

ZIN

Studia informacyjne
Information studies

VOL. 55 2017 NO. 1(109)

p-ISSN 0324-8194

e-ISSN 2392-2648



WDiB UW



STOWARZYSZENIE
BIBLIOTEKARZY
POLSKICH

**ZAGADNIENIA
INFORMACJI
NAUKOWEJ**
Studia informacyjne

**ISSUES IN
INFORMATION
SCIENCE**
Information Studies

VOL. 55 2017 NO. 1(109)
p-ISSN 0324-8194
e-ISSN 2392-2648



WDIiB UW



STOWARZYSZENIE
BIBLIOTEKARZY
POLSKICH



Warszawa 2017

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

Tibor Koltay

Beyond Literacies: The Evolving Landscape of Library Support to Research 2.0 [Poza edukacją kompetencji: rozwój sektora bibliotecznego wsparcia dla badań 2.0] 7

Radosław Lipiński

Information Flows in Transformation and Change Management Supporting Models [Przepływy informacji w modelach zarządzania transformacją i zmianą w organizacjach] 19

Łukasz Opaliński

Bibliometryczna metodologia prognozowania i oceny rozwoju dyscyplin naukowych. Analiza piśmiennictwa. Część 1: Publikacje pionierskie, metoda powiązań bibliograficznych, metoda współcytowań i metoda współwystępowania specjalistycznej terminologii naukowej [Bibliometric Methods to Foresee and Assess the Development of Scientific Disciplines. Literature Analysis. Part 1: Trailblazing Publications, Bibliographic Coupling Method, Co-citation Analysis, and Co-word Methods] 34

Anna Małgorzata Kamińska

ProBIT – prospektywna metoda tworzenia trawersowalnych indeksów cytowań a współczesne problemy organizacji przestrzeni informacji w tradycyjnych bibliograficznych bazach danych [ProBIT – The Prospective Method of Creating the Traversable Citation Index versus Contemporary Issues in Information Space Organization in Traditional Bibliographic Databases] 66

Dorota Siwecka

Historia i współczesność programu CIP (Cataloging in Publication) [The History and Contemporary Issues in Cataloging in Publication (CIP) Program] 83

Piotr Malak, Veslava Osińska, Bożena Bednarek-Michalska

Jakość wybranych elementów metadanych stosowanych w polskich bibliotekach i repozytoriach cyfrowych [Analyzing of Metadata Quality in Polish Digital Libraries] 97

Joanna Ucińska

Humanistyka cyfrowa: problemy definiowania pojęcia [Digital Humanities: Problems in Defining the Concept] 124

Magdalena Zych <i>Przekaz symboliczny i podejście User Experience na przykładzie serwisów internetowych teatrów krakowskich</i> [Symbolic Messages and User Experience Approach Exemplified with the Websites of Cracow Theaters]	146
--	-----

RECENZJE I OMÓWIENIA | REVIEWS

<i>Tibor Koltay, Sonja Špiranec, László Z. Karvalics (2016): Research 2.0 and the future of information literacy.</i> [Badania 2.0 a przyszłość kompetencji informacyjnych]. Amsterdam [etc.]: Elsevier: Chandos Publishing 2016 (Zuzanna Wiorogórska)	159
<i>Wśród zagranicznych książek</i> [Foreign Publications] (Jacek Wojciechowski)	163
<i>Przegląd polskich nowości wydawniczych</i> [New Polish Publications] (Anna Stanis) ..	173

KRONIKA | CHRONICLE

<i>Międzynarodowa Konferencja „Wizualizacja informacji w humanistyce” (Toruń, 23–24 marca 2017)</i> [International Conference “Information Visualisation in Humanities” (Torun, March 23–24, 2017)] (Veslava Osińska, Bożena Bednarek-Michalska)	178
<i>IV Międzynarodowa Konferencja Naukowa „Nauka o informacji w okresie zmian: Innowacyjne usługi informacyjne (Warszawa, 15–16 maja 2017)</i> [Fourth International Scientific Conference “Information Science in the Age of Change: Innovative Information Services” (Warsaw, May 15–16, 2017)] (Piotr Tańkowski)	181
<i>I Ogólnopolska Konferencja „Organizacja i recepcja treści w środowisku cyfrowym” (Katowice, 23 maja 2017)</i> [First National Conference „Content Organization and Reception” (Katowice, May 23, 2017)] (Anna Matysek)	186
<i>IV Konferencja Naukowa „Biblioteka w przestrzeni edukacyjnej”. Książki w plikach – publikowanie, udostępnianie i użytkowanie (Kraków, 24–25 maja 2017)</i> [Fourth Scientific Conference “Libraries in Educational Space”. Books in Files – Publishing, Circulation, and Use] (Agnieszka Folga, Joanna Kołakowska)	189
<i>III Konferencja Naukowa Konsorcjum BazTech. Bibliograficzne bazy danych: perspektywy i problemy rozwoju (Kraków, 26–27 czerwca 2017)</i> [Third Scientific Conference of the BazTech Consortium. Bibliographic Databases: Perspectives and Problems of Development] (Jadwiga Woźniak-Kasparek)	192
WSKAZÓWKI DLA AUTORÓW GUIDELINES FOR AUTHORS	197

Beyond Literacies: The Evolving Landscape of Library Support to Research 2.0¹

Tibor Koltay

*Institute of Learning Technologies
Eszterházy Károly University Jászberény, Hungary*

Abstract

Purpose/Thesis: This paper identifies some of the tasks and roles that academic libraries have to perform in order to react to the emergence of Research 2.0.

Approach/Methods: The argument is based on a non-exhaustive review of the recent literature. Results and conclusions: Academic libraries should respond to the emergence of Research 2.0 by filling niches in services provided by other academic units.

Research Limitations: As a rule, only the literature of the second half of the 2010s was taken into consideration.

Practical implications: The tasks identified in this paper may not seem urgent today, but the likelihood that they will become an imperative in the future is high.

Originality/Value: The issues identified in this paper are already a part of everyday best practices in several countries.

Keywords:

Data citation. Data literacy. Data quality. Research data services. Research management.

Received: 29 April 2017. Reviewed: 19 July 2017. Accepted: 30 August 2017.

1. Introduction

Research 2.0 (Science 2.0, or eScience) is understood as research in the sciences, social sciences and the humanities. It is making use of the Internet's power, which provides the possibility to access and manipulate massive amounts of data. As a result, it brings in a potential to enable different new forms of networking and encourages openness of scientific research, even though many of its aspects remain debated (Pasquetto et al., 2015).

The practical attainment of open science is data sharing that leads to open data, the benefits of which in academia are manifold because it enables data to be reused by other researchers and applied in various other contexts, therefore it fosters transparency in the research process and data collection methods. Nonetheless, one of the most difficult challenges is convincing researchers to accept the idea of data sharing.

¹ This text is the edited version of the paper presented at the 4th International Scientific Conference: Information Science in the Age of Change, Innovative Information Services, in Warsaw, 15th–16th May 2017. Its preparation was supported by the EFOP-3.6.1–16-2016–00001 project “Complex Development of Research Capacities and Services at Eszterházy Károly University”.

The authenticity and trustworthiness of the data is reaffirmed by usage. Data sharing also saves time because the researchers who reuse data do not have to “reinvent” many of the results already attained (Patel, 2016).

Besides the intensive use and openness of research data, Research 2.0 is determined by the growing adoption of social media tools by researchers (Koltay et al., 2015). There is a palpable growth of interest in academic social media sites (social media services, designed for researchers), such as Academia.edu (www.academia.edu), ResearchGate (www.researchgate.net), Mendeley (www.mendeley.com) and Kudos (www.growkudos.com).

The allegedly best known ResearchGate offers services for building, showcasing, and measuring reputation as it provides 10 reputational mechanisms, even though it struggles with the deployment of engagement metrics. In general, it has the potential to become a major deliverer of scholarly reputation, but it did not achieve this yet (Nicholas et al., 2016).

Nonetheless, empirical studies of natural and social science researchers in Britain and the United States show that the reliability and credibility remain central for them and are guaranteed by peer review. As authenticity is ensured by the participants, messages can be transmitted on blogs or social networking sites, but usually they are not cited as publishing articles in peer-reviewed journals (sometimes publishing books) remains the condition of researchers’ advancement (Watkinson et al., 2016).

The examination of 116 early career researchers in seven countries (Britain, China, France, Malaysia, Poland, Spain, and the United States) shows that they have no choice but to adapt to the rules, i.e. to publish in journals that have high prestige. Social media is used mainly to obtain PDF copies of articles, to keep contacts with other researchers and build a digital presence of themselves and of their publications, while active collaboration is not present (CIBER, 2016).

There are several tasks and roles that academic libraries may fulfil in order to react in this environment, last but not least to change the views that see libraries only as dispensaries of books and articles (Jahnke et al., 2012). Some of the possible tasks have been identified by Koltay (2015b; 2016a).

While some teaching staff members and researchers may be reserved against the involvement of librarians, interviews also indicate their willingness to accept data-related library services, provided that they are adequate (McLure et al., 2014).

This paper intends to inform about both theoretical considerations and practical experiences reflected in the literature. As such, it addresses only a selection of issues, thus it does not try to be exhaustive.

Awareness of the core processes of research and the dynamics of change within scholarly communities has to become integral part of researchers’ mid-range strategies. Such awareness should result in the provision of a variety of targeted services within research data services, information literacy and data literacy education, raising awareness of teaching staff members on different issues and providing individual support to them.

In accordance with the above, this paper addresses the following issues and activities:

- Research data services;
- Data literacy education;
- Supporting individual teaching staff members;
- Changing roles for the information professional.

2. Research Data Services

Research data services (RDS) provide an overall service framework that is relevant to the variety of players and related processes of managing and curating research data. RDS should be provided by academic libraries to help researchers to meet the challenges of the data-intensive part of Research 2.0. Some academic libraries are already offering research data services that mainly are extensions of their traditional services, which consist of consulting with teaching staff or students on data management plans and metadata standards; providing reference support for finding and citing data sets; or providing web guides and finding aids for data or data sets. Technical services in the form of providing support for using data repositories, preparing data sets for repositories bear considerable potential, but are provided to a lesser extent than informational ones (Tenopir et al. 2016).

2.1. RDS concepts

Research data management (RDM) is the main constituent of RDS. It is a set of general activities not specifically attached, but potentially performed by the library. RDM means caring for research data, facilitating access to it, preserving and adding value to this data throughout its lifecycle. It is one of the key determinants of open data publication and data curation, bringing the benefits of enabling findability and understandability of data, avoiding unnecessary duplication, and validating results (University of Edinburgh, 2015).

Research data management is not always clearly differentiated from *data curation*. Its goal is enabling data reuse and preventing data loss. Therefore, it is broader than data management, and involves maintaining, preserving and adding value to digital research data throughout its lifecycle and is embedded into the larger meaning of curation as taking care of resources (DCC, 2015; Partlo et al., 2015; Kouper, 2016; Poole, 2016).

Besides data management and data curation, there is also data stewardship that sets somewhat different but not entirely disparate accents by being directed towards maintaining the integrity of and preserving access to data (York et al., 2016).

Even though *data governance* is usually a subject of interest for the business sector, being knowledgeable of its principles and practice is crucial also for the management of research data. Data governance can be defined as the exercise of decision-making and authority that comprises a system of decision rights and accountabilities that is based on agreed-upon models, which describe who can take what actions, when and under what circumstances, using what methods. It enables better decision-making because it concerns rules, policies, standards, decision rights, accountabilities, and methods of enforcement (DGI, 2017). It also deals with availability, access, provenance, meaning, and trustworthiness (ECAR, 2015). These qualities make it to be the starting point for managing data with the distinction that data management involves determining the actual criteria employed for data quality, while data government is about designating who should make these decisions (Khatri & Brown, 2010). Therefore, it is not only one of the necessary conditions for providing open data, but brings substantial knowledge to the work of the academic librarian (Koltay, 2016b).

All above processes require distinguished attention to *data quality* that is determined by trust, authenticity, reliability, understandability, discoverability, accessibility, and integrity (Giarlo, 2013). Besides of these, it is characterised by availability, acceptability, accuracy,

integrity, completeness, and usability (IBM, 2012). Credibility depends – among others – on the coverage of data, authority, format expectations, quality control processes and the opposition between being readable by humans or machines. Examining the quality of research data also should mean considering relevance to the given research question and format that comprises the fields and units used, naming conventions, dates of creation and update (Zilinski & Nelson, 2014). Issues, related to copyright, data licensing, data security, data privacy and to the danger of the erroneous interpretation of data also are tied to quality (Patel, 2016).

2.2. RDS processes and services

Data citation and metadata concern both researchers and librarians. Data citation is a potential source of motivation for researchers to share and publish their data, while they may not have the time and expertise to provide standard metadata (Mooney & Newton, 2012). There are several initiatives and guidelines on providing standardized forms for data citation (e.g. DATACITE, 2016a; ANDS, 2017). A metadata schema is offered, e.g. by DATACITE, (2016b). Having recognized its importance, Clarivate, a commercial information provider offers its *Data Citation Index*.

Supplying metadata for data sets is one of the most important technical RDS activities. Libraries are experienced in this field of activities as they have been providing “traditional” metadata to books and journal articles. By creating metadata to datasets, academic libraries fill in a niche in the support chain. This is especially true because everyone, including funding bodies, recognises that the generation of metadata by researchers is a time and resource-intensive activity, which nonetheless is vital for the curation of research data. According to Chao (2015), in providing metadata, distinguished attention should be given to documenting information about the implemented method.

When students face the difficulties related to exploring the boundaries of their disciplines and doing cross-disciplinary research, they find it extremely hard to know what datasets are available in a given field of study. Established researchers are in a better situation because they are often familiar with at least the main data sources and providers. In both cases, information professionals can gain insight through *data reference interviews* (consultations) that are rooted in traditional librarianship, and may be guided by accepting some of the following guidance, originating both from the traditions of reference work and the needs for data-related information. These factors make the data reference interview process much more complicated than the traditional one:

- Data reference interviews may consist of more questions than a traditional one;
- Making assumptions about the user’s requirements, prior knowledge or viewpoint should be avoided;
- Acronyms and jargon used by the user may require explanation;
- It may be impossible to find the perfect source for the user. Nonetheless, the librarian can probably find some useful starting points (Rice & Southall, 2016).

RDS may include *data management instruction*. In the long term, such education should be integrated into the disciplinary curricula to allow learning within the context of the given fields. Nonetheless, several existing programs are already available, making it unnecessary to create new training programs from scratch (Eaker, 2014). The findings of a survey of

researchers and teaching staff members, working in academic institutions, research organizations, federal agencies and non-profit organizations in the United States, representing hard sciences, social sciences, humanities, and law supported the idea that assistance in data management instruction provided by trained librarians can help science educators, who lack time and expertise (Tenopir et al. 2016).

2.3. *Opinions on RDS*

According to the League of European Research Universities (LERU), supporting research data management is a new role for academic libraries. Nonetheless, they can provide help in increasing the visibility of research data and they are well placed to advocate best practices in data management and data citation. The main reason of involving libraries into these processes is the need to minimize the time that researchers have to spend on technical and administrative processes. According to the LERU report, successful RDM requires cooperation between librarians and researchers, especially in clarifying what is expected of researchers when describing and citing data (LERU, 2013).

The involvement of libraries is a process that has multiple stages and could include some of the actions recommended by the Association of European Research Libraries (LIBER), for example:

- (1) Offering research data management support, including assistance in preparing data management plans;
- (2) Liaising with researchers, data archives and data centres to foster an interoperable infrastructure for data access, discovery and data sharing;
- (3) Providing services for storage, discovery and permanent access;
- (4) Promoting data citation (Christensen-Dalsgaard, 2012).

As indicated also in the above recommendations, there is no need for every institution to implement the same RDS components. Services should depend on several factors, including the size of the institution and the vision of its management. In any case, engaging in strategic planning is an ideal start either to establishing RDS or further developing the existing services (Briney, 2016).

Another type of guidance is offered by Erway et al. (2015) who outline the building blocks, needed for starting data services in libraries that have very little, if anything in place, and have limited resources at their disposal. They also advise on how to build on these foundations as needs become evident and as resources allow.

Besides of this, librarians who want to educate themselves in RDM can consult a selection of relevant resources, available in the webliography, compiled by Goben and Raszewski (2015).

On the other hand, it is necessary to understand that RDM needs to be a researcher-owned endeavour. Having recognised this, the University of Cambridge made efforts to create a community of Data Champions among researchers that provides training and advocacy for good RDM practice (Higman et al., 2017).

3. Data literacy education

Data literacy itself is not without precedents, while providing data literacy education to different groups of (graduate and undergraduate) students, as well as to researchers (including teaching staff members, non-teaching researchers, postdoctoral researchers, and doctoral students) is in many regards new task for academic librarians.

Data literacy can be defined as a specific skill set and knowledge base which empowers individuals to transform data into information and into actionable knowledge by enabling them to access, interpret, critically assess, manage, and ethically use data (Koltay, 2015a). Further definitions are reviewed by Crusoe (2016).

Data literacy education should incorporate both the social and technical aspects of data (Sharma & Qin, 2014). This is demonstrated by the list of abilities that define data literacy, conceived by Calzada, Prado & Marzal (2013), as follows:

- Knowing how to select and synthesize data and combine them with other information sources and prior knowledge;
- Identifying the context in which data is produced and reused;
- Recognizing source data value, types and formats;
- Determining when data is needed;
- Accessing data sources appropriate to the information needed;
- Critically assessing data and their sources;
- Determining and using suitable research methods;
- Handling and analysing data;
- Presenting quantitative information;
- Applying results to learning, decision making or problem-solving;
- Planning, organizing and assessing ourselves throughout the process.

Data quality, as well as data citation and metadata also remain foundational for data literacy education (Ridsdale et al., 2015).

Data literacy is clearly tied to information literacy education that is constantly changing. One of the most evident signs of this transformation is the fact that *Information Literacy Competency Standards for Higher Education*, conceived in the United States of America by the Association of College and Research Libraries (ACRL, 2000) is being supplanted by the *Framework for Information Literacy for Higher Education* (ACRL, 2015), albeit there are already opinions that the Standards should remain as the principal foundation for the understandings of information literacy practice (Jackman, & Weiner, 2017).

As to the possible methods of data literacy education Maybee and Zilinski (2015) introduced *data informed learning* as an approach that – instead of teaching generic data-related skills – focuses on how to use data in disciplinary contexts and is based on the idea of informed learning (Bruce, 2008).

4. Supporting individual researchers

Although supporting individual researchers is different from general RDS activities, this kind of services should be strengthened. As individual's needs are often difficult to articulate, to achieve success in this field, two questions should be asked and answered.

- (1) What is the perception of librarians about their own role in relation to the research activities of teaching staff and what how do they want to achieve?
- (2) What is the perception of teaching staff members about the supporting role of librarians in their research?

The involvement in the support of individual research can be grouped into the following categories:

- Providing informal alerting services;
- Purchasing requested resources;
- Answering in-depth reference questions;
- Creating visual representations of data;
- Co-researching about scholarly publishing in a specific area of knowledge;
- Co-researching the scholarship of teaching;
- Consulting about searching (Brydges & Clarke, 2015).

5. Awareness-raising

There are tasks that academic libraries are expected to address, albeit they are non-traditional, i.e. not offered directly by the libraries themselves, but the librarians are the ones who can make researchers aware of their existence.

Despite problems and controversies around Open Access to scientific publications (OA), it is still important. In particular, there should be ample information about the Diamond, the Gold and the Green routes of OA, not forgetting about issues of article processing charges (APCs). A relatively new and to an extent controversial problem is the presence of *predatory* journals and publishers. Predatory journals' main goal is to seize article processing charges that are paid by authors, who publish their papers in the form of Gold Open Access. APCs themselves are the "protagonists", but the real problem is that predatory journals tend to publish substandard or low-quality papers as they display false identities that mimic established journals, without offering substantial added value. This manifest itself – among others – in the fact that these journals offer unrealistically fast turnaround times of publication, partly because the level of the peer-review process is usually extremely low or it is simply non-existent (Nelson & Huffman, 2015). At the moment, the best tool to sift out predatory journals, is the *Directory of Open Access Journals* (DOAJ, <http://doaj.org>) that is a whitelist of quality open-access publishers and journals.

Alternative metrics (often called *altmetrics*) of scientific output also pertain to Research 2.0 and exhibit great variety (Peters & Bar-Ilan, 2014). Recently, social media sites, first of all ResearchGate, have advanced in showcasing and measuring reputation by using alternative metrics (Nicholas et al., 2016). This is one more reason why librarians have to raise awareness of academic social media sites, even if they have relatively restricted impact on researcher's perception and behaviour.

6. Changing roles for the information professional

Maintaining datasets requires knowledge of both the given domain and advanced technological knowledge. This raises questions whether academic librarians possess the necessary skills and qualifications for this task. Behind this question, there is also a more profound one, if there are conceptual differences between managing data and managing documents. If the answer is yes, librarians have to fulfil new roles, for which they have to be well-prepared (Nielsen, & Hjørland, 2014).

Research 2.0 and especially open data are connected to big data. Approaches to big data are often accompanied by a rhetoric that is both utopian and dystopian, i.e. big data is seen as a panacea to many problems and a tool that enables decreasing privacy and civil freedoms, while the changes that it causes are more subtle shifts (Boyd & Crawford, 2012). However, debates seem to respond more to the popular media's hype about these issues, while practicing data analysts are neither proponents, nor critics of these (Carter & Sholler, 2016).

If there is a relative lack of commonality even in the requirements between the roles of data librarians, data stewards or data curators, the disparity is more sizeable between librarians and data scientists. At the same time, there are intersections between them in data management (Lyon et al., 2015). In addition to this, the roles within this domain are often distributed among professionals because they are different, but still show features of interdependence. Having varied names, data analysts, data journalists or data engineers require high quality, curated data because they work with it. Data librarians show their value by having the data in their care and presenting it to be used and analysed (Lyon & Mattern, 2016). The strongest commonality between these roles and the most powerful connection to data science is the significant attention to data quality and to creativity, intelligence in exploring the nonstandard, personalized and often not automated environment that has to be discovered for finding evidence and indicators for decision-making (Cao, 2016).

On a deeper, conceptual level it seems to be important to heed the warning of Robinson (2016) who reminds us that academics and practitioners should not to be specialised either in data, i.e. examining pattern and syntax and using quantitative methodology, or in information, i.e. interested in meaning, semantics and using qualitative methods because becoming entrenched in being data scientists or information professionals would be harmful for both professions.

7. Conclusion

This paper addressed several issues related to research data services, providing a patchy picture that intended to give insight into several essential and intriguing questions. None of the answers to these is final, but – even if slowly – researchers gradually will have to get accustomed to publishing not only manuscripts, but making research data open (Patel, 2016). Academic libraries may take the lead in this process, or – at least – it is necessary for them to reflect on new developments by filling the niches that appear in services, provided for researchers by other academic units.

References

- ACRL (2000). *Information Literacy Competency Standards for Higher Education*. Chicago, IL.: Association of College and Research Libraries.
- ACRL (2015). *Framework for Information Literacy for Higher Education*. Chicago, IL.: Association of College and Research Libraries.
- ANDS (2017). *Data citation* [online]. Australian National Data Service, [09.04.2017], <http://www.ands.org.au/working-with-data/citation-and-identifiers/data-citation>
- Boyd, D.; Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679.
- Briney, K. (2016). Strategic planning for research data services. *Bulletin of the Association for Information Science and Technology*, 42(4), 39–41.
- Bruce, C. S. (2008). *Informed learning*. Chicago, IL: American Library Association.
- Brydges, B.; Clarke, K. (2015). Is it time to re-envision the role of academic librarians in faculty research? [online] *Library Connect*, 13(7), [09.04.2017] <http://libraryconnect.elsevier.com/articles/2015-07/it-time-re-envision-role-academic-librarians-faculty-research>
- Calzada Prado, J.; Marzal, M. Á. (2013). Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents. *Libri*, 63(2), 123–134.
- Cao, L. (2016). Data science: nature and pitfalls. *IEEE Intelligent Systems*, 31(5), 66–75.
- Chao, T. (2015). Mapping methods metadata for research data. *International Journal of Digital Curation* 10(1), 82–94.
- Carter, D.; Sholler, D. (2016). Data science on the ground: Hype, criticism, and everyday work. *Journal of the Association for Information Science and Technology* 7(10), 2309–2319.
- Christensen-Dalsgaard, B. et al (2012), *Ten recommendations for libraries to get started with research data management*. [online], LIBER, [09.04.2017] <http://www.libereurope.eu/news/ten-recommendations-for-libraries-to-get-started-with-research-data-management>
- CIBER (2016). *Early career researchers: the harbingers of change? Final report from CIBER. Year 1* [online], CIBER-Research.UK, [09.04.2017], http://ciber-research.eu/download/20161120-ECR_Year_1_final_report_071116.pdf
- Crusoe, D. (2016). Data Literacy defined pro populo: To read this article, please provide a little information. *The Journal of Community Informatics*, 12(3), 27–46.
- DATA CITE (2016a). *Why is it so important to cite data?* [online]. DataCite, [09.04.2017], <https://www.datacite.org/cite-your-data.html>
- DATA CITE (2016b). *Metadata Schema 4.0* [online]. DataCite, [09.04.2017], <https://schema.datacite.org/>
- DCC (2015). *What is digital curation?* [online]. London: Digital Curation Centre [09.04.2017] <http://www.dcc.ac.uk/digital-curation/what-digital-curation>
- DGI (2017). *The DGI Data Governance Framework*. [online]. London: Data Governance Institute, [09.04.2017], <http://www.datagovernance.com/dgi-data-governance-framework/>
- ECAR (2015). *The Compelling Case for Data Governance*. [online]. EDUCAUSE ECAR Working Group. [09.04.2017], <http://www.educause.edu/library/resources/compelling-case-data-governance>
- Eaker, C. (2014). Educating researchers for effective data management. *Bulletin of the American Society for Information Science and Technology*, 40(3), 45–46.
- Erway, R; Horton, L.; Nurnberger, A.; Otsuji, R.; Rushing, A. (2015). *Building Blocks: Laying the Foundation for a Research Data Management Program*. Dublin, Ohio: OCLC Research [online], OCLC [09.04.2017], <http://www.oclc.org/content/dam/research/publications/2016/oclc-research-datamanagement-building-blocks-2016.pdf>
- Giarmo, M. (2013). Academic Libraries as Quality Hubs. *Journal of Librarianship and Scholarly Communication*, 1(3), 1–10.

- Goben, A.; Raszewski, R. (2015). Research data management self-education for librarians: a webliography [online]. *Issues in Science and Technology Librarianship* 82, [09.04.2017], <http://www.isrl.org/15-fall/internet2.html>
- Higman, R.; Teperek, M.; Kingsley, D. (2017). Creating a Community of Data Champions. *bioRxiv*, 104661, DOI 10.1101/104661
- IBM (2012). *Successful information governance through high-quality data*. Somers, NY: IBM Corporation
- Jackman, L. W.; Weiner, S. A. (2017). The rescinding of the ACRL 2000 Information Literacy Competency Standards for Higher Education—Really? *College & Undergraduate Libraries*, 24(1), 117–119.
- Jahnke, L.; Asher, A.; Keralis, S. D. (2012). *The problem of data*. Washington, DC: Council on Library and Information Resources.
- Khatri, V.; Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1), 148–152.
- Koltay, T. (2015a). Data literacy: In search of a name and identity. *Journal of Documentation*, 71(2), 401–415.
- Koltay, T. (2015b). Research 2.0, research data services and academic libraries. *Praktyka i Teoria Informacji Naukowej i Technicznej*, 23(4), 3–12.
- Koltay, T. (2016a). Are you ready? Tasks and roles for academic libraries in supporting Research 2.0. *New Library World*, 117(1–2), 94–104.
- Koltay, T. (2016b). Data governance, data literacy and the management of data quality, *IFLA Journal*, 42(4), 303–312.
- Koltay, T.; Špiranec, S.; Z. Karvalics, L. (2015). The shift of information literacy towards research 2.0. *The Journal of Academic Librarianship*, 41(1), 87–93.
- Kouper, I. (2016). Professional participation in digital curation. *Library & Information Science Research*, 38(3), 212–223.
- LERU (2013). *LERU Roadmap for Research Data*. Leuven: League of European Research Universities Research Data Working Group [online]. LERU [29.08.2017], <http://www.leru.org/index.php/public/publications/year/2013/>
- Lyon, L.; Mattern, E. (2016). Education for real-world data science roles (Part 2): A translational approach to curriculum development. *International Journal of Digital Curation*, 11(2), 13–26.
- Lyon, L.; Mattern, E.; Acker, A.; Langmead, A. (2015). Applying translational principles to data science curriculum development [online]. In: *iPres Conference Proceedings*, Chapel Hill, North Carolina, 2–6 November 2015, D-Scholarship, Institutional Repository of the University of Pittsburgh [29.08.2017], <http://d-scholarship.pitt.edu/27159/>
- Maybee, C.; Zilinski, L. (2015). Data informed learning: A next phase data literacy framework for higher education. *Proceedings of the Association for Information Science and Technology*, 52(1), 1–4.
- McLure, M.; Level, A.V.; Cranston, C.L.; Oehlerts, B.; Culbertson, M. (2014). Data curation: A study of researcher practices and needs. *portal: Libraries and the Academy*, 14(2), 139–164.
- Mooney, H.; Newton, M. P. (2012). The anatomy of a data citation: Discovery, reuse, and credit. *Journal of Librarianship and Scholarly Communication*, 1(1), 1–14.
- Nelson, N.; Huffman, J. (2015). Predatory journals in library databases: How much should we worry? *The Serials Librarian* 69(2), 169–192.
- Nicholas, D.; Clark, D.; Herman, E. (2016). ResearchGate: Reputation uncovered. *Learned Publishing*, 29(3), 173–182.
- Nielsen, H. J.; Hjørland, B. (2014). Curating research data: the potential roles of libraries and information professionals. *Journal of Documentation*, 70(2), 221–240.
- Partlo, K., Symons, D., & Carlson, J. D. (2015). Revolutionary or evolutionary? Making research data management manageable. In: B. L. Eden (ed.) *Creating Research Infrastructures in the 21st-Century Academic Library: Conceiving, Funding, and Building New Facilities and Staff*. Lanham (MD): Rowman & Littlefield. 175–201.

- Pasquetto, I. V.; Sands, A. E.; Borgman, C. L. (2015). Exploring openness in data and science: What is "open," to whom, when, and why? *Proceedings of the Association for Information Science and Technology*, 52(1), 1–2.
- Patel, D. (2016). Research data management: a conceptual framework. *Library Review*, 65(4–5), 226–241.
- Peters, I.; Bar-Ilan, J. (2014). Informetrics, bibliometrics, altmetrics: What is it all about? *Proceedings of the American Society for Information Science and Technology*, 51(1), 1–4.
- Poole, A. H. (2017). The conceptual ecology of digital humanities. *Journal of Documentation*, 73(1), 91–122.
- Rice, R.; Southall, J. (2016). *The data librarian's handbook*. London: Facet Publishing.
- Ridsdale, Ch.; Rothwell, J.; Smit, M., Ali-Hassan, H.; Bliemel, M.; Dean, I.; Kelley, D.; Matwin, S.; Wuetherick, B. (2015). *Strategies and best practices for data literacy education: Knowledge synthesis report*. Halifax, NS: Dalhousie University [online], Mike Smit, [09.04.2017], http://www.mikesmit.com/wp-content/papercite-data/pdf/data_literacy.pdf
- Robinson, L. (2016). Editorial. *Alexandria*, 26(2), 73–76.
- Sharma, S.; Qin, J. (2014). Data management: Graduate student's awareness of practices and policies. *Proceedings of the Association for Information Science and Technology*, 51(1), 1–3.
- Tenopir, C.; Allard, S.; Sinha, P.; Pollock, D.; Newman, J.; Dalton, E.; Frame, M.; Baird, L. (2016). Data management education from the perspective of science educators. *International Journal of Digital Curation*, 11(1), 232–251.
- Tenopir, C.; Pollock, D.; Allard, S.; Hughes, D. (2016). Research data services in European and north American libraries: Current offerings and plans for the future. *Proceedings of the Association for Information Science and Technology*, 53(1), 1–6.
- University of Edinburgh (2015). *Why is data management important?* Edinburgh: University of Edinburgh [online], [29.08.2017], <http://www.ed.ac.uk/schools-departments/information-services/research-support/data-management/why-manage-data>
- Watkinson, A.; Nicholas, D.; Thornley, C.; Herman, E.; Jamali, H. R.; Volentine, R.; Allard, S.; Levine, K.; Tenopir, C. (2016). Changes in the digital scholarly environment and issues of trust: An exploratory, qualitative analysis. *Information Processing & Management*, 52(3), 446–458.
- York, J.; Gutmann, M.; Berman, F. (2016), *What Do We Know About The Stewardship Gap?* [online]. Ann Arbor, MI: University of Michigan, [29.08.2017] <http://hdl.handle.net/2027.42/122726>
- Zilinski, L. D.; Nelson, M. S. (2014). Thinking critically about data consumption: Creating the data credibility checklist. *Proceedings of the American Society for Information Science and Technology*, 51(1), 1–4.

Poza edukacją kompetencji: rozwój sektora bibliotecznego wsparcia dla badań 2.0

Abstrakt

Cel/Teza: W artykule wskazano niektóre zadania i role, które biblioteki akademickie muszą spełnić w odpowiedzi na pojawienie się tzw. badań 2.0.

Koncepcja/Metody badań: Argumentacja oparta jest na niewyczerpującym przeglądzie najnowszej literatury.

Wyniki i wnioski: Biblioteki akademickie powinny reagować na pojawienie się badań 2.0, wypełniając nisze w usługach, które temu typowi badań zapewniają inne jednostki akademickie.

Ograniczenia badań: Zasadniczo w przedstawionych rozważaniach wykorzystano jedynie piśmiennictwo, które ukazało się po 2015 r.

Zastosowania praktyczne: Zadania wskazane w niniejszym artykule nie muszą wydawać się pilne dzisiaj, ale prawdopodobieństwo, że w przyszłości staną się imperatywem jest wysokie.

Oryginalność/Wartość poznawcza: Zagadnienia wskazane w niniejszym artykule już dzisiaj są częścią codziennych dobrych praktyk w wielu krajach.

Keywords:

Cytowanie danych. Data literacy. Jakość danych. Usługi dotyczące danych badawczych. Zarządzanie badaniami.

Dr. habil. TIBOR KOLTAY, PhD Professor, Chair of the Institute of Learning Technologies Eszterházy Károly University in Jászberény, Hungary. Since 1981 he worked as an information officer, then as director of different academic libraries. From 1992 to 2004 he was also involved in part-time LIS teaching activities in different institutions. From 2004 he has been working as a full-time LIS teaching staff member at Szent István University and (until 2012) the University of West Hungary. Since 2004 he won eight research scholarships, including a Fulbright Research Scholarship to the School of Library and Information Science of Kent State University, Kent, Ohio. Prof. Koltay also served as ERASMUS guest lecturer at five European universities. After a longstanding engagement with the study of abstracting, he turned towards information literacy and other related literacies, including data literacy, still maintaining his interest in the interdisciplinary relationships between linguistics and information science. His most important recent publications are Abstracts and Abstracting. A genre and set of skills for the twenty-first century. Oxford: Chandos Publishing, in 2010; The media and the literacies: media literacy, information literacy, digital literacy. Media Culture & Society, 33(2), 211–221, in 2011, Data literacy: in search of a name and identity. Journal of Documentation, 71(2), 401–415 in 2015, Are you ready? Tasks and roles for academic libraries in supporting Research 2.0. New Library World, 117(1–2), 94–104, Data governance, data literacy and the management of data quality. IFLA Journal, 42(4), 303–312, Research 2.0 and the Future of Information Literacy (with Sonja Špiranec and László Z. Karvalics) Oxford: Chandos Publishing, Library and information science and the digital humanities: perceived and real strengths and weaknesses. Journal of Documentation, 72(4), 781–792 in 2016 and Data literacy for researchers and data librarians. Journal of Librarianship and Information Science, 49(1), 3–14 in 2017.

Information Flows in Transformation and Change Management Supporting Models

Radosław Lipiński

Abstract

Purpose/Thesis: The purpose of this article is to analyze the use of the information flow perspective in existing, popular models that support transformation and change.

Approach/Methods: The author identified the concept of information flows and the criterion for evaluating transformation models. He analyzed (evaluated) five of most popular contemporary models supporting transformation process management.

Results and conclusions: Using the perspective of information flows in change processes can improve the efficiency of the transformation process. Currently used models support this area in a very limited manner.

Practical implementation: The results of the analysis can be used to build a practical change supporting model based on information flows and method to evaluate transformation models.

Originality/Value: Defining the concept of information flows in the domain of change and transformation. Conducting research on existing transformation models from information process perspective.

Keywords

Information flows. Process. Transformation. Change. Model evaluation.

Received: 21 May 2017. Reviewed: 19 July 2017. Accepted 28 August 2017.

1. Introduction

Change is something that accompanies man perpetually from birth to death, practically in every aspect of life. Governing changes and transformations is a key area for managing organizations. One of the subject's researchers, in his considerations, summarized it in the following words: "It is evident from this article that change is an ever-present element that affects all organizations. There is a clear consensus that the pace of change has never been greater than in the current continuously evolving business environment. Therefore, the successful management of change is a highly required skill" (Todnem, 2005, 378). The above is confirmed by data provided by Hany Wells regarding further development of project management methods, so called Project Management Methodologies:

Since 1999, PMMs are firmly placed as one of the top ten contributing factors towards project failure, according to the Standish Group (2010). In the 2008 version of the CHAOS report, it was stated that although improvements in the rate of project success (35%) are obtained, the rate of failure (19%), and challenged (46%) project performance remain at levels that deserve further attention (Wells, 2012).

Organizations that possess the ability to make effective changes, build stronger and more efficient structures (organizational, business, personal and social), and are capable of using them in an ever-changing environment.

This article attempts to analyze how information processes and its flows affect change processes, and how they are supported in the existing transformation management models. Information processes remain simultaneously the object and the subject of transformation. Having taken into account to what extent change and its effective management affects our reality, including the dimension of the flow of information to the scope of the discussion might prove beneficial to analyzing this topic.

2. Information Flows – definition and criterion

The concept of information flow in contemporary world has been frequently used in different contexts and meanings. However, before we attempt to define it, let us consider the very concept of information. Its unequivocal definition is still subject to numerous discussions. The concept itself has many meanings on various levels. The philosopher, Professor Marek Hetmański, sums it up in an interesting manner:

Information, since it is not a concrete thing, does not take a uniform shape, nor is it homogeneous; it can be recognized in many things. It has the nature of a relational event. In abstract terms, it takes the shape of an opportunity which is realized in specific situations, in the actions of systems of things and processes occurring in particular situations. Somehow it always “operates”, “does” something particular which brings various results and effects. [...] It is present in these aspects of reality, where some difference marks its presence, and manifests itself where concrete action and change take place. It is not present in unchanging continuity and stagnation, although it is potentially rooted everywhere as a form ready to communicate and interact. It has a causative and dynamic nature. It applies equally to the physical and natural world as well as to human and social matters (Hetmański, 2013, 15).¹

For the purpose of this paper, a detailed analysis of this concept will not bring additional value, but it is important to understand how wide topic it is.

Let us return to the discussion on the question of information flows. At times the use of this term refers to the description of the process of communication between people, and sometimes concerns the flow of information in a wider sense. This concept may also be encountered in connection with knowledge flows. It is used in the context of various information processes, including IT, logistics and medicine. It is reflected perfectly in the results of using the Google Scholar search engine. The enquiry of the term “information flow” performed on September 10th, 2016² returned articles and works from the categories mentioned by the author. The query conducted in Polish³ revealed that the understanding of the very concept of information flow as well as the variety of responses were equally high, but the returned categories were different: business communication, organization management, philosophy, logistics, quality management, etc. Undoubtedly, enquires in both languages have shown a very broad understanding of this concept and its applicability in many areas.

Bearing in mind the gravity of this concept for the thesis of these considerations, participants in qualitative research – seven experienced managers and leaders of organizations

¹ Translation from Polish.

² https://scholar.google.pl/scholar?q=information+flow&btnG=&hl=pl&as_sdt=0%2C5&as_vis=1

³ https://scholar.google.pl/scholar?hl=pl&as_sdt=0,5&as_vis=1&q=przep%C5%82yw+informacji

(the qualitative research conducted by the author of this article is described in more detail further) – were asked to provide their definitions of the term information flows. The received answers are listed in the Table 1.

Tab. 1. Definitions of qualitative research respondents regarding the term information flow (own data, based on the results of qualitative research, 2016)

	Definition of the term: information flow
Respondent 1	For me this is an abstract concept, just like the Platonic concept ... shadows on the wall of the cave. The flow of information in an organization reflects the way people communicate with each other.
Respondent 2	Passing knowledge or expectations in organizational structures in different directions, from one place of an organization to another, from one point to another.
Respondent 3	Most frequently with silo mind-set, that is, with top-down information transfer, and in instances when we talk about passing information between two separate silos, going through the top to reach down again. There is also a horizontal information transfer, and it is relatively uncommon in organizations, but more efficient and prompt.
Respondent 4	First, the flow of information in organizations is the flow of information from top to bottom and from bottom to top. We can also talk about the flow of horizontal information, which is not as crucial as well-structured communication – from top to bottom and from bottom to top.
Respondent 5	This is the most important aspect, the basic factor when it comes to the ability to accomplish anything, any given project. It seems to me that among the people involved in project implementation, the truly involved ones, the lack of information flow between them may lead to tremendous difficulties in project realization. I am not saying that it makes it impossible, but quite difficult.
Respondent 6	They are verbal, written, computer, emotional means of communication of information from man to man, from machine to man, from man to machine, or machine to machine.
Respondent 7	This is passing information among members of this organization – information concerning not only the organization, but also important aspects of the performed tasks.

The answers gathered are already more convergent and located within the area of the organization, its structures, its flow directions, and its components. Respondents' definitions allow to give direction to the understanding of the concept of information flows and link them with the subject of the article (change, transformation, information processes). In the context of the above article, we shall define the concept of information flows that will be used for its purposes. Information flow is to be understood as a process in which information (whatever form it shall assume) propagates in an environment and affects the future state of this environment.

The concept of the environment used in the aforementioned definition needs to be understood as a combination of three elements, which are: (1) people together with their personalities (emotions, motivations, beliefs, experiences, etc.), and relationships between

them; (2) environment-specific static structures – for the purposes of this paper, they will be organizational structures such as: teams, groups, subgroups, departments, etc.; (3) defined and undefined processes occurring in the environment – information, business, communication, legal, administrative, IT, etc.

What is of particular vitality in information flows and transformations is the fact that while considering flows in a given change, the process has to be perceived holistically, with the emphasis on the aspects of the process, the environment, the human factors as well as the relationships in which all of the above is taking place, additionally taking into account the final result of the transformation. The abovementioned holism can be compared to the system concept and information flow in a system that is undergoing changes. Since the focus of this work is on organizations and people, the sociological approach seems most appropriate:

The concept of a system entails a complex whole, consisting of many elements interconnected by mutual relations and separated from the external environment by a clear boundary. The apparent examples of systems are living organisms, but also: molecules, buildings, planets, galaxies. Such a general concept can be applied to the analysis of society at various levels of generality (Sztompka, 2005, 20).

2.1. Assessment of the support of information flows by transformation models

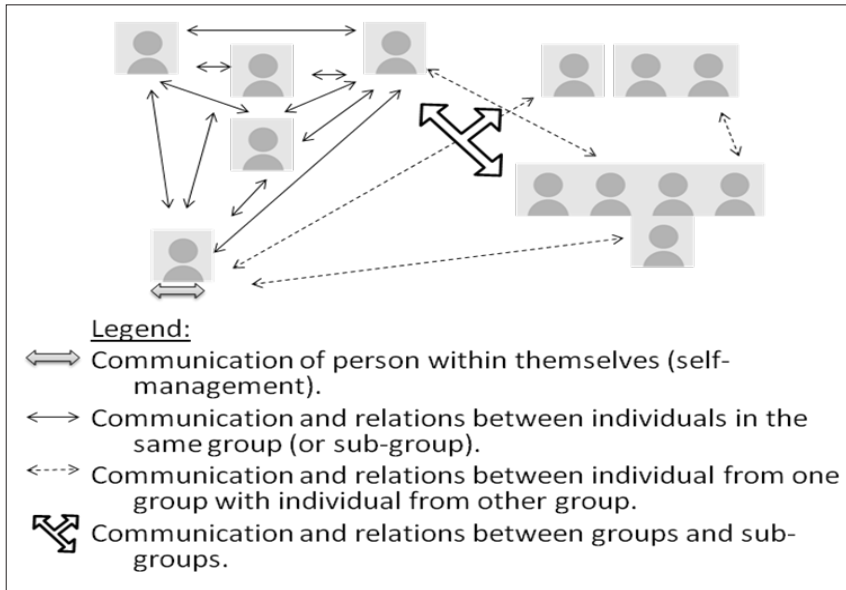
Another question might be posed here of how to evaluate models for managing transformational processes in relation to their support to information flows. It needs to be stressed that the field of change management is characterized by limited maturity of research standards. The research is focused on discovering new concepts, not on building solid foundations and structures that can serve as means of assessing and analyzing already existent knowledge. This proposition was made by Barends and others in 2013 research:

This review indicates that the field's research is limited by the dominance of one-shot studies and seldom addresses the same intervention more than once. As this may be appropriate for a field where theory is nascent and topics have attracted little research, it is unacceptable for a mature field with a research tradition of 50+ years. The field's lack of replication is particularly disturbing. (...) Research activities in OCM seem to be isolated, unrelated, and fail to build on previous studies. Instead, the pursuit of novelty in research, the development of new conceptual frameworks, and the pursuit of new-fangled constructs appear to drive the research agenda. As a result, we know increasingly less about more (Barends et al., 2013, 17).

Therefore, let us propose a model evaluation criterion that can be used to answer the question of how a particular model supports the management of flows during the transformation.

Organizations can be perceived as a set of units and various types of groups (operating in the environment of information processes). Between these elements a continuous flow of information occurs. The basic channels in interpersonal interaction are the existing relationships between individuals. At the group level, its members build a particular group culture, which is related to possessing a common vision of the world and a specific language. The above leads to a situation in which information flows are not just mere relationships between people, but also relationships between groups (at different level). This complexity of interaction is presented in the Figure 1.

Fig. 1. The model of communication flows through relationships within an organization



According to research results from IBM (2008) which involved 1,500 people implementing in-company changes in 21 industries, only 41% of the changes can be considered successful. While conducting this study, it appeared that people and communication remained decisive factors of success. The literature is abundant with instances of works that tie the success of transformation to the efficiency of communication (Kotter, 1994; Mulcahy, 2009; Wagner, 2016). The findings of this study demonstrate how crucial the appropriate management of relationships and communication flows are to the success of performing transformation. There are numerous volumes describing how to increase efficiency in managing communication and influencing others (Cialdini, 2001; Cialdini et al., 2008; Grenny et al., 2013). Including these tools and techniques in a transformational model is very helpful in moderating flows and improves the effectiveness of performance. It needs to be mentioned that at present the factor of relationships and its use for the benefit of the organization is so important for operating in complex structures that it has been introduced into everyday corporate life as one of the official development elements for employees (so-called “networking”). It is often treated equally with other “conservative” skills development, and sometimes regarded as being even more important.

After these initial considerations about information flows and transformations within an organization, let us attempt at the definition of the criterion itself. A transformational model which has inherent structural mechanisms for supporting information flow management (communication and relations) should be the model better assessed by this criterion. Assessments that may be in use for a given criterion are related to the answer to question whether the model supports the management of communication flows. They may take on values: no (1), partial (2), yes (3). A detailed definition of ratings can be found in the Table 2.

Tab. 2. Detailed assessment levels for the information flow criterion

Final assessment result	Definition
yes (3) (best result, expected)	The model in the majority of its components includes support for communication flows. It provides tools and techniques directly, and makes it easy to use tools and techniques existing outside of this model.
partially (2)	The model partially include support for communication flows. It might provide some tools and techniques directly, or make it easy to use tools and techniques that exist outside of this model.
no (1)	The model provides little or none support for communication flows. It does not propose tools or support techniques.

2.2. Validation and verification of the proposed criterion

For the validation of the criterion, an in-depth interview (IDI) method has been selected, commonly referred to as casual interview. It belongs to a set of qualitative measures used in social research:

Qualitative interview is the interaction between the interviewer and the respondent. The interviewer is in possession of a general plan of the research which is not a specific set of questions that need to be asked in particular order with specific words (Babbie, 2013, 242).

This particular method allows the researcher to obtain information and feedback from people who participate in the processes under investigation, or have the knowledge regarding their functioning. In this study, the selection of respondents was of particular importance. Seven persons from different organizations were selected, all with diverse experience, different age, and professional history. What all respondents had in common was their great commitment and experience in the processes of change and transformation. The interviews were conducted in the second quarter of 2016 in Warsaw by the author of the article.

The criterion for evaluating the perspective of information flows received raving reviews from respondents. In additional questions section, virtually all the respondents mentioned how crucial the role of the management of communication flows in the transformation was. The question regarding the competence that influences effectiveness in the current role best got the following answers: communication, understanding people, psychology, ability to cooperate in a group and team. Communication – this is the key to success in dispersed, international organizations where project teams through which projects are implemented, as well as customers are scattered all over the world and in all time zones. Communication is definitely the key. Half of the respondents mentioned the criterion of communication support in response to a question regarding their own favorite indicator of transformational models.

Interestingly enough, several respondents tied the area of information flows with cultural or emotional factors, an exemplary quote:

Communication is key. For me, these two elements [culture, emotions] belong to the component of communication and relationships. You have verbal, non-verbal communication, it evokes certain emotions, and you have to manage that as well. For me, it's still part of communication, the way you communicate, it gives rise to other emotions depending on the culture, the society, the place where one was raised.

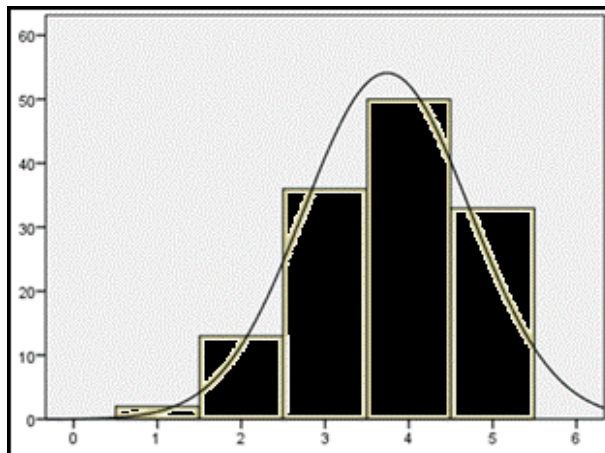
Summarizing, the criterion for information flows was evaluated well during interviews. In the additional questions section, practically all respondents emphasized how crucial the role of the information flow management during transformation was, which was reflected in the quotations. It clearly displays that the respondents confirmed the nature of the influence of the information flow factor in the transformation processes.

Moreover, the accuracy and validity of the use of information flows in evaluating transformation models were confirmed in the quantitative study. The quantitative study is defined as

numerical representation and processing of data to describe the phenomena that this data concerns (Babbie, 2013, 544).

It was completed by the author in August-September 2016 with the aid of the online survey tool⁴. The questionnaires had been sent to about 200 people (excluding those who had participated in qualitative research), who participate in transformational processes (performing managerial and expert roles) in their organizations in Poland. 134 different people completed the questionnaire in full, and only these complete results were taken into account in the final analysis. In this study, the criterion of information flows, on a scale of 1 to 5 (5 is maximal mark), was assessed on an average of 3.74. The median of responses was 4.89% of the respondents gave positive answers to this criterion (values 5 and 4). The Figure 2 presents the histogram of the breakdown of respondents' ratings.

Fig. 2. Histogram of numerical results for the criterion



⁴ The survey was conducted using an online tool <https://www.limeservice.com/en/>

3. Transformation landscape

Transformation and change remain an enormous interdisciplinary field that embraces many issues, concepts and models that support operating in this environment. One can distinguish many possibilities of deconstructing and analysis of this sphere. Despite the large number of materials and papers available in contemporary information resources, it is difficult to pinpoint a single, universally accepted taxonomy for the problem. Due to the sheer scale of the environment of ongoing changes, the mere attempt to create such a description reaches far beyond the scope of this work, and remains a matter for separate considerations and analysis.

However, while analyzing available literature, one may encounter some holistic studies describing change and transformation which propose the division into the following areas: unit, team, organization, leadership (Cameron & Green, 2015, 7). Other studies analyzing transformational models from the perspective of their application in the organization, reveal the division built on management levels, i.e. the area of the strategy of an organization, tactical and, finally, operational (Assen et al., 2014, 1, 81, 193). One may also find studies that propose entirely different divisions (Todnem, 2005).

While defining areas, one must certainly take into account the environment from the perspective of organizations of different types. Having entered company, corporate or non-profit organization field, three main areas of operations can be established:

- complex and well-structured transformation processes – projects, programs, and their portfolios;
- delivery of services, that is execution of processes that continuously transform the organization's resources into defined value for customers and users;
- maintaining and developing the information and knowledge resources of an organization which in contemporary world is, in the majority of cases, reduced to data, systems and IT infrastructure.

These three areas of activity refer directly to the transformational model groups used in organizations, and are consistent with organizational hierarchies, functions, or development paths common in companies. However, this division appears to be considerably technocratic and disregards the enormous area of transformation processes existing in the holistic understanding of the environment – meaning the transformational processes in individuals and groups of people that take place within the organization, as well as outside, in the society. Therefore, the definition of these areas typical for complex organizational structures needs to be extended by two further areas:

- transformation processes that take place at the level of individuals and the environment which can be called human management in change (task);
- transformational processes occurring in groups of people (models focused on human factors).

These two areas coherently expand our perspective and provide an assessment of transformational models in broadly understood (holistic) transformational areas.

3.1. Choice of model representatives in defined areas

Having defined the transformation areas, the representatives of transformation models need to be chosen. Each area will identify one model that can be considered the most popular. Let us assume that popularity, understood as the end-user's desire to use it, is the final measure of the usefulness of transformation model used by users in a given area of the environment. It is presumed that the desire of individuals (or organizations) to spend their time and money to learn about and use the model is a good indicator that the model is useful to them. This is based on the premise "I invest because it brings me a profit".

Another question is how to choose a model among those currently functioning on the market in a particular area. Let us use the method of triangulation for the purpose, defined as the method of utilizing data and research methods from various sources (both quantitative and qualitative), so that the results are analyzed in a broader perspective and reduce measurement errors (Olsen, 2009, 3). The input data for selecting models will come from several sources such as:

- the number of certified professionals active on the market (certificates of the same difficulty level when several of them appear in the models);
- the number of members of the official expert groups associated with the particular transformation models on the professional community website⁵.

The data collected from these sources can be found in the Table 3. The models (candidates) used in the table were selected by the author on the basis of literature research and discussions with transformation practitioners.

Having analyzed and taken into account the results from the sources, the models ranked (Tab. 3) first in their category were selected for further analysis. These models are: PMBOK, ITIL, TOGAF, SLII, Kotter – 8 steps of change management. Next section contains more information about chosen models.

4. Transformation models and their evaluation for information flows support

The main purpose of this section is to evaluate five selected transformation models from the perspective of their support of information flows in the change. Therefore, a very general description of each model is provided. In order to fully understand the analyzed models, the indicated sources need to be consulted.

4.1. PMBOK

PMBOK is the official name of the currently most popular project management model. The main principles of the model are described in the book *Project Management Body of Knowledge* (commonly referred to as PMBOK). The model described in the PMBOK book has been created and is being developed by the Project Management Institute (PMI). The

⁵ Based on www.linkedin.com. This social network for professionals has 450 millions of users (information from July 2016).

PMI proposal is a set of good practices in project management, the use of which increases the likelihood of the success of the performed change (PMI, 2013).

Tab. 3. Data regarding market transformation models in the defined areas

Name of the area	Transformation Model	Number of expert group members at linkedin.com ¹	Number of certified specialists ²
Complex and well-structured transformation processes	Project Management Professional (PMP)	PMI 202 000 PMP 82 000	PMP – approximately 713 000
	Prince 2	OGC – none Prince2 13 500	Prince 2 Practitioner – lack of data
	Scrum	Scrum Alliance – 46 000 Scrum – 75 000	Scrum PSM II and III – approximately 1000
Delivery of services	Information Technology Infrastructure Library (ITIL)	ITIL Service Manag. – 153 000 ITIL – 75 000	ITIL Foundation – over 1 500 000
	Microsoft Operations Framework (MOF)	MOF V4 – 300	Lack of data
	ISO/IEC 20000	ISO 20000 – 4000	Lack of data
	Control Objectives for Information and related Technology (COBIT)	COBIT – 18 000	Lack of data
Maintaining and developing the information and knowledge resources	The Open Group Architecture Framework (TOGAF)	TOGAF – 34 000	Certified – 40 868
	Microsoft Solutions Framework (MSF)	Lack of group	Lack of data
	Federal Enterprise Architecture	FEAF – 500	Lack of data
Transformation processes that take place at the level of individuals	Situational Leadership II (SLII)	Situational Leadership – 3500	Lack of program
	Virginia Satir model	Lack of group	Lack of program
Transformational processes occurring in groups of people	Kotter 8 steps of change	Change Management Practitioners – 8 000	Lack of program
	Lewin, 3-step model (defrost, change, freeze)	Lack of group	Lack of program
	Sange System Model	Lack of group	Lack of program
¹ Data collected in August 2016. ² Data collected in August 2016 from official web pages of discussed transformation models.			

4.2. ITIL

ITIL (Information Technology Infrastructure Library) is a service management model developed by a government agency (Central Communications and Telecommunications Agency) in the United Kingdom. Its first version was published in 1989. The third version of the library is currently in use. This model was built for IT services, but most of the features proposed in this model can be successfully applied to other types of services as well. The ITIL model is not intended solely for managing transformations (it covers a much larger scope). However, the management of service transformation and change is an important part of the ITIL model (one of the five parts of the ITIL library is devoted to this topic) (TSO, 2007).

4.3. TOGAF

TOGAF (The Open Group Architecture Framework) is a model designed to manage the Enterprise Architecture. The TOGAF has been prepared and is being developed by the Open Group consortium. TOGAF introduces methods for: design, planning, implementation and maintenance of architecture. It is a model that describes chiefly the field of information architecture but transformation management remains one of the most important parts of this model. What is of vital importance is that TOGAF considers the transformation from the perspective of the organization's architecture (including information architecture) (TOGAF, 2011).

4.4. *Situational Leadership (SLII)*

Another model chosen for the analysis is Ken Blanchard's situational leadership model (also known as Situational Leadership II). This model is used to manage employees in tasks. However, this model has no limits regarding the size of the tasks which may be different types of changes and transformations. The final version of SLII was presented by Ken Blanchard in the early 1980's. Since then, it has been successfully implemented in many organizations of various types (Blanchard, 2009).

4.5. *Kotter – 8 steps of managing change*

The 8-step change management model was created by John Paul Kotter. It was initially described in the book *Leading Change* (Kotter, 1994). The model was created as a result of the analyses of numerous transformations performed in different organizations. Kotter began with the definition of the most common reasons for the failure of transformation, and, using it as the foundation, crafted a model that counteracts these causes. After the publication, the model gained considerable popularity among managers of organizational change, and was intensively developed at a later stage (Kotter, 2012; Kotter et al., 2006).

4.6. *Assessment of described models with the proposed criterion*

After a brief description of selected models, it is vital to analyze how they support the information processes that take place during the transformation. For this purpose the proposed criterion will be used, accompanied by the defined rating scale.

The PMBOK model does not directly describe the manner of how to manage communication flows and relationships during the transformation. However, some prominent tools and mechanisms existing within this model support the abovementioned areas. The first one is process-based design of the model constructed, and information flows between these processes. Many existing processes indirectly support the flow of information. The above phenomenon is particularly evident in the project initiation phase, and in the artifacts related to the completion of certain processes. An adequate example is a project charter or a communication plan. The second mechanism includes the knowledge areas existing in the model: communication management and stakeholder management. These knowledge areas contain tools that are strongly associated with information flows for communication and relationships, and their use increases the efficiency of the flows. Taking into account the abovementioned elements, the PMBOK model was assessed in the criterion: partially (2).

ITIL has a couple of in-built mechanisms which support the information flows processes. These mechanisms are primarily comprehensive definitions of roles for particular processes and services. Their scope of responsibilities includes communication and building relationships with relevant stakeholders of the process and services. In addition, an essential element in the life cycle of *Service Operation* is communication and establishing relationships with the client (e.g. special request type – request for information, customer response requirement within allotted time, etc). We must note, however, that these elements are related to processes and, to a smaller extent, to information flows, but they do not support the area of interpersonal relationships which is a key factor in information flows; they are only mentioned in the transition lifecycle (dedicated for transformation and change management areas). Taking into account the aforementioned features of this model, we can assume that ITIL describes the area of information flows partially (2).

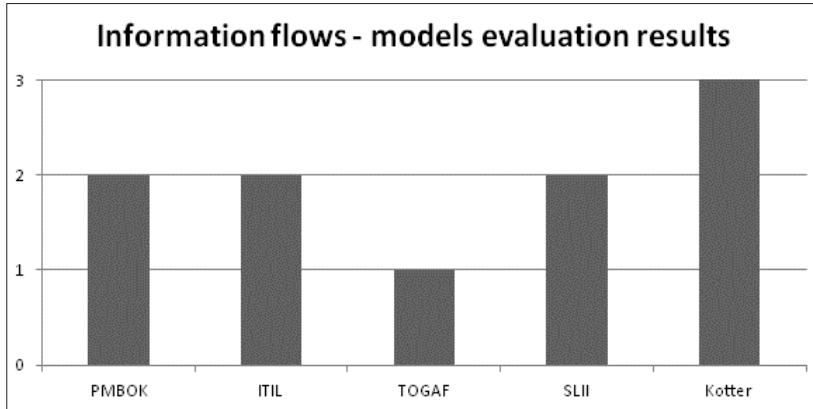
The TOGAF model transformation does not take into account information flows between transformation participants. Some of the tools and techniques specified in the TOGAF documentation relate to relationships and information flows (such as architectural points of view, stakeholders' concerns, etc), but they are not sufficiently significant for the model to support areas relevant to this criterion. Therefore, the TOGAF rating is: no (1).

The SLII model has some elements that support the concept of information flows. One of these is to clearly define the relationship between the superior (the person delegating the task) and the person performing the task. This mechanism provides the adjustment of appropriate communication and management style between two key sides of change. An additional mechanism is to identify supporting roles in a delegated task. By specifying roles, the manager has to establish appropriate communication and relationship-based channels with these people. Unfortunately, the model does not include other tools supporting the flows, but these two elements allow to evaluate the SLII model in this criterion at the level: partially (2).

The Kotter model to a large extent refers to the management of information flows in transformation. From the very beginning it emphasizes how important it is to make the members of the organization aware of the urgency of change through communication flows and the need to work through appropriate relationships. Step two of the model (create a change coalition) boils down to establishing a strong relationship between the key members of the organization to support transformation. The next steps include a variety of elements and tools to support information flows: creating a collective vision for transforming together

through the coalition of change, its effective communication, informing about successes, etc. Practically, each of 8 steps displays references to this perspective. Kotter model by criterion can be considered as fully supporting the flow of information in the change. Therefore, the rating given is: yes (3).

Fig. 4. Results of analysis of all evaluated models by information flows criterion



5. Conclusion

It needs to be stressed that this article is a unique attempt at analyzing the process of change and transformation from the perspective of information processes, and the unquestionable impact these processes have on the final result of each change. The most vital aspect of this paper was the analysis of models supporting transformation processes, and establishing the extent to which they support the information flow processes that occur during change. The final results of the analysis prove that only one of the models visibly supports information flows. This model (Kotter) is a relatively rarely used in organizations (compared to other models analyzed), and covers an area frequently neglected in everyday life of the enterprise. Other popular and widely used models (PMP, ITIL, TOGAF) do not devote sufficient attention to the question of information flows during change. Important output of this article is the definition of the criterion that allows any researcher to measure to what extent transformation model supports information flow processes in change. Another interesting question to be raised is how moderating flows of information affects the efficiency of transformation processes, and whether a model based on such flows can be created.

