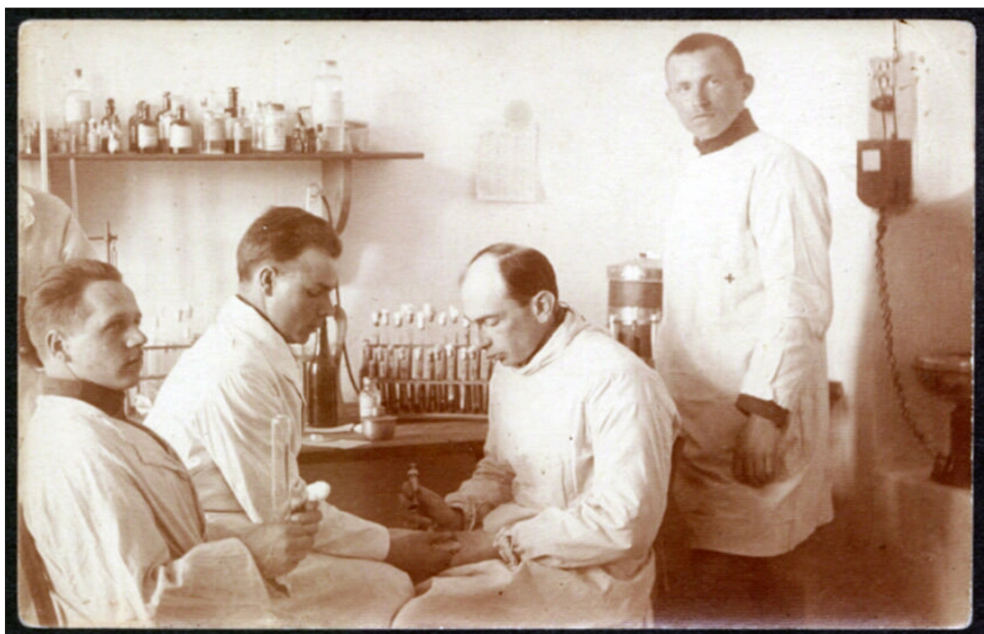
Zbiory ikonograficzne są cennym, choć nie priorytetowym, typem materiałów gromadzonych w bibliotekach, gdzie najczęściej są przechowywane, opracowywane i udostępniane w osobnych działach, jako specjalistyczne kolekcje na przykład o charakterze artystycznym (grafika, rysunek, fotografia). Mogą też wchodzić w skład większych kolekcji, np. dokumentów życia społecznego. „Jeszcze do niedawna zbiory fotograficzne były często lekceważone. Fotografie katalogowano zbiorczo, bardziej jako kolekcje, czasem w ogóle nie opisywano” (Miller & Wombard, 2009, 202). Przechowywano je luzem, w pudełkach, teczkach, bez szczegółowego komentarza. Negatywy fotograficzne – w kopertach czy tekturowych pudełkach „przechowywano [...] razem z dokumentami papierowymi, a jeszcze w latach 80. XX w. twierdzono czasem, że w ogóle nie są archiwaliami” (Miller & Wombard, 2009, 222). Ich tradycyjny opis, dotyczący zawartości, treści przedstawień,



zawsze stwarzał w środowisku bibliotecznym duże problemy, ale też trzeba przyznać, że nie poświęcono im tyle uwagi co dokumentom piśmienniczym. Fotografie w zbiorach bibliotek, archiwów i innych instytucji informacji i kultury są opisywane i porządkowane stosownie do przyjętych zwyczajów i standardów, co nie oznacza, że ich odszukanie jest zawsze łatwe i skuteczne. Przyczyn tego stanu rzeczy jest kilka. Jedną, jak wcześniej wspomnieliśmy, jest preferowanie dokumentów piśmienniczych kosztem dokumentów innych typów, skutkiem czego mechanizmy i narzędzia indeksowania i wyszukiwania dokumentów piśmienniczych są lepiej rozpoznane, bogatsze i skuteczniejsze. Trudności nastręcza sama praca z dokumentem wizualnym, poznawanie go, rozpoznawanie, interpretowanie. Niejednokrotnie brakuje danych tekstowych, dokumentów towarzyszących, zapisków na odwrociach itp., które pomogłyby odpowiedzieć na pytania kluczowe dla opracowania zasobu, brakuje także klarownych instrukcji regulujących sposób tworzenia opisów i generowania metadanych.

Do najważniejszych zadań bibliotekarza, który ma do opracowania fotografię bez żadnego tekstu, należą: identyfikacja i datowanie, określenie miejsca, przedmiotów widocznych na fotografii, zidentyfikowanie portretowanej osoby, autora fotografii itp. Wymaga to szerokiej wiedzy. Aby mieć pewność, że przypuszczenia odnośnie do datowania są słuszne, należy znać techniki i technologie wykonywania fotografii, szczególnie jeżeli mamy do czynienia z fotografią starą, przydatna jest także znajomość historii sztuki, m.in. stylów i panującej wówczas mody, niezbędna jest umiejętność odróżnienia „czystej” fotografii od jej drukowanej wersji itd. Opracowanie fotografii oprócz analizy i nazwania elementów przedstawienia, które się nań składają (ang. *ofness*), oraz „cech fizycznych dokumentu, (...), czym jest zasób, dzieło, jaka jest jego forma” (Rygiel, 2012, 78) (ang. *isness*, co jest bliskie kategorii gatunek/forma), może zawierać odpowiedź na pytanie o symbolikę, o to czego dotyczy, o czym jest, jakie niesie przesłanie itp. (w literaturze anglojęzycznej ta ścieżka opracowania nosi nazwę *aboutness* i eksplorujemy ją, opracowując wiele innych typów dokumentów). „Pojęcia *aboutness* i *ofness* stosowane w opisie obrazów oparte są na teorii Panofsky’ego, dotyczącej trzech warstw znaczeniowych dzieła sztuki oraz koncepcji OF i ABOUT Shatford” (Rygiel, 2012, 78–79). Najniższy poziom, pre-ikonograficzny, idąc śladem Panofsky’ego, lub *ofness* według Shatford tworzą zidentyfikowane i nazwane elementy opracowywanego przedstawienia. Poziom ikonograficzny wymaga interpretacji wyodrębnionych elementów *ofness* i ich konfiguracji z uwzględnieniem wiedzy przedmiotowej oraz konwencji i kodów artystycznych, społecznych, kulturowych. Trzeci poziom, ikonologiczny, wymaga rozpoznania warstwy symbolicznej przedstawienia, zrekonstruowania programu ikonograficznego, co na ogół dla osoby niemającej odpowiedniej wiedzy i doświadczenia może być zadaniem niewykonalnym. Podejście ikonograficzne zakłada czytanie sztuki, w tym przypadku fotografii, czyli interpretowanie wizualnej reprezentacji, przez umieszczanie jej elementów w tradycji, która daje im znaczenie inne niż to, które sugeruje ich wizualny wygląd (Bał, 1991, 177). Zakładając, że etap pracy z dokumentem zostanie zrealizowany z sukcesem, kolejne czynności związane z wyrażeniem treści z wykorzystaniem stosowanego w systemie słownika nie różnią się od tych podejmowanych przy opracowaniu dokumentów piśmienniczych, zakładając, że słownik jest dostatecznie nasycony relewantnym słownictwem. W dalszej części artykułu wrócimy do tych zagadnień, jako że automatyzacja wprawdzie nie gwarantuje bezwzględnie lepszych rezultatów, ale pozwala je wzbogacić i przyspieszyć, co też nie jest bez znaczenia.

Wspomniany wcześniej amerykański historyk sztuki Erwin Panofsky wydzielił trzy poziomy analizy dzieła sztuki: pre-ikonograficzny, ikonograficzny i ikonologiczny. Podział ten można zastosować podczas analizy przedmiotowej. Bibliotekarz opracowujący zdjęcie w modelu tradycyjnym skupia się na dwóch pierwszych poziomach, pomijając znaczenie symboliczne przedstawienia. Uwzględnienie poziomu pre-ikonograficznego pozwala opisać kategorie obiektów przedstawionych na fotografii. Jednak ich poprawne nazwanie jest możliwe dopiero po identyfikacji indywidualnych osób, ich kategorii, obiektów itd. Ideałem byłoby określenie znaczenia symbolicznego przedstawienia, o ile w konkretnym przypadku można o nim mówić, a co w bibliotekach ma miejsce niezwykle rzadko. Weźmy jako przykład poniższą fotografię (Rys. 1) z biblioteki cyfrowej Polona i jej opis, w którym znajdziemy pewne elementy z kategorii *ofness* i *isness*. Autor opracowania nadał fotografii tytuł (Wnętrze laboratorium szpitala wojennego w Kielcach), zrezygnował z próby jej interpretacji symbolicznej, a nawet z głębszej analizy warstwy *ofness*. Charakterystyka wyszukiwawcza dokumentu sporządzona w bibliotece cyfrowej Polona składa się z następujących haseł – słów kluczowych: *Kielce (woj. świętokrzyskie); Lekarze; Szpitale wojskowe*, datacji powstania fotografii: *1914–1918*, określenia gatunku: *Odbitka na papierze srebrowo-żelatynowym; Fotografia polska* oraz opisu fizycznego, tj. *1 fot. : odb. na pap. srebrowo-żelatynowym; 8,9x13,9 cm*. W rezultacie również wyszukiwanie jest możliwe jedynie na poziomie tak wyodrębnionych kategorii. Poniżej przedstawiamy tagi przypisane tej fotografii przez grupę internautów, studentów kierunku architektura informacji.



Rys. 1. Wnętrze laboratorium szpitala wojennego w Kielcach [1914–1918].

Źródło: Polska Biblioteka Cyfrowa Polona [online], <https://polona.pl/item/wnetrze-laboratorium-szpitala-wojennego-w-kielcach,N7C4NjkkOA/0>, [30.01.2019]

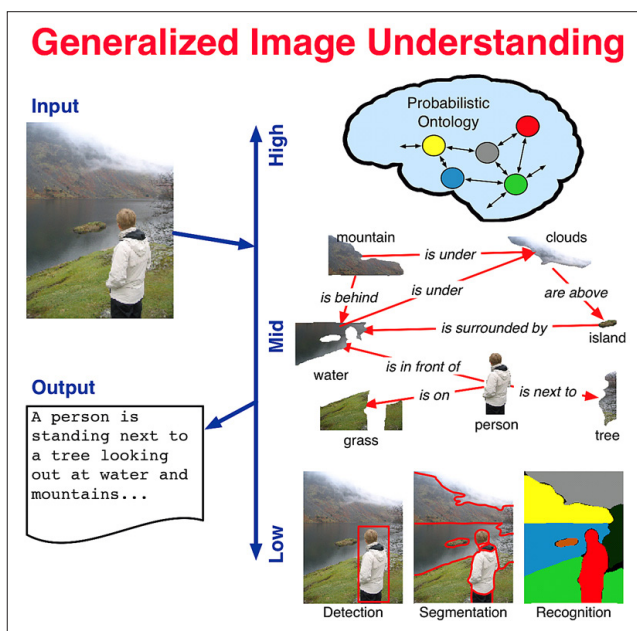
Tab. 1. Zestawienie tagów użytkowników z opisem w bibliotece Polona dla fotografii  
*Wnętrze laboratorium szpitala wojennego w Kielcach [1914–1918]*

Uporządkowane alfabetycznie tagi, nadane przez internautów	Opis w Bibliotece Cyfrowej Polona
<p>ampułki, apteczka, apteka, archiwum, badania, badanie, badanie krwi, badanie lekarskie, białe fartuchy, biały fartuch, biały kitel, biel, biologia, blat, butelka, chemia, chemicy, chemikalia, chirurdzy, choroba, choroby, chory, ciekawość, czarno białe, czarno-biała, czarno-białe, czerwony krzyż, czysto, dawna medycyna, dawne czasy, dawne leczenie, dawno temu, dezynfekcja, doktor, doktorzy, doświadczenie, eksperyment, eksperymenty, faceci w bieli, farmaceuci, fartuch, fartuchy, fiolka, fiołki, fotografia, gabinet, gabinet lekarski, historia, hospital, igła, kitel, kitel lekarski, kitle, kitle lekarskie, krew, kwarantanna, laboratorium, laboratorium medyczne, lecznictwo, lek, lekarstwa, lekarstwo, lekarz, lekarze, lekcja medycyny, leki, ludzie, medycy, medycyna, mężczyzna, mężczyźni, młodzi ludzie, mundur, mundur lekarski, nauka, naukowcy, odbitka, odczynniki, oddawanie krwi, odwaga, opatrunek, operacja, organizm, pacjenci, pacjent, pielęgniarz, pielęgniarze, pierwsza pomoc, pipeta, pipety, pobieranie, pobieranie krwi, pomieszczenie, pomoc, półka, praca, pracownia, pracownicy, praktyki, próbówka, próbówki, próbki laboratoryjne, przyrządy medyczne, ramka, retro, roztwór, sepia, skupienie, spirytus, stare kitle lekarskie, stare narzędzia, stare zdjęcie, staroświeckie, sterylność, stetoskop, stół, strzykawka, strzykawki, studenci, studenci medycyny, substancja, szczepienie, szczepionka, szczepionki, szkolenie, szkoła, szpital, ściana, termometr, testy, ukłucie, uniform, uzdrowienie, vintage, wkłucie, wojsko, zabieg, zafascynowanie, zajęcia, zastrzyk, zdjęcie, zdrowie</p>	<p>adres wydawniczy: [1914–1918]  data powstania dokumentu: [1914–1918]  opis fizyczny: 1 fot.: odb. na pap. srebrowo-żelatynowym ; 8,9x13,9 cm  gatunek: odbitka na papierze srebrowo-żelatynowym  fotografia polska  forma i typ: Fotografie  słowa kluczowe: Szpitale wojskowe  Lekarze  Kielce (woj. świętokrzyskie)  uwagi: Tyt. nadany przez katalogującego.  sygnatura: E.82315  prawa do utworu:  Domena Publiczna. Wolno zwielokrotnić, zmieniać i rozpowszechnić oraz wykonywać utwór, nawet w celach komercyjnych, bez konieczności pytania o zgodę. Wykorzystując utwór należy pamiętać o poszanowaniu autorskich praw osobistych Twórcy.  źródło: Biblioteka Narodowa  rekord bibliograficzny: zobacz w katalogu  uid repozytorium: 5786918</p>

### 3. Fotografia w sieci – rozpoznawanie obrazu, kształt, barwa, jasność

Odczytywanie każdego tekstu, także wizualnego, jest wielopoziomowe, złożone, a gdy czyni to człowiek, dodatkowo mniej lub bardziej subiektywne, interpretatywne, uczestniczące. Inaczej dzieje się w przypadku technik komputerowego rozpoznawania obrazów,

dla których trudnością jest nie identyfikacja kształtów, barw, tekstury itp. czy nazwanie rozpoznanych elementów, ale odkrycie powiązań (relacji, zależności) pomiędzy nimi, zinterpretowanie kontekstu różnych poziomów. Komputer rozpoznaje obiekty w sposób szczegółowy, ale nie zawsze poprawnie je identyfikuje i kategoryzuje (uogólnia). Problemem jest efektywność semantyki wizualnej, łączącej wiedzę językową i pozajęzykową z rozpoznawaniem wzorców przez komputer. Semantyka wizualna ma wesprzeć człowieka w opisywaniu i wyszukiwaniu obrazów, ale na tym etapie jeszcze nie zastąpi ludzkiej inteligencji, choć w pewnych sytuacjach opis wykonany przez człowieka jest nieosiągalnym luksusem dodatkowo naznaczonym zawieszeniem w sieciach znaczeń, które osoba ta sama utkała (Geertz, 2005, 19). Trudności, zarówno dla komputera, jak i dla człowieka, pojawiają się przy podejmowaniu prób interpretacji przedstawienia. Czy osoba na fotografii poniżej (Rys. 2) jest turystą/turystką, czy może właścicielką/właścicielem schroniska z niepokojem spoglądającym na spowite mgłą góry, a może jeszcze kimś innym? Czy jest to zwykłe zdjęcie, czy może zapis chwili przed lub po bardzo ważnym wydarzeniu? I tak dalej.



Rys. 2. Zrozumieć obraz. Źródło: Corso, 2014

Większość operacji w ramach widzenia komputerowego dotyczy rozpoznawania wzorców (ang. *pattern recognition*), czyli identyfikacji struktur w obrębie obrazu. Komputerową analizę obrazu można zatem sprowadzić do takich czynności jak:

- analiza ogólnych cech obrazu bądź jego fragmentu – analiza cech ilościowych (próbkowanie, pomiary kolorów);
- ekstrakcja cech – odseparowanie od siebie kształtów i obszarów w warstwie wizualnej dzięki zbadaniu krawędzi obiektów i ich struktury geometrycznej;
- segmentacja – podział obszaru wizualnego na spójne obszary o podobnych cechach, takich jak kolor, kształt;

- analiza tekstur – wydzielenie z obrazu charakterystycznych i rozpoznawalnych struktur takich jak np. struktura drewna, bruku;
- wydzielenie i śledzenie obiektów semantycznych, tj. takich, które posiadają złożone znaczenia, w których są wykazywane relacje zależności, np. dom, człowiek, samochód;
- śledzenie zmiany położenia tych obiektów ruchem własnym (przemieszczanie się) – przydatne w przypadku badania sekwencji obrazów statycznych, bądź cyfrowych obrazów animowanych typu GIF;
- reprezentacja obrazów i opis kształtów – sporządzanie syntetycznych opisów obszarów wydzielonych podczas procesu segmentacji: opisy kształtów, obiektów płaskich i przestrzennych;
- analiza sceny i zdarzeń – określanie zależności czasowo-przestrzennych między obiektami, analiza ruchu obiektów, klasyfikowanie ujęć i zdarzeń (Domański, 2010, 24).

Współczesne zdjęcia cyfrowe składają się z kilkudziesięciu milionów pikseli. „Porównywanie wszystkich, piksel po pikselu nawet przy obecnych mocach obliczeniowych komputerów jest zadaniem nierealnym” (Staszczak, 2010). Dlatego stosuje się rozwiązania upraszczające, polegające na określaniu tzw. wektorów cech. „Podczas wyszukiwania podobnego zdjęcia proces ten rozpoczyna się od wygenerowania wektora cech na podstawie wybranego identyfikatora, porównania go z innymi wektorami cech obrazów umieszczonych w bazie, a następnie pogrupowania i zobrazowania wyników” (Staszczak, 2010). W praktyce proces wyszukiwania danych graficznych w zbiorach obrazów odbywa się przez odpowiednie indeksowanie wektorów. Indeks w tym przypadku jest cechą obiektu o relatywnie małym stopniu złożoności. Indeksy tworzone są wprost z danych obrazowych, przez wybranie ich określonych własności całościowych bądź fragmentarycznych przy użyciu funkcji matematycznych. „Bardzo często indeks zawiera informacje, które nie są wprost reprezentowane w danych” (Forczmański, 2005). Wprowadzono rozróżnienie indeksów na niski i wysokopoziomowe (opisowe). Pierwsze z nich dotyczą informacji graficznej o cechach percepcyjnych i fizycznych, takich jak kształt, kolor, tekstura itp. Indeksy wysokopoziomowe są z kolei hasłami kategoryzującymi, tworzonymi na podstawie zawartości i metadanych.

Segmentacja obrazu przez wydzielenie w nim charakterystycznych cech obiektów skutkuje powstawaniem deskryptorów punktowych, krawędziowych i liniowych. Najpopularniejszym sposobem reprezentacji obiektów na obrazie jest ich kształt: kontur i krawędź. Rozwiązanie to sprawdza się bardzo dobrze wtedy, „gdy rozpoznawaniu podlegają konkretne, izolowane obiekty, m.in. w analizie obrazów medycznych (np. kształty komórek na obrazach mikroskopowych lub organy wewnętrzne ciała ludzkiego na obrazach tomograficznych) i optycznym rozpoznawaniu znaków (OCR)” (Forczmański, 2005). W tej sytuacji wszystkie inne cechy charakterystyczne są pomijane. Kontrast, jasność i barwa nie są tak stabilnymi cechami jak kształt. Kształt, którego odseparowane elementy przybierają postać binarną, składającą się z konturu i wnętrza są odporne na częściowe zasłonięcie obiektu, złej jakości odwzorowanie jego krawędzi czy szумы wynikające z niedoskonałości urządzeń rejestrujących obrazy. Najlepszym z rozwiązań jest „samouczenie się” algorytmów rozpoznających kształty, począwszy od tych najprostszych, płaskich figur typu elipsa, prostokąt, przez modele trójwymiarowe, a na odstępstwach i wariacjach bazujących na elementach skomplikowanych skończywszy. Rozwiązania tego typu są już częściowo wdrażane. Wszystko to dotyczy jednak warstwy logicznej obrazu, nie jego treści, interpretacji.

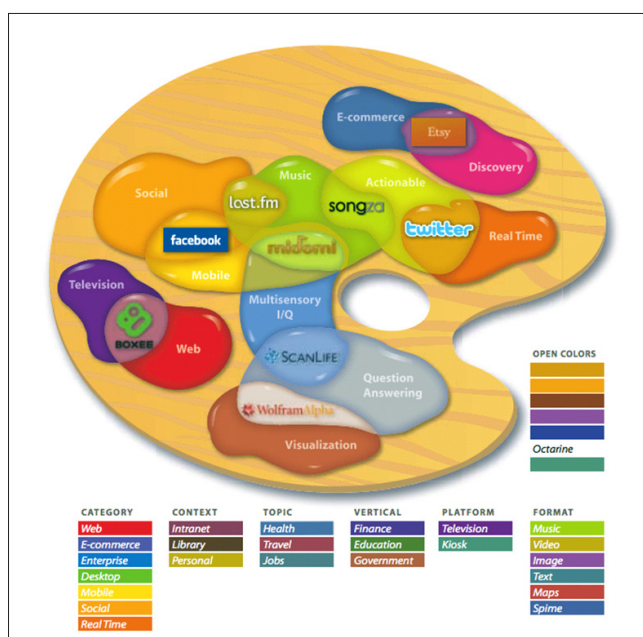
Deskrytory kształtu są wrażliwe na obroty, przesunięcia i zniekształcenia. Dodatkowo, mają niską złożoność obliczeniową. Tą metodą wyszukiwane są kształty obiektów, podobne lub takie same, jak założone na wejściu. Algorytm bada obraz na poziomie pikseli, wyodrębniając kontury przedmiotów za pomocą różnic w przejściach kolorystycznych (gradientów). Dzięki tej metodzie rozpoznawane obiekty charakteryzują się określonym kształtem, rodzajem tego kształtu, ale nie wielkością. „Zbieramy wszystkie właściwości, które oddają rzucające się w oczy szczegóły geometryczne obrazu” (Gevers & Smeulders, 2003). Za pomocą *shape searching* możliwe było stworzenie programów rozpoznających tekst ze skanowanych dokumentów (oprogramowanie typu OCR umożliwiające rozpoznawanie znaków w plikach graficznych). Oczywiście jest, że, aby rozpoznawanie miało określoną skuteczność, trzeba zdefiniować parametry, takie jak: alfabet znaków wejściowych do rozpoznawania, znaki nieliterowe, a także możliwość odróżnienia elementów w postaci grafik, tabel czy wzorów matematycznych. Skuteczność pracy tego typu oprogramowania ciągle rośnie. Na bazie tej technologii powstają zaawansowane rozwiązania, m.in. takie, jak: aplikacja *Retrievr*, *MindFinder* czy wyszukiwarka w serwisie *Modista.com*.

Segmentacja z kolei bazuje na wyborze reprezentatywnych fragmentów obrazów celem ich porównywania z wcześniej wybranym wzorcem. Chodzi tutaj o stworzenie języka obrazu, definiowanie którego opiera się na słowniku wizualnym (ang. *visual vocabulary*). Słownik wizualny ma za zadanie określać autonomiczne wizualnie figury (fragmenty obrazu), które zostają nazwane *wzorcami wizualnymi*. Elementy reprezentatywne są wybierane w trzech płaszczyznach: horyzontalnej, diagonalnej i wertykalnej. Na podstawie tak utworzonych wzorców, ich powiązań i relacji algorytm jest w stanie rozpoznawać kształty bardziej złożone i w konsekwencji odpowiadać na zapytanie wizualne. Słownik wizualny nie odwołuje się do słowa pisanego, skojarzeń czy wyobrażeń o obrazach.

Wdrożeniami tego typu systemów zainteresowały się w szczególności nauki biomedyczne i techniczne. Powstają komercyjne rozwiązania wśród setek tysięcy obrazów odnajdujące te, w których występują zadane fragmenty. W naukach medycznych powstał na przykład system *ViVo* (Bhattacharya, 2005), który jest w stanie rozpoznawać kształty poszczególnych narządów, a także porównywać ich modele trójwymiarowe ze zdjęciami z interwencji lekarskich. Innym przykładem może być wyszukiwanie zamiennych części samochodowych po wskazaniu zdjęć, na których widnieją jedynie kształty albo też fragmenty kształtów konkretnych podzespołów samochodowych. Z racji przepisów prawnych części te nie mogą mieć zastrzeżonych, oryginalnie fabrycznych oznaczeń producentów.

Ważnym elementem obrazu, który ma odzwierciedlenie w jego reprezentacji słownej, wektorowej lub innej, jest barwa. „W wyniku badań stwierdzono, iż w sytuacji wystarczającego oświetlenia człowiek w pierwszej kolejności reaguje na jasność i kolor, następnie na kształt i ruch, a na końcu na teksturę i inne cechy” (Forczmański, 2005). Deskryptorami graficznymi w wyszukiwarkach wizualnych są dominanty barwne, histogramy i modele (przestrzenie) barw. Są trzy możliwości: wyszukiwanie kolorami, które łączymy z nastrojem, wyszukiwanie kolorem, który występuje w poszukiwanym obrazie oraz identyfikacja barwy z elementem wyszukiwania. W przypadku wyszukiwania kolorem, który występuje w poszukiwanym obrazie, narzędziem do wyszukiwania staje się cyfrowa paleta barw, oparta na technikach procentowego zliczania występowania wybranych kolorów w obrazach. Proces wyszukiwania barw polega na wskazaniu z podstawowej palety barw koloru, ewentualnie kolorów, których szukamy w bazie obrazów. Co więcej, niektóre narzędzia

dopuszczają, aby użytkownik sam określił procentowy udział każdego z kolorów, które chce mieć w wyszukany obraz. Takie prototypowe rozwiązanie stworzyła na przykład grupa inżynierów z Toronto. Firma TinEye uruchomiła projekt Multicolor Search, który korzysta z obrazów udostępnionych na licencjach Creative Commons z serwisu Flickr. Wyszukiwarka kolorów ma specjalnie zaprojektowany algorytm, współpracujący z większością formatów plików graficznych dostępnych w Internecie. Są to m.in. JPEG (zamiennie można używać JPG), PNG i GIF. Paleta barw RGB umożliwiającą przeszukiwanie bazy obrazów jest automatycznie przeliczana na wartości szesnastkowe (HEX). W procesie indeksacji obrazów nie jest używana żadna instrukcja znakowania kolorów słowami kluczowymi. Narzędziem wyszukiwawczym jest bezpośrednio wybrany przez użytkownika kod koloru. Przeszukiwany obraz posiada specjalne informacje wizualne, charakterystyczne tylko dla niego. Zbiór tych danych twórcy projektu porównują do niepowtarzalnych linii papilarnych na ludzkich dłoniach. Ta niepowtarzalność odnosi się do sposobu rozmieszczenia wybranych barw i ich odcieni na wyszukiwanych obrazach. Kolejnym krokiem w procesie wyszukiwawczym jest tworzenie mapy kolorów. System określa kody poszczególnych kolorów wraz z ich procentowym udziałem w całym obrazie. Na tym etapie użytkownik ma możliwość ingerencji w wyniki końcowe. Można uwzględnić, albo też usunąć, kolor dominujący powiązany z tłem we wczytywanym obrazie. Dzięki takiemu rozwiązaniu w fazie końcowej system bardzo szybko i trafnie odnajduje obrazy podobne kolorystycznie.



Rys. 3. The infinite colors of search. Źródło: Morville & Callender, 2010, 16

Barwa w obrazie łączy się także z jej identyfikacją z elementem wyszukiwania i ma miejsce w wizualizacji marek i produktów (Rys. 3). Określony kolor kojarzy się z zasobem, do którego się odnosi. Naturalnie, proponowana organizacja informacji za pomocą kolorów

nie będzie kompletna z racji kontekstu, w jakim te dane się znajdują. Innym przykładem będzie kojarzenie konkretnie wyszukiwanego koloru z marką, firmą, produktem czy usługą internetową. W przestrzeni sieciowej dochodzi do nieformalnego zastrzegania barw do konkretnej usługi czy marki internetowej. Przykładem może być kojarzenie barwy ciemnoniebieskiej z portalem społecznościowym Facebook, edukacji internetowej rozumianej jako e-learning z kolorem zielonym, handlu elektronicznego z barwą żółtą, a wyszukiwania – z pomarańczową. Atlas barw internetowych jest otwarty, a identyfikacja wizualna usług i branż internetowych rozwija się bardzo dynamicznie.

Kolejnym parametrem opisującym obraz cyfrowy jest jego jasność. Dominanty jasności można zobrazować za pomocą histogramu, czyli graficznego wykresu jasnych, średnich i ciemnych partii obrazu. Rozkład ten jest podstawowym narzędziem służącym do maszynowego uczenia rozpoznawania kształtów. To na podstawie wyodrębnienia różnych jasności i kontrastów komputer jest w stanie wyróżnić i rozdzielić (posegmentować) elementy obrazu. Zasada ta służy do tworzenia algorytmów rozpoznających i wyszukiwujących obrazy, dokonujących autopoprawek zdjęć wykonanych amatorskimi aparatami, umożliwia zaawansowaną korektę (obróbkę) jasności poszczególnych partii obrazu w programie graficznym. Histogram nie charakteryzuje obrazu w sposób jednoznaczny. Wiele różniących się od siebie obrazów może mieć identyczne wartości przebiegu zmienności światła w całym przedstawieniu obrazowym. Nie jest też możliwe odtworzenie obrazu cyfrowego tylko i wyłącznie na podstawie jego histogramu. Technicznie rzecz ujmując, histogramy mogą przybierać zwizualizowane formy dwu- bądź trójwymiarowe. Jeśli nałożymy na siebie trzy histogramy składowych RGB jednego pliku cyfrowego, to umożliwimy komputerowi porównanie obrazów z innymi, które posiadamy w bazie danych. Zasada porównywania trójkanałowych histogramów jest najprostszą metodą wyszukiwania obrazów wizualnie podobnych do wyjściowego. Zaletą histogramów jest ich niska wrażliwość na operacje obrotów, przesunięć i skalowania oraz niska złożoność obliczeniowa. Wyszukiwanie odbywa się według atrybutów formalnych (kolorowe piksele), a nie z wykorzystaniem metadanych czy warstwy zawartości.

#### 4. Wyszukiwanie kontekstowe

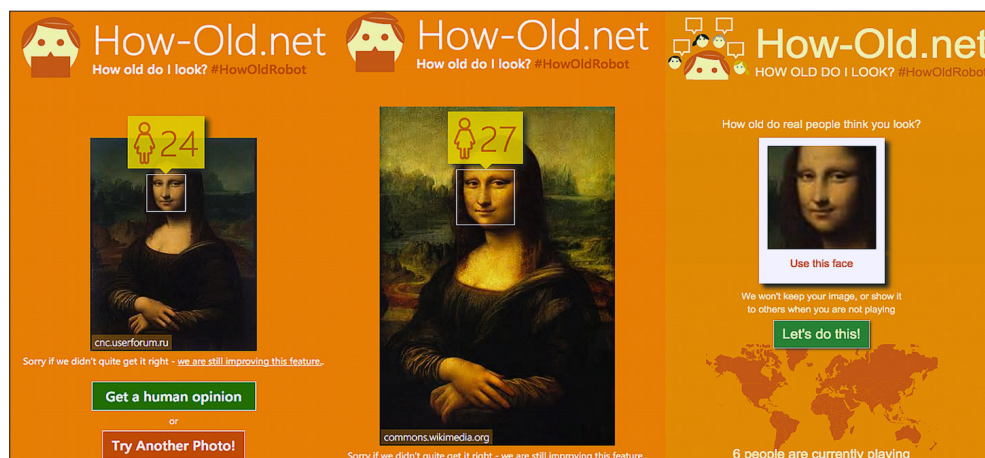
Najbardziej zaawansowanym sposobem wyszukiwania materiałów wizualnych jest wyszukiwanie w kontekście. Na podstawie wspomnianego wcześniej słownika wizualnego powstaje nowy typ relacji między obiektami widocznymi na materiałach wizualnych. Wyszukiwanie kontekstowe jest modelem, który sam się rozwija przez wdrażanie coraz to bardziej zaawansowanych zapytań w postaci obrazów. Można zaryzykować stwierdzenie, że wyszukiwanie kontekstowe jest przyszłością wyszukiwania materiałów wizualnych w ogóle. Odpowiada na problem masowego wzrostu popularności obrazów, fotografii, grafik itp. funkcjonujących w sieci, bo nie przekłada języka obrazu na mowę. Na bazie tych rozważań i udoskonalonej wizji Visual Vocabulary powstał nowy sposób opisu treści pod nazwą Visual Memex. W języku tym koduje się zarówno wygląd obiektów, jak i relacje między nimi bez podania źródła, tak jak jest to stosowane w obecnych rozwiązaniach. Algorytm jest w stanie dokładnie wskazać elementy na zdjęciu, ich wzajemne relacje i kontekst przestrzenny, w którym się znajdują. Dzięki temu możliwe jest rozpoznawanie obrazów malarskich, zdjęć



fotograficznych wykonanych w różnej technice, obiektów muzealnych zarówno dwu- jak i trójwymiarowych, grafik, szkiców i rysunków jako jednorodnych obiektów posiadających tę samą treść dostępną w określonym kontekście. Jest to kluczowe podejście w rozumieniu tego, co jest na zdjęciu oraz jaki obiekt przedstawia dany obraz z pominięciem techniki wykonania (sposobu wykonania) czy stopnia transformacji oryginalnego materiału wizualnego. Określenie podobieństwa wizualnego obrazów z różnych dziedzin nie jest możliwe techniką porównywania pikseli. W tym wypadku analizowana jest tzw. zawartość wizualna wyższego poziomu, na przykład w przypadku konkretnej sceny abstrahuje się od pory roku, w jakiej zostało wykonane zdjęcie, natężenia światła, formy – czy jest to obraz malarski, czy też grafika itp. (Shrivrasava et al., 2011).

## 5. Inicjatywy wspomagające rozwój wyszukiwarek wizualnych

Wszystkie wskazane wcześniej techniki komputerowego analizowania i reprezentowania obrazów mimo ich licznych i bezdyskusyjnych zalet zyskują, gdy do pracy z zasobem zostaje włączony człowiek. Im więcej wzorców wizualnych, tym doskonalszy system, stąd inicjatywy angażowania społeczności internautów do wzbogacania bazy wiedzy wyszukiwarek wizualnych. Przykładem takiej praktyki są choćby dwa projekty – Flickr i „Kto jest na Foto”. Innym przykładem jest system How Old<sup>3</sup> (Rys. 4) będący samouczącym się systemem eksperckim określającym wiek i płeć osoby przedstawionej na zdjęciu i wykorzystującym do nauki przykłady wczytywane przez internautów. W polu *Get a human option*, jeśli użytkownik ma odmienne zdanie co do wyniku, który podaje system, można do wczytanego przez siebie zdjęcia dodać poprawny wiek osoby, która na nim widnieje. Jeśli chodzi o identyfikację płci, to praktycznie jest ona rozpoznawana bezbłędnie.



Rys. 4. Szacowane wyniki wieku osoby na wczytanym do systemu portrecie. W tym przypadku jest to słynna *Mona Lisa*. Wczytanie dwóch bardzo podobnych zdjęć skutkuje podaniem zbliżonych, choć nie identycznych wyników. Źródło: <https://how-old.net> [30.01.2019]

<sup>3</sup> Zob. <http://how-old.net>

Inną aktywnością jest uruchamianie „gier metadanych”, na przykład Zen Tag czy Portrait Tag stworzone dla Muzeum Sztuki w Monachium. Aplikacje te są dostępne zarówno w postaci przeglądarkowej, jak i na urządzenia mobilne. Mają one w interaktywny sposób pobudzić kreatywność użytkowników, spowodować chęć rywalizacji w tworzeniu coraz bogatszych treści opisujących cyfrowe obrazy. Internauci najczęściej nadają słowa kluczowe (tagi) opisujące fotografię (kolekcję fotografii), tłumaczą fragmenty zeskanowanego tekstu na inny język lub „nazywają” wyświetlane przez komputer fragmenty obrazów. Projekty przeznaczone dla społeczności internetowej takie jak Open Image Collection czy Library Flickr Commons oferują z kolei możliwości sugerowania użytkownikowi tagów, które już ktoś wcześniej dodał. Mechanizm ten jest także dostępny w aplikacji Artigo<sup>4</sup>, gdzie użytkownik sprawdza sugerowane opisy obrazów. Dodatkowo, aplikacja pyta o relacje treściowe elementów obrazu z określeniami słownikowymi. Użytkownik ma za zadanie odpowiedzieć twierdząco bądź zanegować wybór spośród możliwych odpowiedzi.

Innym przykładem jest aplikacja o nazwie Metadata Games, gdzie to użytkownik informuje algorytm aplikacji, co widzi na ilustracji (Rys. 5). Komputer rozróżnia złożoność fraz wpisywanych przez internautę. Zadawane przez niego pytania początkowo są bardzo proste. Użytkownik, po udzieleniu poprawnych odpowiedzi (pokrywających się z posiadaną bazą odpowiedzi robota), przechodzi do etapu trudniejszego, czyli sugerowania fraz bardziej złożonych, których nie ma jeszcze w bazie danych. Za każdą odpowiedź gracz otrzymuje punkty zgodnie z zasadą: im mniej powtarzalna odpowiedź, tym większa liczba punktów.



Rys. 5. Metadata Games. Źródło: <http://play.metadatagames.org/stupidrobot> [30.01.2019]

## 6. Podsumowanie

Wydaje się, że dziś optymalnym rozwiązaniem problemów wyszukiwania fotografii byłby system hybrydowy, łączący opis tworzony przez profesjonalistę (archiwistę, bibliotekarza, muzealnika) z treściami tworzonymi oddolnie (kolektywnie) przez internautów (tagowanie) oraz z rezultatami komputerowej analizy obrazu. Opis obiektu wizualnego byłby tworzony

<sup>4</sup> Zob. <http://www.muenchen.de/int/en/events/art.html>

równoległe na różnych płaszczyznach, co pozwoliłoby na zmniejszenie ryzyka powstawania błędnych i niezrozumiałych opisów oraz na wzrost możliwości wyszukiwania wizualnego z uwzględnieniem potrzeb użytkowników. Sformułowany przez nas wniosek wprawdzie nie jest odkrywczy, ale zawiera ważną presupozycję, iż podejście tradycyjne, metadane i paradygmat sieciowy nie tworzą alternatywy, a przeciwnie – są komplementarne. Usprawnienie technologii komputerowych zależy przede wszystkim od bogactwa bazy wzorców dostarczanych do rozpoznawania. Dlatego tak istotne jest projektowanie systemów, które w sposób ciągły będą się uczyły nowych figur systematycznie pozyskiwanych z przykładowych obrazów. Współczesna technologia cyfrowa dysponuje dużym potencjałem adaptacyjnym. Kreatywny sposób wykorzystania narzędzi analizujących obrazy cyfrowe może pozwolić na znaczący postęp w dziedzinie wyszukiwania, zwłaszcza że fotografia cyfrowa staje się nie tylko rodzajem ikony, ale też znakiem czasów, wpisując się w kontekst komunikacji wizualnej w wielokanałowej komunikacji międzyludzkiej. Incepcjonizm (rozpoznawanie wyuczonych wzorców obrazowych na innych obrazach) przyspiesza w miarę nabierania przez algorytmy (na przykład Tensor Flow) „doświadczenia” w rozpoznawaniu i uczeniu się wzorców. Eksperymenty z zakresu sztucznej inteligencji prowadzone w przestrzeni wizualnej Internetu dają coraz bardziej interesujące i spektakularne rezultaty. Pozostaje mieć nadzieję, że ten nurt przemian nie ominie bibliotek i ich zasobów fotografii.

## Bibliografia

- Baca, M., ed. (2002). *Introduction to art image access: Issues, tools, standards, strategies* [online], [01.02.2019]. [http://www.getty.edu/research/publications/electronic\\_publications/intro\\_aia/index.html](http://www.getty.edu/research/publications/electronic_publications/intro_aia/index.html)
- Baca, M., Harpring, P., Lanzi, E., McRae, L., Whiteside, A. B. (2006). *Cataloging cultural objects: A guide to describing cultural works and their images* [online], [01.02.2019], [http://cco.vrafoundation.org/index.php/toolkit/cco\\_pdf\\_version/](http://cco.vrafoundation.org/index.php/toolkit/cco_pdf_version/)
- Bal, M. (1991). *Reading "Rembrandt": Beyond the word-image opposition: the Northrop Frye lectures in literary theory*. New York, Cambridge, Amsterdam: University Press.
- Beghtol, C. (1986). Bibliographic classification theory and text linguistics: Aboutness analysis, intertextuality and the cognitive act of classifying documents. *Journal of Documentation*, 42(2), 84–113.
- Bhattacharya, A., Ljosa, V., Jia-Yu Pan, Verardo, M. R., Hyungjeong Yang, Faloutsos, C., Singh, A. K. (2005.). ViVo: Visual Vocabulary Construction for Mining Biomedical Images. In: *Fifth IEEE International Conference on Data Mining (ICDM'05)* (50–57). Houston: IEEE.
- Clough, P. D., Ireson, N., Marlow, J. (2009). Extending domain-specific resources to enable, semantic access to cultural heritage data. *Journal of Digital Information*, 10 (6), [online], [01.02.2019], <https://journals.tdl.org/jodi/index.php/jodi/article/view/698/578>
- Corso, J. (2014). *Semantic knowledge. Generalized image understanding with probabilistic ontologies and dynamic adaptive graph hierarchies* [online], [30.01.2019], <http://www.cse.buffalo.edu/~jcorso/r/career/index.html>
- Domański, M. (2010). *Obraz cyfrowy: Reprezentacja, kompresja, podstawy przetwarzania. Standardy JPEG i MPEG*. Warszawa: Wydaw. Komunikacji i Łączności.
- Forczmański, P., Szeptycki, P. (2005). *Indeksowanie dużych zbiorów obrazów* [online], [30.01.2019], [http://pforczmanski.zut.edu.pl/pub/2005\\_wi.pdf](http://pforczmanski.zut.edu.pl/pub/2005_wi.pdf)
- Geertz, C. (2005). *Interpretacja kultur. Wybrane eseje*. Kraków: Wydaw. UJ.
- Gevers, T., Smeulders, A. W. M. (2003). *Image search engines: An overview*, [online], [30.01.2019], <http://staff.science.uva.nl/~gevers/pub/overview.pdf>

- Jørgensen, C. (2003). *Image retrieval: Theory and research*. Lanham: Scarecrow Press.
- Jørgensen, C., Jaimes, A., Benitez, A. B., Chang, S-F. (2001). A conceptual framework and empirical research for classifying visual descriptors. *Journal of the American Society for Information Science and Technology*, 52(11), 938–947.
- Jørgensen, C. (1999). Access to pictorial material: a review of current research and future prospects. *Computer and the Humanities*, 33(4), 293–318.
- Jørgensen, C. (1998). Attributes of images in describing tasks. *Information Processing & Management*, 34(2/3), 161–174.
- Miller, M., Wornbard, M. (2009). Fotografie w zbiorach cyfrowych – problemy z opracowaniem formalnym i rzeczowym na przykładzie Biblioteki Cyfrowej Politechniki Warszawskiej. *Przegląd Biblioteczny*, 77(2), 201–218.
- Morville, P., Callender, J. (2010). *Search Patterns: Design for Discovery*. Sebastopol: O'Reilly Media.
- Rygiel, P. (2012). Teoria i praktyka opracowania rzeczowego dokumentów ikonograficznych w dużych kolekcjach bibliotecznych. *Zagadnienia Informatyki Naukowej*, 98(1), 77–94.
- Shatford-Layne, S. (1994). Some issues in the indexing of images. *Journal of the American Society for Information Science*, 45(8), 583–588.
- Shatford, S. (1986). Analyzing the subject of a picture: A theoretical approach. *Cataloging & Classification Quarterly* 6(3), 39–62.
- Staszczak, R. Znajdź mnie, jeśli potrafisz – czyli jak znaleźć obrazy w Sieci [online], [30.01.2019], <http://www.chip.pl/artykuly/porady/2010/06/znajdz-mnie-jesli-potrafisz#ixzz47pHZv6c5>
- Svenonius, E. (1994). Access to non-book materials: the limits of subject indexing for visual and aural languages. *Journal of the American Society of Information Science*, 45(8), 600–606.
- Tabakowska, E. (2006). *Ikoniczność znaku: słowo – przedmiot – obraz – gest*. Kraków: Universitas.
- Talbierska, J. (2009). Zbiory artystyczne, zbiory specjalne, czy zbiory ikonograficzne? Typologia i funkcje. *Annales Universitatis Wratislaviensis*, 3167, *Bibliotekoznawstwo*, 28, 19–38.
- Wilson, T. D. (2000). Human information behaviour. *Informing Science. The International Journal of an Emerging Transdiscipline*, 3(2), 49–56.
- Włodarczyk, B., Woźniak-Kasperek, J. (2017). *Język haseł przedmiotowych Biblioteki Narodowej: od analizy dokumentu do opisu przedmiotowego*, Warszawa: Wydaw. SBP.

---

## Photography in the Library and on the Web: Two Perspectives on Representation and Retrieval

### Abstract

**Purpose/Thesis:** The purpose of the article is to present two approaches to photographs representation and retrieval: the traditional one, library and metadata-based and the automatic one, known as the network approach. The discussion is focused on indicating the possibility of the mutual complementation of both approaches in order to create a new information and search/retrieval quality in multi-channelled human communication.

**Approach/Methods:** The method of literature analysis and critical literature review was used.

**Results and conclusions:** Today the optimum solution for the problems of photographs representation and retrieval would be a hybrid system combining descriptions created by professionals with the content generated by Internet users (tagging) and the results of computer image analysis. The description of a visual object created at different levels would reduce the risk of producing erroneous, incomprehensible and incomplete descriptions while enhancing the possibility of visual objects retrieval, with the individual needs of users taken into consideration.

**Originality/Value:** The article discusses the issues of computer image analysis, in particular photographs analysis, and its usefulness for information and search purposes: these issues are still underrepresented in Polish book and information science literature.

**Keywords**

Computer image analysis. Image representation. Library. Photographs. Retrieval. Searching.