References

- Assen, M.; Berg, G.; Pietersna, P. (2014). *Key Management Models* (2nd edition). London: Financial Timem Prentice Hall.
- Babbie, E. (2013). *Podstawy badań społecznych*. Warszawa: Wydawnictwo Naukowe PWN.
- Barends, E.; Janssen, B.; ten Have, W.; ten Have, S. (2013). Effects of change interventions: What kind of evidence do we really have? [online] [03.06.2016] *The Journal of Applied Behavioral Science*:

- http://www.tenhavecm.com/wp-content/files_mf/barendsetal.2013effectsofchangeinterventionswhatkindofevidencedowereallyhave.pdf
- Blanchard, K. (2009). *Przywództwo wyższego stopnia*. Warszawa: Wydaw. Naukowe PWN.
- Cameron, E.; Green, M. (2015). *Making Sense of Change Management* (4th edition). London: Kogan Page Limited.
- Cialdini, R. B. (2001). *Wywieranie wpływu na ludzi – Teoria i praktyka*. Gdańsk: Wydawnictwo Psychologiczne.
- Cialdini, R. B.; Goldstein, N. J.; Martin, S. J. (2008). *Yes 50 Scientifically Proven Ways to Be Persuasive*. London: Free Press.
- Grenny, J.; Patterson, K.; Maxfield, D.; McMillan, R.; Switzler, A. (2013). *Influencer The New Science of Leading Change*. New York: Mc Graw Hill Education.
- Hetmański, M. (2013). *Epistemologia informacji*. Kraków: Copernicus Center Press.
- IBM (2008). *Making Change Work* [online]. IBM [24.07.2016], http://www-07.ibm.com/au/pdf/making_change_work.pdf
- Kotter, J. P. (1994). *Leading Change*. London: Harvard Business Review Press.
- Kotter, J. P. (2012). *The Heart of Change: Real-Life Stories of How People Change Their Organizations*. London: Unabridged.
- Kotter, J. P.; Rathgeber, H.; Mueller, P. (2006). *Our Iceberg Is Melting*. New York: St. Martin's Press .
- Mulcahy, R. (2009). *PMP Exam Prep*. New York: RMC Publication Inc.
- Olsen, W. (2009). *Triangulation in Social Research: Qualitative and Quantitative Methods Can Really Be Mixed, Development in Sociology* [online]. Association for Progressive Communications [28.07.2016], <http://research.apc.org/images/5/54/Triangulation.pdf>
- PMI (2013). *A guide to the project management body of knowledge (PMBOK Guide)* (5th edition). Newtown Square: Project Management Institute.
- Sztompka, P. (2005). *Socjologia zmian społecznych*. Kraków: Znak.
- Todnem, R. (2005). Organizational Change Management: A Critical Review. *Journal of Change Management* , 5(4), 369–380.
- TOGAF (2011). *TOGAF an Open Group standard* [online]. The Open Group Architecture [16.04.2014], <http://www.opengroup.org/subjectareas/enterprise/togaf/>
- TSO (2007). *Service Transition*. London: The Strategy Office.
- Wagner, D. N. (2016). Breakin the Project Wave: Understanding and avoiding failure in project management [online]. *PM Word Journal* 5(1) [21.07.2016], <http://pmworldlibrary.net/wp-content/uploads/2016/01/pmwj42-Jan2016-Wagner-breakin-the-project-wave-featured-paper.pdf>
- Wells, H. (2012). How effective are project management methodologies. *Project Management Journal*, 43(6), 43–58.

Przepływy informacji w modelach zarządzania transformacją i zmianą w organizacjach

Abstrakt

Cel/Teza: Celem artykułu jest przeanalizowanie wykorzystania perspektywy przepływów informacji w istniejących, popularnych modelach wspierania procesów transformacyjnych i zmian w organizacjach.

Koncepcja/Metody badań: Autor zdefiniował pojęcie przepływów informacji i wskazał kryterium ewaluacji modeli transformacji. Przeanalizował (podał ocenie) pięć współcześnie najpopularniejszych modeli wspierania procesu zarządzania transformacją.

Wyniki i wnioski: Wykorzystanie perspektywy przepływów informacji w procesach zmian może poprawić efektywność procesów transformacyjnych. Obecnie stosowane modele uwzględniają ten obszar w bardzo ograniczony sposób.

Zastosowanie praktyczne: Wyniki przedstawionej analizy mogą być wykorzystane do budowy praktycznego modelu wspierania zmian w organizacjach, opartego na przepływach informacji oraz metody ewaluacji modeli transformacyjnych.

Oryginalność/Wartość poznawcza: Zdefiniowanie pojęcia przepływów informacji w dziedzinie zmian i transformacji w organizacjach. Przeprowadzenie badania istniejących modeli transformacyjnych z perspektywy procesów informacyjnych.

Słowa kluczowe

Przepływy informacji. Proces. Transformacja. Zmiana. Model ewaluacji.

RADOSŁAW LIPIŃSKI holds currently role of Team Manager in one of the global financial institution. Previously he was working in various managerial roles in one of the leading IT corporation. Alumnus of Warsaw University of Technology Faculty of Electronics and Information Technology (MEng in Computer Science) & Production Engineering (Master's degree in Business Management). Certified coach (CoachWise), practitioner and specialist for transformation and change processes.

*Contact to the author:
lipinski.radoslaw@gmail.com*

Bibliometryczna metodologia prognozowania i oceny rozwoju dyscyplin naukowych.

Analiza piśmiennictwa.

Część 1. Publikacje pionierskie, metoda powiązań bibliograficznych, metoda współcytowań i metoda współwystępowania specjalistycznej terminologii naukowej

Łukasz Opaliński

Biblioteka Politechniki Rzeszowskiej im. I. Łukasiewicza

Abstrakt

Cel/Teza: Celem pracy jest dokonanie krytycznego przeglądu literatury naukowej dotyczącej metod ilościowych stosowanych w obszarze bibliometrycznego prognozowania przyszłego rozwoju dyscyplin naukowych. Przez „przegląd krytyczny” rozumie się wskazanie i omówienie mankamentów prezentowanych w literaturze przedmiotu podejść do rozpatrywanego problemu.

Koncepcja/Metody badań: Przyjęta w pracy metoda polega na analizie literatury przedmiotu, stworzeniu typologii wykorzystywanych w ramach bibliometrii metod prognozowania i oceny rozwoju dyscyplin, podkreśleniu ograniczeń stosowanych w tym zakresie metod oraz wyciągnięciu wniosków końcowych. Postarano się również zaznaczyć obszary, w których przejawia się praktyczna stosowalność prognoz tempa i kierunku rozwoju dyscyplin nauki.

Wyniki i wnioski: Z uwagi na fakt, że niniejszy artykuł stanowi pierwszą z dwóch części, na jakie podzielono całość analizy piśmiennictwa, przedstawione w jego ramach wyniki mają jedynie wstępny charakter. Podkreślono w ich ramach wiodącą i wciąż aktualną rolę metod zapoczątkowanych w publikacjach pionierskich, wskazano, że towarzyszące im mankamenty i próby ich przezwyciężenia stanowią integralną część ewolucji całego omawianego nurtu, a także zaznaczono istnienie silnego związku pomiędzy bibliometryczną metodologią zorientowaną prognostycznie a algorytmicznymi metodami grupowania publikacji w klastry dokumentów spójnych tematycznie, które reprezentują dyscyplinarne obszary problemowe.

Oryginalność/Wartość poznawcza: Oryginalność i wartość poznawcza pracy polega na podjęciu próby przybliżenia rozległości i stopnia zaawansowania badań prowadzonych nad przewidywaniem i oceną stanu rozwoju dyscyplin na świecie. Badania te są realizowane w polskim środowisku naukowym stosunkowo rzadko w związku z czym wydaje się, że przegląd przyjmowanej w ich ramach metodologii mógłby przyczynić się do wzbudzenia większego zainteresowania tą problematyką. Zaakcentowanie słabych stron poszczególnych metod jest w tym kontekście podstawą do dalszego modyfikowania i udoskonalania poszczególnych metod. W obrębie całego spektrum bibliometrii istnieje bowiem jeszcze wiele przestrzeni dla tego rodzaju przedsięwzięć, ponieważ obecnie o żadnej z metod nie można powiedzieć, że idealnie spełnia ona postawione przed nią zadanie i jest „ostatnim słowem” w ramach całego nurtu prognostycznego.

Słowa kluczowe

Bibliometria. Metody ilościowe w informatologii. Prognozowanie. Rozwój dyscyplin naukowych.

Otrzymano: 3 stycznia 2017. Zrecenzowano: 10 lipca 2017. Zaakceptowano: 5 sierpnia 2017.

1. Wprowadzenie

Za twórcę terminu „bibliometria” (ang. *bibliometrics*) powszechnie uznaje się Alana Pritcharda, który zaproponował go w miejsce używanej dotychczas nazwy „bibliografia statystyczna” (ang. *statistical bibliography*) (zob. np.: Diodato, 1994, 14; Nowak, 2006, 16; Skalska-Zlat, 1988, 259). Określenie „bibliografia statystyczna” zostało z kolei użyte po raz pierwszy przez Edwarda Wyndhama Hulme’a w 1922 r. podczas wykładów na University of Cambridge, wydanych rok później w książce pt. *Statistical bibliography in relation to the growth of modern civilization: two lectures delivered in the University of Cambridge in May, 1922* (Hertzfel, 1987, 150; Nowak, 2006, 16; Pindlowa, 1988, 305; Pritchard, 1969, 348). Niedawno Ronald Rousseau zakwestionował jednak pierwszeństwo Pritcharda w tym względzie i stwierdził, że termin „bibliometria” (fr. *bibliométrie*) pojawił się już wcześniej w książce pt. *Traité de Documentation* autorstwa belgijskiego prawnika i bibliografa, twórcy nauki o dokumentacji – Paula Otleta, wydanej w 1934 r. (Rousseau, 2014, 218). Nazwisko Otleta jest pomimo tego współcześnie kojarzone przede wszystkim z Uniwersalną Klasyfikacją Dziesiętną (UKD).

Pritchard zdefiniował bibliometrię jako „(...) zastosowanie metod matematycznych i statystycznych do książek i innych środków komunikacji [piśmienniczej]”¹ (Pritchard, 1969, 348–349). Z punktu widzenia tematyki niniejszego artykułu ważne wydaje się również uzupełnienie definicji bibliometrii Pritcharda o jeden z jej celów, jakim w opinii Pritcharda było „naświetlenie (...) natury i kierunków rozwoju dyscyplin [naukowych] (...) za pomocą zliczania i analizowania różnorodnych aspektów komunikacji piśmienniczej”² (Pritchard, 1969, 348). Tradycyjne zadania stojące przed bibliometrią to więc m. in. ilościowe badania struktury, koncentracji, rozproszenia czy przyrostu literatury naukowej w czasie, jej starzenia się, porównawcze oceny produktywności indywidualnych naukowców lub instytucji badawczych, ewaluacyjna (wartościująca) analiza cytowań lub ustalanie zależności pomiędzy czynnikami biorącymi udział w procesach komunikowania i transferu wiedzy (np. tzw. prawo Bradforda, prawo Lotki, prawo Zipfa). Poza ogółem tych zadań istnieje jeszcze jedna grupa zagadnień należących do przedmiotu bibliometrii, na którą składają się: ocena dynamiki zmian i prognozowanie dalszego rozwoju dyscyplin naukowych oraz wyodrębnianie nowo powstających i nie w pełni jeszcze ukształtowanych (ang. *emerging*) subdyscyplin (obszarów problemowych czy, inaczej mówiąc, wąskich specjalności dyscyplinarnych³)

¹ Oryg.: „(...) the application of mathematics and statistical methods to books and other media of communication” (Pritchard, 1969, 348–349).

² Oryg.: „(...) to shed light on (...) the nature and course of development of a discipline (...), by means of counting and analysing the various facets of written communication” (Pritchard, 1969, 348).

³ W tym miejscu nasuwa się konieczność doprecyzowania wykorzystywanej w artykule terminologii. Zgodnie z *Rozporządzeniem Ministra Nauki i Szkolnictwa Wyższego z dnia 8 sierpnia 2011 r. w sprawie obszarów wiedzy, dziedzin nauki i sztuki oraz dyscyplin naukowych i artystycznych* (Dz.U. 2011, nr 179,

(Skalska-Zlat, 1993, 20; Skalska-Zlat, 1999, 60). Z tą ostatnią funkcją działalności bibliometrycznej ściśle związane jest tzw. mapowanie (ang. *mapping*), czyli odwzorowywanie i wizualizowanie epistemologicznej struktury różnych dyscyplin naukowych, realizowane obecnie często za pomocą wyspecjalizowanych programów komputerowych, takich jak np. *CiteSpace* (zob. np.: Chen, 2006), *Pajek* (zob. np.: Liu & Lu, 2012, 529) lub *BibExcel* (zob. np.: Liu & Gui, 2016, 162). Wizualizacje te określa się jako „topologiczne mapy nauki” (ang. *topological maps of science*) (zob. np.: Garfield, 2004, 119). Analiza takiej mapy, będącej w rzeczywistości siatką cytowań, bazuje na założeniu, że właściwości strukturalne siatki są związane z jej właściwościami dynamicznymi. Można też, alternatywnie, tworzyć mapy w oparciu o publikacje wydane w określonych przedziałach czasowych (ang. *time slices*), uzyskując w ten sposób serię map i obserwować ich transformacje (zob. np.: Diodato, 1994, 37–38; Shibata et al., 2007, 876; Verbeek et al., 2002, 203; Zitt & Bassecoulard, 1994, 334–335). W szczególności badanie wieku zbioru powiązanych publikacji (tj. określenie średniej wieku publikacji wchodzących w skład tego zbioru) jest jednym ze sposobów na wczesne wykrywanie młodych, nie w pełni jeszcze ukształtowanych obszarów problemowych, jako „zaczątków” nowych subdyscyplin. Dodatkowym wskaźnikiem jest tutaj cytawalność takich zbiorów. Im jest ona wyższa, tym szersze zainteresowanie w środowisku naukowym wzbudza taki zbiór i tym bardziej prawdopodobne jest, że w przyszłości zbiór ten będzie się rozrastać i ewoluować w stronę samodzielnej subdyscypliny lub dyscypliny (Shibata et al., 2008, 765; Small, 2006, 596). Niektórzy badacze twierdzą przy tym, że okres największego zainteresowania i najintensywniejszych prac prowadzonych w ramach tego odgałęzienia bibliometrii przypada na ostatnie 15 lat (Ena et al., 2016, 1015; Wang & Ho, 2016, 482), chociaż początków realizacji takich zadań można dopatrywać się już w latach pięćdziesiątych XX w., zaraz po założeniu filadelfijskiego Institute for Scientific Information (Boyack & Klavans, 2014, 671). Obecnie można nawet spotkać się z opinią, że zadanie „mapowania” nauki, wizualizacja dyscyplin naukowych i rozpoznawanie – mówiąc słowami Dereka de Solli Price’a – „frontów badawczych” (ang. *research fronts*) jest naczelnym, pierwszoplanowym komponentem bibliometrii (Huang & Chang, 2015, 2042; Price, 1967, 114.; Wang & Ho, 2016, 482).

poz. 1065) terminem o najszerszym zakresie znaczeniowym jest „obszar wiedzy” (np. obszar nauk przyrodniczych). Terminem węższym jest „dziedzina nauki” (np. dziedzina nauk o Ziemi), a terminem najwęższym jest „dyscyplina” (np. geologia). Terminy takie jak „domena”, „subdyscyplina”, „specjalność” czy „obszar problemowy” nie zostały zdefiniowane w ramach ww. Rozporządzenia, w związku z czym autor niniejszego artykułu uznał, że należy dodatkowo wyjaśnić sens, w jakim używane są one w ramach niniejszego opracowania oraz, że pod warunkiem zachowania konsekwencji w posługiwaniu się nimi dopuszczalna jest pewna dowolność w ich zdefiniowaniu. Z tego względu zdecydowano, że „domena”, „subdyscyplina” i „specjalność” mogą być używane wymiennie i że wszystkie one desygnują struktury węższe niż dyscypliny w rozumieniu ww. Rozporządzenia MNiSW (np. za specjalności w ramach geologii można byłoby uznać paleoklimatologię lub sedymentologię). Określenie „obszar problemowy” charakteryzuje się natomiast jeszcze węższym znaczeniem, które odnosi się do zespołu hipotez lub propozycji teoretycznych związanych ściśle z jakimś jednym problemem rozpatrywanym w ramach specjalności. Jeśli specjalnością byłaby wspomniana paleoklimatologia, jednym z jej obszarów problemowych mogłyby być np. wyjaśnianie tektonicznych przyczyn zmian klimatycznych zachodzących w epoce wczesnego plejstocenu. Jeżeli taki zespół hipotez czy twierdzeń udałoby się potwierdzić doświadczalnie lub zbudować dla nich spójną konstrukcję teoretyczną wyjaśniającą je dostatecznie wiarygodnie i kompleksowo, obszar problemowy mógłby wtedy przekształcić się w nową specjalność (subdyscyplinę czy domenę).

Przewidywanie dalszego rozwoju dyscyplin naukowych i jego tempa oraz kluczowych kierunków tego rozwoju niesie ze sobą rozliczne korzyści praktyczne. Jest ono na tyle istotne, że istnieją nawet towarzystwa naukowe i cykliczne konferencje, które poświęca się wyłącznie tym zagadnieniom (Small et al., 2014, 1450).

Po pierwsze może być ono cenne z punktu widzenia samych naukowców. Śledzenie najnowszych tendencji dyscyplinarnych i identyfikacja najbardziej perspektywicznej, priorytetowej tematyki i pytań badawczych (ang. *hot topics*) pozwoliłyby naukowcom nadażyć za gwałtownym, wielopłaszczyznowym rozwojem dyscyplin i racjonalnie ukierunkować swoje wysiłki badawcze, a także ułatwiłoby im planowanie przyszłej współpracy naukowej. Ponieważ obecnie w erze „zalewu informacyjnego” (zob. np.: Mabe, 2010, 138; Meadows, 2000; Muraszkiewicz, 2014) do tych celów nie wystarcza już bieżące monitorowanie najważniejszych czasopism dyscyplinarnych (np. czasopism przeglądowych), powstaje potrzeba zaprojektowania efektywniejszych, zautomatyzowanych metod wykrywania aktualnie wiodącej problematyki (Wang et al., 2013, 717; zob. też: Takeda & Kajikawa, 2009, 544; Tu & Hsu, 2016, 2016–2017). Przyczyną, dla której uwagę aktywnych w swojej dyscyplinie badaczy mogą przykuwać nowe („młode”) i obiecujące kierunki badań, jest także możliwość zaistnienia w międzynarodowej społeczności naukowej poprzez np. nawiązanie współpracy z cenionymi badaczami i udział w zespołach publikujących wpływowe prace, mające duże szanse na pozyskanie wielu cytowań (Guo et al., 2011, 421). Publikowanie w ramach specjalności rozwojowej niesie bowiem ze sobą zdecydowanie większe perspektywy, niż publikowanie w obszarze, który społeczność naukowa uznała już za wyeksploatowany i mało progresywny, inaczej mówiąc za taki, który okres rozkwitu ma już za sobą i w którym panuje stagnacja (Sun et al., 2016, 750). Duże znaczenie ma też tutaj szybkość, z jaką poszczególni naukowcy będą w stanie pozyskać, przyswoić i wykorzystać wiedzę naukową rozwijającą się w młodej i dopiero formującej się specjalności. Im szybciej będą w stanie to osiągnąć i im bardziej zdystansują potencjalnych rywali, tym bardziej podniosą swoją konkurencyjność (ang. *competitive advantage*) i zyskają prestiż związany potencjalnie z posiadaniem statusu jednego z współtwórców czy pionierów nowo powstałej domeny (Shibata et al., 2008, 758).

W omawianym kontekście praktycznego znaczenia prognoz rozwoju dyscyplin naukowych wydaje się ponadto, że warto dodatkowo zaznaczyć, iż analizy tego rodzaju mogą okazać się nie bez znaczenia także dla bibliotekarzy, a zwłaszcza dla bibliotekarzy dziedzinowych. Hanning Guo i in. (2011) wskazali bowiem, że także i ta grupa zawodowa może odnieść korzyść z posługiwania się narzędziami ułatwiającymi tworzenie nowych kategorii lub haseł przedmiotowych w trakcie rzeczowego opracowywania zbiorów, uzupełniania istniejących języków informacyjno-wyszukiwawczych (np. tezaurusów) o nowe terminy albo wyodrębnianie specjalnych kolekcji zbiorów bibliotecznych, ukierunkowanych na potrzeby informacyjne badaczy prowadzących studia o profilu zgodnym z nowo powstającymi subdyscyplinami (Guo et al., 2011, 422).

Następnym obszarem, w którym zauważalne są praktyczne korzyści płynące z bibliometrycznego prognozowania przyszłego rozwoju dyscyplin naukowych, jest ogólnie rozumiana polityka naukowa, w skład której wchodzi administrowanie i zarządzanie działalnością naukową oraz kwestie finansowania naukowych przedsięwzięć z budżetu państwa (zob. np.: Braun et al., 2000, 24). Ponieważ badania naukowe są „zaczynem” (ang. *seeds*) innowacyjności przemysłowej, znaczenie prac badawczo-rozwojowych (ang. *R&D – Research and*

Development) jest uznawane za fundamentalne w procesie stymulowania innowacyjności i konkurencyjności narodowego przemysłu i technologii (Shibata et al., 2008, 758–759). W warunkach polskich prace badawczo-rozwojowe obejmują programy rządowe dotyczące naukowo-technologicznego rozwoju państwa, badania marketingowe i strategie rozwoju przedsiębiorstw. Ich celem w tej perspektywie jest proces doskonalenia i usprawniania techniki i technologii oraz kreowanie zrównoważonej polityki w zakresie rozwoju technologicznego (zob. np.: KPB, 2011; MNiI, 2004). Istotną kwestią jest tutaj również racjonalna alokacja funduszy na cele naukowe, która byłaby zgodna z badawczymi priorytetami, a także możliwościami danego kraju lub instytucji (zob. np.: Huang & Chang, 2015, 2041–2042; Kajikawa et al., 2008, 772–773; Lee et al., 2012, 814–815; Tseng et al., 2009, 73–74). Finansowanie badań naukowych w racjonalny sposób przyczyniłoby się do dokonywania odkryć, które mogą znaleźć zastosowania komercyjne (np. w przemyśle) i zaowocować dalszymi odkryciami o naturze interdyscyplinarnej. Interdyscyplinarność jest natomiast postrzegana jako „siła napędowa” innowacyjności (Upham & Small, 2010, 15–17; zob. też: Sun et al., 2016, 751).

Jeszcze jedną sposobnością praktycznego wykorzystania wiedzy o rozwoju dyscyplin naukowych jest zastosowanie zbliżone do omówionych wyżej, ale usytuowane w sektorze prywatnym. Zarządzanie technologiami pozwala bowiem na poprawę efektywności wykorzystania technologicznego i finansowego kapitału, planowanie krótko- lub długoterminowych strategii produkcyjno-rynkowych, jak również przewidywanie cykli życiowych produktów technologicznych i racjonalizację dalszych inwestycji (Chen et al., 2012, 1099). Ponieważ wszelkie strategie inwestycyjne są nieodłącznie związane z ryzykiem, np. co do właściwego ulokowania środków finansowych czy długości cyklu życia produktu (czy popyt na taki produkt będzie trwał wystarczająco długo aby zwróciły się koszty jego produkcji, promocji i rozprowadzania?), pożądanym byłby dostęp do metod pozwalających na zredukowanie tego ryzyka i wskazanie takich obszarów, w których inwestowanie jest potencjalnie najbardziej opłacalne (Érdi et al., 2013, 226; Lee, 2008, 504).

Niejako przy okazji identyfikowania najważniejszych ścieżek historycznego rozwoju dyscyplin (złożonych ze szczególnie wpływowych publikacji) oraz budowania prognoz, pojawia się możliwość przewidzenia tego, które z artykułów mogą stać się w przyszłości tzw. publikacjami kluczowymi, centralnymi dla pewnych subdyscyplin (ang. *core papers*). Na tej samej podstawie możliwe stałoby się również konstruowanie zautomatyzowanych systemów rekomendujących publikacje relewantne do pewnej tematyki (ang. *recommendation systems*). Mogłyby być one pomocne tak dla samych autorów, jak i dla recenzentów oceniających bibliografie załącznikowe publikacji pod kątem adekwatności przywoływanych w nich pozycji (Shibata et al., 2012, 78–79).

2. Cel pracy

Celem pracy jest dokonanie krytycznego przeglądu literatury naukowej, dotyczącej metod ilościowych stosowanych w obszarze bibliometrycznego prognozowania dalszego rozwoju dyscyplin naukowych. W szczególności istotne wydaje się w tym kontekście wskazanie mankamentów poszczególnych podejść wykorzystywanych przez autorów zajmujących się tego rodzaju zagadnieniami. Mankamenty te sprawiają, że wśród wskazanych i omówionych

podejść pozostaje wciąż wiele miejsca na dalsze udoskonalenia, modyfikacje i eksperymenty związane z projektowaniem nowych metod służących realizacji celów prognostycznych.

Literatura przedmiotu dotycząca wyodrębniania nowo powstających specjalności w ramach dobrze ugruntowanych teoretycznie dyscyplin oraz prognozowania ich dalszego rozwoju jest nie tylko bardzo obszerna, ale też wysoce zróżnicowana pod względem stosowanych w tym zakresie metod i samego przedmiotu badań. Przedmiotem tym mogą być bowiem zarówno artykuły naukowe, patenty czy konferencje, jak i autorzy publikacji naukowych i sieć współpracy między nimi. Z uwagi na to autor niniejszej pracy zdecydował, że najkorzystniejsze z punktu widzenia czytelnika oraz z punktu widzenia przejrzystości przeglądu literatury, będzie podzielenie go na podgrupy publikacji pod pewnymi względami podobnych. Podział ten można określić mianem „typologii”, ponieważ nie jest on tak rygorystyczny, jak wymaga tego klasyfikacja, kategoryzacja czy systematyka. Typologia oznacza bowiem podzielenie elementów pewnego zbioru na podzbiory elementów pod pewnymi względami do siebie podobnych. Elementy te nie muszą być więc identyczne, ale musi łączyć je pewien, na ogół różny, stopień podobieństwa. Ponadto wyróżnione podzbiory (typy) nie muszą być rozłączne, ponieważ niektóre elementy mogą spełniać warunek podobieństwa w stosunku do więcej niż jednego elementu wzorcowego (Sosińska-Kalata, 2002, 19). Cechą typologii jest ponadto tzw. infinityzm, czyli akceptacja nieskończonej różnorodności poddawanych podziałowi elementów i dopuszczalność abstrahowania ich cech wspólnych bez konieczności definiowania ściśle jednolitego kryterium wyboru (Sosińska-Kalata, 2002, 26).

Wydzielono w sumie 11 typów publikacji mających pewne cechy wspólne. Należy przy tym wyjaśnić, że z uwagi na wspomnianą rozległość literatury przedmiotu niniejszy artykuł stanowi pierwszą z dwóch części, na jakie podzielono całość analizy piśmiennictwa. Każdy z kolejnych podrozdziałów (od 3 do 7) omawia jedną z metod, które w przeciwieństwie do metod demonstrowanych w części drugiej można określić mianem tradycyjnych (klasycznych, konwencjonalnych, podstawowych). Są nimi po pierwsze, metody wypracowane w publikacjach pionierskich, które zapoczątkowały jakąś technikę badań lub po raz pierwszy postawiły problem, który był następnie rozwijany, modyfikowany lub szeroko komentowany przez kolejnych badaczy. W dalszej kolejności omówiono bardziej szczegółowo każdą z tych metod, tj. wskazano w jaki sposób ewoluowały, jakie zastosowania zostały dla nich odnalezione oraz jakiego rodzaju wyników można oczekiwać w związku z zastosowaniem każdej z nich. W ramach tego omówienia starano się zachować porządek chronologiczny, w jakim wprowadzano do naukowstwa wszystkie metody. Za najwcześniejszą z nich można uznać metodę generowania powiązań bibliograficznych, za niemal równie wczesną – tzw. metodę historiograficzną (zob. podrozdz. 3). Następnie, po około dziesięciu latach, w obszarze bibliometrii pojawiła się metoda współcystowań, która sama z kolei wyprzedziła o około 13 lat metodę badania współwystąpień terminów naukowych Callona i in. (zob. podrozdz. 7). Ostatnią z metod tradycyjnych, której prognostyczny potencjał dostrzeżono jednak z pewnym opóźnieniem, była analiza tzw. głównych ścieżek rozwoju dyscyplin naukowych oparta w znacznej mierze na matematycznej teorii grafów (zob. podrozdz. 3). Pierwsze cztery z pięciu wymienionych (tj. wszystkie metody tradycyjne, podstawowe, z wyłączeniem szczegółów współczesnych zastosowań analizy ścieżek rozwoju dyscyplin) wraz ze wskazaniem na możliwość ich interpretacji w kategoriach ustalania stopnia podobieństwa znaczeniowego publikacji, wyczerpują część pierwszą niniejszego

artykułu. W części drugiej⁴ skupiono się na dalszych modyfikacjach metod podstawowych, na pracach porównujących przyniesione przez nie wyniki, jak również na pracach, które za swój przedmiot obrały pomijane wcześniej odgałęzienia naukowej komunikacji, a przede wszystkim dokumenty patentowe jako szczególny typ publikacji naukowych. W części drugiej rozwinięto także zasygnalizowane w części pierwszej zagadnienie prognostycznego ukierunkowania analiz ścieżek rozwoju dyscyplin naukowych oraz poświęcono więcej uwagi metodom zaczerpniętym wprost ze statystyki matematycznej, a zwłaszcza z tzw. statystycznej analizy dynamiki zjawisk masowych. Na zakończenie części drugiej podsumowano kilka wyjątkowo oryginalnych podejść metodologicznych do problemu prognozowania i oceny rozwoju dyscyplin naukowych, które nie zmieściły się w żadnej z grup omówionych dotychczas. Po ich skrótowym scharakteryzowaniu zaprezentowano wnioski, które wypływają z całokształtu analizy piśmiennictwa oscylującego wokół rozpatrywanego zagadnienia bibliometrycznego prognozowania rozwoju nauki.

Należy także dodać, że w ramach poniższego przeglądu skupiono się przede wszystkim na krytycznej analizie określonych relacji pomiędzy dokumentami, jako nośnikami informacji naukowej. Charakterystyka sieci współpracy między autorami takich dokumentów (ang. *co-authorship network*) została natomiast pominięta, przede wszystkim z uwagi na ograniczoną ilość miejsca oraz – jak się wydaje – nieco mniejszą wagę takiej analizy w zakresie rozpoznawania nowo wykształcających się frontów badawczych. W ramach obydwu części poniższego przeglądu postarano się również zaznaczyć wady, które towarzyszą większości opisywanych podejść.

3. Publikacje pionierskie

W swojej prekursorskiej pracy z 1963 r. Maxwell Kessler zaproponował metodę tzw. generowania powiązań bibliograficznych (ang. *bibliographic coupling*), która początkowo miała służyć jako narzędzie usprawniające proces wyszukiwania informacji (Kessler, 1963, 11). Jej głównym założeniem była możliwość stwierdzenia, że istnieje powiązanie treściowe (znaczeniowe, poznawcze) pomiędzy tymi dokumentami, które mają przynajmniej jedną wspólną pozycję bibliograficzną (odwołują się do co najmniej jednej, tej samej publikacji). Siła tego związku może być zmierzona za pomocą liczby takich wspólnych jednostek bibliograficznych (ang. *coupling units*), figurujących jednocześnie w bibliografiach załącznikowych dwóch publikacji (Kessler, 1963, 10; zob. też: Diodato, 1994, 12–13). Z czasem zastosowanie tej metody wykroczyło poza jej pierwotne przeznaczenie i zaczęło obejmować sposób tworzenia zbiorów (tzw. klastrów⁵ – ang. *clusters*) prac traktujących o zbliżonych zagadnieniach, a więc

⁴ Część druga artykułu zostanie opublikowana w następnym numerze ZIN [red.].

⁵ Angielski termin *data clustering* w znaczeniu „grupowanie danych” (ewentualnie też „gromadzenie danych”) pojawił się w literaturze naukowej po raz pierwszy w 1954 r. w tytule publikacji Anila Jaina dotyczącej analizy danych antropologicznych (Wierzchoń & Kłopotek, 2015, 21). W języku polskim termin *clustering* tłumaczy się jako „grupowanie” i to wyrażenie będzie odąd używane, jako odpowiednik terminu angielskiego, w niniejszym artykule. Rzeczownik *cluster* może być tłumaczony jako „zbiór”, „klastr”, „gromada”, „zbitka” lub „grono”, a w niniejszym artykule autor posługuje się dwoma pierwszymi słowami spośród pięciu wymienionych. W literaturze polskojęzycznej można także spotkać się z tłumaczeniem angielskiego *cluster* jako „gniazdo tematyczne”, które trafnie oddaje jego sens (W. Pindlowa, 1994, 42).

np. zbiorów prac wytyczających granice pewnego obszaru problemowego, włączając w to także fronty badawcze w rozumieniu Price'a (zob. np.: Huang, Chang, 2014, 1723; Jarneving, 2005, 246–248). Obecnie wykrywanie powiązań bibliograficznych jest jedną z najważniejszych metod, na których bazuje odwzorowywanie („mapowanie”) wewnętrznej, kognitywnej struktury dyscyplin naukowych i zachodzących wewnątrz niej relacji.

Rok po ukazaniu się artykułu Kesslera Eugene Garfield, Irving Sher i Richard Torpie wysunęli propozycję stworzenia tzw. historiografii nauki (ang. *historiography of science*). Była ona metodą graficznego obrazowania historii rozwoju danej dziedziny lub dyscypliny naukowej w postaci grafów⁶, w których punkty (tzw. wierzchołki grafu – ang. *nodes*) symbolizują prace naukowe, a łączące je linie – wiążące je ze sobą cytowania. Uzyskuje się dzięki temu poglądową genealogię danej dziedziny lub dyscypliny. Demonstruje ona „rodowód” dziedzin/dyscyplin, kierunki ich rozwoju i kluczowe (najczęściej cytowane) publikacje, które doprowadziły do stanu, w jakim obecnie znajduje się dziedzina/dyscyplina. Metoda uwidacznia bibliograficznych „poprzedników” i „potomków” (ang. *bibliographic antecedents and descendents*) każdego rozpatrywanego punktu (publikacji) połączonego z pozostałymi elementami w obrębie grafu (Garfield et al., 1964, 1–6; Garfield et al., 2003, 400; Garfield, 2004, 119–121). Metoda ta nie została wprawdzie wypracowana wyłącznie z myślą o przewidywaniu dalszych trendów w rozwoju dyscyplin, ale sam pomysł budowania graficznych reprezentacji danych segmentów nauki w oparciu o cytowania literatury, a zwłaszcza w połączeniu z generowaniem klastrów (jako reprezentantów frontów badawczych) jest wykorzystywany do celów prognostycznych (zob. np.: Takeda & Kajikawa, 2009, 547). Rozrysowanie dyscypliny jako „chmury” punktów (symbolizujących publikacje) wraz z łączącą je siecią linii reprezentujących albo bezpośrednio cytowania (ang. *direct citation*), albo też i inne – najczęściej oparte na cytowaniach – miary siły (lub „bliskości”) związków pomiędzy punktami, pomaga bowiem dostrzec np. kognitywne centra (ang. *hubs*) dyscyplin, ich rozmiar (w sensie liczby dedykowanych im publikacji), związki z dyscyplinami pokrewnymi, a przede wszystkim zmiany zachodzące w czasie w strukturze dyscypliny lub ewolucję czy przesunięcia wiodącej tematyki frontów badawczych. Warto dodać, że w literaturze przedmiotu wyróżniono pięć rodzajów czasowej ewolucji frontów badawczych, które dają się zaobserwować za pośrednictwem podejścia wywodzącego się z metody badania zmian struktury grafów w czasie. Są to: fronty nowo powstałe (ang. *emerging fronts*), fronty rozrastające się (ang. *growing fronts*), fronty stabilne (ang. *stable fronts*), fronty zmniejszające swój rozmiar, „kurczące się” (ang. *shrinking fronts*) i fronty zanikające (ang. *exiting fronts*) (Huang & Chang, 2014, 1722; Upham & Small, 2010, 16–19). Każdy taki typ, do którego zakwalifikuje się jakiś front, pozwala więc już na stawianie pewnych hipotez co do jego przyszłej postaci. Aktualnie dostępne jest narzędzie pozwalające na komputerowe przetwarzanie bardzo dużych zbiorów danych w celu wykreowania historiograficznej mapy danej dziedziny/dyscypliny, które nazwano *HistCite* (zob. np.: Garfield et al., 2002).

W 1973 r. Henry Small, równocześnie z Iriną Marshakovą Shaikevich (1973), zainicjowali badania nad nową metodą badania stopnia podobieństwa treści dokumentów, która daje

⁶ Z formalnego punktu widzenia graf to zbiór wierzchołków (punktów), które są połączone „krawędziami” (ang. *edges*) w taki sposób, że każda krawędź kończy się i zaczyna w którymś z wierzchołków. Ścisłą definicję i omówienie podstawowych zagadnień teorii grafów można znaleźć np. w podręczniku Robina Wilsona (2012).

możliwość „mapowania” rozwoju nauki i intelektualnej architektury dyscyplin naukowych (Marshakova Shaikovich, 2005, 1535; Small, 1973, 266). Metodę tę nazwano „współcytowaniem” (ang. *co-citation*) i zdefiniowano ją jako częstość, z jaką pewne dwa dokumenty są cytowane razem (w parze) przez inne, wydane później od nich publikacje naukowe. Siła kognitywnego związku par dokumentów jest równa (niekiedy znormalizowanej z uwagi np. na różnice w średniej długości bibliografii załącznikowych w różnych dyscyplinach) liczbie takich publikacji naukowych, które cytują jednocześnie obie prace składające się na skojarzoną ze sobą (właśnie poprzez współcytowanie) parę. W zbiorze dokumentów utworzonych z użyciem tej metody częstość współcytowań może być naturalnie różna i można ustalić dla niej pewne minimum (Small, 1973, 265–267; zob. też: Diodato, 1994, 42–47). Metoda ta jest niejako odwrotnością metody generowania powiązań bibliograficznych Kesslera. Small zaznaczył także, że w przeciwieństwie do metody generowania powiązań bibliograficznych, metoda współcytowań jest dynamiczna. To znaczy, że wzorce współcytowań i liczba publikacji cytujących jakieś wydane wcześniej pary innych prac może ulegać zmianom z biegiem czasu, co ma czynić metodę zliczania współcytowań bardziej odpowiednią w zakresie odwzorowywania relacji łączących różne dokumenty w dłuższych okresach. Ponadto związki wykryte przez obie te metody są zazwyczaj różne – nie ma pomiędzy nimi symetrii (Small, 1973, 265–267).

W publikacji Smalla i Belvera Griffith’a z 1974 r. po raz pierwszy zastosowano w praktyce (jako eksperyment) metodę tworzenia klastrów publikacji powiązanych ze sobą poprzez współcytowania. Small i Griffith dążyli do odtworzenia struktury takich specjalności naukowych, które cechowały się wtedy wysoką „aktywnością” (innymi słowy, były specjalnościami rozwijającymi się w szybkim tempie) (Small & Griffith, 1974, 17). Punktem wyjścia metody było wyselekcjonowanie grupy publikacji cytowanych w pierwszym kwartale 1972 r. powyżej pewnego minimum (ang. *initial threshold*) równego 10 cytowaniom. Publikacje te musiały być zarazem połączone relacją współcytowania również mającą różne progi (1, 3, 6 i 10 współcytowań) (Small & Griffith, 1974, 20, 22). Generowanie klastra odbywało się poprzez wybór pewnego dokumentu i dołączenie do niego wszystkich innych dokumentów, z którymi był on związany współcytowaniem. Po dodaniu tych nowych dokumentów, utworzony zbiór uzupełniono o prace, z którymi wszystkie publikacje nowo dodane były połączone tą samą relacją. Proces ten był kontynuowany aż do chwili, w której nie pozostał już żaden dokument mogący dołączyć do utworzonego klastra w oparciu o tę samą zasadę⁷. Według Smalla i Griffith’a jest to najprostsza metoda (algorytm) tworzenia klastrów, którą określono mianem „grupowania jednopiętowego” (ang. *single-link clustering*) z uwagi na fakt, że do włączenia w klaster wystarczy tylko jedno połączenie relacją cytowania z którymkolwiek dokumentem będącym już wewnątrz klastra (Small & Griffith, 1974, 23; zob. też: Wierzchoń & Kłopotek, 2015, 35–36). W efekcie zastosowania tej procedury otrzymano szereg klastrów, z których jeden wyróżniał się wyraźnie pod względem swojego dużego rozmiaru i tematycznej heterogeniczności zawartych w nim publikacji (należących głównie do dyscyplin biomedycznych). Okazało się bowiem, że publikacje opisujące pewne szeroko stosowane metody prowadzenia badań były szczególnie wysoko

⁷ Warto w tym miejscu podkreślić, że w tak skonstruowanym klastrze nie mamy do czynienia ze zbiorem publikacji, w którym każda praca jest powiązana (poprzez współcytowanie) z każdą inną, ale ze zbiorem, w którym każda publikacja jest powiązana z przynajmniej jedną inną pracą.

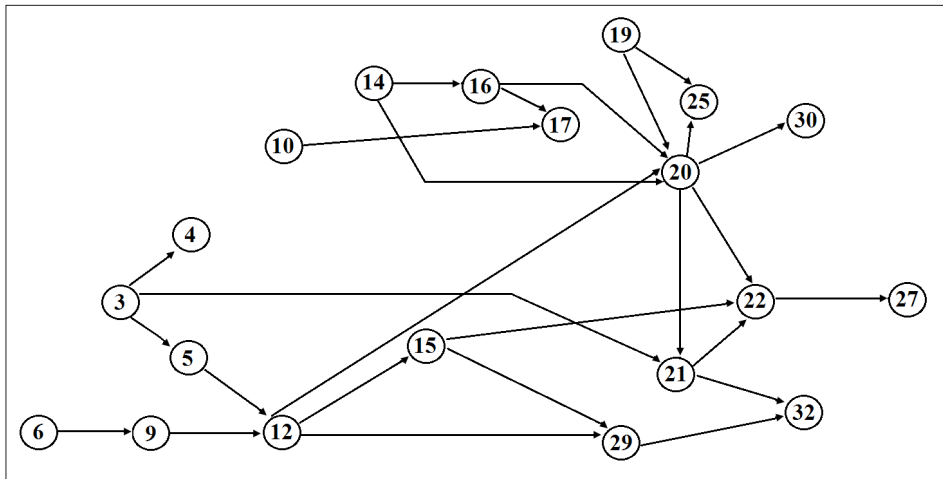
cytowane w obrębie różnych dyscyplin, co wywołało niekorzystną tendencję do łączenia ze sobą bardzo dużej liczby publikacji, które w istocie miały ze sobą niewiele wspólnego pod względem poruszanej w nich problematyki (Small & Griffith, 1974, 38; Small & Sweeney, 1985, 392). Drugim problemem był fakt, że publikacje wysoko współcytowane były zarazem także wysoko cytowane indywidualnie. Z tego powodu pojawiła się kolejna, niepożądana tendencja (ang. *bias*) do faworyzowania dyscyplin, w których wysokie liczby cytowań są normą (np. dyscypliny medyczne), w przeciwieństwie do dyscyplin mniejszych i bardziej wyspecjalizowanych (np. dyscyplin inżynierskich) (Small & Sweeney, 1985, 392). Sprawilo to, że ten szczególny klaster publikacji biomedycznych był wyjątkowo trudny do zinterpretowania z naukoznawczego punktu widzenia. W związku z tym Small i Ernest Sweeney (1985) zdecydowali, że należy nieco zmodyfikować metodę tworzenia klastrów. Stwierdzili oni, że po pierwsze należy wprowadzić procedurę normalizacji liczby cytowań prac wysoko cytowanych podczas pierwszego etapu selekcji publikacji mogących potencjalnie stać się składnikami klastrów. Normalizacja przybrała postać tzw. liczenia ułamkowego (ang. *fractional citation counting*). Przedtem każde cytowanie było uważane za równe każdemu innemu. Obecnie dzięki zliczaniu ułamkowemu, cytowanie otrzymane od publikacji z bibliografią załącznikową zawierającą niewiele pozycji, posiadało większą wagę niż cytowanie otrzymane od publikacji z bardzo liczną bibliografią załącznikową, np. od artykułu przeglądowego (Small & Sweeney, 1985, 392–393). Druga zasadnicza modyfikacja pierwotnej metody to ustalenie górnego limitu wielkości klastra (tj. maksymalnej liczby dokumentów, które mogą wchodzić w jego skład), która zapobiegłaby formowaniu się amorficznych makro-klasterów (Small & Sweeney, 1985, 395–397). Pomimo wprowadzenia tych udoskonaleń, Loet Leydesdorff w 1987 r. podkreślił, że sama metoda grupowania jednopięciowego zawiera w sobie czynniki, które powodują, że w wyniku jej zastosowania zawsze otrzyma się przynajmniej jeden klaster heterogeniczny i trudny do sensownego zinterpretowania. Wynika to z faktu, że w ramach tej metody publikacje marginalne, słabo powiązane z publikacjami centralnymi, utworzą własny klaster leżący niejako na uboczu (Leydesdorff, 1987, 301). Ponadto Leydesdorff sformułował drugi (wynikający z pierwszego) zarzut wobec tej metody, według którego wygenerowane w jej ramach klastry zawsze będą wykazywać skłonność do pewnej niespójności (Leydesdorff, 1987, 302–304). W efekcie „trudno jest tak naprawdę określić, co uzyskujemy jako końcowy rezultat w wyniku generowania [tzw.] map nauki [przy użyciu tej metody]”⁸ (Leydesdorff, 1987, 303). Trzecim zarzutem Leydesdorffa było potraktowanie zbioru danych zawartych w *Science Citation Index* (SCI) w ramach metody grupowania jednopięciowego, jako zbioru („worka” – ang. *garbage can*) indywidualnych i w zasadzie przypadkowych połączeń dokumentów. Zamiast tego, w opinii Leydesdorffa, konceptualizacja danych pochodzących z SCI powinna polegać na przyjęciu, że zbiór ten jest wysoce ustrukturyzowaną siecią zhierarchizowanych relacji między pracami cytowanymi a cytującymi (Leydesdorff, 1987, 320).

W 1989 r. Norman Hummon i Patrick Doreian wysunęli własną propozycję identyfikacji głównego nurtu przepływu informacji (ang. *main path*) w ukierunkowanej (ang. *directed*) sieci cytowań obejmującej pewną dyscyplinę lub subdyscyplinę. Sieć ukierunkowana cechuje się tym, że występuje w niej rozróżnienie na publikacje cytujące i cytowane. Przepływ

⁸ Oryg.: „It becomes difficult to say what one is actually producing in the end in terms of ‘maps of science’” (Leydesdorff, 1987, 303).

informacji ma tu więc wymiar czasowy (z przeszłości w przyszłość). Identyfikacja głównego nurtu przepływu informacji zasadza się na identyfikacji najważniejszych par dokumentów cytowanych i cytujących (ang. *links*). Najważniejszymi są połączenia par wierzchołków o najwyższej tzw. wartości przejściowej (ang. *traversal count*). Innymi słowy są to te połączenia dwóch wierzchołków sieci, które występują najczęściej podczas obliczania największych „długości” (sekwencji kolejnych punktów w sieci) wszystkich możliwych „ścieżek” (ang. *search paths*) łączących te dwa punkty (Hummon & Doreian, 1989, 48–50). Wartość przejściowa określa, ile razy każda z par połączonych punktów pojawiła się na ścieżce łączącej wszystkie kolejne pary w całej sieci cytowań. Rozważmy następujący przykład. Istnieje pewna sieć cytowań o postaci przedstawionej na rysunku 1.

Rys. 1. Przykładowy fragment grafu ukierunkowanego reprezentującego siatkę cytowań publikacji. Źródło: (Hummon & Doreian, 1989, 41)



Strzałki reprezentują istniejące cytowania (przepływ wiedzy od dokumentu cytowanego do cytującego). Np. publikacja nr 3 jest cytowana przez publikacje nr 4, 5 i 21. Z całego grafu można wyodrębnić wszystkie możliwe „ścieżki” lub „podgrafy” (ang. *subgraphs*) leżące na drodze łączącej dowolne pary wierzchołków. Niech będą to np. wierzchołki nr 3 i 22. W tym przypadku istnieją 42 możliwe ścieżki⁹. Są nimi: 3–5; 3–12; 3–15; 3–20; 3–21; 3–22; 5–12; 5–15; 5–20; 5–21; 5–22; 12–15; 12–20; 12–21; 12–22; 15–20; 15–21; 15–22; 20–21; 20–22; 21–22; 5–3; 12–3; 15–3; 20–3; 21–3; 22–3; 12–5; 15–5; 20–5; 21–5; 22–5; 15–12; 20–12; 21–12; 22–12; 20–15; 21–15; 22–15; 21–20; 22–20; 22–21.

⁹ W grafie przedstawionym na rysunku 1 istnieje 7 wierzchołków, które posiadają połączenia prowadzące od wierzchołka nr 3 do wierzchołka nr 22. Są to wierzchołki nr: 3, 5, 12, 15, 20, 21 i 22. Tylko „przechodząc” przez nie można dotrzeć od wierzchołka nr 3 do wierzchołka nr 22. Np. wierzchołek nr 16 jest usytuowany całkowicie poza tymi możliwymi ścieżkami. Jeżeli natomiast wyróżniony w ten sposób „podgraf” ma dokładnie 7 wierzchołków, teoretycznie istnieją $7(7-1) = 7 \cdot 6 = 42$ możliwe połączenia między nimi – tj. każdy wierzchołek z każdym innym wierzchołkiem, oprócz samego siebie (zob. Hummon & Doreian, 1989, 49). Dodatkowo znaczenie ma tu kierunek połączenia (przepływu informacji). Połączenie „3–5” nie jest więc tym samym, co połączenie „5–3”.

Tylko 19 wyróżnionych połączeń, spośród wszystkich 42 możliwych, jest w rzeczywistości zrealizowanych w przykładowym grafie na rysunku 1. Np. połączenie „3–21” (leżące na drodze pomiędzy wierzchołkami nr 3 i nr 22) ma postać następującej ścieżki: 3–5–12–20–21. Kiedy rozpatrzylibyśmy wszystkie możliwe ścieżki dla powyższego przykładu dziewiętnastu par okazałoby się, że połączenie „5–12” występuje najczęściej, tj. pojawia się w największej liczbie połączeń, spośród wszystkich dziewiętnastu faktycznie zrealizowanych w tym konkretnym grafie. Istnieje ono np. w ścieżce łączącej parę „3–15”: 3–5–12–15, w ścieżce łączącej parę „3–22”: 3–5–12–15–22, w ścieżce łączącej wierzchołki „5–21”: 5–12–20–21, a ponadto jeszcze w siedmiu innych ścieżkach. Ogółem występuje więc ono w 10 ścieżkach, podczas gdy np. para „3–21” tylko w dwóch („3–21” i „3–22”). W związku z tym połączenie „5–12” ma najwyższą wartość przejściową (*traversal count*), równą 10 czyli jest połączeniem o najwyższej „sile” w przykładowym podgrafie „3–22” (zob. Hummon & Doreian, 1989, 50). Wartość tę można w analogiczny sposób ustalić dla każdej pary połączonych wierzchołków (dla wszystkich możliwych podgrafów danego grafu), a „główna ścieżka” całego grafu to taka, która odzwierciedla sieć najsilniejszych połączeń wewnątrz niego. Będąc przy każdym wierzchołku mamy bowiem do wyboru jeden lub więcej kolejnych wierzchołków różniących się wartością przejściową o znanej nam wartości. Tym sposobem identyfikuje się publikacje najistotniejsze (o największej sile wpływu) w całej sieci lub, inaczej mówiąc, najważniejszy strumień (ang. *main stream*) przepływu informacji. Należy dodatkowo podkreślić, że publikacje składające się na „główną ścieżkę” niekoniecznie muszą być tym samym publikacjami najwyżej cytowanymi w obrębie sieci (Liu & Lu, 2012, 529). Metoda ta sama w sobie nie została zaprojektowana z myślą o prognozowaniu rozwoju dyscyplin, pomimo tego wydaje się, że posiada ona znaczny potencjał w tym zakresie (zob. np.: Ma & Liu, 2016; Tu & Hsu, 2016).

4. Miary podobieństwa publikacji

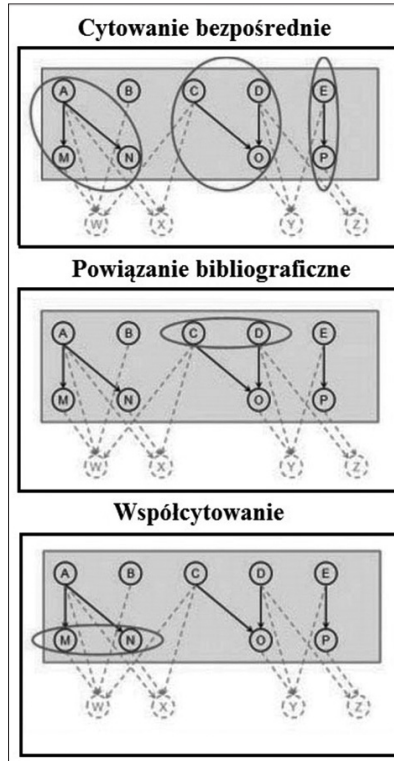
Jak wynika z powyższego przeglądu publikacji pionierskich, większość prób identyfikacji kierunków dalszego rozwoju dyscyplin naukowych opierała się na wyodrębnianiu podzbiorów (klastrów) publikacji naukowych należących do określonej dyscypliny, które stanowią reprezentacje jej subdyscyplin – węższych niż cała dyscyplina specjalności czy obszarów problemowych, które mogą ewoluować w swoim własnym tempie i kierunku. Metody oceny tempa tej ewolucji wymagają posłużenia się specjalnie zaprojektowanymi technikami analizy wyróżnionych już wcześniej klastrów. Ocenia się np. ich rozmiar, średni wiek publikacji wchodzących w ich skład, wiodącą tematykę, a przede wszystkim zmiany zachodzące we wspomnianych cechach (w szczególności w średnim wieku i rozmiarze) w miarę przesuwania się na osi czasu w kierunku przyszłości. Operacja wydzielenia klastrów w ramach dyscypliny jest więc podstawą dalszej analizy ukierunkowanej prognostycznie. Z samej definicji klastra wynika z kolei wprost, że stanowi on zgrupowanie elementów homogenicznych, chociaż niekoniecznie identycznych. Dwa dowolne elementy należące do jednego klastra powinny być jednak podobne do siebie w znacznie większym stopniu, niż do dowolnego elementu należącego do innego klastra (Wierzchoń & Kłopotek, 2015, 17). Powstaje w związku z tym zasadnicze pytanie dotyczące tego, w jaki sposób należy zdefiniować to podobieństwo oraz w jaki sposób posłużyć się zaakceptowaną już miarą

podobieństwa w trakcie samej operacji grupowania. Ponieważ odpowiedzi na te dwa pytania są od siebie niezależne, w efekcie badacze mają do czynienia z ogromną wielością różnorodnych algorytmów służących do wyróżniania klastrów (Wierzchoń & Kłopotek, 2015, 17), co znajduje swoje odbicie także w bibliometrycznych zastosowaniach tej operacji. Z ogółu stosowanych w tym zakresie miar podobieństwa trzy zostały już wspomniane w poprzednim podrozdziale. Były nimi siła powiązania bibliograficznego, siła relacji współcytowania i „siła” (w sensie liczebności) cytowań bezpośrednich. Czwartą miarą podobieństwa jest z kolei omówiona w kolejnym podrozdziale siła (intensywność) relacji współwystępowania specjalistycznych terminów/słów kluczowych (ang. *co-word analysis*). Wszystkie tego rodzaju miary pozwalają na kwalifikowanie publikacji do klastrów jako twórców względnie homogenicznych, stanowiących reprezentacje ewoluujących w czasie subdyscyplin lub specjalności dyscyplinarnych (zob. np.: Soós, 2014, 23). Trzy pierwsze miary w przystępny sposób omówili i przedstawili graficznie Kevin Boyack i Richard Klavans w 2010 r. (zob. Rysunek 2) (Boyack & Klavans, 2010, 2390). Cytowanie bezpośrednie pozwala stworzyć klaster obejmujący prace cytowane wraz z cytującymi. Na rysunku 2 są to np. publikacje „A”, „M” i „N” tworzące trójelementowy klaster {A, M, N}. Powiązane bibliograficznie występuje pomiędzy publikacjami „C” i „D” ponieważ obydwie cytują publikację „O”. W ten sposób powstał dwuelementowy klaster {C, D}. Publikacje „M” i „N” są z kolei połączone relacją współcytowania z uwagi na fakt, że obydwie zostały zacytowane przez dokument „A”. W tym przypadku powstał więc klaster {M, N} (Rys. 2).

W związku z niewielką liczbą prac omawiających relację podobieństwa w kontekście mapowania nauki wydaje się, że w ramach drugiej podgrupy wystarczy wspomnieć o przeglądzie tego rodzaju miar dokonany przez Annę Huang (2008) i zwrócić szczególną uwagę na tzw. podobieństwo kosinusowe (ang. *cosine similarity*) publikacji naukowych. Jest ono bowiem techniką często stosowaną w perspektywie podziału zbioru publikacji na klastry. W dużym skrócie polega ono na zamianie całości treści dokumentu, znormalizowanej w odpowiedni sposób (np. ujednotwiona się wyrażenia synonimiczne, eliminuje tzw. „słowa stopujące” – ang. *stop words* – które są wyrażeniami w rodzaju „i”, „lub”, „także” itd. oraz usuwa prefiksy i sufiksy pozostawiając jedynie semantyczne rdzenie słów) na wektory n -wymiarowe. Wektor ma postać ciągu kolejnych n wyrażen, np.: {*chemic, kinetic, fabric, polymer...*}. Jeżeli po normalizacji w dokumencie pozostaje np. 1000 unikalnych słów, utworzony zeń wektor ma 1000 wymiarów – każdy wymiar odpowiada każdemu, unikalnemu słowu/terminowi. W ten sposób można zbudować tzw. macierze (ang. *matrices*) pokazujące, które znormalizowane terminy i w jakiej liczbie występują w określonych, porównywanych dokumentach. Jeżeli dwa dokumenty wykorzystują wiele tych samych terminów to stopień ich podobieństwa będzie wysoki. Dzięki temu możliwe jest precyzyjne, ilościowe określenie stopnia semantycznej równoważności dwóch lub więcej publikacji naukowych. Podobieństwo kosinusowe przyjmuje wartości z przedziału od 0 do 1. Wartość „1” oznacza, że dokumenty są identyczne. Więcej informacji na ten temat można znaleźć np. w publikacjach Anny Huang (2008, 51–52), Mirosława Wierzchońa i Sławomira Kłopotka (2015, 31–32), Piotra Potiopy (2011, 416–417), Gerarda Saltona i Donny Bergmark (1979, 146–147), a także w poradniku Erica Garcii (2015).

Rys. 2. Graficzna reprezentacja powiązania hipotetycznych publikacji za pomocą trzech relacji wykorzystywanych w obrębie nurtu bibliometrycznego mapowania dyscyplin naukowych.

Źródło: opracowanie własne na podstawie (Boyack & Klavans, 2010, 2390)



5. Metoda współcytowań

Jak już zasygnalizowano w części czwartej, metoda współcytowań to pierwsza z czterech podstawowych metod wyróżniania względnie homogenicznych klastrów publikacji powiązanych ze sobą siecią cytowań, która daje w efekcie swego rodzaju ekwiwalenty frontów badawczych Price'a. Jej odmianę polegającą na analizie współcytowań autorów publikacji naukowych (zamiast samych publikacji) wykorzystał w praktyce Olle Persson (1994). Zarazem powtórzył on zarzut Leydesdorffa wobec metody grupowania jedno-połączeniowego (wytwarzanie zbyt dużych i heterogenicznych klastrów) oraz zauważył, że sama liczba współcytowań autorów jest zbyt prostą miarą podobieństwa i że należy stosować normalizację z uwagi na ogólną cytawalność poszczególnych autorów (Persson, 1994, 34). Persson zademonstrował następnie w jaki sposób normalizację tego rodzaju można wprowadzić do analizy współcytowań autorów artykułów ogłoszonych w czasopiśmie *Journal of the American Society for Information Science* w latach 1986–1990 oraz zaproponował ograniczenie, według którego algorytm nie pozwoli na uformowanie się klastra, który

zawierałby mniej niż trzech autorów. Algorytm eliminował ponadto z dalszego procesu przyłączania do innych klastrow takich par autorów, którzy zostali już wcześniej włączeni do odrębnego klastra. Modyfikacja ta miała zapobiegać powstawaniu klastrow heterogenicznych (Persson, 1994, 32, 34).

W tym samym roku Michael Zitt i Elise Bassecouard (1994) uzupełnili analizę klastrow publikacji współcytowanych jako reprezentacji obszarów problemowych o badanie średniego wieku klastrow celem identyfikacji ich trendów rozwojowych (Zitt & Bassecouard, 1994, 335). Zaznaczyli przy tym, że wadą tego podejścia jest konieczność zaakceptowania faktu, że daty ostatecznej publikacji tekstów są późniejsze niż daty nadesłania ich do redakcji czasopism, które z kolei są późniejsze niż daty prowadzenia badań i formułowania końcowych wniosków lub hipotez. Opóźnienie to (ang. *time lag*) wpływa negatywnie na możliwość oceny dynamiki zmian klastrow, która jest kluczowym elementem podczas identyfikowania bieżących trendów badawczych (Zitt & Bassecouard, 1994, 336). W publikacji Zitta i Bassecouard zwrócono ponadto uwagę na inną wadę metody analizy współcytowań polegającą na zdolności tej metody do jedynie wycinkowej klasyfikacji całości literatury cytowanej przez społeczność naukową. Inaczej mówiąc, te publikacje, które nie są współcytowane (lub są współcytowane w niewystarczającej liczbie) z innymi pracami, pozostają poza nawiasem stosowalności metody. Problematyczna w opinii obojga autorów była również arbitralność ustalania maksymalnej wielkości klastrow (jak uczynili to np. Small i Sweeney w 1985 r.). Według Zitta i Bassecouard pytanie o to, jak duży musi być klastrow, aby mógł on zostać uznany za samodzielną jednostkę, pozostawało otwarte (Zitt & Bassecouard, 1994, 337, 339). Oryginalna metoda szacowania trendu została ostatecznie oparta przez parę autorów o stosunek wieku publikacji tworzących (wchodzących w skład) klastrow do wieku publikacji cytujących dokumenty wchodzące w skład klastra. Trend badawczy dyscypliny konstytuowały w ich opinii tematy progresywne i zakodowane tekstualnie, tj. określone w pewien sposób pod względem tematyki odrębną metodą (Zitt & Bassecouard, 1994, 340). „Progresywne” oznacza tu bycie cytowanym przez publikacje wydane stosunkowo niedawno. Zarazem tematy te muszą być oparte na materiale „młodym” (tj. na klastrach publikacji o niskiej średniej arytmetycznej ich wieku) (Zitt & Bassecouard, 1994, 343–344, 348). Dopiero te dwa warunki spełnione łącznie pozwalają określić, które trendy czy obszary problemowe można uznać za „wschodzące”.

Henry Small (2006) również postawił pytanie o to, czy możliwe jest przewidzenie rozrostu pewnych obszarów badawczych i kierunku postępu w nauce, albo wyłonienia się całkowicie nowych dyscyplin spośród tych już istniejących (Small, 2006, 596). Jego metoda rozwiązania tego problemu opierała się na analizie klastrow publikacji współcytowanych i zarazem takich, które stanowiły górny 1% publikacji cytowanych indywidualnie najwyższej (ang. *top 1%*) w zachodzących na siebie przedziałach czasowych (np. przedziały postaci 2000–2003, 2002–2004 itd.). Small położył przy tym szczególny nacisk na „kontynuowalność” (ang. *continuity*), inaczej mówiąc ciągłość występowania w klastrze tych samych, najwyższej cytowanych publikacji w kolejnych przedziałach czasu. Występowanie takiej ciągłości nazwał on „łańcuchami klastrow” (ang. *cluster strings*) (Small, 2006, 597; Upham & Small, 2010, 19). Predykcji obszarów wzrostu służyła w tym ujęciu tzw. „aktualność klastra” (ang. *currency of cluster*) będąca średnim wiekiem publikacji najwyższej cytowanych wchodzących w jego skład. Publikacje te stanowiły centrum klastra, jego kognitywny „rdzeń”. Obserwując jak zachowywały się „łańcuchy” podczas przejść w coraz bliższe współczesności

przedziały czasu i mierząc ich aktualność, można było wychwycić transformacje jakim ulegał „łańcuch” (rozrost, przyłączanie publikacji z pokrewnych dyscyplin, rozgałęzianie się, zanikanie itd.). Aktualność mieściła się w przedziale od 0 do 1. Aktualność równa „1” oznaczała, że wszystkie publikacje najwyższej cytowane (cały *top 1%*) zostały wydane w ostatnim roku przedziału czasowego, który poddaje się analizie. Aktualność równa zeru oznaczała natomiast, że wszystkie te prace pochodziły z pierwszego roku wyznaczającego dolną granicę przedziału (Small, 2006, 607–608). Następnie zbadano związek rozmiaru klastrów z ich aktualnością. Rozmiar klastra był tu wskaźnikiem tego, jak ożywiona jest aktywność publikacyjna w danym obszarze tematycznym, jak wielu badaczy przyciąga jego tematyka i jaki jest jego ogólny potencjał badawczy. Small zaobserwował w tym względzie pewną korelację. Okazało się np., że z wysoką aktualnością z przedziału 0,8–1 związany był rozrost klastra o około 35%, a z aktualnością równą lub niższą niż 0,2 – spadek liczby artykułów bardzo wysoko cytowanych w klastrze (Small, 2006, 608). Small skonkludował jednak, że korelacja ta była zbyt słaba aby wyjaśnić rozrastanie się klastrów za pomocą samej jedynie aktualności. Poza aktualnością musiały więc istnieć inne, niezidentyfikowane czynniki, które wpływały na powiększanie się obszarów badawczych. Według Smalla należało wobec tego poszukiwać dalszych determinantów krótko – i długoterminowej dynamiki ewoluowania dyscyplin (Small, 2006, 609–610).

Ostatnią publikacją, o której należy powiedzieć w kontekście metody badania współcytowań jest artykuł Samuela Phineasa Uphama i Henry’ego Smalla (2010), w którym wspomniano o m. in. dwóch poważnych ograniczeniach metody w takiej postaci, jaką zaprezentował Small w 2006 r. Pierwszym była jej niezdolność do zidentyfikowania nowego frontu badawczego, jeżeli żaden z należących doń artykułów nie stanie się wysoko cytowany. Ograniczenie to dotyczy przede wszystkim takich dyscyplin nauki, w których normą są raczej niskie wskaźniki cytowań (np. matematyka, nauki społeczne i humanistyczne). Stąd metoda ta nigdy nie pozwoli na identyfikację nowo wyłaniającego się obszaru problemowego czy subdyscypliny natychmiast po jej uformowaniu się. Drugim ograniczeniem metody była niemożność zidentyfikowania wszystkich prac należących do nowo powstałego obszaru problemowego. Zamiast tego, przy jej użyciu możliwa będzie tylko detekcja samego istnienia takiego obszaru i dostarczenie pewnej próbki należących do niego prac wysoko cytowanych, a nie jego precyzyjne wyodrębnienie. Metoda ta nadawała się więc najlepiej do „szybkiego skanowania” rozleglejszych obszarów nauki, a nie do dokładnego rozpoznawania i wyszczególnienia nowych prądów intelektualnych czy nowych tendencji dyscyplinarnych (Upham & Small, 2010, 18; zob. też: Van Den Besselaar & Heimeriks, 2006, 380). Oryginalność omówionego przez Smalla i Uphama podejścia prognostycznego sprowadzała się po pierwsze, do wyróżnienia nowych kategorii klastrów. Były nimi klastry publikacji wysoko cytowanych o zerowej „kontynuowalności”, tj. takie, które zaobserwowano po raz pierwszy dopiero w pewnym (którymś z kolei) przedziale czasowym ze wszystkich przedziałów objętych analizą (Upham & Small, 2010, 19). Ponadto do nowych kategorii klastrów należały klastry „rosnące”, „stabilne”, „kurczące się” i „zanikające” pod względem liczby składających się na nie dokumentów w kolejnych przedziałach czasu (Upham & Small, 2010, 21–22). Po drugie, wprowadzono nowy wskaźnik – tzw. endogeniczność klastra. Określała ona procentowo, jak wiele z dokumentów cytowanych (tworzących „rdzeń” klastra) było zarazem dokumentami cytującymi ten „rdzeń”. Wskaźnik ten informuje więc o tym, do jakiego stopnia klastry „bazuje sam na

sobie”, jak bardzo jest samowystarczalny. Front badawczy o wysokiej endogeniczności jest tworem rozwijającym się w oparciu o swoje własne podstawy, jest mało interdyscyplinarny i będzie prawdopodobnie ewoluował w kierunku samodzielnej subdyscypliny (Upham & Small, 2010, 20). Ogółem, prognoza wyłaniania się nowych subdyscyplin polegała więc ostatecznie na rozważeniu zależności między typem klastra („rosnący”, „stabilny” itd.), zmianami cytawalności publikacji tworzących klastry w kolejnych przedziałach czasu (tzw. „wchłanianie” frontu/klastra przez sąsiednie subdyscypliny), „natychmiastowością” klastra (tj. odsetkiem publikacji tak cytujących, jak i cytowanych w klastrze, które opublikowano w „górnym” brzegowym roku analizowanego przedziału czasowego) oraz stopniem jego endogeniczności (Upham & Small, 2010, 23–26). Podsumowując, wydaje się, że metoda Uphama i Smalla była udoskonaloną i bardziej wyrafinowaną formą metody zaproponowanej przez samego Smalla w 2006 r.

W związku z metodą badania współcytowań należy ponadto wyraźnie podkreślić, że jednym z jej ograniczeń, o którym nieczęsto wspomina się w literaturze zagranicznej, jest konieczność dysponowania indeksem cytowań (takim jak np. *Web of Science* lub *Scopus*). Indeks cytowań, który obejmowałby całość polskojęzycznej literatury naukowej ze wszystkich dziedzin nauki nie jest natomiast na chwilę obecną dostępny. Istniejące polskie indeksy cytowań są bowiem albo dopiero w fazie „embrionalnej” („POL-index”), albo dotyczą tylko wybranych dyscyplin (np. „Arton”, „CYTBIN”), albo prace nad rozwijaniem indeksu nie są już kontynuowane (np. „Indeks Cytowań Socjologii Polskiej”, „Indeks Cytowań Historiografii Mediów Polskich”) (Seweryn & Swoboda, 2013). Jedyną realną ewentualnością wydaje się tu „Google Scholar Citations”, chociaż i w tym przypadku istnieje szereg mankamentów i możliwych nieprawidłowości, do których prowadzi poleganie na tym źródle danych. Przegląd takich mankamentów przedstawili np. Alberto Martín-Martín i in. (2016, 31–47).

6. Metoda generowania powiązań bibliograficznych

Generowanie powiązań bibliograficznych posiada przynajmniej dwie zalety w porównaniu z metodą współcytowań. Po pierwsze, do realizacji badania z użyciem tej metody nie jest potrzebny indeks cytowań, a jedynie dostęp do bibliografii załącznikowych wybranego zbioru publikacji. Po drugie, w metodzie tej nie występuje problem dotyczący metody współcytowań, polegający na fakcie konieczności opóźnienia badania aż do czasu zakułowania się wystarczającej liczby cytowań (a zatem i współcytowań) wokół publikacji reprezentatywnych dla dyscypliny lub funkcjonujących w jej obrębie specjalności (zob. np.: Boyack & Klavans, 2010, 2391; Glänzel, 2012, 195; Glänzel & Thijs, 2012, 400–401; Shibata et al., 2008, 762; Shibata et al., 2009, 577). Ocenia się, że do czasu wystąpienia cytawalności maksymalnej musi minąć przynajmniej 2–3 lata, a niekiedy nawet więcej, w zależności od konkretnej, wybranej jako przedmiot badań dyscypliny (zob. np.: Adler et al., 2009, 5; Briotta Parolo et al., 2015, 736–738; Dorta-González & Dorta-González, 2013, 594–600; Glänzel & Schoepflin, 1995, 39; Glänzel & Schoepflin, 1999, 43; Moed et al., 1998, 394, 396, 399).

Kevin Boyack, Katy Börner i Richard Klavans (2009) wykorzystali metodę powiązań bibliograficznych w celu wizualizacji dynamiki rozwoju dyscyplin chemicznych pomiędzy rokiem 1974 a 2004. Wykorzystaną tu miarę podobieństwa między czasopismami (siłę powiązania bibliograficznego) znormalizowano przy pomocy podobieństwa kosinusowego

(zob. podrozdział 4; Boyack et al., 2009, 47, 51). Głównym przedmiotem analizy była mapa nauki za 2002 r. złożona z 671 wyodrębnionych klastrów oraz połączeń pomiędzy nimi. Mapę uzupełniono dodatkowo o dane dotyczące relacji wzajemnych cytowań czasopism. Relacja występowała między czasopismami A i B, jeżeli czasopismo A cytowało czasopismo B, a czasopismo B cytowało czasopismo A (Boyack et al., 2009, 48, 51). Dzięki temu wszystkie czasopisma nieobecne na mapie za rok 2002 mogły zostać do niej dołączone w oparciu o siłę związku z poszczególnymi klastrami z 2002 r. (ang. *journal-to-cluster-similarity*). To znaczy, że każde czasopismo, które pojawiło się w danych o cytowaniach *Journal Citation Reports* (JCR) w którymkolwiek roku z przedziału 1974–2004 zostało przypisane do jednego z 671 klastrów wykrytych w 2002 r. (Boyack et al., 2009, 51). Dzięki użyciu kategorii tematycznych JCR uzyskano następnie czasowo-tematyczny przekrój wszystkich 671 klastrów dla każdego kolejnego, pięcioletniego przedziału czasowego (Boyack et al., 2009, 52, 54). Pozwoliło to na zaobserwowanie, jak w następujących po sobie przedziałach czasowych zmieniały się ilościowo rozmiary poszczególnych, należących do dyscyplin chemicznych, specjalności oraz ich stopnie wzajemnego przenikania się. Przykładowo przed 30 laty bioinżynieria niemalże nie istniała samodzielnie poza obszarem biochemii, podczas gdy w 2004 r. odgrywała już znaczącą rolę w mikrobiologii. Odnotowano też istnienie przepływu wiedzy z bioinżynierii do chemii polimerów, chemii analitycznej i chemii ogólnej (Boyack et al., 2009, 55–59). Jako słabą stronę swojego podejścia metodologicznego Boyack i in. wskazali natomiast fakt, że przyjęcie czasopism jako jednostki analizy jest w stanie wskazać jedynie duże, silne trendy w badanych obszarach. Obniżenie poziomu agregacji danych do poziomu indywidualnych publikacji pozwoliłoby natomiast na identyfikację także tych mniej widocznych i węższych trendów, a w szczególności takich trendów badawczych, które usytuowane są „na przecięciu” (ang. *at intersection*) szerszych specjalności (Boyack et al., 2009, 59–60).

Z interesującym pomysłem związanym z udoskonaleniem metody powiązań bibliograficznych wystąpił Sandor Soós (2014). Zaproponował on mianowicie uwzględnienie dodatkowego parametru przypisów współdzielonych przez pewne publikacje, jakim był wiek tych przypisów. Im starsze były współdzielone przed dwa współczesne artykuły pozycje bibliograficzne, tym słabszy powinny mieć one wpływ na siłę powiązania bibliograficznego obu artykułów współczesnych. Im z kolei wcześniejsze daty wydania publikacji figurujących w bibliografiach załącznikowych dwu artykułów, tym wyższy stopień intelektualnego czy poznawczego pokrewieństwa między artykułami, w których te pozycje bibliograficzne współwystępują. Modyfikację tę Soós nazwał „powiązaniem bibliograficznym wrażliwym na wiek pozycji bibliograficznych” (ang. *age-sensitive bibliographic coupling*) (Soós, 2014, 24). Celem modyfikacji było doprecyzowanie metody tworzenia takich klastrów publikacji związanych relacją bibliograficznego podobieństwa, które byłyby w stanie lepiej zrekonstruować historyczny rozwój danej struktury tematycznej i uwidocznić bardziej klarownie dynamikę danego obszaru badań (Soós, 2014, 24, 33). Różnica w dacie wydania pozycji bardzo starych (wydanych np. 800 i 1000 lat temu) w mniejszym stopniu determinowała siłę powiązania bibliograficznego, niż różnica w wieku publikacji wydanych np. 8 i 10 lat temu (Soós, 2014, 27). Waga takich dokumentów współcytowanych nie maleje w związku z tym liniowo (wraz ze wzrostem dystansu od czasów terażniejszych), ale wykładniczo. Konsekwencją tego faktu jest to, że przywoływane w dwóch różnych artykułach pozycje antyczne lub średniowieczne mają niemal tę samą wagę, która utrzymuje się na równie

niskim poziomie. Jeżeli np. publikacja P1 ma tę samą liczbę współdzielonych przypisów co publikacja P2 i publikacja P3, ale z P2 dzieli przypisy młode, a z P3 stare, według standardowej metody powiązań bibliograficznych P1 jest tak samo podobne do P2 jak do P3, ponieważ istotna jest tylko sama liczba wspólnych przypisów. W metodzie wrażliwej na wiek P1 będzie natomiast dużo bardziej podobne do P2 niż do P3 (Soós, 2014, 29). Zdaniem Soósa jego metoda pozwala na generowanie klastrów odzwierciedlających podziały i różnicowanie się kierunków badawczych, które wyłaniają się z czasem w obrębie danych dyscyplin nauki. W ramach eksperymentalnego porównania wyników stosowania metody „klasycznej” i metody wrażliwej na wiek, w przypadku tego samego zbioru publikacji, nowa metoda pozwoliła np. rozszcześcić duży, mało spójny i reprezentujący szeroką dyskusję teoretyczną klastery na klastry mniejsze, które reprezentowały odpowiednio: początkowy kontekst dyskursu, jego późniejszy rozwój i wydzielenie się spójnych tematycznie szkół, będących odpowiedziami na problemy podniesione w ramach kontekstu początkowego (Soós, 2014, 49–50). Potencjał metody w zakresie prognozowania leży zatem głównie w możliwości precyzyjniejszego wyodrębniania wąskich trendów badawczych (ang. *strains of research*) w korpusach publikacji współczesnych oraz w dokładniejszym rozróżnieniu tych kierunków rozwoju, które są odbiciem dyskusji nad zagadnieniami klasycznymi (ang. *classical*) i zagadnieniami bieżącymi (ang. *recent*), co jest często niemożliwe przy zastosowaniu tradycyjnej metody generowania powiązań bibliograficznych (Soós, 2014, 48–50).

Mu-Hsuan Huang i Chia-Pin Chang (2014) również wykorzystali metodę generowania powiązań bibliograficznych pomiędzy artykułami cytowanymi wysoko, jako tymi, które odznaczają się wysoką jakością intelektualną, znaczną siłą wpływu czy rangą naukową (Huang & Chang, 2014, 1722). Huang i Chang zastosowali, wzorem Smalla (2006) oraz Smalla i Uphama (2010) pięć zachodzących na siebie okresów obserwacji (ang. *sliding citation window*), aby prześledzić ewolucję pewnej szczególnej specjalizacji z zakresu elektroniki. Okresy obserwacji to lata: 2000–2005, 2001–2006, 2002–2007, 2003–2008 i 2004–2009 (Huang & Chang, 2014, 1723). Okresy te, jako zazębiające się ze sobą, były zdaniem Huang i Chang w stanie zrekonstruować dynamikę specjalności naukowej, w przeciwieństwie do jakiegokolwiek odgórnie ustalonego (statycznego) okresu obserwacji. Po określeniu wymaganego stopnia cytawalności selekcjonowanych publikacji oraz prognozy siły powiązania bibliograficznego (ang. *coupling strenght*) równego co najmniej pięciu wspólnym pozycjom bibliograficznym, posłużono się techniką tzw. grupowania hierarchicznego¹⁰ z wykorzystaniem wskaźnika korelacji Pearsona jako miary podobieństwa między grupowanymi elementami (Huang & Chang, 2014, 1724, 1726; zob. też: Wierzchoń & Kłopotek, 2015, 35–38; nt. wskaźnika Pearsona jako miary podobieństwa między publikacjami zob. np.: Huang, 2008, 52). Warto podkreślić, że zarówno wybór metody generowania klastrów jak i wybór wskaźnika podobieństwa między grupowanymi dokumentami ma duży wpływ na uzyskane wyniki (tj. liczbę i rozmiar uzyskanych klastrów) (Huang & Chang, 2014, 1726).

¹⁰ W tzw. grupowaniu hierarchicznym (ang. *hierarchical clustering*) punktem wyjścia jest potraktowanie całego m-elementowego zbioru wszystkich obiektów jako grupy m jednoelementowych klastrów. Następnie oblicza się tzw. odległości (podobieństwo) każdej pary takich jednoelementowych klastrów i wybiera parę elementów o najwyższym stopniu podobieństwa. Para ta tworzy pierwszy, 2-elementowy klastery. Dalej aktualizuje się obliczenia odległości między klastrem a wszystkimi pozostałymi 1-elementowymi klastrami i powtarza całą procedurę od początku. Nt. dalszych szczegółów grupowania hierarchicznego, łącznie z jego wadami, zob. np. (Wierzchoń & Kłopotek, 2015, 35–38).

Wobec wielości dostępnych w tym zakresie metod i wskaźników, wybór ten nie powinien być zatem przypadkowy i wymaga on uzasadnienia. W wyniku wyboru dokonanego przez Huanga i Chang, w każdym przedziale czasowym uzyskano szereg klastrów reprezentujących różne fronty badawcze i cechujące się różnym rozmiarem (liczbą publikacji) i problematyką (zob. Huang & Chang, 2014, 1729–1739). W celu zakwalifikowania uzyskanych frontów do różnych specjalności badawczych (ang. *knowledge domains*) w ramach wybranego obszaru, zasięgnięto opinii ekspertów. Wykorzystując klasyfikację frontów badawczych Smalla i Uphama (2010) (fronty „rosnące”, „stabilne” itd.), wzbogaconej o jeden nowy typ frontu – frontu „powtórnie wyłaniającego się” (ang. *re-emerging*), wytypowano łącznie 18 frontów badawczych należących do siedmiu różnych specjalności i wskazano specjalność najbardziej rozwojową (ang. *the most cutting-edge sector and trend for future development*). Dokonano tego na podstawie charakterystyk wchodzących w jej skład frontów, tj. oceny tego, ile każda specjalność zawierała frontów nowo powstałych, ile stabilnych, ile rosnących, ile kurczących się itd. (Huang & Chang, 2014, 1739–1740). Jako na ograniczenia zastosowanej metodologii wskazano na wykorzystanie artykułów naukowych o ustalonym progu cytowalności pomimo, że w dyscyplinach zorientowanych technologicznie patenty mogą stanowić relatywnie ważniejsze źródło informacji. Ponadto, w obliczu świadomości tempa rozwoju współczesnej technologii Huang i Chang uznali, że ustalanie okresów obserwacji cytowanych publikacji (ang. *citation window*) mogłoby zostać oparte na wartościach wskaźnika *citing* lub *cited half-life* kluczowych czasopism dyscyplinarnych, zamiast na „prywatnej” intuicji autorów (Huang & Chang, 2014, 1741–1742). Warto w tym miejscu nadmienić, że proces starzenia się literatury naukowej stanowi przeszkodę w efektywnym stosowaniu metody powiązań bibliograficznych. Prawdopodobieństwo, że bibliografie załącznikowe dwóch powiązanych tematycznie artykułów będą zawierać jakieś wspólne pozycje jest bowiem niskie jeżeli między opublikowaniem takich dwóch prac minęło dużo czasu (Glänzel & Thijs, 2012, 400–401).

7. Metoda współwystępowania słów kluczowych

W literaturze przedmiotu podaje się, że metoda analizy współwystępowania terminów/słów kluczowych wykrystalizowała się jako samodzielna technika badawcza w wydanej w 1986 r. książce pt. *Mapping the dynamics of science and technology*, której autorami byli Michel Callon, John Law i Arie Rip (zob. np.: Lee, 2008, 505; Sasson et al., 2015, 620; Van Den Besselaar & Heimeriks, 2006, 380). Najogólniej mówiąc metoda ta polega na wykrywaniu terminów występujących w tytułach, abstraktach lub treści dwóch lub więcej różnych publikacji łącznie. Ma ona zapewnić wgląd w podejmowaną w tych publikacjach tematykę i przedstawiane w nich koncepcje. Zakłada się przy tym, że różne specjalizacje badawcze posługują się różną terminologią. Powstawaniu nowych koncepcji, metod czy teorii towarzyszy powstawanie nowego słownictwa, które jest dla nich reprezentatywne (definiuje je). Wykrycie powtarzających się współwystąpień pewnych fraz w różnych, rozdzielonych okresach pozwala domniemywać, że reprezentowane przez to słownictwo trendy, koncepcje lub teorie zostały zaakceptowane przez społeczność naukową i stały się w pewnych kręgach tej społeczności „obligatoryjne” (Neef & Corley, 2009, 659; zob. też: Diodato, 1994, 54–55; Verbeek et al., 2002, 205–206). Callon i in. w swojej pracy z 1986 r.

zapropowali tzw. indeks jednoznaczności publikacji (ang. *equivalence index*), który jest wartością wprost proporcjonalną do liczby współwystąpień (w jednym dokumencie) dwóch różnych terminów i jednocześnie odwrotnie proporcjonalną do liczby różnych dokumentów zawierających te dwa terminy. Indeks ten mieści się w przedziale od 0 do 1 i im jest większy, tym bardziej jednoznacznie semantycznie są obie publikacje (Neef & Corley, 2009, 664; Sasson et al., 2015, 620). Dzięki temu uzyskuje się dokładną miarę ilościową określającą podobieństwo różnych publikacji. Miara ta jest stopniowalna (podobieństwo zmienia się „płynnie”). Możliwe staje się więc tworzenie „map”, na których znajdują się obszary terminów związanych bardzo blisko, związanych luźniej, słabo lub prawie wcale. W bibliometrii często stosuje się też metodę *co-word* łącznie z innymi technikami analizy struktur sieciowych oraz technikami statystycznymi, które pozwalają na pogłębienie analizy wzorców wykorzystywania określonej terminologii (Neef & Corley, 2009, 659; zob. też: Marshakova Shaikevich, 2005).

Koncepcję współwystępowania słów kluczowych wykorzystał Chaomei Chen (2006) w ramach modelowania szybko zachodzących zmian w problematyce frontów badawczych i zarazem w ich tzw. „bazie wiedzy” (ang. *intellectual base*). Bazę tę stanowiły publikacje cytowane przez artykuły tworzące front badawczy. Rozróżnienie to zostało zaczerpnięte przez Chena od Perssona (1994) (Chen, 2006, 360; Persson, 1994, 31). Zmiany zachodzące w obszarze frontu są na ogół powiązane ze zmianami w jego bazie intelektualnej. Dynamika wzajemnej zależności frontu i jego bazy, przesunięcia w leżącym u podstaw frontu przepływie wiedzy mogą zostać uchwycone właśnie poprzez analizę szczegółowej, dyscyplinarnej terminologii i jej (nawet nieznaczących) przemian (Chen, 2006, 360–361). Chen proponował ponadto włączenie analiz typu *co-word* do specjalnego, zaprojektowanego pod kątem identyfikacji nowo powstających trendów badawczych oprogramowania nazwanego *CiteSpace II*. Zasadą jego działania jest analiza sieci cytowań publikacji w czasie z wykorzystaniem zautomatyzowanych algorytmów oraz twierdzeń i wskaźników pochodzących z teorii grafów. Elementem związanym ściśle z analizą terminologii jest detekcja tzw. „eksplozji” pewnych wiązek terminów/słów kluczowych (ang. *burst detection*). Przez „eksplozję” rozumie się tutaj nagły, gwałtowny wzrost częstotliwości wykorzystywania przez badaczy z danej specjalności określonych słów kluczowych czy fraz, co sygnalizuje wzrost zainteresowania i nasilenie badań prowadzonych nad daną problematyką. Jest to zjawisko znamionujące raczej zmianę orientacji frontu niż jego ogólny, stabilniejszy profil tematyczny (Chen, 2006, 364–365, 374–375). W *CiteSpace II* kluczowe frazy/terminy wykryte w takich wiązках stosuje się poza tym jako „etykiety” (ang. *labels*) frontów badawczych. Ma to zapewnić bardziej precyzyjny opis charakteru frontu niż prosta agregacja terminów najczęściej występujących w klastrach publikacji uzyskanych z wykorzystaniem służących temu, tradycyjnych metod (Chen, 2006, 375). Chen zaznaczył ponadto, że podejście oparte na badaniach terminologii (ang. *word-profile approach*) nie jest też pozbawione pewnych wad. Jedną z nich jest wspomniana już i stosowana niekiedy prosta agregacja terminów, które powtarzają się najczęściej w pewnych klastrach. Znaczenie takich terminów może być bardzo zróżnicowane, „rozbiegające się” wielokierunkowo, co wymaga znacznego wkładu własnego osoby analizującej terminy, aby powiązać je w sensowny sposób. Po drugie, operacja nadawania klastrów nazw często skutkuje nazewnictwem zbyt szerokim (ogólnym), aby mogło być ono użyteczne. Terminy skojarzone ściśle z trendami nowo powstającymi mogą zostać niejako przesłonięte przez terminy należące do obszarów z dłuższą tradycją, lepiej

ugruntowanych, ogólniej – obszarów w danym klastrze pierwszoplanowych (ang. *persistent*) (Chen, 2006, 364). Podejściu tekstualnemu stawiano też podobne zarzuty związane z jego niewystarczającą zdolnością do odróżniania prac należących do odmiennych specjalności, czego przyczyn upatrywano w niskiej wrażliwości metody na różne konteksty posługiwania się zbliżonym nazewnictwem i wynikające stąd „przecenianie” (ang. *overestimate*) stopnia podobieństwa tych prac (Glänzel & Thijs, 2012, 400).

Woo Hyoung Lee (2008) analizował przy użyciu techniki współwystępowania terminów funkcjonującą w informatyce subdyscyplinę bezpieczeństwa informacji. Jego podejście polegało na stworzeniu „mapy wiedzy” (ang. *knowledge map*) dla wybranego 2006 r., która składała się z terminów występujących równocześnie w różnych publikacjach. Uzyskaną w ten sposób mapę tworzyły „gniazda” (centra – ang. *hubs*) terminów związanych ze sobą ściślej i luźniej. Siła tego związku determinowała położenie terminów na mapie (Lee, 2008, 509). Taka struktura jest bliskim odpowiednikiem klastrów publikacji, które również są zbiorami słabiej i silniej połączonych elementów (leżących tak wewnątrz jak i na zewnątrz konkretnego klastra). Dalsza analiza wykorzystywała parametry struktur sieciowych, takie jak „centralność” (ang. *centrality*) i „wskaźnik znaczenia w sieci” lub „miara stopnia pośredniczenia” (ang. *betweenness*) (Lee, 2008, 509, 522). Są to wskaźniki typowe dla analiz klastrów złożonych z wierzchołków i połączeń między nimi („krawędzi”). Konstrukcja metody Lee opierała się na odnalezieniu centrum zbudowanej metodą *co-word* sieci. Centrum to było grupą blisko związanych wierzchołków o niespotykanie dużej liczbie połączeń (współwystąpień) z innymi wierzchołkami w całej sieci. Centrum charakteryzowało się więc wysoką miarą „centralności”, co można porównać do wysokiej siły wpływu, silnego oddziaływania na całą subdyscyplinę (Lee, 2008, 511–512). Centrum było także miejscem o wysokim stopniu „pośredniczenia” na mapie, tj. leżało na drodze łączącej wiele różnych punktów rozsianych w całej strukturze (Lee, 2008, 522). Z użyciem tych parametrów określono właściwości odnalezionego w danym roku centrum, aby następnie „cofnąć się w czasie” (tj. przeanalizować tą samą metodą grupę publikacji wydanych wcześniej), aż do okresu w którym aktualne centrum dopiero zaczynało się formować. Należało teraz ponownie zbudować analogiczną mapę wiedzy dla tamtego (wcześniejszego) okresu i sprawdzić, jak wyglądały parametry aktualnego centrum (tj. centrum z 2006 r.) w okresie jego „młodości”. Jeżeli na aktualnej mapie uda się odnaleźć takie rejony, które mają podobne lub identyczne właściwości strukturalne, jakie wykazywało aktualne centrum w swojej młodości, byłyby one właśnie tymi, zdefiniowanymi leksykalnie rejonami, które mają największe szanse, aby stać się centrum w przyszłości (Lee, 2008, 509). Lee wytypował w ten sposób kryptografię, jako najbardziej perspektywiczną subdyscyplinę informatyki, zarazem zaznaczył jednak, że efektywność jego metody jest trudna (a nawet niemożliwa) do natychmiastowego zweryfikowania (Lee, 2008, 519, 521, 524).

Mark Neef i Elizabeth Corley (2009) posłużyli się metodą współwystępowania terminów zaczerpniętych z tytułów publikacji do rozpoznania trendów badawczych i ewolucji, powstawania i zanikania obszarów problemowych w obrębie ekologii pomiędzy rokiem 1970 a 2005 (Neef & Corley, 2009, 658, 662–663). Jak zaznaczyli Neef i Corley, w przeciwieństwie do niektórych innych przedsięwzięć wykorzystujących tę samą metodę, w ich pracy analizowano nie tylko zestawy słów współwystępujących najczęściej w dłuższych okresach, ale też podjęto próbę wyodrębnienia trendów krótkotrwałych, podlegających gwałtownym zmianom i obecnych niejako na tle trendów stabilniejszych i długotrwałych.

Zamiarem było tutaj wyróżnienie najważniejszych tematów badawczych dominujących w różnym czasie i prześledzenie tego, jak zmieniała się ich rola dla dyscypliny ekologii jako całości (Neef & Corley, 2009, 660). Zastosowano ponadto normalizację słownictwa, polegającą m.in. na scaleniu słów występujących w liczbie pojedynczej i mnogiej, ujednoczeniu synonimicznych wersji różnych wyrażzeń, ujednoczeniu wersji różniących się literacją (np. pisownia amerykańska i brytyjska), albo słów pisanych w wersji z myślnikiem lub bez (Neef & Corley, 2009, 662). Cały okres badania podzielono na dwa podokresy (1970–1987 i 1988–2005), a następnie podzielono częstość wystąpień każdego (znormalizowanego już) słowa w danym roku przez całkowitą liczbę słów użytych we wszystkich tytułach publikacji ogłoszonych w tym samym roku (Neef & Corley, 2009, 664). W dalszej kolejności Neef i Corley posłużyli się indeksem podobieństwa Callona i in. z 1986 r. oraz algorytmem grupującym (tzw. algorytm Warda – zob. np.: Mojena, 2006). Uzyskano klastry zawierające finalnie od 2 do 9 słów i zastosowano do nich miary „centralności” (jak centralna dla całej ekologii jest tematyka danego klastra?) i „gęstości” (ang. *density*). Jeżeli klaster jest „gęsty” oznacza to, że autorzy używając jednego ze słów należących do klastra, bardzo często używają też jednocześnie pozostałych słów składających się na ten klaster (Neef & Corley, 2009, 665–666). Każdy uzyskany klaster umieszczono następnie na diagramie pokazującym zależność „centralności” od „gęstości”. Klastry o wysokich wskaźnikach obu tych miar stanowiły centra dla swoich dyscyplin reprezentując koncepty spójne i dobrze rozwinięte. Klastry o niskiej gęstości i zarazem wysokiej centralności reprezentowały natomiast zagadnienia ważne dla dyscypliny, ale jeszcze słabo rozwinięte – zagadnienia nowo wyłaniające się (ang. *emerging*) (Neef & Corley, 2009, 667–670). Dodatkowo poddano szczegółowej analizie słowa wykazujące największe wahania (wzrost i spadek częstości występowania) w obu wyróżnionych podokresach i powtórzono dla nich cały proces badawczy streszczony powyżej (Neef & Corley, 2009, 673–675). Do podstawowych ograniczeń tej metody wskazano możliwą wieloznaczność terminów, ich zależność od kontekstu i zmienność znaczenia terminów technicznych w czasie. Przywołano w tym kontekście opinię Loeta Leydesdorffa z 1997 r., według której język jest tworem zbyt dynamicznym, aby mógł stać się narzędziem efektywnej analizy typu *co-word* w zastosowaniach naukoznawczych (Neef & Corley, 2009, 667). Wskazano również na możliwą nieadekwatność tytułów niektórych publikacji do ich zawartości treściowej, a także na problem pojawiający się w trakcie analizy niewielkich, wąskich pól badawczych. Polega on na niemożności uchwycenia trendów „szerokozakresowych”, przenikających więcej niż jedną subdyscyplinę lub specjalność. Problem ten może być o tyle istotny, że to właśnie trendy tego rodzaju stanowią na ogół przedmiot zainteresowania polityki naukowej i mogą nieść ze sobą następstwa sięgające poza sam obszar czysto akademicki (Neef & Corley, 2009, 659–660, 667–678; zob. też: Van Den Besselaar & Heimeriks, 2006, 380).

Odmienne podejście do problemu prognozowania rozwoju nauki i technologii przyjęli Takao Furukawa i in. (2015). Skupili się oni na międzynarodowych konferencjach jako na wydarzeniach, podczas których prezentuje się najnowsze odkrycia i osiągnięcia. Konferencje naukowe (często cykliczne) zapewniają – zdaniem Furukawy i in. – najszybszy sposób transferu informacji, a ich podstawową funkcją jest przesuwanie granic wiedzy (ang. *advance the frontiers of knowledge*) (Furukawa et al., 2015, 281). Wykorzystana przez zespół autorów metoda była standardową analizą typu *co-word* zastosowaną do abstraktów wystąpień konferencyjnych dostępnych w Internecie. Na jej podstawie stworzono „sieć

czasową” (ang. *temporal network*) sesji konferencyjnych, która pozwalała zaobserwować ewolucję tematyki poruszanej podczas sesji i wytypować problematykę aktualnie najbardziej wpływową (Furukawa et al., 2015, 283–288). To, co wydaje się najistotniejsze w przypadku metody Furukawy i in. to fakt, że w ich ocenie tematyka sesji konferencyjnych ma silne tendencje do wskazywania przyszłościowych perspektyw rozwoju dyscyplin. Rada programowa konferencji ustala zazwyczaj wiodącą tematykę i selekcjonuje ważniejsze problemy, które powinny zostać omówione, co odzwierciedla sposób postrzegania przez społeczność naukową tego, co w danej dyscyplinie można uważać za pionierskie (Furukawa et al., 2015, 291). Problematyczną stroną metody zespołu autorów mogą stanowić natomiast przypadki, w których na konferencję zostaje zgłoszonych bardzo dużo referatów, które zostają wtedy „sztucznie” rozdysponowane pomiędzy różne sesje. Może to wywołać sytuację, w której w przybliżeniu ta sama tematyka jest omawiana na odrębnych sesjach (noszących inne tytuły) lub odwrotnie, w obrębie jednej sesji konferencyjnej wygłaszane referaty mogą nie pokrywać nawet w przybliżeniu jednakowych zagadnień (Furukawa et al., 2015, 292). Xiaoling Sun i in. (2016, 759) wskazują na analogiczny problem, na jaki napotyka metoda identyfikacji współwystępujących słów kluczowych. Polega on na niezdolności uchwycenia podobieństwa takich publikacji, w których autorzy nie posługiwaliby się podobnym słownictwem, a mimo to zajmowałiby się bliskimi sobie zagadnieniami.

Tianlong Zheng i in. (2016) wykorzystali metodę współwystępowania terminów do analizy aktualnych tendencji badawczych oraz najnowszych prądów intelektualnych (ang. *hotspots*) w międzydyscyplinarnej specjalności zajmującej się badaniem mikro- i nanopercherzyków (ang. *micro/nano-bubbles*) (Zheng et al., 2016, 55). Zbadano ilościowe wzorce (wzrost i spadek) współwystępowania 30 najczęściej pojawiających się słów kluczowych w publikacjach z lat 1991–2014 wyszukanych w Web of Science (WoS). Dla uwydatnienia wymiaru czasowego analizy cały ten okres podzielono na cztery sześciolate przedziały i dla każdego z nich skonstruowano mapę (ang. *co-word network*) (Zheng et al., 2016, 55, 63, 65). Innowacyjnym elementem metody zespołu Zhenga było wykorzystanie nie tylko autorskich słów kluczowych i słów widniejących w tytułach publikacji, ale także tzw. słów „dodatkowych” (ang. *KeyWords Plus*) dostępnych w bazie WoS. Słowa te wyróżniają się pod względem źródła pochodzenia, którym są tytuły publikacji przywoływanych w bibliografiach załącznikowych artykułu, do którego słowa te są dodawane przez specjalny algorytm (Zheng et al., 2016, 65). Metoda, która wyróżnia się uwzględnieniem w analizie „dodatkowych” słów kluczowych (obok słów autorskich i słów z tytułów publikacji) została określona jako „analiza klastrów słownych” (ang. *word cluster analysis*) (Zheng et al., 2016, 65). Zdaniem niektórych badaczy słowa „dodatkowe” są ukierunkowane bardziej niż słowa autorskie na nowatorskość i innowacyjność kierunków podejmowanych i referowanych w publikacjach badań, niż na ich postać tradycyjną i głębiej zakorzenioną w nauce (Li et al., 2009, 700; zob. też: Li et al., 2011, 17).

8. Zakończenie

Podsumowując stwierdzenia i fakty streszczone w ramach powyższych podrozdziałów, można przedstawić kilka ogólnych wniosków dotyczących postaci, korzeni i perspektyw zarysowujących się przed nurtem prognostycznym istniejącym we współczesnej bibliometrii.

Przed wszystkim warto podkreślić, że badania pionierów, prowadzone głównie w latach 60. i 70. XX w. nadały początek i w dużym stopniu ukształtowały aktualną formę, jaką przybiera większość prób oceny i przewidywania dalszego postępu badań naukowych. Są one wciąż żywotne i inspirują teraźniejsze wysiłki badaczy, co świadczy o ich ponadczasowym charakterze i uwidacznia praktyczne potrzeby, na które stanowią one racjonalną odpowiedź. Nie znaczy to zarazem, że metody określone w niniejszym opracowaniu jako tradycyjne, pozbawione są wszelkich wad. Przeciwnie, różnorakie braki metodologiczne zostały dostrzeżone niemal równocześnie z pojawieniem się i pierwszymi próbami stosowania tych metod w praktyce, a krytyka poszczególnych odmian bibliometrycznej metodologii prognostycznej i dążenie do jej udoskonalenia stanowią integralną część i etap rozwoju całej tej gałęzi naukoznawstwa. W tym kontekście należy również zwrócić uwagę na fakt, iż najważniejszymi i najczęściej powtarzającymi się w praktyce badawczej problemami czy zarzutami stawianymi przed metodami podstawowymi są niewątpliwie:

- Arbitralność ustalania granicznych progów siły współcytowania oraz arbitralność definiowania publikacji wysoko cytowanych. Nabiera to obecnie szczególnego znaczenia w czasach współczesnych, bowiem w ogólnosiwiatowych indeksach cytowań widnieją publikacje cytowane po kilkanaście, a nawet kilkadziesiąt tysięcy razy.
- Metoda grupowania jednopiętowego wykazuje silną tendencję do tworzenia klastrów niejednorodnych tematycznie, co w warunkach współczesnej, wysokiej interdyscyplinarności wielu dyscyplin naukowych powoduje trudności z przyporządkowaniem tak zbudowanego klastra do właściwej dziedziny/dyscypliny.
- Z drugiej strony w obrębie bibliometrii i nauk komputerowych ma się obecnie do czynienia z tak dużą różnorodnością zarówno algorytmów grupujących jak i kryteriów podobieństwa publikacji, w oparciu o które przeprowadza się operacje grupowania, że dokonanie i uzasadnienie konkretnego wyboru również stało się poważnym wyzwaniem. Znajomość wszystkich konsekwencji dokonanego wyboru i możliwych, wywołanych przezeń nieprawidłowości wymaga z jednej strony dogłębnej znajomości ściśle technicznych aspektów działania algorytmów grupujących, z drugiej zaś dostępu do badań porównawczych, który to dostęp – jak postarano się zaznaczyć w drugiej części niniejszego artykułu – jest aktualnie raczej ograniczony.
- Powyższe uwagi odnoszą się również do rozmaitych procedur normalizacyjnych wprowadzanych w obszar prognostycznych badań empirycznych na etapie ich projektowania.
- Często podkreślanym problemem w stosowaniu metody współcytowań jest ponadto fakt istnienia opóźnień w pojawianiu się kolejnych cytowań, które wynikają tak z konieczności dotarcia idei przenoszonych przez publikacje do grona ich odbiorców i wywarcia wpływu na ich myśl, jak i z faktu istnienia – nierzadko bardzo długich – opóźnień w samym procesie publikacyjnym (w cyklu wydawniczym).
- Problematyczne pozostaje w dalszym ciągu także pytanie o maksymalny dopuszczalny rozmiar klastra, oraz o jego rozmiar minimalny, tj. o minimalną liczbę składających się na niego publikacji, która pozwoliłaby na uznanie go za twór samodzielny.
- Szereg problemów metodologicznych wynika ponadto z rozbieżności opinii specjalistów nt. tego, które ze źródeł danych stanowią najlepszą odpowiedź na potrzeby związane z monitorowaniem dyscyplin ze względu na tempo i zakres ich rozwoju, tj. jak i w jakich dyscyplinach należy traktować artykuły z czasopism, monografie

- i rozdziały w monografiach, dokumenty patentowe, materiały konferencyjne itd. Innym obliczem tego samego problemu są również kwestie fizycznej dostępności do różnego rodzaju danych źródłowych, które stanowią elementarny wymóg samej możliwości prowadzenia badań, a także warunkują do pewnego stopnia ich jakość.
- Na tym samym poziomie mankamentów metodologicznych usytuowane są kwestie dotyczące wyboru spośród wielu możliwych poziomów agregacji danych empirycznych, jak również kwestie związane z cytowalnością typową dla różnych form wydawniczych oraz różnych dyscyplin czy obszarów badawczych.
 - W odniesieniu do metody generowania powiązań bibliograficznych, w tekście niniejszego artykułu wskazano na jeden z ciekawszych problemów, którym jest zjawisko starzenia się literatury cytowanej stanowiące przeszkodę w porównywaniu publikacji wydanych w różnym czasie pod względem tego, jak wyglądają tworzone w ramach tej metody fronty badawcze (zob. podrozdz. 6).
 - W odniesieniu do metody współwystępowania słów kluczowych można dodatkowo w tym miejscu przypomnieć, że głównymi wysuwanyymi wobec niej zastrzeżeniami są trudności związane z semantyką terminów naukowych – np. zależność ich znaczenia od kontekstu, wieloznaczność, nieadekwatność do niektórych koncepcji, trudności w identyfikowaniu trendów szerokok zakresowych za pośrednictwem terminów o wąskim i wysoce wyspecjalizowanym znaczeniu.

Ponadto ważne w kontekście przedstawionej pierwszej części całej analizy piśmiennictwa, jak i jej części drugiej, która ukaże się w następnym numerze ZIN, wydaje się nawiązujące do punktu 3 spostrzeżenie, według którego istnieje silna korelacja pomiędzy prognostycznymi metodami bibliometrycznymi, a algorytmicznymi metodami grupowania publikacji w spójne tematycznie klastry reprezentujące wąskie, dyscyplinarne „subsposóbności” czy mówiąc inaczej, obszary problemowe poddawane dalszej analizie ilościowej o różnie skonstruowanych profilach. Korelacja ta jest dostrzegalna we wszystkich wyróżnionych podgrupach publikacji, które omówiono w odrębnych podrozdziałach. Nasuwa to myśl o konieczności rozważenia jeszcze jednego, fundamentalnego problemu rozgraniczenia obszarów problemowych czy specjalności, poprowadzenia linii podziału pomiędzy nimi (tzw. delineacja) jeszcze przed przystąpieniem do badań zorientowanych prognostycznie. Innymi słowy, aby móc powiedzieć cokolwiek o perspektywie dalszego rozwoju dyscypliny, konieczne jest jej wcześniejsze zdefiniowanie i wyodrębnienie w sensie wyselekcjonowania wszystkich i zarazem tylko tych publikacji, które faktycznie do niej przynależą i decydują o jej rzeczywistym kształcie (strukturze). Popęlenie błędu na tym etapie procedury badawczej skutkuje więc zawsze zniekształceniem ostatecznego wyniku badania i wyprowadzeniem mylnych wniosków co do tempa lub kierunku rozwoju konkretnej dyscypliny. Problem ten jest jedną z istotniejszych kwestii, którą zajmują się przedstawiciele bibliometrii nie tylko w związku z przewidywaniem rozwoju nauki, ale też w związku z zagadnieniami z zakresu np. strategii wyszukiwania informacji lub ewaluatywnej analizy cytowań. Wskazuje się w tym kontekście na wady istniejących systemów klasyfikacyjnych (np. systemu przyjętego w *Journal Citation Reports* lub w indeksie cytowań *Scopus*) funkcjonujących ponadto głównie na poziomie czasopism, a nie indywidualnych publikacji (zob. np.: Waltman & Van Eck, 2012, 2378–2379). Fakt ten rzuca dodatkowe światło na potrzebę dalszego dopracowywania istniejących metod prognostycznych, która stanowi jeden z wątków przewodnich drugiej części niniejszego artykułu.

Bibliografia

- Adler, R.; Ewing, J.; Taylor, P. (2009). Citation Statistics. A Report from the International Mathematical Union (IMU) in Cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS). *Statistical Science*, 24(1), 1–14.
- Boyack, K.W.; Börner, K.; Klavans, R. (2009). Mapping the Structure and Evolution of Chemistry Research. *Scientometrics*, 79(1), 45–60.
- Boyack, K.W.; Klavans, R. (2010). Co-Citation Analysis, Bibliographic Coupling and Direct Citation: Which Citation Approach Represents the Research Front Most Accurately? *Journal of the American Society for Information Science and Technology*, 61(12), 2389–2404.
- Boyack, K.W.; Klavans, R. (2014). Creation of Highly Detailed, Dynamic, Global Model and Map of Science. *Journal of the Association for Information Science and Technology*, 65(4), 670–685.
- Braun, T.; Schubert, A.; Kostoff, R. (2000). Growth and Trends of Fullerene Research as Reflected in Its Journal Literature. *Chemical Reviews*, 100(1), 23–37.
- Briotta Parolo, P.; Pan, R.K.; Ghosh, R.; Huberman, B.A.; Kaski, K.; Fortunato, S. (2015). Attention Decay in Science. *Journal of Informetrics*, vol. 9(4), 734–745.
- Chen, Ch. (2006). CiteSpace II: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359–377.
- Chen, S-H.; Huang, M-H.; Chen, D-Z. (2012). Identifying and Visualizing Technology Evolution: A Case Study of Smart Grid Technology. *Technological Forecasting and Social Change*, 79(6), 1099–1110.
- Diodato, V. (1994). *Dictionary of Bibliometrics*. New York: London: Norwood: The Haworth Press.
- Dorta-González, P.; Dorta-González, M-I. (2013). Impact Maturity Times and Citation Time Windows: The 2-Year Maximum Journal Impact Factor. *Journal of Informetrics*, 7(3), 593–602.
- Ena, O.; Mikova, N.; Saritas, O.; Sokolova, A. (2016). A Methodology for Technology Trend Monitoring: The Case of Semantic Technologies. *Scientometrics*, 108(3), 1013–1041.
- Érdi, P.; Makovi, K.; Somogyvári, Z.; Strandburg, Ka.; Tobochnik, J.; Volf, P.; Zalányi, L. (2013). Prediction of Emerging Technologies Based on Analysis of the US Patent Citation Network. *Scientometrics*, 95(1), 225–242.
- Furukawa, T.; Mori, K.; Arino, K.; Hayashi, K.; Shirakawa, N. (2015). Identifying the Evolutionary Process of Emerging Technologies: A Chronological Network Analysis of World Wide Web Conference Sessions. *Technological Forecasting and Social Change*, 91, 280–294.
- Garcia, E.K. (2015). *Cosine Similarity Tutorial* [online]. Department of Electrical Engineering, University of Washington, [06.11.2016], <http://www.minerazzi.com/tutorials/cosine-similarity-tutorial.pdf>
- Garfield, E. (2004). Historiographic Mapping of Knowledge Domains Literature. *Journal of Information Science*, 30(2), 119–145.
- Garfield, E.; Pudovkin, A.; Istomin, V.S. (2002). Algorithmic Citation-Linked Historiography-Mapping the Literature of Science. *Proceedings of the American Society for Information Science and Technology*, 39(1), 14–24.
- Garfield, E.; Pudovkin, A.; Istomin, V.S. (2003). Why Do We Need Algorithmic Historiography? *Journal of the American Society for Information Science and Technology*, 54(5), 400–412.
- Garfield, E.; Sher, I.; Torpie, R. (1964). *The use of citation data in writing the history of science. A final report of research for the Air Force Office of Scientific Research under contract AF 49(638)-1256*. Philadelphia, Pennsylvania, USA: Institute for Scientific Information, Inc.
- Glänzel, W. (2012). Bibliometric Methods for Detecting and Visualizing Emerging Research Topics. *El Profesional de la Información*, 21(1), 194–201.
- Glänzel, W.; Schoepflin, U. (1995). A Bibliometric Study on Ageing and Reception Processes of Scientific Literature. *Journal of Information Science*, 21(1), 37–53.

- Glänzel, W.; Schoepflin, U. (1999). A Bibliometric Study of Reference Literature in the Sciences and Social Sciences. *Information Processing and Management*, 35(1), 31–44.
- Glänzel, W.; Thijs, B. (2012). Using 'Core Documents' for Detecting and Labelling New Emerging Topics. *Scientometrics*, 91(2), 399–416.
- Guo, H.; Weingart, S.; Börner, K. (2011). Mixed-Indicators Model for Identifying Emerging Research Areas. *Scientometrics*, 89(1), 421–435.
- Hertzfel, D. (1987). Bibliometrics, History of the Development of Ideas In. In: A. Kent, H. Lancour, W. Nasri (eds.). *Encyclopedia of Library and Information Science*, vol. 42, Suppl. 7. New York: Marcel Dekker, Inc., 144–219.
- Huang, A. (2008). Similarity Measures for Text Document Clustering. In: J. Holland, A. Nicholas, D. Brignoli (eds.). *Proceedings of the Sixth New Zealand Computer Science Research Student Conference (NZCSRSC 2008)*. Christchurch, New Zealand, April 14–18, 2008, 49–56 [online]. University of Canterbury, [30.08.2017], https://nzcsrsc08.canterbury.ac.nz/site/proceedings/NZCSRSC_2008_Proceedings.pdf
- Huang, M-H.; Chang, C-P. (2014). Detecting Research Fronts in OLED Field Using Bibliographic Coupling with Sliding Window. *Scientometrics*, 98(3), 1721–1744.
- Huang, M-H.; Chang, C-P. (2015). A Comparative Study on Detecting Research Fronts in the Organic Light-Emitting Diode (OLED) Field Using Bibliographic Coupling and Co-Citation. *Scientometrics*, 102(3), 2041–2057.
- Hummon, N.P.; Doreian, P. (1989). Connectivity in a Citation Network: The Development of DNA Theory. *Social Networks*, 11(1), 39–63.
- Jarneving, B. (2005). A Comparison of Two Bibliometric Methods for Mapping of the Research Front. *Scientometrics*, 65(2), 245–263.
- Kajikawa, Y.; Yoshikawa, J.; Takeda, Y.; Matsushima, K. (2008). Tracking Emerging Technologies in Energy Research: Toward a Roadmap for Sustainable Energy. *Technological Forecasting and Social Change*, 75(6), 771–782.
- Kessler, M.M. (1963). Bibliographic Coupling Between Scientific Papers. *American Documentation*, 14(1), 10–25.
- KPB (2011). *Krajowy Program Badań. Założenia polityki naukowo-technicznej i innowacyjnej państwa. Krajowy Program Badań opracowany na podstawie art. 4 ust. 1 ustawy o zasadach finansowania nauki. Załącznik do uchwały nr 164/2011 Rady Ministrów z dnia 16 sierpnia 2011 r.* [online]. Ministerstwo Nauki i Szkolnictwa Wyższego, [30.08.2017], http://www.ncbir.pl/gfx/ncbir/user-files/_public/bip/20110816_kpb.pdf
- Lee, W.H. (2008). How to Identify Emerging Research Fields Using Scientometrics: An Example in the Field of Information Security. *Scientometrics*, 76(3), 503–525.
- Lee, L-Ch.; Lee, Y-Y.; Liaw, Y-Ch. (2012). Bibliometric Analysis for Development of Research Strategies in Agricultural Technology: The Case of Taiwan. *Scientometrics*, 93(3), 813–830.
- Leydesdorff, L. (1987). Various Methods for the Mapping of Science. *Scientometrics*, 11(5–6), 295–324.
- Li, J.; Zhang, Y.; Wang, X.; Ho, Y-S. (2009). Bibliometric Analysis of Atmospheric Simulations Trends in Meteorology and Atmospheric Science Journals. *Croatica Chemica Acta*, 82(3), 695–705.
- Li, J.; Wang, M-H.; Ho, Y-S. (2011). Trends in Research on Global Climate Change: A Science Citation Index Expanded-Based Analysis. *Global and Planetary Change*, 77(1), 13–20.
- Liu, Ch.; Gui, Q. (2016). Mapping Intellectual Structures and Dynamics of Transport Geography Research: A Scientometric Overview from 1982 to 2014. *Scientometrics*, 109(1), 159–184.
- Liu, J.S.; Lu, L.Y.Y. (2012). An Integrated Approach for Main Path Analysis: Development of the Hirsh Index as an Example. *Journal of the American Society for Information Science and Technology*, 63(3), 528–542.
- Ma, V.C.; Liu, J.S. (2016). Exploring the Research Fronts and Main Paths of Literature: A Case Study of Shareholder Activism Research. *Scientometrics*, 109(1), 33–52.

- Mabe, M.A. (2010). Scholarly Communication: A Long View. *New Review of Academic Librarianship*, 16(Suppl. 1), 132–144.
- Marshakova Shaikevich, I. (1973). System of Document Connections Based on References. *Nauchno-Tekhnicheskaya Informatsiya. Seriya 2, Informatsionnye Protsessy i Sistemy*, 6(2), 3–8.
- Marshakova Shaikevich, I. (2005). Bibliometric Maps of Field of Science. *Information Processing and Management*, 41(6), 1534–1547.
- Martín-Martín, A.; Orduna-Malea, E.; Allyón, J.M.; Delgado López-Cózar, E. (2016). *The Counting House: Measuring Those Who Count. Presence of Bibliometrics, Scientometrics, Informetrics, Webometrics and Altmetrics in the Google Scholar Citations, ResearcherID, ResearchGate, Mendeley & Twitter* [online]. EC3 Working Papers, vol. 21, 1–60, [30.08.2017], <https://arxiv.org/ftp/arxiv/papers/1602/1602.02412.pdf>
- Meadows, J. (2000). The Growth of Journal Literature. In: B. Cronin & H. B. Atkins (eds.). *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*. Medford, NJ: Information Today, Inc., 87–108.
- MNiI (2004). *Założenia polityki naukowej, naukowo-technicznej i innowacyjnej państwa. Projekt 1.10.2004 r.* [online]. Ministerstwo Nauki i Informatyzacji, [30.08.2017], http://www.nauka.gov.pl/g2/oryginal/2013_05/ac91da4d11aad3d93649c13ed181b70a.pdf
- Moed, H.F.; van Leeuwen, T.N.; Reedijk, J. (1998). A New Classification System to Describe the Ageing of Scientific Journals and Their Impact Factors. *Journal of Documentation*, 54(4), 387–419.
- Mojena, R. (2006). Ward's Clustering Algorithm. In: S. Kotz, C. B. Read, N. Balakrishnan, B. Vidakovic (eds.). *Encyclopedia of Statistical Sciences* [online]. Wiley Online Library, [11.11.2016], <http://onlinelibrary.wiley.com/doi/10.1002/0471667196.ess2887.pub2/abstract>
- Muraszkiewicz, M. (2014). An Essay on Information Overload. *Zagadnienia Informacji Naukowej. Studia Informacyjne*, 52, 1(103), 7–18.
- Nowak, P. (2006). *Bibliometria. Webometria. Podstawy, wybrane zastosowania*. Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.
- Neef, M.W.; Corley, E.A. (2009). 35 Years and 160,000 Articles: A Bibliometric Exploration of the Evolution of Ecology. *Scientometrics*, 80(3), 657–682.
- Persson, O. (1994). The Intellectual Base and Research Fronts of JASIS 1986–1990. *Journal of the American Society for Information Science*, 45(1), 31–38.
- Pindłowa, W. (1988). Bibliometria i jej znaczenie dla badań nad książką. *Studia o Książce*, T. 17, 301–327.
- Pindłowa, W. (1994). *Informetria w nauce o informacji. Metody i problemy*. Kraków: Towarzystwo Autorów i Wydawców Prac Naukowych „Universitas”.
- Potiopa, P. (2011). Metody i narzędzia automatycznego przetwarzania informacji tekstowej i ich wykorzystanie w procesie zarządzania wiedzą. *Automatyka*, T. 15(2), 409–419.
- Price, D. de Solla (1967). *Mała Nauka – Wielka Nauka*. Warszawa: PWN.
- Pritchard, A. (1969). Statistical Bibliography or Bibliometrics? *Journal of Documentation*, 25(4), 348–349.
- Rousseau, R. (2014). Library Science: Forgotten Founder of Bibliometrics. *Nature*, 510(7504), 218.
- Salton, G.; Bergmark, D. (1979). A Citation Study of Computer Science Literature. *IEEE Transactions on Professional Communication*, 22(3), 146–158.
- Sasson, E.; Ravid, G.; Pliskin, N. (2015). Improving Similarity Measures of Relatedness Proximity: Toward Augmented Concept Maps. *Journal of Informetrics*, 9(3), 618–628.
- Seweryn, A.; Swoboda, I. (2013). *Cytowania w polskich bazach bibliograficznych* [online]. XII Krajowe Forum Informacji Naukowej i Technicznej: Społeczeństwo – Informacja – Innowacje. Wyzwania Ery Cyfrowej. Zakopane, 24–27 września 2013 r., [30.08.2017], http://www.ptin.org.pl/konferencje/12forum/repozytorium/Seweryn_Swoboda.pdf
- Shibata, N.; Kajikawa, T.; Matsushima, K. (2007). Topological Analysis of Citation Networks to Discover the Future Core Articles. *Journal of the American Society for Information Science and Technology*, 58(6), 872–882.

- Shibata, N.; Kajikawa, Y.; Sakata, I. (2012). Link Prediction in Citation Networks. *Journal of the American Society for Information Science and Technology*, 63(1), 78–85.
- Shibata, N.; Kajikawa, Y.; Takeda, Y.; Matsushima, K. (2008). Detecting Emerging Research Fronts Based on Topological Measures in Citation Networks of Scientific Publications. *Technovation*, 28(11), 758–775.
- Shibata, N.; Kajikawa, Y.; Takeda, Y.; Matsushima, K. (2009). Comparative Study on Methods of Detecting Research Fronts Using Different Types of Citation. *Journal of the American Society for Information Science and Technology*, 60(3), 571–580.
- Skalska-Zlat, M. (1988). Bibliometria – pojęcia, metody, kierunki badań. *Roczniki Biblioteczne*, T. XXXII (2), 259–283.
- Skalska-Zlat, M. (1993). *Bibliometryczne badania rozwoju dyscypliny naukowej*. Seria: Bibliotekoznawstwo XVII. Wrocław: Wydaw. UWr.
- Skalska-Zlat, M. (1999). Bibliografia jako odbicie rozwoju nauki. *Zagadnienia Naukoznawstwa*, 1(139), 57–63.
- Small, H. (1973). Co-Citation in the Scientific Literature: A New Measure of the Relationship Between Two Documents. *Journal of the American Society for Information Science*, 24(4), 265–269.
- Small, H. (2006). Tracking and Predicting Growth Areas in Science. *Scientometrics*, 68(3), 595–610.
- Small, H.; Boyack, K.W.; Klavans, R. (2014). Identifying Emerging Topics in Science and Technology. *Research Policy*, 43(8), 1450–1467.
- Small, H.; Griffith, B.C. (1974). The Structure of Scientific Literatures I: Identifying and Graphing Specialties. *Science Studies*, 4(1), 17–40.
- Small, H.; Sweeney, E. (1985). Clustering the Science Citation Index Using Co-Citations: I. A Comparison of Methods. *Scientometrics*, 7(3–6), 391–409.
- Sosińska-Kalata, B. (2002). *Klasyfikacja. Struktury organizacji wiedzy, piśmiennictwa i zasobów informacyjnych*. Warszawa: Wydaw. SBP.
- Soós, S. (2014). Age-Sensitive Bibliographic Coupling Reflecting the History of Science: The Case of the Species Problem. *Scientometrics*, 98(1), 23–51.
- Sun, X.; Ding, K.; Lin, Y. (2016). Mapping the Evolution of Scientific Fields Based on Cross-Field Authors. *Journal of Informetrics*, 10(3), 750–761.
- Takeda, Y.; Kajikawa, Y. (2009). Optics: A Bibliometric Approach to Detect Emerging Research Domains and Intellectual Bases. *Scientometrics*, 78(3), 543–558.
- Tseng, Y-H.; Lin, Y-I.; Lee, Y-Y.; Hung, W-Ch.; Lee, Ch-H. (2009). A Comparison of Methods for Detecting Hot Topics. *Scientometrics*, 81(1), 73–90.
- Tu, Y-N.; Hsu, S-L. (2016). Constructing Conceptual Trajectory Maps to Trace the Development of Research Fields. *Journal of the Association for Information Science and Technology*, 67(8), 2016–2031.
- Upham, S.P.; Small, H. (2010). Emerging Research Fronts in Science and Technology: Patterns of New Knowledge Development. *Scientometrics*, 83(1), 15–38.
- Van Den Besselaar, P.; Heimeriks, G. (2006). Mapping Research Topics Using Word-Reference Co-Occurrences: A Method and an Exploratory Case Study. *Scientometrics*, 68(3), 377–393.
- Verbeek, A.; Debackere, K.; Luwel, M.; Zimmermann, E. (2002). Measuring Progress and Evolution in Science and Technology – I: The Multiple Uses of Bibliometric Indicators. *International Journal of Management Reviews*, 4(2), 179–211.
- Waltman, L.; Van Eck, N.J. (2012). A New Methodology for Constructing a Publication-Level Classification System of Science. *Journal of the Association for Information Science and Technology*, 63(12), 2378–2392.
- Wang, Ch-Ch.; Ho, Y-S. (2016). Research Trend of Metal-Organic Frameworks: A Bibliometric Analysis. *Scientometrics*, 109(1), 481–513.
- Wang, X.; Wang, Z.; Xu, S. (2013). Tracing Scientists' Research Trends Realtimely. *Scientometrics*, 95(2), 717–729.