---

*KAMIL STĘPIEŃ, doktor nauk humanistycznych w dyscyplinie bibliologia i informatologia, adiunkt w Zakładzie Informatologii Instytutu Informacji Naukowej i Bibliotekoznawstwa UMCS. Zainteresowania badawcze: wizualizacja informacji, wyszukiwanie wizualne, funkcjonowanie fotografii w przestrzeni sieciowej i bibliotecznej, tagowanie i folksonomia. Ważniejsze publikacje: Folksonomie, czyli społecznościowe opisywanie treści (Warszawa, Wydaw. SBP, 2010); Wyszukiwanie obrazów i wyszukiwanie obrazem w przestrzeni internetowej. [W:] Z. Osiński i in. (red.), Biblioteka, Książka, Informacja i Internet (47–59) (Lublin: Wydaw. UMCS, 2015); Komunikacja fotograficzna. Interpretacja fotografii, Kultura i Historia 2018, nr 34, 162–186.*

*Kontakt z autorem:*

*kamil.stepien@umcs.pl*

*Zakład Informatologii, Instytut Informacji Naukowej i Bibliotekoznawstwa*

*Uniwersytet Marii Curie-Skłodowskiej w Lublinie*

*Plac Marii Skłodowskiej-Curie 4*

*20-031 Lublin*

*JADWIGA WOŹNIAK-KASPEREK, profesor nauk humanistycznych, kierownik Katedry Bibliografii i Dokumentacji na Wydziale Dziennikarstwa, Informacji i Bibliologii UW. Zainteresowania badawcze: informacja terminologiczna; teoria bibliografii; reprezentacja, organizacja i wyszukiwanie informacji. Ważniejsze publikacje: Problemy współpracy narzędzi wyszukiwawczych bibliografii specjalnych z uniwersalnymi systemami organizacji wiedzy, Śląski Kwartalnik Historyczny Sobótka 2016, 71(3), 39–44; z Bartłomiejem Włodarczykiem, Język haseł przedmiotowych Biblioteki Narodowej: od analizy dokumentu do opisu przedmiotowego (Warszawa: Wydaw. SBP, 2017); Kłamstwo i metoda: na przykładach reprezentacji treści w katalogach bibliotecznych i w bibliografiach (Warszawa: BEL Studio, 2018).*

*Kontakt z autorką:*

*jbwozniak@uw.edu.pl*

*Katedra Bibliografii i Dokumentacji, Wydział Dziennikarstwa, Informacji i Bibliologii*

*Uniwersytet Warszawski*

*ul. Nowy Świat 69*

*00-046 Warszawa*

---

# Gatunek *information scraps* w indywidualnym/ grupowym zarządzaniu informacją

Paulina Krzanowska

*CommsBlack Sp. z o.o.*

Marek Nahotko

*Instytut Informacji Naukowej i Bibliotekoznawstwa, Uniwersytet Jagielloński*

---

## Abstrakt

**Cel/Teza:** Celem artykułu jest prezentacja roli gatunków informacji w indywidualnym (nieinstytucjonalizowanym) zarządzaniu informacją na przykładzie systemu gatunków zwanych notatkami lub *information scraps* tworzonymi w sposób tradycyjny (bez stosowania narzędzi elektronicznych).

**Koncepcja/Metody badań:** W badaniach grupy studentów zastosowano dwie komplementarne metody badań jakościowych: swobodny wywiad ukierunkowany oraz analizę zebranego materiału empirycznego (odręcznych notatek). Wywiad pozwolił na zebranie wypowiedzi respondentów nt. celów stosowania notatek. Zebrane notatki były analizowane z punktu widzenia formy i treści. Metodologia ta pozwoliła na dokonanie analizy gatunkowej *information scraps*, która zgodnie z literaturą przedmiotu dotyczyć powinna trzech cech gatunku: celu, treści i formy, w tym formy językowej.

**Wyniki i wnioski:** *Information scraps* stanowią system gatunków wspomagających zarządzanie informacją. Ich forma i treść dostosowane są do realizowanych celów komunikacyjnych: kognitywnych, afektywnych, estetycznych i edukacyjnych. Stosowane są w sytuacjach standaryzowanych i powtarzalnych działań komunikacyjnych.

**Oryginalność/Wartość poznawcza:** Stosowanie notatek w działaniach indywidualnych wspomaga tworzenie wspólnego gruntu stanowiącego podstawę działań grupowych, stąd *information scraps*, jak każdy gatunek, wspomagają społeczne działania komunikacyjne.

## Słowa kluczowe

Analiza gatunkowa. Gatunki informacji. Indywidualne/Grupowe zarządzanie informacją. Notatki odręczne

*Otrzymano: 12 grudnia 2018. Zrecenzowano: 27 grudnia 2018. Zaakceptowano: 22 stycznia 2019.*

---

## 1. Wstęp

Człowiek, podobnie jak każdy organizm żywy, może być traktowany jako system informacyjny, internalizujący informację, przetwarzający ją i eksternalizujący<sup>1</sup>. Z drugiej strony, wszelkie działania ludzkie są społeczne w szerokim tego słowa znaczeniu, a więc związane są z komunikowaniem się z innymi<sup>2</sup>. Każdy z nas zarządza informacją przekazywaną przez

<sup>1</sup> W procesach tych tworzona i modyfikowana jest indywidualna wiedza.

<sup>2</sup> Niektóre działania, takie jak: śpiewanie (na scenie), słuchanie muzyki, czytanie, oglądanie filmów i dzieł plastycznych w sposób oczywisty służą komunikacji, rozumianej jako przekazywanie informacji. Inne, jak pisanie pamiętnika lub notatek tylko pozornie nie służy komunikowaniu się z innymi, co przedstawione zostanie w dalszej części artykułu.

innych ludzi oraz danymi spoza środowiska społecznego<sup>3</sup>. Oznacza to, że każde działanie związane jest w jakimś stopniu z zachowaniami informacyjnymi, rozumianymi jako aktywności, działania i reakcje ludzi pojawiające się w momencie zetknięcia się ze źródłami lub kanałami informacyjnymi, takie jak aktywne i pasywne poszukiwanie informacji i jej wykorzystanie (Wilson, 2000, 49). Zachowania te służą realizacji potrzeb informacyjnych, których posiadanie jest immanentną cechą każdego człowieka. Podczas zarządzania informacją i jej organizacji w specyficzny sposób łączone jest to, co jednostkowe (indywidualne zarządzanie informacją, ang. *Personal Information Management* – PIM) z tym, co społeczne (grupowe zarządzanie informacją, ang. *Group Information Management* – GIM). W procesach tych wykorzystywane są narzędzia dwojakiego rodzaju – mentalne<sup>4</sup> (wewnętrzne) i techniczne (zewnętrzne), łącznie charakteryzujące stosowany gatunek informacji. Pierwsze służą do ujednoczenia znaczeń (tworzenia konwencji społecznych) umożliwiającego komunikację; drugie wykorzystywane są do przekazywania, utrwalania i organizacji komunikowanej informacji. W artykule zostaną przedstawione wyniki badań dotyczących obu aspektów zarządzania informacją na przykładzie specyficznego gatunku informacji, jakim są odręcznie tworzone notatki (ang. *information scraps*). Celem badań jest ujawnienie za pomocą analizy gatunkowej notatek zastosowania tego specyficznego gatunku informacji w procesach organizacji informacji.

## 2. Indywidualne i grupowe zarządzanie informacją

Zainteresowanie PIM i GIM wynika z zauważenia skądinąd oczywistego faktu, że jednostka (osoba) nie staje się użytkownikiem informacji jedynie z chwilą przekroczenia progu biblioteki lub stosowania innego, zinstytucjonalizowanego systemu informacyjnego. Człowiek jest biologiczną maszyną informacyjną, która nieprzerwanie przetwarza zarówno materię, jak i informację, pozwalającą na dostęp do materii i właściwy z niej użytek. Przetwarzanie informacji odbywa się nieustannie, w sposób świadomy i nieświadomy (np. także we śnie). W takim razie odbywa się we wszystkich kontekstach, spośród których w informatologii zwykle wyróżnia się kontekst zawodowy i pozazawodowy (osobisty). W normalnej sytuacji oba zawsze się mocno splatają.

Wielość kontekstów wynika z faktu, że potrzeby informacyjne i zachowania, których potrzeby są przyczyną, to zjawiska społeczne, realizowane w społecznym otoczeniu, bliższym i dalszym. Dlatego podział na PIM i GIM jest umowny, zarządzanie informacją może być mniej lub bardziej indywidualne lub grupowe. Jest to więc pewne kontinuum, od mocno indywidualnego do bardzo grupowego zarządzania informacją, ale nigdy nie jest zupełnie pozbawione kontekstu społecznego. Jednocześnie zawsze jest w pewnym stopniu indywidualne. Remigiusz Sapa podkreślał dwoistość PIM/GIM odpowiadającą temu kontinuum, wskazywał na funkcjonowanie tych procesów w obu kontekstach: indywidualnym

<sup>3</sup> Terminy „dane”, „informacja” i „wiedza” rozumiane są tak, jak to zostało przedstawione w (Nahotko, 2018, 69 i nast.).

<sup>4</sup> Narzędzia mentalne rozumiane są jako struktury mentalne służące do bezpośredniego i celowego oddziaływania na czynności realizowane przez jednostkę, sterując ich wykonywaniem. Mogą one być traktowane jako narzędzia przez analogię do oprogramowania komputerowego. Część narzędzi mentalnych stosowana jest w procesach komunikacyjnych.

i grupowym. W pierwszym przypadku PIM/GIM może być przejawem działalności indywidualnej nakierowanej na potrzeby całej grupy. Kontekst grupowy oznacza postrzeganie grupy jako całości i skupianie się przede wszystkim na zarządzaniu informacją przez tak rozumianą grupę (Sapa, 2017, 369–380). Stąd w dalszej części artykułu będziemy posługiwali się skrótem PIM/GIM oznaczającym to continuum.

Krótkie wprowadzenie do problemów PIM/GIM zaczniemy od indywidualnego końca continuum. William Jones w 2007 r. podał trzy aspekty zarządzania indywidualnego, które w 2016 r. zostały przez niego uzupełnione o kolejne trzy. Pierwszy wskazuje na to, że zarządzanie dotyczy informacji należącej do jednostki zarządzającej, przez nią kontrolowanej i wykorzystywanej do celów osobistych. Drugi odnosi się do posiadanej przez innych informacji o jednostce. Kolejny wskazuje na jednostkę jako odbiorcę informacji, bez względu na to, czy odczuwa ona potrzebę tej informacji (może to być np. informacja dostarczana przez reklamy). Wyróżnił także informację, którą jednostka komunikuje innym, np. wysyłając pocztę elektroniczną. Indywidualne autorstwo sprawia, że takie działanie traktowane jest również jako zarządzane indywidualnie. Jones pisał też o informacji pozyskiwanej poprzez doświadczenia jednostki oraz informację ogólnie nazwaną relewantną względem działań (również przyszłych) i zainteresowań jednostki (Jones, 2007, 461–462; Jones et. al., 2016, 573). Podobne stanowisko w tej sprawie przyjmował wcześniej Mark Lansdale, nie traktując informacji własnej (Tomaszczyk, 2017, 210) jako czegoś prywatnego, dostępnego wyłącznie dla jednostki<sup>5</sup>. Indywidualność oznacza możliwość samodzielnego używania, uzyskiwania (ang. *retrieve*) oraz wielokrotnego wykorzystania tego rodzaju informacji (Lansdale, 1988, 55).

Jones odnosił PIM/GIM do teorii i praktyki wszelkich działań względem informacji wykorzystywanej w kontekstach zawodowym i osobistym oraz w wypełnianiu obowiązków wynikających z pełnionej roli społecznej (np. bycia członkiem rodziny lub innej społeczności). Wymienia on aż osiem rodzajów działań PIM/GIM, do których zaliczyć można m.in.: nabycie (ang. *acquire*) lub stworzenie, organizowanie, użytkowanie i rozpowszechnianie (ang. *distribute*) informacji. Wszystkie wymienione działania służą umieszczeniu właściwej informacji we właściwym miejscu i formie, w celu realizacji potrzeb informacyjnych. Czasami Jones zwięźle określa PIM/GIM jako sztukę stosowania informacji dla realizacji czynności życiowych (Jones, 2007, 453; 2015, 5). Analogiczne określenie indywidualnego zarządzania informacją znaleźć można w publikacjach Briana Detlora oraz Katarzyny Materskiej. Pierwszy autor opisywał działania informacyjne podejmowane dla osiągnięcia indywidualnych celów jednostki. Wyróżnia on sześć kroków: nabycie lub stworzenie, uporządkowanie, przechowywanie, wykorzystywanie na własny użytek, rozpowszechnianie (Detlor, 2010, 107). K. Materska krytycznie odnosi się do tego rodzaju definicji wyliczających. Jej zdaniem należałoby skupić większą uwagę na procesach realizowanych w osobistej przestrzeni informacyjnej jednostki, bowiem jest to miejsce wielu interakcji w kontekście PIM (Materska, 2012, 67). Przestrzeń ta (ang. *personal space of information, PSI*) obejmuje wszystkie elementy (ang. *information items*) będące w posiadaniu jednostki, mające wartość informacyjną, a także wszelkie narzędzia (np. aplikacje) wspomagające procesy zarządzania nimi. Według Mirosława Górnego PSI składa się także z elementów mentalnych, takich

<sup>5</sup> Traktowanie informacji jako indywidualnej własności byłoby sprzeczne z jej definicją przyjętą w artykule, traktującą informację jako reprezentację wiedzy przekazywanej w procesie komunikacji.



jak: wiedza, doświadczenie, umiejętności, postawy, wierzenia, uprzedzenia związane z realizacją zadań informacyjnych (Górny et. al., 2017). Każda jednostka posiada tylko jedną PSI. Elementy materialne w niej zawarte mogą występować w formie tradycyjnej i/lub elektronicznej. Są bezpośrednio i szybko dostępne do wykorzystania w procesach kognitywnych. Przestrzeń ta podzielona jest na mniejsze „wyspy”, zwane osobistymi kolekcjami informacji (ang. *personal information collections* – PIC), podzbiorami wypełnionymi zorganizowanymi w ten sposób elementami informacyjnymi. PIC mogą różnić się liczbą, formą i spójnością treści organizowanych elementów. Organizacja elementów obejmuje ich układ przestrzenny, strukturę, właściwości i atrybuty. Stosowane są w różnych celach, są to np. w różnym stopniu uporządkowane dokumenty firmowe znajdujące się w biurkach i szafkach oraz w folderach komputerów, zbiory elementów informacji specyficznych dla danego projektu, które są pierwotnie umieszczone w folderze, a następnie sukcesywnie opracowywane (Jones, 2007, 460, 462–463; 2015, 129–130). PSI budowaną w procesach PIM można uważać za część publicznej przestrzeni informacyjnej, w której realizowane są procesy GIM.

William Jones podkreśla, że PIM/GIM ma istotny wymiar praktyczny. W kontekście jednostki ma ono znaczny wpływ na gospodarowanie takimi dobrami, jak: czas, finanse oraz własna energia, natomiast w przypadku organizacji zasadnicze znaczenie ma wpływ na wydajność pracowników oraz ich relacje zawodowe w grupie (Jones, 2007, 456–457). Problemem występującym w całym kontinuum PIM/GIM jest tzw. fragmentaryzacja (rozproszenie) informacji, polegająca na stosowaniu wielu formatów i narzędzi do jej materializacji. Informacja zapisana w wersji tradycyjnej może być łatwo skonwertowana na wersję elektroniczną i odwrotnie. Korzystanie z kilku kont poczty elektronicznej, zapisywanie informacji na dyskach komputerów lub pamięciach przenośnych, w osobistych organizacjach, a nawet w telefonach komórkowych, to tylko kilka przykładów wskazujących na problemy fragmentaryzacji informacji i jej wpływu na efektywne zarządzanie informacją (Jones, 2007, 453; Materska, 2012, 71).

Na drugim końcu kontinuum PIM/GIM znajduje się grupowe zarządzanie informacją. W zależności od sposobu traktowania „grupy” wyznacza się trzy możliwe perspektywy. Pierwsza z nich odnosi się do możliwości wywierania wpływu przez grupę na PIM/GIM każdej jednostki wchodzącej w zasięg jej oddziaływania. Kolejna perspektywa odnosi się do sposobów dzielenia się informacją pomiędzy jednostkami w grupie oraz zarządzania wspólnymi zasobami informacyjnymi (Teevan et al., 2008, 104). Są to dwa konteksty tego samego zjawiska. W. Jones, definiując GIM wskazał, że jest to suma działalności pojedynczych jednostek, a ich praktyki stosowane w zarządzaniu informacją mają największy wpływ na funkcjonowanie grupy jako całości (Jones, 2015, 3). Podobne zdanie wyrażali W. Lutters, M. Ackerman i X. Zhou, wskazując, że GIM funkcjonuje także w kontekstach zawodowym i prywatnym, w tym domowym (Lutters et al., 2007, 236–237 i 243). Z kolei T. Erickson traktuje GIM jako PIM, działający w znacznie szerszym, społecznym zakresie. Przedstawił on nawet prosty model GIM zawierający się w jednym zdaniu:

Jednostka generuje informację, którą dzieli się z grupą w celu wspierania wspólnych działań (Erickson, 2006, 74–75)<sup>6</sup>.

<sup>6</sup> Wszystkie cytaty z prac nieopublikowanych dotychczas w języku polskim są tłumaczeniami autorów niniejszego artykułu [red.].

Za pomocą tego modelu cytowany autor prezentował problemy i zagadnienia PIM/GIM, odnoszące się do trzech kwestii: dotyczących informacji bezpośrednio, np. jej organizacji i struktury; dotyczących czynności dzielenia się informacją, np. przepływów informacji i sposobów ich organizacji; oraz związanych ze wspieraniem i wykonywaniem działań, np. problemy zapewnienia prywatności i kontroli własnych zasobów informacyjnych.

### 3. Gatunki informacji jako mentalne narzędzia PIM/GIM

Gatunek informacji rozumiany jest jako typizowana forma komunikacji realizowanej we wszystkich kontekstach aktywności ludzkiej, głównie osobistym i zawodowym, a także we wszystkich praktykach społecznych (Devitt, 2004; Luckmann, 2009). Taka typizacja ma miejsce podczas każdego aktu komunikacji społecznej (wymiany informacji). Nie istnieją wypowiedzi „pozagatunkowe”. Oznacza to, że gatunek jest bardzo wydajnym narzędziem badania i zrozumienia stosowanych form komunikacji dzięki wspomaganemu przez niego wyjaśnianiu celów i sposobów stosowania różnych form informacji i jej materializacji w działalności ludzkiej, a także badania struktur społecznych tworzących gatunki i ich roli w koordynacji różnych form organizacji społecznej, w tym organizacji informacji. Gatunek jest więc rodzajem klamry łączącej społeczne, powszechne stosowanie informacji z kontekstem indywidualnego, lokalnego dostosowania i interpretacji. Dla właściwej koordynacji działań społecznych zasadnicze znaczenie ma oddziaływanie<sup>7</sup> gatunku (ang. *uptake*) w konkretnej sytuacji komunikacyjnej (Freadman, 2002). Ponieważ oddziaływanie gatunku ma zarówno wymiar indywidualny (wiedza jednostkowa), jak i zbiorowy (wspólny grunt), gatunki zawsze służą mediacji pomiędzy komunikacją na poziomach indywidualnym i społecznym. Gatunek jest więc zawsze obciążony sprzecznościami spowodowanymi jednoczesnym funkcjonowaniem w indywidualnym i społecznym kontekście komunikacyjnym. Jest więc narzędziem komunikowania służącym jednostkom do wykonywania rozpoznawalnych działań komunikacyjnych realizowanych w obrębie tych kontekstów lub pomiędzy nimi (Andersen, 2017, 5).

W ten sposób rozumiane gatunki nazywane były dotąd w literaturze różnymi terminami. Pisano więc o gatunkach tekstu (np. Isenberg, 1978; Miller, 1984; Franke, 1987), mowy (Bachtin, 1986), dokumentu (Kwaśnik & Crowston, 2005), wypowiedzi (Gajda, 2009) oraz komunikacji (Luckmann, 1986; Swales, 1990). W tym artykule używany będzie termin *gatunki informacji*. Aktywność komunikacyjna, typowa dla ludzi i tworząca standaryzowaną przestrzeń informacyjną, jest jedną z głównych funkcji życiowych. Komunikacja służy przekazywaniu informacji, natomiast tekst jest stosowany w tym procesie jako najbardziej efektywny, ale nie jedyny sposób materializacji informacji. Jeżeli za Carolyn Miller (1984, 162) przyjmiemy, że cel komunikacyjny jest najważniejszą cechą charakteryzującą gatunek, to musimy zauważyć, że cel ten jest osiąganym za pomocą przekazywania informacji. Forma jej materializacji – tekstowa (piśmienna lub werbalna) albo inna, na przykład gestykularna,

<sup>7</sup> Oddziaływanie definiowane jest jako nieprzewidywalne, improwizowane, wielokierunkowe funkcjonowanie gatunku w rzeczywistym czasie i miejscu. *Uptake* pozwala zrozumieć sposób, w jaki systematyczne, typowe relacje między gatunkami koordynują złożone formy działań społecznych – jak i dlaczego gatunki oddziałują w określony sposób oraz co jest wykonywane lub niewykonywane, jako wynik ich oddziaływania.

jest ważna, ale nie pierwszorzędna. Cel komunikacyjny jest osiągnięty przez przekazywanie informacji z wykorzystaniem form i treści typowych dla gatunku. Gatunek funkcjonuje pomiędzy informacją a jej materializacją, dla której dobierane są narzędzia techniczne zapewniające osiągnięcie celu komunikacyjnego z uwzględnieniem projektowanego odbiorcy informacji. Informacja jest sednem każdego tekstu, bez niej tekst traci sens.

Każda nowa informacja jednocześnie wzmacnia i przemodelowuje pewne aspekty gatunku, każdy odbiór informacji przekształca jej społeczne rozumienie. Użytkownik musi włączyć w swoje struktury mentalne wiedzę o możliwych działaniach informacyjnych związanych ze stosowaniem gatunku. Można więc mówić o gatunku jako mentalnym wzorcu (schemacie) umieszczonym w kompetencji językowej nadawcy i odbiorcy, wynikającym z posiadanej wiedzy (zinternalizowany model językowy) (Gajda, 2009, 138). Wzorce stanowią element wiedzy użytkownika języka i są używane do wyzwalania procesów kognitywnych w celu wykonywania zadań komunikacyjnych (Maciejewski, 2009, 85). Jak wspomniano wcześniej, gatunek stanowi informacyjne ujęcie typowej sytuacji/wydarzenia związanego z działaniem w otaczającej rzeczywistości. Mentalną reprezentacją wydarzenia/działania jest schemat poznawczy, model fragmentu rzeczywistości. Część wydarzenia/działania stanowią obiekty informacyjne wspomagające wymianę informacji pomiędzy uczestnikami. Gatunek należy postrzegać jako rodzaj schematu mentalnego, dobieranego odpowiednio do innych schematów wydarzeń/działań związanych z sytuacją. Schemat ten pozwala na konstruowanie relacji pomiędzy elementami sytuacji komunikacyjnej i innymi elementami, związanymi z innymi sytuacjami, których częścią jest proces komunikacji.

Analiza gatunków opiera się na dwóch głównych założeniach: podobieństwo komunikatów (np. tekstów) zależy od cech związanych ze społecznym kontekstem ich tworzenia i stosowania, a cechy te mogą być opisane w sposób, który wskazuje na relacje badanego tekstu z innymi, podobnymi tekstami oraz z wyborami i ograniczeniami wpływającymi na twórców tekstu (Hyland & Salager-Meyer 2008, 303). Traktowanie gatunku jako działania społecznego powoduje zwracanie uwagi w trakcie analizy na cel komunikacyjny komunikatu jako główną cechę gatunku (Askehave & Swales 2001, 197). Język i jego narzędzia postrzegane są jako należące do rzeczywistości społecznej, a jednocześnie budujące ją, co odbywa się poprzez powtarzalne stosowanie skonwencjonalizowanych form komunikacyjnych, przy pomocy których jednostki tworzą relacje, budują społeczności i realizują swoje zadania. Z tego punktu widzenia gatunki są efektem grupowej aktywności pojedynczych agentów społecznych, działających zarówno w granicach swojej historii, jak i w ograniczeniach narzucanych przez kontekst oraz posiadających wiedzę o istniejących gatunkach. Są więc elementem procesów PIM/GIM.

JoAnne Yates i Wanda Orlikowski (1992, 301) sugerują, że dla celów analizy gatunki mogą być charakteryzowane przez podobieństwo tekstów w zakresie treści (sensu) i formy. Treść dotyczy społecznych motywacji, problemów i tematów wyrażanych podczas komunikacji, a forma określana jest przez widoczne cechy fizyczne i językowe komunikacji, takie, jak struktura, nośnik, język lub system symboliczny. Inne cechy te nazywają celem (zamiarem) i formą (Crowston & Williams 2000, 202; Yates et al., 1997). Podobnie jak poprzednio cel dotyczy intencji komunikacyjnych, a forma cech głównie językowych. Shepherd i Watters dla gatunków tekstów elektronicznych dodają trzecią cechę (oprócz treści i formy), mianowicie funkcjonalność, ze względu na to, że nośnik elektroniczny posiada specyficzne cechy funkcjonalne. Cechy te odnoszą się do nowych możliwości, jakie

oferują media elektroniczne (Shepherd & Watters 1999, 2007). Jan Ljungberg (2008) na tej podstawie zaproponował zestaw cech: treść (rodzaj informacji), forma (format prezentacji treści), funkcjonalność (możliwości związane z nośnikiem) i cel (potrzeby indywidualne realizowane społecznym otoczeniu).

#### 4. *Information scraps* jako gatunek informacji

Pomimo tego, że obecnie dostępna jest spora liczba narzędzi elektronicznych stosowanych w PIM/GIM, nadal tworzona jest informacja, której materializacja pozostaje poza zasięgiem tych narzędzi. Gromadzona jest w postaci notatek na karteczkach typu *Post-it*, zapisków na skrawkach papieru wkładanych pospiesznie do kieszeni lub umieszczanych w widocznym miejscu (pulpit biurka, drzwi lodówki) albo przechowywanych w specjalnych zeszytach, pojemnikach lub inaczej zorganizowanych zbiorach. Jej treść jest różnorodna, od pojedynczych wyrazów i szkiców po wielozdaniowe notatki, listy wieloelementowe, a nawet wiersze. Są obecne we wszystkich sytuacjach życiowych, stale noszone albo wypełniające sprzęty w miejscu pracy, często naklejane na monitory komputerów (Bernstein et. al., 2008, 2).

W literaturze obcojęzycznej znaleźć można kilka terminów określających ten gatunek informacji. Stosuje się następujące nazwy: *micronotes*, *notes-to self*, *post-it-notes*, *sticky notes*, *move-it sticky notes* oraz *information scraps*. Każda z nich odnosi się do tego samego zjawiska, stąd używane są zamiennie. Świadczą o tym definicje słownikowe. *Post-it* według *Cambridge Dictionary* oznacza mały kolorowy skrawek papieru przeznaczony do notowania krótkich wiadomości, który może być prowizorycznie przymocowany do pewnej powierzchni (*Post-it note*, 2017). Istotna jest tu wzmianka o formie notatki, zwrócenie uwagi na to, iż może to być dowolny skrawek kartki lub papieru, niekoniecznie przeznaczona do tego karteczka. W *English Oxford Living Dictionary* ta sama definicja umieszczona jest pod dwoma hasłami: *sticky* oraz *post-it* (*Sticky*, 2017). *Micronotes*, zdaniem Min Lin, Wayne Luttersa oraz Tiny Kim (2004, 687), to krótkie, wyrażane zwykle w kilku słowach, osobiste notatki kierowane do samego siebie, rejestrujące informacje godne uwagi (ang. *notable information*). Ich treść zapisywana jest nie tylko na karteczkach typu *Post-it*, ale także innych łatwo dostępnych nośnikach, czasami również na dłoniach. Celem ich tworzenia jest zapisywanie informacji aktualnych i bieżących, które będą użyteczne w przyszłości.

W polskiej literaturze przedmiotu brakuje publikacji na ten temat, stąd nie ma adekwatnego terminu, który w pełni odpowiadałby terminom angielskim. Przybliżone znaczenie ma termin *notatka*, której definicja w *Słowniku Języka Polskiego* składa się z kilku części: jest to „krótki tekst, uwaga, spostrzeżenie zapisane w celu zapamiętania, zapisek” lub „krótka informacja, wzmianka w prasie; notka” oraz „kartka z notatką, zapiskiem” (*Notatka*, 1979, 395). W definicji wskazuje się na kilka gatunków, wymieniając notatkę, zapisek, notkę. Jednak nie jest ona wystarczająco rozbudowana, brakuje w niej kilku kluczowych aspektów. Stąd w dalszej części artykułu jako główne określenie przyjęto termin *information scraps*, stosowany przez Michaela Bernsteina et. al. (2008, 2–3), używany zamiennie z polskim terminem *notatka*.

Bernstein et. al. (2008) wskazują na trzy cechy *information scraps* jako gatunku: występowanie na skrawkach, świsłkach, fragmentach, wycinkach papieru (forma), niekompletność treści oraz stosowanie niestandardowych narzędzi podczas PIM/GIM. Autorzy mają

na myśli np. zanotowane informacje adresowe nienadające się do utrwalenia w książce adresowej lub listę rzeczy do zrobienia utworzoną bez użycia aplikacji do zarządzania listami zadań. W literaturze często podkreśla się także nieformalny, nieoficjalny charakter notatek, co jest zaletą zwiększającą częstotliwość stosowania. Są regularnie stosowane ze względu na łatwość obsługi, elastyczną formę, dopasowanie formy do treści i możliwość przechowywania w dowolnym PIC (van Kleek, 2011, 2907). W ocenie M. Bernsteina i in. *information scraps* posiadają również inne cechy gatunkowe, przedstawione poniżej:

- szybka i łatwa percepcja treści w nich zawartych (ang. *lightweight capture*);
- możliwość zapisu treści w wybranej formie (ang. *flexible contents and representation*);
- wszechstronne zastosowania, tworzenie własnych systemów organizacji informacji lub adaptacja do narzędzi już wykorzystywanych (ang. *flexible use and organization*);
- wspomaganie pamięci dzięki widoczności/czytelności/dostrzegalności (ang. *visibility and reminding*);
- mobilność wspierająca dostępność (ang. *mobility and availability*) (Bernstein et. al., 2008, 5).

Informacja w gatunku *information scraps* przyjmuje wiele form. Wynika to z cechy gatunkowej notatek, jaką jest łatwe dostosowanie do indywidualnych potrzeb i preferencji. Interesujące spostrzeżenie przedstawili Max van Kleek i in. (2011, 2911), którzy wskazali na dwie postaci *information scraps* związane z celami komunikacyjnymi. Po pierwsze, są to notatki, które służą do zapisu informacji o czymś, co notujący zrobił lub doświadczył, w celu kontrolowania tej sprawy na bieżąco lub śledzenia jej dalszego rozwoju, w trakcie którego stosowane są kolejne gatunki. Zostało to określone jako prowadzenie dziennika (ang. *journaling*) lub kroniki (ang. *chronicling*). Po drugie, celem notatek jest zachowanie informacji ze względu na jej wartość osobistą lub emocjonalny, sentymentalny charakter. Wyróżniono cztery najbardziej typowe formy *information scraps*. Są to: listy zadań, spraw do załatwienia (ang. *to-dos*), notatki tworzone w trakcie spotkań (ang. *meeting notes*), informacje kontaktowe (ang. *name and contact information*) oraz notatki instruktażowe (ang. *how-to guides*). Formy te zostały ustalone na podstawie badań Bernsteina i in., którzy dodatkowo dokonali podziału *information scraps* pod względem nakładu pracy, wysiłku włożonego w organizację notatek. Podzielono je na trzy kategorie: wymagające małego, średniego i dużego wysiłku. Małego trudu wymagają notatki fragmentaryczne, posiadające szczególne znaczenie dla notującego albo odseparowane od innych informacji. Przybierają one formę pojedynczych elementów. Mają być stale dostępne i widoczne, stąd nie ma potrzeby stosowania dla nich specjalnych narzędzi tworzenia i archiwizacji. Wysiłek włożony w ich organizację jest niski, ponieważ ogranicza się tylko do umieszczenia ich w widocznym miejscu. Organizacja notatek zawierających informację powtarzalną w różnych wariantach aktywuje umiarkowany poziom wysiłku. Użytkownik takiej notatki musi włożyć więcej trudu w gromadzenie treści jednej kategorii w jednym miejscu. Jego praca polega na grupowaniu notatek w jednolite, wewnętrznie nieuporządkowane zbiory. Dopiero połączenie archiwizacji i organizacji informacji oznacza najwyższy poziom wysiłku (Bernstein et. al., 2008, 19 i 37–38).

*Information scraps*, jak wszystko na tym świecie, mają swój początek i koniec. M. Lin, W. Lutters i T. Kim opracowali ośmioetapowy „cykl życia” notatek. Rozpoczyna się on wraz z odczuwaniem chęci lub konieczności zapisania informacji. Następny etap obejmuje zarejestrowanie treści notatki. W tym miejscu przewidziano dwie możliwości dalszego

ich „życia”. Może nastąpić transfer treści do innego narzędzia, w tym elektronicznego lub mogą one być zachowane w aktualnej formie w celu dalszego ich wykorzystania. W drugim przypadku muszą być dokładne, przechowywane w widocznym miejscu i uaktualniane, ponieważ ponowne ich wykorzystanie zależy od tego, czy są natychmiast dostrzegalne, czytelne i zrozumiałe. Etap zatrzymania ich dla siebie wiąże się ze stałym odnoszeniem się, nawiązywaniem, sprawdzaniem. Oznacza to posiadanie wiedzy o ich istnieniu oraz powracanie do nich i przypominanie sobie ich treści. Kolejny etap określany jest przez cytowanych autorów jako *complete*; dzięki odpowiednim działaniom w poprzednich etapach notatka spełniła zamierzony dla niej cel. W tym momencie staje się bezużyteczna i można postąpić z nią na dwa sposoby: albo ją zarchiwizować, albo się jej pozbyć. W pewnych sytuacjach są zachowywane przez dłuższy okres, od kilku dni po lata. Niektórzy przechowują je tak długo, jak tylko to możliwe i decydują się na ich pozbycie dopiero w przypadku braku miejsca (Lin et al., 2004, 689–692).

W szeroko pojętym zarządzaniu informacją *information scraps* służą realizacji celów komunikacyjnych. Jest ich tak wiele, jak działań, którym mogą towarzyszyć. Wykorzystywane są w trakcie planowania i realizowania aktywności do zapisu informacji ważnych z punktu widzenia osoby notującej. Mogą być tworzone na podstawie kilku źródeł jednocześnie, stając się swego rodzaju zewnętrzną pamięcią, przydatną wszystkim podstawowym operacjom myślowym (Piolat et al., 2005, 291). W literaturze najczęściej opisywane są jako tymczasowe miejsce przechowywania informacji oraz ewentualne zabezpieczenie przed zapomnieniem. Poza tym odgrywają rolę kognitywnego wsparcia (ang. *cognitive support*), czyli stają się miejscem, w którym umieszczane są informacje dotyczące wszelkich pomysłów, obserwacji, wniosków i innych efektów „trudnych procesów myślowych”. Czasami są swego rodzaju archiwum, jeśli notujący stwierdzi, iż zapisana informacja musi być przechowana na dłużej (Bernstein et. al., 2008, 4–5 i 35–36).

Tworzenie notatek jest czynnością złożoną, która obejmuje zdolności poznawcze oraz procesy twórcze. Takie ich rozumienie skłania do stwierdzenia, iż do ich tworzenia niezbędne są kognitywne kompetencje notującego, rozumiane jako zdolność do aktywności umysłowej o charakterze poznawczym. Dotyczą one percepcji, rozumienia, uczenia się i działania, które nieodłącznie powiązane są ze zbieraniem informacji, formułowaniem hipotez, wyciąganiem wniosków i rozwiązywaniem zadań (Kubiński, 2005, 131; Piolat et al., 2005, 292 i 305). Uzasadnia to silną pozycję notatek w sferze kognitywnej i potwierdza konieczność posiadania umiejętności poznawczych niezbędnych do realizacji procesu notowania. Można przyjąć, że gatunki notatek jako schematy mentalne służą nie tylko wspomaganie pamięci, jak określono w przytoczonych definicjach, ale także ułatwiają tworzenie znaczeń i odnajdywanie sensu w typizowanym i powtarzalnym dialogu, prowadzonym pozornie z samym sobą. Gatunek ten umożliwia bowiem poprawną realizację działań społecznych, zarówno w kontekście zawodowym, jaki prywatnym. Notatki oddziałują bowiem na inne gatunki, stosowane w poprzednich lub przyszłych działaniach.

## 5. Analiza gatunkowa notatek studentów

W badaniu przeprowadzonym na przełomie maja i czerwca 2018 r. uczestniczyło ośmioro studentów studiów stacjonarnych Uniwersytetu Jagiellońskiego w Krakowie, będących na etapie pisania pracy dyplomowej. Trzy badane osoby to studenci studiów licencjackich, natomiast pozostali to studenci studiów drugiego stopnia. Grupę badawczą stanowią osoby, które zadeklarowały, iż tworzą notatki typu *information scraps*, co było głównym kryterium doboru próby. Kryterium kompletowania grupy badawczej był również kierunek studiów. Są to reprezentanci różnych specjalności, przy czym połowa z nich, z zamierzenia, jest studentami kierunku Zarządzanie informacją. W celu właściwej organizacji i prezentacji wyników badań ich uczestnicy zostali oznaczeni kodami, wykazanymi w tabeli 1, gdzie R oznacza respondenta.

Tab. 1. Struktura grupy badawczej

Oznaczenie respondenta	Kierunek studiów	Stopień studiów
R1	Biologia	I stopień
R2	Etnologia i antropologia kulturowa	I stopień
R3	Rachunkowość i zarządzanie finansami	II stopień
R4	Socjologia	I stopień
R5	Zarządzanie informacją	II stopień
R6	Zarządzanie informacją	II stopień
R7	Zarządzanie informacją	II stopień
R8	Zarządzanie informacją	II stopień

Wzorując się na podobnych badaniach przeprowadzonych przez Michaela Bernsteina i in., których przedmiotem były *information scraps* tworzone przez pracowników pięciu organizacji, zdecydowano się na użycie podobnych, wzajemnie komplementarnych, technik i metod badawczych. Są to wywiad oraz analiza gatunkowa zebranego materiału empirycznego w postaci karteczek (Bernstein et. al., 2008, 12). Badania oparte są więc na metodologii jakościowej i podzielone zostały na dwie części, odpowiadające obu wymienionym metodom i technikom.

W pierwszej części badań wykorzystano technikę swobodnego wywiadu ukierunkowanego (zob. Zał. 1). Rozmowy z respondentami trwały od 5 do 15 minut. Zebrany materiałem na tym etapie są spisane wypowiedzi studentów na temat tworzonych przez nich notatek typu *information scraps*. Respondenci byli pytani o częstotliwość ich tworzenia, najczęściej stosowane formy, cel tworzenia, nie tylko w kontekście studiowania, ale również podczas działań komunikacyjnych wynikających z życia codziennego. Jedno z pytań miało na celu zweryfikowanie, czy podział na sferę prywatną i „zawodową” przekłada się także na proces tworzenia *information scraps*. Zwrócono również uwagę na potrzeby informacyjne, które towarzyszyły respondentom podczas tworzenia notatek. Ponieważ wywiad poprzedzał zebranie danych empirycznych w postaci notatek, karteczek i skrawków, jedno z zadanych pytań dotyczyło zawartości przekazywanych materiałów, co było istotne dla dalszej analizy materiału empirycznego. Uzyskano także słowną zgodę na wykorzystanie

przykładów *information scraps*, pod warunkiem usunięcia danych osobowych lub innych danych mogących wskazywać na konkretne osoby.

W trakcie rozmów wyłoniono wiele przykładów wdrażania indywidualnie opracowanych i stosowanych zasad gatunkowych. Wywiad posłużył również wyłonieniu stanów afektywnych, emocji, odczuć towarzyszących badanym podczas notowania. W czasie prowadzenia rozmów notowano wszystkie wnioski i spostrzeżenia, dotyczące zarówno zauważonych prawidłowości, jak i nowo pojawiających się okoliczności. Następnym krokiem była całościowa analiza wywiadów, których treści podzielono na segmenty dotyczące celu stosowania *information scraps* w studiowaniu i życiu codziennym oraz zachowań w zakresie zarządzania *information scraps* przejawianych przez studentów.

Drugą część badań stanowiła jakościowa analiza danych, oparta przede wszystkim na technice analizy treści/zawartości danych zastanych w postaci karteczek i innych notatek zebranych od respondentów, w kierunku ich analizy gatunkowej. Uzyskany materiał to notatki wyłącznie w formie tradycyjnej, papierowej, których treści zapisywane były odręcznie. Część materiału została zarejestrowana na zdjęciach, co pozwoliło na pozostawienie oryginału właścicielowi. W sumie wykonano 42 zdjęcia, z których następnie wyodrębniono pojedyncze *information scraps* w liczbie 126 notatek. Natomiast karteczki w formie fizycznej występowały w 63 egzemplarzach, co łącznie daje materiał 189 oddzielnych notatek. Materiał empiryczny analizowany był kilkakrotnie i wieloaspektowo, co jest typowe dla analizy danych jakościowych. Podczas analizy danych jakościowych (ang. *qualitative data analysis, QDA*), łączenie gromadzenia danych z ich jednoczesną analizą jest sytuacją typową. Zgromadzony materiał przeanalizowano w kontekście rodzajów zawartych w nim treści w celu wyróżnienia grup tematycznych, oznaczonych następnie literami alfabetu. Zadbano przy tym o uporządkowanie karteczek i oznaczenie ich autorstwa. Podczas tematycznej analizy treści powracano do wypowiedzi respondentów w wywiadach, które precyzowały zawartość notatek. Analizę materiału empirycznego przeprowadzono następnie w odniesieniu do formy notatek: form językowych, jakimi posłużono się do zapisania informacji na *information scraps* oraz formy materialnej, kształtu i postaci fizycznej karteczek. Analiza gatunkowa zebranego materiału pozwoliła m.in. na wyróżnienie podgatunków notatek.

### 5.1. Cel komunikacyjny (wywiady standaryzowane)

Notatki powstają w związku z pojawianiem się nowych potrzeb komunikacyjnych związanych z określonymi celami. Częstotliwość ich tworzenia jest sprawą indywidualną. Badani najczęściej określali ją wyrażeniami: „często”, „bardzo często”, „regularnie”. Tylko jeden z respondentów (R3) stwierdził, że powstają one „nie aż tak często, na pewno nie codziennie”, za to inny (R4) przyznał, że zdarza mu się stworzyć nawet kilka w ciągu jednego dnia. Wzrost częstotliwości następuje przede wszystkim w okresie sesji egzaminacyjnej lub w czasie zdobywania zaliczeń czy podczas przygotowań do społecznych działań związanych z obowiązkami na uczelni. Wywiady wykazały, że powstają one również w momentach przypływu inspiracji, nagłego pojawienia się myśli lub wniosków uznanych za cenne. Także codzienne sytuacje, z którymi związane jest planowanie działań społecznych, stanowi cel tworzenia notatek. Są to najczęściej czynności określane przez studentów jako sprawy bieżące, takie jak zrobienie zakupów czy spotkania w określonym miejscu i czasie. To pozwala na stwierdzenie, iż *information scraps* powstają w sposób powtarzalny, zwykle ich tworzenie



jest nieplanowane, ale staje się początkiem planowania innych, wspólnych działań. Te, jako zawierające działania komunikacyjne, powodują stosowanie kolejnych gatunków informacji, służących osiągnięciu wspólnych celów. W tym przejawia się oddziaływanie tych gatunków.

*Information scraps* sprawdzają się podczas wykonywania czynności związanych ze studiami. Cztero respondentów (R1, R2, R3, R6) jednoznacznie stwierdziło, że to właśnie w tym celu powstaje ich najwięcej. Cele związane z sytuacjami codziennymi rzadziej bywają wspomagane przez notatki. Z kolei dla dwóch innych osób (R4, R8) sytuacja jest odwrotna. Tworzone przez nich notatki przeważnie dotyczą działań życia codziennego. Respondent R4 z zaskoczeniem uświadomił sobie, że pomimo zaawansowania jego pracy dyplomowej, notatki jej dotyczące pojawiały się sporadycznie. Przeważnie sięgał po nie w celu zanotowania tytułu lub sygnatury publikacji do wypożyczenia z biblioteki. Zapytany o powód tego zaniechania, stwierdził: „przeważnie do tego aspektu życia wykorzystuję komputer i to tam odbywa się wszystko, co związane ze studiami”. Natomiast pozostali badani (R5, R7) nie potrafili wyraźnie określić, która sfera jest dominująca, gdyż obie grupy celów skłaniają do generowania podobnej liczby notatek.

Respondenci zgodnie uznali, że notatki typu *information scraps* są dla nich przydatne w dużym stopniu. Dwoje z nich (R4, R8) wręcz stwierdziło, że nie wyobraża sobie funkcjonowania bez nich. Przyczyną była rola notatek we wspomaganiu osiągnięcia celów komunikacyjnych. Najczęściej wymienianą zaletą *information scraps* było zabezpieczenie przed zapomnieniem informacji. W drugiej kolejności wspomniano o porządkowaniu wiedzy/informacji, które ma zapobiec poczuciu chaosu i braku organizacji podejmowanych działań. Respondent R6 oświadczył, że *information scraps* są dla niego ochroną przed „bezmyślnością pracy”, co może oznaczać wspomaganie procesów poznawczych, w tym komunikacyjnych. Możliwość wyboru i werbalizacji właściwego sposobu postępowania uspokaja respondentów, a także w znacznym stopniu wpływa na systematyczność i konsekwentność ich postępowania. Kilukrotnie uznawana była także za czynnik motywujący. W takim wypadku *information scraps* traktowane są jako pierwszy krok do wykonania zaplanowanej czynności, przez co ułatwia mobilizację do kolejnych działań i stosowania kolejnych gatunków. W wypowiedziach badanych często wskazywano na oszczędność czasu rozumianą jako ograniczenie konieczności wielokrotnego wykonywania tych samych czynności. Z jednej strony, respondent R3, który wykorzystywał *information scraps* do oznaczania ważnych treści w literaturze<sup>8</sup> związanej z pracą dyplomową, łączył oszczędność czasu z ograniczeniem wysiłku związanego z powtórным wertowaniem wykorzystywanej publikacji w celu znalezienia interesującego fragmentu. Dzięki wstępnej selekcji istotnych fragmentów tekstu zmniejsza się zapotrzebowanie na pamięć roboczą, dodatkowe informacje kodowane są np. przy użyciu kolorów. Z kolei R6, mówiąc o oszczędności czasu, miał na myśli zminimalizowanie czasu poświęconego na przypomnienie sobie zaplanowanych czynności. Lista zadań pozwala objąć wzrokiem całość zaplanowanych działań, co przyczynia się do wzrostu poczucia posiadania kontroli i panowania nad sytuacją oraz własnymi poczynaniami. Badanie wykazało, że *information scraps* występują również w roli brudnopisu, z którego treści następnie przenoszone są np. do aplikacji komputerowych lub bardziej sformalizowanych (strukturyzowanych) gatunków. Postrzegane są w charakterze łącznika z innymi gatunkami oraz osobami i ich działaniami.

<sup>8</sup> Takie wykorzystanie notatek należy traktować jako narzędzie komunikacyjne (gatunek), ułatwiające znalezienie wspólnego gruntu i uzgodnienie znaczeń pomiędzy autorem tekstu i jego czytelnikiem.

Analiza postępowania części respondentów pozwoliła zidentyfikować kilka celów *information scraps* – estetyczny, edukacyjny oraz afektywny. Pierwszy związany był z formą karteczek, przede wszystkim kolorowych i umieszczanych w widocznych miejscach. Respondent R8 stwierdził, iż takie notatki są „przyjemne dla oka” i „ładnie to wygląda jak jest mała karteczka przyklejona” [pisownia oryg.]. Cel edukacyjny przejawia się wykorzystywaniem notatek bądź czynności notowania do nabywania nowych umiejętności. Student R2 przyznał, iż robiąc notatkę typu *to-dos* ze studencką listą zadań, wykorzystuje ją do ćwiczenia różnych stylów pisma odręcznego, połączonego z praktyką posługiwania się piórem. Respondent stara się wykonywać ją starannie, co ułatwia odbiór treści. Natomiast cel afektywny realizowany jest przez notatki, których treść ma szczególny, sentymentalny charakter dla ich autora. Zrobienie notatki pozytywnie oddziałuje na stan emocjonalny i samozadowolenie jednostki. Zwykle zachowywane są na dłuższy czas. Badany R5 ma tendencję do tworzenia takich treści, przeważnie zawierających cytaty lub myśli zapisywane po to, by „odkryć je za kilka lat”. Przechowuje on je w miejscach widocznych, aby mogły inspirować i motywować do działania lub gromadzi je w specjalnie przeznaczonym do tego zeszytcie. Działania takie uznaje za rodzaj dialogu z autorami cytatów, podczas którego czytelnik tworzy własny tekst (Smagorinsky, 2001, 134).

Wypowiedzi respondentów na temat celów tworzenia *information scraps* świadczą o ich powstawaniu w typizowanym procesie komunikacyjnym. Wykorzystując *information scraps* w różnych działaniach i kontekstach porozumiewają się, najczęściej z samym sobą, tworząc notatki na własny użytek. Taki sposób komunikowania można uznać za posługiwanie się własną wiedzą dla jej eksternalizacji w celu modyfikacji lub strukturyzacji (po internalizacji). W tym przypadku intencjonalnym odbiorcą informacji jest jej nadawca, a więc granica między nadawcą a odbiorcą zaciera się. Z drugiej strony, wiele notatek uznaje za rodzaj dialogu z autorami innych tekstów, innych gatunków. Wynika to z faktu, że notatka, tak jak każdy tekst, uzyskuje znaczenie od sieci innych tekstów, chociażby przez fakt wspólnej przynależności do określonego gatunku lub systemu gatunków. Oznacza to, że notatki są częścią systemu gatunków wspomagającego działania społeczne poprzez wielorakie oddziaływanie, towarzyszące procesom komunikacyjnym.

## 5.2. Treść notatek (jakościowa analiza danych)

Zebrany materiał badawczy, w liczbie 189 karteczek, został przeanalizowany pod względem treściowym i skategoryzowany w 15 kategoriach – podgatunkach. Ich oznaczenia literowe oraz liczba notatek przydzielonych do kategorii zostały zaprezentowane w tabeli 2. Kategorie zostały uporządkowane według częstotliwości występowania, od najczęściej do najrzadziej pojawiających się.

Zestawienie to wskazuje trzy najczęściej pojawiające się rodzaje treści *information scraps*, których liczebność znacznie przewyższa pozostałe kategorie. Są to myśli/opinie/wnioski/spostrzeżenia autora notatki, materiały dydaktyczne/definicje oraz *to-dos*. Pierwszą z nich i najliczniejszą (36) tworzą swobodne obserwacje, myśli i konkluzje, które są wytworem autora notatki. Są to treści wspomagające naukę, analizy pewnego zagadnienia, zanotowane ciekawostki lub inne informacje przydatne w kontekście pisanej pracy dyplomowej. Notatki te są więc elementem pracy nad wypowiedzią innego gatunku – pracy dyplomowej oraz dialogu z jej przyszłym czytelnikiem. W większości przypadków treść takich *information scraps* była stosunkowo długa,

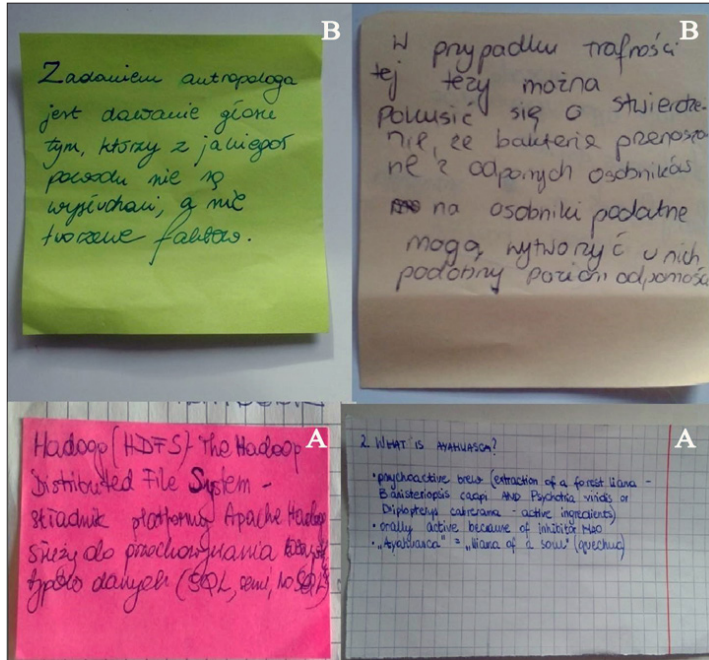
przyjmująca formę jednego, pełnego zdania. Drugą, równie liczną, kategorię (36) stanowią informacje, których zapisanie służy ich utrwaleniu w pamięci. Do tej grupy zaliczane są także definicje terminów, które, jak wynika z przeprowadzonych wywiadów, wspomagają proces nauki. Treść takich notatek najczęściej jest typowa dla klasycznej definicji, składa się z zdefiniowanego terminu (*definiendum*) oraz wyrażenia definiującego (*definiens*). Należy je traktować jako efekt czytania, a więc konwersacji z ich autorem. Przykłady notatek obu kategorii przedstawione zostały na rysunku 1 (w prawym górnym rogu znajduje się kod literowy kategorii).

Tab. 2. Kategorie treściowe badanych *information scraps* wraz z ich liczebnością

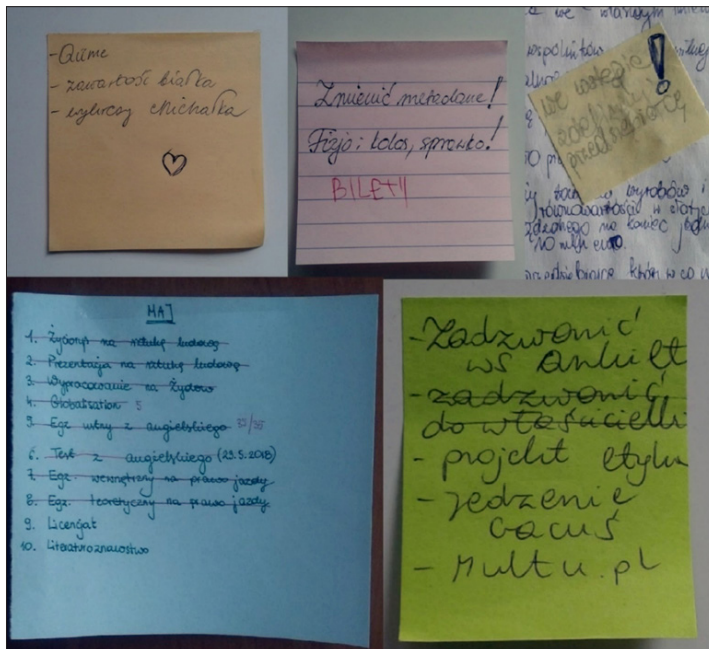
Oznaczenie literowe	Nazwa kategorii	Liczba notatek
A	Materiały dydaktyczne/definicje	36
B	Myśli/opinie/wnioski/spostrzeżenia autora notatki	36
C	<i>To-dos</i>	28
D	Cytaty/„złote myśli”	16
E	Etykiety/oznaczenia większej partii treści/informacji	12
F	Metadane publikacji	12
G	Wątpliwe kwestie do rozwiązania	12
H	Listy zakupów	10
I	Nazwy osobowe (imiona i nazwiska)	7
J	Oznaczenia liczbowe	7
K	Mieszane	4
L	Pojedyncze hasła	4
M	Specjalistyczne/dziedzinowe oznaczenia	2
N	Zaplanowane spotkania	2
O	Wykresy/rysunki	1
<b>Razem:</b>		189

Trzecią grupę, obejmującą 28 notatek, stanowią *information scraps* w formie list *to-dos* (Rys. 2). Zwykle występowały one w postaci wypunktowanych wykazów czynności do wykonania, lecz pojawiały się również notatki bez wyliczeń, które zaliczane były do tej grupy ze względu na ich treść, wskazującą na przynależność do tej kategorii. Czasami były to szybkie, jednowyrazowe notatki zawierające skróty, których stworzenie nie wymaga dużego wysiłku. Listom *to-dos* towarzyszyły także specjalne oznaczenia literowe, graficzne, cyfrowe lub znaki interpunkcyjne (np. wykrzykniki), niekiedy kolorowe lub pogrubione dla zaakcentowania wagi czynności. Kilkakrotnie pojawiały się wykreślenia albo zakreślenia wyrażające wykonanie lub zakończenie jakiegoś działania. Działania te często wiążą się ze stosowaniem innych gatunków informacji, co wskazuje na oddziaływanie na nie gatunku lista *to-dos*. Na przykład działanie *Zadzwoń w sprawie ankiet* (Rys. 2) spowoduje przeprowadzenie rozmowy telefonicznej w gatunku odpowiednim do jej kontekstu.

Należy zwrócić również uwagę na notatki występujące w kategorii D, czyli zapisane cytaty lub „złote myśli” (16 notatek). Treści tego rodzaju chętnie umieszczane są w *information scraps*. Do tej kategorii przyporządkowywano tylko treści, z których jednoznacznie wynikało, że są to cytaty innych osób niż autor notatki. One także są formą dyskusji z ich autorami. Kierowano się cechami formalnymi (cudzysłowy) lub wskazaniem autora wypowiedzi.

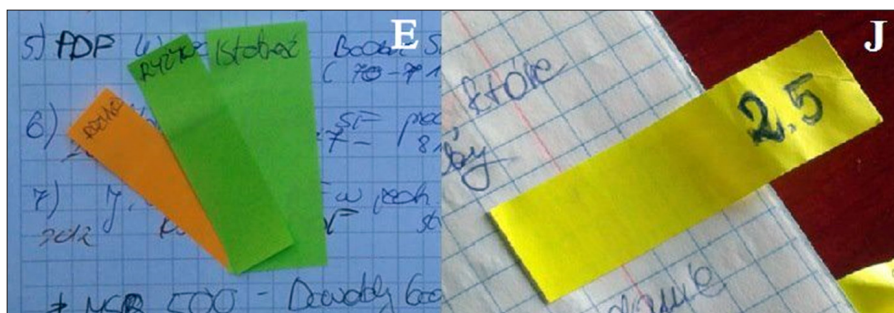


Rys. 1. Przykłady information scraps z kategorii A oraz B



Rys. 2. Sposoby wykorzystywania list to-dos (kategoria C)

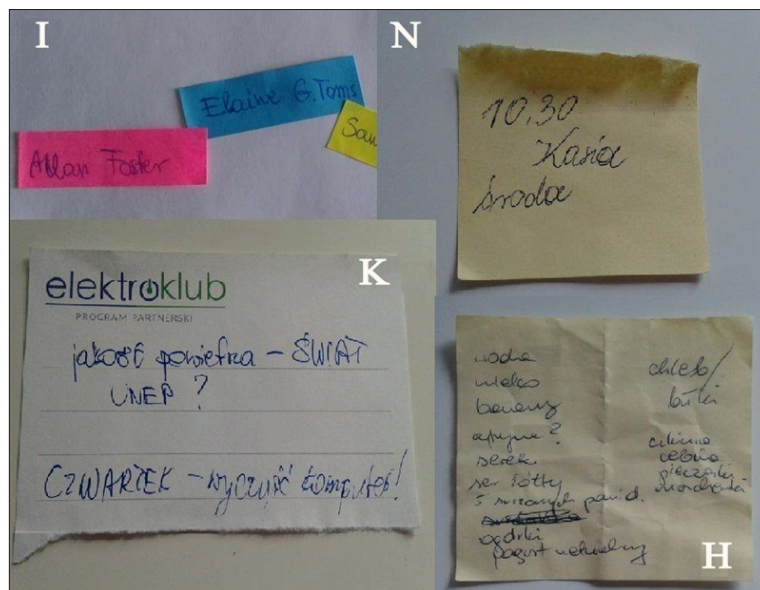
Taką samą liczbę (12) notatek zidentyfikowano w aż trzech kategoriach: E, F oraz G. Wątpliwe kwestie G to grupa notatek, których informacje uzupełniono znakami zapytania, sygnalizującymi wątpliwości autora. Są to nierozstrzygnięte wątpliwości, informacje wymagające sprawdzenia czy dookreślenia. Kategoria F, czyli metadane publikacji to informacje zawierające elementy opisu bibliograficznego: tytuł, tytuł/autor lub rok wydania, wskazujące na planowane działanie uzyskania dostępu do publikacji. Natomiast kategoria E odnosi się do *information scraps*, które pełnią rolę etykiet bądź oznaczeń większych partii innych treści. Notatki z tej kategorii pojawiały się z reguły na marginesie strony, np. przydatnej publikacji w celu zaznaczenia jej istotnego fragmentu. Należy tutaj zaznaczyć, że kategoria J również obejmuje notatki służące segregacji lub wskazywaniu większej partii informacji, jednak do ich oznaczenia wykorzystano cyfry, stąd takie *information scraps* przyporządkowano do oddzielnej grupy. Oba te gatunki stanowią formę dialogu z autorem publikacji. W celu zobrazowania różnic pomiędzy tymi dwiema kategoriami, przedstawiono ich przykłady na rysunku 3. Zbliżoną liczbę do trzech powyższych kategorii (10 przypadków) odnotowano dla list zakupów (podgatunek *to-dos*), oznaczonych literą H. Takie notatki zawierały wyłącznie spisy produktów do nabycia. Te gatunki oddziałują także z gatunkami stosowanymi w kolejnych działaniach, np. etykietami kupowanych produktów. Po 7 przykładów miały kategorie I oraz J. Kategoria I to imiona i nazwiska, najczęściej naukowców (autorów lektur), przez co treści te podobne są do kategorii F. W grupie J znajdowały się wspomniane już oznaczenia liczbowe.



Rys. 3. Różnice pomiędzy kategorią E oraz J

Jedną z ciekawszych *information scraps* ze względu na treść i formę są notatki kategorii K. Są to notatki, które łączą w sobie treści nawet kilku różnych podgatunków. W zebranych materiale empirycznym znaleźć można przykład połączenia listy zakupów z *to-dos* oraz kategorią G. Jeden z respondentów połączył kategorie C z G, jeszcze inny B z D. Na jednej z notatek zauważyć można wykorzystanie aż pięciu różnych kategorii: C, G, I, L i M. Wśród zebranych karteczek znalazły się też zawierające nietypowe treści, których dopasowanie do kategorii było niemożliwe ze względu na ich niepowtarzalny charakter. Były to ciągi cyfrowe bądź zestawienia fachowych danych, stąd decyzja o ich umieszczeniu w osobnej kategorii M. Do kategorii N – *Zaplanowane zadania* przyporządkowano dwie notatki, tyle samo co do kategorii M. Są to krótkie notatki, sprowadzające się do informacji o nadchodzącym spotkaniu, np. miejsce oraz czas. Są one wynikiem dialogu z inną osobą, w którym uzgodniono te informacje. Przykład notatek z tej grupy oraz kilku innych, przedstawiono na rysunku 4.

Analiza materiału wykazała, że *information scraps* w formie obiektów graficznych (rysunków) bez tekstów to pojedyncze przypadki. Tylko jedna z karteczek zawierała wyłącznie element graficzny. Był to naszkicowany wykres, który stanowi osobną kategorię O. Rysunki, np. strzałek, znaków wyboru, emotikonów, pojawiały się sporadycznie, tylko jako element uzupełniający treść tekstową.



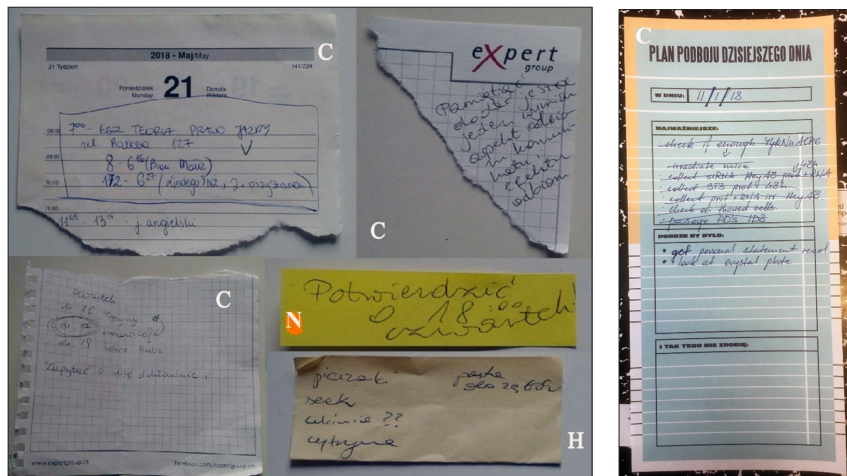
Rys. 4. Przykłady *information scraps* czterech kategorii: H, I, K, N

### 5.3. Forma notatek (jakościowa analiza danych)

Podczas wywiadów respondenci zwracali uwagę na formę notatek, zarówno fizyczną, jak i językową, jako przyczynę ich popularności. Forma umożliwia mobilność notatek, na co uwagę zwrócili respondenci R2 oraz R4, przyznając, że często towarzyszą im w podróży, np. tramwajem. Są one małe, mieszczą się w kieszeni, co traktowane jest jako istotna zaleta. Student R7 podkreślił, że woli tworzyć właśnie takie notatki, ponieważ sięgnięcie po długopis jest dla niego prostsze, szybsze i wygodniejsze niż wybór i zastosowanie narzędzia elektronicznego.

Analiza notatek pod względem ich formy fizycznej, a więc stosowanych rozmiarów oraz kształtów, wykazała, że najwięcej karteczek (161) miało postać *Post-it* i stanowiły one 86% całego zbioru. Ta liczba nie powinna dziwić, ponieważ respondenci, pytani podczas wywiadów o preferowane formy, stwierdzali, iż właśnie z takimi pracuje im się najłatwiej i najwygodniej. Były to karteczki samoprzylepne bądź nie, w różnych kolorach (żółte, różowe, zielone, białe), co miało znaczenie, m.in. podczas lokalizowania notatek w przestrzeni. Miały one różne kształty (kwadratowe, prostokątne) i wielkość. Uczestnicy badania tworzyli *information scraps* także w innych formach. Zaobserwowano notatki wykonywane na pociętych kartkach zeszytu, oderwanych skrawkach stron, a nawet wyrwanych kartkach z kalendarza (Rys. 5). Pojawiały się także kartki zeszytu A5, przeważnie

składane w połowie, służące do tworzenia list *to-dos*, zawierających znaczną liczbę zadań do wykonania. Respondenci wykorzystywali też materiały specjalnie przystosowane do tworzenia notatek. Wiele notatek, szczególnie listy zakupów (H) i kategorii A nosiły ślady intensywnego ich używania i noszenia (zagięcia, pomięcia).



Rys. 5. Przykładowe formy tworzonych *information scraps*

Analiza materiału badawczego pod kątem form językowych, jakim posługiwali się respondenci wykazała, że zależały one od formy fizycznej *information scraps*. Wiele z nich składało się z pojedynczych słów, haseł lub wyrażań, jednak pojawiały się też zawierające dłuższe, dokładniejsze wypowiedzi wyrażane pełnymi zdaniami. Hasłowe treści najczęściej pojawiały się na notatkach typu *to-dos* i to właśnie w tej kategorii język jest najbardziej zróżnicowany. Występują tu słowa kluczowe, pojedyncze wyrazy, które szybko skierowują uwagę na odpowiedni rodzaj czynności. W *information scraps* bardzo często pojawiały się skróty słów, np. *sprawko* zamiast *sprawozdanie* czy *zad* zamiast *zadanie*. Wypunktowane listy niekiedy pozbawione są czasowników, budowane z samych rzeczowników. Wymienić tu można takie słowa jak: *egzamin, esej, jedzenie, wejściówka, projekt, bilety, prezent, mail* oraz *test*. Natomiast na listach czynności znajdowały się takie czasowniki, jak: *poszukaj, dopisz, potwierdź, wziąć, zrobić, zmienić, wysłać*. Tryb rozkazujący i użyta majuskuła podkreślają wagę treści. Dłuższe i dokładniej sformułowane treści zawierały przeważnie notatki z kategorii B i częściowo z kategorii A. Teksty w nich umieszczane były znacznie bardziej rozbudowane, czasem kilkuzdaniowe.

## 6. Wnioski

Zastosowanie teorii gatunków informacji do analizy procesów PIM/GIM przyniosło interesujące efekty. W PIM/GIM stosowane są złożone systemy gatunków informacji wzajemnie na siebie oddziałujących. System taki tworzą również gatunki informacji nazywane w artykule notatkami lub *information scraps*. Informacje przekazywane za pośrednictwem tych gatunków są przedmiotem społecznych procesów komunikacyjnych, wspierają działania

kognitywne takie jak organizacja, analiza, wizualizacja informacji oraz uczestniczą w procesach jej internalizacji i eksternalizacji. Działania te prowadzą do tworzenia PSI jednostki w ścisłym związku z PSI innych jednostek. Gatunki, jako element wiedzy, wierzeń, doświadczenia i uprzedzeń, stanowią niezbędną część zarówno indywidualnej, jak również publicznej przestrzeni informacyjnej. Notatki, z zasady tworzone indywidualnie, dotyczą zwykle działań uwarunkowanych społecznie, więc ich miejsce w kontinuum PIM/GIM, choć zwykle bliskie indywidualnego końca, zawsze zawiera element działań grupowych.

Przedstawione badania gatunku *information scraps* wskazują na nieinstytucjonalne działania zarządzania informacją, które stanowią ważne uzupełnienie działań instytucjonalnych podobnego rodzaju, realizowanych przez biblioteki i inne systemy informacyjne, z wykorzystaniem zupełnie innych gatunków, znacznie bardziej ustrukturyzowanych (standaryzowanych). Badani studenci realizują działania PIM/GIM stale i regularnie, we wszystkich kontekstach, zarówno zawodowym (edukacyjnym), jak i w życiu codziennym, w grupach społecznych różnej liczebności. Wykazują różne schematy zachowań informacyjnych związanych ze stosowaniem notatek, od najprostszyc (cel komunikacyjny osiągnany najmniejszym kosztem), poprzez średnio złożone (np. zróżnicowanie formy i lokalizacji) po najbardziej złożone, mieszane (stosowanie kilku systemów gatunkowych jednocześnie).

Gatunek *information scraps*, ułatwiając indywidualne działania kognitywne, jest jednocześnie elementem strukturyzacji działań grupowych i społecznych służących zarządzaniu informacją. Dialog „z samym sobą” jest częścią budowania wspólnego gruntu i przyswajania wspólnych znaczeń. Na tym gruncie dochodzi do porozumienia w działaniu w sytuacjach standardowych i powtarzalnych. Stąd analiza gatunkowa notatek jest tylko jednym z celów przedstawionych badań. Drugim jest prezentacja działań w zakresie organizacji informacji, które tylko pozornie nie służą komunikowaniu tej informacji. W rzeczywistości *information scraps*, jak i inne podobne gatunki informacji, które przeznaczone są dla ich twórcy (np. pamiętniki), są częścią procesów komunikacyjnych, przeszłych i/lub przyszłych, w których stosowane są gatunki wzajemnie oddziałujące z gatunkiem notatek.

### Załącznik 1. Zestaw pytań zadawanych podczas wywiadu przeprowadzanego ze studentami

Lp.	Pytanie
1	Jak często tworzysz tego typu notatki?
2	Czy tworzysz je wyłącznie podczas pisania pracy dyplomowej oraz innych czynności związanych ze studiami, czy powstają one również w trakcie czynności życia codziennego?
3	Jaką najczęściej przybierają one formę?
4	W jakim celu przeważnie tworzysz takie notatki?
5	Jak bardzo są one dla Ciebie przydatne?
6	Czy podczas tworzenia notatek, kierujesz się własnym systemem bądź metodą, związaną np. z kolorem, szyframi, szczególnym ich umiejscowieniem?
7	Czego dotyczą i co zawierają przekazywane notatki?



## Bibliografia

- Andersen, J. (2017). Genre, organized knowledge, and communicative action in digital culture. In: J. Andersen & L. Skouvig (eds.) *The organization of knowledge: caught between global structures and local meaning* (1–16). Bingley: Emerald Publ. Ltd.
- Askehave, I., Swales, J. (2001). Genre identification and communicative purpose: a problem and a possible solution. *Applied Linguistics*, 22(2), 195–212.
- Bachtin, M. (1986). *Estetyka twórczości słownej*. Warszawa: PIW.
- Bernstein, M., van Kleek, M., Karger, D., Schraefel, M. C. (2008). Information scraps: how and why information eludes our personal information management tools. *ACM Transactions on Information System*, 26(4), art. 24.
- Crowston, K., Williams, M. (2000). Reproduced and emergent genres of communication on the World Wide Web. *The Information Society*, 16(3), 201–215.
- Detlor, B. (2010). Information management. *International Journal of Information Management*, 30, 103–108.
- Devitt, A. (2004). *Writing genres*. Carbondale, IL: Southern Illinois Univ. Press.
- Erickson, T. (2006). From PIM to GIM: personal information management in group contexts. *Communications of the ACM*, 49(1), 74–75.
- Franke, W. (1987). Texttypen – Textsorten – Textexemplare. *Zeitschrift für germanistische Linguistik*, 15(3), 263–281.
- Freadman, A. (2002). Uptake. In: R. Coe, L. Lingard & T. Teslenko (eds.) *The rhetoric and ideology of genre: strategies for stability and change* (39–53). Cresskill: Hampton.
- Gajda, S. (2009). Gatunki wypowiedzi i genologia. W: Z. Bilut-Homplewicz, W. Czachura, M. Smykała (red.) *Lingwistyka tekstu w Polsce i w Niemczech. Pojęcia, problemy, perspektywy* (135–146). Wrocław: Oficyna Wydaw. ATUT.
- Górny, M., Głowacka, E., Kisilowska, M., Osiński, Z. (2017). *Mechanisms of the formation and evolution of personal information spaces in the humanities*. Poznań: Wydaw. Rys.
- Hyland, K., Salager-Meyer, F. (2008). Scientific writing. *Annual Review of Information Science and Technology*, 42(1), 297–338.
- Isenberg, H. (1978). Probleme der Texttypologie. *Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig*, 27(5), 565–579.
- Jones, W. (2015). *Building a better world with our information: the future of personal information management, Part 3*. [Chapel Hill]: Morgan & Claypool Publ.
- Jones, W. (2007). Personal Information Management. *Annual Review of Information Science and Technology*, 41(1), 453–504.
- Jones, W., Thorsteinson, C., Thepvongsa, B., Garrett, T. (2016). Making it real: towards practical progress in the management of personal information. In: *CHI EA '16 Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (571–582). New York, NY: ACM.
- van Kleek, M., Styke, W., Schraefel, M. C., Karger, D. (2011). Finders/Keepers: A Longitudinal Study of People Managing Information Scraps in a Micro-note Tool. In: *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems (CHI '11)* (2907–2916). New York, NY: ACM.
- Kubiński, W. (2005). Dwa słowa o kognitywizmie i kognitywistyce. W: H. Kardela, Z. Muszyński, M. Rajewski (red.) *Kognitywistyka. Problemy i perspektywy* (131–136). Lublin: Wydaw. Uniwersytetu Marii Curie-Skłodowskiej.
- Kwaśnik, B., Crowston, K. (2005). Introduction to the special issue. Genres of digital documents. *Information Technology & People*, 18(2), 76–88.
- Lansdale, M. (1988). The psychology of personal information management. *Applied Ergonomics*, 19(1), 55–56.

- Lin, M., Lutters, W., Kim, T. (2004). Understanding the micronote lifecycle: improving mobile support for informal note taking. In: *CHI '04 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (687–694). New York, NY: ACM.
- Ljungberg, J. (2008). Combining Actor Network Theory and genre theory to understand the evolution of digital genres. *Sprouts: Working Papers on Information Systems*, 8(20).
- Luckmann, T. (2009). Observations on the structure and function of communicative genres. *Semiotica*, 17(1/4), 267–282.
- Luckmann, T. (1986). Grundformen der gesellschaftlichen Vermittlung des Wissens: Kommunikative Gattungen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 27, Sonderhefte, 191–211.
- Lutters, W., Ackerman, M., Zhou, X. (2007). Group information management. In: W. Jones, J. Teevan (eds.) *Personal information management* (236–248). Seattle, WA: University of Washington Press.
- Maciejewski, M. (2009). *Gatunki hipertekstu w perspektywie tekstologicznej. Analiza na przykładzie internetowych prezentacji przedsiębiorstw*. Poznań: WN UAM.
- Materska, K. (2012). Wymiary zarządzania informacją indywidualną. W: B. Sosińska-Kalata, E. Chuchro (red.) *Społeczeństwo i sieć informacyjna: problemy i technologie* (65–79). Warszawa: Wydaw. SBP.
- Materska, K. (2010). Rozwój koncepcji zarządzania informacją. W: D. Pietruch-Reizes, W. Babik (red.) *Zarządzanie informacją w nauce* (11–22). Katowice: Polskie Towarzystwo Informatyki Naukowej.
- Miller, C. (1984). Genre as social action. *Quarterly Journal of Speech*, 70(2), 151–167.
- Nahotko, M. (2018). *Teoria gatunków w organizacji informacji i wiedzy. Podejście informatologiczne*. Kraków: Wydaw. UJ.
- Notatka (1979). W: *Słownik języka polskiego*, T. 2 (395). Warszawa: PWN.
- Piolat, A., Olive, T., Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology*, 19(3), 291–312.
- Post-it (2017, August 8). In: *English Oxford Living Dictionary* [online], [8.08.2017], <http://en.oxford-dictionaries.com/definition/post-it>
- Post-it note (2017, August 8). In: *Cambridge Dictionary* [online], [8.08.2017], <http://dictionary.cambridge.org/dictionary/english/post-it-note?q=Post-it+note>
- Sapa, R. (2017). Grupowe zarządzanie informacją z perspektywy informatologicznej. W: A. Pulikowski (red.) *Kultura książki i informacji: księga jubileuszowa dedykowana Profesor Elżbiecie Gondek* (267–382). Katowice: Wydaw. UŚ.
- Shepherd, M., Watters, C. (1999). The functionality attribute of cybergenres. In: *Proc. of the 32nd Hawaii Intern. Conference on System Science* (2007–2015), Manui, Hawaii, January 5–8, 1999. Washington: IEEE.
- Smagorinsky, P. (2001). If meaning is constructed, what is it made from? Toward a cultural theory of reading. *Review of Educational Research*, 71(1), 133–169.
- Sticky (2017, August 8). *English Oxford Living Dictionary* [online], [8.08.2017], <https://en.oxforddictionaries.com/definition/sticky>
- Swales, J. (1990). *Genre analysis. English in academic and research settings*. Cambridge: Cambridge Univ. Press.
- Teevan, J., Jones, W., Capra, R. (2008). Personal Information Management. *ACM SIGIR Forum*, 42(2), 99–105.
- Tomaszczyk, J. (2017). Metodyka tworzenia i ewaluacji cyfrowych narzędzi indywidualnego zarządzania informacją. W: R. Sapa (red.) *Diagnostyka w zarządzaniu informacją: perspektywa informatologiczna* (209–222). Kraków: Bibl. Jagiellońska.
- Wilson, T. (2000). Human information behavior. *Informing Science*, 3(2), 49–55.
- Yates, J., Orlikowski, W. (1992). Genres of organizational communication: a structural approach to studying communication and media. *The Academy of Management Review*, 17(2), 299–326.
- Yates, J., Orlikowski, W., Rennecker, J. (1997). Collaborative genres for collaboration: Genre systems in digital media. In: *Proc. of the 13th Annual Hawaii Intern. Conference on System Sciences (HICCS 30)* (50–59), Wailea, HA. Washington, DC: IEEE.

## Information Scraps Genre in Individual/Group Information Management

### Abstract

**Purpose/Thesis:** The aim of the article is to present the role of information genres in individual (non-institutionalized) information management on the example of a genre system called notes or information scraps created in a traditional way (without the use of electronic tools).

**Approach/Methods:** The group of students was studied with two complementary methods: an unstructured interview with agenda and the content analysis of collected empirical material (handwritten notes). The interview enabled researchers to collect respondents' opinions on the purposes of using notes. The form and information content of the collected notes was analyzed. This method enabled the information scraps genre analysis, which, according to the literature of the field, should concern three characteristics of the genre: purpose, content and form, including language.

**Results and conclusions:** Information scraps are a genre system supporting information management. Their form and content are adjusted to the communication purposes: cognitive, affective, aesthetic and educational ones. They are used in the case of standardized and repetitive communication activities.

**Originality/Value:** The use of information scraps in individual activities helps create a common ground for the group activities, hence the genre discussed, just as any other genre, supports social communication activities.

### Keywords

Genre analysis. Individual/Group information management. Information genres. Information scraps.

---

*PAULINA KRZANOWSKA, magister zarządzania informacją na Uniwersytecie Jagiellońskim. Autorka pracy dyplomowej Information scraps jako forma indywidualnego zarządzania informacją przez studentów. Zawodowo związana z infobrokeringiem. Na co dzień zajmuje się wyszukiwaniem, analizą i selekcją informacji związanych z Kontraktami Armii USA.*

*Kontakt z autorką:*

*paulinakrzanowska08@gmail.com*

*Dr hab. MAREK NAHOTKO, adiunkt w Instytucie Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego, specjalizuje się w zagadnieniach metadanych dokumentów elektronicznych oraz wykorzystania Internetu w komunikacji naukowej i działalności informacyjnej. Jest autorem książek Teoria gatunków w organizacji informacji i wiedzy (Kraków, 2018), Komunikacja naukowa w środowisku cyfrowym (Warszawa, 2010), Naukowe czasopisma elektroniczne (Warszawa, 2007), Opis dokumentów elektronicznych. Teoretyczny model i możliwości jego aplikacji (Kraków 2006), Metadane: sposób na uporządkowanie Internetu (Kraków, 2004) oraz licznych artykułów publikowanych m.in. na łamach Przeglądu Bibliotecznego, Zagadnień Informacji Naukowej oraz Bibliotheca Nostra.*

*Kontakt z autorem:*

*marek.nahotko@uj.edu.pl*

*Instytut Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego*

*ul. prof. Stanisława Łojasiewicza 4*

*30-348 Kraków*

---

# O rozwoju graficznych języków komunikacji. Przykład wykorzystania UML w obszarze bibliologii i informatologii

Anna Małgorzata Kamińska

*Instytut Bibliotekoznawstwa i Informatyki Naukowej, Uniwersytet Śląski*

---

## Abstrakt

**Cel/Teza:** Celem artykułu jest przybliżenie koncepcji Unified Modeling Language (UML) oraz możliwości jego użycia w dziedzinie bibliologii i informatologii.

**Koncepcja/Metody badań:** Na podstawie analizy piśmiennictwa omówiony został proces kształtowania się języków graficznych oraz ich zalety związane z zapewnieniem zwięzłości i precyzji komunikacji. Analiza przypadku kategoryzacji wydawnictw wykorzystana została do zademonstrowania zasad modelowania pojęć z dziedziny bibliologii i informatologii za pomocą języka UML.

**Wyniki i wnioski:** Przedstawiony przykład modelowania za pomocą UML pojęć z zakresu bibliologii i informatologii pokazuje przydatność tego języka w opracowywaniu nowych metod i narzędzi dla bibliotekarstwa i usług informacyjnych.

**Oryginalność/Wartość poznawcza:** Chociaż UML jest narzędziem szeroko stosowanym w wielu dziedzinach, w piśmiennictwie nauki o informacji nie analizowano dotychczas jego przydatności do modelowania narzędzi i usług informacyjnych.

## Słowa kluczowe

Język graficzny. Komunikacja. Notacja graficzna. UML. Unified Modeling Language.

*Otrzymano: 31 marca 2017. Zrecenzowano: 7 października 2017. Poprawiono: 31 października 2017.  
Zrecenzowano: 19 grudnia 2017. Poprawiono: 4 września 2018. Zaakceptowano: 15 grudnia 2018.*

---

## 1. Wprowadzenie

Mechanizmy komunikacyjne, podobnie jak i inne wykształcone na drodze ewolucji, cechują się lokalnymi kierunkami przystosowania, tzn. ewoluują w stronę zaspokajania potrzeb konkretnych społeczności, które mogą być różne w zależności od ich lokalizacji geograficznej, dostępności surowców i innych zasobów czy też obecności społeczności konkurujących o przestrzeń i surowce. W przypadku ludzi, których duże potrzeby komunikacyjne doprowadziły do wykształcenia różnych języków etnicznych z odmiennym słownictwem, gramatyką i kontekstem, doprowadziło to do problemów komunikacyjnych pomiędzy poszczególnymi społecznościami.

Drugą konsekwencją ewolucyjnego powstawania języka jest redundantność jego przekazu, na co miała wpływ zapewne potrzeba zmniejszenia jego podatności na powstawanie błędów w komunikacji (Piantadosi et al., 2011). Błędy te mogły być powodowane przez osobnicze różnice w budowie aparatu głosowego, hałas lub inne zakłócenia ośrodka przekazu, potrzebę przesyłania komunikatów na większe odległości itp. Szczegółowymi aspektami źródeł tej

redundancji i jej skutków, tak pozytywnych jak i negatywnych, zajmują się językoznawcy, a podstawowe informacje z tej dziedziny przedstawiają np. Ernst Wit oraz Marie Gillette (Wit & Gillette, 1999). Warto zauważyć, że w warunkach kontrolowanej jakości środków przekazu (brak zakłóceń akustycznych, czytelne wydruki, monitory ekranowe o wyraźnym obrazie, itp.) niektóre powody istnienia redundancji językowych tracą rację bytu i niosą niepotrzebny balast, jednak w praktyce codziennej komunikacji międzyludzkiej redundancja jako taka jest niezbędna. Problemem jest zatem jej poziom, a nie jej istnienie lub jej brak.

Języki naturalne powstawały zatem w celu i w warunkach, dla których zapewnienie globalnej zrozumiałości i zwięzłości przekazu nie było istotne. Jednak rozwój cywilizacyjny ludzkości od epoki maszyn parowych, poprzez globalną elektryfikację, aż po rewolucję informacyjną, spowodował wzrost popytu na mechanizmy komunikacji i utrwalania informacji w postaciach zwięzłych i zrozumiałych niezależnie od pochodzenia kulturowego. Czynniki te wpłynęły zapewne w decydującym stopniu na rozwój alternatywnych języków czy też notacji graficznych<sup>1</sup>.

Język graficzny na potrzeby niniejszych rozważań został zdefiniowany jako język, którego wypowiedzi mają formę graficzną i jest to pierwotna forma tych wypowiedzi. Odróżnia to języki graficzne od języków naturalnych, których głównym subkodem jest subkod wokalny, a wypowiedzi w formie graficznej pełnią rolę wtórną. Języki graficzne wykorzystują najczęściej symbole graficzne składające się z elementarnych figur geometrycznych (np. linii, okręgów, wielokątów). Zestawienie i/lub łączenie tych symboli służy tworzeniu wyrażenia niosącego pewną informację. W wielu językach graficznych jest też możliwe wplatanie w strukturę takiego wyrażenia symboli alfanumerycznych, słów z pewnego języka naturalnego, a nawet i dłuższych wypowiedzi.

Choć języki graficzne znajdują obecnie zastosowanie głównie w dziedzinach technicznych, to od pewnego czasu obserwować można ich wykorzystanie w innych gałęziach. Kilka lat temu początkowy brak zainteresowania tymi językami sygnalizowała Małgorzata Stanula-Boroń (Stanula-Boroń, 2001) zauważając, że bywają one niedoceniane. Obecnie języki te usprawniają komunikację w wielu obszarach, np. w medycynie (Venot et al., 2008; Lamy et al., 2005) czy przemyśle (Dietsch et al., 2011). Przykład wykorzystania formalnego języka graficznego (Grafoo) do opisu zbioru ontologii publikacji i ich cytowań SPAR (The Semantic Publishing and Referencing Ontologies) przedstawiłam za jej twórcami przy okazji omówienia usług otwartego dostępu do indeksu cytowań (OpenCitations Corpus) (Kamińska, 2017a). Warto zwrócić uwagę na mnogość technik graficznego definiowania ontologii, w przeglądzie których (Slimani, 2015), zawierającym również odniesienie do języka Grafoo, autor stwierdza, że język UML jest nazbyt zredukowany i że języki bardziej rozbudowane pozwalają na lepszą reprezentację ontologii. Warto zauważyć, że stwierdzenie to stoi w opozycji do opinii wyrażanych przez wielu użytkowników języka UML, którzy jego prostotę postrzegają jako zaletę, która nie musi ograniczać możliwości jego stosowania, gdyż w języku tym istnieją mechanizmy pozwalające na jego rozszerzanie odpowiednio do własnych potrzeb.

<sup>1</sup> W ramach artykułu pojęcia notacja i język graficzny używane są zamiennie, choć w szerszym rozumieniu notacje graficzne dotyczą nie tylko symboli graficznych, ale np. również alfanumerycznych. W odniesieniu do języków graficznych notacja definiowana bywa również jako materializacja jego składni, czyli określenie graficznej reprezentacji poszczególnych elementów danego języka (Olczyk, 2010).

Przykłady zastosowania diagramów klas języka UML dla opisów ontologii, taksonomii oraz do doprecyzowania informacji wyrażanych innymi nieformalnymi notacjami graficznymi znaleźć można w innym moim opracowaniu (Kamińska, 2018b), natomiast wykorzystanie diagramu przypadków użycia do modelowania interakcji użytkownika z bibliograficzną bazą danych przedstawiłam w (Kamińska, 2017b).

## 2. Rozwój języków graficznych

Najstarszymi znanymi formami graficznymi wytworzonymi przez człowieka są rytzy i malowidła naskalne, które datuje się na dziesiątki tysięcy lat. Motywacje ich tworzenia przez ówczesnych ludzi nie są do końca poznane, a tezy zawarte w podręcznikach historii opisujących tamte odległe czasy i ujęcia tego zagadnienia przereklamowano już kilkakrotnie. Rozważając początki języków graficznych można zadać sobie pytanie czy wspomniane formy, określane również często jako jedne z pierwszych przejawów sztuki prehistorycznej, mieszczą się w definicji języka. Biorąc pod uwagę mnogość tych definicji można tutaj przywołać jedną z najszerzych (Drabik & Sobol, 2016):

spół sposób porozumiewania się ludzi pewnego środowiska lub zawodu oraz zapisu i przekazywania informacji w jakiejś dziedzinie wiedzy.

Przyjmując, że najnowsze badania funkcji społecznych tych najwcześniejszych form komunikacji graficznej wskazują na rolę obrazowania poglądu na porządek panujący w naturze, uznać można, że funkcje te mieszczą się w przytoczonej definicji języka.

Analizując ewolucję technik komunikacyjnych gatunku ludzkiego można zauważyć, że początkowo rozwijające się równolegle przekazy głosowe i graficzne w większości przypadków zbiegły się w pewnym momencie, przy czym przekaz głosowy zyskał status nadrzędny, a pismo stało się narzędziem wtórnym, które bywało do niego dostosowane (sytuacja taka widoczna jest w przypadku języków z pismem fonograficznym). Ponieważ utrwalone takim pismem procesy poznawcze muszą przejść dwukrotną transformację formy nośnika informacji, to w niektórych przypadkach mogą powodować rozmycie precyzji opisu. Dla przykładu, rysunek trójkąta i koła o ścisłym położeniu tych figur na płaszczyźnie z określonymi układami współrzędnych, po raz pierwszy przechodzi transformację w przypadku zidentyfikowania i zaklasyfikowania tych obiektów przez człowieka oraz nadania im nazw w języku, którym ten człowiek posługuje się (czyli wytworzeniu jego modelu pojęciowego oznaczonego etykietą językową). Druga transformacja następuje w sytuacji utrwalenia tego modelu pismem opisującym utrwalały model z perspektywy dźwiękowego brzmienia opisującej go wypowiedzi. O ileż prościej i na pewno bardziej precyzyjnie byłoby po prostu skopiować opisywany obrazek w celu przekazania go odbiorcy!

Spostrzeżenia te stały się szczególnie istotne wraz z nastaniem epoki renesansu, a z nią gwałtownym rozwojem architektury oraz początkami innych dyscyplin inżynierii, dając podstawy rysunkowi technicznemu (Tarełko, 2014). Używano go do dokumentowania i projektowania struktur złożonych za pomocą narzędzi kreślarskich, a główne jego formy abstrakcji stanowiły skala oraz rzuty. Powstające modele były bardzo dosłowne i brakowało im elementów opisujących zachowanie czy dynamikę reprezentowanych obiektów i konstrukcji. Sytuacja zaczęła się zmieniać gwałtownie wraz z nastaniem rewolucji

informacyjnej, kiedy to zaistniała potrzeba wyrażania pojęć nie tylko z dziedzin inżynierii, produktów, których obecność może być namacalna, ale idei bardziej ulotnych, wykraczających poza fizyczną percepcję człowieka i niejednokrotnie wchodzących w skład złożonych systemów pojęciowych.

### 3. Współczesne notacje graficzne

Współczesne języki graficzne są językami sztucznymi, tj. utworzonymi celowo do zapewnienia optymalnie sprawnej komunikacji w określonym zakresie. Języki te podlegały jednak także pewnemu procesowi ewolucji. Początek zainteresowania sztucznymi językami graficznymi zbiega się z początkiem światowej rewolucji informacyjnej. Początkowo, w latach 70. XX w., kierunki rozwoju tych języków stymulowane były przez nauki techniczne, w tym potrzebę modelowania struktury otaczającego nas świata i zależności występujących pomiędzy jego bytami, dla potrzeb projektowania struktur baz danych. Naukowcy i inżynierowie potrzebowali zwięzłych i czytelnych metod opisu złożoności modelowanych systemów. Mimo że pierwsze rozwiązania w zakresie tworzenia takich notacji nie dawały tak dużych korzyści jak notacje używane współcześnie, to i tak stanowiły istotny postęp w rozwoju języków graficznych. Często przyjmuje się, że pierwszym ważnym krokiem w ich rozwoju było wprowadzenie pojęć encji i relacji przez Petera Chena (Chen, 1976) z Massachusetts Institute of Technology. Mimo dużego zainteresowania tym nowym wówczas podejściem, jego szybkim rozwojem i niekwestionowaną przydatnością praktyczną (do końca lat 90. XX w. było to najczęściej stosowane podejście w modelowaniu struktur danych, a praktyczne zastosowanie znajduje nawet po dzień dzisiejszy), z perspektywy czasu łatwo zauważyć jego niedostatki i ograniczenia. Polegały one w głównej mierze na koncentrowaniu się na perspektywie strukturalnej (stąd w literaturze podejście to nazywane bywa metodyką strukturalną), z pominięciem aspektów behawioralnych. Drugim niedostatkiem jest zbytne uproszczenie postrzegania relacji zachodzących pomiędzy encjami, gdyż brak tutaj możliwości zobrazowania kierunku relacji, wyróżnienia relacji typu część-całość, a przede wszystkim możliwości wyrażania relacji abstrakcji. Bardziej szczegółowy opis rozwoju tego podejścia przedstawia jego twórca w swoim przeglądowym artykule (Chen, 2002).