- Wierchoń, S.T.; Kłopotek, M.A. (2015). *Algorithms of cluster analysis*. Warszawa: Institute of Computer Science, Polish Academy of Sciences.
- Wilson, R.J. (2012). *Wprowadzenie do teorii grafów*. Warszawa: PWN.
- Zheng, T.; Wang, J.; Wang, Q.; Nie, Ch.; Shi, Z.; Wang, X.; Gao, Z. (2016). A Bibliometric Analysis of Micro/Nano-Bubble Related Research: Current Trends, Present Application, and Future Prospects. *Scientometrics*, 109(1), 53–71.
- Zitt, M.; Bassecouard, E. (1994). Development of a Method for Detection and Trend Analysis of Research Front Built by Lexical or Co-Citation Analysis. *Scientometrics*, 30(1), 333–351.
-

Bibliometric Methods to Foresee and Assess the Development of Scientific Disciplines. Literature Analysis. Part 1: Trailblazing Publications, Bibliographic Coupling Method, Co-citation Analysis and Co-word Methods

Abstract

Purpose/Thesis: The aim of the article is to provide an evaluative review of a literature on the quantitative methods of bibliometric forecasting and assessing the current state of scientific disciplines development. By “evaluative” review it is understood that the author discussed and pinpointed the drawbacks of the literature-proposed ways to assess and predict the development of scientific disciplines.

Approach/Methods: The method used by the author is the analysis of the literature concerning the problems under consideration. The author also proposed a typology of methods applied in the bibliometric predictions area and emphasized the flaws of some previously used methods. Moreover, the possibilities of practical applications of foresight exercises were shown. Finally, the conclusions were drawn with regard to the methods oriented towards the assessment of the rate and direction of scientific disciplines development.

Results and conclusions: Taking into consideration the fact that this article is first part of a larger whole comprising the analysis of the literature in question, the conclusions drawn within the first part have only a tentative character. It was emphasized that the methods proposed in the trailblazing publications are still vital and fulfill a leading role in bibliometrics. Second, it was indicated that the drawbacks accompanying the majority of mentioned methods are an integral part of the evolution of the research mainstream which is being discussed. Third, it is necessary to note that there is a strong correlation between bibliometric predictive methods and an algorithmic means of clustering groups of publications into topic consistent subgroups. These subgroups are usually considered to represent disciplinary subspecialties or, in other words, disciplinary problem areas.

Originality/Value: The originality and value of the article lie in an effort to demonstrate the scope and a high degree of progress of the studies on the possibilities of prediction and assessment of the development of scientific disciplines worldwide. In the Polish scientific environment such studies are relatively rare and it seems that the review of the predictive bibliometric methodology could contribute to the stimulation of interest in the research directed towards the forecasting of future scientific trends. It is also important to stress the disadvantages of particular methods of bibliometric forecasting as in this context they create a substantial amount of space destined for possible improvements and modifications of existing methods. In the field of bibliometrics there is a need for such refinements due to the fact that currently there is no ideal method which could be described as “the last word” on the technique of anticipating future research directions.

Keywords

Bibliometrics. Predictions. Development of scientific disciplines. Quantitative methods in information science

ŁUKASZ OPALIŃSKI uzyskał tytuł magistra w zakresie nauki o języku i komunikacji (jako specjalności wyodrębnionej z kierunku filozofia) na Uniwersytecie Marii Curie-Skłodowskiej w Lublinie w 2005 r. Ukończył studium podyplomowe Technologie Informacyjne i Bibliotekoznawstwo w Wyższej Szkole Informatyki i Zarządzania z siedzibą w Rzeszowie w 2009 r. oraz studium podyplomowe Bibliotekarz Dziedzinowy w Instytucie Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego w 2012 r. Pracuje w Oddziale Informacji Naukowej Biblioteki Politechniki Rzeszowskiej na stanowisku bibliotekarza. Najważniejsze publikacje: Opaliński, Ł. (2012). Niealfabetyczne systemy piśmiennicze a komunikacja językowa w Internecie [online]. *Podkarpackie Studia Biblioteczne*, 1/2012, 1–19, [25. 02. 2017], <http://repozytorium.ur.edu.pl/handle/item/180>; Opaliński, Ł. (2013). Wybrane aspekty metodologii badań cyklu życiowego publikacji naukowych. *Przegląd Biblioteczny*, 81(2), 152–171; Opaliński, Ł., Jaromin, M., Wikiera, J. (2015). Problem stabilności zachowań naukowców w zakresie cytowań w kontekście metodologii badań starzenia się publikacji naukowych i możliwość jego ujęcia ilościowego. *Zagadnienia Informacji Naukowej – Studia Informacyjne*, 53(2/106), 65–83.

Kontakt z autorem:

lopa@prz.edu.pl

Oddział Informacji Naukowej Biblioteki Politechniki Rzeszowskiej

al. Powstańców Warszawy 12

35–959 Rzeszów

ProBIT – prospektywna metoda tworzenia trawersowalnych indeksów cytowań a współczesne problemy organizacji przestrzeni informacji w tradycyjnych bibliograficznych bazach danych

Anna Małgorzata Kamińska

Biblioteka Główna Politechniki Śląskiej w Gliwicach

Abstrakt

Cel/Teza: Artykuł przedstawia autorską metodę budowy bibliograficznych baz danych, sprzyjającą sprawnemu gromadzeniu danych bibliograficznych o wysokiej jakości, które bez dalszych przekształceń stanowią źródło zaawansowanych analiz bibliometrycznych (również tych, które nie ograniczają się jedynie do prostych zestawień ilościowych wyliczanych na podstawie danych zgromadzonych w układzie relacyjnym, ale takich które wymagają trawersowania po ścieżkach grafów reprezentujących sieć cytowań artykułów, czy współpracy między ich autorami).

Koncepcja/Metody badań: Autorka, stojąc przed problemem zgromadzenia znacznego wolumenu danych bibliograficznych do realizacji badań własnych, zapoznała się z szeregiem gotowych narzędzi pozwalających na rejestrowanie danych bibliograficznych, jednak żadne z nich nie oferowało oczekiwanej ergonomii, dbałości o jakość gromadzonych danych ani sposobu ich gromadzenia w układzie umożliwiający prowadzenie na nich dalszych badań bibliometrycznych. Wyciągając wnioski z negatywnych doświadczeń nabytych podczas pracy z tymi narzędziami, zaproponowała własną metodę, którą przedstawia z wykorzystaniem podstawowych notacji języka UML.

Wyniki i wnioski: Przedstawiona metoda okazała się skuteczna i pozwoliła zgromadzić z autopsji w stosunkowo krótkim czasie bardzo dobrej jakości dane o ponad 36 tys. artykułów i ponad 22 tys. autorów, które bezpośrednio posłużyły do realizacji badań bibliometrycznych, które zostaną przedstawione w osobnej pracy.

Zastosowania praktyczne: W pierwszej części artykułu zaprezentowano zagrożenia, jakie nieśie z sobą gromadzenie danych bibliograficznych w tzw. tradycyjnym układzie. Na konkretnych przykładach, dla wybranych krajowych bibliograficznych baz danych wykazano, że zagrożenia te materializują się w postaci występowaniu szeregu anomalii, które zostały przez autorkę skategoryzowane, a naprzeciw którym wychodzi zaproponowana metoda. Dzięki jej zastosowaniu możliwe jest podniesienie jakości danych gromadzonych w bibliograficznych bazach danych oraz obniżenie kosztów ich gromadzenia poprzez usprawnienie całego procesu.

Oryginalność/Wartość poznawcza: W polskim piśmiennictwie bibliologii i informatologii rzadko podejmowana jest tematyka funkcjonowania bibliograficznych baz danych, najczęściej badacze ograniczają się do definiowania „zakresów informacyjnych przetwarzanych danych”. Autorka zwraca uwagę, że na osiągnięty wynik końcowy (jakość danych, koszt ich pozyskania i możliwości dalszego przetwarzania) w równym stopniu wpływ ma również układ, w jakim dane są gromadzone (model informacyjny) oraz odpowiednia organizacja procesu gromadzenia.

Słowa kluczowe

Bibliograficzne bazy danych. ProBIT. Metoda prospektywna. Przestrzeń informacji. Trawersowalny indeks cytowań.

Otrzymano: 19 lutego 2017. Zrecenzowano: 23 czerwca 2017. Zaakceptowano: 30 sierpnia 2017.

1. Wprowadzenie

Dzięki mechanizmom systemów zarządzania bibliograficzne bazy danych umożliwiają gromadzenie, przechowywanie, przetwarzanie, wyszukiwanie i udostępnianie informacji. Ważnymi elementami bibliograficznych baz danych są informacje opisujące cytowania pomiędzy poszczególnymi jednostkami bibliograficznymi. Dzięki tym informacjom możliwe jest wsparcie szeregu działań, spośród których wymienić należy chociażby takie jak:

- ewaluacja prac naukowych,
- ewaluacja dorobku naukowego autora;
- ewaluacja potencjału instytucji afiliujących;
- ewaluacja rangi czasopism i innych wydawnictw naukowych;
- badania nad rozwojem dziedzin nauki i powstawaniem nowych ich obszarów;
- wzbogacenie standardowych mechanizmów wyszukiwawczych OPAC (Woźniak-Kasperek, 2011) o możliwość podpowiadania dokumentów podobnych, gdzie miary podobieństwa mogą bazować na metodach powiązań bibliograficznych, współcytowań i innych (Marszakowa-Szajkiewicz, 2009).

Chociaż pierwsze opinie o istnieniu kognitywnego związku między pracą cytowaną i cytującą datuje się na 1955 r. (Garfield, 1955), to zainteresowanie badaniami tych związków ciągle rośnie i to zarówno pod względem ilościowym (włączanie w badawcze obszary coraz to nowych dziedzin nauki), jak i jakościowym (opracowywanie nowych miar analiz bibliometrycznych oraz bibliometrycznych metod badawczych). Sprzyja temu niewątpliwie rozwój technologii informacyjnych, powszechna komputeryzacja coraz to nowych dziedzin życia, w tym ekosystemu bibliotek oraz ciągły wzrost dostępnych mocy obliczeniowych. Wraz z gwałtownym rozwojem metod badania powiązań w sieciach społecznościowych (ang. *social network analysis*) (Al et al., 2012), przeniesienie na grunt bibliometrii i interpretowanie uzyskanych tą drogą wyników, wydają się być ograniczone jedynie wyobraźnią poszczególnych naukowców (Kretschmer & Kretschmer, 2007).

Jak zauważają badacze krajowi i zagraniczni, w tym Władysław Kolasa, obszernie wyjaśniając to na wielu przykładach, możliwość prowadzenia takich badań oraz wiarygodność ich wyników, zależą przede wszystkim od dostępności i jakości gromadzonych danych, a ta pozostawia wiele do życzenia szczególnie w przypadku bibliograficznych baz danych obejmujących nauki humanistyczne, gdyż te będąc bardziej niż nauki ścisłe czy techniczne spolaryzowane językowo i mając bardziej lokalny charakter współpracy naukowej, są indeksowane przez globalne serwisy w dużo mniejszym stopniu, co prowadzi do powstawania lokalnych bibliograficznych baz danych (Kolasa, 2011). Choć trudno się z tą opinią nie zgodzić, to można zaryzykować postawienie szerszych hipotez. Po pierwsze, przypuszczać można, że niska jakość danych wynika przede wszystkim z lokalnego charakteru gromadzonych informacji bibliograficznych, a w mniejszym stopniu z dziedziny, w ramach której te informacje są gromadzone. Duże i globalne korporacje posiadające

odpowiednie środki na weryfikację jakości rekordów oraz nadążanie za rozwojem technologii, stoją na uprzywilejowanej pozycji wobec dostawców lokalnych (krajowych) baz danych, które często zaprojektowane wiele lat temu, są eksploatowane po dni dzisiejsze bez koniecznych ulepszeń. Wśród wrywkowej kontroli jakości i spójności większości bibliograficznych baz danych dostępnych w Polsce, a realizowanej w ramach badań prowadzonych na potrzeby niniejszych rozważań, stwierdzono problemy, które dotyczą w tym samym stopniu baz humanistycznych jak i technicznych. Przedstawione w dalszej części artykułu przykłady ograniczono do bazy BazTech (<http://yadda.icm.edu.pl/baztech/>), która jest jedną z najbardziej reprezentatywnych baz bibliograficznych z zakresu nauk technicznych oraz bazy CYTBIN (<http://www1.bg.us.edu.pl/bazy/cytbini/>), indeksującej prace z dziedziny bibliologii i informatologii. Jako drugą hipotezę można sformułować przypuszczenie, że możliwość realizacji bardziej zaawansowanych badań bibliometrycznych nie jest warunkowana jedynie jakością gromadzonych danych, ale również organizacją tych danych w konkretnych systemach zarządzania bazami danych. Wiele metod analiz bibliometrycznych operuje na ścieżkach cytowań, tak więc organizacja informacji opisującej bibliografię załącznikową w postaci par „jednostka cytowana – jednostka cytująca” jest niewystarczająca (nawet gdy dane te odzwierciedlają prawdziwe informacje) bez możliwości uwzględnienia faktu, że jednostka cytująca może być jednostką cytowaną dla innych jednostek cytujących.

Więcej informacji o inicjatywach bibliograficznych baz danych realizowanych na obszarze Polski, stanie tych projektów na 2009 r. i planowanych naówczas kierunkach ich rozwoju, uzyskać można ze sprawozdania z konferencji dotyczącej bibliograficznych baz danych (Drabek, 2009), natomiast ocenę stopnia ich dojrzałości i możliwości wykorzystania przez pracowników i jednostki szkół wyższych znaleźć można w (Garczyńska, 2013).

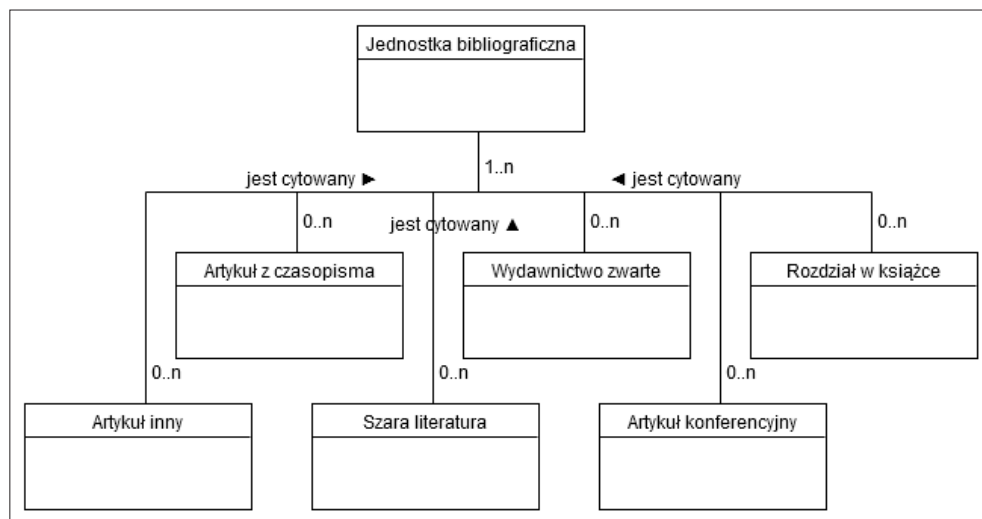
Niniejszy artykuł opisuje problemy, z jakimi styka się większość instytucji udostępniających bibliograficzne bazy danych na terenie Polski, skupiając się na perspektywie konceptualno-technicznej, a dokładniej na architekturze przestrzeni informacyjnej oraz organizacji pracy osób obsługujących systemy bibliograficznych baz danych. Po przedstawieniu i omówieniu tych problemów, w dalszej części artykułu, zaproponowana zostanie metoda organizacji informacji bibliograficznych gromadzonych w bazach, która w połączeniu ze zmianą organizacji pracy osób rejestrujących te informacje, zapewni dużo lepszą jakość gromadzonych danych oraz gromadzenie ich w układzie umożliwiającym realizację zaawansowanych badań bibliometrycznych wprost na zgromadzonych danych w formie relacyjnej. Skuteczność tej metody znalazła praktyczne potwierdzenie podczas realizacji badań własnych, na potrzeby których, rezygnując z wykorzystania narzędzi gotowych, takich jak np. te, które przedstawiono w osobnym opracowaniu (Kamińska, 2017d), zaimplementowano środowisko z zastosowaniem przedstawianej metody. Dane zgromadzone w ramach tych badań posłużyły do realizacji szeregu analiz bibliometrycznych od prostych zestawień ilościowych, poprzez zastosowanie metody powiązań bibliograficznych i metody współcytowań, aż po badania z zastosowaniem szeregu metod analiz sieci społecznościowych.

2. Problemy organizacji przestrzeni informacji w tradycyjnych bibliograficznych bazach danych

Większość inicjatyw bibliograficznych baz danych realizowanych na terenie Polski opiera się na tradycyjnym modelu struktury danych przedstawionym za pomocą diagramu klas języka UML na rysunku 1. Omówienie tej wizualnej techniki komunikacyjnej i propozycję jej wykorzystania na gruncie zagadnień bibliologii i informatologii autorka zaproponowała w pracy pt. *O rozwoju graficznych języków komunikacji* (Kamińska, 2017b). Bardziej obszerne omówienie poszczególnych technik języka UML znaleźć można w monografii samych jego twórców (Booch et al., 2002).

Z modelu wynika, że nadrzędnym elementem tradycyjnej organizacji przestrzeni informacyjnej jest jednostka bibliograficzna, która podlega opisowi. Poza innymi elementami składowymi jej opisu (jak np. tytuł, typ nośnika, autorzy i wiele innych), które nie mają jednak istotnego wpływu na prowadzone tu rozważania, głównym elementem leżącym u podstaw koncepcji bibliograficznych baz danych jest bibliografia załącznikowa.

Rys. 1. Model organizacji danych w tradycyjnych bibliograficznych bazach danych

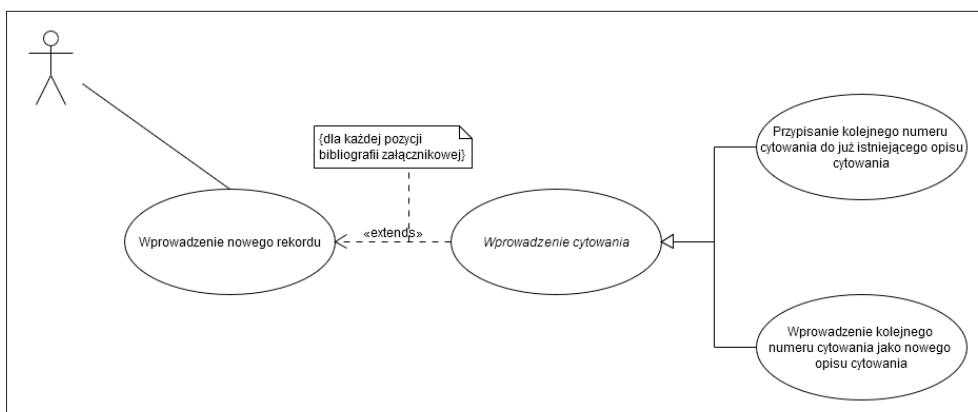


W tradycyjnym modelu opisy bibliograficzne zawarte w bibliografii załącznikowej są gromadzone w składnicach danych oddzielnie dla poszczególnych form wydawniczych (np. artykuł z czasopisma, wydawnictwo zwarte). Na przedstawionym diagramie wyróżniono sześć przykładowych form, choć w konkretnych bazach danych liczba tych kategorii może być różna. Potrzeba wyróżniania różnych kategorii wynika z faktu, że w ramach różnych kategorii dokumentuje się różne cechy dokumentów, co z kolei dyktowane jest potrzebą jak najdokładniejszego opisu jednostek poszczególnych typów. Jak wynika z diagramu, każda jednostka danej kategorii (tj. obiekt klasy Artykuł z czasopisma, obiekt klasy Wydawnictwo zwarte itd.), może być przyporządkowana (tę relację asocjacji możemy nazwać „jest cytowany”) do jednej lub wielu jednostek bibliograficznych. Analizując relacje w odwrotnym kierunku widzimy z kolei, że każda opisywana jednostka bibliograficzna,

w swojej bibliografii załącznikowej, może zawierać zero lub wiele jednostek poszczególnych form wydawniczych. Taka organizacja danych, oparta jedynie na związkach asocjacji, choć na pierwszy rzut oka sprawia wrażenie porządku i dobrej organizacji przestrzeni danych, ma szereg wad i niesie ze sobą wiele zagrożeń, które omówione zostaną w dalszej części artykułu.

Powyżej przedstawiony strukturalny opis przestrzeni tradycyjnych bibliograficznych systemów baz danych jest niepełny, gdyż nie opisuje w żaden sposób jego aspektów behawioralnych, z których dla niniejszych rozważań najistotniejszy jest aspekt interakcji użytkownika takiego systemu z graficznym interfejsem użytkownika (ang. GUI – *graphical user interface*). Aspekt ten ukazuje diagram przypadków użycia (ang. *use case diagram*) wyrażony w języku UML na rysunku 2.

Rys. 2. Model organizacji pracy dla rejestracji jednostek bibliograficznych w tradycyjnych bibliograficznych bazach danych



Modelowanie interakcji użytkownika (interesariusza) z systemem (niekoniecznie musi to być system informatyczny) jest dziedziną samą w sobie, na temat której powstało wiele obszernych opracowań, jak np. (Cockburn, 2004). Dla potrzeb tego badania (jako że mowa o systemach baz danych) przyjęto, że przypadek użycia (ciąg interakcji użytkownika z systemem) jest najmniejszą (atomową) jednostką ciągu czynności, skutek której może zostać utrwalony w składnicy danych. W tym przypadku będzie to wprowadzenie do bazy nowego rekordu jednostki bibliograficznej lub wprowadzenie informacji o pojedynczej pozycji bibliografii załącznikowej dla danej jednostki bibliograficznej. Systemy bibliograficznych baz danych udostępniają wiele innych funkcjonalności (np. możliwość korekty rekordu wcześniej wprowadzonego, edycję słowników typów rozpowszechniania, itp.), jednak nie mają one istotnego znaczenia z punktu widzenia niniejszych rozważań i dlatego zostały w modelu pominięte. Możliwość budowania ograniczonych modeli pozwalających przedstawić system z istotnej w danym momencie perspektywy, z zaniechaniem umieszczania w nim elementów, które mogłyby zaciemnić obraz, jest postrzegane jako dobra praktyka budowy modeli.

Przedstawiony diagram należy „czytać” następująco. Wprowadzany jest nowy rekord do systemu (sposób w jaki jest to robione – kolejność wypełniania pól, itp. zależne będą od konkretnego systemu i mogłyby być wyrażone poprzez opis scenariuszy przypadków

użycia, ale można od nich abstrahować w ramach niniejszych rozważań, podobnie jak można było abstrahować od zawartości informacyjnej poszczególnych klas na poprzednim diagramie). Po wpisaniu danych opisujących nowy rekord, można przejść do opisu jego bibliografii załącznikowej. Oznaczenie «extends» wyraża właśnie, że dany przypadek użycia może, ale nie musi, „rozszerzać” inny przypadek użycia. W tej sytuacji „rozszerzenie” nie nastąpi dla jednostek nie posiadających bibliografii załącznikowej. Notatka nad relacją «extends» oznacza natomiast, że rozszerzający przypadek użycia, będzie realizowany dla każdej z pozycji bibliograficznych z osobna. Rozszerzającym przypadkiem jest „Wprowadzenie cytowania”, które – jak widać na diagramie – opisane jest kursywą. Oznacza to, że ten przypadek użycia jest abstrakcyjny, a więc nie jest możliwe doprecyzowanie przebiegu czynności, które powinny być w jego ramach realizowane, gdyż w tej sytuacji czynności te mogą być realizowane na dwa różne, wzajemnie wykluczające się sposoby, co obrazuje zamknięta strzałka oznaczająca związek „specjalizacji”, poprowadzona z przypadków użycia opisujących odpowiednio wprowadzenie jako pozycję cytowaną istniejącego już wcześniej rekordu z przypisaniem mu numeru z listy bibliografii załącznikowej lub dopisanie nowego rekordu z przypisaniem mu numeru z listy bibliografii załącznikowej.

Przedstawiona powyżej tradycyjna organizacja przestrzeni informacji prowadzi często do gromadzenia danych redundantnych oraz może nie sprzyjać właściwej dbałości o jakość gromadzonych danych, co skutkuje powstawaniem licznych anomalii. Redundancja najczęściej zachodzi w sytuacjach, kiedy dana pozycja bibliografii załącznikowej staje się (lub była wcześniej) również opisywaną jednostką bibliograficzną. Mamy wtedy do czynienia z sytuacją, kiedy ta sama jednostka występuje w roli jednostki bibliograficznej oraz którejś z cytowanych form wydawniczych (dla potrzeby późniejszych rozważań oznaczono ten przypadek jako α). W takich sytuacjach oprócz redundancji dochodzi również często do wewnętrznych sprzeczności zarejestrowanych informacji w przypadkach, kiedy wspólne części zakresu informacyjnego jednostki bibliograficznej oraz którejś z cytowanych form wydawniczych przechowują różne dane reprezentujące te same atrybuty (przypadek β). Z trzecią anomalią mamy natomiast do czynienia w sytuacji, gdy w bazie danych istnieje już rekord opisujący daną cytowaną formę wydawniczą, natomiast operatorowi systemu nie uda się go odnaleźć i zamiast realizacji przypadku użycia „Przypisanie kolejnego numeru cytowania do już istniejącego opisu cytowania” zrealizuje „Wprowadzenie kolejnego numeru cytowania jako nowego opisu cytowania” (przypadek γ). Anomalia ta niesie ze sobą poważne, negatywne konsekwencje dla mechanizmu wyszukiwania jednostek, które cytują zadany artykuł, czyniąc jego działanie mało wiarygodnym. Dwóch z wcześniej wymienionych przypadków nie można uniknąć i wpisują się one wręcz w konstrukcję systemów o wyżej przedstawionej architekturze, jednak ostatni przypadek jest już zależny od rzetelności operatora, która powinna być wspierana przez taką konstrukcję graficznego interfejsu użytkownika, która będzie faworyzować „ponowne użycie” (ang. *reuse*) rekordu, a więc przypisanie istniejącego rekordu cytowanej formy do numeru pozycji bibliografii załącznikowej danej jednostki zamiast dopisywania nowego. Najbardziej oczywistą i najczęstszą (choć nie jedyną) motywacją każdego operatora systemu jest możliwość włożenia jak najmniejszego wysiłku i poświęcenia jak najmniejszej ilości czasu na realizację danej czynności, zatem interfejs graficzny powinien być tak skonstruowany, aby wyszukiwanie wśród rekordów istniejących i przypisanie ich do danego numeru bibliografii załącznikowej było znacznie mniej czasochłonne niż dopisanie rekordu nowego.

Niektóre systemy bibliograficznych baz danych zakładają wręcz zasilanie swoich składnic informacji jedynie poprzez dopisywanie nowych rekordów cytowań (przypadek użycia opcji „wprowadzenie kolejnego numeru cytowania jako nowego opisu cytowania”), pozostawiając porządkowanie zgromadzonych danych realizowanej z opóźnieniem operacji scalania rekordów. Należy tu zauważyć, że występujący często w literaturze termin „scalanie” używany w opisywanym kontekście, nie wydaje się do końca właściwy, gdyż sugeruje, że rekordy wynikowe zawierać będą sumę informacji pochodzącą z rekordów źródłowych, co najczęściej nie jest prawdą. Bardziej właściwym terminem opisującym ten proces wydaje się być „deduplikacja” (Ilyas & Chu, 2015). Owa deduplikacja, realizowana z opóźnieniem w stosunku do czasu zarejestrowania rekordów źródłowych, nie tylko zwiększa sumaryczny nakład pracy na utrzymanie danych w stanie „maksymalnie spójnym”, ale również zmniejsza skuteczność dbałości o jakość gromadzonych danych. I choć jej mechanizmy są i będą użyteczne w każdym systemie bibliograficznych baz danych, to należy dążyć do sytuacji, w której liczba rekordów „nią dotkniętych” będzie jak najmniejsza.

Należy tutaj również wspomnieć, że oprócz wyżej wymienionych czynników sprzyjających powstawaniu anomalii, a leżących po stronie konstrukcji systemu lub rzetelności operatora, istnieje również czynnik od nich niezależny. Są to błędy w materiałach źródłowych popełniane przez samych autorów lub też powstałe w procesie redakcji i druku. Zachodzi wtedy pytanie, które informacje bibliograficzna baza danych powinna uważać za prawdziwe? Te przedstawiane przez egzemplarze drukowane, czy też te, które miał na myśli autor i których poprawność można sprawdzić w innych źródłach? Oczywiście jest, że z punktu widzenia skuteczności realizacji badań bibliometrycznych i wiarygodności uzyskanych tą drogą wyników pozytywnej odpowiedzi należy udzielić na drugie pytanie. Takie też stanowisko przyjęła autorka podczas realizacji swoich badań. Nic nie stoi jednak na przeszkodzie, a nawet wydaje się właściwe, aby docelowa konstrukcja systemów bibliograficznych baz danych przechowywała oba warianty informacji.

3. Przykłady problemów zidentyfikowanych w wybranych ogólnodostępnych bibliograficznych bazach danych

W rezultacie wrywkowej kontroli jakości i spójności danych gromadzonych w bibliograficznych bazach danych opisujących obszary nauki, a realizowanej w ramach prezentowanych tu badań, potwierdzono problemy wcześniej opisane oraz stwierdzono, że dotyczą one w podobnym stopniu baz humanistycznych jak i technicznych. Przedstawione w dalszej części rozdziału przykłady ograniczono do bazy BazTech oraz bazy CYTBIN, której zawartość stanowiła również podstawy do innych badań autorskich dotyczących rozwoju metod bibliometrycznych (Kamińska, 2017c), wizualizacji wyników uzyskanych z ich pomocą (Kamińska, 2017e), poprawy jakości gromadzonych danych bibliograficznych (Kamińska, 2017a) oraz zastosowania metod analizy sieci społecznościowych w bibliometrii (Kamińska, 2018).

Rysunek 3 przedstawia okno systemu CYTBIN w przeglądarce WWW z informacją o opisywanej jednostce (czyli takiej, dla której gromadzona jest informacja o bibliografii załącznikowej), zaś rysunek 4 przedstawia okno z informacją o artykule zarejestrowanym w ramach cytowanych wydawnictw w formie wydawnictwa ciągłego. Jak widać, oba rekordy reprezentujące jeden byt logiczny zawierają sprzeczne informacje o czasopiśmie

i stronach, na których znajduje się artykuł. Przykład ten ilustruje występowanie anomalii typu/ α / oraz/ β /. Faktycznie artykuł, którego dotyczą przedstawione rekordy ukazał się w „Bibliotekarzu” nr 7–8 na stronach 10–14, a nie na stronach 30–34. Widzimy więc, że oba rekordy zawierają błędne informacje oraz reprezentują oddzielne fizyczne jednostki danych.

Rys. 3. Okno systemu CYTBIN wyświetlające informacje o przykładowej opisywanej jednostce

Autor:	Wolff K.
Rodzaj publikacji:	Artykuł
Tytuł:	Biblioteka publiczna w opinii społecznej
Czasopismo:	Przegląd Biblioteczny
Rok publikacji:	2010
Tom:	78
Numer:	7-8
Strony:	10-14
Symbol klasyfikacji tematycznej:	3.12 - Biblioteki publiczne
Liczba cytowań:	0
Liczba instytucji sprawczych:	1
Słowa kluczowe:	Bibliotekoznawstwo, Biblioteka publiczna, Jakość usług bibliotecznych, Czytelnik, Badanie ankietowe
Słowa kluczowe ang.:	Library and information science, Public library
Abstrakt:	Opisano wynik ankiety przeprowadzonej w 2008r. przez Instytut Książki i Czytelnictwa BN. Była ona


Źródło: Baza CYTBIN, dostęp: 6.12.2016

Rys. 4. Okno systemu CYTBIN wyświetlające informacje o przykładowym artykule dla wydawnictw ciągłych

CYTOWANIE - WYDAWNICTWO CIĄGŁE	
Autorzy:	Wolff K.
Rodzaj nośnika:	D
Tytuł:	Biblioteka publiczna w opinii społecznej
Czasopismo:	Bibliotekarz
Rok publikacji:	2010
Numer:	7-8
Strony:	30-34

Źródło: Baza CYTBIN, dostęp: 6.12.2016

Rys. 5. Okno systemu BAZTECH wyświetlające informacje o przykładowym artykule

Artykuł - szczegóły	
Czasopismo	Biuletyn Wojskowej Akademii Technicznej
Tytuł artykułu	Aerozele węglowe otrzymywane z prekursora rezorcynowo-furfuralowego
Autorzy	Kiciński, W.
Treść / Zawartość	
Warianty tytułu	EN Carbon aerogels obtained from resorcinol-furfural precursors
Języki publikacji	PL
Abstrakty	<p>PL Przedstawiono nową metodę otrzymywania aerozeli węglowych poprzez karbonizację organicznych aerozeli rezorcynowo-furfuralowych. Zele rezorcynowo-furfuralowe otrzymywano w wyniku katalizowanej kwasowo kondensacji rezorcyny i furfuralu prowadzonej w wodno-metanolowym roztworze soli nieorganicznych. Następnie sól z sieci żelu wymywano i oczyszczony materiał suszono. Suszenie tradycyjnymi metodami nie prowadziło do zniszczenia struktury porowatej żelu dzięki czemu możliwe było uzyskanie aerozeli organicznych bez konieczności stosowania warunków nadkrytycznych ekstrakcji rozpuszczalnika ze struktury żelu. Karbonizacja aerozeli organicznych pozwalała uzyskać mikro-makroporowate lub mikro-mezo-makroporowate materiały węglowe o strukturze mikropianek i powierzchni właściwej z przedziału 250-580 m²/g. Otrzymane aerozele organiczne i węglowe badano przy użyciu wielu technik, m.in. SEM, XRD, niskotemperaturowej adsorpcji azotu, analizy termogravimetrycznej i elementarnej.</p> <p>EN Acidic condensation of resorcinol with furfural carried out in water-methanol solution of inorganic salts (NaCl or NH₄ClO₄) leads to organic R-F gel. After removing the salt from the gel network, the purified gel can be dried under ambient conditions without significant shrinkage, so the supercritical drying can be avoided, and an organic aerogel can be obtained through a simpler procedure. By changing the H₂O/MeOH ratio and/or the inorganic salt amount, the structure of the created gel can be controlled in the nano and micron-level and as a result polymeric (nano-sized) or colloidal (micron-sized) aerogels can be produced. Carbonization of the colloidal aerogels yields bimodal micro-macroporous carbonaceous materials with the structure of microfoam and low density, while carbonization of the polymeric ones yields bimodal, micro and meso-macroporous carbonaceous materials with high density and mechanical strength. Specific surface area of the carbon aerogels ranged from 250 to 580 m²/g. The obtained organic and carbon aerogels were investigated by the means of SEM, XRD, elemental analysis, N₂ adsorption, and TG.</p>
Słowa kluczowe	<p>PL synteza żol-żel aerozel węglow aerozel koloidalny polikondensacja</p> <p>EN sol-gel synthesis carbon aerogel colloidal structure polycondensation</p>
Wydawca	Wojskowa Akademia Techniczna im. Jarosława Dąbrowskiego
Czasopismo	Biuletyn Wojskowej Akademii Technicznej
Rocznik	2009
Tom	Vol. 58, nr 4
Strony	197--221
Opis fizyczny	Bibliogr. 57 poz., tab., wyk.
Twórcy	autor Kiciński, W. Wojskowa Akademia Techniczna, Wydział Nowych Technologii i Chemii, Instytut Chemii, 00-908 Warszawa, ul. S. Kaliskiego 2
Bibliografia	<p>[1] R. W. Pekala, F. M. Kong, Polym. Prep., 30, 1, 1989, 221.</p> <p>[2] R. W. Pekala, D. W. Schaefer, Macromolecules, 26, 1993, 5487.</p> <p>[3] S. A. Al-Muhtaseb, J. A. Ritter, Adv. Mater., 15, 2, 2003, 101.</p> <p>[4] K. Brodzik, M. Stolarski, J. Walendziwski, Wiadomości Chemiczne, 7-8, 2004, 58.</p>

Źródło: Baza BAZTECH, dostęp: 6.12.2016

Jak już wspomniano, na podstawie danych gromadzonych w takim układzie nie jest możliwe budowanie ścieżek cytowań stanowiących podstawę zaawansowanych analiz bibliometrycznych.

Podobne problemy można zauważyć w bazie BazTech, której przykładowe okno (przełóżarki WWW) przedstawiono na rysunku 5.

Informacje o tym samym artykule, ale występującym w ramach bibliografii załącznikowej przedstawia rysunek 6. Widać tutaj wyraźnie, że dane na rysunku 5 różnią się od danych na rysunku 6 z pozycji [10] w atrybutach opisujących strony, czy numery czasopisma (choć reprezentują te same wartości, to są to jednak inne dane), co świadczy o występowaniu

anomalię typu $\alpha/$ oraz $\beta/$. Na rysunku 6 „wklejono” również fragmenty opisów bibliograficznych dla innych artykułów (pozycje [61] i [79]).

Również tutaj widać różnice w zapisie, co świadczy o redundancji przechowywanych danych, czyli anomalię typu $\gamma/$.

Rys. 6. Okno systemu BAZTECH wyświetlające informacje o bibliografii załącznikowej

Czasopismo	Przegląd Włókienniczy - Włókno, Odzież, Skóra	
Tytuł artykułu	Nowy kierunek zastosowań aerożeli w balistyce	
Autorzy	Litwa, P.	
Warianty tytułu	EN	New fields of applications of aerogels in ballistic
Języki publikacji	PL	
Abstrakty	EN	The paper presents a literature review considering methods of obtaining the carbon aerogels by polycondensation from resorcinol-formaldehyde (RF) precursor and discuss of research results for RF and crosslinked aerogels. The result of own, preliminary research of soft inserts for bullet- and fragment proof vests are presented. They clearly indicate that use of aerogel is effective and increase the ballistic protection. The article includes also through characterization of ability to absorb impact energy during static and dynamic compression tests as well as the ways to control the size and shape of nanoporous. Structure and porosity of porous materials have a significant impact on their properties.
Słowa kluczowe	PL	aerozele balistyka
Wydawca	Wydawnictwo SIGMA-NOT	
Czasopismo	Przegląd Włókienniczy - Włókno, Odzież, Skóra	
Rocznik	2012	
Tom	nr 12	
Strony	23--28	
Opis fizyczny	Bibliogr. 33 poz., fot., rys., tab.	
Twórcy	autor	Litwa, P. Instytut Technologii Bezpieczeństwa MORATEX w Łodzi
Bibliografia	<ol style="list-style-type: none"> 1. Fricke J., Tillotson T: Aerogels: production, characterization and applications, Thin Solid Films 297 (1997); 212 - 223. 2. Zhang R., Li W., Li K.: Effect of concentration of reactants on porosity of hydrogels, organic and carbon aerogels, Microporous and Mesoporous Materials 72 (2004); 167 - 173. 3. Feng J., Feng J., Jiang Y.: Ultra-low density carbon aerogels with low thermal conductivity up to 2000 °C. Materials Letter. 	
[61]	Kiciński W.: Aerozele węglowe otrzymywane z prekursora rezorcynowo- -furfuralowego. Biul. WAT 58 (4) (2009) 197÷221.	
[79]	W. Kiciński, Aerozele węglowe otrzymywane z prekursora rezorcynowo-furfuralowego, Biul. WAT, 58, 4, 2009, 197-221.	
	5. Gupta N., Ricci W.: Processing and compressive properties of aerogel/ epoxy composites, Journal of Materials Processing Technology 198 (2008): 178 -182.	
	7. Yang J., Li S., Yan L. i wsp.: Compressive behaviors and morphological changes of resorcinol-formaldehyde aerogel at high strain rates, Microporous and Mesoporous Materials 133 (2010); 134 - 140.	
	8. Mirzaian M., Hall P. J.: Preparation of controlled porosity carbon aerogels for energy storage in rechargeable lithium oxygen batteries, Electrochimica Acta 54 (2009); 7444 - 7451.	
	9. Czakel O., Szekeley E., Koczka B.: Drying of resorcinol-formaldehyde gels with CO2 medium, Microporous and Mesoporous Materials 148 (2012); 34-42.	
	10. Kiciński W.: Aerozele węglowe otrzymywane z prekursora rezorcynowo-furfuralowego, Biuletyn Wojskowej Akademii Technicznej, Vol. 58, 4(2009); 197-221.	

Źródło: Baza BAZTECH, dostęp: 6.12.2016

Omawiane problemy, choć do tej pory nie tak dokładnie skategoryzowane i powiązane z przyczynami ich powstawania, wynikającymi wprost z przyjętej organizacji przestrzeni informacyjnych, dostrzegane były już od jakiegoś czasu. Ich negatywne skutki rzutują wprost na podstawowe funkcje takich systemów (wyszukiwanie informacji) czyniąc je często niedoskonałymi i mało wiarygodnymi. Ograniczona wiarygodność zgromadzonych danych oraz przechowywanie ich w układzie tradycyjnym, mogą utrudniać, a nawet uniemożliwiać prowadzenie wiarygodnych i bardziej zaawansowanych analiz bibliometrycznych.

Władysław Kolasa, dla potrzeb realizacji bibliometrycznych badań z dziedziny humanistyki, zaproponował metodę „retrospektywnego indeksu cytowań”, którą z powodzeniem zastosował do zgromadzenia potrzebnych mu danych badawczych (Kolasa, 2011). Należy zauważyć, że, jak można domyślać się z opisu zamieszczonego w przytoczonej artykule, celem tej metody nie było zbudowanie „operacyjnej” bibliograficznej bazy danych, która po wdrożeniu mogłaby na bieżąco być zasilana nowymi strumieniami informacji, a jedynie

pozyskanie materiału badawczego o z góry założonym zakresie, drogą „kolejnych przekształceń danych” danych źródłowych pobranych z bazy bibliograficznej, a następnie ich transformacji i migracji do innych składnic danych. Stąd też zapewne wzięła się sama nazwa tej metody. Należy tutaj jednak zauważyć, że, mimo iż niektórzy badacze zaliczają to rozwiązanie do jednej grupy funkcjonalnej np. z bazą CYTBIN (Seweryn & Swoboda, 2014), to powstały „produkt” trudno już nazwać bibliograficzną bazą danych w powszechnym znaczeniu tego określenia, gdyż stanowi on jedynie materializację jej zawartości w wyznaczonym przez badacza okresie. Dodatkowo w bazie tej nie były gromadzone informacje o pozycjach cytowanych niebędących jednocześnie pozycjami cytującymi, gdyż zdaniem autora tej metody wyniki badań bibliometrycznych bez takich informacji będą zbliżone do wyników badań opartych na informacji pełnej.

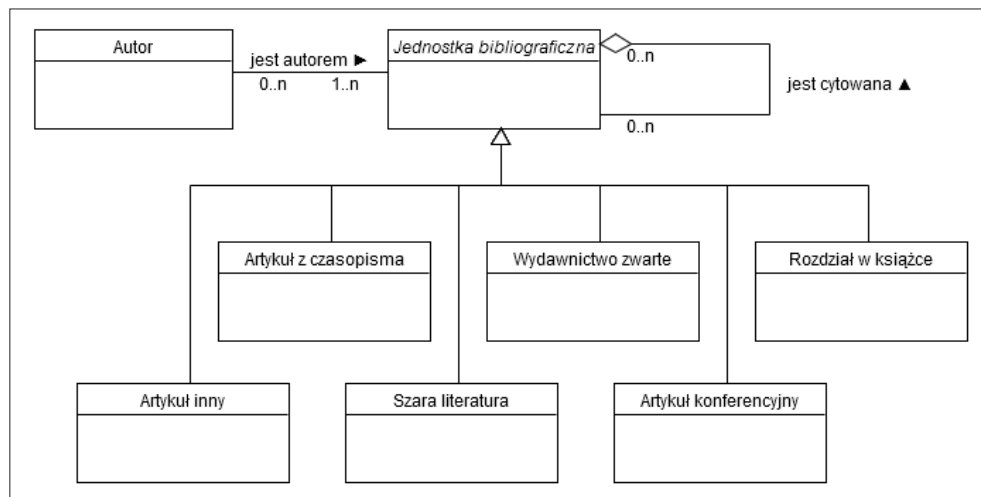
Autorka stała również przed problemem zgromadzenia obszernego wolumenu danych bibliograficznych, których pozyskanie możliwe było tylko z autopsji. Próby wykorzystania gotowych narzędzi od ogólnodostępnych pakietów biurowych do zaawansowanych systemów bibliotek cyfrowych nie zakończyły się zachęcającymi wynikami, ale dostarczyły wielu przemyśleń co do warunków niezbędnych do efektywnego wprowadzania danych o wymaganej jakości. Również próby automatyzacji tego procesu z zastosowaniem mechanizmów OCR okazały się mało skuteczne z powodu różnic w formatach i braku standaryzacji zapisów bibliografii. Stąd ostatecznie autorka przystąpiła do rozwiązania problemu zgromadzenia dużego zbioru danych w sposób odmienny. Nie zakładając z góry, że strumienie napływających danych ustaną w konkretnym momencie oraz chcąc zgromadzić pełną informację o danych zawartych w bibliografiach załącznikowych (co było konieczne chociażby do realizacji badań metodą powiązań bibliograficznych i współcytowań), zaproponowała metodę tworzenia Prospektywnego Bibliograficznego Indeksu Trawersowalnego – ProBIT. Dzięki tej metodzie zbudowana baza bibliograficzna może być zasilana w sposób ciągły (równoległe z realizacją badań bibliometrycznych), a ponieważ gromadzi pełny zakres danych bibliograficznych w układzie relacyjnym, może służyć jednocześnie jako podstawowy system informacyjny. Dodatkowo w dowolnym momencie cyklu życia systemu możliwe jest wykonywanie analiz bibliometrycznych, nawet tych bardziej zaawansowanych i wymagających trawersowania po ścieżkach cytowań, bez potrzeby dodatkowych migracji danych. Nazwę swoją metoda bierze stąd, że (posługując się nomenklaturą algebry relacji używaną w pracy wyżej wspomnianego autora) „klucze relacji” pomiędzy jednostkami cytującymi a cytowanymi określane są już w momencie pierwszego zaistnienia informacji o danym cytowaniu jednostki cytowanej i powinny być one użyte w przyszłości, w przypadku pojawienia się innych źródeł cytujących tę samą jednostkę cytowaną.

4. Propozycja metody tworzenia prospektywnego trawersowalnego indeksu bibliograficznego

Koncepcyjny model klas opisujący organizację przestrzeni informacyjnej dla przedstawianej metody zamieszczono na rysunku 7. U jego podstaw leży obiektowe postrzeganie rzeczywistości, nakazujące dostrzeganie pewnych abstrakcji w postaci związków generalizacji-specjalizacji (a nie tylko związków asocjacji, jak ma to miejsce w modelu tradycyjnym).

Widzimy tutaj, że opisywana jednostka bibliograficzna jest klasą abstrakcyjną co oznacza, że bez jej doprecyzowania („ukonkretnienia”) jej byty nie mogą zaistnieć samodzielnie, tj. jako konkretne obiekty (ang. *concrete objects*). Klasy potomne (np. Artykuł z czasopisma, Wydawnictwo zwarte itp.) dziedziczą wszystkie właściwości tej klasy bazowej, a jednocześnie rozszerzają ją o atrybuty specyficzne dla tych klas.

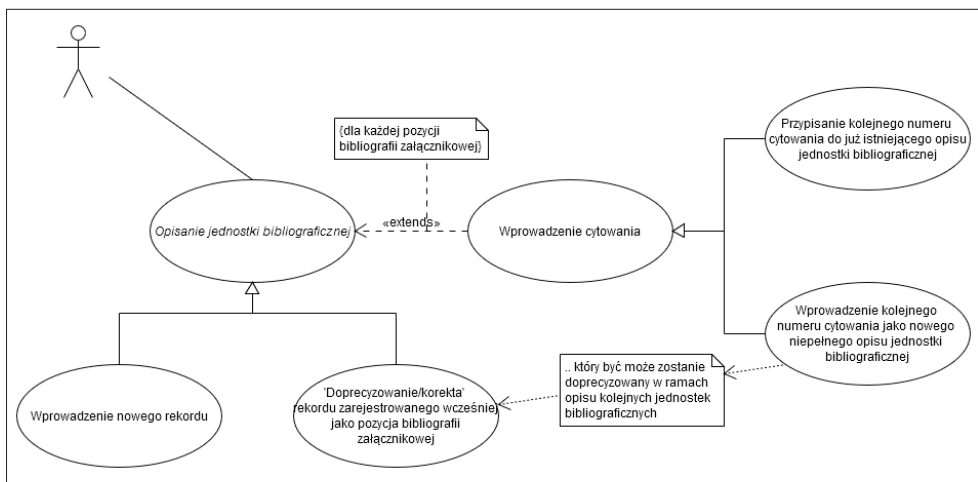
Rys. 7. Model klas w metodzie prospektywnej



Z modelu klas opracowanego dla metody prospektywnej wynika, że potencjalnie opisywaną jednostką bibliograficzną może być jednostka o dowolnej formie wydawniczej. Klasa jednostki bibliograficznej jest związana sama ze sobą relacją agregacji o nazwie „jest cytowana”, a kardynalności tej relacji zdefiniowano na „0..n”. Oznacza to, że bibliografia załącznikowa danej (konkretnej) jednostki bibliograficznej stanowi agregat (zbiór, który w szczególnym przypadku może być zbiorem pustym) innych (konkretnych) jednostek bibliograficznych, zaś dana jednostka bibliograficzna może wchodzić w skład wielu takich agregatów. Tak określony model, mimo że możliwy do zaimplementowania w postaci relacyjnej, może reprezentować całą sieć (a dokładniej skierowany graf) cytowań, czyniąc struktury danych zgodnie z nim zaimplementowane łatwo trawersowalnymi. Dodatkowo wyróżniono na diagramie klasę Autor, która związana jest z jednostką bibliograficzną relacją (0..n,1..n) co oznacza, że każdy autor rejestrowany w systemie powinien być autorem przynajmniej jednej jednostki bibliograficznej, zaś dana jednostka bibliograficzna nie musi mieć przypisanego autora (np. przypadki dokumentów z nieznanym autorem) lub może mieć przypisanego jednego lub więcej autorów. Taki zapis sugeruje, że autorzy są samodzielnymi bytami w systemie (a nie jedynie atrybutami tekstowymi jednostki bibliograficznej), zaś relacja „jest autorem” stanowi tylko wskazanie od konkretnej jednostki bibliograficznej do konkretnych autorów. W ogólnych rozważaniach takich samodzielnnych bytów pojawić się może więcej np. instytucja sprawcza, wydawca, itp., jednak powoływanie ich na potrzeby opisu przedstawianej metody zaciemniłoby tylko obraz. Próby przedstawienia całego zakresu informacyjnego systemu bibliograficznej bazy danych mogącego stanowić „model referencyjny” będą przedmiotem rozważań dalszych prac badawczych.

Na rysunku 8 przedstawiono model przypadków użycia charakteryzujący metodę prospektywną. Podobnie jak w modelu tradycyjnym, podstawową czynnością jest opisanie jednostki bibliograficznej z tym, że tutaj, jako że czynność ta może być realizowana na dwa wykluczające się wzajemnie sposoby, jest ona reprezentowana przez abstrakcyjny przypadek użycia.

Rys. 8. Model organizacji pracy dla rejestracji jednostek bibliograficznych metodą prospektywną



Może on być realizowany poprzez wprowadzenie nowego rekordu opisu lub poprzez doprecyzowanie (a w niektórych przypadkach wręcz skorygowanie) informacji o już istniejącej jednostce bibliograficznej, która wcześniej zaistniała w systemie jako element bibliografii załącznikowej. W trakcie opisywania jednostki bibliograficznej następuje rejestrowanie informacji o poszczególnych pozycjach bibliografii załącznikowej. Różnica w stosunku do modelu tradycyjnego polega na tym, że operuje się tutaj nie na rekordach cytowań, ale na obiektach jednostek bibliograficznych. Możliwe jest więc przypisanie do konkretnego numeru pozycji bibliografii załącznikowej jednostki bibliograficznej wcześniej opisanej lub też zarejestrowanie nowej, która być może w trakcie realizacji kolejnych opisów jednostek bibliograficznych będzie doprecyzowana i stanie się jednostką cytującą – wyraża to odpowiednia notatka zamieszczona na modelu.

Dopowiedzenia wymaga wcześniejszy komentarz o doprecyzowaniu lub wręcz skorygowaniu informacji podczas realizacji przypadku użycia „Doprecyzowanie/korekta”. Wydaje się oczywiste, że informacje o jednostce bibliograficznej, którą właśnie opisujemy (a jeśli robimy to z autopsji, to „mamy ją właśnie w ręce”) są bardziej wiarygodne od informacji zamieszczonych o tej jednostce w bibliografiach załącznikowych innych jednostek. Dlatego jest to najlepszy moment na weryfikację wprowadzonych wcześniej informacji i ich ewentualną korektę. Sytuacja jest mniej oczywista w przypadku rejestrowania informacji o danej jednostce przy okazji realizacji przypadku użycia „Przypisanie kolejnego numeru cytowania do już istniejącego opisu jednostki bibliograficznej”, ale istniejący już w systemie opis tej jednostki „kandydującej” różni się nieco od opisu w bibliografii załącznikowej.

Jeśli opis tej jednostki bibliograficznej został wprowadzony bądź wcześniej skorygowany w ramach przypadku użycia „Opisanie jednostki bibliograficznej”, to dane te oczywiście możemy uznać za wiarygodne. W przeciwnym przypadku nie mamy pewności, czy zarejestrowany wcześniej opis jest bardziej wiarygodny od posiadanego aktualnie. Jest to więc dobry moment na sięgnięcie do innego, bardziej wiarygodnego źródła lub oznaczenie rekordu jako „podejrzanego” i wymagającego późniejszej weryfikacji. Tak czy inaczej zidentyfikowanie takich niezgodności w momencie ich rejestrowania zawsze daje większą szansę na ich wyjaśnienie, gdyż mamy „przed sobą” materiał źródłowy, niż wykrywanie ich i korygowanie *post factum*.

Skuteczność wyszukiwania „kandydujących” rekordów spośród zarejestrowanych wcześniej zależy od przyjazności graficznego interfejsu użytkownika i efektywności jego „mechanizmów sugestii”, podpowiadających operatorowi możliwość ponownego użycia informacji o wcześniej zarejestrowanej jednostce bibliograficznej. Skuteczne mechanizmy nie powinny opierać się jedynie na wyszukiwaniu „dosłownym”, ale powinny implementować bardziej elastyczne algorytmy dopasowania wzorców, np. oparte na miarach podobieństw łańcuchów znakowych, jak np. metody zaproponowane w (Kamińska, 2017a).

Przedstawiony model metody prospektywnej jest modelem obiektowym (wykorzystuje związki generalizacji-specjalizacji). Natomiast większość popularnych systemów zarządzania bazami danych posługuje się modelami relacyjnymi, gdzie nie ma dosłownej implementacji wyżej wspomnianych związków. Zachodzi wobec tego konieczność transformacji modelu obiektowego do modelu relacyjnego. Zadanie to jest jak najbardziej możliwe, co więcej, znanych jest tutaj kilka podejść, których wybór powinien zależeć od wymagań funkcjonalnych i niefunkcjonalnych stawianych budowanemu systemowi (Keller, 1997).

Autorka, budując własną bibliograficzną bazę danych, stanęła przed koniecznością podjęcia wielu decyzji dotyczących architektury gromadzonych informacji, które ostatecznie doprowadziły do powstania logicznego modelu relacyjnego, a następnie modelu danych możliwego do wdrożenia na konkretnej platformie systemu zarządzania bazą danych. Tak powstały system, będący implementacją koncepcji przedstawionej w niniejszym artykule, dostarczył w relatywnie krótkim czasie materiału badawczego o wysokiej jakości. Oznacza to, że dzięki opisanej metodzie można nie tylko zapewnić wysoką jakość gromadzonych danych i rozpocząć na ich podstawie realizację zaawansowanych badań bibliometrycznych, ale również obniżyć ogólne koszty bieżącego zasilania tradycyjnych bibliograficznych baz danych.

5. Wnioski

Artykuł przedstawia problemy z jakimi współcześnie borykają się operatorzy krajowych, zarówno technicznych jak i humanistycznych, bibliograficznych baz danych gromadzących informacje w układzie tradycyjnym. Na konkretnych przykładach wykazano, że niespójności w danych oraz brak możliwości realizacji zaawansowanych badań bibliometrycznych, które bazowałyby na tych danych, nie wynika jedynie z niefrasobliwości operatorów zasilających te bazy, a przede wszystkim z przyjętej organizacji przestrzeni informacyjnej. Zmiana takiego stanu rzeczy nie będzie możliwa poprzez wprowadzanie kolejnych dyrektyw czy zaleceń dla wyżej wspomnianych operatorów, a jedynie poprzez reorganizację modelu danych

i oprogramowania na nim operującego, np. zgodnie z prospektywną metodą tworzenia trawersowalnych indeksów cytowań.

Zaproponowaną w artykule metodę, na tle innych opisywanych w piśmiennictwie badawczym czy implementowanych i stosowanych w bibliograficznych bazach danych/indeksach cytowań, wyróżniają ważne cechy:

- dane gromadzone są bezpośrednio w relacyjnej bazie danych, ale w układzie pozwalającym na łatwą reprezentację tych danych w postaci grafów cytowań o trawersowalnych krawędziach;
- system może realizować funkcje bibliograficznej bazy danych poprzez bieżące gromadzenie, składowanie, przetwarzanie i udostępnianie danych bibliograficznych;
- system może realizować funkcje indeksu cytowań dostępnego wprost z poziomu relacyjnej bazy danych, która może stanowić zarazem podstawowe środowisko analityczno-obliczeniowe;
- funkcje bibliograficznej bazy danych oraz indeksu cytowań mogą być realizowane współbieżnie, bez potrzeby „zamrażania” czy „materializacji” danych przed rozpoczęciem analiz;
- prospektywny sposób wprowadzania danych oszczędza czas operatora, a jednocześnie sprzyja gromadzeniu danych o bardzo dobrej jakości;
- brak potrzeby migracji danych w różnych formatach plików, pomiędzy różnymi platformami w tym brak potrzeby posługiwania się formatem MARC (w środowisku zbudowanym przez autorkę format MARCXML służył jedynie fakultatywnemu udostępnieniu danych dla innych systemów);
- opis pozycji cytujących oraz cytowanych przechowywany jest w jednej strukturze danych (tabeli), co pozwala uniknąć redundancji przechowywanych informacji.

Dalsze badania autorki w tych obszarach koncentrować się będą na opracowaniu ogólnych obiektowych i relacyjnych modeli referencyjnych, transformacji zastanych modeli danych do modeli umożliwiających zastosowanie metody prospektywnej oraz udoskonalaniu funkcji oceniającej podobieństwo rekordów bibliograficznych, która może znaleźć zastosowanie w implementacji przypadku użycia „Przypisanie kolejnego numeru cytowania do już istniejącego opisu jednostki bibliograficznej”.

Na koniec warto zauważyć, że z przedstawionych treści wynika, iż możliwości informacyjne czy analityczne danych systemów bibliograficznych nie zależą jedynie od „zakresów gromadzonych danych”, na których opisie najczęściej poprzestają badacze charakteryzując dane rozwiązanie. Potencjał niedopuszczania do powstawania opisanych wcześniej anomalii oraz możliwości realizacji zaawansowanych obliczeń bibliometrycznych wynika natomiast przede wszystkim z architektury składowanych informacji w poszczególnych systemach, których opis mogłaby ułatwić zaproponowana w artykule notacja. A ponieważ trudno w literaturze znaleźć relewantne informacje na temat opisu poszczególnych systemów z tej perspektywy, mogłyby one prawdopodobnie stanowić wkład w rozwój badań nad tą dziedziną.

Bibliografia

- Al, U.; Taşkın, Z.; Düzyol, G. (2012). Use of Social Network Analysis in Bibliometric Researches. In: *Information on E-motion. Proceedings. BOBCATSSS 2012, 20th International Conference on Information Science, 23–25 January*. Amsterdam: Bock+Herchen Verlag, 40–44.
- Booch, G.; Rumbaugh, J.; Jacobson, I. (2002). *UML przewodnik użytkownika*. Warszawa: Wydaw. Naukowo-Techniczne.
- Cockburn, A. (2004). *Jak pisać efektywne przypadki użycia*. Warszawa: Wydaw. Naukowo-Techniczne.
- Drabek, A. (2009). Konferencja „Bibliograficzne bazy danych: kierunki rozwoju i możliwości współpracy”, Bydgoszcz, 27–29 maja 2009 roku. *Bibliotheca Nostra*, 2(18), 116–121.
- Garczyńska, M. (2013). Bibliograficzne bazy danych w ocenie pracowników i jednostek szkół wyższych w Polsce [online]. *Materiały konferencyjne EBIB*, 24, [06.12.2016], http://open.ebib.pl/ojs/index.php/Mat_konf/article/view/24/164.
- Garfield, E. (1955). Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas. *Science*, 122(3159), 108–111.
- Kretschmer, H.; Kretschmer, T. (2007). A New Centrality Measure for Social Network Analysis Applicable to Bibliometric and Webometric Data. *COLLNET Journal of Scientometrics and Information Management*, 1(1), 1–7.
- Ilyas, I. F.; Chu X. (2015). Trends in Cleaning Relational Data: Consistency and Deduplication. *Foundations and Trends in Databases*, 5(4), 281–393.
- Kamińska, A. M. (2017a). Miary podobieństw łańcuchów znakowych a deduplikacja rekordów w bibliograficznych bazach danych. *Przegląd Biblioteczny*, 4 [w druku].
- Kamińska, A. M. (2017b). O rozwoju graficznych języków komunikacji. *Zagadnienia Informatyki Naukowej*, 2(110) [w druku].
- Kamińska, A. M. (2017c). Potencjał bibliometryczny bibliograficznej bazy danych CYTBIN w świetle prostych i złożonych wskaźników analitycznych. *Bibliotheca Nostra* [w druku].
- Kamińska, A. M. (2017d). Tam, gdzie zaczyna się bibliometria, czyli jak pozyskać materiał analityczny z autopsji [online]. *Biuletyn EBIB*, 3(173), [28.08.2017], <http://open.ebib.pl/ojs/index.php/ebib/article/view/534/698>.
- Kamińska, A. M. (2017e). Wizualizacje wybranych wskaźników bibliometrycznych na przykładzie bibliograficznej bazy danych CYTBIN. *Toruńskie Studia Bibliologiczne*, 2 (19) [w druku].
- Kamińska, A. M. (2018). Zastosowanie metod analizy sieci społecznościowych w bibliometrii i webometrii. Miary i narzędzia. *Nowa Biblioteka. Usługi, technologie informacyjne i media*, 2 (29) [w druku].
- Kolasa, W. M. (2011). Retrospektywny indeks cytowań w humanistyce. Koncepcja, metoda, zastosowanie. *Przegląd Biblioteczny*, 79(4), 466–486.
- Marszakowa-Szajkiewicz, I. (2009). *Badania ilościowe nauki. Podejście bibliometryczne i webometryczne*. Poznań: Uniwersytet im. Adama Mickiewicza.
- Seweryn, A.; Swoboda, I. (2014). Cytowania w polskich bazach bibliograficznych. *PTINT Praktyka i Teoria Informatyki Naukowej i Technicznej*, 1(85), 3–20.
- Keller, W. (1997). Mapping Objects to Tables: A Pattern Language. *Proceedings of the 1997 European Conference on Pattern Languages of Programming (EURO PLoP'97)* [online], Software Research and Industry, Dirk Riehle, [06.12.2016], <http://dirkriehle.com/community-service/hillside-group/europlop-1997/p13final.pdf>.
- Woźniak-Kasparek, J. (2011). *Wiedza i język informacyjny*. Warszawa: Wydawnictwo SBP.

ProBIT – the Prospective Method of Creating the Traversable Citation Indices versus the Contemporary Issues in Information Space Organization in Traditional Bibliographic Databases

Abstract

Purpose/Thesis: The article presents an original method of building bibliographic databases, conducive to the efficient gathering of high quality bibliographic data which, without any further transformations, can serve as a source of advanced bibliometric analysis (i.e. not limited to simple summaries calculated on the basis of data collected in a relational form, but those that need traversing the edges of the graph representing the networks of article citations and cooperation among their authors).

Approach/Methods: Encountering a problem with gathering a significant volume of bibliographic data to conduct her research, the author evaluated a number of tools used to collect the bibliographic data, but none of them offered the expected ergonomics, sufficient care about the quality of the data collected, or data model enabling further advanced bibliometric analysis. Drawing conclusions from the negative experience gained while working with these tools, the author proposed her own original method, which is presented using basic UML notation.