Brak możliwości opisu aspektów behawioralnych nie był traktowany jako istotny problem, ponieważ wcześniej istniały już inne języki graficzne, które tworzono w celu modelowania takich aspektów rzeczywistości. Były to na przykład schematy blokowe (ang. *flowcharts*), których początki datuje się na lata 20. XX w., a ich wprowadzenie przypisuje się Frankowi i Lillian Gilbrethom (Gilbreth & Gilbreth, 1921) oraz diagramy przepływu danych (ang. *data flow diagrams*) umożliwiające graficzne wyrażanie modeli nie tylko przepływów danych w systemach komputerowych, ale również przepływów informacji w instytucjach czy ogólnie – wszelkich podsystemach operujących na zbiorze encji. Powstanie tego podejścia datuje się na lata 70. XX w., kiedy to opisano je po raz pierwszy w artykule *Structured Design* (Stevens et al., 1974).

Poza wcześniej wymienionymi, najbardziej rozpowszechnionymi notacjami, od początku rewolucji informacyjnej rozwijało się również wiele innych języków graficznych, które projektowano, aby rozwiązać problemy opisu rzeczywistości, z którymi nie radziły sobie te

powszechnie stosowane notacje. Były to m.in. diagramy FFBD (ang. *Functional Flow Block Diagram*), mapy procesów (ang. *Business Process Mapping*) i wiele innych, które łączyła wspólna słabość polegająca na wyrażaniu rzeczywistości z perspektywy, do obsługi której zostały stworzone, ale w oderwaniu od opisów tej rzeczywistości w innych perspektywach. Przykładowo, diagramy przepływu danych opisują perspektywę procesową systemów (np. w bibliotece czytelnicy generują strumień danych dotyczących zamawianych publikacji, które są przetwarzane przez bibliotekarzy i skutkują odnotowaniem odpowiednich informacji na karcie/koncie czytelnika), ale nie dają możliwości przedstawienia ich perspektywy strukturalnej (publikacje są bytami abstrakcyjnymi, które posiadają swoje konkretne egzemplarze, z których część jest dostępna, a część właśnie wypożyczona/udostępniona). Warto tutaj zauważyć, że mimo opisanych ograniczeń i związanego z nimi spadku zainteresowania diagramami przepływu danych, okazują się one być jeszcze ciągle wystarczające i pożyteczne w niektórych zastosowaniach (Mejor, 2012).

Wspomnianych wad pozbawione są graficzne języki nowej generacji, które za pomocą różnych tzw. widoków (ang. *view*) umożliwiają opisanie modelowanego bytu z różnych perspektyw. Elementy graficzne użyte w różnych widokach mają nierzadko wspólny mianownik i mogą być stosowane w więcej niż jednym widoku. Do najchętniej używanych języków tej nowej generacji należą BPMN (Chinosi & Trombetta, 2012), IDEF (Hanrahan, 1995) oraz język UML (ang. *Unified Modeling Language*).

UML jest dojrzałym, a jednocześnie ciągle aktywnie rozwijanym (ostatnia jego specyfikacja w wersji 2.5 opublikowana została w czerwcu 2015 r.) graficznym językiem opisu o uniwersalnym zastosowaniu. Początkowo był tworzony głównie na potrzeby informatyków, dzisiaj stanowi doskonale narzędzie do modelowania zjawisk, procesów czy struktur w różnych dziedzinach. Powstał na bazie koncepcji metod obiektowych, które narodziły się we wczesnych latach 90. XX w. w wyniku scalenia kilku istniejących i często używanych notacji zaproponowanych przez Grady'ego Boocha i Jamesa Rumbaugh. Prace rozpoczęte w sierpniu 1994 r. prowadzone w firmie Rational Software Corporation (włączonej w 2003 r. do firmy IBM) zaowocowały najpierw ustaleniem standardu (czyli *de facto* gramatyki) notacji, a następnie połączono metody opisu Grady'ego Boocha (Booch Method), Jamesa Rumbaugh (Object Modeling Technique) i Ivara Jacobsena (Object-Oriented Software Engineering) oraz kilka innych w jeden język, którego specyfikację opublikowano w wersji 0.9 w czerwcu 1996 r. Aktualną specyfikację UML w wersji 2.5, opublikowano online<sup>2</sup> (OMG, 2015) na blisko 800 stronach, przedstawiając ją w sposób formalny, nierzadko przy pomocy samego języka UML.

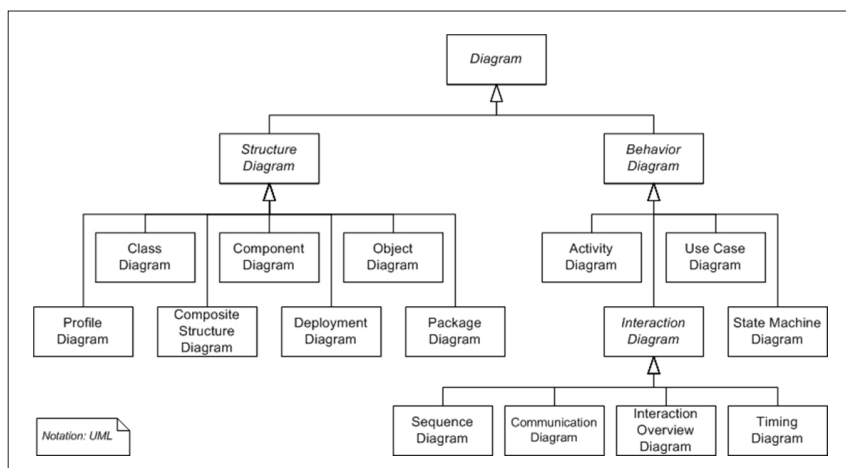
Samo rozwinięcie skrótu UML wskazuje, że jest to uniwersalny język modelowania, który można stosować do opisu dowolnego fragmentu istniejącej rzeczywistości. Uniwersalność UML ilustruje rysunek 1, pokazujący typy diagramów zdefiniowane w standardzie UML.

Wszystkie diagramy dzieli się na te, które opisują strukturę (służą do opisu strukturalnego) oraz te, które opisują zachowanie (służą do opisu behawioralnego). Wśród diagramów strukturalnych znajduje się diagram klas i jest to typ diagramu, który właśnie analizujemy (jest to przykład opisywania językiem UML konstrukcji języka UML). Wśród diagramów opisujących zachowanie znajduje się diagram aktywności (zwany też diagramem czynności), służący najczęściej do opisu kolejności i warunków czynności składających się na

<sup>2</sup> Pod adresem: <http://www.omg.org/spec/UML/2.5/>



realizację danego przepisu/algorytmu. Do diagramów zachowania należą też: diagramy interakcji (pokazujące interakcje, w które wchodzi między sobą obiekty danych klas w celu realizacji określonych celów), diagram przypadków użycia (służący do definiowania celów osób/podmiotów będących beneficjentami/składnikami danego systemu) oraz diagram stanów (opisujący stany, w jakich może znaleźć się obiekt danej klasy w ciągu całego cyklu życia systemu). Wśród diagramów opisujących aspekty strukturalne wyróżniono diagram klas (opisujący zależności między klasami bytów modelu), diagram obiektów (opisujący zależności między konkretnymi bytami), diagram profili (służący wizualizacji semantyki języka), diagram struktur złożonych (pokazujący wewnętrzną strukturę klasy i możliwości jej współpracy z innymi klasami), diagram komponentów (służący do „wysokopoziomowej” wizualizacji składowych modelowanego systemu), diagram wdrożenia (pokazujący fizyczne rozmieszczenie składowych w modelowanym systemie) oraz diagram pakietów (służący do wizualizacji podziału całego systemu na mniejsze w miarę niezależne od siebie fragmenty).



Rys. 1. Typy diagramów zdefiniowane w standardzie UML.

Źródło: [https://upload.wikimedia.org/wikipedia/commons/d/d6/Uml\\_diagram2.png](https://upload.wikimedia.org/wikipedia/commons/d/d6/Uml_diagram2.png)

Szczegółowy opis zastosowania poszczególnych diagramów wykracza poza ramy niniejszego opracowania, ale znaleźć go można w wielu publikacjach w postaci bardziej przyjaznej niż w źródłowej specyfikacji (np. Fowler, 2004).

Mimo że język UML znalazł do tej pory największe zastosowania w różnych dziedzinach nauk informatycznych, to stosuje się go również często w innych dziedzinach inżynierskich. Powodem takiego stanu rzeczy jest łatwość jego przyswojenia, czytelność, stopień rozpowszechnienia i zwięzłość przekazu. Autorka, opierając się na własnych doświadczeniach zdobytych podczas realizacji badań bibliometrycznych, przekonała się o jego praktycznej przydatności i zaletach wynikających z możliwości zastosowania wszędzie tam, gdzie wymagana jest precyzja i zwięzłość przekazu.

## 4. Przykład zastosowania UML do opisu pojęć z dziedziny bibliologii i informatologii

W celu przybliżenia języka UML i zachęcenia czytelnika do jego stosowania, w dalszej części artykułu przedstawione zostały podstawowe idee stojące za opisem złożoności struktury przy pomocy diagramów klas.

Jak piszą Booch, Rumbaugh i Jacobson (Booch et al., 2002, 47):

Klasy to najważniejsze bloki konstrukcyjne wszystkich systemów obiektowych. Klasa jest opisem zbioru obiektów, które mają takie same atrybuty, operacje, związki i znaczenie.

Większość podręczników języka UML właśnie od diagramu klas rozpoczyna omawianie poszczególnych technik notacji graficznych. Na przykładzie identyfikacji form wydawniczych publikacji naukowych omówione zostaną podstawowe konstrukcje gramatyczne, które pozwolą wprowadzić czytelnika w świat klas i relacji.

Jak można wyczytać w *Międzyinstytucjonalnym przewodniku redakcyjnym*<sup>3</sup> (European Union & Office for Official Publications, 2011, 81–82):

### Wydawnictwa zwarte

Zgodnie z normą ISO 2108:2005 (ISBN) wydawnictwo zwarte to wydawnictwo nieperiodyczne, publikowane jako całość w jednej części lub w kilku częściach (wydanych jednocześnie lub niejednocześnie). Wydawnictwa zwarte mogą być rozpowszechniane w dowolnej formie (książki drukowane w twardej lub miękkiej oprawie, książki na kasetach lub płytach CD i DVD, publikacje w języku Braille'a, publikacje w internecie, książki elektroniczne itp.).

Wydawnictwa zwarte wielotomowe składają się z określonej liczby osobnych tomów (wyjątkiem mogą być wydawnictwa zeszytowe). Wydawnictwa te są przewidziane i wydawane jako całość. Poszczególne części mogą mieć własne tytuły i informacje o odpowiedzialności.

Każdemu wydawnictwu zwartemu nadawany jest określony międzynarodowy znormalizowany numer książki ISBN (...).

### Wydawnictwa ciągłe

Zgodnie z normą ISO 3297:2007 (ISSN) wydawnictwo ciągłe to wydawnictwo rozpowszechniane na dowolnym nośniku, publikowane w częściach – zwykle oznaczonych numerycznie lub chronologicznie – ukazujących się w określonych lub nieokreślonych odstępach czasu, bez ustalonego z góry terminu zakończenia. Do wydawnictw ciągłych zalicza się również wydania stale aktualizowane. (...)

Wydawnictwa ciągłe obejmują m.in.:

wydawnictwa periodyczne i seryjne, publikowane w następujących po sobie lub osobnych częściach, bez ustalonego terminu zakończenia, zwykle oznaczone numerami (gazety, czasopisma, magazyny drukowane lub elektroniczne, roczniki, np. sprawozdania czy informatory, serie wydawnictw zwartych); wydania stale aktualizowane, których treść jest na bieżąco uaktualniana, bez ustalonego terminu zakończenia (bazy danych, uaktualniane na bieżąco wydawnictwa luźnokartkowe lub wydania internetowe, np. niniejszy Przewodnik w wersji on-line).

### Uwaga:

Seria wydawnicza to zbiór osobnych publikacji, z których każda posiada własny tytuł, ale wszystkie są opatrzone wspólnym tytułem obejmującym całość serii (tzw. tytułem serii). Poszczególne publikacje mogą być oznaczone numerem lub nie.

Każdemu wydawnictwu ciągłemu nadawany jest określony międzynarodowy znormalizowany numer wydawnictw ciągłych ISSN.

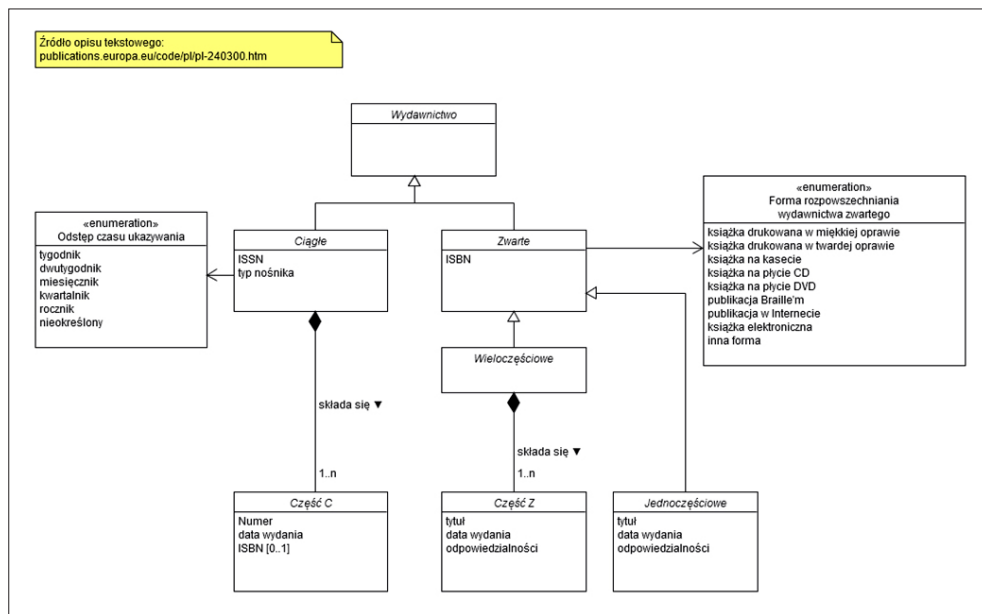
<sup>3</sup> Przewodnik jest również dostępny pod adresem: <http://publications.europa.eu/code/pl/pl-240300.htm>

### Wydawnictwa ciągłe i zwarte

Niektóre publikacje będące częścią wydawnictwa ciągłego (np. roczniki i wydawnictwa zwarte wydane w ramach serii) można dodatkowo traktować jako wydawnictwa zwarte (np. w celu udostępnienia ich zarówno oddzielnie, jak i w prenumeracie).

Publikacje takie należy przede wszystkim uznać za wydawnictwa ciągłe i nadać im numer ISSN (tytuł serii). Następnie należy nadać im numer ISBN (tytuł tomu)”.

Przedstawiony w zacytowanych definicjach i opisach model pojęciowy przedstawiono za pomocą diagramu klas UML (Rys. 2).



Rys. 2. Diagram klas kategoryzacji wydawnictw

Diagram można zacząć „czytać” od prostokąta *Wydawnictwo*. Prostokąty na diagramie klas reprezentują klasy, czyli w tym przypadku pojęcia, których znaczenie chcemy uchwycić graficznie. Nazwa wspomnianej klasy napisana jest kursywą, co oznacza, że jest to klasa abstrakcyjna, czyli taka, która ze względu na swoją ogólność nie może zostać zmaterializowana. Warto tutaj wspomnieć, że fizyczne rozmieszczenie jakichkolwiek elementów graficznych, długości czy sposoby załamania poszczególnych linii, w świetle definicji gramatyki języka UML nie mają znaczenia i nie niosą żadnych informacji. Jednak, podobnie jak w języku mówionym, na zrozumiałość przekazu wpływa nie tylko gramatyka i semantyka, ale również forma wypowiedzi – kolejność użycia poszczególnych części zdania, sposób kładzenia akcentów itp. W tym konkretnym przypadku strzałka skierowana w stronę klasy *Wydawnictwo* i brak strzałek od niej wychodzących sugerują, że jest to najbardziej ogólne pojęcie, od którego warto zacząć analizę diagramu. W demonstrowanym przykładzie wrażenie to pogłębiło stosując dodatkowo konwencję zgodną z przyzwyczajeniem do czytania tekstów z góry do dołu i umieszczając tę klasę na górze diagramu. Do tego typu dodatkowych konwencji należy też użycie kolorowania elementów graficznych.

Analizując tekst źródłowy opisu kategoryzacji wydawnictw widzimy, że wydawnictwa dzielą się na zwarte i ciągle (układ tekstu sugeruje, że są to kategorie rozłączne) i takie właśnie klasy zostały wydzielone na diagramie. Z klasą *Wydawnictwo* łączy je związek specjalizacji (zamknięta strzałka), co oznacza, że klasy te „dziedziczą” wszystkie cechy klasy bazowej, a dodatkowo rozszerzają je o swoje własne cechy. Jak czytamy w tekście źródłowym, wydawnictwa zwarte (informację o nieperiodyczności pominiemy z braku jej doprecyzowania) publikowane są w jednej lub wielu częściach (wydanych jednocześnie lub niejednocześnie). Charakterystykę tę odzwierciedlają dwie klasy *Jednoczęściowe* i *Wieloczęściowe*. Na końcu akapitu opisującego wydawnictwa zwarte czytamy, że każdemu wydawnictwu zwartemu nadawany jest numer ISBN, stąd na diagramie obecność tego atrybutu właśnie w klasie *Zwarte*. Wspomniane zostają formy publikacji wydawnictw, jednak z tekstu nie wynika, czy publikacja następuje w jednej z podanych form, czy też możliwa jest publikacja w kilku formach. Z tekstu nie wynika też czy w przypadku wydawnictw wieloczęściowych każda z części może być opublikowana w innej formie. Na diagramie lista możliwych form publikacji przedstawiona została jako klasa o stereotypie *enumeration*, a wartości tej listy jako lista jej atrybutów. Stereotypy to sposoby nadawania klasom szczególnych znaczeń, a stereotyp *enumeration* jest typowym i najczęstszym sposobem wyrażania typów wyliczeniowych, czyli takich, których wartości stanowią ograniczony zbiór nazwanych elementów. Klasę *Zwarte* łączy relacja nawigacji (strzałka otwarta) z klasą *Forma Rozpowszechniania Wydawnictwa Ciągłego*, co oznacza, że jeden z atrybutów tej klasy wyliczeniowej jest skojarzony z klasą *Zwarte*.

W tekście źródłowym można dalej przeczytać, że wydawnictwa zwarte wielotomowe składają się z określonej liczby osobnych tomów (wprowadzenie terminu tomu znowu jest tutaj mylące i niepotrzebne wobec wcześniejszego używania terminu wydawnictw jedno- i wieloczęściowych – termin ten zatem zostanie pominięty). Stąd na diagramie klasa *Część Z(wartego)* jest w relacji kompozycji (linia zakończona zaczernionym rombem) z klasą *Wieloczęściowe*. Związek kompozycji jest szczególnym przypadkiem związku agregującego, wskazującym na fakt, że obiekt klasy *Część Z* nie może zaistnieć jako samodzielny byt, bez wcześniejszego powołania obiektu klasy *Wieloczęściowe*. Związek ten został opisany wyrażeniem *Składa się* zapożyczonym wprost z tekstu źródłowego, a dla precyzji dodano strzałkę wskazującą kierunek interpretacji tego wyrażenia.

Jak wynika z dalszej części tekstu źródłowego, każda część wydawnictwa zwartego może mieć własne tytuły czy informacje o odpowiedzialności. Tę charakterystykę odzwierciedlają atrybuty dodane do klasy *Część Z*. Warto zauważyć, że tekst nie precyzuje, że wydawnictwa jednoczęściowe powinny zawierać informacje o tytule czy odpowiedzialnościach. Z tekstu wynika, że wydawnictwa ciągle (informację o periodyczności pominięto z tych samych względów co wcześniej) to wydawnictwa publikowane na dowolnym nośniku (dlaczego nie można się tu odwołać do nośników zdefiniowanych już dla wydawnictw zwartych?), publikowane w częściach. Przedstawia to klasa *Części C*, opatrzona, zgodnie z dalszym fragmentem tekstu, atrybutem opisu numerycznego bądź chronologicznego.

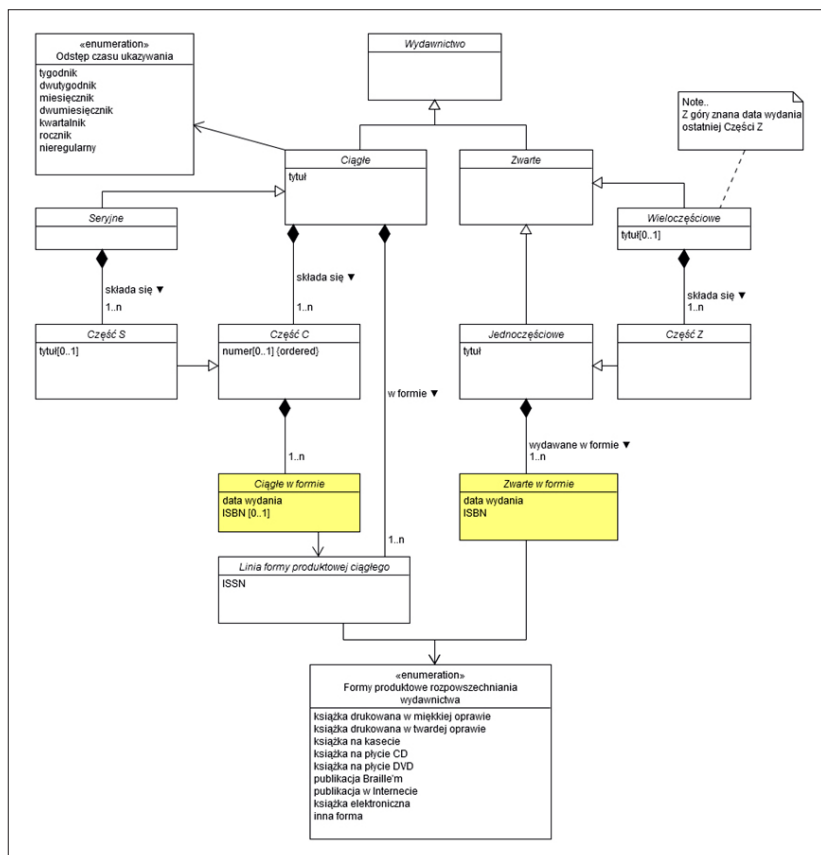
Po charakterystyce wydawnictw periodycznych i stale aktualizowanych następuje uwaga, w której wspomina się o seriach wydawniczych jako o zbiorze osobnych publikacji posiadających tytuł wspólny i tytuł każdej osobnej publikacji – nie wspominając wcześniej nic o sposobie nadawania tytułów dla publikacji ciągłych. Ponieważ każdemu wydawnictwu ciągłemu nadawany jest numer ISSN, na diagramie atrybut go reprezentujący dodany został

do klasy *Ciągłe*. Klasa wyliczeniowa *Odstęp czasu ukazywania* pozwala zakwalifikować wydawnictwo ciągłe do odpowiednich kategorii.

Ostatnia uwaga o możliwości nadawania niektórym częściom publikacji ciągłych numerów ISBN została uwzględniona poprzez dodanie odpowiedniego atrybutu do klasy *Część C* z użyciem notacji [0...1] wskazującej, że wartość tego atrybutu jest opcjonalna (może wystąpić raz lub wcale).

Tak skonstruowany diagram pozwala na łatwą i szybką orientację w zależnościach między pojęciami oraz ich atrybutami. Sprzyja to analizie abstrakcyjnej, porządkowaniu informacji i wprowadzaniu odpowiednich poprawek.

Rysunek 3 zawiera diagram obrazujący jedną z możliwości dokładniejszego opisu przykładowej dziedziny. Przedstawia on poprawiony model pojęciowy, opisujący formy i zasady publikacji wydawnictw ciągłych i zwartych. Należy podkreślić, że u podstaw istoty modelowania leży często potrzeba przekazania obrazu rzeczywistości z pewnej perspektywy, a więc ujawniającego wszystkie istotne szczegóły, a pomijającego detale nieistotne z punktu widzenia celu budowy modelu. Często lepiej zbudować kilka mniejszych diagramów, ukazujących dziedzinę z różnych perspektyw niż próbować umieścić wszystko na jednym diagramie.



Rys. 3. Diagram klas poprawionej kategoryzacji wydawnictw

Przedstawiony diagram nie ujawnia zależności związanych z autorstwem, redakcją, korektą czy innymi rodzajami odpowiedzialności. Nie opisuje również wewnętrznej struktury danego wydawnictwa, na którą składają się rozdziały, spisy treści itp., choć oczywiście opisanie tych zależności podobnym diagramem jest możliwe.

Należy również mieć na uwadze, że opatrzenie wydawnictw numerem ISBN bądź ISSN nie jest w żaden sposób obligatoryjne, a jedynie zalecane w celu uporządkowania rynku wydawniczego. Procedury wynikające z norm czy określone przez poszczególne jednostki zarządzające tymi numerami są jedynie zaleceniami, a poszczególni wydawcy zaopatrujący się z góry w pulę tych numerów, zarządzają nimi samodzielnie nie podlegając w zasadzie żadnej kontroli, co często prowadzi do wielu anomalii. Taki stan rzeczy nie musi wynikać ze złej woli wydawców, a jedynie z bałaganu informacyjnego, który mógłby być uporządkowany przez bardziej precyzyjne opisanie zaleceń, tak w formie słownej jak i graficznej.

Diagram należy „czytać” następująco: wydawnictwa dzielą się na dwie podstawowe kategorie – wydawnictwa ciągłe i wydawnictwa zwarte. Wydawnictwa zwarte mogą być publikowane jako jednoczęściowe o określonym tytule i wydawane w jednej lub wielu różnych formach produktowych rozpowszechniania, z których każda ma odrębny numer ISBN i datę wydania. Wśród form produktowych rozpowszechniania wydawnictw wymienić można:

- druk w miękkiej oprawie;
- druk w twardej oprawie;
- płyta CD;
- inne.

Drugą kategorią wydawnictw zwartych są wydawnictwa wieloczęściowe. Poszczególne ich części wydawane są dokładnie na tych samych zasadach co wydawnictw jednoczęściowych, a różnic je może jedynie istnienie wspólnego tytułu obejmującego wszystkie części składowe. W przypadku wydawnictw wieloczęściowych znana jest z góry przewidywana data wydania ostatniej części (na diagramie opisuje to symbol notatki, którego można użyć do definiowania ograniczeń w sposób bardziej lub mniej formalny).

Wydawnictwa ciągłe natomiast to takie, które ukazują się cyklicznie pod wspólnym tytułem, w określonych bądź nieregularnych odstępach czasu. Każde wydawnictwo ciągłe może być publikowane w jednej lub wielu liniach formy produktowej rozpowszechniania, z których każda powinna posiadać odrębny numer ISSN. Poszczególne części wydawnictwa ciągłego mogą być opatrzone numerem kolejnym (na diagramie wyraża to notka {ordered}) i każda z tych części może być opublikowana w ramach wybranych linii formy produktowej, z określoną datą wydania i możliwością nadania oddzielnych numerów ISBN dla każdej linii z osobna. Szczególnym przypadkiem wydawnictw ciągłych są wydawnictwa seryjne składające się z części wydawanych na dokładnie tych samych zasadach co części wydawnictw ciągłych, ale z możliwością nadania osobnych tytułów dla każdej z części.

## 5. Wnioski

Początkowo niedoskonałe próby wypracowania graficznych narzędzi komunikacji międzyludzkiej zaowocowały ostatecznie powstaniem m.in. języka uniwersalnego – UML, który z wykorzystaniem prostych i intuicyjnych notacji umożliwia opisywanie złożonych systemów czy tworzenie ich modeli w różnych perspektywach.

Celem powstania języka UML nie było zastąpienie żadnych innych używanych powszechnie notacji graficznych, w szczególności tych opisujących ilościowe aspekty systemów czy wyniki i rezultaty badań. W takich przypadkach skuteczniejsze są inne techniki przedstawiania informacji w postaci wizualnej, jak np. różnego rodzaju wykresy czy mapy prezentujące opisywaną rzeczywistość. Odległości pomiędzy poszczególnymi elementami graficznymi reprezentującymi opisywane byty, proporcje ich wielkości czy kolory użyte dla poszczególnych elementów graficznych pozwolą lepiej zobrazować zależności opisywanej i „mierzonej” dziedziny. Wszędzie natomiast tam, gdzie pożądana jest precyzja przekazu, a więc np. w przypadku badań naukowych, prac inżynierskich, organizacji i zarządzania, a także projektowania metod i narzędzi wykorzystywanych w działalności bibliotecznej i informacyjnej, warto rozważyć wykorzystanie UML.

Warto tutaj zauważyć, że język UML, podobnie jak inne języki graficzne służące modelowaniu dziedzin, czy inne notacje graficzne umożliwiające przedstawianie ilościowego czy jakościowego charakteru danych (czyli ich tzw. wizualizacji), czy też wreszcie języki naturalne, stanowią jedynie narzędzia w rękach ich użytkowników. Skuteczność komunikacyjna każdego języka w praktyce zależy nie tylko od jego budowy, ale również od umiejętności jego wykorzystania przez stronę komunikującą oraz poziomu percepcji strony komunikowanej, a te zależą od takich czynników jak:

- stopień znajomości budowy danego języka;
- stopień znajomości dziedziny przedmiotowej, do opisu której użyto danego języka;
- uwarunkowania kulturowe (kierunek czytania komunikatów, sposób interpretacji barw, itp.);
- stopień rozpowszechnienia danego języka;
- przyzwyczajenia wynikające z wcześniej wykorzystywanych metod komunikacji.

Wszystko to sprawia, że bardzo trudno, o ile w ogóle byłoby to możliwe, w sposób formalny dowodzić wyższości jednego narzędzia komunikacyjnego nad drugim, czy też stawiać się w roli arbitra rozsądzającego celowość wykorzystania danego narzędzia w konkretnym przypadku. Duża i nadal rosnąca popularność języka UML w dziedzinach technicznych pozwala sądzić, że język ten może być też użyteczny w opracowywaniu nowych metod i narzędzi dla bibliotekarstwa i usług informacyjnych, zwłaszcza takich jak kategoryzacje, modele semantyczne, taksonomie, modele dziedziny czy ontologie.

Przykłady praktycznego wykorzystania elementów języka UML w dziedzinie naukometrii znaleźć już można np. w opracowaniu autorki dotyczącym definicji struktury modelu analitycznego zastosowanego do budowy narzędzia wspierającego analizę cytowań (Kamińska, 2018a), jak również w artykule poświęconym grafowym reprezentacjom danych opisujących dziedzinę publikowania (Kamińska, 2018c).

## Bibliografia

- Booch, G., Rumbaugh, J., Jacobson, I. (2002). *UML przewodnik użytkownika*. Warszawa: Wydaw. Naukowo-Techniczne.
- Chen, P. (1976). The Entity-Relationship Model: Toward a Unified View of Data. *ACM Transactions on Database Systems*, 1(1), 9–36.
- Chen, P. (2002). Entity-Relationship Modeling: Historical Events, Future Trends, and Lessons Learned. In: M. Broy & E. Denert (eds.) *Software Pioneers* (296–310). Berlin: Springer Verlag.

- Chinosi, M., Trombetta, A. (2012). BPMN: An introduction to the standard. *Computer Standards & Interfaces*, 34(1), 124–134.
- Dietsch, D., Arenis, S. F., Westphal, B., Podelski, A. (2011). Disambiguation of industrial standards through formalization and graphical languages. *2011 IEEE 19th International Requirements Engineering Conference* (265–270). DOI: 10.1109/RE.2011.6051634
- Drabik, L., Sobol, E. (2016). *Słownik języka polskiego*. Warszawa: Wydawnictwo Naukowe PWN.
- European Union, Office for Official Publications. (2011). *Międzyinstytucjonalny przewodnik redakcyjny*. Luksemburg: EUR-OP.
- Fowler, M. (2004). *UML Distilled: A Brief Guide to the Standard Object Modeling Language*. Harlow: Addison-Wesley.
- Gilbreth, F., Gilbreth, L. (1921). Process Charts: First Steps in Finding the One Best Way to do Work. *ASME Transactions*, 43(1818), 1029–1050.
- Hanrahan, R. P. (1995). *The IDEF Process Modeling Methodology* [online]. Software Technology Support Center, [31.10.2017], <http://www.sba.oakland.edu/faculty/mathieson/mis524/resources/readings/idef/idef.html>
- Kamińska, A. M. (2017a). OpenCitations (I4OC) – otwarty indeks cytowań publikacji naukowych [online]. *Biuletyn EBIB*, 6(176), [29.08.2018], <http://open.ebib.pl/ojs/index.php/ebib/article/view/551/746>.
- Kamińska, A. M. (2017b). ProBIT – prospektywna metoda tworzenia trawersowalnych indeksów cytowań a współczesne problemy organizacji przestrzeni informacji w tradycyjnych bibliograficznych bazach danych. *Zagadnienia Informacji Naukowej*, 1(109), 66–82.
- Kamińska, A. M. (2018a). ScientoMiner ICR – moduł importu danych bibliograficznych z zasobów Crossref dla platformy Gephi. *Zagadnienia Informacji Naukowej*, 1(111), 96–113.
- Kamińska, A. M. (2018b). Wykorzystanie języka UML do eliminowania wieloznaczności w komunikatach naukowych. W: B. Sosińska-Kalata, P. Trafiłowski, Z. Wiorogórska (red.), *Nauka o informacji w okresie zmian: innowacyjne usługi informacyjne* (69–85). Warszawa: Wydaw. SBP.
- Kamińska, A. M. (2018c). Zastosowanie struktur grafowych do analiz bibliometrycznych i webometrycznych. Modele i metody. *Nowa Biblioteka. Usługi, Technologie Informacyjne i Media*, 2(29), 47–63.
- Lamy, J. B., Duclos, C., Rialle, V., Venot, A. (2005). Which Graphical Approaches should be Used to Represent Medical Knowledge? *Studies in Health Technology and Informatics*, 116, 719–724.
- Mejor, J. D. (2012). Przepływ danych w placówce informacyjnej call center. *Zagadnienia Informacji Naukowej*, 2(100), 100–119.
- Olczyk, D. (2010). Modelowanie strukturalne – definicje, notacja, techniki i narzędzia. *Zeszyty Naukowe Warszawskiej Wyższej Szkoły Informatyki*, 4, 87–98.
- OMG (2015). *OMG Unified Modeling Language* [online], [06.12.2016], <http://www.omg.org/spec/UML/2.5/PDF/>
- Piantadosi, S. T., Tily, H., Gibson, E. (2011). The communicative function of ambiguity in language. *Cognition*, 122(3), 280–291.
- Slimani, T. (2015). Ontology development: A comparing study on tools, languages and formalisms. *Indian Journal of Science and Technology*, 8(24), 1–12. DOI: 10.17485/ijst/2015/v8i34/54249
- Stanula-Boroń, M. (2001). Informacja, język i wiedza w ujęciu Karla R. Poppera. *Zagadnienia Informacji Naukowej*, 1(77), 3–16.
- Stevens, W., Myers, G., Constantine, L. (1974). Structured Design. *IBM Systems Journal*, 13(2), 115–139.
- Tarełko, W. (2014). Zarys historii pojęć związanych z projektowaniem maszyn. *Zeszyty Naukowe Politechniki Śląskiej. Transport*, 83(1904), 279–286.
- Wit, E., Gillette, M. (1999). What is Linguistic Redundancy? *Technical Report The University of Chicago* [online], [06.12.2016], <http://www.math.rug.nl/~ernst/linguistics/redundancy3.pdf>



Venot, A., Ouard, P., Bar-Hen, A., Duclos, C., Lamy J. B. (2008). An iconic language for the graphical representation of medical concepts. *BMC Medical Informatics and Decision Making*, 8(1). DOI: 10.1186/1472-6947-8-16

---

## On the Development of Graphical Languages of Communication. An Example of the Use of UML in the Field of Book and Information Sciences

### Abstract

**Purpose/Thesis:** The aim of the paper is to present the concept of Unified Modeling Language and discuss the possibilities of its application in the field of book and information sciences.

**Approach/Methods:** Based on the analysis of literature, the process of graphical languages development and their advantages related to the coherence and precision of communication are discussed. The case study of the categorization of publication forms is used to demonstrate the use of UML in modeling terms from the field of book and information sciences.

**Results and conclusions:** The presented example shows usefulness of UML for the development of new methods and tools in library and information services.

**Originality/Value:** Although UML is a tool widely used in many fields, the literature on information science has not analyzed yet its suitability for modeling of information tools and services.

### Keywords

Communication. Graphical language. UML.

---

*Dr ANNA MAŁGORZATA KAMIŃSKA jest adiunktem w Instytucie Bibliotekoznawstwa i Informatyki Naukowej Uniwersytetu Śląskiego w Katowicach oraz pracownikiem Biblioteki Głównej Politechniki Śląskiej w Gliwicach. W 2016 r. obroniła na Wydziale Filologicznym Uniwersytetu Śląskiego w Katowicach rozprawę doktorską pt. Informacja naukowa o górnictwie w świetle wydawnictw ciągłych uczelni technicznych w Polsce (1945–1989). Jej zainteresowania naukowe skupiają się wokół trzech uzupełniających się obszarów: informatologia, graficzne języki komunikacji oraz wizualizacja informacji.*

*Kontakt z autorką:*

*anna.kaminska@us.edu.pl*

*Instytut Bibliotekoznawstwa i Informatyki Naukowej*

*Uniwersytet Śląski w Katowicach*

*Plac Sejmu Śląskiego 1*

*40-032 Katowice*

---

# Architektura informacji w piśmiennictwie zagranicznym 1982–2018. Analiza ilościowa

Anna Matysek

Wydział Filologiczny, Instytut Bibliotekoznawstwa i Informacji Naukowej  
Uniwersytet Śląski w Katowicach

---

## Abstrakt

**Cel/Teza:** Celem artykułu jest przedstawienie procesu kształtowania się zainteresowania problematyką architektury informacji w światowym piśmiennictwie.

**Koncepcja/Metody badań:** Rozwój literatury dotyczącej architektury informacji przebadano na podstawie analizy bibliometrycznej piśmiennictwa zagranicznego na ten temat, opublikowanego do końca 2018 r. Piśmiennictwo to zostało zidentyfikowane na podstawie baz Scopus, Web of Science i Library and Information Science Abstract oraz serwisów Google Scholar, Google Books i księgarni Amazon. Analizę ilościową rozkładu chronologicznego i tematycznego piśmiennictwa uzupełnia wskazanie najważniejszych autorów, książek i czasopism, a także najczęściej cytowanych publikacji.

**Wyniki i wnioski:** Zidentyfikowano 2269 publikacji dotyczących problematyki architektury informacji o bardzo zróżnicowanej tematyce, przede wszystkim z zakresu systemów informacyjnych i nauki o informacji. Wyniki analizy wskazują, że zainteresowanie różnymi aspektami architektury informacji zwiększa się nieprzerwanie od 2000 r., co zostało potwierdzone rosnącą liczbą publikacji i ich cytowań.

**Oryginalność/Wartość poznawcza:** W polskim piśmiennictwie nie podejmowano dotychczas próby analizy literatury z zakresu architektury informacji. Artykuł może być punktem wyjścia do pogłębionej analizy tematycznej oraz stanowić materiał porównawczy dla podobnych badań.

## Słowa kluczowe

Analiza bibliometryczna. Analiza ilościowa piśmiennictwa. Architektura informacji. Architektura systemów informacyjnych. Nauka o informacji.

Otrzymano: 1 lutego 2019. Zrecenzowano: 26 lutego 2019. Poprawiono: 9 marca 2019. Zaakceptowano: 13 marca 2019.

---

## 1. Wprowadzenie

Architektura informacji (AI), rozumiana jako działalność praktyczna związana z porządkowaniem informacji i nadawaniem im logicznych struktur, funkcjonuje odkąd ludzie nauczyli się komunikować (Skórka, 2016, 557). Sam termin jest stosunkowo młody, początkowo w latach 70. XX w. funkcjonował jako „the architecture of information” (León, 2014) w jednym z projektów realizowanych w firmie Xerox. Za ojca terminu „information architecture” i pioniera tej dyscypliny uważa się Richarda Saula Wurmana (Cooper, 2009). Ważną rolę w przybliżeniu omawianego zagadnienia odegrała książka *Information architecture for the World Wide Web*, której pierwsze wydanie ukazało się w 1998 r. (Morville & Rosenfeld, 1998) i w której autorzy w przystępny sposób omówili komponenty AI oraz

proces i metodologię jej projektowania. Rodrigo Ronda León przeanalizował najważniejszą literaturę z zakresu architektury informacji z lat 80. i 90. XX wieku, wskazując, że jest to nowa i praktyczna dyscyplina, w związku z czym niewiele jest tekstów poświęconych tematowi historycznym i teoretycznym (León, 2014). W kolejnej analizie literatury z omawianego zakresu wyszczególniono sześć kluczowych aspektów AI (jakość informacji, metadane, informacja i proces, metodologia, kwestie organizacyjne oraz wyzwania i udane praktyki) oraz wykazano, że wciąż odgrywa ona ważną rolę w zapewnianiu efektywnego dostępu do informacji (Martin et al., 2010).

Wraz z rozwojem technologii informacyjno-komunikacyjnych wzrastało znaczenie architektury informacji jako dyscypliny kształtującej cyfrowe przestrzenie informacyjne zgodnie z potrzebami użytkowników. Rozwój Internetu i różnorodność jego usług spowodowały nie tylko zwiększenie ilości informacji, ale nastąpiły także rewolucyjne zmiany w dostępie do niej, w sposobach gromadzenia, organizowania i prezentowania zasobów informacyjnych, informacja stała się wszechobecna, stawiając nowe wyzwania przed omawianą dyscypliną (Resmini & Rosati, 2011).

W Polsce koncepcję architektury informacji jako pierwszy przedstawił Stanisław Skórka (Skórka, 2002), a Barbara Sosińska-Kalata wykazała, że temat ten stanowi ważny obszar badawczy w nauce o informacji (Sosińska-Kalata, 2007). W ostatnich latach wzrasta zainteresowanie tą dyscypliną w aspekcie kształcenia akademickiego (Tafiłowski, 2016). Jednakże krajowa literatura w tym zakresie nie jest zbyt obszerna. Dominują teksty S. Skórki, jest także kilka artykułów innych autorów. Ponadto na język polski przetłumaczono kilka książek: *Architektura informacji w serwisach internetowych* drugie (Rosenfeld & Morville, 2003) i czwarte wydanie (Rosenfeld, Morville, & Arango, 2017), *Wzorce wyszukiwania* (Morville & Callender, 2011) oraz *Projektowanie nawigacji stron WWW* (Kalbach, 2012).

Ze względu na skromny zasób polskojęzycznych publikacji podejmujących problematykę AI, na co wskazywał też Piotr Tafiłowski (Tafiłowski, 2016), oraz rosnącą popularność kształcenia w tym zakresie, postanowiono przyjrzeć się światowej literaturze zagranicznej na ten temat, poczynając od pierwszych poświęconych mu tekstów, które opublikowano w 1982 roku. Celem artykułu jest przybliżenie procesu kształtowania się zainteresowania architekturą informacji na podstawie analizy ilościowej piśmiennictwa dotyczącego tej tematyki, a także wskazanie najważniejszych dla niej autorów i czasopism oraz najczęściej cytowanych publikacji. Ponadto wyszczególnione zostały dyscypliny, w obrębie których najwięcej publikuje się na temat AI i na tym tle sprawdzono stopień zainteresowania tą problematyką w nauce o informacji.

## 2. Metodologia

Cel artykułu wymagał utworzenia możliwie kompletnej bazy danych publikacji dotyczących architektury informacji. Głównymi źródłami danych zostały bazy Web Of Science Core Collection (WoS) i Scopus, które indeksują po kilkanaście tysięcy tytułów czasopism oraz monografie i materiały konferencyjne, ale różnią się zawartością<sup>1</sup> oraz Proquest Library

<sup>1</sup> O różnicach między Scopus a Web of Science pisali m.in. (Chadegani i in., 2013), (Falagas, Pitsouni, Malietzis, & Pappas, 2008).

and Information Science Abstracts (LISA), będąca jedną z dwóch największych baz bibliograficznych z zakresu nauki o informacji. Opisy bibliograficzne w uwzględnionych bazach, przypisane są do kategorii tematycznych lub dyscyplin, zawierają słowa kluczowe oraz informacje o cytowaniach. Opisy te pozwalają zatem na przeprowadzenie wieloaspektowej analizy ilościowej piśmiennictwa.

Zawartość wybranych baz została przeszukana w dniach 1–2 marca 2019 r. za pomocą terminu „information architecture”. Przeprowadzono też wyszukiwanie za pomocą terminu „architecture of information”, jednak w wyniku uzyskano zbiór tych samych rekordów, co dla wyszukiwania za pomocą pierwszego terminu, uzupełniony o publikacje dotyczące wyłącznie technicznych aspektów architektury systemów informacyjnych. W bazie WoS termin „information architecture” wyszukiwany był w polu Topic (pole to obejmuje przeszukiwanie tytułu, abstraktu i słów kluczowych). Uzyskano w ten sposób 936 rekordów dokumentów. W bazie Scopus wyszukiwanie objęło pola Title-Abstract-Keywords, a w jego rezultacie otrzymano 1835 rekordów dokumentów. Kwerenda w bazie LISA, wszędzie z wyjątkiem pełnego tekstu (NOFT) dała w wyniku 598 rekordów. Po usunięciu zdublowanych opisów z trzech baz otrzymano 2269 rekordów dokumentów. W zbiorze tym znalazły się rekordy 20 książek, wśród których zabrakło kilkunastu znanych autorce tytułów, w związku z czym wykonana została oddzielna kwerenda w celu wyszukania wydawnictw zwartych dotyczących analizowanego tematu. Korzystając z wymienionych już baz oraz Google Books, Google Scholar i księgarni internetowej Amazon.com jako źródeł danych, wyszukano w sumie 76 opisów książek.