Results and conclusions: The presented method has proved effective and helped the author – in a relatively short time – to gather data of very good quality on more than 36 thousand articles and over 22 thousands of authors. These data were used directly for the implementation of bibliometric research, which will be discussed in a separate work.

Practical implications: First part of the article presents risks brought by gathering bibliographic data in so-called traditional form. Using examples from selected national bibliographic databases the author proved that these risks appear as a number of anomalies that have been categorized by the author, and which can be prevented by the proposed method. The method in question enablesto increase the quality of data collected in the bibliographic databases and reduce the cost of data gathering by improving the entire process.

Originality/Value: Polish researchers in the area of book studies rarely pursue the subject of constructing bibliographic databases and usually limit this issue to defining only ‘sets of entities and attributes’ being processed. The author argues that the final result (the quality of data, the total cost of data collecting, the possibility of further analysis) is also affected by the data model (following which the data are collected) and the appropriate organization of the collecting process.

Keywords

Bibliographic database. Information space. ProBIT. Prospective method. Traversable citation index.

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

Kontakt z autorką:

Anna.Kaminska@polsl.pl

Biblioteka Główna Politechniki Śląskiej

ul. Kaszubska 23

44–100 Gliwice

Historia i współczesność programu CIP (Cataloguing in Publication)

Dorota Siwecka

*Institut Informacji Naukowej i Bibliotekoznawstwa
Uniwersytet Wrocławski*

Abstrakt

Cel/Teza: Celem artykułu jest podsumowanie 50-letniej działalności programu katalogowania w procesie wydawniczym (CIP) oraz określenie roli programu w rozwoju informacji bibliograficznej w drugim dziesięcioleciu XXI wieku.

Koncepcja/Metody badań: W pracy posłużono się metodą analizy piśmiennictwa. Analizie poddano zarówno artykuły dotyczące omawianej problematyki, jak i strony WWW instytucji odpowiedzialnych za realizację programu w poszczególnych krajach Europy, Stanach Zjednoczonych, Kanadzie i Australii.

Wyniki i wnioski: Analiza rozwoju programu w badanych krajach pokazuje, że w większości przypadków nie rozwinął się on w pełni. W niektórych został zastąpiony inną formą informacji prospektywnej (np. w Niemczech). Najlepiej program rozwija się w Stanach Zjednoczonych jako narzędzie egzekwowania egzemplarza dla Biblioteki Kongresu.

Oryginalność/Wartość poznawcza: W literaturze pojawiają się okresowo publikacje dotyczące programu CIP. Brakuje natomiast opracowań podsumowujących ostatnie 20-lecie funkcjonowania programu w kontekście rozwoju nowych technologii. W założeniu niniejszy artykuł ma wypełnić tę lukę.

Słowa kluczowe

Katalogowanie w trakcie procesu publikacji. Opis bibliograficzny. Opis CIP.

Otrzymano: 31 maja 2017. Zrecenzowano: 10 lipca 2017. Zaakceptowano: 11 września 2017.

1. Wprowadzenie

Najważniejsze organizacje i instytucje bibliotekarskie na świecie (IFLA, Biblioteka Kongresu, Biblioteka Brytyjska) od wielu lat usiłują wykreować swojego rodzaju model informacji bibliograficznej. Aby tego dokonać, wspólnie starają się ustalić zespół czynników tworzących informację bibliograficzną, które w głównej mierze wpływają na ukształtowanie bibliografii narodowych, a pośrednio również na inne (specjalne) spisy czy bazy bibliograficzne. W ciągu ostatnich dziesięcioleci, obok projektów ściśle związanych z bibliografią, na formowanie się tego modelu miały wpływ również inicjatywy, które przyczyniły się do tworzenia oraz rozpowszechniania bibliografii. Jedną z nich był program katalogowania dokumentów w trakcie procesu wydawniczego (ang. *cataloguing in publication*; CIP). Celem artykułu jest uaktualnienie informacji na temat programu katalogowania w trakcie procesu wydawniczego, który w Polsce nigdy nie stał się elementem szerszej współpracy między środowiskiem bibliotekarskim i wydawniczym, a także wskazanie zmian, jakie nastąpiły w funkcji pełnionej przez CIP w ostatnich latach w krajach, w których program ten bardzo dobrze funkcjonuje.

Na straty czasu spowodowane wielokrotnym opracowywaniem tego samego tytułu w różnych instytucjach zwrócił uwagę już w 1876 r. Justin Winsor, bibliotekarz z Harvard College. Pod koniec XIX w. Henry B. Wheatley w pracy *How to Catalogue a Library* (Wheatley, 1889, 69) zaproponował, by każdy wydawca zamieszczał w swoich książkach kartę katalogową, sporządzoną przez rzetelną instytucję według ujednoczonych reguł. Zdawał sobie jednak sprawę, jak bardzo skomplikowane byłoby to wówczas przedsięwzięcie. Na początku lat 20. XX w. Biblioteka Kongresu rozpoczęła proces centralnego drukowania kart katalogowych dla swoich nabytków. System nie do końca jednak odpowiadał zapotrzebowaniu środowiska bibliotekarskiego. W praktyce sporządzanie opisu, drukowanie kart, a następnie rozsyłanie ich po kraju znacznie opóźniało proces dostarczania publikacji do rąk użytkownika końcowego. Dodatkową trudnością było także zorientowanie się w potrzebach konkretnych bibliotek i dostarczanie odpowiedniej liczby kart katalogowych danej książki, w zależności od liczby i rodzaju używanych w placówce katalogów. W latach 50. i 60. XX w. pojawiła się koncepcja katalogowania książek w trakcie procesu wydawniczego, tak by książka trafiająca do biblioteki posiadała już wydrukowany opis bibliograficzny (Kończak, 2005, 65; Ramlau-Klekowska, 1990, 48). Realizacją tego pomysłu był międzynarodowy program *Cataloguing in Publication*¹. Wynikiem takiego postępowania było umieszczenie w całym nakładzie książki opisu bibliograficznego, który, stworzony według ujednoczonych zasad, pozwalał również na szybkie i poprawne przejmowanie informacji bibliograficznych do innych celów informacyjnych, np. księgarskich.

W trakcie rozwoju programu nazywano tę formę katalogowania także „katalogowaniem u źródła” (ang. *cataloguing in source*), tzn. tam, gdzie książka powstaje (w wydawnictwie) lub „katalogowaniem w czasie druku” (ang. *cataloguing in print*), tzn. w okresie produkcji drukarskiej. Termin „katalogowanie w trakcie procesu publikacyjnego” (ang. *cataloguing in publication*) jest jednak terminem najbardziej ogólnym, ponieważ obejmuje również wydania niedrukowane, np. mikrofilmowe.

2. Rozwój programu CIP

Praktycznie katalogowanie w toku procesu publikacyjnego rozpoczęło się w Związku Radzieckim² w 1961 r. Do zamieszczenia opisu bibliograficznego w książce zobowiązany był wydawca, który miał współpracować z bibliotekami i izbami książki poszczególnych republik, tworząc opis zgodny z klasyfikacją biblioteczną i księgarską oraz odpowiednimi instrukcjami katalogowania. Rezultatem była drukowana w książce makietka karty katalogowej, którą można było skopiować w bibliotece. Zawierała ona opis bibliograficzny, symbol klasyfikacji rzeczowej oraz adnotację informującą o charakterze książki i jej treści (Daszkowski, 1975, 154; Ramlau-Klekowska, 1990, 51).

W Stanach Zjednoczonych pierwszą próbę wprowadzenia katalogowania w procesie publikacyjnym przeprowadzono w latach 1958–1959. „Katalogowaniem u źródła” (ang. *cataloguing in source*) zajęła się wówczas Biblioteka Kongresu (BK). Koszty sporządzania

¹ Pisownia amerykańska: *cataloging in publication*.

² O specjalnych uchwałach i rozporządzeniach dotyczących CIP w Związku Radzieckim, a także o postanowieniach instrukcji wydanej w 1961 r. można przeczytać w artykułach: (Daszkowski, 1975, 153–154; Kończak, 2005, 70; Ramlau-Klekowska, 1990, 51).

opisów (zwłaszcza te leżące po stronie wydawców), czas poświęcony na przesyłanie pocztą makiet książek do BK i opisów bibliograficznych do wydawcy, a także błędy jakie pojawiły się już po wydrukowaniu dzieł (głównie dotyczące opisu fizycznego) oraz wątpliwości co do kierunku rozwoju tego programu, sprawiły, że eksperyment nie powiódł się i na 12 lat zarzucono tę praktykę. Głównym problemem okazało się zachęcenie wydawców do udziału w projekcie. W przeciwieństwie do bibliotek i bibliotekarzy, wydawcy w trakcie realizacji programu ponosili jedynie koszty, a nie widzieli w nim żadnych korzyści (Kończak, 2005, 65).

Kolejną, tym razem udaną, próbę podjęto w 1971 r. Program nazwano wówczas „Cataloging in Publication”, by odróżnić go od poprzedniego nieudanego eksperymentu. Wydawcy tym razem docenili wartość gotowego, sporządzonego przez wyspecjalizowaną instytucję (tzw. centralę CIP) opisu, który mogli wykorzystać m.in. w swoich katalogach prospektywnych i bazach danych i chętnie przystąpili do kolejnej próby wdrożenia programu na terenie Stanów Zjednoczonych (Kończak, 2005, 65; Ramlau-Klekowska, 1990, 49). Dodatkowym argumentem przemawiającym za udziałem w przedsięwzięciu był podkreślany przez biblioteki fakt obniżenia kosztów katalogowania w poszczególnych placówkach, a co za tym idzie wzrost środków na zakup nowości wydawniczych (Daszkowski, 1975, 158).

W sporządzanym opisie BK ograniczyła się do podstawowej informacji katalogowej – tytułu dzieła, nazwy serii, numeru kontrolnego BK (Library of Congress Control Number – LCCN) i numeru ISBN, a także hasła opisu bibliograficznego, uwag, symboli Klasyfikacji Dziesiątej Deweya (KDD) i Klasyfikacji Biblioteki Kongresu (KBBK), haseł przedmiotowych i haseł dodatkowych opisu – pominięto więc opis fizyczny i adres wydawniczy, w których wystąpiło poprzednio najwięcej błędów (Daszkowski, 1975, 159–160). Biblioteka Kongresu dostarczała wydawcy gotowy opis w ciągu 10 dni od uzyskania od niego odpowiednich materiałów³. Opis CIP drukowany jest współcześnie na stronie redakcyjnej książki oraz trafia do bazy BK, którą rozprowadza się raz w tygodniu w postaci elektronicznej (w formacie MARC) poprzez Cataloging Distribution Service (CDS) do dużych bibliotek, serwisów bibliograficznych i księgarzy na całym świecie. Dzięki temu opisy CIP służą też jako forma bibliograficznej informacji prospektywnej, która ułatwia politykę gromadzenia zbiorów oraz proces zamawiania książek w bibliotekach (LC, b.d.-a).

Podobne próby podjęto w latach 70. XX w. w Australii, Brazylii, Republice Federalnej Niemiec, Kanadzie i Wielkiej Brytanii. W Australii od 1974 r. program prowadzony jest przez dział biblioteki narodowej odpowiedzialny za tworzenie bibliografii narodowej, współpracujący z narodową agencją ISBN i narodowym ośrodkiem ISSN. Na podstawie danych dostarczanych przez wydawców centrala CIP w ciągu 10 dni roboczych sporządza opis katalogowy, a następnie dostarcza go do wydawcy. Informacje o publikacjach zamieszcza w narodowej bibliograficznej bazie danych – Australian National Bibliographic Database, która udostępnia swoje dane poprzez platformę Trove bibliotekom, księgarzom i innym użytkownikom na całym świecie. Dane te widoczne są również dla wyszukiwarek internetowych (np. Google) (NLA, b.d.).

W Wielkiej Brytanii opisy CIP tworzone były od 1975 r. przez Bibliotekę Brytyjską, która umieszczała je w bazie komputerowej w formacie zgodnym z formatem British National Bibliography (BNB) (wcześniej format UKMARC, obecnie MARC 21) jako zapowiedzi wydawnicze na dwa miesiące przed deklarowaną datą wydania. Od 1995 r. opisy CIP w Wielkiej

³ Obecnie BK nie podaje, jak długo trzeba czekać na gotowy opis, zaleca natomiast, by przysyłać dane o wydawnictwach na pół roku przed ich opublikowaniem – (LC, b.d.-a).

Brytanii tworzone są przez firmę prywatną Bibliographic Data Services Limited, która na zasadzie porozumienia z Biblioteką Brytyjską zbiera informacje od wydawców na 4 miesiące przed planowaną publikacją. Gotowe rekordy przekazywane są następnie do Biblioteki Brytyjskiej, gdzie trafiają do BNB z dopiskiem „Prepublication record” (British Library, 2014).

W Niemczech program CIP realizowany był w latach 1975–2002. Opisy (obejmujące piśmiennictwo niemieckojęzyczne z terenów Niemiec, Austrii i Szwajcarii) sporządzane w ramach programu publikowane były w serii „N” bieżącej bibliografii narodowej. Program zamknięto po przeprowadzeniu w 1999 r. ankiet wśród wydawców niemieckich, z których wynikało, że opisy CIP w kontekście rozwijających się technologii i dostępu do Internetu nie są już na tyle potrzebne, by zamieszczać je w książkach. Ponadto wydawcy zbyt późno przesyłali dane (często niepełne i błędne) do agencji CIP, co sprawiało, że opis CIP pojawiał się niemal równocześnie z pełnym opisem publikacji w bibliografii narodowej (Hammer, 1999, 11–13).

Oprócz modelu organizacyjnego centrali CIP zlokalizowanej w jednym ośrodku czy instytucji (np. w Australii, Holandii, Nigerii, Niemczech, USA i Wielkiej Brytanii) w niektórych krajach przyjęto model rozproszony. Ze względu na dużą powierzchnię kraju postąpiono tak w Kanadzie, organizując początkowo trzy ośrodki odpowiedzialne za tworzenie opisów CIP: publikacji anglojęzycznych w Toronto i Vancouverze, a francuskojęzycznych w Quebecu. W 2011 r. w Kanadzie funkcjonowały ośrodki w Ottawie oraz w Montrealu (dla wydawnictw francuskojęzycznych), a także oddzielne biura CIP dla publikacji rządowych zlokalizowane w Toronto, Winnipeg i Ottawie (LAC, b.d.). Nad całością przedsięwzięcia czuwała Biblioteka Narodowa Kanady, która zamieszczała wszystkie opisy CIP w bibliografii narodowej *Canadiana*. Od 2016 r. zrezygnowano z programu CIP dla publikacji urzędowych. Obecnie wydawnictwa anglojęzyczne opracowywane są przez Library and Archives Canada (LAC), a francuskojęzyczne przez Bibliothèque et Archives nationales du Québec w ciągu 15 dni roboczych od otrzymania materiałów od wydawcy (LAC, 2015). Dane dotyczące planowanych publikacji rejestrowane są w katalogu AMICUS, w innych katalogach bibliotecznych, a także w *Books in Print* (LAC, 2013).

W Polsce próbę zaprojektowania systemu CIP podjęto już w 1975 r. Początkowo zamierzano przejść model funkcjonujący w ZSRR, ale różnice między wielkością i charakterem instytucji wydawniczych oraz pojawiające się wówczas wątpliwości co do realizacji pomysłu sprawiły, że ostatecznie zrezygnowano z wprowadzania programu. Dopiero na początku lat 80. XX w. stworzono propozycję dwóch modeli: scentralizowanego (z Instytutem Bibliograficznym BN jako główną instytucją) oraz zdecentralizowanego (zblizonego do modelu radzieckiego). Przeprowadzone wówczas referendum wśród bibliotek wojewódzkich pokazało, że większość z nich opowiedziała się za pierwszym rozwiązaniem, ponieważ gwarantowało ono jednolitość i poprawność opisów. Wydawcy nie zostali wówczas poproszeni o przedstawienie własnej opinii. W 1985 r. przystąpiono do komputeryzacji *Przewodnika Bibliograficznego*, „który miał stanowić bazę warsztatową i personalną dla centrali CIP” (Kołacz, 2005, 73). Z uczestnictwa w programie skorzystała jedynie „Nasza Księgarnia”, przystępując do niego pod koniec 1986 r. (Ramlau-Klekowska, 1990, 58–59). W maju 1988 r. minister kultury i sztuki wydał oficjalne zarządzenie w sprawie drukowania opisów katalogowych i adnotacji w wydawnictwach zwartych (*Zarządzenie...*, 1988). Obowiązek sporządzania opisów CIP powierzono Instytutowi Bibliograficznemu BN (IB BN), który opracowywał bibliografię narodową. Programem objęto wszystkich wydawców (około 600) i wszystkie typy publikacji książkowych. Instytut Bibliograficzny zobowiązał

się do tworzenia opisu w ciągu 48 godzin od momentu dostarczenia materiałów. Wydawcy sami musieli zdecydować, w którym momencie produkcji wysłać materiały do Biblioteki Narodowej. Mimo odpowiednich przepisów prawnych w 1989 r. IB BN opracowywał opisy książek jedynie dla 12 firm wydawniczych. Warto jednak odnotować, że w Polsce po raz pierwszy w skali światowej (z inicjatywy wrocławskiego Wydawnictwa Kartograficznego) wprowadzono opisy CIP do map (Kończak, 2005, 74–75; Ramlau-Klekowska, 1990, 59–61).

W swoim artykule z 2005 r. Teresa Kończak (2005, 76) napisała w odniesieniu do polskiej realizacji programu CIP:

Szanse, że będzie lepiej, są raczej nikłe, bo powoli rolę programu CIP przejmują Internet. Coraz więcej bibliotek ma do niego dostęp, a prezentowane tam katalogi wielu dużych bibliotek umożliwiają znalezienie poprawnego opisu bibliograficznego każdej nowej książki.

Słaby rozwój programu CIP w Polsce ma swe źródło w braku zainteresowania nim wydawców (w badaniach z 2004 r. BN podała, że współpracują z 30 wydawnictwami (LC, 2005, 11)). Nie bez znaczenia jest tutaj znaczne rozdrobnienie rynku wydawniczego w Polsce po 1990 r., pogorszenie się sytuacji finansowej bibliotek, wynikające z kryzysu społeczno-ekonomicznego końca lat 80. XX w. oraz brak do 2000 r. odpowiedniej infrastruktury telekomunikacyjnej i powszechnego dostępu do Internetu. Ponadto opisy CIP nie stanowią w Polsce podstawy do tworzenia bibliografii prospektywnych (opisów tych jest zresztą niewiele – ok. 350 rocznie (LC, 2005, 4)), nie są też dostępne z poziomu katalogu komputerowego dla użytkowników Biblioteki Narodowej, co z pewnością nie wpływa korzystnie na zainteresowanie programem wydawców.

2.1. *Badania rozwoju programu*

W 1975 r. sytuację programu w kilku państwach (USA, ZSRR, Brazylii i Australii⁴) przeanalizował i opisał L. R. Swindley (Swindley, 1975). W swoim raporcie omówił warunki, jakie powinny być spełnione w kraju, który planował wdrożyć scentralizowaną formę programu CIP. W 1982 r. zorganizowano międzynarodowe spotkanie ekspertów CIP w Ottawie. Biorący w nim udział bibliotekarze i wydawcy określili wówczas CIP „dzieciństwem Kongresu Paryskiego” (Ramlau-Klekowska, 1990, 51)⁵. Rekomendacje tego spotkania (IFLA, 1986) stały się później podstawą do opracowania przewodnika CIP i poradnika jego realizacji: *Guidelines for Cataloguing-in-Publication* (Anderson, 1986).

W 1990 r. Hope Clement przedstawiła wyniki badań, z których wynikało, że 8 z 12 badanych bibliografii narodowych realizuje program CIP. Każda z tych narodowych agencji bibliograficznych przyznała, że wprowadzenie CIP przyspieszyło proces katalogowania bibliografii narodowych (Clement, 1990, 8). Osiem lat później, w 1998 r., Biblioteka Narodowa Kanady na zlecenie CDNL rozesała ankiety dotyczące organizacji programu CIP na świecie. W odpowiedzi na ankietę 18 z 68 narodowych agencji bibliograficznych oświadczyło, że współpracują ze środowiskiem wydawców poprzez program CIP (NLC, 1998, 7).

⁴ Zamiary wprowadzenia CIP w RFN, Meksyku, Wielkiej Brytanii i Kanadzie autor omówił pobieżnie (Ramlau-Klekowska, 1990, 49).

⁵ Już na wczesnym etapie wprowadzania różnych wersji katalogowania centralnego, okazało się, że nie jest ono możliwe bez uzgodnienia ujednoliconych w skali przynajmniej krajowej przepisów katalogowania. W trakcie Kongresu Paryskiego doszło do pierwszych ustaleń zmierzających do realizacji tej idei.

Badania dotyczące rozwoju programu CIP prowadzi także Biblioteka Kongresu⁶. Wyniki badań przeprowadzonych w latach 2003–2004, opublikowane w 2005 r., wykazały, że programy CIP w poszczególnych krajach⁷ różniły się między sobą zarówno organizacją samego programu (forma zdecentralizowana lub scentralizowana), jak i rejestrowanymi formami dokumentów (książki, dokumenty muzyczne, mikroformy, wydawnictwa ciągłe itd.), formą dostarczania danych przez wydawców (wyłącznie w formie drukowanej, wyłącznie w formie elektronicznej, w obu postaciach, faksem albo osobiście), a także wymaganiami stawianymi wydawcom (nie w każdym kraju wydrukowanie opisu CIP na stronie redakcyjnej jest obowiązkowe i nie każda narodowa centrala bibliograficzna w zamian za sporządzenie opisu wymaga dostarczenia egzemplarza obowiązkowego) (LC, 2005).

W 2011 r. opisy CIP tworzone były m.in. w Australii, Czechach, Estonii, Kanadzie, na Litwie, w Nowej Zelandii, Polsce, Portugalii, Rosji, Słowenii, USA i Wielkiej Brytanii. Do jesieni 2011 r. na kształt programu CIP funkcjonowało także tworzenie opisów w Islandii, ale wyłącznie na potrzeby małych bibliotek, w których nie było odpowiednio przeszkolonych katalogerów⁸. Na Łotwie nie realizowano programu CIP, ale wydawcy, którzy chcieli otrzymać numer ISBN, musieli wysłać do narodowej centrali bibliograficznej odpowiednie metadane w postaci elektronicznej. Były one następnie zapisywane w bazie danych bibliografii narodowej, a po otrzymaniu egzemplarza obowiązkowego odpowiednio uzupełniane⁹.

W trakcie konferencji sekcji IFLA w Warszawie w 2012 r. „Bibliography in a Digital Age” (spotkanie satelickie towarzyszące kongresowi IFLA 2012) potwierdzono, że dane prospektywne, m.in. takie jak opisy CIP, nadal są wartościowe (IFLA WG, 2015).

3. Zalety programu

Program CIP w swoim założeniu miał znacznie obniżyć koszty katalogowania w poszczególnych bibliotekach. Ponadto opracowane w ten sposób wydawnictwa miały szybciej trafiać do rąk użytkowników. Opis CIP znajdujący się w książce miał być przydatny zwłaszcza w bibliotekach, w których brakuje dostępu do Internetu i które nie mogą ściągnąć gotowego opisu np. z bazy biblioteki narodowej albo w sytuacji, kiedy nie byłby możliwy dostęp do tych baz. Zapewne dlatego program szybko rozwinął się w okresie zmniejszania budżetów w bibliotekach i zwiększania się kosztów katalogowania w Stanach Zjednoczonych w latach 70. XX w. (Daszkowski, 1975, 160). Katalogowanie w trakcie procesu wydawniczego pozwalało również na sprawniejsze udostępnienie książki czytelnikowi – w niektórych bibliotekach dzięki opisom CIP można zarezerwować książkę jeszcze zanim trafi ona na biblioteczną półkę (Newlen, 1991, 42).

⁶ Jeden z raportów ukazał się w 1986 r. (Mendenhall, 1982).

⁷ Program CIP funkcjonował w 25 krajach z 42, które odpowiedziały na ankietę.

⁸ Informacje na podstawie ankiety przeprowadzonej przez autorkę w październiku 2011 r. na potrzeby rozprawy doktorskiej. W ankiecie pojawiły się 4 pytania dotyczące programu CIP: 1) Czy w kraju funkcjonuje program CIP? 2) Kto jest odpowiedzialny za tworzenie opisów CIP? 3) Czy opisy te są wykorzystywane przy sporządzaniu bieżącej bibliografii narodowej? 4) W jaki sposób wydawcy mogą wziąć udział w przedsięwzięciu oraz prośba o oszacowanie procentowego udziału wydawców. Ankietę rozesłano do 34 narodowych agencji bibliograficznych, z których 17 odpowiedziało na pytania autorki.

⁹ Informacje na podstawie ankiety przeprowadzonej przez autorkę w październiku 2011 r.

Podczas „Międzynarodowego Kongresu na temat Bibliografii Narodowych”, który odbył się w Paryżu w 1977 r. stwierdzono, że CIP jest pokrewną koncepcją w stosunku do zadań bieżącej bibliografii narodowej, a tworzenie opisów CIP uznano za rozszerzoną działalność narodowych central bibliograficznych. Zauważono wówczas również, że zbiór opisów CIP wydawany w niektórych krajach łącznie z bibliografią narodową, stanowi bibliograficzną informację prospektywną dla bibliotekarzy i użytkowników informacji (IFLA, 1979, 19). Przykładem może być ukazujące się w Czechach elektroniczne czasopismo *O.K. Ohlášené Knihy* (O.K., 2017), niemieckie *Neuerscheinungsdienst*, brytyjskie opisy publikowane w BNB. Dla bibliotek jest to źródło pomocne zwłaszcza przy kształtowaniu polityki gromadzenia zbiorów, a dla wydawców forma reklamy (udział wydawców w programie, który nie zapewnia dystrybucji opisów wydaje się być dla nich dużo mniej atrakcyjny).

Program CIP miał również, poprzez założenie współpracy środowiska wydawców z narodowymi centralami bibliograficznymi, zniwelować problem nieterminowego dostarczania egzemplarzy obowiązkowych (EO) do narodowych central bibliograficznych. W zamian za stworzenie opisu CIP wydawca powinien po opublikowaniu dzieła przesłać do agencji CIP EO w celu sprawdzenia poprawności danych i uzupełnienia opisu o informacje wcześniej niedostępne (takie jak paginacja, ilustracje, format itd.).

Współpraca bibliotekarzy i wydawców w trakcie realizacji programu pokazała również, że wiedza bibliotekarska nie jeden raz okazała się pomocna w pracy wydawniczej – m.in. przy kształtowaniu tytułatury dzieła (zwłaszcza dzieł wielotomowych, adaptacji, wznowień), ustalaniu kolejności wydań czy prawidłowym przydziale numerów ISBN. Ponadto wydawcy mogli wykorzystać gotowe opisy do własnej dokumentacji firmowej, reklamy i sporządzania katalogów wydawniczych (Ramlau-Klekowska, 1990, 59).

Mimo wymienionych zalet program katalogowania w procesie wydawniczym nie cieszy się popularnością, a dane zbierane w jego trakcie rzadko wykorzystuje się przy tworzeniu bieżących bibliografii narodowych (wyjątkami są Australia, Estonia, Litwa, Nowa Zelandia, Słowenia, Stany Zjednoczone czy Wielka Brytania). Najprawdopodobniej szybki rozwój nowych technologii i powszechny dostęp do Internetu i informacji bibliograficznej przyczyni się (podobnie, jak to miało miejsce w Niemczech czy Islandii) do stopniowego odejścia od programu również w innych krajach. Umocnienia pozycji programu CIP można się natomiast spodziewać w krajach, w których rozszerza się funkcje, jakie ma spełniać (np. w USA czy Kanadzie, gdzie służy m.in. do egzekwowania egzemplarzy dla bibliotek).

4. Opis bibliograficzny CIP

Według międzynarodowych zaleceń opis CIP powinien zawierać minimalną liczbę ustalonych elementów oznaczonych zgodnie z ISBD, m.in. międzynarodowy znormalizowany numer wydawnictwa (ISBN, ISSN) (Cybulski, 1978, 228). Ma to swoje odzwierciedlenie w elementach obowiązkowych, jakie wydawca musi podać przy wypełnianiu formularza (m.in. w USA, Australii, Wielkiej Brytanii).

W rekomendacjach ze spotkania w Ottawie w 1982 r. zamieszczono postanowienia dotyczące dwóch typów opisu CIP – zamieszczanych w wykazach prospektywnych i umieszczanych w książkach. W drugim przypadku, bardziej rygorystycznym, wymaga się przestrzegania międzynarodowych zasad opisu bibliograficznego i jednocześnie pomijania elementów,

które nie są znane w momencie tworzenia opisu. Zalecano również, by opis CIP zawierał, oprócz opisu bibliograficznego, hasło opisu bibliograficznego i opis rzeczowy (IFLA, 1986).

W przypadku pierwszego rodzaju opisu (trafiającego do bazy) wymagana jest możliwa kompletność danych oraz jednoznaczna ich identyfikacja jako opisu CIP poprzez dodanie znormalizowanej formuły składającej się ze skrótu CIP oraz identyfikatora kraju albo nazwy jednostki odpowiedzialnej za tworzenie opisu, np. „CIP – Biblioteka Narodowa” dla polskich publikacji. Ponadto opis taki powinien zostać jak najszybciej zastąpiony prawidłowym opisem bibliograficznym wykonanym z autopsji w momencie opublikowania dzieła oraz powinien ukazać się w rejestrze każdej centrali CIP, jeśli dany wydawca bierze udział w więcej niż jednym programie. Rekomendacje nie mówiły jednak, jakie materiały tworzą podstawę do budowy opisu. Bazowano wówczas na odbitkach korektorskich, zwłaszcza stron tytułowych, redakcyjnych, kserokopiach projektów okładek i obwolut, makietach, odpisach danych z metryki drukarskiej itp. (Kończak, 2005, 67–68; Ramlau-Klekowska, 1990, 54–56).

Współcześnie opisy CIP tworzone są na podstawie odbitek korektorskich publikacji (kopie strony tytułowej i strony redakcyjnej, reprodukcje niektórych lub wszystkich stron stanowiących preliminaria – przedmowa, spis treści, obwoluta) oraz specjalnych formularzy¹⁰, które muszą wypełnić wydawcy (przeważnie zamieszczane są na stronach internetowych central sporządzających opisy) i w których podawane są dodatkowe informacje dotyczące publikacji, takie jak: planowana data wydania, opis fizyczny i rzeczowy publikacji, szczegóły dotyczące związku publikacji z innymi wydaniami tego samego dzieła (Kończak, 2005, 67). Zawartość formularzy zmienia się w czasie i dostosowywana jest do nowych technologicznych rozwiązań i zmian w przepisach katalogowania¹¹.

Na podstawie dostarczanych informacji centrala tworzy skrócony opis katalogowy drukowany później w książce. Warto jednak zauważyć, że w tym punkcie program CIP nie do końca zgadzał się z założeniami programu „Universal Bibliographic Control”, który mówił o tym, że narodowe centrale bibliograficzne opracowują pełne autorytatywne opisy bibliograficzne na podstawie egzemplarza obowiązkowego. Opisy wykonywane w trakcie procesu publikacyjnego nie mogły więc zastąpić opisu bibliograficznego tworzonego na potrzeby narodowej rejestracji bibliograficznej. Mimo to, dzięki zastosowaniu zautomatyzowanych systemów już od lat 70. XX w., opis CIP stanowi gotowy fragment opisu bibliograficznego, który albo jest poprawiany i uzupełniany (np. w Australii), albo wymieniany na nowy – pełny opis wraz z wpływem gotowego egzemplarza do centrali.

5. CIP a rozwój nowych technologii – ECIP

Ciągły rozwój nowych technologii w zakresie produkcji książki i form przekazywania informacji sprawia, że program CIP musi być na bieżąco dostosowywany do aktualnego

¹⁰ W 1985 r. sekcja katalogowania IFLA przedstawiła trzy wersje ujednoliconego formularza. Jednak ze względu na zbyt dużą szczegółowość informacji zrezygnowano z wprowadzenia go na skalę międzynarodową i określono, że o jego zawartości decydować ma centrala CIP (zlokalizowana najczęściej w ramach narodowej centrali bibliograficznej).

¹¹ W 2014 r. BK przeprowadziła ankietę wśród wydawców. Jej wyniki przyczyniły się do zmian wprowadzonych w październiku 2015 r. w formularzu CIP oraz dostosowania przesyłanych informacji do nowych przepisów katalogowania RDA (LC, 2015).

poziomu rozwoju technologicznego, tak by był nadal atrakcyjny dla środowisk bibliotekarskich, wydawniczych i księgarskich. W 1993 r. Biblioteka Kongresu przygotowała projekt pilotażowy elektronicznej wersji programu CIP o nazwie „Electronic CIP”. System wdrożono oficjalnie w 1999 r. Nowy program znacznie przyspieszył realizację procesu tworzenia opisów CIP, zmieniając formę przekazywania informacji od wydawcy do agencji bibliograficznej i z powrotem z mailowej na elektroniczną wymianę danych (*Electronic Data Interchange – EDI*). Ponadto wprowadzenie elektronicznej wersji programu i możliwości skopiowania i wklejenia informacji zmniejszyło liczbę pojawiających się wcześniej błędów literowych, jakie popełniano podczas ręcznego przepisywania zawartości opisu CIP (McCroskey & Turvey, 2004, 39).

W 2007 r. BK na dobre zrezygnowała z tradycyjnej (papierowej) formy nadsyłania przez wydawców materiałów do tworzenia opisów CIP. Wydawcy przesyłają obecnie niezbędne informacje poprzez specjalny formularz znajdujący się na stronie WWW Biblioteki Kongresu oraz pełny tekst dzieła drogą elektroniczną. BK tworzy opis CIP i przesyła go wydawcy pocztą elektroniczną. Wersję tradycyjną programu utrzymano jedynie dla książek wydanych w innym języku niż angielski, książek, w których na stronie tytułowej pojawiają się znaki diakrytyczne języków innych niż zachodnioeuropejskie oraz książek, na których zawartość składają się głównie obrazy, tabele, wykresy, formuły matematyczne lub chemiczne itp. (z wyjątkiem książek dla dzieci) (LC, 2006).

6. Alternatywa dla programu CIP

Ograniczenia, jakie narzuciła BK w stosunku do przyjmowania zgłoszeń do programu ECIP zaowocowały nowymi przedsięwzięciami w amerykańskim środowisku wydawców, bibliotekarzy i pracowników informacji. Także w środowisku europejskim w ciągu ostatniego półwiecza pojawiły się inicjatywy dające możliwość kształtowania bibliograficznej informacji prospektywnej.

6.1. EPCN (*Electronic Preassigned Control Number*)

Wydawcom, którzy nie są w stanie skorzystać z elektronicznej wersji programu ECIP, a także tym, którzy nie spełniają warunków uczestnictwa w programie CIP w Stanach Zjednoczonych¹², Biblioteka Kongresu zaproponowała alternatywną formę: program EPCN – Electronic Preassigned Control Number (elektronicznie nadawany numer kontrolny). W 2010 r. liczba wydawców biorących udział w tym programie (47 945) prawie dziesięciokrotnie przewyższała liczbę uczestników programu CIP (5008) (Debus-López, 2011).

PCN, czyli Preassign Control Number jest programem, który pozwala na przypisanie każdemu dziełu (jeszcze nieopublikowanemu przez wydawcę) specjalnego numeru kontrolnego BK (LCCN), który informuje użytkowników, że BK najprawdopodobniej nabędzie tę książkę. Do każdego wydawnictwa tworzony jest również opis bibliograficzny – w przeciwieństwie

¹² Do programu CIP w USA przystąpić mogą tylko wydawcy, którzy opublikowali minimum trzy dzieła różnych autorów. Ponadto do programu nie kwalifikują się publikacje wydawane nakładem autora oraz publikowane na życzenie – (LC, b.d.-a).

jednak do programu CIP opisy te nie są rozprowadzane w postaci elektronicznej ani drukowane w książkach. Ponadto wydawca nie może brać udziału w obu przedsięwzięciach jednocześnie, ponieważ programy wzajemnie się wykluczają. Warto zauważyć, że wydawca umieszczając w książce jedynie numer kontrolny BK, znacznie ułatwia procesy związane z katalogowaniem w bibliotekach i księgarniach. Numer LCCN jest bowiem wykorzystywany przez biblioteki amerykańskie do wyszukiwania rekordów bibliograficznych w bazach danych i zamawiania kart katalogowych w BK lub u innych komercyjnych dostawców (LC, b.d.-b).

6.2. P-CIP (*Publishers Cataloging in Publication*)

Ponieważ w Stanach Zjednoczonych program CIP skierowany jest wyłącznie do wydawców, którzy rozprowadzają swoje publikacje na szeroką skalę, powstał projekt utworzenia podobnego programu dla publikacji ukazujących się w niskich nakładach.

W odróżnieniu od opisu CIP, który przygotowujący jest zawsze przez Bibliotekę Kongresu, opisy P-CIP sporządzają sami wydawcy, ich agenci (m.in. Quality Books Inc., Five Rainbows – Services for Authors & Publishers) lub wyspecjalizowane firmy (np. Cassidy Cataloguing Services Inc., The Donohue Group, Inc.). Za wykonanie takiego opisu firmy te pobierają jednak opłatę. Ponadto do stworzenia opisu wymagany jest zawsze numer uzyskany w ramach PCN. Każdy opis P-CIP przekazywany jest do katalogu amerykańskiego OCLC WorldCat, z którego bibliotekarze mogą pobrać gotowy opis w formacie MARC.

6.3. BIBLINK

W połowie lat 90. XX w. z inicjatywy Unii Europejskiej powołano program BIBLINK, którego przedmiotem były publikacje elektroniczne, a głównym zadaniem miało być zapewnienie przepływu informacji pomiędzy narodowymi agencjami bibliograficznymi a wydawcami i użytkownikami. Wydawcy mieli za zadanie wprowadzić do systemu podstawowe dane, które następnie przesyłano do agencji bibliograficznych, gdzie były uzupełniane o dodatkowe informacje (m.in. hasła rzeczowe). Zmodyfikowane informacje miały wracać do wydawców, którzy mogli kodować je w wydawanych przez siebie publikacjach elektronicznych. Projekt zakładał wykorzystanie różnego rodzaju składni, m.in. HTML, SGML, UNIMARC i formatów narodowych MARC, tak by możliwa była wymiana informacji między wszystkimi uczestnikami programu. Po przeprowadzeniu wersji pilotażowej projektu, zamknięto go w lutym 2000 r. (Patel & Clayphan, b.d.; Smith, 2001, 59).

6.4. Neuerscheinungsdienst

Biblioteka Niemiecka wprowadziła w 2003 r. darmowy serwis „Informacja o nowościach” („Neuerscheinungsdienst”), który zastąpił zamknięty w 2002 r. program CIP (Rinn, 2004, 47–51). Nowy serwis nie stanowi już członu bibliografii narodowej, ale opisy w nim zamieszczane uzupełniane są o symbole KDD stosowane również w bibliografii bieżącej (DNB, 2016). Dane o książkach pochodzą z katalogu „Verzeichnis Lieferbarer Bücher” (VLB) (niemiecki odpowiednik „Books in Print”), do którego wydawcy zgłaszają nowe tytuły za pośrednictwem Internetu, z wykorzystaniem formatu ONIX lub w wersji papierowej. Dane

te są następnie w ciągu dwóch dni przekazywane do Niemieckiej Biblioteki Narodowej, która porządkuje je według treści zgodnie z KDD i publikuje je w kilku wersjach: jako tygodnik online w formacie PDF, za pomocą interfejsów OAI (Open Archives Initiative) i SRU (Search / Retrieve via URL), a także jako dane surowe (DNB, 2017).

7. Perspektywy rozwoju programu CIP

Już w międzynarodowych rekomendacjach z 1986 r. poruszono kwestię wątpliwej przyszłości programu, zwracając uwagę, że

stałemu wzrostowi liczby publikowanych tytułów nie będzie zapewne towarzyszył proporcjonalny wzrost możliwości ich zakupu i gromadzenia w bibliotekach (...) zapotrzebowanie na wielokrotnie opracowywanie tego samego tytułu w wielu bibliotekach może się zmniejszać, natomiast popyt na terminową i sprawną informację będzie rósł. Stąd przewidywane dalsze przesuwanie się punktu ciężkości ze znaczenia opisu umieszczonego w książce na opis wchodzący szybko do obiegu informacji, zwłaszcza wspomaganą techniką informatyczną (Ramlau-Klekowska, 1990, 55).

Przewidywania te potwierdziły się w przypadku niektórych krajów (np. Niemiec).

Współcześnie wiele ośrodków tworzących opisy CIP wypracowało własne rozwiązania dotyczące stopnia kompletności opisów, formy ich wykorzystywania w bibliografiach narodowych, a także sposoby kontrolowania za pomocą programu dopływu egzemplarza obowiązkowego od współpracujących wydawców. W rozwiązaniu tego ostatniego problemu w Stanach Zjednoczonych pomocą ma nowy proces, zapoczątkowany w BK w lutym 2010 r. – zautomatyzowany proces dopominania się przez e-mail o książki, jakie wydawcy są zobowiązani przekazać BK. Wydawca będzie otrzymywał informację o zaległościach raz w miesiącu, a czas na odesłanie egzemplarza do BK będzie wynosił 90 dni. Nieprzesłanie egzemplarza z wyznaczonym terminie ma skutkować wykluczeniem wydawcy z programu (LC, b.d.-a).

W kontekście rozwoju nowych technologii i tworzenia nowego typu zasobów – e-booków, publikacji hybrydowych (posiadających komponenty tradycyjne i cyfrowe) – pojawiają się też pytania o kierunek dalszego rozwoju programu CIP oraz o sposób usprawnienia i ułatwienia przepływu danych między wydawcami a BK i innymi partnerami. Biblioteka Kongresu w raporcie opracowanym w 2009 r. przedstawiła zalecenie dotyczące wyeliminowania dublowania opisów m. in. poprzez usprawnienie wymiany danych i pełną automatyzację programu CIP (LC, 2008, 15). W celu realizacji tego zamierzenia BK w czerwcu 2009 r. rozpoczęła pilotażowy projekt konwersji danych dostarczanych przez wydawców w formacie ONIX do programu ECIP. Dane konwertowane są do formatu MARC z zachowaniem zawartości takich elementów jak „spis treści” i „streszczenie”, występujących w formacie ONIX (LC, b.d.-a).

8. Zakończenie

Na podstawie danych wskazanych w artykule można stwierdzić, że informacja prospektywna publikowana jest tylko w niektórych krajach (USA, Wielka Brytania, Niemcy). W dobie nowoczesnych technologii, kiedy opisy publikacji trafiają do bibliografii narodowych dostępnych online w ciągu kilku dni od daty wydania, informacja ta wydaje się być zbędna

dla bibliotekarzy jako źródło autorytatywnego opisu. Pozostaje jednak istotnym źródłem wpływającym na kształtowanie polityki gromadzenia zbiorów (planowania zakupów).

Opisy CIP zamieszczane na odwrocie strony tytułowej, mogą być ponadto wykorzystywane przez czytelników do spisywania podstawowych, a co najważniejsze – autorytatywnych danych o publikacji, którą właśnie przeglądają. Informacje te mogą wykorzystać m.in. przy tworzeniu przypisów czy bibliografii załącznikowych, katalogowaniu własnego księgozbioru, korekcie informacji zaczerpniętych z różnych źródeł m.in. do menedżerów bibliografii.

Nie można też nie zauważyć problemu egzekwowania EO w wielu krajach na świecie. Zapewne stąd zrodził się pomysł na wykorzystanie programu do monitorowania o przesyłanie egzemplarza do BK¹³.

Nie ulega jednak wątpliwości, że przedsięwzięcia typu CIP miały największą szansę rozwoju w krajach, w których wysoko ceni się informację bibliograficzną, zwłaszcza w środowisku wydawców. Dodatkowym atutem jest zapewne ukazywanie się informacji prospektywnej. W Polsce co prawda funkcjonuje program CIP, ale opisy tworzone w jego ramach nie są udostępniane publicznie.

Śledząc historię rozwoju programu, zauważyć można, że ogromne znaczenie miała dla niego rewolucja technologiczna, jaka rozpoczęła się w połowie ubiegłego wieku. Pojawienie się elektronicznych formatów zapisu danych, nowoczesnych nośników informacji oraz zautomatyzowanie procesów w ośrodkach informacyjnych, bibliotekarskich, a także wydawniczych i księgarskich, sprzyjało nawiązywaniu współpracy między tymi środowiskami m.in. w zakresie realizacji programu CIP.

Od momentu powstania opisywanych inicjatyw nastąpiły też zmiany w obrębie funkcji opisu CIP – od ułatwiającego prace katalogowe w bibliotekach, poprzez podkreślenie funkcji informacji prospektywnej w środowisku bibliotek, księgarzy i użytkowników, po ostatni kierunek – testowany przez Bibliotekę Kongresu – wykorzystania programu do kontroli wpływu egzemplarza obowiązkowego. Zmiany te sprawiają, że obawy, jakie pojawiają się od lat 80. XX w., związane z funkcjonowaniem programu w kontekście nowych technologii (m.in. powstaniem formatu ONIX) oraz dużej aktualności bieżących bibliografii narodowych, a tym samym z zanikiem zainteresowania tego typu informacją, okazują się być, przynajmniej w niektórych krajach, na razie niepotrzebne.

Bibliografia

- Anderson, D. (1986). *Guidelines for Cataloguing-in-Publication*. Paris: IFLA.
- British Library (2014). *The Cataloguing-in-Publication Programme* [online]. The British Library, [29.05.2017], <http://www.bl.uk/bibliographic/natbib.html>
- Clement, H. E. A. (1990). National bibliographic agencies cataloguing survey. *International Cataloguing and Bibliographic Control*, 19(1), 6–10.
- Cybulski, R. (1978). Międzynarodowy Kongres nt. Bibliografii Narodowych. Paryż, 12–15 września 1977 r. *Przegląd Biblioteczny*, 46(2), 223–230.
- Daszkowski, Z. (1975). Katalogowanie książek w toku procesu publikacyjnego (na podstawie doświadczeń radzieckich i amerykańskich). *Przegląd Biblioteczny*, 43(2), 149–161.

¹³ Należy też pamiętać, że problem egzekwowania EO wiąże się również z liczbą egzemplarzy, jaką należy przesłać uprawnionym placówkom (w Polsce 17 egzemplarzy jeśli nakład przekracza 100 egzemplarzy, 2 jeśli nakład jest mniejszy niż 100 egzemplarzy), co dla wielu wydawców stanowi spory wydatek.

- Debus-López, K. (2011). CIP Advisory Group Meeting. In: *ALA 2011 Summer Conference* [online]. New Orleans, [13.02.2017], <http://www.loc.gov/publish/cip/cag>
- DNB (2016). *Neuerscheinungsdienst* [online]. Deutsche Nationalbibliothek, [13.02.2017], http://www.dnb.de/DE/Service/DigitaleDienste/ND/nd_node.html
- DNB (2017). *Neuerscheinungsdienst – Meldeverfahren* [online]. Deutsche Nationalbibliothek, [04.09.2017], http://www.dnb.de/DE/Erwerbung/Meldeverfahren/Meldeverfahren_node.html
- Hammer, N. (1999). CIP-Eindruck in Verlagspublikationen. *Dialog mit Bibliotheken*, 11(3), 11–13.
- IFLA (1979). IFLA International Office for UBC, *Guidelines for the national bibliographic agency and the national bibliography*. Paris: UNESCO.
- IFLA (1986). *Recommended standards for Cataloguing-in-Publication: the CIP data sheet and the CIP record in the book*. London: IFLA International Programme for UBC.
- IFLA WG (2015). *Best Practice for National Bibliographic Agencies in a Digital Age* [online]. IFLA Working Group on Guidelines for National Bibliographies [24.04.2017], <https://www.ifla.org/node/7858>
- Kołacz, T. (2005). Z historii programu CIP – Cataloguing in Publication. *Zagadnienia Informacji Naukowej*, 1(85), 64–80.
- LAC (2013). *Scope and coverage* [online]. Library and Archives Canada, [29.05.2017], <http://www.bac-lac.gc.ca/eng/services/cip/Pages/scope-coverage.aspx>
- LAC (2015). *Frequently asked questions* [online]. Library and Archives Canada, [25.05.2017], <http://www.bac-lac.gc.ca/eng/services/cip/Pages/faq.aspx>
- LAC (b.d.). *List of CIP Agents for Canadian Publishers* [online]. Library and Archives Canada [22.02.2017], http://www.collectionscanada.gc.ca/cip/041003-4000-e.html#national_office
- LC (2005). *Survey of CIP Programs Report* [online]. Library of Congress, [24.04.2017], www.loc.gov/catdir/cipsurvey/IFLA_CIP_Survey_Report.pdf
- LC (2006). *ECIP Replaces Conventional CIP Program* [online]. Library of Congress, [24.04.2017], <http://cip.loc.gov/endofpaperCIP.html>
- LC (2008). *On the record: report of The Library of Congress Working Group on the Future of Bibliographic Control* [online]. Library of Congress. [13.02.2017], <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>
- LC (2015). *CIP Data: Cataloging in Publication Program – Publishers* [online]. Library of Congress, [25.05.2017], <https://www.loc.gov/publish/cip/techinfo/>
- LC (b.d.-a). *Cataloging in Publication Program* [online]. Library of Congress, [24.04.2017], <http://www.loc.gov/publish/cip/>
- LC (b.d.-b). *Preassigned Control Number Program* [online]. Library of Congress, [24.04.2017], <https://www.loc.gov/publish/pcn/>
- McCroskey, M., & Turvey, M. (2004). Processing new CIP materials at the Library of Congress. *Knowledge Quest*, 32(3), 38–39.
- Mendenhall, K. (1982). *A survey of the cataloging in publication program. Final report*. Washington, D.C.: Library of Congress.
- NLA (b.d.). *CiP user guide* [online]. National Library of Australia, [07.03.2012], <http://www.nla.gov.au/cip/user-guide>
- NLC (1998). *Comparative study of the role of national bibliographic agencies in achieving bibliographic control of their national publishing output. Final report on stage two prepared for CDNL*. National Library of Canada.
- Newlen, R. R. (1991). Read the fine print. The power of CIP. *Library Journal*, 116(12), 38–42.
- O.K. (2017). *O.K. Ohlášené knihy* [online]. Praha: Národní knihovna České republiky. [29.05.2017], <http://aleph.nkp.cz/web/ok.htm>
- Patel, M., & Clayphan R. (b.d.). *Project BIBLINK: Linking Publishers and National Bibliographic Agencies* [online], [24.04.2017], <http://www.ukoln.ac.uk/metadata/publications/biblink/proj-biblink.html>

- Ramlau-Klekowska, K. (1990). Program „Cataloguing in Publication” w Polsce i na świecie. *Katalogowanie książki w trakcie procesu wydawniczego*. *Editor*, (3), 47–61.
- Rinn, R. (2004). Der Neuerscheinungsdienst der Deutschen Bibliothek. *Dialog mit Bibliotheken*, 16(2), 47–51.
- Smith, R. (2001). Distribution and exchange. In: *ICNBS Copenhagen 25–27 November 1998. Proceedings of International Conference on National Bibliographic Services*. Copenhagen: The Royal Library, 53–60.
- Swindley, L. R. (1975). *Cataloguing in Publication: An International Survey*. Paris: UNESCO.
- Wheatley, H.B. (1889). *How to Catalogue a Library*. 2nd ed. London: Elliot Stock.
- Zarządzenie nr 9 Ministra Kultury i Sztuki z dnia 16 maja 1988 r. w sprawie drukowania opisu katalogowego oraz adnotacji w wydawnictwach zwartych. (1988). *Dziennik Urzędowy Ministerstwa Kultury i Sztuki*, 4 (poz. 15).

The History and Today of Cataloging in Publication (CIP) Program

Abstract

Purpose/Thesis: The purpose of this article is to summarize the 50-year activity of the cataloging in publication (CIP) program and to define the role of the program in the development of bibliographic information in the second decade of the 21st century.

Approach/Methods: The literature analysis method was used. The analysis covered articles on the subject being discussed and the web sites of institutions responsible for the implementation of the program in Europe, the United States, Canada, and Australia.

Results and conclusions: The analysis of the program development in the surveyed countries shows that in most cases it has not been fully developed. In some countries it has been replaced by another form of prospective information (e.g. in Germany). The program is most advanced in the United States as a tool to enforce a copy for the Library of Congress.

Originality/Value: The literature shows some publications on CIP issues. However, there are not any reports compiling the last 20 years of the program in the context of the development of new technologies. This article is intended to fill a gap on the topic.

Keywords

Bibliographic description. Cataloging in Publication. CIP description

Dr DOROTA SIWECKA jest absolwentką Instytutu Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Wrocławskiego. W 2013 r. uzyskała stopień naukowy doktora w zakresie bibliologii i informatologii oraz kwalifikacje zawodowe bibliotekarza dyplomowanego. Obecnie pracuje jako wykładowca w IINiB UW. Jej najważniejsze publikacje to m.in. Prace IFLA nad międzynarodowymi przepisami katalogowania. Przegląd Biblioteczny 2012, 319–336 oraz uhonorowana Nagrodą im. A. Łysakowskiego SBP monografia Światowy model informacji bibliograficznej. Programy i projekty (1950–2010). Warszawa: Wydaw. SBB, 2015.

Kontakt z autorką:

dorota.siwecka@uwr.edu.pl

*Instytut Informacji Naukowej i Bibliotekoznawstwa
Uniwersytet Wrocławski*

pl. Uniwersytecki 9/13

50–137 Wrocław

Jakość wybranych elementów metadanych stosowanych w polskich bibliotekach i repozytoriach cyfrowych¹

Piotr Malak²

*Institut Informacji Naukowej i Bibliotekoznawstwa
Uniwersytet Wrocławski*

Veslava Osińska

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

Bożena Bednarek-Michalska

*Biblioteka Uniwersytecka
Uniwersytet Mikołaja Kopernika w Toruniu*

Abstrakt

Cel/Teza: Niniejszy artykuł prezentuje wyniki badań nad oceną jakości metadanych w polskich bibliotekach cyfrowych oraz możliwości wykorzystania tych danych do masowego przetwarzania zawartości cyfrowych repozytoriów oraz wyodrębnienia współczesnych prac naukowych.

Koncepcja/Metody badań: Badania zostały przeprowadzone hybrydowo – poprzez automatyczną analizę metodami inżynierii lingwistycznej oraz analizę ekspercką.

Wyniki i wnioski: W wyniku przeprowadzonych badań zidentyfikowano szereg nieścisłości i nieprawidłowości występujących w metadanych dla dokumentów w polskich bibliotekach cyfrowych. Wskazano również propozycje poprawy tego stanu.

Zastosowanie praktyczne: Wnioski płynące z badań mogą przyczynić się do znaczącej poprawy jakości metadanych i umożliwienia wykorzystania metadanych jako wartościowego materiału badawczego.

Oryginalność/Wartość poznawcza: Według najlepszej wiedzy autorów badania tego typu, zarówno w zakresie tematyki jaki i skali, nie były dotychczas prowadzone w Polsce.

Słowa kluczowe:

Analiza metadanych. Biblioteki cyfrowe. Jakość metadanych. Dublin Core.

Otrzymano: 9 czerwca 2016. Zrecenzowany: 22 lipca 2016 oraz 5 września 2016.

Poprawiony: 17 sierpnia 2016 oraz 19 stycznia 2017. Zaakceptowany: 10 marca 2017.

1. Wprowadzenie

W ramach projektu „Badanie struktury i dynamiki cyfrowych zasobów wiedzy za pomocą metod wizualizacji” (ang. *Information Visualization methods in digital knowledge structure*

¹ Badania przeprowadzono w ramach grantu NCN 2013/11/B/HS2/03048. Badanie przeprowadzono przy współpracy z realizatorami projektu *Polska część infrastruktury naukowej CLARIN ERIC CLARIN PL*.

² Kontakt z autorami artykułu za pośrednictwem dr. Piotra Malaka: piotr.malak@uwr.edu.pl

and dynamics study), finansowanego w latach 2014–2017 przez Narodowe Centrum Nauki, a realizowanego na Uniwersytecie Mikołaja Kopernika (UMK), prowadzony jest szereg badań nad dokumentami cyfrowymi i ich opisami bibliograficznymi jako zasobami wiedzy³. Jednym z ważnych tematów badawczych podjętych w ramach tego grantu jest **analiza i ocena jakości metadanych** w polskich bibliotekach cyfrowych. Badanie to prowadzono w celu sprawdzenia możliwości wykorzystania metadanych do masowego przetwarzania zawartości cyfrowych repozytoriów i wyodrębnienia na ich podstawie współczesnych prac naukowych. W wyniku przeprowadzonych badań i analiz dokonano oceny jakości metadanych w polskich bibliotekach cyfrowych, wskazano grupy błędów występujących w opisach oraz zgłoszono postulaty działań zmierzających do poprawy jakości metadanych. Celem niniejszego artykułu jest zaprezentowanie wyników badań nad jakością metadanych w polskich bibliotekach cyfrowych, wniosków płynących z analizy oraz propozycji poprawy bieżącego stanu.

Badaniami opisanymi w niniejszym artykule objęto wszystkie biblioteki cyfrowe w Polsce, w których do opisu dokumentów stosuje się schemat metadanych Dublin Core⁴. Na potrzeby omawianych badań przyjęto następujące założenia:

- (1) analizie poddane są metadane używane w opisach publikacji dostępnych w polskich bibliotekach cyfrowych;
- (2) prace opublikowane przed rokiem 1945 nie spełniają warunku współczesności;
- (3) obiekty oznaczone jako należące do domeny publicznej nie spełniają wymogu współczesności. Założenie to wynika z faktu, że do domeny publicznej trafiają w większości prace starsze, niepodlegające już ochronie prawnej, 70 lat od śmierci autora lub dacie publikacji (w przypadku czasopism).

Spełnienie powyższych założeń związane jest z celami grantu badawczego, jest również wystarczające dla wstępnej selekcji i kwalifikacji obiektów do analizy jakości metadanych na potrzeby niniejszego artykułu.

W omówionych w artykule badaniach jakości metadanych przyjęto następujące hipotezy:

- (1) Duża część metadanych stosowanych w polskich bibliotekach cyfrowych nie spełnia kryteriów jakości określonych na potrzeby omawianych badań.
- (2) Analiza wartości elementów opisu Dublin Core: dc:rights, dc:type, dc:date umożliwi wyselekcjonowanie współczesnych zasobów naukowych z całości zasobów polskich bibliotek cyfrowych, w których stosowany jest ten standard.
- (3) Opisy wydawnictw ciągłych, oznaczonych jako czasopisma, gazety, itp. nie zawierają informacji wystarczających do oceny czy wydawnictwo to ma charakter naukowy, ani tym bardziej do określenia tematyki badawczej poruszanej w publikowanych artykułach. W metadanych opisujących czasopisma i gazety brak jest odniesienia do ich zawartości (zob. Tab. 1). Są to zazwyczaj opisy wydawnictwa zbiorowego, czasem uzupełnione o spisy treści występujące jako samodzielne dokumenty.

Zgodnie z hipotezą 2. do analiz wybrano zawartość następujących elementów metadanych standardu *Dublin Core*:

³ O badaniach związanych z grantem zob. m.in. V. Osińska, P. Malak (2016a; 2016b); V. Osińska, P. Malak, B. Bednarek-Michalska (2016).

⁴ Opisany m.in. w PN-ISO 15836:2006 Informacja i dokumentacja – Zestaw elementów metadanych Dublin Core.

- *dc:date* – pole zawierające najczęściej datę, na dowolnym poziomie szczegółowości, powstania lub udostępnienia obiektu, norma *The Dublin Core Metadata Element Set* (DCMES) zaleca stosowanie zapisu dat w postaci zgodnej z normą ISO 8601⁵, czyli w postaci: RRRR-MM-DD. Jak podaje M. Nahotko (2000), możliwe jest użycie innej formy zapisu dat, ale powinna ona zostać jednoznacznie zidentyfikowana. Dobrym przykładem takiego podejścia do notowania dat jest opracowanie przygotowane przez Bibliotekę Uniwersytecką we Wrocławiu (BUWr) (*Interpretacja schematu Dublin Core*, 2006).
- *dc:type* – pole opisujące rodzaj dokumentu pod względem kategorii lub gatunku. Standard *DCMI Metadata Terms*⁶ zaleca korzystanie ze słownictwa kontrolowanego, zalecanego w dokumencie *DCMI Type Vocabulary*. Są to następujące terminy: Collection, Dataset, Event, Image, Interactive Resource, Service, Software, Sound, Text⁷. Również opracowanie BUWr, w oparciu o Polską normę PN-92 N-01227, podaje słownik kontrolowany typów zasobów. Na dziewięć wyróżnionych w nim kategorii typów formy dokumentu podane są aż trzy synonimy.
- *dc:rights* – informacja o prawach własności intelektualnej, prawach autorskich lub prawach własności.

Tab. 1. Przykładowe metadane czasopisma dostępnego w bibliotece cyfrowej [data dostępu 20 kwietnia 2016]

Przegląd Biblioteczny 2000, z. 4.
 Dublin Core wer.1.1 :
 Tytuł : Przegląd Biblioteczny 2000, z. 4.
 Wydawca : H. Dobrzycki ; Wydawnictwo SBP ; Zakład Narodowy im. Ossolińskich
 Miejsce wydania : Warszawa
 Współtwórca : Sordyłowa, Barbara. Red.
 Instytucja sprawcza : Polska Akademia Nauk. Biblioteka w Warszawie ; Stowarzyszenie Bibliotekarzy Polskich
 Data wydania : 2000
 Typ zasobu : czasopismo
 Format : image/x.djvu
 Identyfikator zasobu : oai:bbc.uw.edu.pl:112
 Język : pol

Źródło: <http://bbc.uw.edu.pl/publication/148>

1.1. Przyjęte kryteria jakości metadanych

Metadane w systemach informacyjno-wyszukiwawczych powinny z założenia spełniać kryteria jakości, czy to wynikające z zastosowanego schematu, czy też z wewnętrznych regulacji instytucji prowadzącej dany system. Do podstawowych funkcji metadanych opisowych należy dostarczenie informacji jednoznacznie identyfikujących konkretny obiekt w zbiorze. Aby

⁵ Date and Time Formats – W3CDT, <https://www.w3.org/TR/NOTE-datetime>

⁶ *DCMI Metadata Terms*, <http://dublincore.org/documents/dcmi-terms/>

⁷ Za *DCMI Type Vocabulary*, <http://dublincore.org/documents/2000/07/11/dcmi-type-vocabulary/>

ta funkcja była realizowana właściwie, w przypadku dokumentów tekstowych w bibliotekach cyfrowych zazwyczaj wystarczy prawidłowo podać dane identyfikacyjne dokumentu, jak *tytuł, autor, identyfikator w systemie*. Metadane schematu Dublin Core dostarczają również informacji przybliżających treść opisywanego dokumentu. W tym przypadku standardy jakości wyznaczone są przez przyjęty system opisu. Na potrzeby omawianych badań przyjęto kryteria oceny jakości metadanych wynikające z założeń badawczych całego projektu grantowego. Podstawowym kryterium było więc dostarczenie informacji, które pozwolą efektywnie wyszukać w dostępnym zbiorze publikacji dokumenty o zadanych parametrach. Parametrami tymi były: współczesność dokumentu (*dc:date*), określona na mniej niż 70 lat od daty powstania utworu (*dc:rights*), naukowy charakter treści dokumentu (*dc:type*) oraz tematyka związana z zakresem nauk humanistycznych. Pole *dc:rights* pozwoliło wyeliminować wszystkie obiekty z zapisem *domena publiczna* jako stare, niewspółczesne – zawartość tego pola może być traktowana komplementarnie do zawartości pola *dc:date*, szczególnie w sytuacjach, gdy data publikacji nie została podana. Ze względu na przyjęte założenia, istotnymi elementami metadanych są zapisy w polach *dc:type*, *dc:rights*, *dc:date*. Proces automatycznego wyszukiwania i klasyfikowania informacji na podstawie zawartości tych pól wymaga, aby dane w nich zawarte pozwalały efektywnie wyodrębnić dokumenty z poszukiwanej kategorii. Aby spełnić te wymagania opisy powinny spełniać następujące kryteria:

- spełniać zalecenia norm;
- jednoznacznie opisywać dokument;
- zachować konsekwencję w stosowanych wyrażeniach.