### 3. Analiza literatury z zakresu architektury informacji

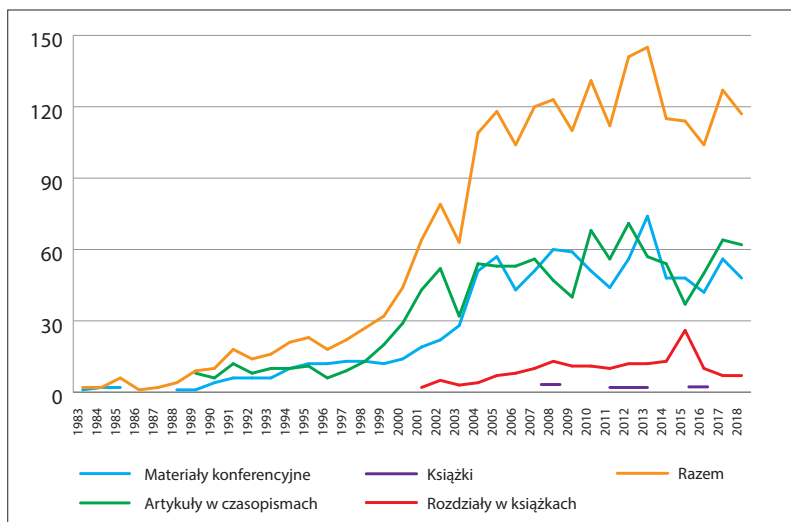
#### 3.1. Zasoby Scopus, WoS i LISA

W zbiorze wyszukanych 2269 publikacji dotyczących problematyki AI znalazło się 1100 artykułów z czasopism naukowych i branżowych, 975 tekstów w materiałach konferencyjnych, 174 rozdziały w książkach oraz 20 książek. Dominują dokumenty w języku angielskim. Najstarsze teksty zostały opublikowane w 1982 r., a pierwszy artykuł z „information architecture” w tytule wydany został w 1983 r., co jest zgodne z informacją podaną przez R. León (León, 2014). W tabeli 1 zestawiono liczbę publikacji w kolejnych dekadach, z podziałem na ich rodzaje. Zagadnienia związane z projektowaniem przestrzeni informacyjnych podejmowane są nieprzerwanie od lat 80. XX w., a zainteresowanie tematem wciąż wzrasta.

Tab. 1. Typy publikacji o architekturze informacji w kolejnych dekadach

	Książki	Rozdział w książce	Materiały konferencyjne	Artykuł w czasopiśmie
1983–1989	0	1	10	17
1990–1999	1	1	94	105
2000–2009	7	64	404	459
2010–2018	12	108	467	519

Wykres na rysunku 1 prezentuje, jak kształtowała się liczba publikacji o architekturze informacji w kolejnych latach z podziałem na ich rodzaje. Do 1999 r. nie przekraczała ona 30 pozycji. Można zauważyć znaczący wzrost liczby dokumentów od 2000 r., wtedy też zaczęły pojawiać się rozdziały poświęcone problematyce AI w książkach. Najwięcej – 145 tekstów, z czego ponad połowę stanowią materiały konferencyjne – wydano w 2013 r. Niemal w całym analizowanym okresie wśród form wydawniczych publikacji o AI przeważają artykuły, od 2001 r. publikowano ich minimum 29 rocznie, a najwięcej – 71, w 2012 r. Wyliczenia za 2018 r. mogą nie być pełne, gdyż w marcu 2019 r. część danych mogła jeszcze nie zostać zarejestrowana w bazach, w związku z czym mniejsza liczba publikacji nie musi oznaczać spadającego zainteresowania tematem. Książki publikowane były przede wszystkim w latach 2005–2016, ale nie więcej niż 3 rocznie.



Rys. 1. Liczba publikacji w kolejnych latach

W tabeli 2 znajduje się wykaz 24 autorów wraz z ich afiliacjami, którzy w wyszukanym zbiorze mają więcej niż 5 publikacji. Wśród nich znalazło się ośmiu przedstawicieli krajów europejskich, pięciu ze Stanów Zjednoczonych, po trzech z Australii, Chin i Malezji oraz dwóch z Brazylii. Autorem największej liczby tekstów jest Davide Bolchini. Jest on profesorem w Indiana University School of Informatics and Computing w Stanach Zjednoczonych i pełni funkcję kierownika katedry Human-Centered Computing. Realizuje projekty z zakresu interakcji człowiek-komputer (HCI) i doświadczeń użytkownika (UX). Kolejny z autorów – Marijn Janssen – ma w zbiorze 13 publikacji. Jest on profesorem w Delft University of Technology i kierownikiem sekcji the Information and Communication Technology. Jego zainteresowania koncentrują się na projektowaniu architektury technologii informacyjno-komunikacyjnych wykorzystywanych we współpracy między organizacjami zarówno publicznymi, jak i prywatnymi. Kolejne 11 publikacji jest autorstwa Martina White’a, założyciela Intranet Focus Ltd oraz profesora wizytującego w Information School na Uniwersytecie w Sheffield w Wielkiej Brytanii. Jego prace ukazały się przede wszystkim w magazynie *EContent* (ISSN: 1525–2531), podejmującym tematykę

publikowania cyfrowego, mediów i marketingu. W tabeli 2 umieszczono także nazwiska autorów popularnych książek o architekturze informacji: Andrea Resmini i Luci Rosatiego – po osiem ich publikacji znalazło się w analizowanym zbiorze.

Tab. 2. Autorzy, którzy opublikowali najwięcej prac o architekturze informacji

Lp.	Autor	Afiliacja	Liczba publikacji
1	Bolchini, Davide	Indiana University, Stany Zjednoczone	16
2	Janssen, Marijn	Technische Universiteit Delft, Holandia	13
3	White, Martin	Intranet Focus Ltd, Wielka Brytania	11
4	Isa, Wan Abdul Rahim Wan Mohd	Universiti Teknologi, Malezja	10
5	Victor E. Sanvido	Pennsylvania State University, Stany Zjednoczone	9
6	Summers, Tyler	The University of Texas at Dallas, Stany Zjednoczone	9
7	Garcia, Roberto	Universitat de Lleida, Hiszpania	8
8	Yu, Changbin B	Australian National University, Australia	8
9	Tan, Chekfoung	University of Reading, Wielka Brytania	8
10	Zhang Jinjiang	Zhejiang University, Chiny	8
11	Rosati, Luca	Università per Stranieri di Perugia, Włochy	8
12	Noor, Nor Laila Md	Universiti Teknologi, Malezja	8
13	Resmini, Andrea	Jönköping University, Szwecja	8
14	Anderson, Brian D O	Australian National University, Australia	8
15	Dubey, Abhishek	Vanderbilt University, Stany Zjednoczone	8
16	Lima-Marques, Mamede	Universidade de Brasília, Brazylia	7
17	Karsai, Gabor	Vanderbilt University, Stany Zjednoczone	7
18	Mehad, Shafie	Universiti Teknologi, Malezja	7
19	Rocha, Alvaro	Universidade do Porto, Portugalia	7
20	Vidotti, Silvana Aparecida Borsetti Gregorio	São Paulo State University, Brazylia	7
21	Brunetti, Josep M	Universitat de Lleida, Hiszpania	7
22	Burford, Sally	University of Canberra, Australia	7
23	Liu Yun	Beijing Institute of Technology, Chiny	6
24	Shieh, Jiann-Cherng	National Taiwan Normal University, Chiny	6

Tabela 3 zawiera wykaz 21 czasopism, w których opublikowano co najmniej siedem artykułów na temat architektury informacji. Dominują czasopisma amerykańskie (dziewięć tytułów), ponadto jest tam pięć periodyków brytyjskich, trzy brazylijskie, i po jednym z Chin, Hiszpani, Kanady i Kuby. Najwięcej – 25 artykułów – opublikowano w hiszpańskim czasopiśmie *Profesional de la Informacion*, co stanowi niewiele ponad 2% wszystkich wyszukanych artykułów. W brazylijskim *Informacao & Sociedade-Estudos* opublikowano

21 tekstów. Kolejne czasopismo na liście to wydawany w Stanach Zjednoczonych *Journal of the American Society for Information Science and Technology* (od 2013 r. ukazuje się pod zmienionym tytułem *Journal of the Association for Information Science and Technology*). W następnych, także amerykańskich czasopismach na wykazie – *EContent* oraz *Information Today*, opublikowano 18 tekstów, a w *Technical Communication* ukazało się 17 tekstów.

Tab. 3. Czasopisma z największą liczbą artykułów o architekturze informacji

Lp.	Tytuł czasopisma	ISSN	Liczba artykułów
1	Profesional de la Informacion	1386-6710	25
2	Informacao & Sociedade-Estudos	0104-0146	21
3	Journal of the American Society for Information Science and Technology / Journal of the Association for Information Science and Technology	1532-2882 / 2330-1635	19
4	EContent	1525-2531	18
5	Information Today	8755-6286	18
6	Technical Communication	0049-3155	17
7	Bulletin of the American Society for Information Science and Technology	2373-9223	15
8	International Journal of Information Management	0268-4012	11
9	Journal of the China Society for Scientific and Technical Information	1000-0135	9
10	ACIMED	1024-9435	9
11	Online	0146-5422	9
12	Bulletin of the American Society for Information Science	0095-4403	9
13	Aslib Proceedings / Aslib Journal of Information Management	0001-253X / 2050-3806	9
14	Journal of Web Librarianship	1932-2909	8
15	Library Hi Tech	0737-8831	8
16	Perspectivas Em Ciencia Da Informacao	1413-9936	8
17	Journal of Medical Internet Research	1438-8871	7
18	Government Information Quarterly	0740-624X	7
19	Informação & informação	1981-8920	7
20	The Electronic Library	0264-0473	7
21	Library Journal	0363-0277	7

Większość tytułów (18) obejmuje swym zakresem naukę o informacji (z czego 15 wyłącznie tę dyscyplinę), trzy podejmują jednocześnie tematykę informatyczną, dwa dotyczą nauk o zdrowiu, a jeden tytuł przypisany jest do zagadnień komunikacji masowej. Spośród periodyków wymienionych w tabeli 3 w bazie Journal Citation Index jest indeksowanych 14, a 12 z nich przypisanych jest w niej do nauk o informacji i bibliotece, w tym *Journal of the American Society for Information Science and Technology* i *Aslib Proceedings* (od

2014 r. *Aslib Journal of Information Management*) jednocześnie do informatyki/systemów informacyjnych. Czasopismo *Technical Communication* reprezentuje jedną dyscyplinę (komunikacja), a *Journal of Medical Internet Research* opiekę zdrowotną i informatykę medyczną. Zakres tematyczny wymienionych 21 czasopism świadczy o tym, że zagadnienia AI poruszane są w periodykach głównie w obszarze nauki o informacji.

W przeszukiwanych bazach nie było indeksowane czasopismo *Journal of Information Architecture*<sup>2</sup> (ISSN 1903–7260), którego redaktorem naczelnym jest Andrea Resmini. Jest ono poświęcone badaniom i rozwojowi architektury informacji we współdzielonych środowiskach informacyjnych z różnych perspektyw – technicznej, kulturowej, społecznej i komunikacyjnej. Jest to półrocznik, w latach 2009–2013<sup>3</sup> ukazało się osiem numerów. Razem z artykułami wstępnymi rozpoczynającymi każdy numer, opublikowano w nim łącznie 35 tekstów.

Architektura informacji jest interdyscyplinarna, wśród jej kluczowych obszarów A. Resmini i L. Rosatti wymieniają m.in.: projektowanie informacji, projektowanie wizualne, naukę o informacji i bibliotekoznawstwo, psychologię poznawczą i architekturę (Resmini & Rosati, 2011, 22). W zbiorze 2269 publikacji znajdują się teksty zróżnicowane tematycznie, dlatego podjęto próbę ustalenia głównych dyscyplin wykorzystując w tym celu klasyfikacje i tematy stosowane w indeksowaniu w bazach stanowiących źródło danych badawczych. Bazy WoS i Scopus, ze względu na rejestrowanie publikacji z zakresu nauk ścisłych, społecznych i humanistycznych, dobrze prezentują przekrój różnych dyscyplin, w obrębie których podejmowana jest tematyka architektury informacji.

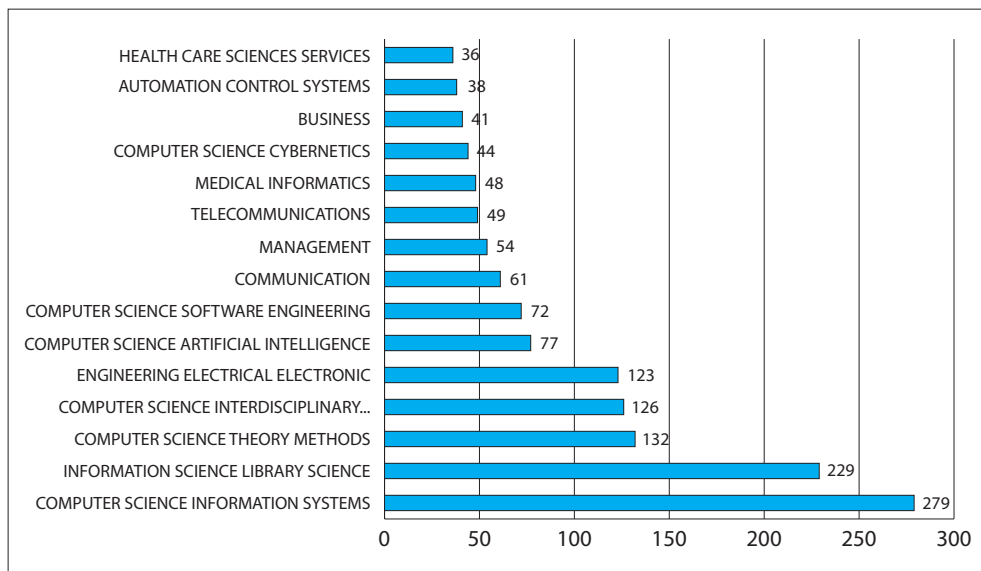
Wykres na rysunku 2 przedstawia zestawienie 15 kategorii dziedzinowych Web of Science (Web of Science Categories), do których należy największa liczba publikacji podejmujących tematykę AI. Najliczniej reprezentowana jest kategoria „Nauki komputerowe i systemy informacyjne”, w której znalazło się 30% wyszukanych tekstów; kolejne 25% piśmiennictwa o AI należy do kategorii „Nauka o informacji i nauka o bibliotece”. Dominują publikacje z zakresu nauk komputerowych (łącznie 765), czyli 84% całego zbioru wyodrębnionego z WoS przypisanych jest do różnych obszarów informatyki. Należy jednak nadmienić, że jeden dokument może być przypisany do więcej niż jednej kategorii dziedzinowej. Trzy pozostałe kategorie, do których przypisane są publikacje dotyczące problematyki AI, to obszary nauk społecznych: komunikacja (61 publikacji), zarządzanie (54) i biznes (41).

W bazie Scopus każdy dokument ma przypisany co najmniej jeden obszar tematyczny (ang. *subject area*). Także w tym źródle tematyka architektury informacji związana jest przede wszystkim z naukami ścisłymi. Do obszaru informatyki należą 1092 publikacje, tj. niemal 60% wyszukanych rekordów. Kolejne licznie reprezentowane kategorie to inżynieria, do której należy 529 dokumentów oraz nauki społeczne (442 dokumenty). Wśród najliczniej reprezentowanych dziedzin znalazły się jeszcze: matematyka (223 rekordy), biznes, zarządzanie i księgowość (153), nauki o decyzjach (119) oraz medycyna (113). Wymienione kategorie tematyczne są bardziej ogólne niż w bazie WoS, dlatego w Scopus dodatkowo analizie poddane zostały słowa kluczowe.

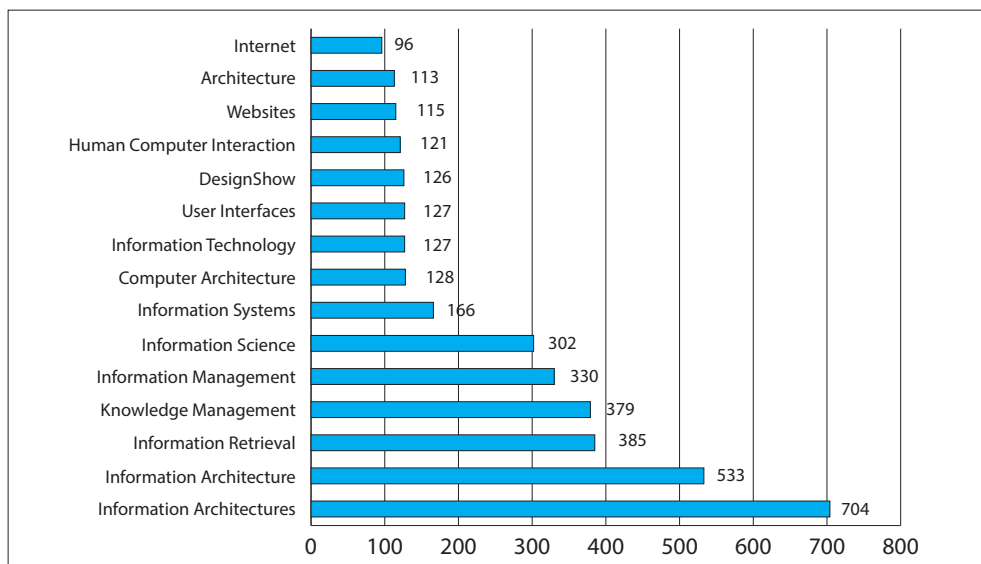
<sup>2</sup> Journal of Information Architecture <http://journalofia.org/>

<sup>3</sup> Z informacji pozyskanych od Jacka Tomaszczyka z bezpośredniej rozmowy z Andream Resmini, redaktorem naczelnym czasopisma, podczas konferencji *Architektura informacji jako dyscyplina akademicka* (AIDA – Kraków 5–6 czerwca 2018 r.) wiadomo, że czasopismo nie zostało zlikwidowane, ale redakcja nie posiada zasobów ludzkich, czasowych i finansowych, aby je regularnie wydawać, mimo że ma materiał do publikacji.





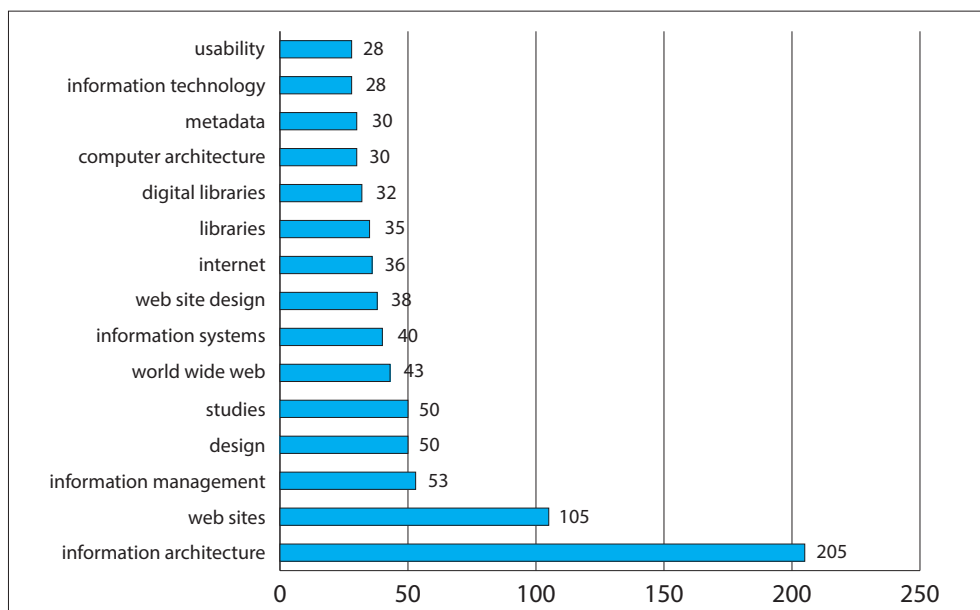
Rys. 2. Kategorie w bazie Web of Science, do których zakwalifikowane są publikacje z zakresu architektury informacji



Rys. 3. Słowa kluczowe występujące w bazie Scopus w opisach publikacji dotyczących architektury informacji

Rysunek 3 przedstawia wykres, w którym zestawiono 15 słów kluczowych najczęściej występujących w indeksowaniu piśmiennictwa o AI. Najwięcej wystąpień zarejestrowano dla dwóch wariantów gramatycznych frazy „information architecture” (704 wystąpienia w liczbie mnogiej i 533 wystąpienia w liczbie pojedynczej). Kolejne cztery słowa kluczowe, występujące więcej niż 300 razy, można uznać za reprezentujące terminologię naukę o informacji: przetwarzanie informacji, zarządzanie wiedzą, zarządzanie informacją, nauka o informacji. Pozostałe słowa kluczowe, charakteryzujące treść minimum 100 dokumentów, związane są przede wszystkim z technologiami informacyjnymi oraz różnymi aspektami projektowania stron internetowych.

Ze względu na wykazane już związki architektury informacji z nauką o informacji, sprawdzono, czy można dla nich ustalić kluczowe obszary badawcze. W tym celu prześledzono najpopularniejsze tematy w bazie LISA, a 15 najliczniej reprezentowanych wyszczególniono na wykresie (Rys. 4). Najczęściej występującym tematem jest architektura informacji – przypisano jej 36% publikacji. Kolejne 18% związane jest z tematyką stron internetowych, a 9% z zarządzaniem informacją. Wśród często poruszanych zagadnień znalazło się projektowanie, badania, World Wide Web, systemy informacyjne, czy projektowanie stron internetowych. Liczne są też teksty związane z bibliotekami tradycyjnymi i cyfrowymi, architekturą komputerów i technologią informacyjną, metadanymi oraz użytecznością. Wykaz tematów potwierdza, że w nauce o informacji tematyka AI podejmowana jest w kontekście praktycznym, przede wszystkim projektowania stron WWW i systemów informacyjnych.



Rys. 4. Tematy publikacji dotyczących architektury informacji, zarejestrowanych w bazie LISA

Kolejnym aspektem analizy były najczęściej cytowane publikacje. Wyniki ponownie przeszedzono oddzielnie dla trzech baz. W tabeli 4 wykazano 19 dokumentów z bazy LISA, które były cytowane co najmniej 10 razy. Większość z nich to artykuły opublikowane w latach 2001–2010. Dwa są z 2012 r., a najnowszy z 2016 r. (czwarte miejsce wśród najczęściej cytowanych). Ich tematyka jest bardzo zróżnicowana. Dwa teksty, na które powoływano się najczęściej, pochodzą z czasopisma z zakresu e-zdrowia. Ponadto wśród najczęściej cytowanych znalazły się artykuły dotyczące: zarządzania bezpieczeństwem informacji, narzędzi wspierających AI (metadane, tezaury, itp.), modelowania nawigacji, serwisów zawierających dane i informacje medyczne, czy badań użyteczności stron internetowych bibliotek akademickich.

W tabeli 5 wymienionych jest 20 najczęściej cytowanych dokumentów pochodzących z bazy WoS. W tym źródle odnotowana została znacznie większa liczba powołań. Aż 80 publikacji cytowanych było co najmniej 41 razy. Wynika to przede wszystkim z charakteru bazy, która z założenia jest indeksem cytowań, a nie bazą bibliograficzną. Spośród wyszukanych w WoS dokumentów, 441 pozycji, czyli 47% wszystkich publikacji cytowane jest co najmniej jeden raz. W sumie zostały one zacytowane 4352 razy. Wśród tekstów, na które powoływano się najczęściej, aż 16 zostało opublikowanych w latach 2000–2009.

Tab. 4. Najczęściej cytowane publikacje dotyczące architektury informacji wg danych bazy LISA

Lp.	Opis bibliograficzny dokumentu	Liczba cytowań
1	2	3
1	Danaher, B. G., McKay, G., Seeley, J. R. (2005). The information architecture of behavior change Websites. <i>Journal of Medical Internet Research</i> , 7(2).	59
2	McKay, H. G., Danaher, B. G., Seeley, J. R., Lichtenstein, E., Gau, J. M. (2008). Comparing Two Web-Based Smoking Cessation Programs: Randomized Controlled Trial. <i>Journal of Medical Internet Research</i> , 10(5).	58
3	Hanseth, O., Lyytinen, K. (2010). Design Theory for Dynamic Complexity in Information Infrastructures: The Case of Building Internet. <i>Journal of Information Technology</i> , 25(1), 1–19.	56
4	Soomro, Z. A., Shah, M. H., Ahmed, J. (2016). Information security management needs more holistic approach: A literature review. <i>International Journal of Information Management</i> , 36(2), 215. <a href="https://doi.org/10.1016/j.ijinfomgt.2015.11.009">https://doi.org/10.1016/j.ijinfomgt.2015.11.009</a>	33
5	Garshol, L. M. (2004). Metadata? Thesauri? Taxonomies? Topic maps! Making sense of it all. <i>Journal of Information Science</i> , 30(4), 378–391.	30
6	Miller, C. S., Remington, R. W. (2004). Modeling information navigation: implications for information architecture. <i>Human-Computer Interaction</i> , 19(3), 225–271.	25
7	Chen, C., Haddad, D., Selsky, J., Hoffman, J. E., Kravitz, R. L., Estrin, D. E., Sim, I. (2012). Making Sense of Mobile Health Data: An Open Architecture to Improve Individual – and Population-Level Health. <i>Journal of Medical Internet Research</i> , 14(4). <a href="https://doi.org/10.2196/jmir.2152">https://doi.org/10.2196/jmir.2152</a>	22
8	McGillis, L., Toms, E. G. (2001). Usability of the Academic Library Web Site: Implications for Design. <i>College and Research Libraries</i> , 62(4), 355–367.	18

1	2	3
9	Pak, R., Price, M. M., Thatcher, J. (2009). Age-Sensitive Design of Online Health Information: Comparative Usability Study. <i>Journal of Medical Internet Research</i> , 11(4), 11. <a href="https://doi.org/10.2196/jmir.1220">https://doi.org/10.2196/jmir.1220</a>	18
10	Watters, C., Duffy, J., Duffy, K. (2003). Using large tables on small display devices. <i>International Journal of Human-Computer Studies</i> , 58(1), 21–37.	15
11	Salvo, M. J. (2002). Critical engagement with technology in the computer classroom. <i>Technical Communication Quarterly</i> , 11(3), 317–337.	13
12	Fisher, L., Bennion, L. (2005). Organizational implications of the future development of technical communication: fostering communities of practice in the workplace. <i>Technical Communication</i> , 52(3), 277–288.	13
13	van den Haak, M. J., de Jong, M. D. T., Schellens, P. J. (2009). Evaluating Municipal Websites: A Methodological Comparison of Three Think-Aloud Variants. <i>Government Information Quarterly</i> , 26(1), 193–202.	13
14	Kees-Jan van Dorp. (2001). Tracking and tracing: A structure for development and contemporary practices. <i>Logistics Information Management</i> , 15(1/2), 24–33.	12
15	Pant, S., Ravichandran, T. (2001). A framework for information systems planning for e-business. <i>Logistics Information Management</i> , 14(1/2), 85–98.	11
16	Iannella, R. (2001). Digital Rights Management (DRM) architectures. <i>D-Lib Magazine</i> , 7(6).	11
17	Klischewski, R., Scholl, H. J. (2008). Information Quality, as Capstone in Negotiating e-Government Integration, Interoperation and Information Sharing. <i>Electronic Government: An International Journal</i> , 5(2), 203–225.	11
18	Savolainen, R., Talja, S., Tuominen, K. (2003). Multiperspective digital libraries: the implications of constructionism for the development of digital libraries. <i>Journal of the American Society for Information Science and Technology</i> , 54(6), 561–569.	10
19	Yang, T., Linder, J., Bolchini, D. (2012). DEEP: Design-Oriented Evaluation of Perceived Usability. <i>International Journal of Human-Computer Interaction</i> , 28(5), 308–346. <a href="https://doi.org/10.1080/10447318.2011.586320">https://doi.org/10.1080/10447318.2011.586320</a>	10

Największym zainteresowaniem cieszył się artykuł prezentujący narzędzie do pomiaru elastyczności infrastruktury IT. Kolejny tekst przedstawia rozwiązania z zakresu AI dla zarządzania jakością łańcucha dostaw. A trzecią w wykazie jest jedna ze starszych publikacji, prezentująca architekturę systemów informacyjnych w organizacjach. Powyżej 100 cytowań ma jeszcze artykuł dotyczący analizy logów wyszukiwania i zastosowania wyników tej analizy do projektowania lepszej architektury informacji, interfejsów i systemów informacyjnych.

Tematyka czasopism, z których pochodzą najczęściej cytowane artykuły jest zróżnicowana. Nauki komputerowe reprezentuje 16 tytułów, a nauki społeczne pięć (jedno z czasopism obie dziedziny). Najwięcej artykułów (5) opublikowano w czasopismach z zakresu nauki o informacji, a po cztery w periodykach podejmujących problematykę systemów informacyjnych (informatyka) i informatyki medycznej. W tabeli 5 znalazły się także artykuły, których tematyka została zakwalifikowana do fizjologii, fizyki i biotechnologii, co

wiąże się z dziedzinową specyfiką systemów i serwisów stanowiących przedmiot analiz. Potwierdza to, że problematyka projektowania przestrzeni informacyjnych jest obecna w różnych dyscyplinach i analizowana w różnych ujęciach. Ogólny trend wskazuje przewagę cytowań tekstów z zakresu nauk komputerowych (informatyki), ale wyraźnie widać obecność nauki o informacji, mimo iż jest to dyscyplina dość wąska, która należąca do nauk humanistycznych i społecznych, jest słabiej reprezentowana w WoS.

Tab. 5. Najczęściej cytowane publikacje o architekturze informacji wg danych Web of Science

Lp.	Opis bibliograficzny dokumentu	Liczba cytowań
1	2	3
1	Byrd, T. A., Turner, D. E. (2000). Measuring the flexibility of information technology infrastructure: Exploratory analysis of a construct. <i>Journal of Management Information Systems</i> , 17(1), 167–208.	231
2	Xu, L. D. (2011). Information architecture for supply chain quality management. <i>International Journal of Production Research</i> , 49(1), 183–198. <a href="https://doi.org/10.1080/00207543.2010.508944">https://doi.org/10.1080/00207543.2010.508944</a>	152
3	Allen, B. R., Boynton, A. C. (1991). Information architecture – in search of efficient flexibility. <i>MIS quarterly : management information systems.</i> , 15(4), 435–445. <a href="https://doi.org/10.2307/249447">https://doi.org/10.2307/249447</a>	109
4	Jansen, B. J. (2006). Search log analysis: What it is, what's been done, how to do it. <i>Library &amp; information science research</i> , 28(3), 407–432. <a href="https://doi.org/10.1016/j.lisr.2006.06.005">https://doi.org/10.1016/j.lisr.2006.06.005</a>	101
5	Marceglia, S., Foffani, G., Bianchi, A. M., Baselli, G., Tamma, F., Egidi, M., Priori, A. (2006). Dopamine-dependent non-linear correlation between subthalamic rhythms in Parkinson's disease. <i>Journal of Physiology-London</i> , 571(3), 579–591. <a href="https://doi.org/10.1113/jphysiol.2005.100271">https://doi.org/10.1113/jphysiol.2005.100271</a>	97
6	Garshol, L. M. (2004). Metadata? Thesauri? Taxonomies? Topic maps! Making sense of it all. <i>Journal of Information Science</i> , 30(4), 378–391.	91
7	Hucul, D., Inlek, I. V., Vittorini, G., Crocker, C., Debnath, S., Clark, S. M., Monroe, C. (2015). Modular entanglement of atomic qubits using photons and phonons. <i>Nature physics</i> , 11(1), 37–42. <a href="https://doi.org/10.1038/nphys3150">https://doi.org/10.1038/nphys3150</a>	90
8	Byrd, T. A., Turner, D. E. (2001). An exploratory examination of the relationship between flexible IT infrastructure and competitive advantage. <i>Information &amp; Management</i> , 39(1), 41–52. <a href="https://doi.org/10.1016/s0378-7206(01)00078-7">https://doi.org/10.1016/s0378-7206(01)00078-7</a>	89
9	Summers, T. H., Akella, M. R., Mears, M. J. (2009). Coordinated Standoff Tracking of Moving Targets: Control Laws and Information Architectures. <i>Journal of guidance, control, and dynamics: a publication of the American Institute of Aeronautics and Astronautics devoted to the technology of dynamics and control</i> , 32(1), 56–69. <a href="https://doi.org/10.2514/1.37212">https://doi.org/10.2514/1.37212</a>	84
10	McKay, H. G., Danaher, B. G., Seeley, J. R., Lichtenstein, E., Gau, J. M. (2008). Comparing Two Web-Based Smoking Cessation Programs: Randomized Controlled Trial. <i>Journal of Medical Internet Research</i> , 10(5).	73

1	2	3
11	Stead, W. W., Miller, R. A., Musen, M. A., Hersh, W. R. (2000). Integration and beyond: Linking information from disparate sources and into workflow. <i>Journal of the American Medical Informatics Association: JAMIA</i> , 7(2), 135–145. <a href="https://doi.org/10.1136/jamia.2000.0070135">https://doi.org/10.1136/jamia.2000.0070135</a>	70
12	Large, A., Beheshti, J., Rahman, T. (2002). Design criteria for children's web portals: the users speak out. <i>Journal of the American Society for Information Science and Technology</i> , 53(2), 79–94.	68
13	Danaher, B. G., McKay, G., Seeley, J. R. (2005). The information architecture of behavior change Websites. <i>Journal of Medical Internet Research</i> , 7(2).	68
14	Hoogervorst, J. (2004). Enterprise architecture: Enabling integration, agility and change. <i>International Journal of Cooperative Information Systems</i> , 13(3), 213–233. <a href="https://doi.org/10.1142/s021884300400095x">https://doi.org/10.1142/s021884300400095x</a>	60
15	McGillis, L., Toms, E. G. (2001). Usability of the Academic Library Web Site: Implications for Design. <i>College and Research Libraries</i> , 62(4), 355–367.	57
16	Richardson, G. L., Jackson, B. M., Dickson, G. W. (1990). A principles-based enterprise architecture – lessons from texaco and star enterprise. <i>MIS quarterly: management information systems.</i> , 14(4), 385–403. <a href="https://doi.org/10.2307/249787">https://doi.org/10.2307/249787</a>	56
17	Yasnoff, W. A., Overhage, J. M., Humphreys, B. L., LaVenture, M. (2001). A national agenda for public health informatics: Summarized recommendations from the 2001 AMIA Spring Congress. <i>Journal of the American Medical Informatics Association: JAMIA</i> , 8(6), 535–545. <a href="https://doi.org/10.1136/jamia.2001.0080535">https://doi.org/10.1136/jamia.2001.0080535</a>	53
18	Seringhaus, M. R., Gerstein, M. B. (2007). Publishing perishing? Towards tomorrow's information architecture. <i>BMC bioinformatics</i> , 8, 5. <a href="https://doi.org/10.1186/1471-2105-8-17">https://doi.org/10.1186/1471-2105-8-17</a>	45
19	Shen, W. M., Lang, S. Y. T., Wang, L. H. (2005). iShopfloor: An Internet-enabled agent-based intelligent shop floor. <i>Ieee Transactions on Systems Man and Cybernetics Part C-Applications and Reviews</i> , 35(3), 371–381. <a href="https://doi.org/10.1109/tsmcc.2004.843224">https://doi.org/10.1109/tsmcc.2004.843224</a>	44
20	Miller, C. S., Remington, R. W. (2004). Modeling information navigation: implications for information architecture. <i>Human-Computer Interaction</i> , 19(3), 225–271.	41

Aby dać pełniejszy obraz cytowań dokumentów dotyczących architektury informacji, prześledzono również dane z bazy Scopus. Tutaj odnotowano ich jeszcze więcej niż w opisanych już bazach. Publikacje w niej wyszukane, począwszy od 1988 r., zacytowane zostały 12 564 razy. Spośród 1836 dokumentów 58%, czyli 1076 publikacji zacytowano co najmniej raz. W tabeli 6 znajduje się wykaz 20 najczęściej cytowanych publikacji, wśród których tylko jedna ma mniej niż 100 cytowań. Wśród tekstów, na które powoływano się najczęściej, 13 zostało opublikowanych w latach 2000–2009. Najwięcej cytowań ma artykuł, który ukazał się w 1991 r. w czasopiśmie *Harvard Business Review*. Dotyczy on możliwości wykorzystania nowych technologii do pomiaru niefinansowych działań firmy, w tym opracowania architektury informacji serwisu realizującego taki pomiar. Najczęściej cytowany artykuł dotyczący AI i zarejestrowany w bazie WoS w zestawieniu danych z bazy

Scopus znalazł się na miejscu drugim, ale tutaj odnotowano niemal dwukrotnie więcej jego powołań. Kolejny tekst, z liczbą 317 powołań, opublikowany został w czasopiśmie *Science*. Traktuje on o nowych rozwiązaniach w architekturze informacji systemów zorientowanych na usługi, umożliwiających projektowanie interfejsów i protokołów do publikowania i dostępu do cennych danych i programów w nauce nie tylko specjalistom. Najczęściej cytowane są artykuły z czasopism naukowych, ale w wykazie znalazły się też cztery teksty opublikowane w materiałach konferencyjnych oraz jedna książka.

Tab. 6. Najczęściej cytowane dokumenty o architekturze informacji wg danych bazy Scopus

Lp.	Opis bibliograficzny dokumentu	Liczba cytowań
1	2	3
1	Eccles, R. G. (1991). The performance measurement manifesto. <i>Harvard Business Review</i> , 69(1), 131–137.	653
2	Byrd, T. A., Turner, D. E. (2000). Measuring the flexibility of information technology infrastructure: Exploratory analysis of a construct. <i>Journal of Management Information Systems</i> , 17(1), 167–208.	405
3	Foster, I. (2005). Service-oriented science. <i>Science</i> , 308(5723), 814–817. <a href="https://doi.org/10.1126/science.1110411">https://doi.org/10.1126/science.1110411</a>	317
4	Geerts, G. L., McCarthy, W. E. (2002). An ontological analysis of the economic primitives of the extended-REA enterprise information architecture. <i>International Journal of Accounting Information Systems</i> , 3(1), 1–16. <a href="https://doi.org/10.1016/S1467-0895(01)00020-3">https://doi.org/10.1016/S1467-0895(01)00020-3</a>	175
5	Xu, L. D. (2011). Information architecture for supply chain quality management. <i>International Journal of Production Research</i> , 49(1), 183–198. <a href="https://doi.org/10.1080/00207543.2010.508944">https://doi.org/10.1080/00207543.2010.508944</a>	174
6	Jarvenpaa, S. L., Ives, B. (1993). The global network organization of the future: Information management opportunities and challenges. <i>Journal of Management Information Systems</i> , 10(3), 25–57.	173
7	Allen, B. R., Boynton, A. C. (1991). Information architecture – in search of efficient flexibility. <i>MIS quarterly : management information systems.</i> , 15(4), 435–445. <a href="https://doi.org/10.2307/249447">https://doi.org/10.2307/249447</a>	172
8	Jansen, B. J. (2006). Search log analysis: What it is, what's been done, how to do it. <i>Library &amp; information science research</i> , 28(3), 407–432. <a href="https://doi.org/10.1016/j.lisr.2006.06.005">https://doi.org/10.1016/j.lisr.2006.06.005</a>	160
9	Qin, Z., Denker, G., Giannelli, C., Bellavista, P., Venkatasubramanian, N. (2014). A software defined networking architecture for the internet-of-things. <i>W IEEE/IFIP NOMS 2014 – IEEE/IFIP Network Operations and Management Symposium: Management in a Software Defined World</i> . IEEE Computer Society. <a href="https://doi.org/10.1109/NOMS.2014.6838365">https://doi.org/10.1109/NOMS.2014.6838365</a>	157
10	Byrd, T. A., Turner, D. E. (2001). An exploratory examination of the relationship between flexible IT infrastructure and competitive advantage. <i>Information &amp; Management</i> , 39(1), 41–52. <a href="https://doi.org/10.1016/s0378-7206(01)00078-7">https://doi.org/10.1016/s0378-7206(01)00078-7</a>	150

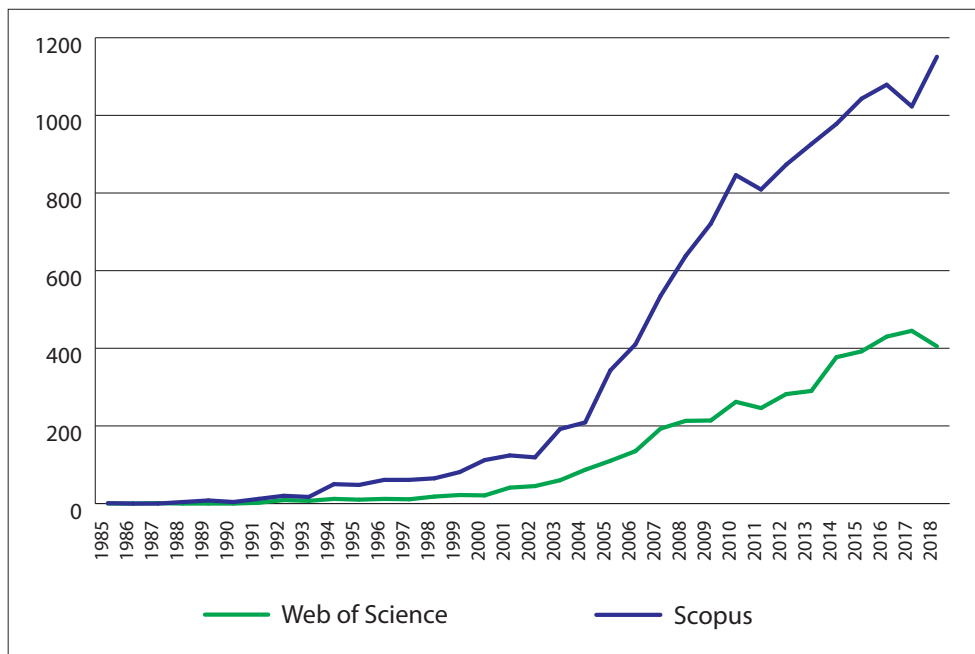
1	2	3
11	Newman, M. W., Landay, J. A. (2000). Sitemaps, storyboards, and specifications: A sketch of web site design practice. <i>Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (DIS 2000)</i> , 263–274.	128
12	Summers, T. H., Akella, M. R., Mears, M. J. (2009). Coordinated standoff tracking of moving targets: Control laws and information architectures. <i>Journal of Guidance, Control, and Dynamics</i> , 32(1), 56–69. <a href="https://doi.org/10.2514/1.37212">https://doi.org/10.2514/1.37212</a>	126
13	Weber, R. H., Weber, R. (2010). <i>Internet of things: Legal perspectives</i> . Springer Berlin Heidelberg. <a href="https://doi.org/10.1007/978-3-642-11710-7">https://doi.org/10.1007/978-3-642-11710-7</a>	123
14	Klemmer, S. R., Newman, M. W., Farrell, R., Bilezikjian, M., Landay, J. A. (2001). The Designers' Outpost: A tangible interface for collaborative web site design. W <i>UIST (User Interface Software and Technology): Proceedings of the ACM Symposium</i> (1–10).	123
15	Arsanjani, A., Zhang, L.-J., Ellis, M., Allam, A., Channabasavaiah, K. (2007). S3: A service-oriented reference architecture. <i>IT Professional</i> , 9(3), 10–17. <a href="https://doi.org/10.1109/MITP.2007.53">https://doi.org/10.1109/MITP.2007.53</a>	117
16	Richardson, G. L., Jackson, B. M., Dickson, G. W. (1990). A principles-based enterprise architecture – lessons from texaco and star enterprise. <i>MIS quarterly: management information systems.</i> , 14(4), 385–403. <a href="https://doi.org/10.2307/249787">https://doi.org/10.2307/249787</a>	114
17	Marceglia, S., Foffani, G., Bianchi, A. M., Baselli, G., Tamma, F., Egidi, M., Priori, A. (2006). Dopamine-dependent non-linear correlation between subthalamic rhythms in Parkinson's disease. <i>Journal of Physiology-London</i> , 571(3), 579–591. <a href="https://doi.org/10.1113/jphysiol.2005.100271">https://doi.org/10.1113/jphysiol.2005.100271</a>	101
18	Garshol, L. M. (2004). Metadata? Thesauri? Taxonomies? Topic maps! Making sense of it all. <i>Journal of Information Science</i> , 30(4), 378–391.	101
19	Large, A., Beheshti, J., Rahman, T. (2002). Design criteria for children's web portals: the users speak out. <i>Journal of the American Society for Information Science and Technology</i> , 53(2), 79–94.	100
20	Kim, Y., Schmid, T., Charbiwala, Z. M., Friedman, J., Srivastava, M. B. (2008). NAWMS: Nonintrusive autonomous water monitoring system. W <i>SensSys'08 – Proceedings of the 6th ACM Conference on Embedded Networked Sensor Systems</i> (s. 309–321). <a href="https://doi.org/10.1145/1460412.1460443">https://doi.org/10.1145/1460412.1460443</a>	97

Po porównaniu danych umieszczonych w tabelach 4, 5 i 6, znaleziono dziewięć publikacji powtarzających się wśród najczęściej cytowanych w Scopus i Web of Science. Jeden z tych artykułów, publikacja Larsa M. Garshola (2004), wystąpił się we wszystkich trzech tabelach, z liczbą 101 cytowań w Scopus, 91 w WoS i 30 w LISA. Prezentuje on tradycyjne narzędzia nauki o informacji, jak metadane, tezaury, taksonomie i mapy tematyczne, które jako niewidoczne elementy wspierają architekturę informacji.

Wykres na rysunku 5 prezentuje liczbę cytowań piśmiennictwa o AI wg danych baz Scopus i Web of Science – od pierwszego, odnotowanego w 1985 r., przez kolejne dwa cytowania, które wystąpiły w 1987 r., do 2018 r. Według danych z bazy Scopus, liczba



powołań stopniowo zwiększała się od 1988 r., a znaczny ich wzrost, powyżej 100 cytowań, można zaobserwować od 2000 r. Najwięcej cytowań (1151) odnotowano w tej bazie dla 2018 r. W bazie WoS, pierwsze cytowanie pojawiło się w 1987 r., w latach 1988–1990 nie było powołań na teksty z badanego tematu, ale później następował systematyczny wzrost aż do 445 w 2017 r. Niewielki spadek cytowań w tej bazie, zauważalny za 2018 r., może być uzasadniony tym, że zestaw danych za ostatni rok mógł być niekompletny w momencie pisania tego artykułu. Wykres pokazuje stale rosnące zainteresowanie publikacjami z zakresu architektury informacji, odnotowane dla obu źródeł.