Zachowanie wymienionych kryteriów, w odniesieniu do każdego pola opisu w metadanych, niezależnie od przyjętego systemu, leży w interesie każdej biblioteki cyfrowej, ponieważ m.in. ułatwia wymianę danych między systemami, a także efektywne wyszukiwanie informacji o dokumentach we własnym systemie. Należy tutaj podkreślić, że we współczesnych bibliotekach cyfrowych zazwyczaj obiektami przeszukiwanymi nie są pełne teksty dokumentów, lecz właśnie metadane opisujące te dokumenty (Baca, 2016). Metadane są również uniwersalnym nośnikiem informacji w procesach wymiany danych pomiędzy różnymi bibliotekami, np. Federacja Bibliotek Cyfrowych (FBC)⁸ prezentuje wyłącznie metadane pozyskiwane z polskich bibliotek cyfrowych. Stąd dbałość o jakość i poprawność danych wprowadzanych do poszczególnych pól opisu powinna być istotnym elementem polityki udostępniania zasobów.

1.2. Przegląd piśmiennictwa

Problematyka jakości metadanych wpisuje się w ogólny nurt rozważań nad jakością bibliotek cyfrowych, nabierając szczególnego znaczenia w kontekście szerokiego udostępniania i promowania danych z zakresu dziedzictwa kulturowego, w tym m. in. poprzez propagację w Europie. W literaturze światowej i polskiej problem metadanych oraz ich jakości pojawia się regularnie. Tym bardziej dziwi fakt, że wiele z podnoszonych przy tej okazji postulatów albo nie budzi zainteresowania twórców bibliotek cyfrowych, albo z innych powodów nie jest uwzględniane w ich doskonaleniu. Poniżej na podstawie wybranych prac krótko scharakteryzowane zostały przykłady problemów podejmowanych w polskich publikacjach

⁸ <http://fbc.pionier.net.pl/>

poświęconych temu tematowi. Ponieważ problem dotyczy polskich bibliotek cyfrowych, przegląd piśmiennictwa został ograniczony do prac polskich badaczy i praktyków.

A. Domagalska (2006) przeanalizowała dostępne w latach 2002–2006 trzy rozwiązania standaryzacyjne z zakresu tworzenia bibliotek cyfrowych (NISO 2004, Strategia Europejska I2010, Strategia Zespołu ds. Standardów dla Bibliotek Naukowych) konstatując, że w żadnym z nich nie poświęcono wcale uwagi kwestiom oceny jakości bibliotek cyfrowych, a tym samym pominięto również kwestie jakości metadanych. W pracy podniesiony został problem wagi metadanych w procesie wyszukiwania treści cyfrowych (w tym zdigitalizowanych) oraz wpływu jakości tych metadanych na efektywność wyszukiwawczą danego systemu, co z kolei przekłada się na poziom satysfakcji użytkowników. Podano tam również trzy typy metadanych: opisowe, administracyjne oraz strukturalne. Autorka pracy podniosła również kwestię konieczności stworzenia, przynajmniej na poziomie instytucji, systemu autorytatywnej kontroli metadanych oraz zwróciła uwagę na fakt, że za błędne metadane odpowiada człowiek – podczas ich tworzenia, lub maszyna – w przypadku automatycznego generowania na podstawie dostarczonych danych.

W pracy zbiorowej pod redakcją G. Płoszajskiego (2008) również rozważano temat kategorii metadanych. Wskazano podstawy prawne regulujące ich definicje oraz różnice pomiędzy metadanymi opisującymi cyfrowe kopie obiektów fizycznych a tymi, które opisują dokumenty utworzone oryginalnie jako cyfrowe. W pracy tej podniesiono również problem metadanych dla archiwów, bibliotek i muzeów oraz rozrózniono „obiekt oryginalny i odpowiadający mu obiekt cyfrowy”, a także wskazano na możliwe różnice w zawartości metadanych opisujących oba typy obiektów. W publikacji tej Europeana jest wyróżniona ze względu na stosowanie metadanych opisowych jako metadanych wyszukiwawczych. Stawia się tam również postulat przyjęcia konkretnej listy typów dokumentów. Największy jednak nacisk położony został na omówienie metadanych technicznych i administracyjnych. Publikacja stanowi zbiór opisów dobrych praktyk związanych z tworzeniem i udostępnianiem kolekcji cyfrowych, popartych przykładami istniejących wzorcowych standardów oraz wdrożeń. Nie dokonuje jednakże oceny bieżącego stanu metadanych w polskich bibliotekach cyfrowych.

Praktycznym problemom wykorzystania metadanych z polskich bibliotek cyfrowych poświęcona została praca M. Werli (2010). Jej autor zaprezentował doświadczenia Poznańskiego Centrum Superkomputerowo-Sieciowego (PCSS) – twórcy FBC, w wykorzystaniu metadanych do tworzenia graficznych interfejsów eksploracji zasobów bibliotek cyfrowych. Autor opisał proces analizy metadanych pod kątem określenia daty oraz miejsca wydania publikacji, czyli w zakresie częściowo pokrywającym się z tematyką badań opisywanych w niniejszym artykule. M. Werla opowiada się za postulatem stosowania słowników lub kartotek haseł wzorcowych przy tworzeniu metadanych w celu osiągnięcia odpowiedniego poziomu ich interoperacyjności. Twierdził, że w przypadku danych typu data, gdzie trudno wykorzystać słowniki zamknięte, należy wdrożyć normalizację ich zapisu, jako przykład podając notację RRRR-MM-DD. Autor postulował również takie zaprojektowanie schematu metadanych, aby możliwe było automatyczne wyodrębnienie określeń przestrzennych czy czasowych. Na podstawie analiz danych przeprowadzonych na potrzeby niniejszego artykułu, można niestety stwierdzić, że stan tego typu metadanych w ciągu sześciu lat od ukazania się pracy Werli nie uległ zmianie. FBC wprowadziła system automatycznego poprawiania i uzupełniania metadanych przesyłanych przez biblioteki cyfrowe, m.in. w zakresie ujednolicania zapisów dat, ale system taki nie zastąpi rzetelnego wprowadzania danych przez biblioteki.

W odczuciu autorów niniejszego artykułu smutną wymowę tego braku poprawy wzmocnia fakt, że polskie biblioteki cyfrowe udostępniają swoje dane za pośrednictwem FBC, gdzie M. Werla pełni funkcję Kierownika Działu Bibliotek Cyfrowych i Platform Wiedzy, a mimo to nie wdrażają wysuwanych przez niego propozycji poprawy jakości metadanych.

Z kolei artykuł M. Nahotko (2010) ma charakter przeglądowy, opisuje badania krajowe oraz międzynarodowe w zakresie automatycznego generowania metadanych oraz przybliża aplikacje stosowane do pracy z metadanymi do 2010 r. Autor duży nacisk położył na wskazanie możliwości wykorzystania nowoczesnych metod indeksowania treści do generowania metadanych opisujących dokumenty.

Warto też wspomnieć o serwisie *Digitalizacja.pl*, poświęconym problemom digitalizacji zbiorów bibliotecznych, muzealnych i archiwalnych. Znajdują się tam wartościowe poradniki oraz dyskusje, w tym dotyczące metadanych i ich jakości. Należy również odwołać się do powstałego w 2012 r. raportu pt. *Metadane, zagadnienia słowników kontrolowanych* (2012), który porusza problematykę rozwiązań dotyczących digitalizacji obiektów muzealnych i ich udostępniania, omawiając również kwestie metadanych oraz dostarczając propozycje rozwiązań w tym zakresie.

Temat metadanych dyskutowany jest również w opracowaniach poświęconych generalnej ocenie jakości bibliotek cyfrowych. Analizy te jednak dotyczą zazwyczaj bibliotek cyfrowych jako serwisów internetowych, ich budowy, dostępności oraz funkcjonowania (Kazan, Skubała, 2008; Potęga, 2009; Głowacka, 2011; Derfert-Wolf, 2011; Żernicka, 2014). Metadane są tam wymieniane jako jeden ze składników oceny jakości biblioteki cyfrowej jako całości, bez poświęcania uwagi jakości samych metadanych. J. Potęga (2009) podaje dodatkowo wyniki analiz ilościowo jakościowych wykorzystania w opisach bibliograficznych pól schematu Dublin Core oraz zawartości tych pól. Z wyjątkiem pracy K. Żernickiej (2014), pozostałe tu wymienione opisują stan do 2011 r.

2. Metodyka badań

Przedmiotem badań były metadane obiektów udostępnianych w bibliotekach cyfrowych bez względu na oryginalną formę wydawniczą. Uwzględniono dokumenty upowszechniane od 2007 r., od kiedy funkcjonuje Federacja Bibliotek Cyfrowych, która pełni funkcje agregatora metadanych z polskich bibliotek cyfrowych i repozytoriów naukowych. Warto tutaj nadmienić, że w ramach projektu „Badanie struktury i dynamiki cyfrowych zasobów wiedzy za pomocą metod wizualizacji” badania prowadzone są nie tylko na danych udostępnianych za pośrednictwem FBC, ale także na zasobach udostępnianych przez inne instytucje – Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego (ICM UW), które prowadzi agregator polskich repozytoriów naukowych i Repozytorium Centrum Otwartej Nauki (CEON)⁹ oraz przez Ośrodek Przetwarzania Informacji (OPI)¹⁰, zbierający dane osobowe o naukowcach polskich¹¹.

⁹ <http://agregator.ceon.pl>

¹⁰ <http://www.opi.org.pl>

¹¹ Ze względu na odmienny charakter danych z OPI, nie przeprowadzono ich analizy jakościowej. Dane te nie zostaną również omówione w niniejszym artykule, ponieważ nie mieszczą się w jego zakresie tematycznym.

Biblioteki i repozytoria cyfrowe gromadzą różne zasoby, zarówno ze względu na tematykę, jak i formę publikacji, w tym także prace naukowe, które są głównym przedmiotem badań wspomnianego projektu badawczego. Różnorodność zgromadzonych w nich materiałów wymagała więc znalezienia sposobu wyodrębniania spośród nich współczesnych publikacji naukowych, do czego planowano wykorzystać metadane trzech wyróżnionych typów. Metadane zarówno w repozytoriach, jak i w bibliotekach cyfrowych prezentowane są w tym samym schemacie, tj. Dublin Core, co pozwala na zastosowanie do ich oceny jednakowych metod badawczych.

Badanie metadanych można przeprowadzić na wiele sposobów, korzystając z metod automatycznych albo za pomocą analiz jakościowych wykonywanych przez eksperta. Oba podejścia mają swoje zalety oraz ograniczenia. Wykorzystanie technik automatycznych jest mniej pracochłonne w dłuższej skali czasowej, jednakże problemem może być niejednorodny sposób zapisu metadanych, co utrudnia ich automatyczne przetwarzanie do celów badawczych. Spełnienie warunku jednolitości zapisu, jako silnie zależnego od człowieka, niestety nie zawsze jest możliwe. Różna pragmatyka indeksowania i staranność wprowadzania metadanych przez różne instytucje i indeksatorów skutkuje znacznym zróżnicowaniem metadanych opisujących te same własności dokumentów. Z kolei analizy eksperckie są zdecydowanie bardziej czasochłonne i nie gwarantują sukcesu ze względu na rozmiary zbiorów danych do opracowania. Dlatego na potrzeby opisywanych badań zastosowano podejście zintegrowane. Dane badawcze były pobierane oraz przetwarzane automatycznie, zaś analizy wykonywano hybrydowo, z wykorzystaniem automatycznych metod NLP (ang. *Natural Language Processing*), wizualizacji informacji oraz analiz eksperckich.

W procesie badawczym zaplanowano następujące kroki:

- (1) pobieranie wszystkich metadanych z wymienionych wcześniej platform agregujących zasoby cyfrowe;
- (2) przetwarzanie metadanych i ocena ich jakości.

3. Pobieranie metadanych

Do realizacji etapu pierwszego zdecydowano się wykorzystać istniejące rozwiązania informatyczne takie jak eksport z bazy danych oraz automatyczne pobieranie danych. Metadane z FBC zostały udostępnione dzięki uprzejmości pracowników PCSS, zaś dane z Agregatora CEON pobrano za pomocą protokołu OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*), co pozwala na automatyczne uaktualnianie tych danych w przyszłości. Ogółem, na początku projektu pobrano 1 714 571 rekordów ze 101 bibliotek i repozytoriów cyfrowych dostępnych za pośrednictwem FBC (w tym 87 bibliotek cyfrowych) oraz 53 899 rekordów dostępnych w Agregatorze CEON¹².

Dane zostały poddane operacjom usunięcia duplikatów, konwersji zapisu kodowego na UTF-8 oraz wyodrębnienia informacji z pól wskazanych we wcześniejszej części niniejszego artykułu.

¹² W chwili oddania artykułu do druku Agregator CEON rejestrował 71 606 dokumentów.

4. Analiza metadanych

W niniejszym podrozdziale zostanie zaprezentowany opis stanu metadanych w polach odpowiadających trzem wybranym elementom standardu Dublin Core, tj. *dc:type*, *dc:rights* oraz *dc:date*.

4.1. Element TYP

Ze względu na fakt, że do badań wykorzystano jedynie metadane dokumentów cyfrowych, bez ich pełnej treści, wstępną klasyfikację zasobów na naukowe oraz nienaukowe przeprowadzono na podstawie zawartości pola opisu typu dokumentu (*dc:type*). W tym polu pojawiają się określenia wskazujące czy dany dokument jest np. dysertacją naukową, artykułem, pracą dyplomową czy monografią.

W metadanych polskich bibliotek cyfrowych można zaobserwować brak jakiegokolwiek normalizacji, co powoduje, że w poszczególnych polach opisów pojawiają się hasła niejednolite w treści (wykorzystanie różnych synonimów tego samego hasła podstawowego) i formie gramatycznej (niekonsekwentne stosowanie w opisach liczby pojedynczej i mnogiej). Dotyczy to również pola *dc:type*, także w przypadku określeń wskazujących na potencjalną pracę naukową. W polu tym pojawiają się na przykład następujące wartości:

- *e-rozprawa habilitacyjna*,
- *e-rozprawa habilitacyjna PŁ*,
- *habilitacja*,
- *habilitacje*,
- *rozprawa habilitacyjna*,
- *dysertacja*,
- *dysertacje*.

Sama niekonsekwencja w stosowaniu liczby pojedynczej i mnogiej nie stanowi dużego problemu podczas automatycznej analizy danych, ponieważ stosunkowo łatwo można było takie opisy ujednoczyć, np. za pomocą lematyzacji¹³. Wymaga to jednak dodatkowego nakładu pracy analitycznej oraz odpowiedniego zaprojektowania systemu przetwarzania danych. Należy natomiast podkreślić, że występowanie tego zjawiska w metadanych z tej samej biblioteki cyfrowej świadczy o nieprzestrzeganiu lub braku standardów jakości przy wprowadzaniu opisów zasobów.

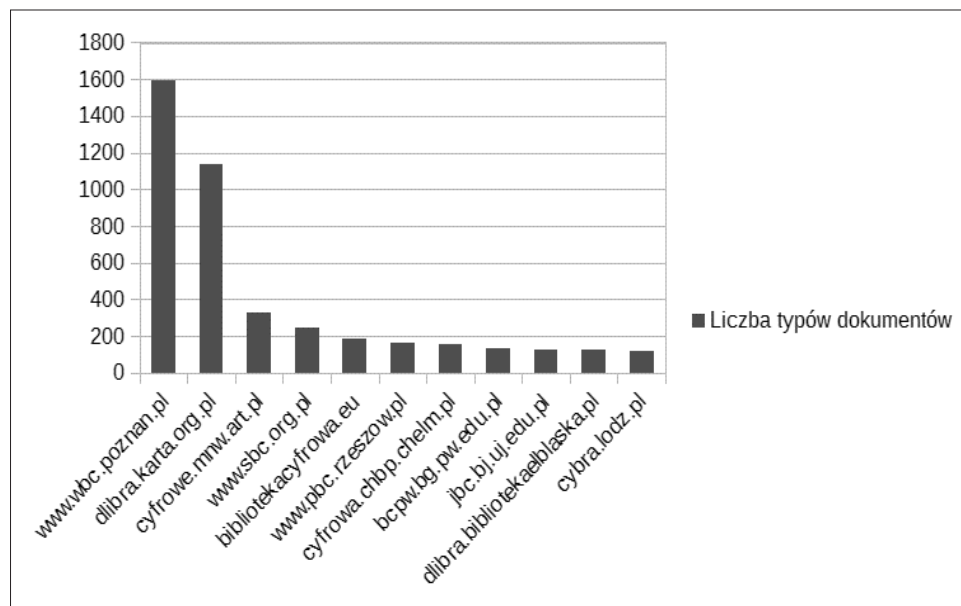
Ze względu na wymienione niekonsekwencje w opisach typów dokumentów stosowanych przez różne biblioteki cyfrowe problematyczne okazało się sformułowanie zapytania wyszukiwawczego, które uwzględniłoby wszystkie potencjalne typy prac naukowych stosowane przez różne biblioteki cyfrowe. Zastany brak jednolitości i konsekwencji w stosowaniu terminów w polach opisu schematu Dublin Core wymusza zastosowanie kilku zapytań wyszukiwawczych dla tego samego typu dokumentu (dotyczy również np. takich samych praw dostępu, itp.).

Dodatkowym utrudnieniem w klasyfikowaniu dokumentów na podstawie metadanych były umieszczane w polu *dc:type* wartości całkowicie nieprawidłowe, jak na przykład: tytuł, wymiary fizyczne oryginału, data publikacji, autor.

¹³ Lematyzacja – proces wyznaczania podstawowej formy gramatycznej, na przykład mianownika liczby pojedynczej, dla wyrazów w tekście.

W wielu przypadkach zaobserwowano nadreprezentację typów dokumentów. Sporo bibliotek, niebędących bibliotekami specjalistycznymi, zbyt rozdrobniło dane, co także utrudnia klasyfikację dokumentów wyłącznie na podstawie ich typów. Na 87 analizowanych zbiorów danych z bibliotek cyfrowych, które były dostępne na platformie FBC na początku badań, aż 11 rejestrowało powyżej 100 typów dokumentów włączając w tę liczbę typy opisane zarówno w liczbie pojedynczej, jak i w mnogiej oraz określenia podane w językach obcych. Te ostatnie traktowane były jako wyrażenia wielowyrazowe, zawierające termin w języku polskim i obcym, przy czym często w ramach opisów jednej biblioteki występowało kilka form zapisu. Zestawienie tych bibliotek oraz liczby rejestrowanych w nich typów dokumentów prezentuje rysunek 1.

Rys 1. Biblioteki rejestrujące powyżej stu typów dokumentów



Źródło: opracowanie własne.

W Wielkopolskiej Bibliotece Cyfrowej¹⁴ odnotowano rekordową liczbę 1594 stosowanych określeń dla typów dokumentów (wliczając w to tłumaczenia nazw typów na język angielski oraz niemiecki), w tym m.in. różne nazwy o tym samym znaczeniu. Różnorodność określeń typów nie dotyczy zatem tylko odrębnych bibliotek, ale występuje także w ramach zasobu pochodzącego z jednej biblioteki cyfrowej, co bardzo utrudnia badanie zasobów, ich grupowanie, czy ocenę jakości metadanych. W tabeli 2 zaprezentowano zestawienie wartości pola *dc:type* dla typu dokumentu: *artykuł* stosowanego w różnych bibliotekach cyfrowych (po przecinku podane są zawartości powielonego pola *dc:type*: tłumaczenia lub synonimy). Należy przy okazji zauważyć, że samo określenie *artykuł* jest niejednoznaczne i nie pozwala określić czy opisuje pracę naukową czy publicystyczną.

¹⁴ <http://www.wbc.poznan.pl>

Tab. 2. Zestawienie wartości pola *dc:type* dla typu dokumentu *artykuł* w różnych bibliotekach cyfrowych

article
article, artykuł, press article
artikel, artykuł, article
artikel, artykuł, press article
artykuł konferencyjny
artykuł naukowy
artykuł z czasopism ogólnopolskich
artykuł z czasopisma
artykuł, article, art
art z czasopisma
artykuł, articles
artykuł, postprint
artykuły
artykuły historyczno-prawne, zeszyt naukowy
artykuły z czasopism ogólnopolskich
artykuły, czasopisma
artykuły, maszynopisy
artykuły, opracowania historyczne
artykuły, polemiki
artykuły, recenzje
e-artykuł
fragment artykułu

Dużą różnorodność form typu publikacji może zobrazować przykład Bałtyckiej Biblioteki Cyfrowej¹⁵. W metadanych z tej biblioteki odnotowano 185 form zapisu typów dokumentów, wliczając w to połączone w listę zapisy z pól *dc:type* dla języków polskiego, niemieckiego oraz angielskiego :

- *magazin, czasopismo, journal;*
- *zeitschrift, czasopismo, journal;*
- *zeitschriften, czasopismo, journal;*
- *zeitschrift, gazeta, journal;*
- *czasopismo, journal;*
- *zeitschrift, czasopismo, periodical.*

O ile dla języka angielskiego można stwierdzić pewną jednolitość formy, to dla języków niemieckiego i polskiego widać dużą swobodę w doborze terminów opisujących typ dokumentu. W zapisie oryginalnym metadanych pobranych z FBC przykładowy rekord z BBC ma następujący zapis:

```
<dc:type xml:lang="de"><![CDATA[magazin]]></dc:type>
<dc:type xml:lang="pl"><![CDATA[czasopismo]]></dc:type>
<dc:type xml:lang="en"><![CDATA[journal]]></dc:type>
```

Bezpośrednio na stronie Bałtyckiej Biblioteki Cyfrowej, w indeksie *Typ zasobu* można znaleźć m. in. hasła w liczbie pojedynczej i mnogiej, tzw. literówki w hasłach, itp. Zostały one zaprezentowane na rysunku 2.

¹⁵ <http://bibliotekacyfrowa.eu>

Rys. 2. Indeks Typ zasobu, Bałtycka Biblioteka Cyfrowa,
<http://bibliotekacyfrowa.eu/dlibra/keywordindex?dirids=1&attId=9>



Należy jednak podkreślić, że podobna sytuacja braku kontroli i konsekwencji w stosowaniu nazw typów dokumentów panuje w zdecydowanej większości bibliotek cyfrowych dostępnych poprzez platformę FBC. Wyniki wstępnego etapu badań, polegającego na analizach danych zawartych w opisach cyfrowych dokumentów, pozwalają wysnuć pesymistyczny wniosek, że typy dokumentów w polskich bibliotekach cyfrowych opisywane są w wielu przypadkach z pominięciem jakichkolwiek norm, czy to polskich (np. PN-92-N-01227) czy międzynarodowych (np. ANSI/NISO Z39.85–2012 The Dublin Core Metadata Element Set).

Nawet w sytuacji, gdy opisy typów dokumentów są ujednolicone i stosowane konsekwentnie, nie ułatwia to wyodrębnienia prac naukowych. W zestawieniu prezentującym szczegółową listę typów dokumentów stosowanych w metadanych z Kujawsko-Pomorskiej Biblioteki Cyfrowej¹⁶ (KPBC), niewiele typów można zaliczyć do opisujących publikacje naukowe (wyróżnione w zestawieniu). Poważny problem interpretacyjny powstaje przy typie *książka* oraz *artykuł*, ponieważ pierwsze określenie może oznaczać zarówno monografię naukową, jak i beletrystykę, zaś drugie artykuł naukowy lub publicystyczny. Wartości

¹⁶ <http://kpbc.umk.pl>. W KPBC wdrożono restrykcyjne zasady stosowania nazw typów, ograniczając tym samym ich redundancję, por. (Derfert-Wolf, 2016).

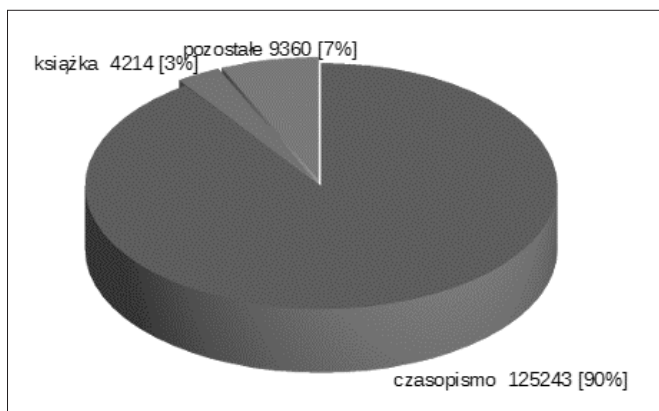
liczbowe umieszczone przy nazwach typów to łączne liczby udostępionych publikacji zawierających w opisie podaną wartość atrybutu *dc:type*.

- | | |
|--|------------------------------------|
| (1) czasopismo 125170 | (40) postprint 13 |
| (2) książka 4214 | (41) plakat 12 |
| (3) grafika 1420 | (42) obwieszczenie 11 |
| (4) ekslibris 1035 | (43) śpiewnik 10 |
| (5) mapa 749 | (44) maszynopis powielony 9 |
| (6) druk muzyczny 593 | (45) atlas 9 |
| (7) materiał aktowy 590 | (46) album 9 |
| (8) pocztówka 548 | (47) wystawa 8 |
| (9) artykuł 543 | (48) raport 8 |
| (10) odbitka 499 | (49) film dokumentalny 8 |
| (11) afisz 409 | (50) maszynopis 7 |
| (12) malarstwo 404 | (51) kalendarz 7 |
| (13) broszura 400 | (52) teka 6 |
| (14) fotografia 304 | (53) skrypt 5 |
| (15) rysunek 270 | (54) reprint 5 |
| (16) nadbitka 188 | (55) wydawnictwo okolicznościowe 4 |
| (17) starodruk 180 | (56) partytura 4 |
| (18) zaproszenie 147 | (57) atlas starodruczny 4 |
| (19) rękopis muzyczny 112 | (58) ulotka wyborcza 3 |
| (20) inkunabuł 111 | (59) postinkunabuł 3 |
| (21) rękopis 94 | (60) plan miasta 3 |
| (22) preprint 94 | (61) ebook 3 |
| (23) druk 87 | (62) cennik 3 |
| (24) nagranie dźwiękowe 67 | (63) legitymacja 2 |
| (25) przewodnik 36 | (64) księga pamiątkowa 2 |
| (26) prezentacja 36 | (65) znaczek okolicznościowy 1 |
| (27) jednodniówka 32 | (66) widok 1 |
| (28) katalog wystawy 28 | (67) telegram 1 |
| (29) ulotka 26 | (68) tekst do mapy 1 |
| (30) odezwa 25 | (69) skład osobowy 1 |
| (31) informator 22 | (70) praca magisterska 1 |
| (32) pamiętnik 21 | (71) plan twierdzy 1 |
| (33) dyplom 21 | (72) pieśń 1 |
| (34) kopia rękopiśmienna druku muzycznego 20 | (73) partytura chóralna 1 |
| (35) program 16 | (74) nekrolog 1 |
| (36) list 16 | (75) mikrofilm 1 |
| (37) statut 14 | (76) komiks 1 |
| (38) druk reklamowy 14 | (77) hasło słownikowe 1 |
| (39) program teatralny 13 | (78) fotokopia 1 |
| | (79) film 1 |

Zaprezentowana różnorodność typów oraz problemy z jednoznacznym wskazaniem prac naukowych tylko na podstawie typu publikacji podanego w opisach dokumentów są wspólne dla wszystkich polskich bibliotek cyfrowych.

Na rysunku 3 przedstawiono udział dokumentów określanych jako książka, czasopismo lub za pomocą nazw innych typów w zasobach KPBC. Podobny rozkład typów dokumentów stwierdzono dla większości polskich bibliotek cyfrowych.

Rys. 3. Ilościowy rozkład typów dokumentów w zasobach KPBC, dane z 2 lutego 2016, źródło: <http://kpbc.umk.pl/dlibra/pubstats>



Przeważający typ zasobów w polskich bibliotekach cyfrowych stanowią wydawnictwa ciągle, opisane zazwyczaj jako *czasopisma*. Na podstawie praktyki oraz analizy dat powstania zasobów można stwierdzić, że są to przeważnie dokumenty z domeny publicznej, czyli, zgodnie z założeniem opisywanych badań, niespełniające wymogu współczesności. Ponadto wiele bibliotek – w tym także KPBC – udostępnia czasopisma naukowe zdigitalizowane w całości, w zeszytach, bez rozbicia na poszczególne artykuły, co nie ułatwia analiz i selekcji danych. Ze względu na brak opisów indywidualnych artykułów, prace zawarte w tak zdigitalizowanych periodykach nie są dostępne do automatycznego przetwarzania tylko na podstawie metadanych. W celu ich analizy konieczne jest wyodrębnianie poszczególnych jednostek ze zdigitalizowanego zeszytu, rocznika, itp. Wymagałoby to użycia narzędzi rozpoznawania granic artykułów oraz wskazania ich części nagłówkowej (tytuł, autorzy, słowa kluczowe, itd.). W ramach omawianych badań tego rodzaju działań nie przeprowadzono.

Odmienną zawartość pod względem typów oraz wieku dokumentów reprezentują cyfrowe repozytoria naukowe. Zawierają one tylko współczesne prace naukowe. Z powodu małej liczby tego typu kolekcji cyfrowych w Polsce, w porównaniu do bibliotek cyfrowych, nie można było badań zawęzić jedynie do zasobów udostępnianych w repozytoriach. Ponadto analiza metadanych z bibliotek cyfrowych pozwala na wypracowanie bardziej uniwersalnych metod selekcji i analizy danych.

Ostatecznie, w wyniku analiz eksperckich, przeprowadzonych na materiale przygotowanym w procesie automatycznego przetwarzania i grupowania typów dokumentów, występujących w metadanych polskich bibliotek cyfrowych, wyodrębniono zbiór około 400 typów, które z dużym prawdopodobieństwem wskazują na pracę naukową. Poniżej zaprezentowany został fragment zestawienia (opracowanie własne na podstawie analizy i grupowania metadanych z polskich bibliotek cyfrowych):

- *art z czasopisma*
- *artykuł konferencyjny*

- *artykuły historyczno-prawne, zeszyt naukowy*
- *book, peerreviewed*
- *czasopismo naukowe polskie*
- *czasopismo, artykuł*
- *diplomarbeit, praca dyplomowa, diploma paper*
- *doktorat*
- *proceedings, e-lecture, audiovisual document*
- *dokument elektroniczny, dokument naukowo-dydaktyczny*
- *dokument elektroniczny, rozprawa habilitacyjna*
- *dokument naukowo-dydaktyczny*
- *dokument piśmienniczy, zeszyty naukowe politechnika łódzka, zeszyty naukowe pt*
- *e-książka, rozprawa habilitacyjna*
- *e-zeszyty naukowe pt, dokument elektroniczny, electronic resource, e-tul scientific bulletins*
- *fragment artykułu*
- *habilitacja*
- *hasło słownikowe*
- *komentarz do artykułu*
- *książka, rozprawa habilitacyjna*
- *książka, rozprawa habilitacyjna, book, dissertation.*

4.2. Element – PRAWA

Poprawne i prawidłowe rozpoznanie oraz oznaczenie praw do udostępnianego cyfrowo utworu jest niezwykle istotne. Pozwala użytkownikom rozpoznać zakres w jakim mogą z danego utworu korzystać (Bednarek-Michalska, 2014), a w przypadku badań przeprowadzanych na dużych zbiorach danych, za pomocą technik automatycznych, pozwala dokonać selekcję zasobów według kryteriów dostępności, ale także np. wieku publikacji. W trakcie opisywanych badań poddano zatem analizie również zawartość pola *dc:rights*. Jak wspomniano wcześniej, dla potrzeb projektu badawczego, którego celem było zbadanie struktury i dynamiki polskich naukowych zasobów cyfrowych przyjęto założenie, że obiekty oznaczone jako należące do domeny publicznej nie spełniają wymogu współczesności. Ocena współczesności dokumentu na podstawie zawartości pola *dc:rights* jest komplementarna w stosunku do oceny na podstawie zawartości pola *dc:date* – pozwala z jednej strony zweryfikować poprawność decyzji podjętej na podstawie podanej daty publikacji, a z drugiej strony pozwala sklasyfikować dokument, dla którego nie podano daty publikacji. Dzięki takiemu założeniu możliwe byłoby ignorowanie podczas badań wszystkich utworów trafiających do domeny publicznej (DP) z racji ich wieku.

W wyniku analizy zawartości pól *dc:rights* w danych z polskich bibliotek cyfrowych okazało się, że hipotezy tej nie można jednoznacznie uznać za potwierdzoną, ponieważ w większości bibliotek cyfrowych nieprawidłowo oraz niekonsekwentnie oznaczany jest status prawny utworu. Bardzo często w polu *dc:rights* spotykane są nieprawidłowości zarówno w treści, jak i formie wpisu. Analiza zawartości pól *dc:rights* wykazuje, że w większości zasobów polskich bibliotek cyfrowych, podobnie jak w przypadku pól *dc:type* oraz *dc:date*, panuje wielka dowolność oraz brak jakiegokolwiek kontroli treści i formy wpisów.

W wyszukiwarce FBC na zapytanie o *domenę publiczną* w polu PRAWA uzyskano w marcu 2015 r. liczbę 114 985 dokumentów. Jednakże na podstawie analizy zawartości zbiorów większości polskich bibliotek cyfrowych można stwierdzić, że było to zdecydowanie za mało (Derfert-Wolf, 2011). Do domeny publicznej powinno być przypisanych, zgodnie z ich faktycznym stanem prawnym, znacznie więcej zasobów z 2.5 mln, jakie były w badanym okresie udostępniane przez FBC.

W odpowiedzi na uwagę, że wyszukiwarka FBC może działać niepoprawnie, PCSS przekazało informację, że narzędzie to jest ciągle rozwijane i jeszcze nie obsługuje wielu typów zapytań. Twórcy FBC obiecali, że wprowadzą poprawki w systemie wyszukiwania. W styczniu 2016 r., po wdrożeniu przez PCSS poprawek w funkcjonowaniu wyszukiwarki FBC (m.in. wynikających ze zgłoszonych przez autorów niniejszego artykułu sugestii), ponownie zadano zapytanie: *dc_rights:(domena publiczna)* i uzyskano w odpowiedzi 1 206 874 dokumenty. Wartość ta wydaje się być zbliżona do faktycznej liczby tego typu dokumentów, jednakże nadal niedoszacowana, prawdopodobnie z powodu omówionych nieprawidłowości występujących w metadanych przekazywanych przez poszczególne polskie biblioteki cyfrowe.

Ewaluacje metadanych, przeprowadzone na potrzeby omawianych badań, dla zasobów bibliotek cyfrowych (z wyłączeniem repozytoriów) dostępnych w agregatorze FBC na początku 2015 r. wykazały, że na 1 650 075 rekordów w 1 429 267 (86% zasobów) z nich jest wypełnione pole *dc:rights*. W przypadku 762 588 rekordów (46% zbioru) pole to ma wartość *domena publiczna* w różnej, ale wciąż rozpoznawalnej jednoznacznie formie. Aż 13 spośród 87 analizowanych bibliotek cyfrowych nie wypełniało pola *dc:rights* w swoich metadanych. Natomiast łącznie w metadanych 35 bibliotek nie pojawiło się określenie „domena” w polu *dc:rights*.

Analiza zawartości pola wskazującego na prawa własności w bibliotekach cyfrowych dostępnych przez FBC wykazała aż 224 wzorce stosowane dla oznaczenia dzieł z domeny publicznej. Oprócz zalecanego określenia *domena publiczna*, można spotkać również m.in. wpisy:

- *domena publiczna (public domain)*,
- *domena publiczna (public domain) / dla wszystkich bez ograniczeń*,
- *dostępne publicznie bez ograniczeń – domena publiczna*.

Niestety, większość wpisów obarczona była błędami. Należały do nich wskazywanie okresu powstania publikacji (np. XIII w., XVII–XVIII w.) lub daty śmierci autora czy też tłumacza, z zachowaniem różnych zasad stosowania znaków interpunkcyjnych, np.:

- *Domena Publiczna – Adam Mickiewicz zm. 1855*
- *Domena Publiczna – Adam Mickiewicz zm. 1855.*
- *Domena Publiczna – Adam Mickiewicz, zm. 1855*

Kolejnymi błędami występującymi we wpisach w polu *dc:rights* były tzw. literówki, np. Ulrich i Urlich (Leon Ulrich), czy wręcz błędy w dacie śmierci:

- Domena Publiczna – Bolesław Leśmian zm. 1937*
- Domena Publiczna – Bolesław Leśmian zm. 1938*

Na podstawie analizy wpisów w polu *dc:rights* w metadanych zbiorów cyfrowych można stwierdzić, że niektóre instytucje udostępniające zasoby cyfrowe rzetelnie wypełniają zawartość tego pola. W zakresie poprawności treści w opisach wyróżniają się np. zbiory Polskiej Akademii Nauk (PAN), dostępne na platformie Repozytorium Cyfrowego Instytutów

Naukowych (RCIN)¹⁷. Większość obiektów udostępnianych przez PAN poprzez platformę RCIN stanowią współczesne prace naukowe, zazwyczaj oznaczone jako chronione prawem autorskim: „copyright”.

Kolejny przykład instytucji konsekwentnie i poprawnie oznaczającej prawa do obiektów stanowi Kujawsko-Pomorska Biblioteka Cyfrowa (KPBC). Zgodnie z wewnętrznymi wytycznymi KPBC każdy obiekt cyfrowy musi mieć pole „prawa” wypełnione wg przyjętego jednego wzorca, zaś bibliotekarze zostali w tym zakresie przeszkoleni¹⁸. Niżej przedstawiono przykład opisu dla czasopisma naukowego, które jest własnością UMK i wszelkie prawa zostały przez poszczególnych autorów przeniesione na uczelnię:

Tytuł: Acta Universitatis Nicolai Copernici. Nauki Matematyczno-Przyrodnicze. Geografia

Prawa: Wszystkie prawa zastrzeżone

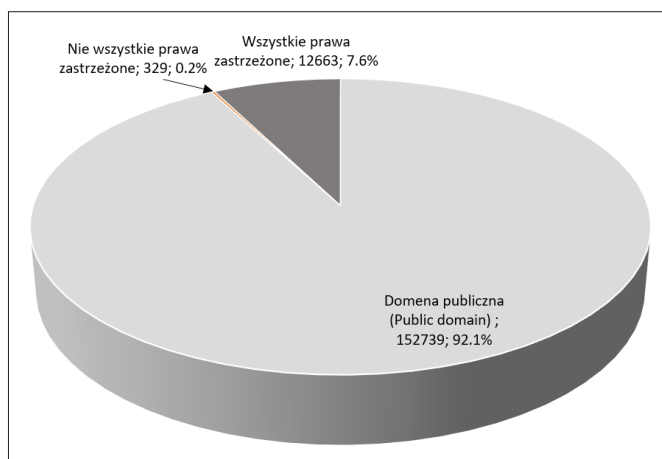
Właściciel praw: Uniwersytet Mikołaja Kopernika w Toruniu

Licencja: Licencja UMK

Prawa dostępu: Dla wszystkich w zakresie dozwolonego użytku

Rysunek 4 prezentuje udział dokumentów o różnych prawach dostępu w zbiorach KPBC. Zaprezentowane proporcje są typowe dla polskich bibliotek cyfrowych (z wyłączeniem repozytoriów).

Rys. 4. Rozkład wartości pola PRAWA w KPBC.
Dane z 2 lutego 2016, źródło: <http://kpbc.umk.pl/dlibra/pubstats>



Analiza pola PRAWA pokazuje, że nie można bezkrytycznie wykorzystywać tych danych do selekcji dokumentów współczesnych. Przyczyną tego stanu jest niewypełnianie pola *dc:rights* przez wiele instytucji, lub też brak kontroli słownictwa używanego w opisach w przypadku większości bibliotek, które podawały status prawny publikacji. Opisy praw do obiektu mogą zostać wykorzystane w automatycznym przetwarzaniu metadanych pod warunkiem odpowiedniego ich wstępnego przetworzenia, np. poprzez uwzględnienie różnych potencjalnych form zapisu oznaczającego domenę publiczną.

¹⁷ <http://rcin.org.pl>

¹⁸ Źródło: informacja od dyrekcji KPBC.

4.3. Element – DATA

Podobnie jak w przypadku pól *dc:type* i *dc:rights* problemy z niejednorodnością opisu występowały również w zasobach tej samej biblioteki cyfrowej w polu *dc:date* oznaczającym datę lub okres czasu związany ze zdarzeniem w cyklu istnienia zasobu. Można wskazać zastosowane różnorodne standardy zapisu daty, nie zawsze zgodne z zaleceniami DCMES, oraz nieściśłości w oznaczaniu dat publikacji. Sam zapis daty także przybierał rozmaite formy:

- RRRR-MM-DD
- DD-MM-RRRR
- RRRR.MM.DD
- DD.MM.RRRR

gdzie RRRRR oznacza rok w postaci pełnej (czterocyfrowej), MM – miesiąc, zaś DD – dzień.

Oprócz różnej kolejności elementów opisujących części daty, pojawiają się różne znaki oddzielające te części (np.: „, „-“). W polach *dc:date* występują też inne sformułowania, nie zawsze konsekwentnie używane w obrębie metadanych z jednej biblioteki, jak:

- „po RRRR”
- „przed”
- „post ante”
- „ok.”
- pierwsza połowa ... wieku
- „ca...”

Ponieważ jednak sam rok jako ciąg czterech następujących po sobie cyfr łatwo poddaje się automatycznemu rozpoznawaniu, stosunkowo prosto udało się z pól *dc:date* pozyskać roczne daty publikacji dokumentów. Zawartość tego właśnie pola stała się głównym wyznacznikiem współczesności publikacji w dalszych etapach badań.

5. Analiza ekspercka danych

W celu weryfikacji skuteczności zastosowanych w projekcie metod automatycznego przetwarzania i analizy statystycznej metadanych zdecydowano się przeprowadzić analizę ekspercką zasobów wybranych bibliotek. Decyzja ta wynikała również z faktu stosowania różnych zasad zapisu metadanych w poszczególnych bibliotekach cyfrowych, z czym wiązała się konieczność zastosowania bardzo ogólnych zasad wyboru i klasyfikacji dokumentów podczas etapu analizy automatycznej. Podczas analizy należało uwzględnić sytuację wydzielenia, przez niektóre biblioteki cyfrowe, prac naukowych jako odrębnej kolekcji. W takich przypadkach podana również była liczba tych prac. Upowszechnienie tej praktyki podniosłoby wartość metadanych jako surowych danych badawczych oraz ułatwiłoby dalsze automatyczne analizy, m. in. naukometryczne.

Jako przykład wydzielonej kolekcji o charakterze naukowym może posłużyć zasób KPBC zatytułowany „Materiały dydaktyczne”. Zgodnie z opisem: *Kolekcja zawiera artykuły, monografie i podręczniki akademickie wydane lokalnie po 1945 r. oraz opracowania historyczne niezbędne dla nauki i dydaktyki. Opiera się o prace pracowników nauki i wydawców naszego regionu, którzy zdecydowali się powierzyć nam prawa do ich udostępnienia. Liczba publikacji w kolekcji: 10946. Liczba publikacji w kolekcji i we wszystkich podkolekcjach: 11038 (KPBC, 2016).*

W przypadku braku wydzielonych kolekcji o charakterze naukowym, konieczne było przeprowadzenie bardziej szczegółowych i wielopoziomowych analiz. Zastosowano metodę analizy jakościowo-ilościowej i podjęto następujące kroki, analogiczne do etapów przetwarzania i analizy automatycznej:

- (1) przeglądanie danych statystycznych udostępnionych przez biblioteki cyfrowe w celu ustalenia:
 - a. typów zasobów (analiza pola opisu TYP z metadanych);
 - b. praw do zasobów (analiza pola PRAWA przeprowadzona w celu ustalenia liczebności zasobów w domenie publicznej);
- (2) jeśli statystyki dostępne na stronach poszczególnych bibliotek nie podawały tych informacji korzystano z zaawansowanego wyszukiwania i przeszukiwano wskazane wyżej pola opisu;
- (3) jeśli wyszukanie zaawansowane nie przynosiło oczekiwanych rezultatów, przeglądano dostępne kolekcje w celu określenia liczebności publikacji współczesnych (zasoby dydaktyczne, artykuły, doktoraty oraz starych (dziedzictwo kulturowe). Niekiedy pomocny był własny podział kolekcji wdrożony przez poszczególne biblioteki cyfrowe, nazwy kolekcji często sugerują, która część zawiera jakie publikacje. Wybór kolekcji powoduje wyświetlenie informacji o jej liczebności.

W tabeli 3 przedstawiono wyniki wybranych biblioteki poddanych badaniom i problemy, jakie rodzi podobna analiza. Do próby badawczej wybrano tylko te biblioteki, które mają większą liczbę zasobów. Wykluczono biblioteki, w których działanie wyszukiwarki wskazywało na ewidentne błędy techniczne.

Tab. 3. Wyniki analizy zasobów naukowych w wybranych bibliotekach. Stan na marzec 2015 r.

Nazwa biblioteki cyfrowej	Liczba dostępnych obiektów cyfrowych	Liczba gazet i czasopism ¹	Liczba obiektów należących do domeny publicznej ²	Liczba domniemanych współczesnych prac naukowych wg analiz szczegółowych ³
I	II	III	IV	V
Jagiellońska Biblioteka Cyfrowa	270 875	248 973 opisane jako: czasopismo	255 356 opisane jako: domena publiczna (public domain)	Odnaleziono zapis w polu Prawa: copyright 6036; licencja CC 200. To może wskazywać, że są to pozycje naukowe. Zidentyfikowano kolekcję UJ: 4442. Razem 6036.
Wielkopolska Biblioteka Cyfrowa	251 896	gazeta: 59 559 czasopismo: 48 679 gazety: 45 976 czasopisma: 45 389	jako domena publiczna oznaczono jedynie 25 078 opisów, inne obiekty nie mają w ogóle oznaczeń statusu prawnego.	Zidentyfikowano kolekcję materiały dydaktyczne: 4327. Znalezione typy: monografia 1575; monografie 631; rozprawa doktorska 1054; artykuł 779; artykuły 391; oprawa habilitacyjna 583; podręczniki akademickie 288; podręczniki 152. Razem 5453.

I	II	III	IV	V
e-biblioteka Uniwersytetu Warszawskiego	204005	201 367 czasopismo	140 193 domena publiczna	W polu prawa znaleziono zapis: wszystkie prawa zastrzeżone dla 348 pozycji oraz około 150 na innych licencjach, co może wskazywać na prace współczesne i naukowe. Jest wyodrębniona naukowa kolekcja e-publikacje 45. Razem 443.
Śląska Biblioteka Cyfrowa	137099	czasopismo 108 525	domena publiczna: 60000 obiektów, nie wszystkie jednak mają opis	Jest kolekcja materiałów dydaktycznych i naukowych 9964. Znaleziono typy: monografia historyczna 190; opracowanie naukowe 181; opracowanie statystyczne 4148; artykuł 1151; rozprawa doktorska 436.
Małopolska Biblioteka Cyfrowa	88 521	czasopismo 78 323	jako domena publiczna oznaczono jedynie 17 817, nie wszystkie obiekty są opisane	Oznaczono typ: artykuł 137. Oznaczono także osobno kolekcję naukową i dydaktyczną 23 691. Zakładamy, że typ artykuł znalazł się w kolekcji naukowej.
Kujawsko-Pomorska Biblioteka Cyfrowa	80 548	czasopismo 68 890	domena publiczna 69 534	Oznaczono także pola prawa z wartościami: wszystkie prawa zastrzeżone 9925; nie wszystkie prawa zastrzeżone 312 czyli licencje CC. Wydzielona kolekcja materiały dydaktyczne 11 038.
Biblioteka Cyfrowa Uniwersytetu Wrocławskiego	52 427	czasopisma 22 614 wydawnictwa ciągłe 8 691	domena publiczna 50 174	Znaleziono około 1000 obiektów, które mają oznaczenia copyright. Jest kolekcja publikacje współczesne 711; e-czytelnia, obiekty współczesne zamknięte 164, materiały edukacyjne 195, e-książki 147. Razem 1217.
Repozytorium Cyfrowe Instytutów Naukowych	52 197	czasopismo 593	domena publiczna 9 350	Odnaleziono opis w polu PRAWA: zastrzeżone – dostęp ograniczony 21 131; prawa zastrzeżone – dostęp nieograniczony 21 087; prac na licencjach CC jest około 300. Zakłada się, że w tym repozytorium wszystkie prace są naukowe. Należy odjąć obiekty z domeny publicznej. Razem 42 847.
Bałtycka Biblioteka Cyfrowa	43 910	czasopismo 25 760 gazeta 2 892	Domena publiczna 5 293 Domena Publiczna 43 684	Odnaleziono kolekcję artykuły 126. W opisach także są określenia: Licencja instytucjonalna 412; licencja instytucjonalna 30 (mała litera). Razem: 126.

I	II	III	IV	V
Mazowiecka Biblioteka Cyfrowa	31 218	czasopism 27 914	domena publiczna 30 745	Nie ma wydzielonej kolekcji naukowej, pojedyncze prace naukowe, artykuły: 3 Razem: 3
Zachodniopomorska Biblioteka Cyfrowa „Pomerania”	31 168	czasopismo 21 015	domena publiczna 6 994 spora część zasobów została nieprawidłowo przypisana poszczególnym bibliotekom	Znaleziono typy: dokument naukowo-dydaktyczny 148; praca doktorska 91; materiały konferencyjne 10; prezentacja multimedialna 17. Są kolekcje: nauka, dydaktyka 424; doktoraty i habilitacje 105; mat. konf. 19. Razem: 548.
Dolnośląska Biblioteka Cyfrowa	25 153	czasopismo 9 784	brak określenia domena publiczna	Znaleziono typy: rozprawa doktorska 472; doktorat 6; materiały konferencyjne 52, artykuły 3857; rozprawy habilitacyjne 74; recenzje 27; raporty 3; bibliografie 15; książki (wiele współczesnych opracowań) 3986; wykład 1. Razem: 8493
Biblioteka Cyfrowa UMCS	16 889	czasopismo 11 29	domena publiczna 10 417	Odnaleziono kolekcje: nauka i dydaktyka 4560; Zidentyfikowano typy: artykuł 1923; rozprawa doktorska 5; monografia 1; praca mgr 1; dyplom 12; referat 4; dokument elektroniczny 7.
Zielonogórska Biblioteka Cyfrowa	14 266	czasopismo 5 928 gazeta 1 951	brak danych	Istnieje kolekcja nauka i dydaktyka 830; habilitacje i doktoraty 128; Odnaleziono typy: rozprawa doktorska 69; artykuł 65. Razem: 958.
Repozytorium Uniwersytetu im. Adama Mickiewicza	10 741	czasopisma 8 601	brak danych	Repozytorium instytucjonalne. Wszystkie prace są naukowe współczesne, jest to dorobek UAM: 10 741 pozycji ⁵ .
<p>¹ W celu wykluczenia ich z dalszych badań, z powodów wyjaśnionych w końcowej części podrozdziału 4.1. Przetwarzanie metadanych – pole TYP.</p> <p>² J.w.</p> <p>³ W bardzo dużym zaokrągleniu i przy dużej niepewności.</p> <p>⁴ Literówka (duża litera P) powoduje, że wyszukiwanie trzeba prowadzić dwiema ścieżkami.</p> <p>⁵ Zawartość repozytoriów instytucjonalnych łatwo poddaje się analizie, ponieważ występują tam wyłącznie współczesne prace badawcze.</p>				

W analizie eksperckiej przebadano 29 największych bibliotek cyfrowych (w dwóch brak było danych powiązanych z celem badania) oraz repozytoriów naukowych, wśród których wskazano 176 147 domniemanych obiektów naukowych. Uwagi przedstawione w kolumnie V: *liczba domniemanych współczesnych prac naukowych* pokazują, że nawet taka pogłębiona analiza nie daje pewności co do rezultatów poszukiwań, za dużo w nich danych o niepewnych wartościach. Nie ma pewności, czy sumując zasoby nie popełniono błędu rachunkowego, wynikającego z tego, że bibliotekarze przypisali jeden obiekt do dwóch różnych cytowanych kolekcji czy typów. Nie wszystkie określenia typów są jednoznaczne, np. książka czy artykuł, mogą oznaczać prace naukowe, ale mogą też opisywać, np. beletrystykę. Ponadto w kolumnie II, gdzie wyodrębniono czasopisma, wiele z nich należy do domeny publicznej, ale sporo z nich jest naukowych. Zawierają one artykuły, których w metadanych nie można zidentyfikować, w większości nie brano ich zatem pod uwagę. Liczba ta jest tylko szacowana, nie można jej przyjąć za pewnik.

Na podstawie wyników badań eksperckich zaprezentowanej próbki polskich bibliotek cyfrowych można przedstawić następujące wnioski:

- (1) Zdecydowana większość zasobów dostępnych za pośrednictwem FBC (ok. 80%) stanowią utwory, które trafiły do domeny publicznej z racji wieku, co niestety nie znajduje odzwierciedlenia w metadanych z powodu występowania nieprawidłowości w wypełnianiu przez poszczególne biblioteki pola *dc:rights*. Wśród zasobów z domeny publicznej ogromną część stanowią czasopisma, skanowane ze względu na stan zachowania i regionalną ważność, co pokazują dość precyzyjnie niektóre dane statystyczne dostarczane przez biblioteki.
- (2) Z pozostałych 20% zasobów przeważają te o charakterze nienaukowym (dokumenty życia społecznego, fotografie, grafika, mapy i inne).
- (3) Większość spośród prac naukowych stanowią prace spoza zakresu nauk humanistycznych i społecznych (w zasobach politechnik zdecydowaną większość utworów stanowią prace z zakresu nauk technicznych i ścisłych).

Według danych za luty 2016 r. FBC (2016a) udostępniało ok. 1.2 mln obiektów dostępnych bez żadnych ograniczeń, co oznacza, że mogą one należeć do domeny publicznej. Kolejny milion dokumentów miało status prawny albo nieoznaczony, albo było chronione. W całym zasobie znajdowało się ponad 1 740 000 czasopism. Artykułów było nieco ponad 110 000, rozpraw doktorskich 12 700, książek 138 000, rozpraw habilitacyjnych 1540, zaś prac dyplomowych 247 (wszystkie dane wyodrębniono na podstawie *dc:type*). Oznacza to, że potencjalnych prac naukowych w FBC było w lutym 2016 r. co najmniej 14 487, zaś maksymalnie mogło ich być około 26 2487.

6. Wnioski

Analiza metadanych udostępnianych przez polskie biblioteki i repozytoria cyfrowe, zarówno automatyczna jak i ekspercka, jest bardzo trudna. Jednym z powodów tej sytuacji jest niezadowolająca dostępność danych. Dla przykładu, największy polski agregator metadanych z bibliotek cyfrowych – FBC udostępnia dane w postaci plików wyeksportowanych z bazy danych, zaś przez protokół OAI-PMH „może udostępniać wybranym podmiotom dane gromadzone z polskich instytucji nauki i kultury” (FBC, 2016b). Utrudnia to automatyczną aktualizację wyników przeprowadzanych badań.

Najważniejszym jednak problemem, jaki spotykają badacze chcący oprzeć swoje badania na metadanych polskich bibliotek cyfrowych, jest niekonsekwencja panująca w opisach bibliograficznych. Większość zasobów polskich bibliotek cyfrowych i repozytoriów opisana jest za pomocą standardu Dublin Core, ale istnieją również biblioteki, które stosują swój własny model i nie zamierzają go zmieniać, ponieważ wymagałoby to wiele pracy, szczególnie przy retrokonwersji. Również w zakresie danych z bibliotek stosujących w opisie dokumentów standard Dublin Core, nie ma zachowanej zgodności we wprowadzaniu danych do poszczególnych pól. W znacznej części dane nie są spójne nie tylko na poziomie pól opisu, ale i zawartości poszczególnych pól.

Pomimo znaczącego przyrostu liczby udostępnianych zasobów cyfrowych od 2009 r. stwierdzono, że w zakresie jakości metadanych nie zaszła znacząca zmiana, w porównaniu do ocen, wniosków i postulatów zaprezentowanych wcześniej przez A. Kazana i E. Skubałę (2008) oraz J. Potęgę (2009).

W niniejszym artykule zaprezentowano wyniki analiz zawartości tylko trzech pól schematu metadanych Dublin Core. Wskazane błędy oraz brak konsekwencji, różne formy językowe wpisów, czy brak kontroli jakości (np. literówki) występują również w pozostałych polach metadanych, co zostało potwierdzone w innej części badań, niezaprezentowanej w niniejszym artykule. Dla przykładu w polu *dc:coverage* spotkać można opis dokumentu, który powinien znajdować się w polu *dc:description* lub słowa kluczowe, które powinny widnieć w polu *dc:keywords*.

W kontekście publikacji M. Werli (2010) warto zasygnalizować, że autorzy niniejszego artykułu, na podstawie analizy metadanych z polskich bibliotek cyfrowych, doszli do podobnych dotyczących jakości opisów w polach *Data*.

W przypadku pola *dc:type* utworzenie wzorcowego zbioru haseł opisujących potencjalne prace naukowe wymagało analizy metadanych z wszystkich dostępnych bibliotek cyfrowych. Uwzględniając skończoną liczbę typów dokumentów, które rzeczywiście opisują prace naukowe, jest to sytuacja kuriozalna, ponieważ w różnych bibliotekach różnie nazywane są takie same typy dokumentów.

Ogromna różnorodność i niekonsekwencje w stosowaniu terminów wpisywanych w pola standardu DublinCore w opisach bibliograficznych z polskich bibliotek cyfrowych wynika, według oceny autorów, z faktu, że bibliotekarze samodzielnie definiują zawartość pól opisu bibliograficznego, wypełniając te pola bez użycia słownictwa kontrolowanego. W Polsce nie zaadaptowano do tej pory słowników kontrolowanych na użytek bibliotek cyfrowych i repozytoriów czy archiwów. Powstałe wiele lat temu Centrum Kompetencji Digitalizacji dla bibliotek, działające w Bibliotece Narodowej, nie potrafi skoordynować działań w zakresie wypracowania standardów, choć należy to do jego obowiązków statutowych i dostaje na ten cel środki finansowe. Warto przypomnieć, że do zadań Centrów Kompetencji należą (MKiDN, 2016):

- *wdrażanie zmian technologicznych dotyczących digitalizacji i przechowywania danych cyfrowych;*
- *koordynacja w zakresie gromadzenia i przechowywania zasobów cyfrowych;*
- *edukacja kadr instytucji kultury prowadzących digitalizację;*
- *udostępnienie materiałów zdigitalizowanych;*
- *wypracowanie standardów;*
- *promocję zasobów cyfrowych.*

Niestety Centrum nie podjęło dotąd żadnych konkretnych inicjatyw w zakresie koordynacji działań bibliotek cyfrowych.

Na niską jakość metadanych polskich bibliotek cyfrowych ma niewątpliwie wpływ brak konsekwencji w stosowaniu obowiązujących w danej bibliotece standardów. Również dopuszczanie do występowania synonimów na listach słownictwa kontrolowanego jest działaniem niedopuszczalnym. Listy takie powinny obowiązywać w każdej bibliotece oraz podawać terminy jednoznaczne, bez synonimów. Z punktu widzenia automatycznego przetwarzania i analizy danych istotne jest, aby dane reprezentujące takie same wartości były w ten sam sposób zapisywane. Można przygotować system automatycznej analizy tak, aby rozpoznawał wszystkie możliwe wariacje typów zapisu danych w poszczególnych polach, ale wymaga to wnikliwej, eksperckiej analizy zbioru metadanych, dodatkowo wykonywanej odrębnie dla każdego osobnego źródła danych, co niepotrzebnie podnosi koszt analiz automatycznych.

Analizy kolejnych pól metadanych obiektów cyfrowych udostępnianych przez polskie biblioteki cyfrowe wiodą do pesymistycznej refleksji, że instytucje odpowiedzialne za budowanie i udostępnianie cyfrowego dziedzictwa wydają się być nieświadome kwestii jakości danych, koncentrując się głównie na istnieniu oraz liczebności tych danych. Niestety, niepoprawne dane, nawet dostępne w największych ilościach, są badawczo nieprzydatne – na ich podstawie nie można metodami analizy automatycznej oraz statystyki wyciągnąć żadnych wiarygodnych, wartościowych wniosków. Oprócz wniosku dotyczącego jakości danych, oczywiście.

Wydaje się, że pracownicy polskich bibliotek cyfrowych nie mają jeszcze świadomości wartości metadanych w erze BigData, a opisy obiektów cyfrowych są przygotowywane nie jako niezależne dokumenty cyfrowe, mające wartość badawczą, ale nadal jako katalogowe reprezentacje dokumentów, jak ma to miejsce w tradycyjnych bibliotekach.

Sytuację ratuje fakt bardzo powszechnego stosowania systemu dLibra¹⁹, który promuje wykorzystanie standardu Dublin Core oraz zapewnia jednolity zapis plików z metadanymi, co znacznie ułatwia ich automatyczne przetwarzanie i analizy.

7. Podsumowanie

Autorzy niniejszego artykułu zdają sobie sprawę z tego, że wnioski wyciągnięte z dotychczasowych badań nie są zadowalające. Poszukują więc nadal metod i narzędzi, które mogą poszerzyć zakres analiz i wnioskowania oraz wnieść nowe rozwiązania. Dotychczasowe analizy pokazują wielorakość problemów wynikającą z jakości pozyskanych danych, problemy te rozwiązywane są przy pomocy metod automatycznych albo za pomocą eksperckich analiz jakościowych. W celu przeprowadzenia automatycznych analiz na dużych zbiorach danych, oprócz stosowania metod autorskich, podjęto współpracę z twórcami projektu CLARIN.PL, m.in. w zakresie automatycznego klasyfikowania dokumentów z zakresu nauk humanistyczno-społecznych na podstawie ich opisów bibliograficznych (CLARIN PL, 2016).

Dotychczasowe analizy metadanych polskich bibliotek cyfrowych wykazały, że twórcom tych bibliotek brakuje jeszcze świadomości wagi poprawnych opisów oraz wymagań dla ich

¹⁹ <http://dlibra.psnc.pl/>

przydatności. Na jakości danych z bibliotek cyfrowych zaważył również brak wypracowanych metod i standardów opisu dokumentów, powodujący dowolność i niekonsekwencję w opisach nawet w ramach jednej biblioteki. Stan bieżący oraz problemy z automatycznych wskazaniem, na podstawie opisu, dokumentów mogących być pracami naukowymi dał autorom impuls do rozpoczęcia prac nad przygotowaniem podręcznego słownika typów dokumentów na użytek bibliotek cyfrowych ale z uwzględnieniem potrzeb badawczych, m.in. naukometrycznych.

Obiecującym rozwinięciem metody grupowania dokumentów na podstawie ich typów jest powiązanie dokumentu z afiliacją autora, jako wspólna wskazówka, że mamy do czynienia z pracą naukową. Badania będą kontynuowane w kierunku usprawnienia metod identyfikacji dokumentów naukowych, a wśród nich praz z zakresu nauk humanistycznych i społecznych. Prace przygotowawcze w tym kierunku zostały już poczynione. Kolejnym etapem są analizy danych z repozytoriów naukowych, dla których z większą dozą pewności można założyć, że gromadzą wyłącznie prace naukowe.

Opracowania przytoczone w rozdziale 1.2 dotyczyły głównie okresu 1995–2011, podczas gdy niniejszy artykuł prezentuje wyniki badań na materiale współczesnym, publikacjach i ich metadanych z lat 2012–2014. W tym okresie można zaobserwować znaczący przyrost dokumentów cyfrowych dostępnych za pośrednictwem FBC: ok. 396 000 do ok. 1320000. Znamienny jest fakt, że pomimo podnoszenia kwestii wagi poprawności metadanych w dostępie do publikacji cyfrowych nadal trzeba wysuwać te same postulaty.