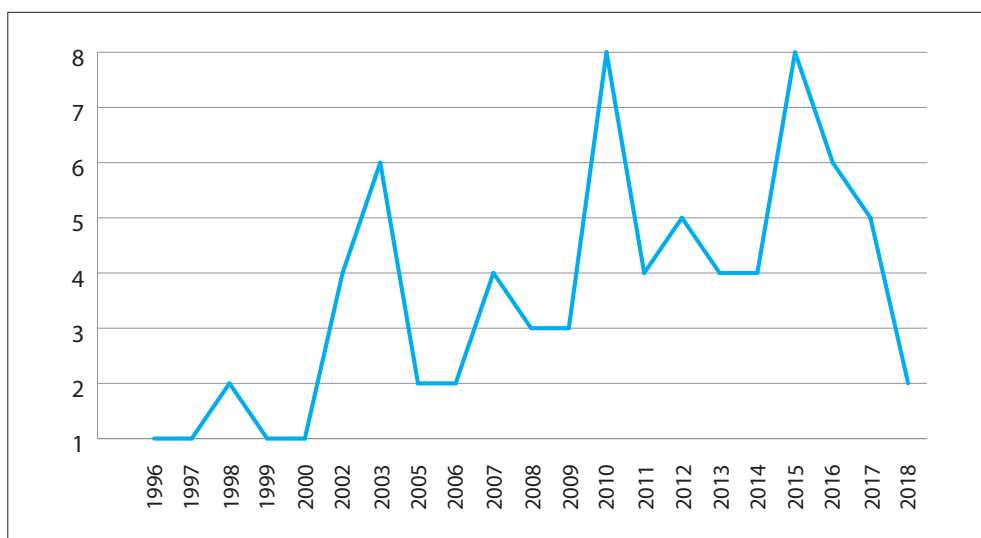


Rys. 5. Rozkład chronologiczny cytowań publikacji dotyczących architektury informacji w bazach Scopus i Web of Science

### 3.2. Książki z zakresu architektury informacji

Gromadząc dane do tej części artykułu, poza książkami znalezionymi w bazie Scopus, skupiono się na poszukiwaniu wydawnictw zwartych, w których tytule, opisie lub w spisie treści występował termin „architektura informacji” (ang. *information architecture*). Pomińnięte zostały książki z zakresu User Experience i User Intereface Design, gdyż tematyka ta nie była obecna w pierwszej części badania. Wykaz 76 wyselekcjonowanych publikacji znajduje się w załączniku 1. Zestawiając je z liczbą wszystkich publikacji, stanowią one 3% zgromadzonej przez autorkę literatury. Pierwsza książka w wykazie (Cook & Books, 1996) została wydana w 1996 r., rok później R.S. Wurman zdefiniował zadania architektów informacji (Wurman, 1997), a w 1998 r. wydana została jedna z najbardziej znanych książek dotyczących architektury informacji (Morville & Rosenfeld, 1998).

Rysunek 6 prezentuje chronologiczny rozkład publikacji książek dotyczących architektury informacji. Poza latami 2001 i 2004 ukazywała się co najmniej jedna książka rocznie, a najwięcej, po osiem tytułów, wydano w 2010 i 2015 r. Kilka publikacji doczekało się kolejnych wydań. Najwięcej, bo aż cztery wydania ma *Information Architecture for the World Wide Web: Designing Large-Scale Web Sites* (Morville & Rosenfeld, 1998, 2002, 2006), ale ostatnie wydanie, najbardziej zmienione, przygotowane zostało z trzecim współautorem i ukazało się pod zmodyfikowanym tytułem *Information Architecture: For the Web and Beyond* (Rosenfeld et al., 2015). Trzy inne książki miały po dwa wydania. Kilku autorów opublikowało więcej niż jedną książkę. Uwzględniając wszystkie wydania, zauważyć można, że najbardziej aktywnymi autorami publikacji książkowych są: Peter Morville (sześć książek), Louis Rosenfeld (pięć książek) oraz Gerard Blokdyk (cztery książki). Po dwie książki wydali: Jorge Arango, Dan Brown, Wei Ding, Xia Lin, Andrea Resmini, Donna Spencer i Christina Wodtke.



Rys. 6. Chronologiczny rozkład publikacji książek na temat architektury informacji

Ostatnią z interesujących kwestii dotyczących wydawnictw zwartych była próba oceny ich popularności mierzona liczbą cytowań. W tym celu każda książka została sprawdzona w Google Scholar (GS), gdyż było to jedyne źródło indeksujące większość zgromadzonych tytułów. Cytowania dla kilku wydań tego samego tytułu zostały zsumowane, gdyż wybrane narzędzie nie uwzględniało oddzielnych danych dla różnych wydań. Spośród 76 książek, dziewięć nie zostało zaindeksowanych przez GS, a pięć tytułów nie miało cytowań. Pozostałe publikacje miały minimum po dwa powołania. Łącznie książki zacytowano 9078 razy, co stanowi tylko o 1/4 mniej niż liczba wszystkich cytowań piśmiennictwa o AI zarejestrowanych w bazie Scopus i dwa razy więcej niż suma cytowań wykazywanych dla tego piśmiennictwa w Web of Science. Najwięcej, 3128 cytowań, ma książka *Information Architecture for the World Wide Web: Designing Large-Scale Web Sites* (Morville & Rosenfeld, 1998 i kolejne dwa wydania), na czwarte wydanie od 2015 r. powoływano się 147 razy. Pierwsza dziesiątka najczęściej cytowanych książek przedstawiona jest w tabeli 7

(w opisach uwzględniono pierwsze wydanie). Jak widać, dwie najczęściej cytowane prace są autorstwa lub współautorstwa P. Morville'a. Z liczbą przekraczającą 500 cytowań jest książka dotycząca badań użyteczności – technik gromadzenia informacji o wymaganiach użytkowników w projektowaniu produktów cyfrowych. W tabeli znajduje się wspomniana już wcześniej książka R. S. Wurmana o architektach informacji. W wykazie są również publikacje dotyczące architektury systemów informacyjnych przedsiębiorstw oraz książki o charakterze bardziej praktycznym, przedstawiające metodę sortowania kart, projektowania nawigacji czy tworzenia użytecznych treści dla użytkowników Internetu, a jeden z tytułów poświęcony jest perspektywom prawnym Internetu Rzeczy.

Tab. 7. Najczęściej cytowane książki

Lp.	Opis książki	Liczba cytowań
1	Morville, P., Rosenfeld, L. (1998). <i>Information Architecture for the World Wide Web: Designing Large-scale Web Sites</i> (1 edition). Cambridge ; Sebastopol, CA: O'Reilly Media.	3078
2	Morville, P. (2005). <i>Ambient Findability: What We Find Changes Who We Become</i> (1 edition). Beijing ; Sebastopol, CA: O'Reilly Media.	833
3	Courage, C., Baxter, K. (2005). <i>Understanding Your Users. A Practical Guide to User Requirements: Methods, Tools, and Techniques</i> . Amsterdam: Elsevier. <a href="https://doi.org/10.1016/B978-1-55860-935-8.X5029-5">https://doi.org/10.1016/B978-1-55860-935-8.X5029-5</a>	525
4	Wurman, R. S. (1997). <i>Information Architects</i> (First Edition). Zürich: Graphis Inc.	487
5	Cook, M., Books, H.-P. P. (1996). <i>Building Enterprise Information Architectures: Reengineering Information Systems</i> (1 edition). Upper Saddle River, NJ: Prentice Hall.	309
6	Minoli, D. (2008). <i>Enterprise Architecture A to Z: Frameworks, Business Process Modeling, SOA, and Infrastructure Technology</i> . Boca Raton: Auerbach Publications. <a href="https://doi.org/10.1201/9781420013702">https://doi.org/10.1201/9781420013702</a>	285
7	Weber, R. H., Weber, R. (2010). <i>Internet of things: Legal perspectives</i> . Springer Berlin Heidelberg. <a href="https://doi.org/10.1007/978-3-642-11710-7">https://doi.org/10.1007/978-3-642-11710-7</a>	284
8	Spencer, D. (2009). <i>Card Sorting: Designing Usable Categories</i> (1 edition). Rosenfeld Media.	264
9	Kalbach, J. (2007). <i>Designing Web navigation</i> (1st ed). Beijing ; Sebastopol: O'Reilly.	249
10	Redish, J. G. (2007). <i>Letting Go of the Words</i> . Elsevier Inc. <a href="https://doi.org/10.1016/B978-0-12-369486-7.X5000-9">https://doi.org/10.1016/B978-0-12-369486-7.X5000-9</a>	204

#### 4. Zakończenie

Architektura informacji jest młodą dyscypliną. Pierwsze publikacje dotyczące tego tematu pochodzą z 1982 r., a w tytule termin „information architecture” po raz pierwszy użyty został w 1983 r. Od 2000 r. obserwuje się rosnące zainteresowanie problematyką AI – znacząco wzrosła liczba poświęconych jej publikacji, artykułów w czasopiśmie i tekstów

w materiałach konferencyjnych. Także liczba cytowań publikacji z tego zakresu z roku na rok jest większa. Mimo iż w świetle zgromadzonych danych książki stanowią zaledwie 3% piśmiennictwa o AI, to liczba ich cytowań wskazuje na ich intensywne wykorzystywanie przez kolejnych autorów. Zainteresowanie jest globalne, nie można ustalić kluczowych ośrodków prowadzących badania w zakresie architektury informacji, autorzy są z całego świata, a pod względem liczby publikacji wyróżniają się przede wszystkim: Davide Bolchini, Marijn Janssen i Martin White. Biorąc pod uwagę liczbę książek, jako najbardziej produktywnych autorów należy wskazać: Petera Morville'a, Gerarda Blokdyka i Luisa Rosenfelda. W badaniu ustalono zaledwie osiem periodyków, w których opublikowano więcej niż 10 artykułów dotyczących AI. Poza kilkoma numerami *Journal of Information Architecture* (ukazującego się regularnie w latach 2009–2013), nie ma wiodącego czasopisma poświęconego temu tematowi. Dominują czasopisma z nauki o informacji, systemów informacyjnych (informatyka) i komunikacji, co potwierdza szerokie, interdyscyplinarne zainteresowanie problematyką projektowania przestrzeni informacyjnych. Z kolei materiały konferencyjne prezentują przede wszystkim aspekty informatyczne AI. Wśród najczęściej cytowanych artykułów przeważają publikacje z czasopism informatycznych, ale licznie reprezentowane są też czasopisma z nauki o informacji. Analiza literatury zgromadzonej w omówionych badaniach nie wykazała związków AI z architekturą, projektowaniem wizualnym czy psychologią. Zebrane piśmiennictwo reprezentuje przede wszystkim architekturę informacji jako przedmiot informatyki (badań nad systemami informacyjnymi), inżynierii oraz nauki o informacji.

Podjęto próbę uchwycenia kierunków i trendów badań w zakresie architektury informacji. Przedstawiona analiza nie jest kompletna, bo uwzględniła tylko piśmiennictwo zarejestrowane w wykorzystanych bazach bibliograficznych i bazach cytowań oraz serwisach internetowych udostępniających informacje o publikacjach monograficznych. Wykorzystane źródła można jednak uznać za najbardziej kompletne rejestry piśmiennictwa badanej dziedziny, jakkolwiek nie uwzględniają one wszystkich publikacji ukazujących się na świecie. Dalsze poszukiwania literatury, która nie została zarejestrowana w tych źródłach wymagałoby dużego nakładu czasu, a jego rezultaty z dużym prawdopodobieństwem nie wpłynęłyby znacząco na końcowe wnioski przedstawionego badania. Ponadto warto byłoby przeprowadzić bardziej pogłębioną analizę tematyczną zgromadzonej literatury, na co jednak zabrakło miejsca w tym artykule.

## Załącznik 1. Wykaz książek zgromadzonych do analizy

- Arango, J. (2018). *Living in Information: Responsible Design for Digital Places*. Brooklyn, NY: Two Waves Books.
- Bakalov, F. (2010). *Designing Information Architecture of Websites*. Lambert Academic Publishing.
- Baofu, P. (2008). *The future of information architecture: conceiving a better way to understand taxonomy, network, and intelligence*. Oxford, UK: Chandos Publishing.
- Batley, S. (2007). *Information Architecture for Information Professionals*. Oxford: Chandos Publishing.
- Blokdyk, G. (2015). *Information architecture – simple steps to win, insights and opportunities for maxing out success*. COMPLETE Publishing.
- Blokdyk, G. (2017a). *Enterprise Information Architecture EIA: End-To-End Data Analysis*. CreateSpace Independent Publishing Platform.

- Blokdyk, G. (2017b). *Information Architecture Complete Self-Assessment Guide*. CreateSpace Independent Publishing Platform.
- Blokdyk, G. (2018). *Enterprise Information Architecture: Third Edition*. CreateSpace Independent Publishing Platform.
- Boiko, B. (2017). *Information Architecture from the Info Out: Book two in the infoOut series*. Independently published.
- Boucadair, M., Binet, D. (2013). *Solutions for Sustaining Scalability in Internet Growth*. IGI Global. <https://doi.org/10.4018/978-1-4666-4305-5>
- Brandon, D. M. (2008). *Software engineering for modern web applications: Methodologies and technologies*. IGI Global. <https://doi.org/10.4018/978-1-59904-492-7>
- Brown, D. M. (2006). *Communicating Design: Developing Web Site Documentation for Design and Planning*. Berkeley, CA: New Riders.
- Brown, D. M. (2010). *Communicating Design: Developing Web Site Documentation for Design and Planning*. Berkeley, CA: New Riders.
- Brown, P. (2003). *Information Architecture with XML: A Management Strategy*. Chichester, West Sussex, England ; Hoboken, NJ, USA: Wiley.
- Caprio, L., Ghiglione, B. (2003). *Information architecture*. Milano: Tecniche nuove.
- Cook, M., Books, H.-P. P. (1996). *Building Enterprise Information Architectures: Reengineering Information Systems*. Upper Saddle River, NJ: Prentice Hall.
- Courage, C., Baxter, K. (2005). *Understanding Your Users*. Elsevier Inc. <https://doi.org/10.1016/B978-1-55860-935-8.X5029-5>
- Covert, A. (2014). *How to Make Sense of Any Mess: Information Architecture for Everybody*. CreateSpace Independent Publishing Platform.
- Dade-Robertson, M. (2011). *The Architecture of Information: Architecture, Interaction Design and the Patterning of Digital Information*. Abingdon, Oxon; New York: Routledge.
- Davis, N. (2015). *Information Architecture: Theory, Science and Practice*. Elsevier Science & Technology.
- Dijck, P. V. (2003). *Information Architecture for Designers: Structuring Websites for Business Success*. Mies, Switzerland Hove: Rotovision.
- Ding, W., Lin, X. (2010). *Information Architecture: The Design and Integration of Information Spaces*. Morgan & Claypool Publishers.
- Ding, W., Lin, X., Zarro, M. (2017). *Information Architecture: The Design and Integration of Information Spaces*. San Rafael: Morgan & Claypool Publishers.
- Evernden, R., Evernden, E. (2003). *Information First: Integrating Knowledge and Information Architecture for Business Advantage*. London; New York: Butterworth-Heinemann.
- Gilchrist, A., Mahon, B., eds. (2003). *Information Architecture: Designing Information Environments for Purpose*. New York: Facet Pub.
- Glushko, R. J., ed. (2013). *The Discipline of Organizing*. Cambridge, MA: The MIT Press.
- Godinez, M., Hechler, E., Koenig, K., Lockwood, S., Oberhofer, M., Schroeck, M. (2010). *The Art of Enterprise Information Architecture: A Systems-Based Approach for Unlocking Business Insight*. Upper Saddle River, NJ: IBM Press.
- Goldschmidt, H. M. J., Cox, M. J. T., Grouls, R. J. E., Van De Laar, W. A. J. H., Van Merode, F. G. G. (1998). *Reference information model for clinical laboratories: RILA as laboratory management toolbox* (T. 55). IOS Press. <https://doi.org/10.3233/978-1-60750-899-1-i>
- Gotz, R. (2012). *Practical SharePoint 2010 Information Architecture*. New York, NY: Apress.
- Governor, J., Hinchcliffe, D., Nickull, D. (2009). *Web 2.0 Architectures: What entrepreneurs and information architects need to know*. Sebastopol, CA: O'Reilly.
- Greenberg, J., Méndez, E. (2013). *Knitting the Semantic Web*. Taylor and Francis. <https://doi.org/10.4324/9780203052051>

- Hausman, K. K., Cook, S. L. (2010). *IT Architecture For Dummies*. Hoboken, NJ: For Dummies.
- Hedden, H. (2016). *The Accidental Taxonomist*.
- Hinton, A. (2014). *Understanding Context: Environment, Language, and Information Architecture*. Sebastopol, CA: O'Reilly Media.
- Information Resources Management Association. (2015). *Web design and development: Concepts, methodologies, tools, and applications*. IGI Global. <https://doi.org/10.4018/978-1-4666-8619-9>
- Jamil, George Leal. (2015). *Handbook of Research on Information Architecture and Management in Modern Organizations*. (J. Poças-Rascão, F. Ribeiro, A. M. da Silva, eds.). Hershey, PA: IGI Global.
- Jamil, G.L., Poças-Rascão, J., Ribeiro, F., & da Silva, A. M. (2015). *Handbook of research on information architecture and management in modern organizations*. IGI Global. <https://doi.org/10.4018/978-1-4666-8637-3>
- Kalbach, J. (2007). *Designing Web navigation*. Beijing; Sebastopol: O'Reilly.
- Kauhanen-Simanainen, A. (2007). *Corporate Literacy: Discovering the Senses of the Organisation*. Elsevier Ltd. <https://doi.org/10.1016/C2013-0-16637-8>
- Mahon, B., Hourican, R., & Gilchrist, A. (2002). *Research into information architecture: the roles of software, taxonomies and people : a report*. London: Tfpl Ltd.
- McKee, J. (2012). *Applying principles from IT architecture to strategic business planning*. IGI Global. <https://doi.org/10.4018/978-1-4666-2527-3>
- Minoli, D. (2008). *Enterprise architecture A to Z: Frameworks, business process modeling, SOA, and infrastructure technology*. CRC Press. <https://doi.org/10.1201/9781420013702>
- Morrogh, E. (2002). *Information Architecture: An Emerging 21st Century Profession*. Upper Saddle River, NJ: Pearson.
- Morville, P. (2005). *Ambient Findability: What We Find Changes Who We Become*. Beijing; Sebastopol, CA: O'Reilly Media.
- Morville, P., Callender, J. (2010). *Search Patterns: Design for Discovery*. Sebastopol, CA: O'Reilly Media.
- Morville, P., Rosenfeld, L. (1998). *Information Architecture for the World Wide Web: Designing Large-scale Web Sites* (1 edition). Cambridge; Sebastopol, CA: O'Reilly Media.
- Morville, P., Rosenfeld, L. (2006). *Information Architecture for the World Wide Web: Designing Large-Scale Web Sites, 3rd Edition* (3rd edition). Sebastopol, CA: O'Reilly Media.
- O'Brien, H., & Cairns, P. (2016). *Why engagement matters: Cross-disciplinary perspectives of user engagement in digital media*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-27446-1>
- Österberg, M. (2016). *Web Strategy for Everyone*. Intranätverk. Pobrano z <http://webstrategyforeveryone.com/>
- Powell, M. (2003). *Information Management for Development Organisations*. Oxfam.
- Redish, J. G. (2007). *Letting Go of the Words*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-369486-7.X5000-9>
- Reiss, E. L. (2000). *Practical Information Architecture: A Hands-On Approach to Structuring Successful Websites*. Harlow: Addison-Wesley Professional.
- Resmini, A., ed. (2014). *Reframing Information Architecture*. Springer International Publishing
- Resmini, A., Rosati, L. (2011). *Pervasive Information Architecture: Designing Cross-Channel User Experiences*. Burlington, MA: Morgan Kaufmann.
- Rosenfeld, L. (2011). *Search Analytics for Your Site: Conversations with Your Customers*. Brooklyn, N.Y: Rosenfeld Media, LLC.
- Rosenfeld, L., Morville, P. (2002). *Information Architecture for the World Wide Web: Designing Large-Scale Web Sites, 2nd Edition*. Cambridge, Mass: O'Reilly Media.
- Rosenfeld, L., Morville, P., Arango, J. (2015). *Information Architecture: For the Web and Beyond* (4 edition). Sebastopol, CA: O'Reilly Media.
- Russell-Rose, T., Tate, T., eds. (2013). *Designing the search experience: the information architecture of discovery*. Amsterdam: Elsevier, Morgan Kaufmann.

- Sabharwal, A. (2015). *Digital Curation in the Digital Humanities: Preserving and Promoting Archival and Special Collections*. Elsevier Ltd. <https://doi.org/10.1016/C2014-0-02808-0>
- Saravanan, T. (2012). *Information Architecture: A Multidisciplinary Text Book*. Ess Ess Publications.
- Schmitt, G. (1999). *Information Architecture: Basics and Future of Caad*. Springer Science & Business Media.
- Shanley, C. (2016). *Cracking Information Architecture: The definitive field guide for designers, business managers and project teams*. eBook: Colin Shanley: Books.
- Shepard, D. (2016). *The Information Architecture Handbook – Everything You Need to Know about Information Architecture*. Emereo Pty Limited.
- Shivakumar, S. K. (2016). *Enterprise Content and Search Management for Building Digital Platforms*. Wiley-IEEE Press. <https://doi.org/10.1002/9781119206842>
- Spencer, D. (2009). *Card Sorting: Designing Usable Categories*. Rosenfeld Media.
- Spencer, D. (2010). *A Practical Guide to Information Architecture*. Victoria, Australia: UX Mastery.
- Stackowiak, R., Licht, A., Mantha, V., Nagode, L. (2015). *Big Data and The Internet of Things: Enterprise Information Architecture for A New Age*. New York: Apress.
- Steenson, M. W. (2017). *Architectural Intelligence: How Designers and Architects Created the Digital Landscape*. Cambridge, MA: The MIT Press.
- Tupper, C. (2011). *Data Architecture: From Zen to Reality*. Burlington, MA: Morgan Kaufmann.
- Valverde, R., & Talla, M. R. (2012). *Information systems reengineering for modern business systems: ERP, supply chain and e-commerce management solutions*. IGI Global. <https://doi.org/10.4018/978-1-4666-0155-0>
- Wachter-Boettcher, S. (2012). *Content Everywhere: Strategy and Structure for Future-Ready Content*. Brooklyn, N.Y: Rosenfeld Media.
- Weber, R. H., Weber, R. (2010). *Internet of things: Legal perspectives*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-11710-7>
- Wodtke, C. (2002). *Information Architecture: Blueprints for the Web*. Boston: New Riders.
- Wodtke, C., Govella, A. (2009). *Information Architecture: Blueprints for the Web* (2 edition). Berkeley, CA: New Riders.
- Wurman, R. S. (1997). *Information Architects*. Zürich: Graphis Inc.
- Xu, L. D. (2014). *Enterprise Integration and Information Architecture: A Systems Perspective on Industrial Information Integration*. Boca Raton: Auerbach Publications.

## Bibliografia

- Chadegani, A. A., Salehi, H., Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., Ebrahim, N. A. (2013). A comparison between two main academic literature collections: Web of science and scopus databases. *Asian Social Science*, 9(5), 18–26, [9.03.2019], <https://doi.org/10.5539/ass.v9n5p18>
- Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *The FASEB Journal*, 22(2), 338–342, [9.03.2019], <https://doi.org/10.1096/fj.07-9492LSF>
- Garshol, L. M. (2004). Metadata? Thesauri? Taxonomies? Topic maps! Making sense of it all. *Journal of Information Science*, 30(4), 378–391.
- Kalbach, J. (2012). *Projektowanie nawigacji strony WWW: Optymalizacja funkcjonalności witryny*. Gliwice: Helion.
- León, R. R. (2014). Information Architecture: historical-conceptual analysis. Rodrigo Ronda León. UX, Graphic & Web Designer [9.03.2019], [http://www.rodrigorondaleon.com/documents/Information\\_architecture\\_historical-conceptual\\_analysis.pdf](http://www.rodrigorondaleon.com/documents/Information_architecture_historical-conceptual_analysis.pdf)
- Morville, P., Callender, J. (2011). *Wzorce wyszukiwania: projektowanie nowoczesnych wyszukiwarek*. Gliwice: Helion.

- Morville, P., Rosenfeld, L. (1998). *Information Architecture for the World Wide Web: Designing Large-scale Web Sites*. Sebastopol, CA: O'Reilly.
- Resmini, A., Rosati, L. (2011). *Pervasive Information Architecture: Designing Cross-Channel User Experiences*. Burlington, MA: Morgan Kaufmann.
- Resmini, A., Rosati, L. (2012). A Brief History of Information Architecture. *Journal of Information Architecture*, 3(2) *Information Architecture for the World Wide Web: Designing Large-scale Web Sites* [9.03.2019], <http://journalofia.org/volume3/issue2/03-resmini/>
- Rosenfeld, L., Morville, P., Arango, J. (2017). *Architektura informacji w serwisach internetowych i nie tylko* (4. wyd.). Gliwice: Helion.
- Rosenfeld, L., Morville, P. (2003). *Architektura informacji w serwisach internetowych*. Gliwice: Helion.
- Skórka, S. (2002). Architektura informacji Nowy kierunek rozwoju informacji naukowej. *EBIB Elektroniczny Biuletyn Informacyjny Bibliotekarzy*, 40(11), [9.03.2019], <http://www.ebib.pl/2002/40/skorka.php>
- Skórka, S. (2016). Architektura informacji. W: W. Babik (red.), *Nauka o informacji* (s. 555–575). Warszawa: Wydaw. SBP.
- Sosińska-Kalata, B. (2007). Współczesne oblicze nauki o informacji w Polsce i za granicą. W: E. Gondek, D. Pietruch-Reizes (red.), *Studia z informacji naukowej i dyscyplin pokrewnych. Prace dedykowane Profesor Barbarze Stefaniak* (93–119). Katowice: Wydaw. Uniwersytetu Śląskiego.
- Tafiłowski, P. (2016). Architektura informacji jako problem badawczy informatologii. *Zagadnienia Informacji Naukowej*, 54(2), 37–47.
- 

## Information Architecture in Foreign Literature 1982–2018. Quantitative Analysis

### Abstract

**Purpose/Thesis:** The aim of the article is to present the process of shaping interest in information architecture in world literature.

**Approach/Methods:** A bibliographic analysis of subject literature published until the end of 2018 was carried out to examine the development of information architecture literature. The literature was identified in Scopus, Web of Science and Library and Information Science Abstract, and the Google Scholar, Google Books services and Amazon Bookstore. The quantitative analysis of the chronological and thematic distribution of literature is complemented by the identification of the most important authors, books and journals, as well as the most frequently cited publications.

**Results and conclusions:** 2269 publications on information architecture were identified. They varied in topics, covering mainly information systems and information science issues. The results of the analysis indicate that interest in various aspects of information architecture has been steadily increasing since 2000, as confirmed by the growing number of publications and their citations.

As a result of the conducted query, a collection of 2269 publications on a wide range of topics, mainly in the field of information systems and information science, was created. The results of the analysis indicate that interest in various aspects of information architecture has been continuously increasing since 2000, which was confirmed by the growing number of publications and their citations.

**Originality/Value:** No attempts in Polish publications have been made to analyze literature in the field of information architecture. The article can be a starting point for in-depth thematic analysis and can serve as a comparative material for similar studies.



### Keywords

Bibliometric analysis. Information architecture. Information science. Information systems architecture. Quantitative analysis of literature.

---

*Dr ANNA MATYSEK jest adiunktem w zakładzie Zarządzania Informacją w Instytucie Bibliotekoznawstwa i Informacji Naukowej Uniwersytetu Śląskiego w Katowicach. Jej zainteresowania badawcze skupiają się wokół problematyki normalizacji w działalności informacyjnej, ilościowych badań informacji, architektury informacji i wyszukiwania informacji. Ważniejsze publikacje: Normalizacja europejska w zakresie informatologii, Katowice 2014. Cytowania dokumentów elektronicznych w wydawnictwach naukowych. Nowa Biblioteka 2017, nr 4(27) s. 53–68. Otwarte czasopisma w naukach humanistycznych. Analiza ilościowa na przykładzie bibliologii i informatologii. Nowa Biblioteka 2018, nr 2 (29) s. 65–81.*

*Kontakt z autorką:*

*anna.matysek@us.edu.pl*

*Adres do korespondencji:*

*Instytut Bibliotekoznawstwa i Informacji Naukowej*

*Uniwersytet Śląski w Katowicach*

*pl. Sejmu Śląskiego 1*

*40-032 Katowice*

**Renata Frączek (2017).** *Upowszechnianie wyników badań naukowych w międzynarodowych bazach danych. Analiza bibliometryczna na przykładzie nauk technicznych ze szczególnym uwzględnieniem elektrotechniki.* Katowice: Wydaw. Uniwersytetu Śląskiego, ss. 648. ISBN 978-83-226-3062-4

Od wielu lat w środowisku naukowym utrzymuje się pogląd o braku odpowiedniej reprezentacji polskich prac w międzynarodowym obiegu naukowym. Przyczyn takiego stanu rzeczy upatruje się m.in. w braku spójnych metod promocji badań naukowych czy stosowaniu nieskutecznych mechanizmów widoczności ich efektów. W nurt rozważań na temat obecności polskich badaczy w zasobach światowych wpisuje się także obszerna publikacja autorstwa Renaty Frączek. Przedmiotem rozważań autorki jest analiza dorobku badaczy z afiliacją polskich wyższych uczelni technicznych przeprowadzona na podstawie zasobów bazy Scopus oraz kolekcji baz Web of Science. Poza identyfikacją źródeł zaindeksowanych w bazach, publikacja dokumentuje także obecność w ich zasobach polskich czasopism z dziedziny elektrotechniki.

Przedstawiony przedmiot i cele badań wpłynęły na przyjęcie metod badawczych oraz bazę źródłową recenzowanej książki, a także na konstrukcję jej całości. R. Frączek, przyjmując perspektywę bibliologiczno-informatologiczną dla teoretycznej części pracy, zastosowała metodę analizy i krytyki piśmiennictwa, dla części empirycznej – metody ilościowe. Baza źródłowa objęła 241 publikacji, zarówno polsko-, jak i anglojęzycznych. W realizacji programu badawczego autorka sięgnęła po metody ilościowe. Narzędziem analiz bibliometrycznych uczyniła bazy Scopus i kolekcję baz Web of Science, a pośrednio bazy ARIANTA i BazTech. Pierwsza z baz krajowych miała posłużyć do wyodrębnienia listy czasopism z dziedziny elektrotechniki, druga natomiast – pełnić funkcję weryfikatora zarejestrowanych artykułów z tych czasopism. Materiał badawczy objął swoim zasięgiem chronologicznym lata 1970–2014, choć w przypadku bazy Scopus wykazano także publikacje wydane przed II wojną światową.

Całość rozważań ujęto w czterech rozdziałach. W pierwszym, na podstawie analizy piśmiennictwa fachowego, autorka omówiła kwestie zastosowania metod badań nauko- i bibliometrycznych do oceny nauki, a także przywołała przykłady analiz bibliometrycznych publikacji z zakresu nauk technicznych, społecznych, biologiczno-medycznych i humanistycznych. Mimo że przytoczone przykłady stanowiły właściwą egzemplifikację metody naukometrycznej, w tej części pracy zabrakło jednak odwołania do wkładu w rozwój naukometrii dorobku takich czasopism jak choćby *Nauka*, *Journal of Informetrics*, *Journal of Scientometrics and Information Management*, *Scientometrics*, *Journal of Scientometrics Research* itp.

W rozdziale drugim R. Frączek przybliżyła istotę baz bibliograficznych i baz cytowań. Ponadto wnikliwie scharakteryzowała krajowe i międzynarodowe bazy danych rejestrujące dorobek nauk technicznych. Odniosła się także do celów tworzenia baz dorobku naukowego poszczególnych uczelni technicznych i uporządkowała wiedzę na temat wybranych wskaźników oceny dorobku naukowego.

W rozdziale trzecim autorka nakreśliła historię rozwoju elektrotechniki w Polsce, wskazała jej miejsce w systemie klasyfikacji nauk i eksplorowane obszary badawcze oraz nakreśliła sylwetki jej twórców. Omówiła także pierwsze polskie naukowe czasopisma z tego zakresu i poruszyła zagadnienia kształcenia w zakresie elektrotechniki na studiach wyższych w Polsce.

Rozdziałem najobszerniejszym, a zarazem najciekawszym jest rozdział czwarty. Autorka zaprezentowała w nim wyniki przeprowadzonych przez siebie analiz. Wewnętrznie rozdział ten podzielony został na dwie części: część 1. poświęconą widoczności czasopism z elektrotechniki w bazach danych

Scopus i Web of Science oraz część 2. dokumentującą obecność w tych bazach publikacji badaczy reprezentujących wyższe uczelnie techniczne w Polsce. W części dotyczącej czasopism autorka w pierwszej kolejności przytoczyła tytuły polskich czasopism z zakresu różnych obszarów nauki indeksowane w obydwu bazach, następnie zaprezentowała podstawowe informacje na temat polskich czasopism z elektrotechniki w nich uwzględnionych, przytoczyła ich wskaźniki wpływu oraz dokonała analizy cytowań artykułów z nich pochodzących. Wydaje się, że w kontekście analiz cytowań interesującym byłoby jednak nie tylko wskazanie, ile razy teksty te były cytowane przez innych badaczy, z jaką częstotliwością na przestrzeni lat i w jakich obszarach naukowych zostały wykorzystane, ale przede wszystkim czyjego autorstwa były oraz jakich zakresów tematycznych dotyczyły. Pozwoliłoby to na uchwycenie głównych kierunków badań oraz ich popularność w obiegu naukowym. Takie badanie musiałoby być jednak uzupełnione przynajmniej analizą treściową abstraktów artykułów, co wymagałoby dodatkowych pogłębionych badań, a ponadto zastosowania metod jakościowych.

Druga część rozdziału czwartego przyniosła analizę źródeł informacji zaindeksowanych w bazach Web of Science i Scopus, w których uwzględniono teksty polskich badaczy afiliowanych w uczelniach technicznych. Identyfikacja rekordów z polską afiliacją pozwoliła autorce w pierwszej kolejności na ustalenie wszystkich rodzajów publikacji i tytułów źródeł, w których publikowali polscy badacze z uczelni technicznych. Wyekscerpowanie ze zbioru rekordów prac z zakresu elektrotechniki umożliwiło z kolei określenie udziału procentowego w materiale badawczym publikacji z obszaru elektrotechniki, a w konsekwencji: (1) zidentyfikowanie tytułów czasopism (278 dla baz Web of Science i 51 dla bazy Scopus), w których znalazły się teksty dotyczące elektrotechniki z polską afiliacją, (2) uwidocznienie w rankingu czasopism miejsca polskich periodyków z tego zakresu, (3) wytypowanie ośrodków naukowych z najliczniejszą reprezentacją badaczy publikujących na łamach wszystkich czasopism, (4) wskazanie sieci współpracy autorów publikacji, (5) opisanie wygenerowanego zbioru publikacji pod względem językowym, (6) wyłonienie okresów wzmożonej aktywności publikacyjnej z zakresu elektrotechniki względem częstotliwości podejmowania innych zagadnień z dziedziny nauk technicznych, (7) wskazanie szczegółowych obszarów badawczych (kontekstów tematycznych), w które wpisywały się analizowane publikacje. W przypadku zbiorów danych wyekscerpowanych z obydwu baz danych autorka zastosowała dwie różne struktury i dwa różne stopnie szczegółowości opisu. Dla danych z baz Web of Science podała: wszystkie zidentyfikowane tytuły czasopism (963 tytuły), w których znalazły się polskie teksty z zakresu elektrotechniki, następnie dane szczegółowe (jak: kraj wydania, wydawca, zakres chronologiczny materiału, ogólną liczbę rekordów, w tym rekordów z dotyczących tematycznie elektrotechniki), dotyczące czasopism należących do obszaru tematycznego i kategorii „Engineering” (278 tytułów), a na koniec liczbę rekordów pochodzącą z tych czasopism (278 tytułów) przypisaną do poszczególnych polskich uczelni technicznych. W przypadku rekordów pochodzących z bazy Scopus zastosowała inną kolejność i strukturę opisu: najpierw wymieniła tytuły czasopism (ogólnych i elektrotechnicznych), w których autorzy z afiliacją wyższych uczelni technicznych w Polsce publikowali najczęściej (51 tytułów), następnie w odrębnych tabelach podała tytuły czasopism (202 tytuły), materiałów pokonferencyjnych (160 tytułów), wydawnictw zwartych (160 tytułów) i seryjnych (159 tytułów) tylko z zakresu elektrotechniki, w których ukazało się najwięcej tekstów polskich badaczy, a wreszcie – analogicznie jak dla bazy Web of Science – wskazała liczbę rekordów, pochodzą z czasopism (840 tytułów), w których publikowali autorzy afiliowani w polskich uczelniach technicznych, tym razem nie z elektrotechnicznych, lecz wszystkich zidentyfikowanych w bazie. Mimo pewnych niekonsekwencji w strukturze opisu danych pochodzących z obydwu baz, które łączyć można z faktem, że autorka ze statystyki uczyniła narzędzie rozpoznania i opisu zjawiska oraz że operowała obszernym i zróżnicowanym co do swojej proveniencji materiałem, pozyskane dane pozwoliły Jej na ocenę widoczności publikacji polskich reprezentantów nauk technicznych w międzynarodowych bazach danych. Spośród wykorzystanych przez autorkę baz danych najliczniej prace polskich autorów reprezentowane były w bazie Scopus. Przeważały w niej publikacje przedstawicieli Politechniki Warszawskiej i Wrocławskiej oraz Akademii Górniczo-Hutniczej. Chociaż prace

polskich autorów stanowiły zaledwie 1,3% publikacji zamieszczanych w czasopismach z zakresu elektrotechniki na świecie, badania R. Frączek dowiodły, że od 2000 r. liczba publikacji sukcesywnie wzrasta. W przypadku czasopism polskich prace autorstwa Polaków stanowiły 72% wszystkich wydanych tekstów. Wśród publikacji zaindeksowanych w bazach światowych zdecydowanie dominował język angielski (90%). Na drugim miejscu uplasował się język polski, pozostałe języki odgrywały mniej znaczącą rolę. Badacze afiliowani w polskich uczelniach technicznych najchętniej publikowali w czasopismach wydawanych w Stanach Zjednoczonych i w Wielkiej Brytanii.

Podsumowując, można stwierdzić, że publikacja Renaty Frączek w interesujący sposób udokumentowała obecność polskich czasopism elektrotechnicznych i badaczy z afiliacją polskich uczelni technicznych w obiegu międzynarodowym. Badania te mogą stanowić materiał wyjściowy do krajowej oceny działalności naukowej i stać się inspiracją do podnoszenia rangi polskich periodyków oraz prowadzenia dalszych analiz w innych dyscyplinach z obszaru nauk technicznych, tak w kontekście badań indywidualnych, jak i porównawczych.

**Małgorzata Kowalska-Chrzanowska**

*Instytut Informacji Naukowej i Bibliologii  
Uniwersytet Mikołaja Kopernika w Toruniu*

*Nadesłano: 7 stycznia 2019.*

## Wśród zagranicznych książek

Jacek Wojciechowski

Swoje zainteresowanie nauką o informacji zacząłem pomniejszać po śmierci prof. Marii Dembowskiej, kiedy całkiem rozumną nazwę *dokumentalistyki* zastąpiono *informatologią*, z sukcesywnym coraz większym lekceważeniem bibliotekoznawstwa i bibliotekarstwa, a z czasem również przeważającej części informacyjnej praktyki. Użyteczność spadła do poziomu rzek podczas suszy, pojawiły się natomiast liczne teksty o niczym – z niewielkimi wyjątkami. U nas poza tym obrodziło opasłymi tomami o charakterze autotelicznym, które skupiły się na opisywaniu tego, czym zajmuje się ta dyscyplina.

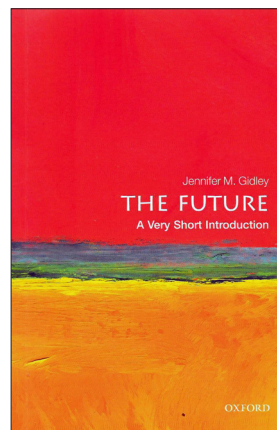
Jeszcze większe zamieszanie wprowadziły z czasem nieweryfikowane prognozy tworzenia *Artificial Intelligence* (pol. sztuczna inteligencja – AI) oraz przesadne radości z powodu postępującej automatyzacji procesów informacyjnych. Trudno to zrozumieć. Jeśli bowiem automatyzacja w ogóle służy różnym rozwiązaniom praktycznym, bo ułatwia i standaryzuje wiele czynności, to odnoszona do procedur intelektualnych, w tym informacyjnych, stara się redukować je do poziomu przełykania albo trawienia, zaś skomplikowane urządzenia przedstawia w ujęciu maksymalnie prostym. Co zaskakuje dodatkowo, ponieważ informatologia nie należy do rodziny dyscyplin inżynierskich, zatem w tych okolicznościach kwalifikuje się tam raczej do biernej obserwacji.

W świetle różnych sygnałów wygląda na to, że zapędziła się na jednokierunkową ulicę, która może prowadzić na zamknięty parking. Otóż to nie jest pomysł dobry, już nawet obecnie, nie mówiąc o ujęciu perspektywnym. Być może nie ja powinienem snuć takie niezbyt miłe refleksje, jednak ktoś musi. Dlatego powołuję się na innych i wybrałem do zasygnalizowania akurat takie teksty. Może ktoś zwróci na nie uwagę i zechce je przeczytać. Albo nie zechce, a jednak przeczyta. Dlaczego nie?

### Jutro to już dzisiaj [\*\*\*\*\*]

Jennifer M. Gidley (2017). *Future. A very short introduction*. Oxford: Oxford University Press, pp. 164 s. ISBN 978-0-19-873528-1

Oto książka, która próbuje zdetonować albo przynajmniej pokruszyć, prognozy przyszłej egzystencji komunikacji: już to zmechanizowanej bądź w pełni automatycznej (zatem ekspulsji bibliotekarstwa też) – a zwłaszcza prognozy opanowania świata przez sztuczną inteligencję – które zawaładnęły potoczną propagandą medialną, ale również naukową, z udziałem nauki o informacji włącznie. Gdyby treść tej książki odnieść do najważniejszych, ściśle eksperckich, wypowiedzi na temat AI, to pozostaje ona stosunkowo bliska sceptycznie umiarkowanej refleksji prof. M. Boden, natomiast zachowuje duży dystans wobec apologetycznych dla AI, ale niekonsekwentnych, opinii prof. N. Bostroma<sup>1</sup>.



<sup>1</sup> M. A. Boden (2016). *AI. Its nature and future*. Oxford: Oxford University Press; N. Bostrom (2016). *Superintelligence. Paths, dangers, strategies*. Oxford: Oxford University Press.

Autorką jest doktor psychologii **Jennifer M. Gidley**, adiunkt na University of Technology – Institute for Sustainable Futures – w australijskim Sydney i zarazem przewodnicząca *World Future Studies Federation*. Zaś do swoich zamierzeń używa wszelkich dostępnych narzędzi, od racjonalnej i dobrze udokumentowanej argumentacji dowodowej, poprzez dodatki demagogiczne oraz jeszcze złośliwości, starając się wyklarować mglisty pejzaż przyszłych nieuchronności.

Z tym, że argument główny jest nie do podważenia. Otóż my wszak nie wiemy w szczególach, jaka ta przyszłość może być i będzie: to tylko antycypacja. Rzeczywistość jutrzejsza nie spadnie z nieba jak manna, trzeba ją dopiero wykreować, natomiast nie ulega **żadnej** wątpliwości, że nie istnieje wyłącznie jeden i jednolity jej wizerunek. Możliwości wariantywnych jest więcej niż dużo.

Tymczasem cierpimy na rozmaite fobie oraz utopie prognostyczne, przejęte z opowieści literackich, nieraz wstrząsających (teraz też filmowych i serialowych), ale przecież opartych na fantazjowaniu, zmyśleniach i fikcji. Czemu towarzyszą jeszcze przeświadczenia o linearnym rozwoju świata, konsekwentnie na mocy wynikania: przeszłość > teraźniejszość > przyszłość. Co da się zanegować w dowolny sposób – no bo niby dlaczego miałyby tak być – jakkolwiek niekoniecznie trzeba. Ale jeśli nie, to pozostaniemy na etapie utopii, który dominuje aktualnie w propagandzie obiegowej.

Bo też wyważona refleksja naukowa nad przyszłością, tak naprawdę ma krótką historię. Mianowicie samo pojęcie wyodrębnionej dyscypliny naukowej, **futurologii**, wylansował nie tak znów dawno niemiecki profesor **Ossip Flechtheim**, równoległe do powszechnie używanego określenia **prognoza**, a regularne studia akademickie nad przyszłością realizuje się mniej więcej dopiero od 50 lat. Wcześniej były swobodne dywagacje na temat samej AI. Natomiast luźne przypuszczenia iluminacyjne, jaka będzie przyszłość, pojawiają się od tysięcy lat, jednak stopień ich racjonalizacji był i jest niski.

Autorka przypomina, że ewolucyjne koncepcje rozwoju świata istniały od dawna. Wspierane świadomością upływającego czasu oraz intelektualną syntezą, układały się w wizerunki przyszłości, jednak zróżnicowane. I nadal tak jest.

Leonardo da Vinci, oprócz tworzenia modeli przyszłych urządzeń, miał pomysł na idealne miasta. Thomas More zapowiadał społeczeństwo wyżej ceniące wartości zbiorowe aniżeli indywidualne. Natomiast nutę optymizmu zamroził był Thomas Malthus, strasząc nadmiarem populacji na świecie. Te oraz inne niepokoje zostały potem podchwyczone przez teksty literackie i w końcu, zamiast techno-optymizmu, rozpowszechniły się rozmaite obawy. Nie lepiej uzasadnione logicznie niż oczekiwania pomyślności, lecz bardziej emocjonalne.

W racjonalnym podejściu przednaukowym wypracowano metodę **predykcji**, czyli statystyczny sposób wnioskowania o możliwych zmiennych w przyszłości, szacowanych na zasadzie prawdopodobieństwa. Ale, mówiąc po ludzku, to są tylko przypuszczenia, chociaż możliwe zrationalizowane i uporządkowane, w sumie jednak loteryjne. Początkowo, wychodząc z przesłanek pozytywistycznej jedności, zakładano tunelowe kształtowanie się przyszłości bezalternatywnej: jednej. Chociaż z próbami oderwania się od rozwojowego obrazu przeszłości, bo predykcja prognostyczna to nie historyczna. Ale dopiero w drugiej połowie ubiegłego stulecia pojawiły się opinie, że wersje przyszłości mogą być rozmaite. Stąd określenie **futuribles** (ang. *possible futures*) – różniące się ponadto w odniesieniach globalnych od ewentualnie wytyczonych indywidualnie.

W niezbyt długim okresie refleksji **naukowej** nad przewidywaniem i kształtowaniem przyszłości wypracowano pięć kierunków prospekcji, niekoniecznie równoważnych, ale też niewykluczających się wzajemnie. Dominująca dotychczas i nadal szeroko akceptowana jest **ewolucyjna**, prosto rozwojowa koncepcja przyszłości, wywodząca się z jednoznacznych koncepcji pozytywizmu. To za jej sprawą prognoza *technologiczna* uchodzi za jedyną wartą zainteresowania. Wobec tego to ona wpływa zasadniczo na wizerunek przyszłego obiegu informacji oraz kształtuje niedobre prognozy dla bibliotekarstwa.

Z nieco bardziej krytycznej refleksji wywodzą się nastawienia na przyszłość **pożądaną** (jednak nie wiadomo: przez kogo) i na dostosowanie zmian do pojawiających się okoliczności. Istnieje również koncepcja, nie wiadomo dlaczego nazywana *kulturalną*, a przewidująca głównie rozwój industrializmu

i konsumeryzmu. Najbliższe idei bibliotekarstwa aktywnego wydaje się założenie **uczestniczącego** kreowania przyszłości, czyli zaangażowanego – w zgodzie z pojawiającymi się okolicznościami. A ponadto istnieje pomysł **integralny**, holistyczny: zebrania oraz implementacji wszystkich tych przesłanek razem, dla uaktywnienia. No i z tego miałyby się wyłonić ramowy zarys pejzażu przyszłości. Jaki? Tego na razie do końca nie wiadomo. Dlatego wszystkie wizje prognostyczne, zarówno apokaliptyczne, jak też nadmiernie radosne, nie mają w tej chwili poważniejszego uzasadnienia.

Jak stwierdziła przywołana (niestety, błędnie) w tej książce znana specjalistka, **Eleanora Masini**: coś się zmieni, bo zmienić się musi, ale także coś zmienić można i trzeba. Otóż nie brzmi to wystarczająco konkretnie, dlatego postępować trzeba bardzo ostrożnie. Natomiast nie ulega wątpliwości, że przyszłość rozpoczyna się już teraz. Dlatego od zaraz konieczny jest mądry namysł.

Tymczasem bowiem dokonano się gigantyczne zamieszanie. Opinie o przyszłej rzeczywistości, generowane głównie w trybie *iluminacji*, niewiele dostarczyły na jej temat argumentów **twardych**, rozmnożyły się więc przede wszystkim spekulacje, a w ślad za nimi następują ewentualnie przyszłościowe decyzje fragmentaryczne, podejmowane *w ciemno*. Wsparte medialną mitologizacją, w której (baśniowe) naddania supertechnologiczne roją się ponad miarę.

Gidley sarkastycznie wypowiada się o mocach *nadprzyrodzonych*, przypisywanych robotom i dronom, a przede wszystkim neguje rozdętą utopię sztucznej inteligencji. Która – nawiasem mówiąc – nie może być sztuczna: to oksymoron, jak zimne gorąco. Wmawianie wszystkim, że maszyny staną się intelektualnie sprawniejsze niż ludzie, to transhumanistyczny kit, obliczony na epatowanie niezwykłością. Z kolei nastawienia posthumanistyczne (są takie) przewidują ingerencje technologiczne w procesy nie tylko manualne, ale także intelektualne właśnie, w ramach idei perfekcjonizmu. W sumie zaś pomysł, że technologia rzeczywiście będzie w stanie wszystkim zapewnić wszystko co potrzebne, także w zakresie dostawy treści, opiera się na przesłankach irracjonalnych. Poza tym nic nie dokonuje się bez konsekwencji ani bezkarnie. Każde nowe narzędzie musi amputować jakąś czynność dotychczasową, bo nie byłoby nowe. Czy to zawsze dobrze? Holistyczny układ komunikacji publicznej i globalny obieg treści mogą świadczyć, że niekoniecznie.

Najbardziej autorkę irytuje, że, w myśleniu o przyszłości, technologiczne utopie całkowicie prześlониły odczucia humanistyczne. I namawia ona do sprzeciwu wobec cybernetycznej wizji rozwoju inteligencji, bo to nie znajduje żadnego solidnego wsparcia dowodowego. Bierze się zaś z mocno kontrowersyjnego przywiązania do bezalternatywnej, linearnej koncepcji dziejów, dopełnionej nastawieniem na industrializm. Otóż to jest tylko jedna z szeregu ewentualności prospektywnych, w dodatku stosunkowo słabo uargumentowana.

W autorskim ujęciu J. M. Gidley program pracy na rzecz przyszłości powinien być inny: łącząc ochronę wartości humanistycznych z rozumną mechanizacją (automatyzacją – w mniejszym zakresie) tego, co można i warto. Ale sam rozwój społeczny musi być wieloraki oraz proporcjonalny, w szerokim wymiarze: psychiczny, duchowy i estetyczny. Trzeba chronić odczucia, empatię, uczciwość, a nade wszystko dbać o stałe wzbogacanie świadomości. Natomiast przeciwstawiać się wszystkim zabiegom destrukcyjnym, zwłaszcza techno-utopijnym.

Maszyny są wyłącznie replikacyjne. Tylko człowiek ma ogromne i niedoszacowane możliwości rozwoju intelektualnego oraz kreatywnego. Jednak pod warunkiem aplikacji myślenia ponadformalnego i nawet paradoksalnego, z wykorzystaniem intuicji, wyobraźni, jak też twórczej relatywizacji.

Na razie wygląda na to, że prawie wszyscy czują respekt przed przyszłością i różnymi z jej strony zagrożeniami. Oczywiście: nie można z góry wykluczyć najrozmaitszych kryzysów i na ogół trudno je przewidzieć. To w każdym razie nie znaczy, że na pewno ich nie będzie. Otóż zapobieganie przedzające polega na kreowaniu społecznego **pożytku**, choćby przez dobrze zorganizowane i produktywne rozpowszechnianie użytecznych informacji, jednak oparte nie tylko na transmisyjnej mechanizacji, lecz w równym stopniu na czynniku ludzkim.

Gidley nie opowiada bajek i nie sugeruje, że wszystko zrobi się samo. Bez dobrego przygotowania nie uda się nic! Obecnie fundament preparacyjny wyznacza edukacja: szkolna i poszkolna. Jeśli jednak

realizowana w trybie niby zmodernizowanym, czyli mocno digitalnym, to tylko udaje modyfikację. Serwowane treści są ciągle te same, tylko inaczej opakowane. Właśnie tak, imitacyjnie, wygląda pseudouniwersalny program kształcenia UNESCO *Education for All*, w istocie wyzuty z rozwiązań perspektywnych. Mamy zatem do czynienia nadal z *futuribles*. Które trzeba starannie przemyśleć, unikając błędów. Nic nie jest jeszcze przesądzone.

## Infoignorancja

Tom Nichols (2017). *The death of expertise*. New York: Oxford University Press, pp. 252. ISBN 978-0-19-040911-2

Pomimo rozmaitych złośliwość i nieco demagogicznych komentarzy, szczególnie w odniesieniu do sztucznej inteligencji oraz do stechnicyzowania przyszłej rzeczywistości, na ogół J. Gidley utrzymała się w tonacji obowiązujących kanonów dyskusji naukowej. Natomiast nie przejął się nimi nadmiernie Tom Nichols, profesor Naval War College oraz Harvard Extension School w amerykańskim Cambridge (Newport). Nie dość, że papla i gaduła, to jeszcze – jak przystało na wykładowcę wojskowej szkoły morskiej – dokonał zmasowanego ataku rakietowego na cały obecny system komunikacyjno-informacyjny, osiągając spustoszenie wręcz apokaliptyczne. A co najdziwniejsze: często w tych swoich szarżach ma rację albo jest tej racji bliski. Na pewno nie wszyscy się cieszą.

Nichols uważa mianowicie za kompletną bzdurę całą tę gadaninę o populacji informacyjnej oraz o społeczeństwie, rzekomo opartym na wiedzy. Żadnej wiedzy nie ma, więc na czym się tu opierać? Wprawdzie nie napisał tego wprost, ale postrzega jedynie społeczeństwo oparte na ignorancji.

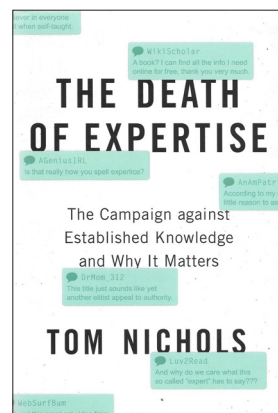
Dzisiaj każdy, kto nabył klawiaturę i wtyczkę do sieci, doskonale wie, jak zbawić świat oraz co pomaga na odciski, a także jak zwrócić Krym Ukrainie, mimo braku orientacji, gdzie ta Ukraina leży. Jednocześnie na uczelniach młodociani faceci z laptopami instruują profesorów, czego mają uczyć i jak, bo skoro masz laptop, to wiesz lepiej niż oni. Wprawdzie wciąż jeszcze, przez nieuwagę, nazywa się profesorów ekspertami, lecz wszyscy boczą się na nich, burzą i uważają za przemądrzałków, przecież każdy sam wie wszystko lepiej. Stąd taki a nie inny tytuł opracowania. Zmierzch epoki ekspertów sugerowali zresztą również R. i D. Susskindowie w bardzo głośnej książce bestsellerowej<sup>2</sup>.

Aktualna zaś sytuacja jest taka, że każdy czegoś nie wie i wszyscy są z tego niebawale dumni. Króluje zadufanie, narcyzm oraz bezgraniczna pewność siebie.

Nichols uważa, że serwowana przez sieć stechnicyzowana kultura jest jeszcze gorsza niż faktyczna ignorancja. Internet serwuje bowiem wiedzę pozorną, a więc dezinformuje, zamazując informację rzeczywistą. Stwarza mit równości wszystkich w zakresie wiedzy, toteż każdy wypowiada się autorytatywnie na dowolne tematy, w przeświadczeniu o własnym znanstwie. Uważając jednak za wiarygodne wyłącznie to, co jest zgodne z indywidualnym przeświadczeniem.

Każdy zatem wie wszystko, każdy zna się na wszystkim i wygląda na to, że każda opinia może uchodzić za równowążną innym. Wobec tego wypowiedzi eksperckie tracą na znaczeniu i zamierają: mało kto chce ich słuchać. Oczywiście eksperci też myślą się niekiedy, ale to są raczej odstępstwa od reguły, aniżeli zdarzenia dominujące. Jakkolwiek z kolei istnieje też nadmierne zadufanie eksperckie: kiedy znawcy, wiedząc *coś o czymś*, mają przeświadczenie, że wiedzą *wszystko o wszystkim*.

W ogóle nadmierne mniemanie o sobie oraz o swoich kompetencjach przybrało rozmiary epidemii. Niby bowiem poprawiło się kształcenie, lepszy jest podobno dostęp do informacji, a media rzeczywiście



<sup>2</sup> R. Susskind, D. Susskind (2017). *The future of the professions. How technology will transform the work of human experts*. Oxford: Oxford University Press.



stały się ogólnodostępne. Ale w rezultacie to pogorszyło wymianę opinii oraz spłycało dyskusje: są kłótnie, zamiast dialogów. Opiniodawcy upierają się przy stanowiskach bezdowodowych, przyjętych na wiarę, a w stosunku do kontrowersji mnożą się zarzuty oraz teorie spiskowe.

Nichols rozszalał się szczególnie, opiniując obecne studia wyższe, które jego zdaniem uległy deprecjacji za sprawą ich urynkowania oraz w następstwie przyznawania studentom racji na każdym kroku. To są bowiem teraz klienci, wnoszący do uczelnianych biznesów to, co tam najważniejsze: opłaty. W nastawieniu na pozyskiwanie środków, uniwersytety akceptują byle co – zdevaluowała się już sama uczelniana profesja. Natomiast utrzymał się mit wartości wyższego wykształcenia. Absolwenci uważają się za świetnie wyedukowanych i nikogo nie interesuje, że nikt ich nigdzie nie wdraża do myślenia krytycznego.

Nie bez racji dodaje też – u nas również to widać – że w uczelnianych relacjach zaczynają przeważać odniesienia mailowe. Niby to nic złego, ale w następstwie pojawia się przekonanie, że wykładowcy oraz studenci są w zakresie wiedzy równi. Poza tym studenckie oceny wykładowców, sformułowane anonimowo i schematycznie, dewaluuja studia: mają však oświecać profesorów! Oraz pod takie dyktando ma też miejsce modyfikacja i przebrzmowanie programów nauczania. W rzeczywistości zaś okazuje się, że to jest „smutny kit”.

W aranżowanych niby dialogach pojawiają się nerwowe sondaże akceptacji: jak bardzo lubisz swoje studia bądź jak bardzo ich nie lubisz. Uczelnia upodabnia się więc do restauracji. Także rozpoznanie studenckiej aktywności nie jest budujące. Ujawniają się przywary, jakby zachowane sprzed stu lat: nietolerancja, dogmatyzm, a nawet przemoc, to atrybuty częste, jeśli nie powszechne. W związku z taką wylizanką, nie nazwałbym autora przyjacielem studentów.

On zresztą sieci również przyprawia kły – czasami zasadnie, a niekiedy nie. W Internecie, zdaniem wielu, jest wszystko o wszystkim; nie ma tam kwestii nierozwiązywalnych. No to po co pytać ludzi, wszak często omylnych, skoro można przepytac komputer? W rezultacie każdy uważa, że coś wie, ale to są głównie dyrdymały (półdyrdymały?) z Wikipedii. Czyli to jest komunikacja instant, w której na dodatek anonimowość lokuje wszystkich na tym samym poziomie ignorancji.

W zapale inwektyw, Tom Nichols sam wpada w pułapkę pustostłowa. Powiada mianowicie, że jakoby już **Theodore Sturgeon** ponad sto lat temu sugerował, że 90% całej naukowej wiedzy to śmieci. No więc bajdurzy. Przywołany w ten śmieciowy sposób został szwajcarski profesor chirurgii, który naprawę nazywał się **Theodore Billworth** (1829–1894) i nie ma żadnego dowodu, żeby twierdził właśnie coś podobnego, ani że był w takiej opinii kompetentny. Chirurg to chirurg – nawet jeśli bardzo dobry<sup>3</sup>.

Mój pogląd w tym zakresie jest taki, że jeśli ktoś chce krytykować czyjś dorobek naukowy, to powinien przedtem sam sobie narzucić rygory wiarygodności, ujawniając starannie kto, gdzie i co napisał albo powiedział. Bez tego tworzy się plotkarnia. No i Nichols w czymś takim uczestniczy.

Owszem: media potrafią każdemu wcisnąć dowolny idiotyzm, zatem ogłupiają. Ale też, a nie wyłącznie – nie da się uogólniać bezwarunkowo. Szczególnie dają popalić niekompetentne medialne dziennikarstwo, oparte w Internecie głównie na beztreściowym klikaniu. Zaś odbiór jest taki, że każdy wierzy w to, co odnotowuje i widzi. Tak utrwala się ignorancja.

Nichols nie ma wątpliwości, że obiegu informacji (treści) nie można rzucić na żywioł i pozostawić samemu sobie. Pośrednictwo musi być: koniecznie biblioteczne i mądre. Ale jak je zorganizować, zapewnić oraz wykorzystać – o tym już nie dopisuje ani słowa. Ucieka w ogólne określenie **ekspert** – lecz ich charakterystyka jest w książce mętna.

<sup>3</sup> T. Nichols odwołuje się do tzw. prawa Sturgeona (ang. *Sturgeon's Law*), tj. aforyzmu, według którego większość każdego opublikowanego materiału, wiedzy, a najogólniej – wszystkiego, jest bezwartościowa (zob. np. Oxford Dictionaries, <https://www.oxforddictionaries.com/>). Wyrażenie tego poglądu przypisuje się amerykańskiemu pisarzowi i wydawcy fantastyki naukowej, Theodorowi Sturgeonowi (1918–1985, właśc. Edward Hamilton Waldo), którego powieści tłumaczone były również na język polski; miał on stwierdzić, iż „90 per cent of everything is crap” [red.].

Eksperci nie zawsze przyznają, że czegoś nie wiedzą, także myślą się niekiedy i zresztą mają do tego prawo – byle nie za często. Niepewność jest integralnie wpisana w strukturę poznania, materia badań ma ograniczenia, toteż ostrożność konkluzyjna stanowi warunek pierwszy. A czy w każdym wypadku przestrzegany? Główna zasada nauki: powtórzenie procesu dociekań z tym samym rezultatem końcowym – nie zawsze daje się spełnić, a jeżeli nawet, to przeważnie z ogromnym trudem. Występują też błędy metodologiczne; niekiedy ustalenia z jednej dyscypliny bywają bezzasadnie implementowane do innej, a bywa i tak, że, zamiast objaśnień zjawiska, pojawiają się gołosłowne sugestie. Żeby już nie wspominać o udawaniu kompetencji oraz o zamierzonych fałszerstwach, których trochę jest. Życie naukowe to też przecież życie.

Mimo wszystko, bez ekspertów świat nie da sobie rady, trzeba ich zatem ogólnie akceptować, natomiast uważniej patrzeć na ręce. Bądź co bądź są tylko doradcami, a nie decydentami. Jednak publiczność, zauroczona siecią i zadufana w sobie, bardzo ekspertów nie lubi. Podtrzymywana w tej swojej niechęci przez szkoły, które wmawiają młodzieży, że najważniejsze są jej **odczucia**, a jeszcze bardziej przez media, gdzie zamiast sygnalizacji co jest ważne, padają pytania: czego ludzie chcą się dowiedzieć. I potem pojawiają się byle jakie odpowiedzi.

Trudno byłoby ukryć, że wywody Toma Nicholasa nie zawsze są na profesorskim poziomie. Mimo to częściej ma jednak rację, aniżeli jej nie ma. Natomiast niekiedy dostosowuje się stylistycznie do – bardzo od naukowości odległego – stylu argumentacji okołoinformacyjnej, promedialnej i ogólnokomunikacyjnej. Dla polemiki, a także dla rozbudzenia wątpliwości, to może być tryb użyteczny. Natomiast dla ustalenia całej prawdy i wyłącznie prawdy, to chyba jednak nie wystarczy. Przydałoby się stanowisko bardziej eksperckie. Ponieważ na razie wobec pytania o granice między ignorancją a rozeznaniem, kompetentna odpowiedź nie pada.

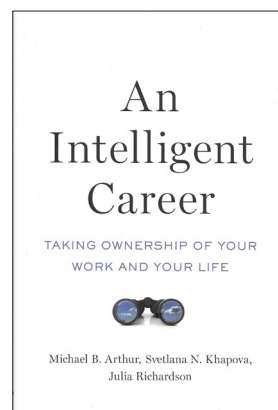
### A jednak: pochwała intelektu [\*\*\*\*]

Michael B. Arthur, Svetlana N. Khapova, Julie Richardson (2017). *An intelligent career. Taking ownership of your work and your life*. New York: Oxford University Press, pp. 264. ISBN 978-0-1904-94131

Trudno naukowe wydawnictwa szeregować w rankingowych tabelach, jak piłkarskie kluby, ale oczywiście są wśród nich firmy lepsze i gorsze, świetne oraz beznadziejne. Oxford University Press (UOP), z siedzibami w różnych punktach świata, jest jak futbolowy Bayern Monachium. Zdecydowanie dba o wysoki poziom treściowy, chociaż zdarzają się pojedyncze wpadki. Zakres tematyczny oferuje niebywały oraz rozpowszechnia w każdy obecnie możliwy sposób, więc przez publikację drukowane, cyfrowe i hybrydowe czasopisma. Marka OUP ma swoją wartość.

To jest także wydawnictwo, które nabywcę traktuje jak mało kto, mianowicie poważnie. Oferuje dobrą informację oraz promocję niezbyt nachalną, ceny są (powiedzmy) znośne, a transakcje niekłopotliwe, zatem jako taki komfort jest. Będąc klientem OUP od wielu lat, opinię wystawiam korzystną. Dla porównania: wielu innych naukowych edytorów zachowuje się rozmaicie, czasami wręcz po bandycku – bywa, że nóż sam się otwiera w kieszeni.

Poza tym ma ta oficyna ciągoty ku półnaukowej popularyzacji różnych obszarów wiedzy i radzi z tym sobie skutecznie. Otóż właśnie: w niejkiej kontynuacji poprzednich, przywołanych tu wypowiedzi, pojawiła się tam książka o charakterze doradczym – namawiająca do wysiłku **intelektualnego** oraz do skupienia się na indywidualnym rozwoju umysłowym i zawodowym. Motywująca **każdego**. Na przekór całej tej paplaninie o przyszłej mechanizacji życia i sztucznej inteligencji.



Szkoda tylko, że nie obeszło się bez gadulstwa i rozmaitych uproszczeń. Każdy zakres książkowych roztrząsań – a jest ich kilka – poprzedza swoista przypowieść bajkowa, nieraz tak infantylna, że chce się wyć. Ale potem skojarzenia z kierunkowymi dyrektywami uzasadniają się nie najgorzej. A zmierzają do wmówienia wszystkim, że pieczę nad swoją zawodową karierą i nad własnym życiem najlepiej powierzyć samemu/samej sobie.

Nie wszystko w tej książce jest rzeczywiście odkrywcze, lecz czy musi być? Są powtórzenia rozmaitych porad pradawnych i dawnych, ale są również dodane nowe, podpowiadające jak racjonalnie oraz skutecznie kształtować swój własny rozwój. Z czym, gdzieś na marginesie, kojarzy się również wsparcie biblioteczne. A nad całością góruje pochwała **inteligencji**, pojmowanej jako zdolność opanowania i zrozumienia rozległej wiedzy oraz umiejętności uczenia się nowych treści. Przez ludzi, a nie przez cyborgów.

Autorski skład jest w tej książce barwny oraz globalny, właśnie na miarę OUP. Oto prof. **Michael Arthur** z Suffolk University w amerykańskim Bostonie, znawca problematyki zarządzania. Oraz prof. **Svetlana Khapova** z Vrije Universiteit w holenderskim Amsterdamie: zajmuje się organizacją. A do tego prof. **Julia Richardson** z Curtius Business School w australijskim Perth – wyspecjalizowana w zakresie kierowania ludźmi. Oczywiście taki globalny triumwirat dało się stworzyć tylko za sprawą relacji sieciowych. Czy ja mówię, że nie?

Razem stworzyli użyteczny oraz ciekawy podręcznik, nawet jeżeli miejscami nafaszerowany trywialnościami, a niekiedy nie brak tam również naiwności. Ale ogólnie – tak wyobrażam sobie produktywny serwis informacyjny. Przy czym zdecydowanie ukierunkowany humanistycznie: bez leżenia płackiem przed technologią. Mianowicie cała trójka autorów nie ma wątpliwości, że również w przyszłości leczyć nas będą lekarze, a nie ślusarze, bądź mechanicy.

Zdaniem autorów obecnie rosną możliwości podjęcia pracy wysoko kwalifikowanej. Akurat takiej sugestii nie sformułowałby Tom Nichols, ale być może oni znają się na tym lepiej. Jest bowiem taka dopowiedź, że teraz firmy nawet same szykują i przygotowują pracowników do nowych kompetencji. Projektują odpowiedni nabór, a potem sprzyjają rozwojowi umiejętności. A skoro tak, to trzeba stosownych okazji szukać. Oczywiście że – jak napisano – także przez sieć. Ale być może kompetentne pośrednictwo bibliotek też byłoby szczególnie pożądane.

W takich okolicznościach, zabiegając o dobrą pracę, trzeba o sobie wiedzieć możliwie dużo i potraktować tę orientację jako azymut kierunkowy. Skojarzyć z tym, czego potencjalny pracodawca oczekuje, w myśl zasady podwójnej lojalności: wobec siebie i swoich zamierzeń oraz wobec wybranego pracodawcy i jego zadań. Niekiedy to nie jest możliwe albo z czasem przestaje być możliwe. No to wtedy trzeba zmienić otoczenie.

Autorzy odnoszą się oczywiście do edukacji (to istotny powód dla sygnalizowania tutaj tej książki) jako głównego narzędzia dostarczania treści informacyjnych oraz nabywania umiejętności, jednak czynią to dosyć schematycznie. Są mianowicie bezgranicznie przychylni nowym ofertom kształceniowym, w tym również sieciowym oraz zautomatyzowanym i sugerują (niekoniecznie w zgodzie z prawdą), że teraz w nauczaniu zaczyna przeważać nastawienie na uczących się, a nie (jak dawniej) na nauczających. Jak tu zasygnalizowałem, T. Nichols ma na ten temat całkowicie odmienną opinię i trudno odmówić mu racji: dobrze nie jest.

Zresztą w tej książce również znalazła się garść ostrzeżeń. Choćby przestroga, że samo przyjęcie informacji to jeszcze nic nie jest, trzeba bowiem opanować umiejętności procesualne, a zatem nieodzowna bywa praktyka wykonawcza, możliwa dopiero u pracodawcy. Tam więc ma miejsce finalizacja, lub dalszy etap, uczenia się i trzeba to wykorzystać dla własnego rozwoju, również intelektualnego. W tym zakresie autorzy nie trzymają się zresztą technoutopijnej apologii, lecz podkreślają wartości ludzkiego myślenia.

Takie proedukacyjne opisanie pracodawców oraz firm zatrudniających tworzy wokół nich klimat sprzyjający. Autorzy są zdania, że dokonało się znaczne ucywilizowanie rynku pracy wykwalifikowanej. Nie ma już – podobno – dawnego drenażu mózgowi ani kaperowania geniuszy, natomiast ukształtował się model cyrkulacji wiedzy pogłębionej. To znaczy firmy **wspólnie** inwestują w ośrodki

profesjonalnego (w istocie eksperckiego) doskonalenia, a później korzystają razem z rezultatów. Co zresztą wymaga skoordynowanego systemu informacji, lecz o tym nie ma w książce ani słowa.

Nowość polega na tym, że oto powstały specjalne obszary wysoko kwalifikowanego kształcenia. W Stanach Zjednoczonych istnieje **Silicon Valey**. W Europie Zachodniej natomiast – obszar od Manchesteru i Londynu, przez Brukselę i Zurich, aż po Mediolan. No i w tamtejszych ośrodkach właśnie wylawia się specjalistów co lepszych, a zwłaszcza lepszych z najlepszych.

W ślad za takim sposobem edukowania samodzielnych (do pewnego stopnia) w istocie ekspertów, zmienia się częściowo forma zatrudniania. Coraz więcej osób, tak właśnie przygotowanych, podejmuje pracę w trybie **kontraktowym**, jeżeli tylko to jest możliwe. Pozostają więc z pracodawcami tylko w ogólnej relacji pracobiorczej, natomiast większość lub wszystkie powinności realizują na własną rękę, często poza biurem, w samochodach, na parkingach i w ogóle w ruchu. Ceną jest wzmożony stres, wyczerpanie oraz poczucie intensywniej eksploatacji, lecz odwrotu nie ma.

W takich okolicznościach rzeczywiście wzrasta niepomierne rola Internetu jako narzędzia nie tylko łączności, lecz i wykonawstwa. Ale wobec tego można również wyobrazić sobie powinność sieciowego wspierania zawodowego ekspertów i wykwalifikowanych agentów ze strony **bibliotek** – w zakresie doradczym oraz pośrednictwa, a nawet do pewnego stopnia stabilizacyjnym. Jednak takich inicjatyw nie zauważyłem.

Autorzy zwracają też uwagę na zasadnicze zmiany w obecnej praktyce wybierania profesji, zwłaszcza przy założeniu długotrwałej kontynuacji zawodu. Kiedyś było tak, że dzieci przejmowały zawód często po rodzicach, a niekiedy także ich firmy, bez szczególnego namysłu: po prostu kontynuacyjnie. Obecnie tak nie jest lub jest bardzo rzadko. Konieczna więc bywa staranna samoanaliza – rozpoznanie cech własnych, w zestawieniu z charakterystyką przewidywanej specjalności zawodowej do podjęcia.

W tekście pojawia się sugestia praktyczna. Zastanów się szczegółowo, jaki/jaka jesteś i poszukaj (zdaniem autorów: tylko w sieci – i to uważam za grubą błąd) ofert pracy, które dają się z tym kojarzyć. Rejestr wariantów analitycznych wygląda znajomo, wręcz nawet stereotypowo; nie widzę w nim nic nowego, co zresztą nie znaczy, że jest do luzu. Jaki umysł: praktyczny, analityczny, kreatywny, kooperacyjny czy kompetencyjny? Hm. A jak wszystkie razem? Zwłaszcza, że taka klasyfikacja nie pochodzi z najwyższej półki teoretycznej. Oraz jakie przeważają nastawienia własne – głównie na siebie bądź również na innych lub ogólnie elastyczne. Następnie zaś trzeba porównać te swoje inklinacje, z tym czego potencjalny pracodawca oczekuje.

Z wyborem miejsca pracy można trafić dobrze albo mniej dobrze, bądź w ogóle kiepsko, dlatego nie należy decydować o tym w ciemno. Potrzebna jest informacja o rodzaju pracy i o pracodawcy, ale na taką natrafić trudno, dominuje bowiem kamuflaż, w sieci dodatkowo wzmożony, bo poza kontrolą. Natomiast kiedy już zapadnie decyzja na tak, to trzeba jeszcze rozliczyć się z własnych intencji. Czy popracujesz głównie dla zarobku, czy dla instytucjonalno-strukturalnej kariery oraz awansów, czy też przede wszystkim dla spełnienia siebie. Każdy z tych kierunków wymaga nieco innego podejścia, chociaż można założyć sobie wszystkie trzy razem, jakkolwiek realizacja nie jest wtedy łatwa.

Taka autoanaliza ma sens również dlatego, że ambitni i wymagający pracodawcy coraz częściej testują kandydatów i szukają wśród nich osób najszybciej. Można z tego skorzystać też dla celów własnych, ale przede wszystkim pokazać się w świetle możliwie korzystnym – jeżeli ma miejsce wcześniejsze przygotowanie. Potrzebna jest zatem inwencja i kreatywność.

Na to, co człowiek może wnieść do procesu pracy, składa się wiedza nabyta, przetworzona w umiejętności – z wykorzystaniem doświadczenia oraz obserwacji innych (w tekście przesadnie uwypukla się ogląd sieciowy). Oryginalne rozróżnienie polega na tym, że mowa o koniecznym opanowaniu umiejętności w trybie **twardym**, zwartym – mianowicie technologicznych i analitycznych. To jest konieczne, ale nie wystarczy. Istnieją bowiem również kompetencje **miękkie**, określane inaczej jako inteligencja emocjonalna – tym bardziej niezbędne do opanowania. No i tego nie da się osiągnąć w trybie sieciowym, ani tym bardziej automatycznym, bez wsparcia przez mediacyjnych, ale **żywych** pośredników profesjonalnych.

W sumie wydaje mi się, że autorzy w życzliwości dla sieci zachowują jednak umiar i konieczny rozsądek. Proponują rozpoznanie tam ofert pracy oraz zawodowych opinii rozmaitych osób – z przyrównaniem do własnych umiejętności oraz inklinacji. To, z grubsza biorąc, jest sensowna porada: przemyślenie ewentualnego azymutu zawodowego. Natomiast napisano też wyraźnie, że **relacje** sieciowe są (jednak!) iluzoryczne. Jest wśród nich szansa **ogólnego** rozpoznania innych oraz **wstępnej** prezentacji siebie, lecz niewiele więcej. Trzeba mieć taką świadomość. Również: egzystencji w dwóch różnych rzeczywistościach, mianowicie realnej oraz wirtualnej, które jednak nie są całkiem odrębne. To środowiskowa hybryda, z dalekim od naturalności obszarem technologicznym. Nad którym trzeba panować, bo jest niebezpieczny. Już teraz, a podobno będzie jeszcze gorzej, jeśli wierzyć w Prawo Moore'a sugerujące, że ilość tranzystorów podwaja się co dwa lata. Ode mnie: wolne żarty – kto to i jak obliczył? Moje własne, nie gorsze, prawo jest takie, że liczba much w bibliotecznych magazynach redukuje się proporcjonalnie do redukcji magazynów.

W książce napisano, żeby imać się przede wszystkim takiej pracy, która dostarczy także nowej wiedzy – na przyszłość. Tego często nie widać na pierwszy rzut oka, toteż trzeba szukać w trybie pogłębionym i w żadnym wypadku nie przeoczyć szansy, bo takich w nadmiarze nie ma. Ciekawostka: rozdział na ten temat przywołuje motto z wiersza **Wisławy Szymborskiej** *Nic dwa razy się nie zdarza* – w angielskim przekładzie. Trafne.

Odwrotną stronę zawodowej kariery wyznaczają niepowodzenia, które zdarzają się zawsze, ale nie muszą być górą. Z autorskich wypowiedzi wynika, że trzeba je sobie uświadomić, spisać i wmontować w wykres indywidualnego rozwoju. Żeby oderwać się od nich, uniezależnić oraz przemyśleć różne sposoby reagowania.

Przy postawie biernej powstaje coś na podobieństwo negatywnej spirali reakcji. Dominuje bezradność i stałe pogorszenie sytuacji oraz nastroju. Można też (inaczej) przyjąć postawę neutralną: trwać bez poszukiwania rozwiązań pożytecznych, zatem bez sukcesów, ale też bez postępującej autodestrukcji. Jednak najlepiej przełamać inercję i podjąć aktywnie poszukiwania rozwiązań lepszych – prospektywnych. Niby racja, ale i tutaj, i w książce brzmi to bardzo drętwo.

Rzecz jasna to pracodawcy określają warunki zatrudnienia i funkcjonowania, trzeba je zatem rozpoznać, przeanalizować, zrozumieć i wykorzystać, dla pożytku obu stron. Autorzy twierdzą, że to da się zrobić przez sieć, byle tylko skupić się na istocie wskazanych tam zadań do wykonania i potraktować ją kreatywnie. A następnie stworzyć zarys produktywnej autoprezentacji, eksponując zwłaszcza to, co ma się przydatnego w związku z tym do zaferowania.

Podjęcie pracy to ważny krok, niekoniecznie jednorazowy, natomiast dobrze, jeżeli za każdym razem w jakiś sposób rozwojowy. Dlatego wymaga możliwie szerokiego rozeznania, intuicji oraz spokojnej analizy. Trzeba sporo pytać, a dopiero potem formułować własne wnioski. Korzystne dla siebie – co zresztą nie znaczy, że przeciw innym.

Nawet odwrotnie: dobrze byłoby odszukać inne osoby o podobnych zamierzeniach i pofunkcjonować razem, w takim stopniu i tak długo, jak to jest możliwe. W książce pojawiają się opinie, że bywa w tym pomocna informacja sieciowa, ale sporo w tej sugestii legendziarstwa, nie ma natomiast konkretów. W praktyce nic nie jest takie proste.

Ta monografia w sposób jednoznaczny lansuje koncepcję pracownika kreatywnego: doinformowanego, umiejącego co trzeba oraz rozwijającego się nieustannie. Dla dobra własnego, dla pożytku pracodawcy i dla wartości ogólnych. Podstawę stanowi opanowanie wiedzy, umiejętności oraz nabranie wprawy. Nie tylko w oparciu o przejęte treści, lecz także odczucia i przemyślenia, z czym przekazy automatyczne nie mają nic wspólnego.

Nie jest to książka bezbłędna. Obok sugestii nowych i produktywnych, przywołuje również dyrektywy odległe w czasie, być może częściowo już przestarzałe. Ale zwraca uwagę już nawet swoją formą: monografii drukowanej o charakterze poradnikowym. Inicjatorzy założyli najwidoczniej, że w tej postaci zdoła spełnić swoje powinności w stopniu przynajmniej dostatecznym. I jest to wysoce prawdopodobne.

*Nadesłano: 28 kwietnia 2018.*

## Przegląd polskich nowości wydawniczych

Anna Stanis

*Biblioteka Uniwersytecka w Warszawie*

A. Januszko-Szakiel (2017). *Archiwistyka cyfrowa: długoterminowa ochrona dziedzictwa nauki i kultury*. Warszawa: Wydaw. SBP, ss. 240. Seria „Nauka, Dydaktyka, Praktyka” 181. ISBN 978-83-65741-02-8

Książka ukazuje problemy długoterminowej archiwizacji zasobów cyfrowych z punktu widzenia teoretycznego, prawnego, technicznego, ekonomicznego i organizacyjnego. Dotyczy głównie dokumentów, które stanowią dziedzictwo narodowe i są określane jako narodowy zasób cyfrowy. W związku z brakiem jednoznacznej, uzgodnionej definicji terminu narodowy zasób cyfrowy autorka proponuje własną definicję, opracowaną na potrzeby książki. Omawia także definicje pojęć podstawowych z zakresu długoterminowej archiwizacji zasobów cyfrowych.

Opisuje model standardowy ISO dotyczący organizacji i funkcjonowania archiwów cyfrowych (Open Archival Information System) oraz system DSEP (Deposit System for Electronic Publication) wdrożony w Bibliotece Narodowej w Holandii.

W obszarze technicznym dominuje temat trwałości nośników cyfrowych, formatu zapisu danych i metadane zapisów cyfrowych. Ekonomiczne zagadnienia trwałej ochrony zasobów cyfrowych związane są z ustaleniem źródeł ich finansowania.

W książce szczegółowo omówiono także podstawy prawne trwałej ochrony zasobów cyfrowych, prawo autorskie i prawa pokrewne. Przybliżono zagadnienie *preservation policy*, czyli zbioru dokumentów o charakterze umów, zarządzeń, ustaw regulujących czynności związane z długoterminową ochroną zasobów cyfrowych. Każdy kraj tworzący narodową *preservation policy* powinien, obok narodowych aktów prawnych, uwzględnić też dyrektywy nadrzędne, np. te dla krajów Unii Europejskiej.

Organizacyjne zagadnienia długoterminowej archiwizacji zasobów cyfrowych wiążą się z badaniami dotyczącymi poziomu świadomości pracowników polskich instytucji pamięci konieczności podjęcia działań na rzecz ochrony polskiego dziedzictwa cyfrowego. Omówiono polskie wybrane inicjatywy na rzecz trwałej ochrony polskiego dziedzictwa cyfrowego.

W ostatniej części publikacji przedstawiono podstawowe założenia autorskiej propozycji programu długoterminowej ochrony polskich zasobów cyfrowych.

Publikację uzupełnia obszerna bibliografia na stronach 221–240.



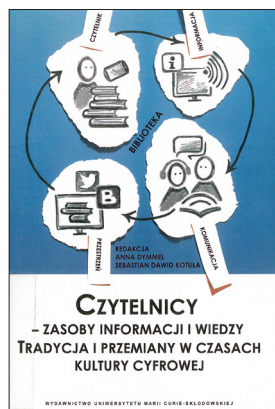
A. Dymmel, S. D. Kotuła, red. (2017). *Czytelnicy – zasoby informacji i wiedzy. Tradycja i przemiany w czasach kultury cyfrowej*. Lublin: Wydaw. Uniwersytetu Marii Curie-Skłodowskiej, ss. 367. ISBN 978-83-7784-997-2

Tematem książki jest wpływ, jaki wywierają na funkcjonowanie współczesnych bibliotek stosowane w nich nowe technologie. Przedstawione zagadnienia ilustrują kierunki zmian zachodzących w bibliotekach zarówno w zakresie porządkowania informacji, jak i wykorzystania osiągnięć rewolucji cyfrowej w narzędziach komunikacji.

Problemy podjęte w publikacji były przedmiotem konferencji „Czytelnicy – zasoby informacji i wiedzy. Tradycja i przemiany w czasach kultury cyfrowej i Internetu”, zorganizowanej przez Instytut Informacji Naukowej i Bibliotekoznawstwa oraz Bibliotekę Główną Uniwersytetu Marii Curie Skłodowskiej w Lublinie, która odbyła się 6–7.10.2016 r.

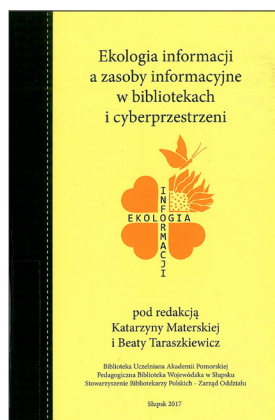
Praca zawiera 23 teksty napisane na podstawie referatów wygłoszonych na konferencji.

Grupując zagadnienia bliskie sobie tematycznie, zawartość tomu podzielono na cztery części: (1) Książka, czytelnik/użytkownik. Badania i rekonesanse badawcze. W tej części znalazły się rozważania teoretyczne poświęcone przemianom form książki i czytelnictwa w epoce cyfrowej, tekst opisujący potrzeby informacyjne studentów Biblioteki Uniwersytetu Łódzkiego na podstawie przeprowadzonych badań. Przedstawiono również dostępność *Kultury* paryskiej w bibliotekach w czasach PRL i zainteresowania czytelnicze lubelskich Żydówek w XIX w. (2) Zasoby informacji i wiedzy – sposoby organizacji. Sposoby organizacji informacji i wiedzy w dobie cyfryzacji zaprezentowano na przykładzie Politechniki Poznańskiej, Centrum Informacji i Dokumentacji Europejskiej w Bibliotece Uniwersytetu Szczecińskiego oraz Centrum Informacji Naukowej i Biblioteki Akademickiej w Katowicach. Znalazły się tu również teksty dotyczące klasyfikacji piśmiennictwa, analizy wybranych stron WWW, serwisu YouTube. (3) Komunikacja naukowa w dobie cyfrowej. Ta część tomu zawiera rozważania nad kompetencjami pracowników informacji, możliwościami porozumiewania się z użytkownikami, ofertę cyfrowej usługi wypożyczeń *Academica*. (4) Czytelnicy, użytkownicy bibliotek i informacji – edukacja, usługi, przestrzeń. W ostatniej części tomu omówiono zagadnienia kompetencji informacyjnych użytkowników, rolę blogów w promowaniu literatury, usługę biblioteczną *Book Club to Go* w Monterey Country, a także „nowe przestrzenie” dla książek (centra handlowe, lotniska) i ich organizację udostępniania zbiorów.



K. Materska, B. Taraszkiewicz, red. (2017). *Ekologia informacji a zasoby informacyjne w bibliotekach i cyberprzestrzeni*. Słupsk: SBP, Zarząd Oddziału w Słupsku; Biblioteka Uczelniana Akademii Pomorskiej, Pedagogiczna Biblioteka Wojewódzka w Słupsku, ss. 248. Seria „Ekologia informacji”. ISBN 978-83-88783-24-1

W znanej serii „Ekologia informacji” ukazała się już szósta obszerna publikacja, zawierająca 17 artykułów dotyczących zagadnień z zakresu ekologii informacji w kontekście zasobów informacyjnych w bibliotekach i cyberprzestrzeni. Teksty pogrupowano w trzech rozdziałach poświęconych antropoinfosferze, ekologii informacji w bibliotekach, cyberprzestrzeni.



W grupie tekstów dotyczących wybranych aspektów antropoinfosfery omówiono zagadnienia smogu informacyjnego, fonoholizmu, zarządzania wiedzą, ochrony praw autorskich, świadomego odbioru dzieł literackich i filmowych.

W rozdziale *Biblioteczne i kulturowe środowisko informacyjne* znalazły się teksty omawiające badania stron internetowych bibliotek, różne typy zasobów bibliotek (prawne, kolekcje dziedzictwa kulturowego).

Ostatnia grupa tekstów poświęcona jest perspektywom i zagrożeniom wynikającym z funkcjonowania człowieka w cyberprzestrzeni. Przedstawiono kontekst społeczno-kulturowy crowdsourcingu internetowego, wartość edukacyjną memów internetowych, cyberstalking i łamanie praw człowieka w cyberprzestrzeni, a także manipulowanie informacją. Tom zamyka tekst poświęcony pojęciom ekologii, ekosystemu, niszy ekologicznej, autekologii dla celów zhumanizowania sztucznej inteligencji komputera.

Artykuły poprzedzono tekstem poświęconym działalności SBP z okazji setnego jubileuszu Stowarzyszenia Bibliotekarzy Polskich.

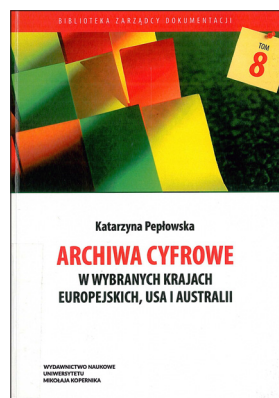
**K. Peplowska (2017).** *Archiwa cyfrowe w wybranych krajach europejskich, USA i Australii*. Toruń: Wydaw. Naukowe Uniwersytetu Mikołaja Kopernika, ss. 193. Seria „Biblioteka Zarządcy Dokumentacji”; t. 8. ISBN 978-83-231-3980-5

Inspiracją do przeprowadzenia badań nad światowymi archiwami cyfrowymi było dla autorki książkę powołanie do życia w 2008 r. polskiego Narodowego Archiwum Cyfrowego, które realizuje zadanie zarządzania zasobem cyfrowym. Przedmiot badań objął sześć archiwów: The National Archives w Wielkiej Brytanii, National Archives w Australii, National Archives and Records Administration w Stanach Zjednoczonych, Schweizerisches Bundesarchiv w Szwajcarii, Rahvusarhiiv w Estonii i Riksarkivet w Norwegii.

Odpowiednio do przeprowadzonych badań w książce w sześciu rozdziałach przedstawiono analizę funkcjonowania poszczególnych archiwów wg określonego klucza. W ramach każdego rozdziału przedstawiono okoliczności powstania archiwum narodowego, podstawę prawną jego działalności oraz strukturę organizacyjną. Omówiono także zasady gromadzenia, przechowywania, opracowania i udostępniania elektronicznego zasobu. Rozdziały poprzedza wstęp, w którym przeprowadzono analizę wieloznacznego terminu „archiwum cyfrowe” (digital archive).

W monografii archiwum cyfrowe jest „utożsamione z systemem teleinformatycznym wdrożonym w strukturze archiwum wieczystego zarządzającym cyfrowym zasobem (do którego należy zaliczyć przede wszystkim dokumentację elektroniczną)” (s. 14). W każdym z wymienionych powyżej archiwów dla nowego typu dokumentacji stworzono nowe komórki organizacyjne w dotychczasowej strukturze. Wdrożono system teleinformatyczny (z wyjątkiem National Archives w Australii), zbudowany na podstawie standardu Open Archival Information System (OAIS).

Książkę uzupełnia obszerna bibliografia: źródła (18 poz.), opracowania (208 poz.).





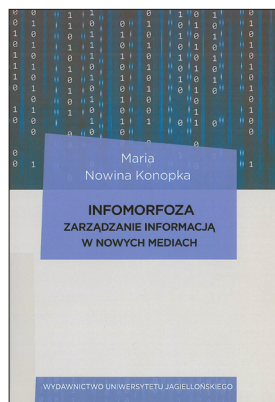
**M. Nowina Konopka (2017).** *Infomorfoza. Zarządzanie informacją w nowych mediach*. Kraków: Wydaw. Uniwersytetu Jagiellońskiego, ss. 309. Seria „Dziennikarstwo, Media i Komunikacja Społeczna”. ISBN 978-83-233-4367-7

Wychodząc od koncepcji teoretycznej mediamorfozy (transformacji środków komunikowania/mediów) Rogera Fidlera, autorka proponuje przyjęcie pojęcia infomorfozy rozumianej jako „proces transformacji mediów implikowany oddziaływaniem szerokiego spektrum zróżnicowanych potrzeb społecznych, politycznych i ekonomicznych dokonujących się w związku ze wzrastającym znaczeniem informacji oraz upowszechnieniem dostępu do niej” (s. 10). Na transformację nowych mediów istotny wpływ mają procesy zarządzania informacją stosowane przez rządy i globalne korporacje, kształtujące nowy geoinformacyjny podział świata. Autorka postawiła szereg pytań badawczych: o zasady funkcjonowania nowych mediów i specyfikę zarządzania informacją w tym środowisku, o realizację potrzeb informacyjnych i sposoby zarządzania informacjami w celu dotarcia do odbiorcy, o sterowanie informacją przez rządy narodowe i wpływ geografii Internetu na stworzenie nowej mapy świata według kryterium dostępu i reprezentacji informacji. Odpowiedzi na te pytania są zawarte w czterech rozdziałach książki.