Źródła finansowania:

Badania przeprowadzono w ramach grantu NCN 2013/11/B/HS2/03048.

Badanie przeprowadzono przy współpracy z realizatorami projektu Polska część infrastruktury naukowej CLARIN ERIC CLARIN PL.

Bibliografia

- ANSI/NISO Z39.85–2012 The Dublin Core Metadata Element Set [online] NISO, Baltimore 2013. [05.08.2017], http://www.niso.org/apps/group_public/download.php/10256/Z39-85-2012_dublin_core.pdf
- Baca, M., ed. (2016). *Introduction to Metadata*. 3rd ed. [online] Los Angeles: Getty Publications, 2016. [05.08.2017] <http://www.getty.edu/publications/intrometadata>
- Bednarek-Michalska, B. (2014). Prawo autorskie i jego ograniczenia dla polskich bibliotek cyfrowych.. W: A. Puławski (red.) *Znaczenie udostępniania kopii cyfrowych regionalnych zbiorów bibliotecznych w sieci* : materiały z konferencji, Stargard Szczeciński, 5 września 2014 r. Stargard Szczeciński, 5. September 2014, 51–73.
- CLARIN PL (2016). *Polska część infrastruktury naukowej CLARIN ERIC*. [online]. CLARIN PL, [05.08.2017], <http://clarin-pl.eu/>
- DCMI Metadata Terms*, [online], 2012 [05.08.2017] <http://dublincore.org/documents/dcmi-terms/>
- DCMI Type Vocabulary*, [online], 2000, [05.08.2017] <http://dublincore.org/documents/2000/07/11/dcmi-type-vocabulary/>
- Derfert-Wolf, L. Jak posługiwać się biblioteką cyfrową? W: H. Hollender (red). *Cyfrowy świat dokumentu – wydawnictwa, biblioteki, muzea, archiwa*. Warszawa, CPI 2011, 188–237.
- Digitalizacja.pl* [online] [05.08.2017] <http://www.digitalizacja.pl/>

- Domagalska, A. (2006). Problemy jakości metaopisów w bibliotekach cyfrowych – II Krajowa Konferencja Naukowa Technologie Przetwarzania Danych [online] [05.08.2017], http://www.cs.put.poznan.pl/kkntpd/tpd_pliki/publikacja/pub/55.pdf
- FBC (2016a). *Moduł analityczny FBC*. [online]. Federacja Bibliotek Cyfrowych, [05.08.2017], <http://fbc.pionier.net.pl/pro/wp-content/plugins/baza-fbc/pivot/>
- FBC (2016b). *Otwarte dane FBC – API*. [online]. Federacja Bibliotek Cyfrowych, [05.08.2017], <http://fbc.pionier.net.pl/pro/wspolpraca/api/>
- Głowacka, E. (2011). Jakość bibliotek cyfrowych – aspekty i kryteria oceny. e-mentor. Dwumiesięcznik Szkoły Głównej Handlowej w Warszawie [online], 2011, 2 (39), [05.08.2017], <http://www.e-mentor.edu.pl/artukul/index/numer/39/id/828>
- Kazan, A., Skubała, E. (2008). Polskie biblioteki cyfrowe na platformie dLibra – zasób w kontekście tworzenia nowoczesnych kolekcji źródeł informacji dla nauk technicznych, W: H. Ganińska, (red.) *Informacja dla nauki a świat zasobów cyfrowych*, Poznań 2008, 21–33.
- Interpretacja schematu Dublin Core wraz z materiałami pomocniczymi dla redaktorów zasobów cyfrowych Biblioteki Cyfrowej Uniwersytetu Wrocławskiego [online] Wrocław 2006 [05.08.2017] http://www.bu.uni.wroc.pl/sites/default/files/images/doc/bc/eporadnik_redaktora_bcuwr.pdf
- ISO 8601. Date and Time Formats. [online]. W3C. [05.08.2017] <https://www.w3.org/TR/NOTE-datetime>
- KPBC (2016). *Opis kolekcji: Materiały dydaktyczne*. [online]. Kujawsko-Pomorska Biblioteka Cyfrowa, [05.08.2017], <http://kpbc.umk.pl/dlibra/collectiondescription?dirids=1>
- Metadane, zagadnienia słowników kontrolowanych* (Kołpanowicz M., red.) [online] Narodowy Instytut Muzealnictwa i Ochrony Zbiorów, 2012. [05.08.2017] http://nimoz.pl/upload/digitalizacja/Raport_Metadane_NIMOZ_2012.pdf
- MKiDN (2016). *Digitalizacja. Działalność Centrów Kompetencji*. [online]. Ministerstwo Kultury i Dziedzictwa Narodowego, [05.08.2017], <http://www.digit.mkidn.gov.pl/pages/zasoby/centra-kompetencji.php>
- Nahotko, M. (2010). *Automatyczne tworzenie metadanych*. Bibliotheca Nostra: śląski kwartalnik naukowy 2/2, 13–31 [online] [05.08.2017], http://bazhum.muzhp.pl/media/files/Bibliotheca_Nostra_slaski_kwartalnik_naukowy/Bibliotheca_Nostra_slaski_kwartalnik_naukowy-r2010-t2-n2/Bibliotheca_Nostra_slaski_kwartalnik_naukowy-r2010-t2-n2-s13-31/Bibliotheca_Nostra_slaski_kwartalnik_naukowy-r2010-t2-n2-s13-31.pdf
- Nahotko, M. (2000). *Metadane* W: EBIBB 6/2000(14) [online] [05.08.2017], <http://www.oss.wroc.pl/biuletyn/ebib14/nahotko.html>
- Osińska, V.; Malak, P. (2016a). Dynamiczne sieci społeczne. W: A. Kwiatkowska, M. Sysło. (red.) *Informatyka w edukacji*. Toruń: UMK, 2016.
- Osińska, V.; Malak, P. (2016b). Maps and Mapping in Scientometrics. W: M. Górska, A. Wendel (red.): *Metody i narzędzia badań piśmiennictwa cyfrowego i jego użytkowników*, Wrocław 2016, 59–73.
- Osińska, V.; Malak, P.; Bednarek-Michalska, B. (2016). Badanie struktury i dynamiki zasobów cyfrowej wiedzy przy pomocy metod wizualizacji – projekt realizowany na UMK W: R. Bomba, A. Radomski, E. Solska (red.) *Humanistyka Cyfrowa. Badanie tekstów, obrazów i dźwięku*. Lublin e-naukowiec.eu. 8–18, [online] [05.08.2017], http://e-naukowiec.eu/wp-content/uploads/2016/05/Humanistyka_cyfrowa.pdf
- Płoszajski, G., red. (2008). Standardy w procesie digitalizacji obiektów dziedzictwa kulturowego, [online] Biblioteka Główna Politechniki Warszawskiej [05.08.2017] http://bcpw.bg.pw.edu.pl/Content/1262/BG_Stand_w_proc_digit.pdf
- PN (1992). PN-92-N-01227 — *Bibliotekarstwo i bibliografia. Typologia dokumentów. Terminologia*. Warszawa PKN.
- PN-ISO 15836:2006 Informacja i dokumentacja – Zestaw elementów metadanych Dublin Core (2012). Warszawa PKN.

- Potęga, J. (2009). Metadane w polskich bibliotekach cyfrowych W: *Cyfrowość bibliotek i archiwów*. Warszawa, 26–27 listopada 2009 [online]. Biblioteka Narodowa [05.08.2017], <http://www.bn.org.pl/download/document/1260454699.ppt>
- Werla, M. (2010). Wykorzystanie metadanych z polskich bibliotek cyfrowych [online]. W: C. Mazurek, M. Stroiński, J. Węglarz (red.). *Polskie Biblioteki Cyfrowe 2010. Materiały z konferencji zorganizowanej w dniach 20–21 października 2010 roku przez: Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe*. Poznańskie Centrum Superkomputerowo-Sieciowe, Poznań 2011, 125–129 [05.08.2017] <http://lib.psnk.pl/Content/376/BC-22-Werla.pdf>
- Woodward, E. *Metadata for image collection. Inverse proportions: the quantity vs. quality conundrum* [online] American Libraries Magazine – July 21, 2014 [05.08.2017] www.americanlibrariesmagazine.org/article/metadata-image-collections

Analyzing the Metadata Quality in Polish Digital Libraries

Abstract

Purpose/Thesis: The authors discussed the research on the quality of metadata in Polish digital libraries and the evaluation of the possibility to use metadata in automatic processing of digital repositories to detect current scientific documents.

Approach/Methods: The research was conducted in a hybrid way: automatic NLP analysis and expert analysis.

Results and conclusions: The result of the research in question is the identification of the set of failures and misstatements in the metadata of digital objects available in Polish digital libraries.

Practical implications: The improvements to the current metadata state were proposed. The results and conclusions of this research may contribute to the improvement of the quality of metadata and open the possibility of using metadata as valuable research data.

Originality/Value: To the authors' best knowledge, there is not any similar research in Poland as for the scope and scale.

Keywords

Digital libraries. Dublin Core. Metadata analysis. Metadata quality.

dr PIOTR MALAK – adiunkt w Instytucie Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Wrocławskiego. Wcześniej, w latach 2001–2016 pracował jako asystent, a później adiunkt w Instytucie Informacji Naukowej i Bibliologii UMK w Toruniu. Doktor nauk humanistycznych w zakresie bibliologii. Jego zainteresowania badawcze dotyczą inżynierii lingwistycznej, zarządzania informacją, wyszukiwania informacji oraz zarządzania czasem i zadaniami. Stypendysta Szwajcarskiego Funduszu Stypendialnego SCIEX. Członek Polskiego Towarzystwa Informatycznego, International Society for Knowledge Organization, recenzent projektów COST. Autor m.in.: Indeksowanie treści. Porównanie skuteczności metod tradycyjnych i automatycznych, Warszawa: Wydawnictwo SBP 2012; Malak P. Problemy lingwistyczne we współczesnej informacji naukowej; Babik W. (red) Nauka o Informacji. Warszawa: Wydaw. SBP 2016, 469–491; Malak P., Pawłowski A.: Ewaluacja skuteczności systemów wyszukiwania informacji. Od eksperymentu Cranfield do laboratoriów TREC i CLEF. Genezą i metody: Toruńskie Studia Bibliologiczne 2015, 8(2), 137–156.

Kontakt z autorem:

piotr.malak@uwr.edu.pl

*Instytut Informacji Naukowej i Bibliotekoznawstwa
Uniwersytet Wrocławski*

pl. Uniwersytecki 9/13
50–137 Wrocław

*dr hab. VESLAVA OSIŃSKA – adiunkt w Instytucie Informacji Naukowej i Bibliologii UMK w Toruniu. Doktor nauk humanistycznych w zakresie bibliologii. Zainteresowania badawcze skupia wokół metod i technik wizualizacji informacji i wizualizacji nauki, a w szczególności analizy dynamiki rozwoju nauki Polskiej. Kierownik grantu badawczego z NCN pt. „Badanie struktury i dynamiki cyfrowych zasobów wiedzy za pomocą metod wizualizacji” (wizualizacja.nauki.umk.pl). Członek Management Committee sieci projektów COST Action TD1201 (knowscape.org). Redaktor portalu www.wizualizacja.informacji.pl. Członek Polskiego Towarzystwa Informatycznego, International Society of Knowledge Organization oraz Stowarzyszenia Naukowców Polaków Litwy. Więcej na stronie: www.umk.pl/~wieo. Autorka: V. Osinska, G. Osinski & B. Kwiatkowska. *Visualization in Learning: Perception, Aesthetics and Pragmatism*. In A. Ursyn (Ed.) *Maximizing Cognitive Learning through Knowledge Visualization*. Hershey, PA: IGI Global 2015, pp. 381–414; *Visual mining czyli eksploracja informacji za pomocą graficznych reprezentacji. Praktyka i Teoria Informacji Naukowo Technicznej* 2013, t. 3; Osinska V. *Prezentacja informacji: Babik W. (red) Nauka o Informacji, SBP, Warszawa 2016, s. 577–598.**

Kontakt z autorką:

wieo@umk.pl

Instytut Informacji Naukowej i Bibliologii
UMK w Toruniu
ul. Bojarskiego 1
87–100 Toruń

*mgr BOŻENA BEDNAREK-MICHALSKA – starszy kustosz dyplomowany, zastępca Dyrektora Biblioteki Uniwersyteckiej w Toruniu ds. Informacji i Innowacji. Pracuje jako ekspert doradzający MNiSW w zakresie nowoczesnych technologii informacyjnych, modeli naukowych open access. Była członkini zespołu ds. digitalizacji przy MKiDN oraz zespołu interdyscyplinarnego do spraw działalności upowszechniających nauk przy MNiSW i zespołu ds. otwartego dostępu do nauki. Interesuje się głównie tendencjami rozwojowymi w bibliotekach akademickich, nowoczesną komunikacją naukową, bibliotekami cyfrowymi, technologiami informacyjnymi, otwartą nauką i nowymi modelami publikowania naukowego oraz prawem autorskim. Zrealizowała wiele grantów i projektów dla Biblioteki Uniwersyteckiej w Toruniu. Działała w polskich i zagranicznych organizacjach pozarządowych promujących otwartość w nauce (EBIB, KOED, EIFL, E-LIS, SPARC EUROPE) i podnoszenie kwalifikacji bibliotekarzy (Stowarzyszenie EBIB, SBP). Od 1998 r. redaktor naczelna Biuletynu EBIB – czasopisma fachowego open access dla specjalistów informacji i bibliotekarzy, a także redaktorka i wiceprzewodnicząca elektronicznego serwisu EBIB dla bibliotekarzy. Autorka: *Ocena jakości bibliotekarskich serwisów informacyjnych udostępnianych w Internecie*. *Ebib* 2, 31, 2002; et. all: *Przewodnik po otwartej nauce*, ICM UW 2009.*

Kontakt z autorką:

Bozena.Bednarek-Michalska@bu.umk.pl
Biblioteka Uniwersytecka, UMK w Toruniu
ul. Gagarina 13
87–100 Toruń

Humanistyka cyfrowa: problemy definiowania pojęcia¹

Joanna Ucińska

*Kolegium Międzywydziałowych Indywidualnych Studiów Humanistycznych
Uniwersytet Warszawski*

Abstrakt

Cel/Teza: Celem artykułu jest próba wskazania kluczowych cech pojęcia *humanistyki cyfrowej* oraz zrozumienie, co naprawdę jest rozumiane przez to pojęcie.

Koncepcja/Metody badań: Autorka analizuje najbardziej charakterystyczne cechy pojawiające się w definicjach *digital humanities* w naukowej literaturze anglojęzycznej oraz zestawia je z cechami najczęściej pojawiającymi się w polskich definicjach pojęcia *humanistyka cyfrowa*.

Wyniki i wnioski: Wynik badań jest negatywny – nie jest możliwe stworzenie jednego zestawu cech HC, ale możliwe jest wskazanie takich właściwości humanistyki cyfrowej, co do których większość badaczy jest zgodna.

Oryginalność/Wartość poznawcza: Największą wartością poznawczą pracy stanowi fakt, że jest to pierwsza systematyczna analiza oraz zestawienie definicji polskich i anglojęzycznych humanistyki cyfrowej.

Słowa kluczowe

Analiza. Definicje. Cyfrowa. Humanistyka. Humanistyka cyfrowa.

Otrzymano: 10 maja 2017. Zrecenzowano: 19 lipca 2017. Zaakceptowano: 4 września 2017.

1. Wstęp

W drugim numerze *Zagadnień Informatyki Naukowej* z 2016 r. ukazał się artykuł Marii Przystek-Samokowej, w którym autorka kreśli obraz pojęcia *humanistyka cyfrowa* (HC) w polskim piśmiennictwie naukowym (rozumianym nietradycyjnie – badaczka analizuje nie tylko artykuły naukowe, ale i teksty popularnonaukowe, zaproszenia na konferencje lub wykłady specjalne, strony internetowe i wpisy na blogach) oraz przedstawia zarys pola semantycznego HC (Przystek-Samokowa, 2016). Przedstawiona w tym artykule klasyfikacja w trafny sposób podsumowuje zróżnicowane sposoby postrzegania *humanistyki cyfrowej* przez polskich naukowców, jednak nie wyczerpuje spektrum problemów ani nie zgłębia podejścia normatywnego ukazanego w polskich tekstach. Poszukując w swoich badaniach uniwersalnych cech definicji *humanistyki cyfrowej*, na podstawie analizy literatury zagranicznej oraz polskiej wyprowadziłam nieco inne wnioski niż konkluzje Marii Przystek-Samokowej,

¹ Artykuł powstał na podstawie pracy magisterskiej autorki pt. *Humanistyka cyfrowa w Polsce. Kierunki badań, ośrodki, projekty*, napisanej pod kierunkiem prof. dr hab. Barbary Sosińskiej-Kalaty na Wydziale Dziennikarstwa, Informatyki i Bibliologii Uniwersytetu Warszawskiego.

dlatego też ze względu na złożoność i wielostronność zagadnienia warto ponownie podjąć temat rozumienia pojęcia/zakresu znaczeniowego określenia *humanistyka cyfrowa*.

Termin *humanistyka cyfrowa* nie został jeszcze jednoznacznie zdefiniowany, mimo że określa się w ten sposób stale rozszerzające się pole badań. Co więcej, sama dyscyplina również nie ma wyraźnie oznaczonych granic ani jasno zarysowanej metodologii; jak zauważa Patrik Svensson w artykule *The digital humanities as humanities project*, charakteryzując ją za to nieustannie toczące się negocjacje oraz ścieranie się tradycji epistemicznych i radykalnie odmiennych wizji rozwoju (Svensson, 2012, 42). Do takiego stanu rzeczy w największej mierze doprowadziła zmiana terminologiczna, która nastąpiła w latach 2004–2006 i na którą kluczowy wpływ miały trzy wydarzenia: publikacja książki *A Companion to Digital Humanities* (2004)², założenie organizacji parasolowej dla branżowych stowarzyszeń naukowych Alliance of Digital Humanities Organization (ADHO) oraz inauguracja Digital Humanities Initiative przez National Endowment for the Humanities, niezależną agencję federalną rządu amerykańskiego (Kirschenbaum, 2010).

Pojęcie *digital humanities* było używane w literaturze naukowej przed 2004 r., ale badania dziś określane mianem humanistyki cyfrowej zaliczono do dziedziny *humanities computing*, która ówczesnie miała już ugruntowaną tradycję, sięgającą lat 50. XX w. (Vanhoutte, 2013). Niemniej jednak przywołana wyżej publikacja, wydana w 2004 r. i zredagowana przez Susan Schreibman, Raya Siemensa oraz Johna Unswortha, doprowadziła do zmiany postrzegania pola badawczego i przyczyniła się do ustanowienia *digital humanities* jako nazwy dla całej dyscypliny (Vanhoutte, 2013). Jak wspomina w swoim eseju Matthew Kirschenbaum, wybór tytułu dla tego zbioru tekstów był wynikiem negocjacji między redaktorami a działem marketingu, który proponował nazwę *A companion to digitized humanities*; John Unsworth zaproponował *A companion to digital humanities*, aby nie ograniczać pola skojarzeń tylko do digitalizacji (Kirschenbaum, 2010, 57).

Status nowej nazwy został ugruntowany rok później, gdy w wyniku połączenia Association for Computers in the Humanities i Association for Literary and Linguistic Computing w 2005 r. powstało Alliance of Digital Humanities Organization (ADHO) – grupa patronacka dla organizacji i stowarzyszeń zajmujących się humanistyką cyfrową. Ostatnim z wymienionych wydarzeń zasadniczych dla ukonstytuowania się określenia *digital humanities* jako nazwy dziedziny było powołanie w 2006 r. Digital Humanities Initiative – inicjatywy National Endowment for the Humanities (NEH), jednego z najważniejszych publicznych podmiotów w Stanach Zjednoczonych udzielających dofinansowania w dziedzinie nauk humanistycznych, która miała na celu przyznawanie grantów na projekty z zakresu humanistyki cyfrowej. Dwa lata później inicjatywa została przekształcona w Office of Digital Humanities – biuro dysponujące stałym budżetem, co podniosło prestiż nowopowstałej dyscypliny (Kirschenbaum, 2010).

Zmiana nazwy zapoczątkowała serię dyskusji i prób zdefiniowania czym jest humanistyka cyfrowa; pojawiły się również inne, choć o wiele mniej liczne refleksje, dotyczące statusu humanistyki cyfrowej jako dziedziny odrębnej od *humanities computing*. Patrik Svensson w pracy *Humanities computing as digital humanities* udowadnia, że transformacja, która się dokonała, ma następstwa nie tylko w dyskursie, ale i na płaszczyznach dyscyplinarnej,

² W styczniu 2016 r. wydawnictwo Wiley opublikowało drugie wydanie tego opracowania pt. *A New Companion to Digital Humanities*.

instytucjonalnej oraz społecznej (w odniesieniu do społeczności akademickiej) (Svensson, 2013, 160). Jego zdaniem nie powinno się utożsamiać tych dwóch pól badawczych: pojęcie *digital humanities* ma szerszy zakres i zawiera w sobie *humanities computing* (Svensson, 2013, 175). Sytuację ponadto komplikuje fakt, że w literaturze przedmiotu wyróżnia się dwie fale humanistyki cyfrowej. *The Digital Humanities Manifesto 2.0* zwraca uwagę na ilościowy charakter pierwszej z nich i dążenie do automatyzacji pracy z tekstem; drugą zaś określa jako jakościową, empiryczną, generatywną oraz skoncentrowaną na interpretacji (DH Manifesto 2.0, 2009). David M. Berry wyodrębnia także trzecią falę humanistyki cyfrowej, skupioną na kodzie komputerowym uwikłanym we wszystkie aspekty kultury i pamięci. Według badacza odmiana ta prowadzi do wskazywania przez nowoczesne technologie nieprawidłowości pojawiających się w badaniach humanistycznych oraz zakwestionowania przyjętych założeń nie wprost, takich jak uznane kanony czy humanizm liberalny (Berry, 2012). Teoria Berry'ego nie jest szeroko przyjmowana, jednak warto ją odnotować w kontekście problemów z określeniem dyscypliny.

Liczba definicji pojęcia *digital humanities* nie została ustalona. Temat nakreślenia ram dyscypliny i opisanie metodologii był podnoszony wielokrotnie, zarówno jako główny problem tekstu, jak i jedynie element jego ramy teoretycznej, o czym wspomina Matthew Kirschenbaum w artykule *What Is Digital Humanities and What's It Doing in English Departments?* (Kirschenbaum, 2010). Co więcej, proces definiowania nie jest ograniczony do tradycyjnego medium, rozumianego w tym przypadku jako publikacja naukowa – ma miejsce również w internetowych kanałach komunikacji naukowej. Od 2009 r. uczestnicy realizowanego na łamach mediów społecznościowych projektu Day of Digital Humanities są zapraszani nie tylko do blogowania, ale i do odpowiadania na pytanie „What is Digital Humanities?” (Terras et al., 2013b, 279). Jason Hepler zarchiwizował 817 takich wypowiedzi z lat 2009–2014, które następnie umieścił w bazie danych połączonej ze stroną internetową <http://whatisdigitalhumanities.com>, na której za każdym razem po odświeżeniu wyświetlana jest inna definicja. Na podstawie odpowiedzi zebranych podczas Day of DH w 2011 r. Fred Gibbs zaproponował własną kategoryzację definicji humanistyki cyfrowej, w której najliczniejsze grupy stanowią „variation on «the application of technology to humanities work»”, „working with digital media or a digital environment” oraz „minimize the difference between DH and humanities” (Gibbs, 2013, 290).

Sytuacja ta stała się na tyle symptomatyczna, że artykuły, w których próbuje się zdefiniować pojęcie *digital humanities* są określane jako *genre pieces* (Kirschenbaum, 2010). Systematyczne podejście do problemu zostało zaprezentowane w książce *Defining Digital Humanities* z 2013 r. – publikacja przedstawia wybór 21 artykułów naukowych i wpisów z blogów podejmujących tematykę definiowania humanistyki cyfrowej. Celem redaktorów nie była jednak próba normatywnego ujęcia zjawiska, ale zaprezentowanie przekroju dyskusji wyznaczających granice postrzegania dyscypliny (Terras et al., 2013a). Podobne ujęcie przedstawione zostało w *Debates in the Digital Humanities*, książce wydanej rok wcześniej, w której problemowi definicji poświęcono jedną z sześciu części, a zawarte w niej artykuły także miały przedstawić różne stanowiska w debatach toczących się wokół definiowania humanistyki cyfrowej.

2. Próba oceny popularności definicji humanistyki cyfrowej najszerzej przyjmowanych w literaturze anglojęzycznej

2.1. Metodologia

Ze względu na brak zgody co do tego w jaki sposób powinna być rozumiana humanistyka cyfrowa oraz w celu ustalenia co w dyskursie naukowym faktycznie oznacza *digital humanities* zdecydowałam się przeanalizować cztery definicje najszerzej przyjmowane w literaturze anglojęzycznej. Za wskaźnik popularności oraz poziomu akceptacji w obrębie dyscypliny uznałam liczbę cytowań artykułu, w którym została zawarta dana definicja – nawet jeśli niektórzy z cytujących ją autorów odnieśli się do niej polemicznie, szeroki odzew świadczy o wpływie, który wywarła. Liczba cytowań została ustalona na podstawie bazy danych wyszukiwarki Google Scholar ze względu na fakt, że baza Scopus nie podaje ich wiarygodnej liczby przy tekstach naukowych z dziedziny humanistyki cyfrowej – m.in. artykuł *What is Digital Humanities and What's It Doing in English Departments* według bazy Scopus był cytowany 2 razy, podczas gdy w Google Scholar widnieje liczba 179 odwołań; analogiczny problem powtarzał się w przypadku pozostałych anglojęzycznych tekstów naukowych, które cytuję. Przyjęłam, że minimalna liczba cytowań, niezbędna do określenia oddziaływania danej publikacji, wynosi 50 – wydaje się, że poniżej tej granicy trudno mówić o szerokim oddźwięku artykułu. Ponieważ jedynie trzy publikacje zawierające definicję humanistyki cyfrowej i zaindeksowane przez Google Scholar spełniły to kryterium, pod uwagę wzięłam także definicję podaną przez czasopismo *Digital Humanities Quarterly*. W analizie nie zostały uwzględnione natomiast artykuły z blogów – chociaż niektóre z nich spotkały się z głośną reakcją środowiska cyfrowych humanistów, to jednak nie istnieją odpowiednie narzędzia, które pozwoliłyby ocenić ten wpływ.

2.2. Definicja z monografii *Digital Humanities*

Najczęściej cytowana definicja humanistyki cyfrowej pochodzi z książki autorstwa Anne Burdick, Johanny Drucker, Petera Lunenfelda, Todda Presnera oraz Jeffreya Schnappa *Digital Humanities*, która ukazała się w 2012 r. nakładem MIT Press. Pozycja ta była cytowana 251 razy; definicja *digital humanities* została umieszczona w ostatnim rozdziale książki, stanowiącym krótki przewodnik po humanistyce cyfrowej. Zdaniem jej autorów,

Digital Humanities refers to new modes of scholarship and institutional units for collaborative, transdisciplinary, and computationally engaged research, teaching, and publication.

Digital Humanities is less a unified field than an array of convergent practices that explore a universe in which print is no longer the primary medium in which knowledge is produced and disseminated (Burdick et al., 2012, 122).

Taką propozycję eksplikacji omawianego określenia można zaliczyć do definicji kontekstowych; ponadto zauważalny jest w niej podział na dwa człony, z których każdy mógłby stanowić oddzielną definicję. Niemniej możliwa jest również spójna interpretacja – części nie wykluczają się wzajemnie.

Pierwszy fragment dotyczy punktu odniesienia humanistyki cyfrowej, tutaj rozumianego jako nowe formy prowadzenia studiów oraz nowe typy jednostek akademickich (i nie tylko), przeznaczonych do prowadzenia interdyscyplinarnych badań we współpracy

z innymi instytucjami, z wykorzystaniem metod obliczeniowych (komputerowych) oraz do nauczania i działalności publikacyjnej. Ogólne wskazanie na nowatorstwo, sugerowane przez zwrot *new modes of*, wydaje się uzasadnione w sytuacji, gdy dyscyplina znajduje się w fazie gwałtownych przemian; ciekawość mogłoby wzbudzić natomiast bardzo konkretne określenie charakteru badań realizowanych w jednostkach badawczych zajmujących się humanistyką cyfrową, który ma być interdyscyplinarny, wykorzystywać nowe technologie i opierać się na współpracy. Być może tak precyzyjne wskazanie podyktowane było koniecznością ustanowienia standardów, które powinny obowiązywać w tej konkretnej działalności naukowej i które w tradycyjnej humanistyce nie są rozpowszechnione. Ostatni element pierwszej części definicji, odwołujący się do działalności edukacyjnej i wydawniczej świadczy o świadomości tego, że dla rozwoju tego nowego modelu humanistyki konieczne jest wychowywanie następných pokoleń humanistów cyfrowych oraz popularyzowanie wyników prowadzonych badań.

W drugim członie następuje doprecyzowanie pojęcia *humanistyki cyfrowej*, którą autorzy *Digital Humanities* definiują jako gamę zbieżnych praktyk umożliwiających badanie uniwersum, w którym druk nie jest już głównym medium tworzenia i upowszechniania wiedzy. Ten fragment jest dosłownym zapożyczeniem ze wspomnianego wyżej *The Digital Humanities Manifesto 2.0* z 2009 roku i w oryginale brzmi w następujący sposób:

Digital Humanities is not a unified field but an array of convergent practices that explore a universe in which: a) print is no longer the exclusive or the normative medium in which knowledge is produced and/or disseminated; instead, print finds itself absorbed into new, multimedia configurations; and b) digital tools, techniques, and media have altered the production and dissemination of knowledge in the arts, human and social sciences (DH Manifesto 2.0, 2009, akapit 9).

Zacytowanie tego passusu nie jest zaskakujące – część autorów *The Digital Humanities Manifesto 2.0* jest związana z Uniwersytetem Kalifornijskim, podobnie jak niektórzy autorzy omawianej książki, istnieje więc pewne prawdopodobieństwo, że któryś z nich przyczynił się do powstania manifestu. Definicja z *Digital Humanities* jest okrojona z jednego fragmentu względem pierwowzoru, ponadto ma złagodzoną formę, o czym świadczy zastosowanie innych spójników – w oryginale jest to „[Digital Humanities] is not...but” (Burdick et al., 2012, 122), a w zapożyczeniu „[Digital Humanities] is less...than” (DH Manifesto 2.0, 2009, akapit 9). Mimo zauważalnych zmian, w obu przypadkach ten pierwszy człon jest ważny – wskazuje na to, czym według autorów nie jest humanistyka cyfrowa, mianowicie nie jest jednolitą dziedziną naukową.

Główną konsekwencją takiego rozróżnienia jest zaprzeczenie statusu humanistyki cyfrowej jako nauki tradycyjnej, z czym wiąże się otwarty charakter środowiska cyfrowych humanistów – mogą nimi być również naukowcy, którzy dotychczas nie afiliowali się w ten sposób; w tym sensie zniesiony zostaje warunek instytucjonalizacji, który w naukach tradycyjnych uznaje się za jedno z tzw. zewnętrznych kryteriów oceny stopnia autonomizacji dyscypliny naukowej (por. Kamiński, 1998, 254–257). Mogłoby się wydawać, że drugi fragment definicji staje przez to w sprzeczności z pierwszym, w którym wspomina się o nowych instytucjach, ale prawdopodobnie należy rozumieć tę myśl szerzej: o ile nowe formy prowadzenia studiów mogą wymagać tworzenia nowych jednostek akademickich, o tyle instytucjonalizacja i określanie humanistyki cyfrowej mianem dziedziny czy dyscypliny naukowej może prowadzić do zamknięcia jej na międzyobszarowość, nowości i prowadzić do skostnienia, co byłoby zaprzeczeniem tego, w jaki sposób autorzy *Digital Humanities* ją pojmują.

W kontekście instytucjonalizacji oraz wyznaczania granic dziedzin naukowych warto zwrócić uwagę na rozumienie drugiego członu w terminie *digital humanities* – w obu przytoczonych fragmentach „humanistyka” w odniesieniu do obszaru nauk humanistycznych jest pojmowana dosyć szeroko. Burdick, Drucker, Lunenfeld, Presner oraz Schnapp w swojej definicji nie piszą o bezpośrednim powiązaniu *digital humanities* z humanistyką – można uznać, że założenie takiego związku zostało przyjęte *implicite*, niemniej w samej definicji mówi się jedynie o tym, że HC dotyczy badań naukowych, w których wykorzystywane są nowoczesne technologie. Może to prowadzić do błędnych wniosków – np. w obszarze nauk ścisłych cyfrowe instrumentarium badawcze jest wykorzystywane od dawna, jednak nie sposób określić tych badań mianem *digital humanities*.

W drugim cytowanym fragmencie humanistyka cyfrowa odnosi się nie tylko do badań związanych z tradycyjnie rozumianym obszarem problemowym humanistyki, ale również do dziedziny nauk społecznych oraz do domeny sztuki. Takie ujęcie znacznie rozszerza rozumienie terminu HC oraz również może prowadzić do niewłaściwych wniosków.

2.3. Definicja Matthew Kirschenbauma

Drugą, najliczniej cytowaną publikacją, dotyczącą definiowania pojęcia *digital humanities* jest wspomniany wyżej artykuł Matthew Kirschenbauma *What Is Digital Humanities and What's It Doing in English Departments?*. Opublikowany pierwotnie w 150 numerze czasopisma *ADE Bulletin* w 2010 r., tekst był kilkakrotnie przedrukowywany (m.in. w *Debates in the Digital Humanities* oraz *Defining Digital Humanities: A Reader*) oraz doczekał się 179 cytowań. Poza przywołaniem definicji innych badaczy oraz namysłem nad dziedziną jako taką, Kirschenbaum wysuwa własną propozycję:

Whatever else it might be then, the digital humanities today is about a scholarship (and a pedagogy) that is publicly visible in ways to which we are generally unaccustomed, a scholarship and pedagogy that are bound up with infrastructure in ways that are deeper and more explicit than we are generally accustomed to, a scholarship and pedagogy that are collaborative and depend on networks of people and that live an active 24/7 life online (Kirschenbaum, 2010, 60).

Choć ta wypowiedź nie jest definicją w ścisłym znaczeniu, to badacz ukazuje w niej własny sposób rozumienia humanistyki cyfrowej, co ma szczególne znaczenie ze względu na szeroki oddźwięk wywołany przez publikację. W zacytowanym fragmencie można wyróżnić trzy podstawowe płaszczyzny – przedmiotową, relacyjną oraz jakościową.

Pierwsza z nich dotyczy tego, co należy uznać za główny obiekt w wypowiedzi Kirschenbauma. Są to „a scholarship and a pedagogy”, które można przetłumaczyć jako „studia i sposoby przekazywania wiedzy” (przekład na podstawie Oxford Dictionary oraz Nowego Słownika Fundacji Kościuszkowskiej). Badacz powtarza je konsekwentnie w tym zestawieniu, przez co można rozumieć, że odnoszą się także do czegoś bardziej ogólnego, czyli nauki, a co za tym idzie – humanistyce cyfrowej zostaje przyznany status osobnej dyscypliny. Nie oznacza to jednak, że *scholarship* można przetłumaczyć tu bezpośrednio na *dyscyplinę* – takie odczytanie wynika z celowości zastosowania tego terminu przez autora.

Płaszczyzna relacyjna odnosi się do grona obiorców, którym dla autora *What Is Digital Humanities and What's It doing in English Departments?* jest społeczność naukowa, do której zaliczają się również osoby nie prowadzące profesjonalnie badań, a „jedynie” zainteresowani nauką. Odbiorcy nauki są nieprzyzwyczajeni do nowych form komunikacji,

sposobów, w jakie humanistyka cyfrowa staje się widoczna dla większej publiczności – choć nie jest to wyrażone wprost, można przypuszczać, że Kirschenbaum miał na myśli m.in. blogi naukowe i Twittera – oraz że wiąże się ona z infrastrukturą (którą w tym kontekście można rozumieć również jako technologię) bardziej wszechstronnie i wnikliwie, niż wynikałoby to ze stereotypu badacza-humanisty, którego większość pracy polega na studiowaniu tekstu w odosobnieniu. Na tej płaszczyźnie uwidacznia się również związek innego typu – studia i sposoby przekazywania wiedzy są zależne od grup osób często korzystających z Internetu oraz aktywnych w mediach społecznościowych, domyślnie: społeczności humanistów cyfrowych.

Sfera jakościowa obejmuje cechy przypisywane humanistyce cyfrowej przez Kirschenbauma; nawet jeśli nie wszystkie zostały wyrażone wprost, można je określić na podstawie analizy dwóch poprzednich płaszczyzn. Kluczową właściwością jest niezależność – badacz mówiąc o *digital humanities*, przyznaje dyscyplinie pozycję samodzielnej nauki. Następną ważną cechą to nowatorskość, odkrywczość na tyle duża, by zaskakiwać odbiorców; wiąże się to również z wychodzeniem poza schematy. Cały opisywany aspekt połączony jest również ze wspomnianą widocznością dla audytorium nie zaznajomionego głębiej z humanistyką cyfrową oraz powiązanie z technologią. Kolejną cechą jest otwartość na współpracę oraz pewne uzależnienie od Internetu – cyfrowi humaniści często pracują nad projektami wspólnie, ale w rozproszeniu, za pośrednictwem sieci.

Kirschenbaum w swojej „definicji” buduje obraz humanistyki cyfrowej jako innowacyjnej dyscypliny, w której stosowane są niestandardowe rozwiązania oraz w której szczególnie nacisk kładziony jest na aspekt społecznościowy – rozumiany jako współpraca, praca zespołowa oraz tworzenie społeczności cyfrowych humanistów, komunikujących się ze sobą za pośrednictwem nowych mediów.

2.4. Definicja N. Katherine Hayles

Trzecią pod względem popularności publikacją zawierającą definicję pojęcia *digital humanities* jest artykuł N. Katherine Hayles *How We Think: Transforming Power and Digital Technologies*, który ukazał się w wydanej w 2012 r. książce *Understanding Digital Humanities* i według bazy danych wyszukiwarki Google Scholar był cytowany 96 razy. Badaczka przygląda się w nim następstwom praktyk stosowanych w humanistyce cyfrowej i niektórym ważnym dla niej kwestiom, takim jak skala w kontekście badań ilościowych oraz procedury *close reading*, możliwości wypracowywania teorii czy współpraca. Zależy jej na możliwie szerokim rozumieniu terminu „humanistyka cyfrowa”:

I posit the digital humanities as a diverse field of practices associated with computational techniques and reaching beyond print in its modes of enquiry, research, publication, and dissemination. In this sense, the digital humanities includes text encoding and analysis, digital editions of print works, historical research that re-creates classical architecture in virtual reality formats such as Rome Reborn and The Theater of Pompey, archival and geospatial sites, and since there is a vibrant conversation between scholarly and creative work in this field, electronic literature and digital art that draws on or remediates humanities traditions (Hayles, 2012, 45).

Hayles proponuje definicję zakresową, podzieloną na dwie warstwy – ogólną i szczegółową, z których każda zawiera element wyliczenia. Przeciwnie niż u Kirschenbauma, nie da się tutaj wyróżnić płaszczyzny relacyjnej i jakościowej; w sferze rzeczowej humanistyka

cyfrowa jest nazwana „diverse field of practices”, co można przetłumaczyć jako „różnicowana dziedzina praktyk”. Nacisk nie został położony na samoistność dyscypliny jako takiej, ale na jej aspekt funkcjonalny – po określeniu generalnym następuje doprecyzowanie stosowanych technik („computational techniques”).

W warstwie ogólnej autorka artykułu poza określeniem rzeczowym i powiązaniem dziedziny ze stosowaniem metod obliczeniowych wyszczególnia domeny, w których humanistyka cyfrowa wykracza poza możliwości druku. Należą do nich sposoby: formułowania pytań badawczych, prowadzenia badań, publikowania oraz udostępniania prac. Określenie tożsamości dziedziny zatem dokonuje się tutaj poprzez ustanowienie pewnej opozycji (alternatywy) w stosunku do „tradycyjnej” humanistyki związanej z medium druku; są to zarazem założenia metodologiczne, które nie dotyczą konkretnych postulatów, ale w planie ogólnym sugerują nowoczesność i otwartość.

W warstwie szczegółowej N. Katherine Hayles wymienia, co jej zdaniem zawiera się w humanistyce cyfrowej: kodowanie tekstu i jego analiza, cyfrowe edycje dzieł wydanych wcześniej w druku, badania historyczne odtwarzające klasyczną architekturę w rzeczywistości wirtualnej, archiwalne badania geoprzestrzenne, literatura elektroniczna oraz sztuka cyfrowa czerpiąca z tradycji nauk humanistycznych. Mimo deklaracji dotyczącej chęci włączenia możliwie szerokiego spektrum zjawisk, z dzisiejszej perspektywy zakres praktyk wskazanych przez badaczkę może wydawać się raczej skromny, jednak w momencie ukazania się tekstu był zapewne dosyć nowatorski i odzwierciedlał rozpiętość istotnych ówczasie tematów dla dziedziny.

Pojawia się tutaj również aspekt networkingu – zwrócona została uwaga na korzyści wynikające z połączenia pracy naukowej i kreatywnej („since there is a vibrant conversation between scholarly and creative work in this field”). Ten wymiar nie jest szczególnie eksponowany, jednak już samo uwzględnienie go w definicji jest znaczące i świadczy o uznaniu za integralny element dyscypliny.

Hayles proponuje definicję, która ma być możliwie inkluzywna, choć paradoksalnie niektóre jej elementy wydają się ograniczać tę otwartość. Niemniej jednak sformułowana przez badaczkę propozycja tworzy spójną całość, która po pewnym uaktualnieniu dalej mogłaby służyć do szerokiego opisu humanistyki cyfrowej.

2.5. Definicja *Digital Humanities Quarterly*

Redaktorzy *Digital Humanities Quarterly*, najstarszego czasopisma specjalistycznego z dziedziny humanistyki cyfrowej, zadeklarowali w artykule otwierającym pierwszy numer, że odraczają odpowiedź na pytanie „Czym jest humanistyka cyfrowa?”, ponieważ z czasem ramy dyscypliny zostaną nakreślone przez tematykę publikacji ukazujących się na łamach *DHQ* (Flanders et al., 2007). W zamian postanowili zapytać jak można kształtować humanistykę cyfrową, a dobór tekstów ukazujących się w *Digital Humanities Quarterly* sugeruje, że szukanie odpowiedzi na to pytanie dalej stanowi ich myśl przewodnią.

Na stronie internetowej periodyku *Digital Humanities Quarterly* widnieje jednak definicja pojęcia *digital humanities*, co stoi w sprzeczności z oświadczeniem Julii Flanders, Wendella Pieza i Melissy Terras, ale przybliży problematykę zagadnień poruszanych przez publikowane artykuły. Mając na uwadze fakt, że jest to propozycja uznanego czasopisma, można przyznać jej status normatywny, ponadto jest to jedna z niewielu definicji dziedziny, które można uznać za klasyczne (w rozumieniu systematyki definicji):

Digital humanities is a diverse and still emerging field that encompasses the practice of humanities research in and through information technology, and the exploration of how the humanities may evolve through their engagement with technology, media, and computational methods (Flanders et al., 2007).

W tym ujęciu można wyróżnić dwa istotne elementy, typowe dla definicji klasycznej: *gatunek* i *różnicę gatunkową*. W pierwszej części cytowanego fragmentu – odnoszącej się do *gatunku* – humanistyka cyfrowa została nazwana zróżnicowaną i ciągle rozwijającą się dziedziną, co określa jej położenie względem innych nauk – *Digital Humanities Quarterly* uznaje ją za osobną dyscyplinę. Cechy *zróżnicowana* i *ciągle rozwijająca się* oddają charakter dziedziny, która nie wypracowała jeszcze własnej metodologii ani nawet nie została do końca określona i wciąż znajduje się w fazie intensywnych zmian. Takie przedstawienie humanistyki cyfrowej mogłoby ją deprecjonować, w rzeczywistości jednak działa na jej korzyść – w ramach *młodej* dyscypliny można dokonać więcej, przyczyniając się przy tym do kształtowania jej zakresu.

Druga część definicji doprecyzowuje *różnicę gatunkową*, która obejmuje zarówno praktyki badań humanistycznych wykorzystujące technologie informacyjne, jak i badania nad tym, w jaki sposób humanistyka może rozwijać się poprzez wykorzystanie technologii, mediów i metod obliczeniowych. Pierwszy człon tego fragmentu jest dosyć ogólny – nie zostały wymienione konkretne narzędzia ani sposoby ich użycia, chodzi po prostu o zastosowanie narzędzi informatycznych, co pozwala na określenie mianem *humanistyki cyfrowej* szerokiego spektrum projektów. Drugi człon natomiast tworzy przestrzeń dla metarefleksji o humanistyce – w jaki sposób ją rozwijać przy pomocy nowoczesnych instrumentów.

Taki sposób definiowania osłabia podział na (*tradycyjną*) *humanistykę* i *humanistykę cyfrową*, zmienia perspektywę widzenia tej pierwszej – być może w przyszłości wszystkie badania humanistyczne będą prowadzone przy pomocy metod wykorzystywanych w humanistyce cyfrowej – oraz osadza tę drugą w kontekście szeroko rozumianej tradycji humanistycznej.

2.6. Podsumowanie

Omawiane definicje są dosyć różnorodne – każda z nich inaczej określa humanistykę cyfrową i inny jej aspekt stawia w centrum zainteresowania. Niemniej niektóre właściwości się powtarzają, co umożliwia wskazanie kluczowych cech i ustalenie bazowej definicji humanistyki cyfrowej – a przynajmniej podstawowego sposobu jej pojmowania. Płaszczyzna porównawcza została przedstawiona w tabeli 1.

W tabeli zostały przytoczone cechy humanistyki cyfrowej wymieniane przez każdą z definicji; są uszeregowane według częstotliwości występowania. Poszczególne pozycje mogły być sformułowane w nieco inny sposób, zależnie od definicji, ale na potrzeby tabeli nazewnictwo zostało ujednoczone. Znak + oznacza uwzględnienie danej cechy w definicji.

Jedynym elementem, który pojawia się we wszystkich definicjach, są powiązania z technologią. W propozycjach N. Katherine Halyses oraz *Digital Humanities* zostało wyjaśnione, co należy przez nie rozumieć – autorzy piszą o *computational techniques* (lub *computationally engaged research*), *text encoding* oraz wykorzystaniu rzeczywistości wirtualnej. Definicja z *Digital Humanities Quarterly* również wspomina o *computational methods*, jednak odnosi się to do rozwijania humanistyki w ogóle; w kontekście humanistyki cyfrowej występuje sformułowanie *information technology*. Włączenie tej cechy do wszystkich przytoczonych

prób wyjaśnienia czym jest humanistyka cyfrowa pokazuje, jak duże ma ona znaczenie i że jest konstytutywną częścią składową rozumienia pojęcia *digital humanities*.

Tabela 1. Porównanie występowania kluczowych cech humanistyki cyfrowej w jej definicjach w literaturze anglojęzycznej

Cecha	<i>Digital Humanities</i>	Kirschenbaum	N. Katherine Hayles	<i>Digital Humanities Quarterly</i>
Powiązania z technologią	+	+	+	+
Status dziedziny		+	+	+
Nowatorstwo	+	+	+	
Współpraca/Networking	+	+	+	
Praktyki badawcze	+		+	+
Sposoby prowadzenia studiów	+	+		
Sposoby przekazywania wiedzy	+	+		
Działalność publikacyjna	+		+	
Wykraczanie poza możliwości druku	+		+	
Niezależność		+		
Otwarcie na innych odbiorców niż naukowcy		+		
Uzależnienie od Internetu		+		
Opozycja do tradycyjnej humanistyki			+	
Ciągły rozwój dyscypliny				+
Metarefleksja				+
Osadzenie humanistyki cyfrowej w ogólnym kontekście humanistyki				+

Do najczęściej pojawiających się cech należy status dziedziny, który odnosi się do tego, czy autor lub autorzy danej definicji przyznają humanistyce cyfrowej samodzielny status ontologiczny jako dyscyplinie. Kluczowe dla tej pozycji pojęcie *field* pojawia się we wszystkich propozycjach, nie we wszystkich jednak w pozytywnym kontekście – redaktorzy *Digital Humanities* nie opierają na nim swojego rozumienia humanistyki cyfrowej, o czym wspominałam w interpretacji. Mimo tego oparcie trzech innych definicji na koncepcie dziedziny jest wystarczająco znaczące – samodzielny status ontologiczny wiąże się z większym prestiżem oraz możliwością funkcjonowania jak pozostałe nauki.

Następnym, również często pojawiającym się określeniem jest nowatorstwo, nieobecne jedynie w propozycji czasopisma *Digital Humanities Quarterly*. Ta cecha nie musi wiązać się z odkrywczością samych badań – jest obecna raczej w sferze metodologii i podejścia badawczego; ze względu na stosowane narzędzia łączy się z technologią, ale nie dotyczy jej w ścisłym sensie, jest czymś ponad nią. Może odwoływać się do samego konceptu

interdyscyplinarności lub korzystania z instrumentów, o których wcześniej nie myślano, że mogą mieć zastosowanie w humanistyce.

Współpraca/networking to cecha, która także pojawiła się w trzech spośród czterech przytoczonych definicji. Odnosi się zarówno do współpracy międzyinstytucjonalnej (jak w propozycji z *Digital_Humanities*), współpracy naukowców (jak u Kirschenbauma), jak i do współpracy branży naukowej i kreatywnej (jak u N. Katherine Hayles). Wskazuje to na charakter samej dziedziny – projekty realizowane w jej ramach opierają się na podziale kompetencji oraz nierzadko tzw. pracy w chmurze; można to uznać za jeden z kluczowych elementów dystynktywnych humanistyki cyfrowej. Trzeba równocześnie pamiętać, że zespolowość badań, networking i komunikacja badaczy za pomocą technologii sieciowych są cechami charakterystycznymi dla całej tzw. Nauki 2.0, a więc są to własności, które łączą HC z Nauką 2.0, odróżniając ją równocześnie od tradycyjnie uprawianej humanistyki.

Ostatnim określeniem, jakie pojawia się w kontekście humanistyki cyfrowej w trzech definicjach równocześnie, są praktyki badawcze, przez które należy rozumieć sposoby działania lub procedury dotyczące prowadzenia badań. *Digital_Humanities* opiera rozumienie humanistyki cyfrowej na tym pojęciu; wprowadzie N. Katherine Hayles oraz *Digital Humanities Quarterly* nie stawiają tego terminu w centrum, ale wciąż czynią z niego ważny element swoich propozycji.

Do cech, które pojawiają się w co najmniej dwóch definicjach, należą: sposoby prowadzenia studiów, sposoby przekazywania wiedzy, działalność publikacyjna oraz wykraczanie poza możliwości druku. Można uznać je za istotne, ale nie kluczowe – fundamentalne dla rozumienia humanistyki cyfrowej są określenia pojawiające się w większości źródeł, w tym przypadku w co najmniej trzech tekstach. Cechy lub aspekty występujące tylko w jednej definicji można odebrać jako wariacje jednostkowe, które nie mają jednak zasadniczego znaczenia w szerszym kontekście.

Z tabeli oraz powyższej analizy wynika, że najwięcej określeń które można uznać za najważniejsze zawiera definicja skonstruowana przez N. Katherine Hayles. Nie musi to oznaczać, że jest najlepsza, niemniej sposób rozumienia autorki charakteryzuje pewna wszechstronność, którą warto podkreślić.

Znamienne jest, że wszystkie opisywane cechy odnoszą się do członu „digital” – żadna z definicji nie odnosi się wprost do tego, czym jest humanistyka. W większości przypadków przyjmowane interpretacje pierwszego członu opisują humanistykę jako pole badań wykorzystujące dotychczas druk lub pismo drukowane; jedynie N. Katherine Hayles odnosi się także do sztuki i architektury. *Humanistyka* w „cyfrowym kontekście” oznaczałaby więc badania zwyczajowo związane z literaturą i innymi dziedzinami bazującymi na medium słowa pisanego lub drukowanego przeniesione do komputerowego (cyfrowego) środowiska. Nie zmienia to jednak kluczowych cech humanistyki cyfrowej przyjmowanych najszerzej w literaturze anglojęzycznej, którymi są: powiązania z technologią, status dziedziny, nowatorstwo oraz współpraca/networking.

3. Definicje humanistyki cyfrowej w literaturze polskiej

Kwestia definicji humanistyki cyfrowej w polskiej literaturze naukowej przedstawia podobny stopień skomplikowania jak w literaturze anglojęzycznej. Polscy naukowcy mierzą

się z przeniesieniem pojęcia *digital humanities* na rodzimy grunt od 2012 r. – pierwszej krajowej konferencji poświęconej problematyce humanistyki cyfrowej „Zwrot cyfrowy w humanistyce” (Bomba & Radomski, 2013). Nie oznacza to, że wcześniej nie były prowadzone w Polsce projekty, które można by zakwalifikować w ten sposób, czego przykładem jest choćby Narodowy Korpus Języka Polskiego (NKJP, 2008), nad którym rozpoczęto prace w 2008 r. Długą już historię mają w Polsce w szczególności badania nad tzw. inżynierią lingwistyczną (zob. Mykowiecka, 2006), niemniej konferencja z 2012 r. sprawiła, że zaakceptowano humanistykę cyfrową jako osobny nurt badań oraz zwrócono uwagę na możliwości, jakie nowoczesne technologie oferują naukom humanistycznym. To wydarzenie można uznać za swoisty punkt zwrotny dla rozwoju HC w Polsce – od tego momentu możemy mówić o świadomych badaniach nad nią, prowadzonych na gruncie rodzimych instytucji, o czym świadczy operowanie nazwą HC zarówno w nazwach podejmowanych projektów badawczych, jak i w nazewnictwie tworzonych jednostek badawczych.

W artykułach powstających w polskim środowisku naukowym daje się zauważyć pewna niechęć do formułowania definicji – wielu badaczy próbuje odpowiedzieć na pytanie czym jest humanistyka cyfrowa, ale większość z nich skupia się na zakreśleniu jej ram, często w możliwie najszerszy sposób, i rzadko kiedy pojawia się konkluzja zawierająca konkretne sformułowanie problemu. Zdarzają się również sytuacje, w których badacz wręcz odmawia sformułowania definicji, np. Andrzej Radomski w artykule *Humanistyka cyfrowa w praktyce – analiza i wizualizacja obrazów* napisał otwarcie:

(...) chciałbym zaznaczyć, że nie będę tutaj definiował pojęcia humanistyki cyfrowej. Nie jest to wynik jakiegoś «lenistwa intelektualnego». Po prostu badacze unikają (jak na razie) takiej procedury – ze względu na rozległość pola badawczego oraz ciągle zmieniające się metody i narzędzia. W tym wypadku to praktyka zakreśla sposób rozumienia i uprawiania humanistyki cyfrowej (Radomski, 2015, 6).

Takie ujęcie sprawy wywołuje wrażenie jakby zdefiniowanie humanistyki cyfrowej nie było w ogóle możliwe. Co więcej, na stronach internetowych polskich instytucji zajmujących się HC zazwyczaj nie ma informacji czym jest humanistyka cyfrowa – tak jest np. w przypadku konsorcjum DARIAH-PL (2015), CHC IBL PAN (2017), Centrum Humanistyki Cyfrowej Instytutu Badań Literackich PAN (CHC IBL PAN, 2017) czy nawet Katolickiego Uniwersytetu Lubelskiego, który w ramach swojej oferty edukacyjnej od roku akademickiego 2016/2017 proponuje kierunek studiów drugiego stopnia *Humanistyka Cyfrowa* (KUL, 2016).

Problem braku definicji humanistyki cyfrowej w polskim dyskursie naukowym został poruszony w 2016 r. w trzech artykułach: wspomnianym *Co to jest humanistyka cyfrowa? Pole semantyczne pojęcia (zarys)* Marii Przystek-Samokowej, *Nowa «respublica litteraria»? Humanistyka cyfrowa jako metaorientacja współczesnych badań humanistycznych* Ewy Solskiej oraz w *Humanistyka cyfrowa a socjologia cyfrowa. Nowy paradygmat badań naukowych* Magdaleny Szpunar. Ostatnia z wymienionych badaczek do wyżej przytoczonych słów Andrzeja Radomskiego odnosi się wprost:

Nie do końca można się zgodzić z Radomskim, który twierdzi, że owo definiowanie utrudniają ciągle zmieniające się narzędzia. Rzeczywiście zgodnie z hipotezą Ogburna warstwa kulturowa nie nadąża za zmianami dokonującymi się w warstwie technologicznej, ale zaznaczyć wyraźnie należy, że metody i narzędzia wcale nie zmieniają się w zawrotnym tempie. Ważniejszą sprawą jest tutaj brak rozumienia i umiejętności analizowania otrzymanych danych (Szpunar, 2016, 358).

Szpunar pisze również o tym, że dalsza rezygnacja z wysiłków określenia humanistyki cyfrowej „skazuje ją na utrwalające się traktowanie jej w kategoriach fanaberii cyfrowych

apologetów” (Szpunar, 2016, 358). Podobnego zdania jest Ewa Solska, według której „w Polsce nie istnieje humanistyka cyfrowa jako programowe przedsięwzięcie” (Solska, 2016, 107). Zdaniem badaczki powinna istnieć debata dotycząca teorii, metodologii i kwestii instytucjonalnych. Obie te wypowiedzi świadczą o tym, jak ważne jest zdefiniowanie terminu „humanistyka cyfrowa” i że brak jasno określonych granic jego zakresu już prowadzi do negatywnych konsekwencji.

Wydawać by się mogło, że ogólna sytuacja nie odbiega tak bardzo od tej, która ma miejsce w świecie literatury anglojęzycznej, niemniej jednak zachodzi kilka istotnych różnic, nawet jeśli podobne wrażenie sprawia chaos związany z brakiem porozumienia między badaczami. Pierwsza z tych różnic dotyczy faktycznego istnienia definicji – w literaturze anglojęzycznej jest ich bardzo dużo, wystarczy chociażby przypomnieć przywoływaną już stronę <http://whatisdigitalhumanities.com>; w przypadku literatury polskiej należy mówić przede wszystkim o „próbach definicji”, nie o definicjach. Następnie w przypadku literatury angielskiej możliwe jest określenie stopnia oddziaływania definicji i przyjęcie wymiernych wyznaczników tego wpływu; w odniesieniu do literatury polskiej jest to bardzo trudne, ponieważ większość artykułów nie jest indeksowana w bazach podających liczbę cytowań. Trzecią, bardziej ulotną różnicą, która jednak wciąż ma istotne znaczenie, jest „podejście normatywne” – naukowcy publikujący w języku angielskim nie obawiają się konkretyzowania swoich poglądów, podczas gdy polscy badacze, jak wspominałam, sprawiają wrażenie uprzedzonych do tworzenia precyzyjnych sformułowań.

3.1. *Metodologia*

Opisane wyżej różnice między podejściem obecnym w literaturze angielskojęzycznej i polskiej obrazują problemy wiążące się z przyjęciem metodologii służącej do zbadania polskich definicji. Podstawowa trudność wiąże się z kwestią wyodrębnienia materiału badawczego – tylko w nielicznych przypadkach za podstawową jednostkę można uznać skonkretyzowaną definicję. W związku z tym zastosowałam metodę podsumowania i przeglądu – sposoby definiowania humanistyki cyfrowej w polskich artykułach zostaną omówione pokrótce, ze wskazaniem najważniejszych cech przypisywanych analizowanemu terminowi przez autora danego tekstu.

Próba badawcza, obejmująca siedem tekstów naukowych (oraz jeden artykuł, który zawiera próbę podsumowania stanowisk polskich badaczy) została pozyskana przez wyszukiwanie proste w bazach CEEOL oraz BazHum, jak również w wyszukiwarce Google, na podstawie terminu „humanistyka cyfrowa”. Następnie w wyniku selekcji teksty zostały ograniczone do tych, w których autor podejmując problematykę humanistyki cyfrowej, stara się zaproponować własną definicję lub przedstawia swoje rozumienie tego terminu, co pozwoliło wykluczyć artykuły zawierające jedynie tłumaczenia definicji. Jest to dosyć istotny element procedury selekcyjnej, ponieważ polscy badacze często cytują wypowiedzi innych naukowców lub po prostu przekładają definicje z języka angielskiego, co sprawia, że nie można uznać ich za autorskie sformułowania. Podsumowania zostały uszeregowane chronologicznie względem roku ukazania się publikacji; jeśli kilka tekstów zostało wydanych w tym samym roku i nie było możliwe ustalenie, który z nich został opublikowany jako pierwszy, zastosowano porządek alfabetyczny według nazwisk badaczy.

3.2. Radosław Bomba, *Narzędzia cyfrowe jako wyznacznik nowego paradygmatu badań humanistycznych*

Tak jak sugeruje tytuł, tekst Radosława Bomby jest poświęcony narzędziom cyfrowym i ich roli we współczesnej humanistyce. Opisana została w nim również humanistyka cyfrowa – badacz przedstawił ją w dosyć szeroki sposób, nazywając zarazem „nowym paradygmatem badań”, „szczególną perspektywą badawczą na gruncie współczesnej humanistyki” (Bomba, 2013, 57) oraz „nową dyscypliną” (Bomba, 2013, 69). Użycie tylko tych trzech określeń pokazuje, że według R. Bomby granice HC są nieostre na tyle, by jednoznaczne określenie jej statusu ontologicznego było dosyć trudne lub wręcz niemożliwe. Takie podejście staje się jednak bardziej zrozumiałe w kontekście kwestii instytucjonalizacji. Mimo tego, że jej korzenie sięgają m.in. digitalizacji przeprowadzanej przez różnego rodzaju instytucje, humanistyka cyfrowa jest

zjawiskiem, które daleko wykracza poza tradycyjne instytucje naukowe, czyli uniwersytety, szkoły, biblioteki, archiwa i muzea (Bomba, 2013, 68),

a jej przyswojenie

wymaga przemyślenia instytucjonalnych ram funkcjonowania humanistyki (Bomba, 2013, 70).

HC wyrasta ponad instytucjonalne ramy i nie może być „tylko” dyscypliną lub nowy podejściem do badań; stając się „czymś więcej”, traci ostrość pozwalającą na zdefiniowanie jej istoty.

Do kluczowych cech humanistyki cyfrowej, według tego badacza, należy oparcie na narzędziach cyfrowych, które mogą ewoluować na jej gruncie; szczególną rolę wśród nich odgrywa wizualizacja. Wiąże się to jednocześnie ze stawianiem badaczom pewnych wymagań – muszą mieć odpowiednie kompetencje, pozwalające na wykorzystywanie i tworzenie oprogramowania używanego w badaniach (Bomba, 2013). Innymi aspektami, na które Bomba zwraca uwagę są: popularność modelu autorstwa zbiorowego, korzystanie z praktyk nawiązujących do norm otwartej nauki, takich jak udostępnianie zasobów i danych w Internecie, dopuszczenie użytkowników do oddolnej produkcji oraz możliwość rejestrowania zachowań ludzkich na szeroką skalę (Bomba, 2013).

Mimo że autor poruszył w swoim artykule wiele wątków dotyczących omawianej tematyki, kluczowe było skoncentrowanie uwagi tylko na treściach związanych z definiowaniem humanistyki cyfrowej. Tym, co najbardziej zwraca uwagę w postawie demonstrowanej przez autora jest niemożność zaproponowania jednoznacznej definicji HC oraz dosyć stanowczy stosunek względem zagadnień instytucjonalnych; pozostałe właściwości przypisywane przez niego HC nie odbiegają zbytnio od tych, które pojawiały się w artykułach anglojęzycznych.

3.3. Radosław Bomba, Andrzej Radomski «*Wstęp. Zwrot cyfrowy w humanistyce*»

Andrzej Radomski i Radosław Bomba we *Wstępie* do pracy zbiorowej *Zwrot cyfrowy w humanistyce. Internet / Nowe media / Kultura 2.0*, z której pochodzi kilka cytowanych w niniejszej pracy artykułów, nie tylko nakreślają ogólną problematykę tomu, ale i przedstawiają własne rozumienie terminu „humanistyka cyfrowa”. Według tych autorów jest to

idea połączenia refleksji humanistycznej z nowymi narzędziami cyfrowymi, które wykorzystać można do przetwarzania, wizualizowania, prezentowania i popularyzowania wyników badań naukowych (Bomba & Radomski, 2013, 7).