W pierwszym rozdziale dokonano przeglądu koncepcji nowych mediów i określenia terminów nowe media, scharakteryzowano społeczeństwo sieci. W drugim podjęto próbę zdefiniowania pojęcia informacji, opisano morfologię sieci i zarządzanie informacją. W rozdziale trzecim przedstawiono potrzeby i typy zachowań informacyjnych internautów. W rozdziale czwartym opisano informozę nowych mediów, efekt Snowdena, geografie Internetu, przepływ informacji.

Upowszechnienie dostępu do informacji spowodowało pojawienie się zróżnicowanych potrzeb społecznych, kulturowych, politycznych i ekonomicznych, które wywarły wpływ na charakter infomorfozy. Zdaniem autorki infomorfoza jest efektem wydarzeń politycznych, rozwiązań gospodarczych i zmian społecznych.

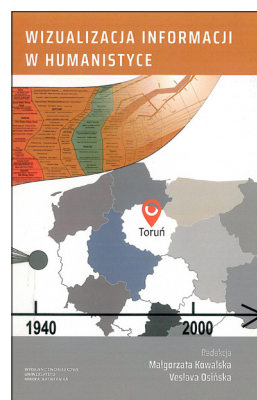
Praca stanowi cenne źródło informacji dla osób interesujących się komunikacją i nowymi mediami, zwłaszcza, że uzupełniono ją bogatą bibliografią (s.283–309).



**M. Kowalska, V. Osińska, red. (2017).** *Wizualizacja informacji w humanistyce*. Toruń: Wydaw. Naukowe Uniwersytetu Mikołaja Kopernika, ss. 326. ISBN 978-83-231-3884-6

Tematem publikacji jest przedstawienie wizualizacji informacji w naukach humanistycznych, a szczególnie wizualizacja wiedzy i wizualizacja w kulturze i sztuce. Oba te zagadnienia poprzedza część pierwsza zatytułowana *Wizualna analiza danych*, obejmująca osiem artykułów, w których autorzy opisują własne rozwiązania przedstawienia różnego typu danych w postaci graficznej, np. *Działalność bibliotek naukowych w Polsce* Barbary Marii Morawiec, pokazuje wizualizację danych statystycznych opracowanych przez GUS dotyczących funkcjonowania bibliotek naukowych w Polsce w 2015 r.

Część druga, zatytułowana *Wizualizacja wiedzy*, składa się z pięciu artykułów omawiających możliwości wizualizacji w prezentacji informacji pochodzących z różnych obszarów wiedzy. W tej części m.in. Małgorzata Janik w artykule *Wizualizacja wizualizacji w procesie dydaktycznym* opisała



wizualizację terminu wizualizacja dokonaną przez studentów kierunku Zarządzanie informacją Instytutu Informatyki i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego.

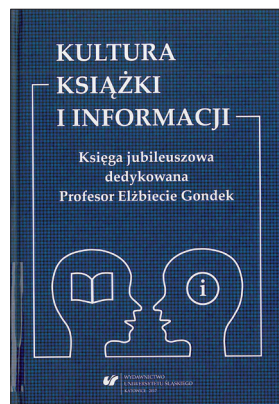
Ostatnia część, trzecia, zatytułowana *Wizualizacja w kulturze i sztuce*, zawiera sześć artykułów poświęconych dużemu oddziaływaniu efektów wizualizacji w dziedzinie sztuki, komunikacji i kultury.

Celem publikacji, jak twierdzą redaktorki tomu, jest „zwrócenie uwagi na różne możliwości, jakie niesie ze sobą wizualizacja informacji oraz wywołanie dyskusji na temat jej potencjalnych zastosowań w różnych obszarach badawczych nauk humanistycznych” (s.14).

**A. Pulikowski, red. (2017).** *Kultura książki i informacji. Księga jubileuszowa dedykowana Profesor Elżbiecie Gondek*. Katowice: Wydaw. Uniwersytetu Śląskiego, ss. 455. Seria „Prace naukowe Uniwersytetu Śląskiego w Katowicach”; nr 3594. ISBN 978-83-226-3095-2

Publikacja przygotowana z okazji jubileuszu pracy naukowej i dydaktycznej Profesor Elżbiety Gondek (Dyrektor Instytutu Bibliotekoznawstwa i Informatyki Uniwersytetu Śląskiego w Katowicach w latach 2005–2012). Zawiera 27 tekstów dotyczących zagadnień z zakresu bibliologii i informatologii. Artykuły pogrupowano w trzech rozdziałach: (1) *Kultura książki*, (2) *Ludzie książki*, (3) *Kultura informacji*. Dwa początkowe teksty w książce poświęcono pracy naukowo-dydaktycznej Profesor Gondek i bibliografii publikacji Pani Profesor za lata 1980–2016.

*Nadestano: 8 listopada 2018.*



## ECIL – „Europejska konferencja edukacji informacyjnej”

(Oulu, Finlandia, 24–27 września 2018 r.)

W dniach 24–27 września 2018 r. w Saint-Malo odbyła się 6. edycja „European Conference on Information Literacy” (ECIL), zorganizowana przez Wydział Nauk o Informacji i Komunikacji Uniwersytetu w Zagrzebiu, Wydział Zarządzania Informacją Uniwersytetu Hacettepe w Ankarze oraz Information Literacy Association (InLitAs), we współpracy z Wydziałem Informacji i Komunikacji Uniwersytetu w Oulu.

W konferencji uczestniczyło 270 osób z 44 krajów. ECIL, to obok angielskiego LILAC (Librarians’ Information Literacy Annual Conference), największe i najważniejsze odbywające się w Europie coroczne spotkanie bibliotekarzy i badaczy zajmujących się kompetencjami informacyjnymi i edukacją informacyjną. Podczas czterodniowych obrad odbyło się osiem warsztatów, wygłoszono 134 referaty (w sesjach równoległych), 11 krótkich prezentacji (tzw. *pecha-kucha*), zaprezentowano 19 posterów oraz trzy prezentacje z badań doktoranckich. Dla doktorantów i innych zainteresowanych zorganizowano także dwie sesje dyskusyjno-warsztatowe z udziałem ekspertów (ang. *senior scholars coaches*) – międzynarodowego grona profesorów.

Wybrane referaty zostały opublikowane w recenzowanym tomie pokonferencyjnym, nakładem wydawnictwa Springer w ramach serii *Communications in Computer and Information Science*<sup>1</sup>.

Referaty w sesjach zostały podzielone na 19 bloków tematycznych:

- Information literacy and education,
- Information literacy good practices,
- Information literacy and lifelong learning,
- Information literacy in different cultures and countries,
- Information literacy and active citizenship,
- Information literacy and digital empowerment,
- Information literacy research and information literacy in theoretical context,
- Information literacy and libraries,
- Information literacy in the workplace and information literacy and employability,
- Information literacy for different groups,
- Information seeking and information behaviour,
- Information literacy instruction,
- Information literacy and health and well-being,
- Information literacy and aspects of education,
- Information literacy and related concepts,
- Information literacy in different contexts,
- Information literacy across disciplines,
- Information literacy research,
- Information literacy in different contexts of everyday life.

Szczegółowy program konferencji oraz abstrakty dostępne są na stronie ECIL 2018, <http://ecil2018.ilconf.org/>.

<sup>1</sup> S. Kurbanoglu, J. Boustany, S. Spiranec, E. Grassian, D. Mizrachi, D., eds. (2019). *Information Literacy in the Everyday Life: 6th European Conference, ECIL 2018, Oulu, Finland, September 24–27, 2018, Revised Selected Papers*. Cham: Springer International Publishing.

Organizatorzy zaprosili jako *keynote speakerów*: Fransa Mäyrä, Kristiinę Kumpulainen, Karen Fisher oraz Petera Batha.

Frans Mäyrä z Uniwersytetu w Tampere wygłosił referat podczas sesji otwierającej. Mówił w nim o doświadczeniach Centre of Excellence in Game Culture Studies – laboratorium badawczego finansowanego przez Fińską Akademię Nauk, które bada rolę, jaką odgrywają interaktywne technologie w życiu społecznym i w komunikacji międzyludzkiej, a także o tym jak, pod wpływem technologii cyfrowych, ludyfikacja i gamifikacja zakorzeniły się w codzienności. Przekonywał o potrzebie kształcenia kompetencji w zakresie umiejętnego korzystania z gier (ang. *gaming literacy*), opartego na krytycznym myśleniu.

Wystąpienie Kristiiny Kumpulainen z Uniwersytetu w Helsinkach dotyczyło promowania wielo-kompetencyjności (*multiliteracy*) w fińskiej edukacji wczesnoszkolnej. „The Joy of Learning Multiliteracies” to program wdrażany przez fińskie Ministerstwo Edukacji i Kultury, którego zadaniem jest wykształcenie u dzieci umiejętności wyszukiwania, interpretowania i wykorzystania różnych form tekstu zarówno w środowisku szkolnym, jak i poza nim. Sprzyjać ma temu nowatorskie podejście pedagogiczne, oparte na współtworzeniu programu nauczania przez uczniów czy aktywnym uczestniczeniu dzieci w życiu lokalnych wspólnot (poza szkołą).

Karen Fischer z Uniwersytetu w Waszyngtonie opowiedziała z kolei o potrzebach informacyjnych i kształtowaniu kompetencji informacyjnych wśród mieszkańców obozów dla uchodźców, na przykładzie swojej pracy w Za’atari – największym na świecie obozie dla uchodźców syryjskich w Jordanii, gdzie przebywa ok. 83 tys. ludzi, z których 80% jest poniżej 25. roku życia. Oboz jest obecnie 4. co do wielkości jordańskim miastem. Nie ma w nim ani jednej księgarni, do niedawna nie działała także żadna biblioteka. Prelegentka opowiedziała o warsztatach organizowanych głównie dla dziewcząt i kobiet (grupy defaworyzowanej) mających wzmocnić ich kompetencje informacyjne, a także o organizowaniu pierwszej biblioteki w obozie.

Peter Bath z Uniwersytetu w Sheffield przedstawił wyniki projektu pn. „A Shared Space and a Space for Sharing”, wpisującego się w szeroką tematykę *health literacy*, a dotyczącego wspierającego wpływu internetowych forów o zdrowiu na zachowania wyszukiwawcze i informacyjne pacjentów i członków ich rodzin. Badania dotyczyły zaufania i empatii podczas dzielenia się informacjami na forach i tego, jak emocje wpływają na dzielenie się informacjami i ich wykorzystywanie. W badaniach wykorzystano analizę tematyczną postów i komentarzy na dwóch forach, a także przeprowadzono wywiady z ich użytkownikami.

Sesje plenarne i równoległe były podporządkowane określonym obszarom badań i praktyki *information literacy*. W sprawozdaniu zostały omówione jedynie wybrane referaty dotyczące dwóch obszarów: kompetencji informacyjnych w miejscu pracy i edukacji informacyjnej w szkolnictwie wyższym.

José Apolinário Teixeira, Helena Karsten i Gunilla Widén z Åbo Akademi University w Finlandii omówili praktyki w zakresie zarządzania wiedzą w *OpenStack* – open source’owej chmurze obliczeniowej dla *big data*, rozwijanej przez społeczność złożoną zarówno z firm, jak i osób prywatnych. Autorzy badań zidentyfikowali zestaw dziesięciu praktyk wykorzystywanych na co dzień podczas rozwijania oprogramowania w trybie pracy zdalnej.

Kolejny reprezentant tego samego uniwersytetu, Farhan Ahmad, poświęcił swój referat zagadnieniu kompetencji informacyjnych w wielojęzycznym miejscu pracy i dzieleniu się wiedzą w kontekście językowym innym niż języka rodzimego. Autor przeprowadził wywiady z pracownikami korporacji międzynarodowych w Finlandii. Przekonywał, że różnorodność językowa, coraz częściej obecnie spotykana w środowisku pracy, wprowadza trudności informacyjne i wymaga kształtowania kompetencji w zakresie interakcji międzyjęzykowej, komunikacji językowej oraz międzykulturowej, a także wypracowania wspólnych schematów dzielenia się wiedzą. Różnice są widoczne pomimo wspólnego języka pracy, jakim jest zazwyczaj angielski, a kompetencje informacyjne wykształcone w języku rodzimym często są niewystarczające do pracy w środowisku międzynarodowym.

Jeszcze inny typ kompetencji informacyjnych w miejscu pracy przedstawiła Joumana Boustany z francuskiego Uniwersytetu Paris-Est Marne-la-Vallée. Omówiła ona wyniki badań dotyczących zarządzania danymi badawczymi wśród kadry naukowej francuskich uczelni. Potwierdziły one

wyniki podobnych badań przeprowadzonych rok wcześniej na wielu uczelniach europejskich (także polskich). Znaczna część badanych nie orientowała się, czym są dane badawcze i jak ważne jest właściwe nimi zarządzanie, zwłaszcza w kontekście wymogów stawianych przez europejskie instytucje finansujące projekty naukowe.

Połączenie badań prowadzonych nad kompetencjami informacyjnymi w miejscu pracy i w szkolnictwie wyższym przedstawiły Anmar Salman i Sirje Virkus z Uniwersytetu w Tallinie. Omówiły wyniki badań dotyczących wpływu zachowań kadry zarządzającej uczelnią na kulturę informacyjną w tym środowisku. Przeanalizowały trzynaście aspektów zachowań decydentów wpływających na efektywność działania poszczególnych jednostek uczelni wyższych w Estonii. W badaniu wykorzystaly metody jakościowe – podejście konstruktywistyczne, wywiady oraz teorię ugruntowaną.

O zaangażowaniu bibliotek akademickich w edukację młodzieży licealnej mówili Anne-Lisa Van der Meulen i Paul Buschmann z Uniwersytetu w Gent. Flamandzkie Ministerstwo Edukacji wymaga od absolwentów liceów m.in. podstawowych umiejętności badawczych, czyli zebrania informacji dotyczących danego zagadnienia, usystematyzowania tych informacji, ich oceny, syntezy oraz przedstawienia wyników i wniosków. Ponieważ są to umiejętności ściśle związane z edukacją informacyjną, biblioteka uniwersytecka w Gent opracowała program wsparcia nauczycieli i uczniów i oferuje w swojej siedzibie warsztaty wprowadzające licealistów w świat informacji naukowej. Uczestnicy warsztatów na wstępie zapoznają się z różnymi źródłami informacyjnymi oraz strategiami wyszukiwawczymi, żeby w dalszej części zajęć przejść do samodzielnego wyszukiwania informacji na zadany temat. Korzyści z warsztatów są obopólne – nie tylko uczniowie zwiększają swoje kompetencje informacyjne, ale także biblioteka i uniwersytet budują swój pozytywny wizerunek i promują się jako miejsce przyjazne użytkownikom, w którym warto studiować.

Im wcześniej zaczną oswajać się przyszłych studentów z zasobami informacyjnymi, tym będą oni bardziej pewnymi siebie, zmotywowanymi użytkownikami informacji, lepiej przygotowanymi do procesu studiowania. Danica Dolničar i Bojana Boh Podgornik z Uniwersytetu w Ljublanie przeprowadziły badanie wśród studentów pierwszego roku. Jego wyniki pokazały, że mimo faktu, iż studenci byli biegli w wykorzystywaniu technik informacyjnych, wysoko samo oceniali swoje umiejętności korzystania z Internetu i motywację do nauki, to jednak ich wyniki dotyczące poziomu kompetencji informacyjnych, efektywności wyszukiwania informacji i dzielenia się wiedzą były niskie. Do badania wykorzystano standardy *information literacy* opracowane przez Association of College & Research Libraries (ACRL). Na podstawie wyników badań przygotowano program z zakresu edukacji informacyjnej dla tej grupy studentów.

O dostosowywaniu oferty edukacyjnej biblioteki do potrzeb studentów kierunków medycznych mówiły także Ann De Meulemeester, Renaat Peleman i Heidi Buisse z Uniwersytetu w Gent. Po kilku latach prowadzenia szkoleń z zakresu edukacji informacyjnej w różnych formatach (szkolenie tylko dla pierwszego roku, szkolenie online) postanowiono wprowadzić kompleksową ofertę obowiązkowych szkoleń stacjonarnych na każdym z trzech lat studiów licencjackich. Pierwszy rok to wprowadzenie, każdy kolejny rozbudowuje zdobytą wcześniej wiedzę. Do ewaluacji postępów studentów (przed i po szkoleniu) wykorzystywana jest pięciostopniowa skala samooceny kompetencji informacyjnych w zakresie medycyny, opracowana przez autorów kursu.

O uwzględnianiu preferencji i przyzwyczajzeń studentów dotyczących korzystania z podręczników akademickich mówiły Almuth Gastinger z Uniwersytetu Naukowo-Technologicznego w Trondheim i Ane Landøy z Uniwersytetu w Bergen. Przedstawiły one wyniki norweskiej części międzynarodowego badania pn. *The Academic Reading Format International Study* (ARFIS). Okazało się, że studenci nadal wybierają podręczniki w wersji drukowanej mimo wielu zalet wersji elektronicznych (dostęp 24/7, oszczędność miejsca, możliwość robienia notatek na tekście, kopiowania, drukowania czy nawet aspektów ekologicznych). Autorki badania zasugerowały, że niskie wykorzystanie e-podręczników może być spowodowane nieświadomością studentów, o ile efektywniej można z nich korzystać niż z formatów drukowanych. Dlatego wystąpiły z postulatem, aby bibliotekarze akademicy byli szkoleni z funkcjonalności e-podręczników tak, aby mogli do ich wykorzystania zachęcać studentów; natomiast

studenci, w ramach szkoleń z zakresu edukacji informacyjnej, powinni odbyć kursy z wykorzystania akademickiej literatury cyfrowej, z naciskiem na jej rozbudowaną funkcjonalność.

O wspieraniu studentów w procesie studiowania przez biblioteki akademickie mówiła także Zuza Wiorogórska z Uniwersytetu Warszawskiego, która omówiła wyniki badań jakościowych (wywiadów) przeprowadzonych wśród studentów wietnamskich studiujących w Warszawie. Podkreśliła, że w kontekście jej badań, różnice kulturowe mogą wpływać na kompetencje informacyjne i wskazała, jakie środki powinny podjąć uniwersytety europejskie, coraz bardziej otwierające się na studentów z Azji, aby pomóc tej grupie studentów pokonać bariery kulturowo-informacyjne i wesprzeć w procesie studiowania. Jednym z takich środków jest dostosowanie oferty edukacji informacyjnej do potrzeb studentów azjatyckich.

Zagadnieniu kompetencji informacyjnych w codziennym życiu poświęcony był referat Zabloni Pingo i Bhuvy Narayana z University of Technology w Sydney, dotyczący kompetencji w zakresie ochrony prywatności (ang. *privacy literacy*) podczas korzystania z technologii społecznych, np. mediów społecznościowych czy kart programów lojalnościowych. *Privacy literacy* to umiejętne korzystanie z technologii, umiejętność oceny ryzyka oraz korzyści, jakie niesie ze sobą dzielenie się prywatnymi informacjami.

Autorka niniejszego sprawozdania wzięła też udział w jednym z ośmiu warsztatów zorganizowanych podczas konferencji. Był to warsztat metodologiczny pt. *What's my approach? Deciding on the approach to use for your research*, moderowany przez Sheilę Webber i Pamelę McKinney z Uniwersytetu w Sheffield. W pierwszej części przedstawiono metody jakościowe przydatne w badaniach z zakresu nauki o informacji, a zwłaszcza edukacji informacyjnej, takie jak: studium przypadku, badanie w działaniu, metody etnograficzne, metody autoetnograficzne czy metody fenomenograficzne oraz scharakteryzowano proces badawczy, uwytklając wagę wyboru właściwej metody do zamierzonego badania wpływ tej metody na cały proces badawczy. W drugiej części warsztatu uczestnicy, podzieleni na grupy, mieli za zadanie opracować pytania badawcze i nakreślić proces badawczy (gromadzenie, analiza, opracowanie i prezentowanie danych), wykorzystując losowo wybraną metodę oraz narzucony scenariusz.

Podczas sesji zamykającej konferencję Albert K. Boekhorst wygłosił mowę ku pamięci zmarłego w grudniu 2017 r. Profesora Foresta Windhama „Woodiego” Hortona Jr., osoby niezwykle zasłużonej dla światowej edukacji informacyjnej, autora książki *Understanding Information Literacy. A primer*, współtwórcy serii warsztatów „Training for trainers in information literacy” organizowanych z ramienia UNESCO w wielu krajach; warsztaty te stanowiły przełom w dyskusji i praktyce *information literacy*. Prof. Horton był też twórcą projektu *Overview of Information Literacy Resources Worldwide* – wydawanego przez UNESCO zestawu publikacji dotyczących dobrych praktyk w edukacji informacyjnej na całym świecie, w 42 językach (w tym polskim).

Ciekawym dodatkiem do sesji konferencyjnych była możliwość zwiedzenia miejskiej biblioteki publicznej w Oulu, a także zaparkowany przed centrum konferencyjnym kolorowy bibliobus, jeden z pięciu należących do biblioteki miejskiej, obsługujący małe szkoły podstawowe, nieposiadające własnej biblioteki oraz osoby starsze i niepełnosprawne. Na pokładzie kolorowego bibliobusa mieści się ok. 2 tys. woluminów książek, materiały audiowizualne, jest kącik do czytania oraz skomputeryzowana wypożyczalnia.

Polskę na ECIL 2018 reprezentowały: Justyna Jasiewicz, Anna Jupowicz-Ginalska, Małgorzata Kisilowska, Anna Mierzecka i Zuza Wiorogórska (UW).

Od tej edycji ECIL będzie organizowany co dwa lata. Kolejna konferencja odbędzie się w terminie 21–24 września 2020 r. w niemieckim Bambergu.

Zuza Wiorogórska

Katedra Informatologii

Wydział Dziennikarstwa, Informacji i Bibliologii

Uniwersytet Warszawski

## Wskazówki dla autorów

Redakcja *Zagadnień Informatyki Naukowej – Studiów Informatycznych* przyjmuje wyłącznie teksty wcześniej nieopublikowane i niezłożone do druku w innych czasopismach lub pracach zbiorowych. Przyjmowane są: oryginalne rozprawy i prace badawcze, recenzje oraz sprawozdania z konferencji i innych wydarzeń naukowych.

Teksty artykułów są recenzowane zgodnie z zasadami *double-blind peer review*. Zapewnienie anonimowości tekstów przekazywanych do recenzji wymaga, aby w tekście artykułu w żadnym miejscu nie była umieszczona informacja umożliwiająca identyfikację autora.

Każdy artykuł recenzowany jest na podstawie jednolitego formularza przez dwóch recenzentów dobranej spośród specjalistów problematyki w nim poruszanej. Każda recenzja zawiera jednoznaczne wskazanie czy tekst rekomendowany jest do publikacji w *Zagadnieniach Informatyki Naukowej*. Podstawowymi kryteriami oceny artykułu są: zgodność tematu z profilem czasopisma, wartość merytoryczna, organizacja logiczna i forma językowa tekstu.

O przyjęciu tekstu do publikacji autorzy informowani są w ciągu 10 tygodni od otrzymania go przez Redakcję. Redakcja przyjmuje wyłącznie teksty przygotowane zgodnie z zasadami przedstawionymi poniżej. Teksty należy nadsyłać na adres e-mail: [zin@uw.edu.pl](mailto:zin@uw.edu.pl)

### 1. Zasady ogólne

#### 1.1. Format

Wszystkie pliki (tekst artykułu, materiały ilustracyjne) należy przysyłać jako dokumenty edytora MS WORD w formacie RTF. Zaleca się stosować w tekście czcionkę Times New Roman 12 pkt, interlinię 1.5. Tytuł artykułu należy wyróżnić czcionką Times New Roman 16 pkt. **Nie należy używać automatycznych stylów.**

Materiały ilustracyjne, wstawione w treść artykułu, dodatkowo należy przysyłać również w formacie JPG. Załączniki powinny być ponumerowane według kolejności występowania w tekście oraz zawierać nazwę, np.: *1. Tab. 1. Poziomy metadanych* albo *3. Rys. 1. Mapa myśli*.

#### 1.2. Długość tekstu

Artykuł nie powinien przekraczać 40 000, a recenzja lub sprawozdanie 14 000 znaków (ze spacjami).

#### 1.3. Strona tytułowa

Autorzy artykułów proszeni są o przygotowanie odrębnej strony tytułowej, zawierającej:

- tytuł artykułu (w językach polskim i angielskim)
- dane autora (imię i nazwisko, afiliacja – w językach polskim i angielskim)
- adres e-mail
- adres do korespondencji
- notę biograficzną autora (patrz niżej)
- abstrakt ustrukturyzowany (patrz niżej)
- słowa kluczowe (patrz niżej)
- oświadczenie o oryginalności tekstu (patrz niżej).

Zgodnie z zasadami przeciwdziałania zjawiskom *ghostwritingu* i *guest authorship* Redakcja prosi również, aby na tej stronie ujawnione zostały nazwiska i afiliacje wszystkich osób, które przyczyniły się do powstania artykułu, ich rola i udział w przygotowaniu publikacji (kto jest autorem koncepcji, założeń, metod itp. wykorzystywanych w pracy zgłoszonej do druku; procentowy udział w przeprowadzonych badaniach i opracowaniu artykułu). Redakcja prosi także o podanie informacji o źródłach finansowania publikacji, wkładzie instytucji naukowo-badawczych, stowarzyszeń i innych podmiotów (*financial disclosure*).

#### 1.4. Nota biograficzna autora / autorów

Na stronie tytułowej należy umieścić zwięzłą notę biograficzną (ok. 70 słów) każdego autora artykułu. Nota powinna zawierać następujące informacje: tytuł / stopień naukowy lub zawodowy autora, aktualne miejsce pracy i zajmowane stanowisko; specjalności naukowe lub zawodowe, najważniejsze publikacje (max. 3). Opisy publikacji powinny być sporządzone zgodnie z zasadami APA Style 6th.

### 1.5. Abstrakt ustrukturyzowany

Na stronie tytułowej należy umieścić abstrakt w języku polskim o objętości ok. 100 słów (ok. 1 tys. znaków) oraz jego przekład na język angielski. W abstrakcie należy wyróżnić co najmniej cztery spośród następujących kategorii informacji:

- Cel/Teza | Purpose/Thesis (*obowiązkowo*)
- Koncepcja/Metody badań | Approach/Methods (*obowiązkowo*)
- Wyniki i wnioski | Results and conclusions (*obowiązkowo*)
- Ograniczenia badań | Research limitations (*opcjonalnie*)
- Zastosowanie praktyczne | Practical implications (*opcjonalnie*)
- Oryginalność/Wartość poznawcza | Originality/Value (*obowiązkowo*)

### 1.6. Słowa kluczowe

Na stronie tytułowej artykułu należy umieścić od 4 do 10 słów kluczowych, w formie fraz nominalnych w mianowniku liczby pojedynczej, których pierwszy wyraz zapisany jest wielką literą, uporządkowanych alfabetycznie, rozdzielonych kropkami. Słowa kluczowe należy podać w językach polskim i angielskim.

### 1.7. Oświadczenie o oryginalności tekstu

Na stronie tytułowej artykułu należy umieścić oświadczenia autora /autorów, że tekst przedstawiany Redakcji *Zagadnień Informatyki Naukowej – Studiów Informacyjnych* nie był dotychczas opublikowany ani zgłoszony do publikacji w żadnym innym czasopiśmie lub pracy zbiorowej. Jeśli tekst był prezentowany na konferencji, należy podać jej szczegółowe dane wraz z ewentualnymi informacjami o publikacji materiałów konferencyjnych. Jeśli artykuł jest częścią przygotowywanej do druku książki, należy podać jej dane oraz planowany termin publikacji.

## 2. Zasady opracowania artykułu

### 2.1. Organizacja i podział tekstu

Tekst artykułu powinien być podzielony na podrozdziały zaopatrzone w tytuły. W pierwszej części pod nagłówkiem **Wprowadzenie** zaleca się umieścić informacje wprowadzające w problematykę prezentowaną w artykule. W części ostatniej – pod nagłówkiem **Wnioski** lub **Zakończenie** – wnioski końcowe i podsumowanie przedstawionych rozważań.

Dopuszcza się stosowanie do trzech poziomów podziału tekstu, każdy wyodrębniony własnym śródtytułem i opatrzonego oznaczeniem numerycznym zgodnie z następującymi regułami:

#### 1. Pierwszy poziom podziału

##### 1.1. Drugi poziom podziału

###### 1.1.1 Trzeci poziom podziału

### 2.2. Przypisy

**Nie stosuje się przypisów bibliograficznych.** Odesłania do wykorzystanej literatury należy przygotować zgodnie z edytorskimi standardami tekstu naukowego APA 6<sup>th</sup> (patrz niżej).

Przypisy zawierające komentarze, dygresje, objaśnienia i inne dodatkowe informacje należy umieszczać na dole strony i numerować liczbami arabskimi; zaleca się ograniczenie liczby przypisów do niezbędnego minimum.

### 2.3. Pisownia tytułów w tekście artykułu

Tytuły wystaw, konferencji, programów itp. powinny być zapisane w cudzysłowie. Tytuły publikacji (książek, czasopism, artykułów itp.) należy wyróżnić kursywą.

### 2.4. Wyróżnienia w tekście

W tekście można stosować wyróżnienia za pomocą czcionki półgrubej (bold).



## 2.5. Materiały ilustracyjne i ich oznaczanie w tekście

Materiały ilustracyjne (tabele, wykresy itp.) powinny być przygotowane w odcieniach szarości lub kolorystyce czarno-białej. Wszystkie tego typu materiały należy oznaczyć wskazaniem rodzaju materiału (np. Tabela, Rysunek, Fotografia, Wykres), jego numeru w tekście oraz jego tytułu (np. Tab. 1. Poziomy metadanych). W odpowiednich miejscach tekstu artykułu należy umieścić odesłania do informacji prezentowanych w formie ilustracji, używając w tym celu skrótu określenia rodzaju ilustracji oraz jej numeru (np. zob. Tab. 1, zob. Wykr. 5).

## 2.6. Cytowanie wykorzystanej literatury w tekście i bibliografia załącznikowa

Cytowania w tekście i bibliografię załącznikową należy przygotować zgodnie ze standardami edytorskim publikacji naukowych APA 6<sup>th</sup>. W bibliografii załącznikowej mogą być umieszczone wyłącznie opisy publikacji cytowanych w tekście artykułu.

Publikacje należy cytować w tekście używając odsyłaczy w formie: (nazwisko, rok wydania), np. (Dembowska, 1991); gdy publikacja ma dwóch autorów należy podać obydwa nazwiska połączone znakiem ampersand (nazwisko1 & nazwisko2, rok), np. (Cisek & Sapa, 2007); gdy publikacja ma trzech i więcej autorów należy podać nazwisko pierwszego autora, skrót *et. al.* i rok wydania (nazwisko1 et al., rok), np. (Berners-Lee et al., 2001); gdy publikacja jest pracą zbiorową, należy podać nazwisko redaktora, skrót red. i rok wydania (nazwisko, red., rok), np. (Kocójowa, red., 2005). Jeśli w publikacji nie wskazano nazwiska autora lub redaktora, należy podać pierwszy wyraz tytułu i rok wydania (Wyraz, rok), np. (Biblioteki, 1976). Odwołania do określonych stron cytowanych tekstów należy podawać w formie: (Dembowska, 1991, 15), albo (Cisek & Sapa, 2007, 40–42), (Dervin & Nilan, 1986, 3) albo (Kocójowa, red., 2005, 18).

Opisy bibliograficzne wykorzystanych publikacji należy umieścić na końcu tekstu w układzie alfabetycznym, bez numeracji pozycji, pod nagłówkiem **Bibliografia**.

Opisy autorskich książek i artykułów umieszcza się pod nazwiskiem pierwszego autora. Opisy prac zbiorowych należy umieszczać pod nazwiskiem redaktora, po którym podaje się skrót *red.* lub *ed.* Jeśli w publikacji nie wskazano autora lub redaktora pracy zbiorowej, jej opis należy umieścić pod pierwszym wyrazem tytułu.

Tytuły książek i czasopism należy zapisać kursywą, tytuły artykułów w czasopismach i artykułów lub rozdziałów w książkach – czcionką prostą.

W opisach artykułów w pracach zbiorowych stosuje się oznaczenie skrótu „W” dla publikacji w języku polskim i „In” dla publikacji w językach obcych.

Opisy prac tego samego autora powinny być uporządkowane według chronologii wstępującej, a w każdym z nich należy powtórzyć nazwisko i inicjał (inicjały) autora. Prace tego samego autora opublikowane w tym samym roku należy uporządkować w kolejności alfabetycznej tytułów i oznaczać wg zasady:

Dembowska, M. (1976a) ...,

Dembowska, M. (1976b) ..., itd.

### 2.6.1 Przykłady redagowania opisów bibliograficznych

#### KSIĄŻKA

Breslin, J.G., Passant, A., Decker, S. (2009). *The Social Semantic Web*. Berlin: Heidelberg: Springer Verlag.

Dembowska, M. (1991). *Nauka o informacji naukowej: organizacja i problematyka badań w Polsce*. Warszawa: IINTE.

#### PRACA ZBIOROWA

Bellardo Hahn, T., Buckland, M., eds. (1998). *Historical Studies in Information Science*. Medford, NJ: Information Today.

Biblioteki (1976). *Biblioteki publiczne województwa toruńskiego: informator*. Toruń: Wojewódzka Biblioteka Publiczna i Książnica Miejska im. M. Kopernika.

Kocójowa, M., red. (2005). *Profesjonalna informacja w Internecie*. Kraków: Wydaw. UJ.

#### ARTYKUŁ W CZASOPIŚMIU

Dervin, B., Nilan, M. (1986). Information Needs. *Annual Review of Information Science and Technology*, 21, 3–31.

Osińska, V. (2010). Rozwój metod mapowania domen naukowych i potencjał analityczny w nim zawarty. *Zagadnienia Informatyki Naukowej*, 96(2), 41–51.

## ARTYKUŁ W PRACY ZBIOROWEJ

- Rayward, W.B. (1998). Visions of Xanadu: Paul Otlet (1868–1944) and Hypertext. In: T. Bellardo Hahn & M. Buckland (eds.). *Historical Studies in Information Science* (65–80). Medford, NJ: Information Today.
- Gawrysiak, P. (2000). W stronę inteligentnych systemów wyszukiwawczych. W: Cz. Daniłowicz (red.) *Multimedialne i sieciowe systemy informacyjne* (59–69). Wrocław: Oficyna PWR.

## ARTYKUŁ W CZASOPISIMIE ELEKTRONICZNYM

- Berners-Lee, T., Hendler, J., Lassila, O. (2001). The Semantic Web. *Scientific American* [online], May, [30.06.2013], <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>
- Hollender, H. (2013). SYNAT: dziesiątki dużych i małych pomysłów na informację naukową. *Biuletyn EBIB* [online], 135(8), [15.07.2013], [http://www.ebib.pl/?page\\_id=413#art6](http://www.ebib.pl/?page_id=413#art6)
- Miller, H. (2013). Big-Data in Cloud Computing: A Taxonomy of Risks. *Information Research* [online], 18(1), [15.07.2013], <http://informationr.net/ir/18-1/paper571.html>

## HASŁA ENCYKLOPEDYCZNE

- Psychology of Culture Contact (1926). *Encyclopaedia Britannica*, Vol. 1, 13th ed. (765–771). London and New York, NY: Encyclopaedia Britannica.
- Iluminatorstwo (1971). *Encyklopedia Wiedzy o Książce* (911–952). Wrocław – Warszawa – Kraków: Zakł. Narod. im. Ossolińskich.
- Big Data (2013, November 12). *Wikipedia, The Free Encyclopedia* [online] [12.11.2013], [http://en.wikipedia.org/w/index.php?title=Big\\_data&oldid=581347727](http://en.wikipedia.org/w/index.php?title=Big_data&oldid=581347727)

Autorskie artykuły encyklopedyczne należy opisywać tak jak artykuły w pracach zbiorowych.

## DOKUMENT Z WITRYNY INSTYTUCJI, ORGANIZACJI LUB OSOBY PRYWATNEJ

- Aristotle (2009). *Organon*. From 1a to 164a according to Bekker numbers [online]. Translated under the editorship of W.D. Ross. Internet archive [29.10.2013], [http://archive.org/stream/AristotleOrganon/AristotleOrganon-collectedWorks\\_djvu.txt](http://archive.org/stream/AristotleOrganon/AristotleOrganon-collectedWorks_djvu.txt)
- MNiSW (2011). *Narodowe Centrum Nauki w Krakowie. Nadchodzi czas nauki* [online]. Ministerstwo Nauki i Szkolnictwa Wyższego, [15.07.2013], <http://www.nauka.gov.pl/?id=2268>
- Smith, B. (2004). *Ontology and Information Systems* [online]. The Buffalo University, Department of Philosophy, [15.07.2013], <http://ontology.buffalo.edu/ontology.doc>
- US NLM (2004). *Medical Subject Headings* [online]. US National Library of Medicine. National Institutes of Health, [15.07.2013], <http://www.nlm.nih.gov/mesh/meshhome.html>

## Guidelines for Authors

ZIN – *Studia Informacyjne* (ZIN – *Information Studies*) accepts only manuscripts that have not been published before and are not under consideration for publication anywhere else. Following types of paper may be submitted for publication: original papers, book reviews, conference (and other events) reports.

Each manuscript is reviewed under a double-blind peer review process. In order to ensure the anonymity of the review process, please do not place any information in the text that could be used to identify the author.

Each manuscript is reviewed by two referees, selected on the basis of necessary expertise in the subject area under review. The review report is based on standard form containing a statement whether the manuscript is recommended for publication. Criteria for acceptance include appropriateness to the field of the Journal, scientific merit, proper text organization and correct language use.

The final decision about publication of manuscript will be sent to Author within 10 weeks after text submission. Manuscript should be formatted according to guidelines listed below and submitted via e-mail: [zin@uw.edu.pl](mailto:zin@uw.edu.pl)

### 1. General guidelines

#### 1.1. Format

All files should be submitted in RTF (Rich Text Format) files, including text and illustrative content. All pages must be typed and 1.5 spaced using 12-point Times New Roman font. The title of the manuscript should be typed 14-point font. Please do not use any preformatted styles.

Illustrative content inserted in the article, should be send also in JPG format. Attachments should be numbered in order of occurrence and include the title, for example: *1. Tab. 1. List...* or *3. Fig. 1. System...*

#### 1.2. Extent

Manuscript should be no longer than 40,000 characters (including spaces), review and report no longer than 14,000 characters.

#### 1.3. Title page

Authors should prepare **separate title page**, which include:

- **title of the paper,**
- **the name(s) of the author(s) with appropriate affiliations,**
- **he e-mail address of the corresponding author,**
- **address for correspondence,**
- **biographic note (see below),**
- **structured abstract (see below),**
- **keywords (see below),**
- **statement of originality (see below).**

According to the Journal policy against *ghostwriting* and *guest authorship*, authors are requested to list on title page names and affiliations of each person that contributed to the text (author of the idea, methods, etc. used in the submitted manuscript; percentage of contribution to the research process and text compilation). Authors are also requested to describe sources of founding that have supported the work and the financial involvement of research institutes, associations and other entities (*financial disclosure*).

#### 1.4. Author(s) biographic note

Title page should include concise biographic notes (about 70 words) of each author : academic degree or professional position, current place of work and position, area of interest, the most important publications (max. 3).

#### 1.5. Structured abstract

An abstract (about 100 words or 1000 characters) should be included with each submission and placed on the title page. Abstract should be formatted according to categories listed below. Author should identify at least four mandatory sections:

- **Purpose/Thesis** (*mandatory*)
- **Approach/Methods** (*mandatory*)
- **Results and conclusions** (*mandatory*)
- **Research limitations** (*optional*)
- **Practical implications** (*optional*)
- **Originality/Value** (*mandatory*)

### **1.6. Keywords**

Title page should include keywords (4 to 10) as a noun phrases in singular form, where first element is capitalized. Keywords in alphabetical order should be delimited by full stop.

### **1.7. Statement of originality**

Author(s) should include on title page statement that submitted text has not been published before and is not under consideration for publication anywhere else. If the paper was presented at a scientific meeting, provide detailed information about the event and the conference proceedings. If the paper will be the part of the author's book, provide its details and planned publishing date.

## **2. Manuscript format and preparation**

### **2.1. Body of the paper**

The text should be organized into entitled sections and subsections. Text should start with **Introduction**, giving an overview and stating the purpose and end with **Conclusion**, giving the summary of the author contributions to the study.

Author may use three levels of headings. Each heading should have its own title and number according to the following pattern:

#### **1. First-level heading**

##### **1.1. Second-level heading**

###### **1.1.1 Third-level heading**

### **2.2. References**

Bibliographic citations are not allowed in footnotes. The reference list should be prepared according to APA 6-th Edition citation style (see below). Footnotes can be used only to give additional information or commentary. Footnotes to the text are numbered consecutively with Arabic numerals. It is recommended to limit the amount of footnotes per page.

### **2.3. Titles in the body of the text**

Titles of exhibitions, conferences, programmes, etc should be written within double quotation marks. Use italics for publication titles (books, journals, papers, etc.).

### **2.4. Emphasis**

Bold face should be used to emphasize certain words or passages.

### **2.5. Illustrative content**

All illustrations (tables, charts, figures etc.) should be converted to greyscale. All illustrations should be cited in the text properly to their form (Table, Figure, Photograph, etc.) and have title and consecutive number (e.g. Tab. 1. Metadata levels). Use abbreviation in the text when refereeing to the illustrative content (e.g. see Tab. 1, see Fig. 5).

### **2.6. Citations and reference list**

Use APA 6-th Edition as a citation and reference list format. The references list should only include works that are cited in the text.

Cite references in the text by name of the author(s) and year of publication in parentheses: (Name, Year of publication), eg. (Dembowska, 1991). If there are two authors, put their names with ampersand (&) mark

between: (Name & Name, Year of publication), eg. (Cisek & Sapa, 2007). If there are more than two authors, put the name of the first one followed by abbreviation *et al.*: (Name et al., Year of publication), eg. (Berners-Lee et al., 2001). Edited books are cited by the name(s) of the editor(s) followed by abbreviation *ed(s)*: (Name, ed., Year of publication), eg. (Bellardo Hahn & Buckland, eds., 1998). If there is no author or editor information, put the first word from the title and the year of publication: (Word, Year of publication), eg. (Biblioteki, 1976). Use the following pattern when referring to specific pages in the cited publications: (Dembowska, 1991, 15) or (Cisek & Sapa, 2007, 40–42) or (Bellardo Hahn & Buckland, eds., 1998, 18).

Place the reference list at the end of the text under the heading **References**. Reference list should be in alphabetical order without numbering.

List the references (books and journal articles) in alphabetical order by authors' last names. Citations of edited books list under the name of editor followed by abbreviation *Ed.*. If there is no author or editor information, list the publication under the first word from the title.

Use italics for book titles and regular font for titles of papers and book chapters. Use abbreviation *In*: when referring to book chapters in citations.

If there are two or more items by the same author(s), list them in order of year of publication (reverse date order). If two or more works are by the same author(s) within the same year, list them in alphabetical order by title and distinguish them by adding the letters a, b, c, ... to the year of publication:

Dembowska, M. (1976a) ....

Dembowska, M. (1976b) ....., etc.

### 2.6.1 References List Examples

#### BOOK

Breslin, J.G., Passant, A., Decker, S. (2009). *The Social Semantic Web*. Berlin: Heidelberg: Springer Verlag.

Dembowska, M. (1991). *Nauka o informacji naukowej: organizacja i problematyka badań w Polsce*. Warszawa: IINTE.

#### BOOK (EDITED)

Bellardo Hahn, T., Buckland, M., eds. (1998). *Historical Studies in Information Science*. Medford, NJ: Information Today.

Biblioteki (1976). *Biblioteki publiczne województwa toruńskiego: informator*. Toruń: Wojewódzka Biblioteka Publiczna i Książnica Miejska im. M. Kopernika.

#### JOURNAL ARTICLE

Osińska, V. (2010). Rozwój metod mapowania domen naukowych i potencjał analityczny w nim zawarty. *Zagadnienia Informatyki Naukowej*, 96(2), 41–51.

Dervin, B., Nilan, M. (1986). Information Needs. *Annual Review of Information Science and Technology*, 21, 3–31.

#### BOOK CHAPTER

Rayward, W.B. (1998). Visions of Xanadu: Paul Otlet (1868-1944) and Hypertext. In: T. Bellardo Hahn & M. Buckland (eds.). *Historical Studies in Information Science* (65–80). Medford, NJ: Information Today.

#### ELECTRONIC JOURNAL ARTICLE

Berners-Lee, T., Hendler, J., Lassila, O. (2001). The Semantic Web. *Scientific American* [online], May, [30.06.2013], <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>

Hollender, H. (2013). SYNAT: dziesiątki dużych i małych pomysłów na informację naukową. *Biuletyn EBIB* [online], 135 (8), [15.07.2013], [http://www.ebib.pl/?page\\_id=413#art6](http://www.ebib.pl/?page_id=413#art6)

Miller, H. (2013). Big-Data in Cloud Computing: A Taxonomy of Risks. *Information Research* [online], 18(1), [15.07.2013], <http://informationr.net/ir/18-1/paper571.html>

#### ARTICLE IN ENCYCLOPEDIA

Psychology of Culture Contact (1926). *Encyclopaedia Britannica*, Vol. 1, 13th ed. (765–771). London and New York, NY: Encyclopaedia Britannica.

Iluminatorstwo (1971). *Encyklopedia Wiedzy o Książce* (911–952). Wrocław – Warszawa – Kraków: Zakł. Narod. im. Ossolińskich.

Big Data (2013, November 12). *Wikipedia, The Free Encyclopedia* [online] [12.11.2013], [http://en.wikipedia.org/w/index.php?title=Big\\_data&oldid=581347727](http://en.wikipedia.org/w/index.php?title=Big_data&oldid=581347727)

Article in encyclopedia with author information describe as book chapter.

ELECTRONIC DOCUMENT FROM WEBSITE

MNiSW (2011). *Narodowe Centrum Nauki w Krakowie. Nadchodzi czas nauki* [online]. Ministerstwo Nauki i Szkolnictwa Wyższego, [15.07.2013], <http://www.nauka.gov.pl/?id=2268>

Smith, B. (2004). *Ontology and Information Systems* [online]. The Buffalo University, Department of Philosophy, [15.07.2013], <http://ontology.buffalo.edu/ontology.doc>

US NLM (2004). *Medical Subject Headings* [online]. US National Library of Medicine. National Institutes of Health, [15.07.2013], <http://www.nlm.nih.gov/mesh/meshhome.html>

---

**Adres Wydawnictwa**

ul. Konopczyńskiego 5/7

00-335 Warszawa, tel. 22 827 52 96

**Prenumerata i sprzedaż**

Dział Promocji i Kolportażu SBP

Al. Niepodległości 213, 02-086 Warszawa, tel. 22 825 50 24

**Cena prenumeraty na 2019 r. – 118 zł**

Wydawnictwo SBP – Warszawa 2018. Nakład 350 egz.

Ark. wyd. 10. Ark. druk. 8,75.

Skład i łamanie: Justyna Grzymała-Łuszcz

Druk i oprawa: Mazowieckie Centrum Poligrafii

ul. Piłsudskiego 2A, 05-270 Marki, [www.c-p.com.pl](http://www.c-p.com.pl)

e-mail: [biuro@c-p.com.pl](mailto:biuro@c-p.com.pl), tel. 22 497 66 55