W kontekście charakteryzowania HC używają oni również określenia „nurt”, co ma swoje konsekwencje, które opiszę później.

W przytoczonej definicji kluczową rolę odgrywają trzy komponenty: refleksja humanistyczna, narzędzia cyfrowe oraz popularyzowanie wyników badań. Pierwszy z nich sytuuje HC w generalnym kontekście humanistyki, wskazując, że humanistyka cyfrowa i humanistyka „tradycyjna” nie są bytami oderwanymi od siebie. Narzędzia cyfrowe to powtarzający się element również definicji anglojęzycznych; w rozumieniu Bomby i Radomskiego stanowią fundament humanistyki cyfrowej. Mimo pewnej oczywistości w tym kontekście, zwraca uwagę położenie nacisku na wizualizację, która jest uważana za kluczową wśród dostępnych instrumentów. Popularyzowanie wyników badań przy użyciu narzędzi cyfrowych kieruje zaś uwagę w stronę odbiorcy, zyskującego w takim ujęciu na znaczeniu; być może nie musi on nawet należeć do społeczności naukowej.

Wszystkie te pojęcia są zogniskowane wokół *nurtu* – humanistyka cyfrowa nie jest zatem dyscypliną, ale wiązką opisanych wyżej tendencji, co nadaje jej uniwersalny charakter. Unika się w ten sposób zamknięcia w ramach konkretnych instytucji i poniekąd zaprasza badaczy z całego spektrum humanistycznych dziedzin – HC mogłaby być użyteczna dla każdego z nich.

3.4. *Anna Nacher, «Poza cyfrowość w zwrocie cyfrowym – od humanistyki cyfrowej do spekulatywnej komputacji»*

Anna Nacher w swoim artykule porusza szereg problemów związanych z humanistyką cyfrową: jej historią, przekształceniami czy związkami z pokrewnymi dziedzinami; niewiele pisze jednak o HC „samej w sobie”. Zwraca uwagę na brak krytycznego podejścia do tego zjawiska; wśród konsekwencji, które się z tym wiążą, wymienia ryzyko sprowadzenia humanistyki cyfrowej do wykorzystywania multimediiów w praktyce akademickiej, wizualizacji danych i modelowania, digitalizacji archiwów, projektów związanych z wykorzystaniem big data oraz zarządzania informacją (Nacher, 2013). Choć wyraźnie zaznacza, że humanistyka cyfrowa to nie tylko te instrumenty, to jednak fakt wymienienia właśnie ich można uznać za pewien najbardziej podstawowy sposób rozumienia humanistyki cyfrowej. Inne zastrzeżenia względem HC to „predylekcja do wyjaśnień monokausalnych (...) oraz fetyszycacja jednorodności metodologicznej” (Nacher, 2013, 85), jak również spostrzeżenie, że wnioski z badań prowadzonych pod szyldem humanistyki cyfrowej mają często charakter tautologii (Nacher, 2013).

Badaczka ta zauważa jednak, że HC może być postrzegana jako szansa dla humanistyki na podniesienie jej statusu; zdaniem Nacher,

największą wartością humanistyki cyfrowej jest konieczność przemyślenia na nowo ni mniej, ni więcej, tylko istoty refleksji o charakterze humanistycznym (Nacher, 2013, 85).

Wśród opisywanych cech HC znajduje się również wewnętrzna różnorodność.

W podejściu A. Nacher najbardziej widoczny jest sceptycyzm, ale nie ma on ostatecznego charakteru – badaczka wskazuje na możliwość wykorzystania HC do „promocji”

humanistyki w ogóle. Jej wypowiedź można odczytywać przede wszystkim jako ostrożne przypomnienie, że do żadnego ze zjawisk nie można podchodzić bezkrytycznie – nawet jeśli budzi znaczny entuzjazm wśród naukowców.

3.5. *Maciej Maryl, F5: Odświeżanie filologii*

Maciej Maryl jest badaczem związanym z Instytutem Badań Literackich Polskiej Akademii Nauk, który problematyce definiowania humanistyki cyfrowej poświęcił artykuł wstępny do jednego z numerów *Tekstów Drugich*. Odnosi się w nim do projektu <http://whatisdigitalhumanities.com/> oraz pisze o zwrocie cyfrowym; zasadniczą część jego tekstu stanowią jednak rozważania na temat tego, co należy rozumieć przez HC.

Maryl pisze wprost:

w humanistyce cyfrowej chodzi, jak sądzę, przede wszystkim o sposób uprawiania nauki, a nie o określony zestaw zagadnień badawczych (Maryl, 2014, 12);

stwierdza również, że

[HC] to nie dyscyplina ani kierunek, tylko raczej ruch czy zespół wspólnych wartości, który przyświeca badaczom z różnych dyscyplin humanistycznych (Maryl, 2014, 11).

W artykule została także podjęta próba wskazania głównych założeń humanistyki cyfrowej; jako pierwsze z nich Maryl wymienia narzędzia, które ze względu na aplikacyjny charakter HC

stają się (...) infrastrukturą badawczą wymagającą ciągłych uaktualnień i stałego utrzymywania (Maryl, 2014, 13).

W tym kontekście istotna jest ponadto wizualizacja, którą badacz rozumie jako narzędzie oferujące możliwość analizy i badania zgromadzonych danych oraz jako sposób prezentacji wyników (Maryl, 2014). Dane same w sobie to kolejne z opisywanych założeń:

teksty stają się narzędziami, a informacje w nich zawarte – danymi (Maryl, 2014, 13)

– tak autor artykułu opisuje wkraczanie instrumentów ilościowych do metodologii badań literackich. Następne z analizowanych aspektów to: budowanie interdyscyplinarnych zespołów jako sposób prowadzenia badań, otwarcie na nowe formy komunikacji naukowej, działania prowadzone w duchu otwartej nauki, nadanie badaczowi nowej roli oraz poszerzenie i zmiana roli odbiorców rezultatów badań naukowych (Maryl, 2014).

Chociaż w ujęciu Maryla humanistyka cyfrowa jest uznawana za nowy sposób prowadzenia badań, a nie za osobną naukę, to dzięki takiemu podejściu zyskuje mocne osadzenie w ogólnym dyskursie humanistycznym. Jak pisze Maryl

humanistyka cyfrowa to naukowy styl życia, który staje się udziałem nas wszystkich (Maryl, 2014, 18)

– można to odczytywać jako prognozę sytuacji, w której tożsamość dziedzinowa, zwłaszcza w obrębie humanistyki, ulegnie zatarciu i każdy naukowiec, zamiast określać się jako literaturoznawca, antropolog czy kulturoznawca będzie po prostu cyfrowym humanistą. Cechy HC wymienione przez M. Maryla wynikają z takiej optyki i wszystkie są wzajemnie powiązane.

3.6. *Barbara Sosińska-Kalata, «Wstęp do pracy zbiorowej «Nauka o informacji w okresie zmian : informatologia i humanistyka cyfrowa»*

W książce *Nauka o informacji w okresie zmian: informatologia i humanistyka cyfrowa* zostały zebrane artykuły stanowiące pokłosie konferencji o tym samym tytule; ich autorzy jednak zrezygnowali z prób zdefiniowania humanistyki cyfrowej. Swoje podejście zaproponowała tylko Barbara Sosińska-Kalata we *Wstępie* do omawianej publikacji:

W badaniach humanistycznych proces ten [cyfryzowania warsztatu badawczego – przyp. aut.] nabiera przyspieszenia dopiero w ostatnich kilkunastu latach, prowadząc do coraz bardziej dynamicznego kształtowania się dziedziny określanej jako humanistyka cyfrowa, czy też e-humanistyka. Najogólniej mówiąc, humanistyka cyfrowa wykorzystuje narzędzia i metody informatyczne do badania człowieka i jego twórczości w środowisku cyfrowym oraz rozwija humanistyczną refleksję dotyczącą technologii cyfrowych i ich oddziaływania na kulturę i twórczość człowieka (Sosińska-Kalata, 2016, 11).

W przytoczonej definicji trzy najważniejsze elementy to uznanie HC za dziedzinę nauki, zwrócenie uwagi na narzędzia informatyczne oraz wzmianka dotycząca refleksji o charakterze humanistycznym. O ile status HC jako dziedziny oraz cyfrowe instrumenty badawcze mogą wydawać się oczywiste w kontekście wyżej omawianych artykułów, o tyle „humanistyczna refleksja dotycząca technologii cyfrowych i ich oddziaływania na kulturę i twórczość człowieka” (Sosińska-Kalata, 2016, 11) nie należy do komponentów definicji HC, które pojawiają się szczególnie często. Jest to o tyle istotne, że cytowany „namysł” lokuje humanistykę cyfrową w szerszej tradycji humanistycznej, jak również nadaje jej – jako dziedzinie – pewną zdolność do autorefleksji.

3.7. *Ewa Solska, «Nowa respublica litteraria? Humanistyka cyfrowa jako metaorientacja współczesnych badań humanistycznych»*

Ewa Solska rozpoczyna swoje rozważania od kwestii tego czym jest humanistyka, osadzając w tym kontekście rozróżnienie na humanistykę cyfrową i humanistykę słowa; swój artykuł poświęca pierwszej z nich. Zdaniem badaczki humanistyka cyfrowa nie jest odrębną dziedziną przedmiotową – do statusu takiej brakuje jej własnej metodologii – ale można ją uznać za „nowy sposób i reguły ich [badań – przyp. JU] uprawiania, którego celem jest adekwatne usytuowanie szeroko pojętych nauk o człowieku w obecnych kontekstach nauki w ogóle” (Solska, 2016, 102). Ze względu na te uwarunkowania,

[HC] pozostaje metonimiczną formułą retoryczno-perswazyjną, za pomocą której postulatorzy dokonują obiektywizacji zdarzenia zwanego zwrotem cyfrowym w humanistycę, sugerując tym samym, że mamy do czynienia z nowym programem dla nauk humanistycznych (Solska, 2016, 108).

– ten komentarz zbliża tekst Solskiej do sceptycznego tonu A. Nacher; można uznać to za przestrożę przed szerokim etykietowaniem różnych projektów mianem *humanistyki cyfrowej*.

Ewa Solska postuluje określanie HC

w kategoriach metaorientacji jako rozszerzenia pola badań humanistycznych w stronę cyberprze-strzeni i sieci intermedialnych (Solska, 2016, 108).

W artykule pojawia się również odwołanie do instrumentów badawczych – w opinii autorki humanistyka cyfrowa zaspokaja potrzebę narzędzi, dzięki którym można badać

to, czego poznanie nie było możliwe wcześniej przez brak odpowiednich metod, zwłaszcza ilościowych; do ich grona zalicza metody analizy statystycznej, metody analizy w badaniach filologicznych oraz w *game studies*, a także wizualizację (Solska, 2016).

Podsumowując, w rozumieniu autorki tekstu *Nowa «respublica litteraria»? Humanistyka cyfrowa jako metaorientacja współczesnych badań humanistycznych*, kluczowe wyrażenia dotyczące humanistyki cyfrowej to: formuła retoryczno-perswazyjna, metaorientacja i narzędzia. Badaczka w zdecydowany sposób opowiada się przeciw nadawaniu HC osobnego statusu ontologicznego, sytuuje ją jednak w kontekście nauk i refleksji humanistycznej.

3.8. *Magdalena Szpunar, «Humanistyka cyfrowa a socjologia cyfrowa. Nowy paradygmat badań naukowych»*

Magdalena Szpunar w artykule *Humanistyka cyfrowa a socjologia cyfrowa. Nowy paradygmat badań naukowych* porusza spektrum zagadnień związanych z humanistyką cyfrową – pisze m.in. o jej obecnej sytuacji, nawiązuje do zwrotu cyfrowego, analizuje wypowiedzi innych badaczy oraz omawia kwestię socjologii cyfrowej. Nie proponuje własnej definicji HC, choć w tekście sugeruje jak można by ją określić w czasach masowego ucyfrowienia. Sama określa humanistykę cyfrową jako „paradygmat w nauce” (Szpunar, 2016, 356); wśród jej cech wskazuje

otwartość i gotowość do dzielenia się wiedzą, a także nieustanne uczenie się – także od innych, niekoniecznie profesjonalistów i specjalistów w danej dziedzinie (Szpunar, 2016, 364).

Pojawia się więc tutaj otwarcie społeczności naukowców na osoby spoza świata nauki, co ma dosyć doniosłe znaczenie – oznacza odejście od elitaryzmu. Autorka wspomina również o deinstytucjonalizacji i pracy zespołowej, zaznacza też jednak, że „humanistyka cyfrowa nie jest tożsama z otwartą nauką” (Szpunar, 2016, 365), co pozwala przyjąć pewne ramy ograniczające.

3.9. *Podsumowanie*

Obraz, jaki wyłania się z przeanalizowanych tekstów, jest dosyć niejednorodny; można wyróżnić tylko kilka słów-kluczy, które pojawiają się jednocześnie w kilku artykułach. Tabelaryczne zestawienie cech HC wyodrębnianych przez polskich autorów, jak to uczyniono w przypadku literatury anglojęzycznej, byłoby więc mało użyteczne ze względu na duże rozproszenie używanych określeń, jednak możliwe jest stworzenie pewnego podsumowania stanowisk prezentowanych przez polskich badaczy.

W każdym z omawianych tekstów pojawia się nawiązanie do narzędzi humanistyki cyfrowej. Niektórzy autorzy szczegółowo je dyskutują, inni jedynie o nich wspominają, niemniej dla wszystkich są one istotne. Należy je interpretować jako kluczowy komponent humanistyki cyfrowej – w rozumieniu polskich naukowców konstituują one jej byt. Najczęściej wymienianym narzędziem jest wizualizacja – została wyszczególniona aż w czterech artykułach.

Większość badaczy jest zgodna, że humanistyka cyfrowa nie jest dyscypliną naukową – takie stwierdzenie, nawet jeśli wyrażone nie wprost, pojawia się w czterech z siedmiu omawianych tekstów: Radosław Bomba i Andrzej Radomski piszą o *nurcie*, Ewa Solska

o *metaorientacji*, Magdalena Szpunar o *paradygmacie w nauce*, a Maciej Maryl określa HC przede wszystkim jako *sposób uprawiania nauki*. Pojęcie *paradygmatu* pojawia się także w artykule Bomby, który jednak nazywa HC także *dyscypliną*, podobnie jak Barbara Sosińska-Kalata, która używa terminu *dziedzina*; jedynie Anna Nacher nie odnosi się do tej kwestii. Taka perspektywa przekłada się na mocniejsze usytuowanie humanistyki cyfrowej na gruncie „tradycyjnej” humanistyki – niezależnie od przyjętego stanowiska, wszyscy cytowani naukowcy żywią przekonanie o współzależności obu „humanistyk”.

Następne punkty styczne to otwarcie społeczności naukowców na nową publiczność, nawiązywanie do norm otwartej nauki oraz praca zespołowa – piszą o tym Radosław Bomba, Maciej Maryl oraz Magdalena Szpunar. Pierwsze z wymienionych zagadnień jest związane z przemianą modelu komunikacji naukowej – naukowcy, korzystając z blogów oraz innych platform w obrębie mediów społecznościowych mają możliwość dotarcia do osób spoza świata nauki, które mogą być zainteresowane tematyką prac danego badacza. Korzystanie z otwartych licencji jest z kolei pewną tendencją, która zaczyna być zauważalna w naukach humanistycznych, a praca zespołowa to nierzadko konsekwencja prowadzenia interdyscyplinarnych projektów, do zrealizowania których potrzebne są kompetencje z różnych dziedzin.

Dwie pozostałe kwestie, występujące w więcej niż jednym artykule, to deinstytucjonalizacja oraz refleksja humanistyczna. Problematyka wyjścia poza sztywne instytucjonalne ramy pojawia się w tekstach Radosława Bomby oraz Magdaleny Szpunar. O refleksji humanistycznej piszą z kolei Radosław Bomba i Andrzej Radomski we wspólnym tekście oraz Barbara Sosińska-Kalata; w obu przypadkach jest to namysł nad połączeniem humanistyki i nowych technologii oraz efektami takiego zespolenia.

Trudno byłoby określić przeanalizowane wyżej cechy jako fundamentalne dla polskiego rozumienia *humanistyki cyfrowej*, przede wszystkim ze względu na fakt, że rozpatrywane były nie definicje, a całe artykuły, których autorzy zastanawiali się nad naturą omawianego zjawiska. Niemniej, celem tej części pracy nie było poszukiwanie definicji wspólnej, którą można byłoby uznać za pewną normę, a jedynie przegląd stanowisk autorów podejmujących rozważania o naturze humanistyki cyfrowej.

4. Definiowanie humanistyki cyfrowej w piśmiennictwie anglojęzycznym i polskim

Między sposobami definiowania humanistyki cyfrowej w piśmiennictwie anglojęzycznym i polskim można zauważyć kilka istotnych różnic, o czym wspominałam przy okazji pisania o ogólnej sytuacji normatywnej terminu „humanistyka cyfrowa” w pracach polskich naukowców. Przeprowadzona analiza rodzimych tekstów umożliwia jednak nakreślenie bardziej szczegółowej sytuacji porównawczej.

Najczęściej powtarzany w tekstach polskich naukowców element związany z humanistyką cyfrową – narzędzia cyfrowe – jest zbieżny z najpopularniejszą cechą z artykułów anglojęzycznych, którą są powiązania z technologią. Oba sformułowania można uznać za odnoszące się do tego samego zjawiska, którym jest wprowadzenie instrumentów cyfrowych do humanistyki. Drugą z cech wspólnych stanowi praca zespołowa, istotna zarówno dla naukowców publikujących po polsku, jak i po angielsku. Pozostałe właściwości, które pojawiły się zarówno w literaturze polskiej, jak i w angielskiej – otwarcie społeczności naukowców na

nową publiczność oraz refleksja humanistyczna – są mniej znaczące jako punkty wspólne, ponieważ w literaturze anglojęzycznej wspomniano o nich tylko jednokrotnie.

Większość anglojęzycznych autorów jest zgodna co do przyznania HC statusu dziedziny, natomiast polscy naukowcy w większości mają całkowicie odmienne zdanie – dla nich humanistyka cyfrowa nie jest dyscypliną. Chociaż proponują różne jej interpretacje (m.in. jako paradygmat czy metaorientacja), niezgoda na autonomiczny status ontologiczny wyznacza elementarną różnicę między sposobem pojmowania HC demonstrowanym w piśmiennictwie polskim i tym, który dominuje w piśmiennictwie anglojęzycznym. Inną rozbieżność stanowi fakt, że w polskim piśmiennictwie humanistyka cyfrowa jest osadzona w tradycji humanistycznej; w artykułach anglojęzycznych taka koncepcja pojawia się tylko u jednego z najczęściej cytowanych autorów, a kolejny sytuuje humanistykę cyfrową w opozycji do tradycyjnej. Pozostałe dwie cechy HC, najczęściej pojawiające się w polskich tekstach naukowych, odwoływanie się do standardów otwartej nauki oraz deinstytucjonalizacja, nie pojawiły się w piśmiennictwie anglojęzycznym w ogóle.

Porównanie sposobów definiowania humanistyki cyfrowej przez autorów rodzimych i anglojęzycznych uwidacznia kluczowe podobieństwo i kluczową różnicę. Podobieństwo leży w oparciu definicji na pojęciu narzędzi cyfrowych, które siłą rzeczy wiążą humanistykę z technologią i które są niezbędne do uprawiania humanistyki cyfrowej. Różnica widoczna jest przede wszystkim w podejściu do statusu HC jako dziedziny/dyscypliny: polscy naukowcy w większości nie uznają HC za takową, podczas gdy anglojęzyczni zgadzają się na jej odrębny status ontologiczny. Warto zauważyć również, że polscy autorzy o wiele częściej odwołują się do definicji humanistyki samej w sobie, co może wpływać na szerszy ogląd sytuacji.

Bibliografia

- Berry, D. M. (2012). Introduction: Understanding the Digital Humanities. In: D. M. Berry (eds.), *Understanding Digital Humanities*. Basingstoke: Palgrave Macmillan UK, 1–20.
- Bomba, R. (2013). Narzędzia cyfrowe jako wyznacznik nowego paradygmatu badań humanistycznych. W: R. Bomba, A. Radomski (red.), *Zwrot cyfrowy w humanistyce. Internet / Nowe Media / Kultura 2.0*. Lublin: E-Naukowiec, 57–72.
- Bomba, R.; Radomski, A. (2013). Wstęp. Zwrot cyfrowy w humanistyce. W: R. Bomba, & A. Radomski (red.), *Zwrot cyfrowy w humanistyce. Internet / Nowe Media / Kultura 2.0*. Lublin: E-naukowiec, 7–9.
- Burdick, A.; Drucker, J.; Lunenfeld, P.; Presner, T.; Schnapp, J., eds. (2012). *Digital Humanities*. Cambridge, Massachusetts: The MIT Press.
- CHC IBL PAN (2017). *Centrum Humanistyki Cyfrowej IBL PAN* [online]. Centrum Humanistyki Cyfrowej IBL PAN, [15.03.2017], <http://chc.ibl.waw.pl/pl/>
- DARIAH-PL (2015). *Konsorcjum Dariah-PL* [online] Konsorcjum Dariah-PL, [15.03.2017], <http://dariah.pl>
- DH Manifesto 2.0*. (2009). A Digital Humanities Manifesto [online], A Digital Humanities Manifesto, UCLA [3.03.2017], <http://manifesto.humanities.ucla.edu/2009/05/29/the-digital-humanities-manifesto-20/>
- Digital Humanities Quarterly. *About DHQ* [online]. Digital Humanities Quarterly, [28.02.2017], <http://www.digitalhumanities.org/dhq/about/about.html>
- Flanders, J.; Piez, W.; Terras, M. (2007). Welcome to Digital Humanities Quarterly [online]. *Digital Humanities Quarterly*, March [1.03.2017], <http://www.digitalhumanities.org/dhq/vol/1/1/000007/000007.html>
- Gibbs, F. (2013). Digital Humanities Definitions by Type. In: M. Terras, J. Nyhan, E. Vanhoutte (eds.), *Defining Digital Humanities: A reader*. Farnham; Burlington: Ashgate, 289–298.

- Hayles, N. K. (2012). How We Think: Transforming Power and Digital Technologies. In: D. M. Berry (ed.), *Understanding Digital Humanities*. Basingstoke: Palgrave Macmillan UK, 42–66.
- Kamiński, Stanisław (1998). *Nauka i metoda. Pojęcie nauki i klasyfikacja nauk*. Lublin: Towarzystwo Naukowe KUL.
- Kirschenbaum, M. G. (2010). What Is Digital Humanities and What's It Doing in English Departments? *ADE Bulletin*, (150), 55–61.
- KUL (2016). *Humanistyka cyfrowa* [online]. Katolicki Uniwersytet Lubelski. Serwis dla kandydata, [15.03.2017], http://www.kul.pl/humanistyka-cyfrowa,art_65557.html
- Maryl, M. (2014). F5: odświeżanie filologii. *Teksty Drugie*, 2, 9–20.
- Mykowiecka, A. (2006). *Inżynieria lingwistyczna. Komputerowe przetwarzanie tekstów w języku naturalnym*. Warszawa: PJWSTK.
- Nacher, A. (2013). Poza cyfrowość w zwrocie cyfrowym – od humanistyki cyfrowej do spekulatywnej komputacji. W: R. A. Bomba Radosław (red.), *Zwrot cyfrowy w humanistyce. Internet / Nowe Media / Kultura 2.0*. Lublin: E-naukowiec, 83–97.
- NKJP (2008). *O projekcie NKJP* [online]. Narodowy Korpus Języka Polskiego, IPI PAN; IJP PAN; Wydawnictwo Naukowe PWN; Zakład Językoznawstwa Komputerowego i Korpusowego UŁ [3.03.2017], <http://nkjp.pl/>
- Przastek-Samokowa, M. (2016). Czym jest humanistyka cyfrowa? Pole semantyczne pojęcia (zarys). *Zagadnienia Informacji Naukowej*, 54(2), 82–93.
- Radomski, A. (2015). Humanistyka cyfrowa w praktyce – analiza i wizualizacja obrazów. *Roczniki kulturoznawcze*, 6(4), 5–15.
- Solska, E. (2016). Nowa „respublica litteraria”? Humanistyka cyfrowa jako metaorientacja współczesnych badań humanistycznych. *Roczniki Kulturoznawcze*, 7 (1), 99–118.
- Sosińska-Kalata, B. (2016). Wstęp. W: B. Sosińska-Kalata (red.), *Nauka o informacji w okresie zmian: informatologia i humanistyka cyfrowa*. Warszawa: Wydaw. SBP, 9–13.
- Svensson, P. (2013). Humanities Computing as Digital Humanities. In: M. Terras, J. Nyhan, & E. Vanhoutte (Redaktorzy), *Defining Digital Humanities: A reader*. Farnham; Burlington: Ashgate, 159–186.
- Svensson, P. (2012). The digital humanities as a humanities project. *Arts and Humanities in Higher Education*, 11(1–2), 42–60.
- Szpunar, M. (2016). Humanistyka cyfrowa a socjologia cyfrowa. Nowy paradygmat badań naukowych. *Zarządzanie w Kulturze*, 17(4), 355–369.
- Terras, M.; Nyhan, J.; Vanhoutte, E. (2013a). Introduction. In: M. Terras, J. Nyhan, & E. Vanhoutte (eds.), *Defining Digital Humanities: A reader*. Farnham; Burlington: Ashgate, 1–10.
- Terras, M.; Nyhan, J.; Vanhoutte, E. (2013b). Selected Definitions from the Day of Digital Humanities: 2009–2012. In: M. Terras, J. Nyhan, & E. Vanhoutte (eds.), *Defining Digital Humanities: A Reader*. Farnham; Burlington: Ashgate, 278–288.
- Vanhoutte, E. (2013). The Gates of Hell: History and Definition of Digital | Humanities | Computing. In: T. Melissa, N. Julianne, & V. Edward (eds.), *Defining Digital Humanities: A Reader*. Farnham; Burlington: Ashgate, 119–156.

Digital Humanities: Problems in Defining the Concept

Abstract

Purpose/Thesis: This article attempts to indicate a set of the most common features of digital humanities and to determine what the term DH really means.

Approach/Methods: The author examines English and Polish scientific articles discussing the definitions of digital humanities and juxtaposes them.

Results and conclusions: The result is negative as it is not possible to create one set of digital humanities features, nevertheless it is possible to find some most often occurring features.

Originality/Value: This article provides the first systematic analysis and compilation of Polish and English-language definitions of the digital humanities.

Keywords

Analysis. Definitions. Digital. Humanities. Digital humanities.

JOANNA UCIŃSKA ukończyła studia magisterskie na kierunku Informacja naukowa i bibliotekoznawstwo, specjalizacja: Architektura Informacji i Wiedzy, którą studiowała w ramach studiów w Kolegium Międzywydziałowych Indywidualnych Studiów Humanistycznych i Społecznych na Uniwersytecie Warszawskim. Do zainteresowań badawczych autorki należą: humanistyka cyfrowa, wykorzystanie nowych technologii w naukach humanistycznych, użytkownicy informacji, kompetencje cyfrowe (ang. data literacy), inżynieria wiedzy oraz funkcjonowanie nowych technologii w kontekście społecznym.

Kontakt z autorką:

joanna.ucinska@student.uw.edu.pl

Przekaz symboliczny i podejście User Experience na przykładzie serwisów internetowych teatrów krakowskich¹

Magdalena Zych

Wydział Zarządzania i Komunikacji Społecznej, Uniwersytet Jagielloński

Abstrakt

Cel/Teza: Głównymi celami były opis przekazu symbolicznego oraz ocena User Experience serwisów teatrów krakowskich. Serwisom WWW przypisano możliwość wspomagania realizacji pozainstrumentalnych potrzeb użytkownika.

Koncepcja/Metody badań: Zastosowano metodologię jakościową. Materiał gromadzono za pomocą TAP i techniki inspirowanej procedurą ZMET. W badaniu wzięło udział dziesięciu użytkowników związanych ze środowiskiem teatralnych. Zbadano serwisy dziewięciu teatrów. Materiał poddano analizie tematycznej.

Wyniki i wnioski: Przekaz symboliczny opisano na przykładzie serwisu Teatru Scena STU. Za czynniki wpływające na przekaz symboliczny uznano cechy indywidualne użytkownika, nastawienie emocjonalne i elementy wnoszone przez interakcję użytkownik-serwis. Stwierdzono, że pozytywna ocena użyteczności niekoniecznie przekłada się na oczekiwany przekaz symboliczny.

Oryginalność/wartość poznawcza: Artykuł stanowi przyczynek do badań nad zależnościami między przekazem symbolicznym, przekazem informacyjnym oraz użytecznością serwisów WWW, dotąd nieporuszonymi na gruncie informatologii.

Słowa kluczowe

Funkcja serwisu WWW. Przekaz symboliczny. Serwis WWW teatru. User Experience. Użyteczność.

Otrzymany: 26 czerwca 2017. Zrecenzowany: 7 lipca 2017. Zaakceptowany: 31 sierpnia 2017.

1. Wprowadzenie

Jednym z popularnych, choć wciąż niespójnych, podejść w badaniu i projektowaniu serwisów WWW jest User Experience. Część badaczy kwestionuje nawet możliwość projektowania User Experience, twierdząc, iż można co najwyżej projektować dla User Experience (Bargas-Avila & Hornbæk, 2011, 1; Diefenbach & Hassenzahl, 2011, 461). W tym artykule postanowiono skorzystać z podstawowych osiągnięć UX, nie zamykając się jednak na związany z tym aparat pojęciowy i metodologiczny. W centrum badań ustawiono przekaz symboliczny serwisu WWW. Jest to konsekwencją założenia, że serwisy WWW mogą

¹ Artykuł powstał na podstawie pracy magisterskiej autorki pt. *Przekaz symboliczny i podejście User Experience na przykładzie serwisów internetowych teatrów krakowskich* napisanej w 2016 r. pod kierunkiem dr hab. Marii Próchnickiej w Instytucie Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego. Praca ta uzyskała II nagrodę w konkursie na najlepszą pracę magisterską Nagroda Młodych SBP w okresie obejmującym trzy lata akad.: 2013/14, 2014/15 i 2015/16 [red.].

wspierać realizację poza-instrumentalnych potrzeb użytkownika. Przekaz symboliczny stanowi zagadnienie niezbadane na gruncie informatologii, zaś przeprowadzone badania empiryczne mogą stanowić przyczynek do dalszych badań w kierunku określenia relacji przekazu symbolicznego z użytecznością oraz przekazem informacyjnym.

2. Przedmiot i cel badań

Przedmiotem badań jest podejście User Experience (UX) oraz przekaz symboliczny serwisów internetowych. Badanie miało charakter eksploracyjny i zostało przeprowadzone na przykładzie serwisów WWW teatrów krakowskich.

Głównymi celami badań były: opis przekazu symbolicznego oraz ocena User Experience wybranych serwisów internetowych. W niniejszym artykule User Experience rozumiane jest jako system obejmujący wartości pragmatyczne i hedonistyczne. Przyjęto, że wartości pragmatyczne charakteryzowane są m.in. przez funkcjonalność (ang. *functionality*) oznaczającą zakres dostępnych operacji i poprawność ich działania (Laskowski, 2011, 2192) oraz użyteczność (ang. *usability*), która w tej pracy oznacza łatwość obsługi czegoś i zakładaną możliwość zaspokojenia instrumentalnych potrzeb użytkownika. O wartościach hedonistycznych piszą m.in. Sarah Diefenbach oraz Marc Hassenzahl, których zdaniem idea wartości hedonistycznych wciąż ewoluuje (Diefenbach & Hassenzahl, 2011, 461). W piśmiennictwie naukowym wartości hedonistyczne UX, zwane również wartościami nie-instrumentalnymi, opisywane są za pomocą następujących haseł:

- afilacja/przynależność/interakcje społeczne (Diefenbach & Hassenzahl, 2011, 461; O'Brien, 2010),
- eskapizm (O'Brien, 2010, 3; Wu & Holsapple, 2014, 82),
- estetyka (Bargas-Avila & Hornbæk, 2011, 1; Coursaris et al., 2010, 1; Nawaz, 2013, 49),
- ewokacja – odnosi się do symbolicznego znaczenia oraz potencjału w przywoływaniu wspomnień (Partala & Kallinen, 2012, 26),
- identyfikacja – możliwość ekspresji siebie za pomocą przedmiotów, które się posiada i/lub z których się korzysta (Helfenstein, 2012, 191; Karapanos et al., 2009, 730; Partala & Kallinen, 2012, 25; Wu & Holsapple, 2014, 82),
- piękno (Bargas-Avila & Hornbæk, 2011, 2; Diefenbach & Hassenzahl, 2011, 462; Karapanos et al., 2009, 735),
- przyjemność (Karapanos et al., 2009, 735; O'Brien, 2010; Schuitema et al., 2013, 40),
- radość z użytkowania (Bargas-Avila & Hornbæk, 2011, 2; Schuitema et al., 2013, 40),
- reakcje afektywne/emocje (O'Brien, 2010; Schuitema et al., 2013, 40; Wu & Holsapple, 2014, 82),
- rozwój osobisty (Bargas-Avila & Hornbæk, 2011, 2; Karapanos et al., 2009, 730; Partala & Kallinen, 2012, 25),
- samorealizacja (Bargas-Avila & Hornbæk, 2011, 1; Nawaz, 2013, 49),
- stymulacja, wyzwanie (Bargas-Avila & Hornbæk, 2011, 2; Diefenbach & Hassenzahl, 2011, 461; O'Brien, 2010; Karapanos et al., 2009, 730; Partala & Kallinen, 2012, 25),
- zabawa, rozrywka (Diefenbach & Hassenzahl, 2011, 462; O'Brien, 2010; Coursaris et al., 2010, 1–2; Helfenstein, 2012, 191; Wu & Holsapple, 2014, 82),
- zaskoczenie (Bargas-Avila & Hornbæk, 2011, 2).

Przekaz symboliczny serwisu internetowego rozpatrywany jest w niniejszym artykule bardzo wąsko, tj. jako możliwość oddawania charakteru teatru przez jego serwis WWW. Poza ogólnymi wynikami badań przeprowadzonych na serwisach WWW dziewięciu wybranych krakowskich teatrów, przedstawiono opis przekazu symbolicznego jednego z analizowanych serwisów – Krakowskiego Teatru Scena STU.

3. Stan badań

Analizę i krytykę piśmiennictwa naukowego przeprowadzono na podstawie publikacji w językach polskim lub angielskim, opublikowanych w latach 2009–2015. W uzasadnionych przypadkach włączono też starsze publikacje naukowe, których wartość merytoryczna przydatna była do zaprojektowania procedury badawczej lub czynienia ustaleń terminologicznych. Gromadzone były publikacje dotyczące User Experience, oceny serwisów internetowych, zwłaszcza serwisów teatrów oraz instytucji zbliżonych, jak również publikacje poświęcone komunikacji teatru z interesariuszami. Największą wagę przypisano publikacjom zawierającym wyniki badań empirycznych. Łącznie uwzględniono 81 publikacji.

Kwerendą wyszukiwawczą objęto źródła komercyjne i Open Access. Wśród wykorzystanych źródeł znalazły się m.in.: bibliografie (np. Polska Bibliografia Bibliologiczna (1995-); Przewodnik Bibliograficzny (MARC 21, 1973), komercyjne bazy piśmiennictwa naukowego (np. Academic Search Complete; Library, Information Science & Technology Abstracts), wyszukiwarki (np. Google Scholar; OAIster), archiwa dziedzinowe i repozytoria instytucjonalne (np. E-LIS; Queensland University of Technology ePrints Archive), katalogi czasopism (Directory of Open Access Journals; Portal Czasopism Naukowych), biblioteki cyfrowe (Federacja Bibliotek Cyfrowych) oraz CEON Biblioteka Nauki. Preferowanym trybem dostępu do zasobów było przeszukiwanie. Nie zawsze było ono możliwe ze względu na funkcjonalności dostępne w poszczególnych źródłach. Wyszukiwanie najczęściej prowadzono za pomocą zapytań w formie różnych wariacji składających się z następujących słów kluczowych: „user experience”, „theater website”, „website”, „theatre”, „usability”, „hedonic”, „hedonic quality”, „hedonic value”, „symbolic”, „symbolic quality”, „performing arts”, „evaluation”, „funology”, „branding” oraz ich polskich odpowiedników. Podczas tworzenia zapytań, w miarę oferowanych w źródle możliwości, korzystano z pomocy wyszukiwawczych, takich jak np. słowniki kontrolowanej leksyki oraz łączenie wyrażen/słów kluczowych operatorami logicznymi.

W wyniku analizy piśmiennictwa zidentyfikowano dwie luki. Po pierwsze niewiele jest publikacji poświęconych empirycznym badaniom serwisów WWW teatrów, zwłaszcza prowadzonych z udziałem użytkowników, toteż gromadzono publikacje dotyczące oceny innych serwisów internetowych, przy czym starano się, aby owe serwisy były zbliżone tematycznie do serwisów teatrów lub należały do instytucji podobnych do teatrów. Pod tym względem przeważają badania serwisów muzeów, o których pisali m.in. Jessie Pallud, Detmar W. Straub (Pallud, Straub, 2014), Mohd S. Yamin i Eswati A. Jaafar (Yamin & Jaafar, 2013) oraz Debra A. Riley-Huff (Riley-Huff, 2009). Natomiast za najbardziej wartościową publikację, ze względu na badany serwis internetowy, uznano *The Royal Opera House Website Evaluation Report* (Trosheva, 2013). W projektowaniu własnej procedury badawczej przydatne okazały się publikacje dotyczące komunikacji teatru z widzami, w tym badanie

nawyków konsumenckich związanych z nabywaniem biletów na spektakle oraz rola emocji w tym procesie (Troilo et al., 2014), komunikacja teatru z publicznością za pomocą Internetu (Turrini et al., 2012), strategia pozyskiwania widzów teatralnych z wykorzystaniem narzędzi komunikacji cyfrowej (Kossecki & Świerczyńska-Kaczor, 2014) oraz badanie publiczności i wizerunku teatrów warszawskich (Siwiak et al., 2013).

Druga luka dotyczy przekazu symbolicznego serwisów internetowych, chociaż odszukać można publikacje, których autorzy analizują m.in. serwisy WWW, odwołując się do symbolicznych kategorii. Nie tłumaczą jednak, jak definiują bądź interpretują kategorie symboliczne, którymi się posługują. Nieliczni autorzy, którzy stosowali określenie „symboliczny” w odniesieniu do przedmiotu swych badań, charakteryzowali go za pomocą następujących określeń: znaczenie barw (Hernik, 2012), wizerunek (Helfenstein, 2012, 191), prestiż/duma wynikające z użytkowania marki lub produktu (Helfenstein, 2012, 191; Saucken & Gomez, 2014, 653), a także tożsamość społeczna (Schuitema et al., 2013, 40). Dodatkowo część atrybutów (np. ewokacja, ekspresja siebie, samorealizacja) bywa przypisywana zarówno do sfery symbolicznej, jak i do wartości hedonistycznych UX. W wyniku analizy piśmiennictwa zidentyfikowane trzy publikacje, w których omówione zostały badania układu składającego się z wartości hedonistycznych, pragmatycznych i symbolicznych (Bhui & Ibrahim, 2013; Hernik, 2012; Schuitema et al., 2013). Na ich podstawie nie udało się jednak wyłonić spójnego modelu uwzględniającego UX i przekaz symboliczny.

4. Zastosowana metodologia

W omówionych w artykule badaniach zastosowano metodologię jakościową. Materiał empiryczny gromadzono za pomocą symultanicznej wersji *think-aloud protocol* (TAP) oraz techniki inspirowanej procedurą Zaltman Metaphor Elicitation Technique (ZMET). W badaniu wzięło udział dziesięciu użytkowników związanych ze środowiskiem teatralnym. Byli wśród nich aktorzy, pracownicy teatrów oraz widzowie.

W piśmiennictwie naukowym odnotowuje się trzy warianty TAP: *concurrent* (wersja symultaniczna), *retrospective* (retrospektywna) oraz *constructive interaction*, zwana również *co-discovery learning* (AlRoobaea et al., 2013, 92; Elling et al., 2012; Haak et al., 2009, 194; Szóstek 2011, 97). W omawianych badaniach zastosowano symultaniczną wersję TAP, co oznacza, iż każdy z użytkowników proszony był o werbalizację myśli, odczuć i działań w trakcie indywidualnej interakcji z analizowanymi serwisami internetowymi.

ZMET wywodzi się z badań marketingowych i służy do mapowania modeli mentalnych użytkowników, także z uwzględnieniem stanów afektywnych. Istnieje wiele wariantów ZMET, które różnią się liczbą i ułożeniem poszczególnych etapów badania. Zdaniem Geralda Zaltmana, twórcy ZMET, kluczowym elementem jest multisensoryczność badania, osiągnięta m.in. przez bazowanie na opowieści o materiałach wizualnych (Zaltman, 1996, 13). W badaniach autorki ZMET posłużyła jako inspiracja do wypracowania własnej procedury, w której każdy z użytkowników proszony był o krótką opowieść o wcześniej zgromadzonych przez siebie obrazach (od 8 do 12), przedstawiających wyobrażenie o wybranym teatrze krakowskim. Następnie zadawane były pytania dotyczące zaprezentowanych obrazów, dotyczących np. najbardziej znaczących obrazów oraz skojarzeń, których nie udało się oddać za pomocą materiałów wizualnych.

Wybór serwisów WWW poddanych badaniu składał się z kilku etapów. Za kryteria konieczne przyjęto posiadanie przez serwis polskiej wersji językowej oraz jego przynależność do teatru krakowskiego, rozumianego jako teatr z siedzibą w Krakowie, niezależnie od organizatora oraz posiadania własnych scen teatralnych.

Pierwszą kwerendę przeprowadzono w marcu 2015 r. Korzystając z wyszukiwarki Google, w polu wyszukiwania prostego wpisano dwa terminy: „teatr” i „Kraków”. Zapoznano się z trzema pierwszymi stronami wyników. Następnie kompletowane zestawienie serwisów uzupełniono na podstawie zawartości internetowych repertuarów teatralnych oraz katalogów DMOZ i PKT. Opisaną kwerendę powtórzono w listopadzie 2015 r. Po tej aktualizacji na liście znalazło się 21 serwisów internetowych. Ostatecznie doboru dziewięciu serwisów poddanych badaniu empirycznemu dokonano nakładając na otrzymaną listę serwisów następujące filtry: typologię widowisk teatralnych według Christophera Balme'a (Sajkiewicz, 2010, 29), aktualną strukturę teatralnej mapy Krakowa oraz interfejs serwisu WWW, w tym układ serwisu.

Decyzja, które serwisy, spośród dziewięciu zaproponowanych, były uwzględniane podczas sesji należała już do użytkownika. Najpierw użytkownik wybierał jeden teatr, dla którego kompletował obrazy oddające subiektywne skojarzenia związane z tym teatrem. Ten wybór teatru determinował pierwszy serwis, który później poddany był badaniu. Następne dwa serwisy wybierane były przez użytkownika za pomocą planszy, na której znajdowały się niepodpisane, kolorowe zrzuty ekranowe stron głównych wszystkich dziewięciu serwisów, każdy w formacie A4. Przyjęto zasadę: użytkownik dokonywał wyboru tylko jednego serwisu z każdej z trzech grup (pierwsza liczba w nawiasie oznacza, ile razy użytkownicy wybrali teatr do części z obrazami; druga odnosi się do wyboru serwisu z planszy):

- Gr. 1. Teatry dramatyczne: Krakowski Teatr Scena STU (3+1), Narodowy Stary Teatr im. Heleny Modrzejewskiej (2+3), Teatr im. Juliusza Słowackiego (0+1).
- Gr. 2. Teatry dramatyczne: Teatr Barakah (1+1), Teatr Łaźnia Nowa (1+4), Teatr Odwrócony (1+2).
- Gr. 3. Teatry muzyczne i lalki: Krakowski Teatr Variété (0+4), Opera Krakowska (1+1), Teatr Groteska (1+3).

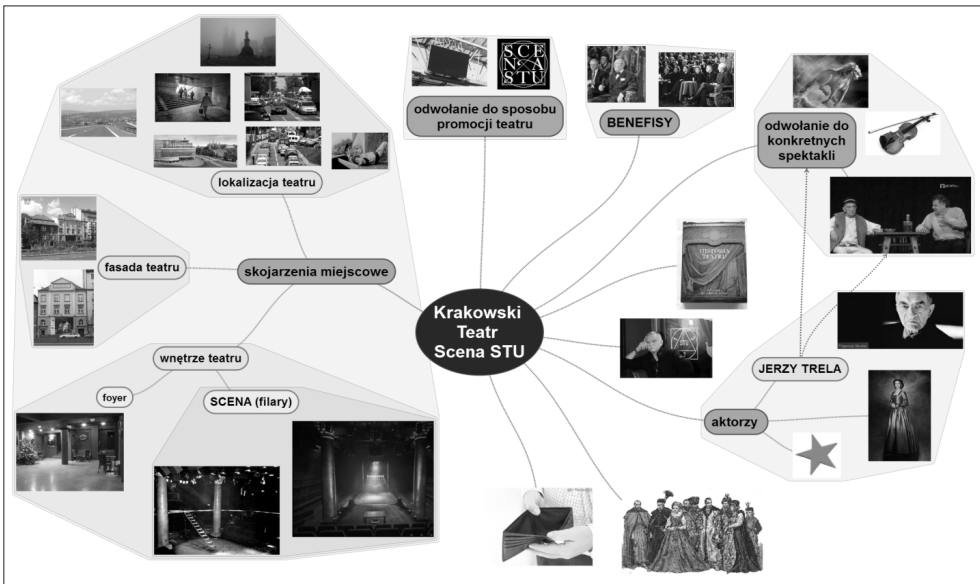
Na początku badania użytkownik opowiadał o każdym ze zgromadzonych wcześniej materiałów wizualnych. Po zakończeniu wypowiedzi zadawane były pytania do przedstawionego zbioru obrazów. Po wybraniu dwóch pozostałych serwisów, użytkownik wskazywał kolejność, w której chciałby dokonywać interakcji z serwisami. W każdym z serwisów użytkownik wykonywał 12 zadań. Po zbadaniu trzech serwisów zadawane były ogólne, podsumowujące pytania. Sesję z każdym z użytkowników rejestrowano w formie audio, od momentu prezentowania obrazów, do udzielenia odpowiedzi na ostatnie z pytań podsumowujących. Nie przewidziano limitu czasowego. Całość zgromadzonego materiału poddano transkrypcji oraz analizie tematycznej, którą usprawniono programem QDA Miner.

5. Opis przekazu symbolicznego serwisu Krakowskiego Teatru Scena STU

Serwis Teatru STU analizowany był czterokrotnie, przy czym trójka użytkowników zdecydowała się zbierać obrazy oddające wyobrażenia związane z Krakowskim Teatrem Scena STU. Wyodrębniono cechy wspólne tego zbioru, które przedstawiono w sposób graficzny (Rys. 1). Są to odwołania do konkretnych spektakli oraz „Benefisów”, lokalizacji teatru,

fasady budynku, wnętrza teatru, a także osób związanych z Teatrem, wśród których powtarzającym się nazwiskiem był Jerzy Trela. Powtarzało się także nawiązywanie do sposobu promocji Teatru, w tym odwołania do telebimu znajdującego się na fasadzie Teatru, na którym wyświetlane są zapowiedzi spektakli. Przywołany został też logotyp i jego konsekwentne stosowanie na materiałach promocyjnych Teatru. Natomiast najbardziej znaczącymi obrazami w kontekście Teatru Scena STU były obrazy przedstawiające Jerzego Trełę i historię Teatru – w kontekście realizacji marzeń aktorskich (dla użytkownika 1A1) oraz fasada budynku (dla użytkowników 4W2 i 10W5).

Rys. 1. Mapa myśli przedstawiająca obrazy do Krakowskiego Teatru Scena Stu zgromadzone przez użytkowników 1A1, 4W2 i 10W5



Na podstawie wypowiedzi użytkowników oceniających serwis Krakowskiego Teatru Scena STU można przyjąć, iż powierzchowne skojarzenia z Teatrem udało się dobrze oddać w serwisie. Benefisom poświęcona została osobna zakładka w menu, użytkownik 10W5 wskazał elementy potwierdzające w serwisie skojarzenia ze znanymi aktorami oraz wysokimi cenami biletów, zaś na stronie głównej pojawiają się zdjęcia sceny, co jest dosłowną reprezentacją tego elementu świata realnego w rzeczywistości wirtualnej, a – jak zauważyli użytkownicy 4W2 oraz 10W5 – układ serwisu także nawiązuje do wyglądu sceny. Z opisu sceny Teatru przedstawionego przez użytkownika 4W2 wyłania się obraz kwadratu oraz bliskiego kontaktu widzów z aktorami. W serwisie nie udało się odszukać nawiązania do wspomnianego bliskiego kontaktu z aktorami, o którym podczas opowieści o Teatrze STU mówił także użytkownik 7W4, natomiast silnie odwzorowana została geometria sceny, do której odnoszą się wypowiedzi Q1 i Q2 oraz Q3 dotyczące układu serwisu.

(Q 1) „Bo scena jest charakterystyczna w Teatrze Stu, prawda? To znaczy widzowie mają bardzo bliski kontakt z aktorami. Siedzą niemalże koło nich i otaczają ich z trzech, tą scenę z trzech stron, więc ten kontakt jest taki niemalże namacalny.

To jest taki kwadracik, w którym się wszyscy spotykają” – użytkownik 4W2 o Krakowskim Teatrze Scena STU.

(Q 2) „Jest to teatr krakowski. Jest to teatr kameralny, moim zdaniem, ale tego tutaj nie widzę, aczkolwiek ta strona... ona, jej ciemne tło powoduje, że no moje skojarzenia są ewidentne właśnie z tym wnętrzem Teatru, z tym kwadratem, z tą publicznością otaczającą scenę – to, jak najbardziej budzi moje skojarzenia” – użytkownik 4W2 o serwisie Krakowskiego Teatru Scena STU.

(Q 3) „[C]ały serwis jest, tak jakby, według mnie... ładnie poukładany w takie... prostokąty... (...) cały serwis jest oparty na takich prostokątach, przynajmniej ja tak widzę: jest prostokąt, nazwijmy to z nagłówkiem; prostokąt ze zmieniającymi się reklamami najprawdopodobniej najbliższych spektakli; jest kolejny prostokąt – aktualności; prostokąt z możliwością zamawiania biletów, a także newsletter – też prostokąt. I stopka – prostokąt, także... Ten kwadrat, który *de facto* też jest prostokątem – idealnie pasuje” – użytkownik 10W5 o serwisie Krakowskiego Teatru Scena STU.

Odnotowano trudność z odwzorowaniem głębszych skojarzeń. W serwisie nie udało się odnaleźć elementów wskazujących na krakowski charakter Teatru, o którym mówił użytkownik 4W2. Natomiast wątpliwości użytkownika 7W4, co do charakteru Teatru STU budził zamieszczony w serwisie WWW opis:

(Q 4) „Rodzina STU, komuna teatralna, czołówka światowej alternatywy, enklawa kontestacji – no, to <śmiech> To pierwsze zdanie jest wątpliwe. <śmiech> W moim odbiorze, więc ja mówiłem raczej co innego. Tak, bo czołówka światowej alternatywy – no, początki takie były... początki takie były. Enklawa kontestacji, kontrkultury – no to, co tam by Klata musiał pisać, no?” – użytkownik 7W4.

Dla użytkownika 1A1 szczególnie znaczące były dwa filary znajdujące się na scenie, które przywołały myśl o teatrze antycznym i znajdują odzwierciedlenie w zdjęciach prezentowanych w serwisie:

(Q 5) „To te... Bo w środku są takie właśnie, o takie! Takie filary. I bardzo mi się to kojarzy z jakimś starym, jakimś greckim, antycznym teatrem. Miejscem, właściwie sceną jakąś taką antyczną. Takie coś mam, takie mam sobie też skojarzenie. Po prostu ten środek. Środek Teatru Stu. Te filary. Tak mi się fajnie kojarzą, tak. Ale bardzo lubię w ogóle, ja... jakoś tak ten grecki teatr mi się zawsze strasznie podobało. Takie... Także pozytywne, pozytywne bardzo skojarzenie” – użytkownik 1A1 o Krakowskim Teatrze Scena STU.

Równocześnie ów antyczny teatr, a przez to i Teatr STU, był przez użytkownika 1A1 bardzo pozytywnie kojarzony, na co nałożyły się inne skojarzenia zobrazowane m.in. przez wytworne, eleganckie kostiumy. Natomiast ogólne wrażenie z interakcji z serwisem WWW, z powodu napotkanych trudności z zakresu użyteczności, pozostawiło użytkownika 1A1 z poczuciem dysonansu:

(Q 6) „Właśnie od teatru bym oczekiwał takiego klasycznego przedstawienia, klasycznego teatru, klasycznej sceny, z takimi filarami greckimi, co jest w Teatrze Stu. Po serwisie.... <śmiech> ciężko jest mi to później określić, bo jest taki niedopracowany ten serwis. (...)” – użytkownik 1A1.

Gdyby natomiast rozpatrywać serwis Teatru STU tylko w zakresie estetyki, podstawowe skojarzenia z tym Teatrem, zgodnym zdaniem użytkowników zostały oddane.

- (Q 7) „Mnie taka – tak samo sam teatr określiłem jako wyrafinowany, tak samo strona też mi się wydaje taka stonowana, bardzo z klasą” – użytkownik 10W5 o serwisie Krakowskiego Teatru Scena STU.
- (Q 8) „Tak. No jest trochę, można powiedzieć, że jest ciemno. Ciemno, smutno. No, ale tak właśnie teatralnie. W jakiś inny świat się można przenieść... Brązowa, stonowana. Taka elegancka. Myślę, że do takiego teatru, który się kojarzy jednak z elegancją, a nie z jakimiś takimi... mówiąc [w] cudzysłowie, nowoczesnymi spektaklami, to jest ok, tak. Tak kolorystyka odpowiada” – użytkownik 1A1 o serwisie Krakowskiego Teatru Scena STU.
- (Q 9) „Ogólnie rzecz biorąc, nie lubię czarnego tła, ale akurat w tym przypadku w ogóle mi nie przeszkadza. Dłatego, że jest to teatr, czy...-& Dłatego, że to jest TEN teatr. Bardzo pasuje, moim zdaniem. Ta strona została przygotowana w sposób przemyślany. Chce Pani to rozwinąć jakoś, czy już na tym chce Pani zakończyć? No właśnie jak widzę, jak widzę wewnątrz tego teatru, właśnie sfotografowane w taki sposób, no to widać wyraźnie, że tam jest właśnie tak: czarno i tylko jedna smuga światła puszczona na środek sceny i ja właśnie tak widzę tę stronę internetową, że ona jest czarna i ma w sobie tylko ten jeden, jedno okno, w którym wyświetlają się informacje. I to budzi moje skojarzenia z tym teatrem. Nie wiem – może nie być zachęcające dla ludzi, którzy nie chodzą do tego teatru, właśnie przez to, że jest taka czarna i ponura” – użytkownik 4W2 o serwisie Krakowskiego Teatru Scena STU.
- (Q 10) „[N]o ale to „Wyzwolenie” jakoś bardziej do mnie przemawia, (-), bo tam mamy jeszcze taki jakiś układ, taki tutaj sceniczny, (-)” – użytkownik 7W4.

6. Czy serwis WWW musi oddawać charakter teatru?

Badania własne oparto na wstępnym założeniu – serwisom WWW przypisano możliwości wspomagania realizacji poza-instrumentalnych potrzeb użytkownika, stąd, oprócz użyteczności i funkcjonalności, postanowiono uwzględnić w badaniu wartości hedonistyczne oraz przekaz symboliczny. W celu potwierdzenia słuszności tego założenia, zadano użytkownikom pytanie o funkcje, które ich zdaniem powinien realizować serwis WWW. W efekcie otrzymano katalog sześciu funkcji:

- (1) archiwalnej/dokumentacyjnej
- (2) edukacyjnej/kulturotwórczej
- (3) informacyjnej
- (4) oddawania charakteru teatru
- (5) pragmatycznej
- (6) promocyjnej.

W tym artykule uwagę poświęcono znaczeniu przekazu symbolicznego, rozumianego jako oddawanie charakteru teatru.

Opinie użytkowników odnośnie istotności przekazu symbolicznego serwisów WWW teatrów krakowskich były zróżnicowane. Można je ująć w ramach następujących postaw:

- (1) sceptycyzm wobec możliwości oddawania charakteru teatru za pomocą serwisu WWW (np. Q11)

- (2) kompromis – serwis może oddawać charakter teatru, ale nie powinno to przesłaniać funkcji pragmatycznych i/lub informacyjnych serwisu (np. Q12)
- (3) konieczność – serwis powinien lub nawet musi oddawać charakter teatru (np. Q13).
- (Q 11) „Ciężko powiedzieć na podstawie samego serwisu, jakiego rodzaju jest to teatr, co... W żaden sposób serwis nie może tego powiedzieć. Normalnym jest, że marketing musi działać niezależnie od formy... w jaką ubrane są spektakle, jakiego rodzaju są tam robione produkcje. Marketing a scena często niewiele ma ze sobą wspólnego. Po prostu mamy opis spektaklu, mamy zdjęcia ze spektaklu i cała otoczka zrobiona przez marketing, a to, co jest opowiedziane często nie musi mieć żadnego związku, zresztą tak samo, jak z trailerami filmowymi – tutaj sobie coś mruga, najlepsze zdjęcia powybijane, najbardziej charakterystyczne, zrobiony jakiś krótki opis... <odgłos przewijania strony> więc no ciężko mi odpowiedzieć na to pytanie na podstawie serwisu. Moim zdaniem się nie da odpowiedzieć na takie pytanie na podstawie serwisu” – użytkownik 9A5.
- (Q 12) „Myślę, że powinien na pewno w jakiś sposób oddawać charakter teatru, tak jak ten Teatr Odwrócony, ale nie powinno to przysłaniać tych funkcji... dla których tak naprawdę, wydaje mi się, wchodzimy na stronę teatru, czyli że chcemy kupić bilet, sprawdzić terminy spektakli itd., więc... No, jest fajnie, kiedy to pasuje, ten serwis do danego teatru, ale powinno się też dać właśnie zrobić to, co najważniejsze, tak” – użytkownik 2A2.
- (Q 13) „Musi oddawać charakter, musi – w takim mieście, jak Kraków, musi oddawać charakter teatru, dlatego że teatrów w Krakowie jest więcej i ta strona po prostu nie może sobie tak wisieć. Ona musi do kogoś przemawiać i teatry no wyrabiają sobie swoją taką, taką wierną publiczność wśród ludzi z Krakowa, chociaż... chociaż wiem, że coraz mniej akurat krakowian chodzi do teatru, a coraz więcej osób przyjeżdżnych” – użytkownik 4W2.

Użytkownicy poszukali elementów potwierdzających wyobrażenia o danym teatrze m.in. w opisach, zarówno teatru, jak i spektakli oraz w materiałach wizualnych, takich jak zdjęcia i/lub multimedia. Dla niektórych użytkowników ważne było, gdy materiały wizualne przedstawiały wnętrze i fasadę teatru, inni natomiast zwracali większą uwagę na materiały ukazujące fragmenty spektakli, gdyż na ich podstawie próbowali rozpoznać konwencję, jaką posługuje się dany teatr.

Zidentyfikowano różne poziomy wyobrażeń – począwszy od powierzchniowych, które związane są m.in. z lokalizacją teatru oraz konkretnymi osobami, przez abstrakcyjne (jak np. ryzyko i nieokreśloność kojarzące się użytkownikowi 2A2 z Teatrem Odwróconym), aż do elementów na tyle silnie zakorzenionych w świadomości odbiorców (np. wybrane logotypy), że funkcjonują one niezależnie od całości serwisów i ocena ich spójności z serwisem WWW staje się niemożliwa. Użytkownicy najczęściej odszukiwali w badanych serwisach WWW powierzchowne skojarzenia, stąd otwarte pozostaje pytanie o to, na ile możliwe jest oddanie charakteru teatru za pomocą serwisów WWW w przypadku bardziej abstrakcyjnych wyobrażeń.

Ocena trafności przekazu symbolicznego serwisów WWW zależna była od wiedzy i doświadczenia użytkownika, co szczególnie widoczne było na przykładzie serwisów Teatru Odwróconego oraz Opery Krakowskiej. Kolejnym z indywidualnych czynników wpływających na możliwość analizy przekazu symbolicznego serwisu było nastawienie

emocjonalne użytkowników. Zaobserwowano kilka kategorii wrażeń emocjonalnych i poznawczych, doświadczanych przez użytkowników podczas badania serwisów WWW teatrów krakowskich: odniesienia do wrażeń dotyczących świata realnego, wrażenia wynikające z interakcji z serwisem oraz wrażenia wywołane sytuacją badania i wrażenia nie wyrażone podczas werbalizacji. Tylko część z nich można zakwalifikować do UX serwisów WWW.

Nie sposób omówić w tym artykule trudności napotykanych przez użytkowników podczas interakcji z serwisami. Wiązały się one m.in. z procesem zakupu i informowaniem o dostępności biletów, obecnością i poprawnością działania wyszukiwarki, przejrzystością i klarownością etykiet i opisów, aktualizacją treści, nadmiarem bądź niekompletnością informacji, a wreszcie nieprzewidywalnością serwisu. Istotne jest to, iż użyteczność bywała wiązana z przekazem symbolicznym (np. Q14):

(Q 14) „Na pewno nie Teatr Variété – róbmy drogą eliminacji – strasznie mi się nie podoba to, że nie widzę dwóch pierwszych literek, bo ktoś genialnie umieścił biały napis na jasnym tle. Od razu mi się wydaje, że skoro napisu dobrze nie potrafią zrobić, nie zadowolą mnie, jeżeli chodzi o sztukę” – użytkownik 10W5.

W wypowiedzi Q14 oraz w opisie przekazu symbolicznego Teatru STU zaznaczono, że negatywne wrażenia związane z użytecznością serwisu, mogą być włączone przez użytkownika w ocenę przekazu symbolicznego serwisu i sprzyjać niekorzystnej dla teatru interpretacji (Q6). Nie jest to zaskakujące twierdzenie. W trakcie badań zauważono również zależność w drugą stronę, tj. pozytywnej ocenie wartości pragmatycznych towarzyszyła luka w przekazie symbolicznym. Najwyraźniej zauważono to w wypowiedziach użytkowników 4W2, 6A3 i 2A2. Dwie pierwsze dotyczyły serwisu Teatru Variété, zaś trzecia serwisu Teatru Odwróconego. W każdym z tych przypadków użyteczność serwisu została przez użytkowników oceniona pozytywnie, natomiast w ogólnym odbiorze serwisu brakowało jakiegoś elementu oddającego wyobrażenia użytkownika o danym teatrze. Użytkownik 2A2, podczas całego badania, podkreślał swój sceptycyzm wobec możliwości pełnienia przez serwis WWW funkcji oddawania charakteru teatru, zaś jedynym odstępstwem od tej tezy, była właśnie ogólna ocena serwisu Teatru Odwróconego. Zdanie uzasadniał odwołaniem do teorii czystej formy Witkacego, na której opiera się Teatr Odwrócony (Q15, Q16):

(Q 15) „Teoria czystej formy to jest taka teoria, która mówi, że widz, który ogląda spektakl, ogląda dramat... powinien doznać takich wewnętrznych jakichś uczuć, że po wyjściu ze spektaklu... powinien mieć takie poczucie, że brał udział w jakimś takim nieuchwytnym, bardzo dziwnym zdarzeniu, które miało go, w jakiś sposób tam, właśnie, oczyścić. Na tej zasadzie” – użytkownik 2A2 o Teatrze Odwróconym.

(Q 16) „Korzysta się z niego bardzo wygodnie, ale... Właśnie przez to, że korzysta się z niego tak bardzo wygodnie, to czegoś mi tutaj brakuje. Teraz mi przyszło do głowy, że gdyby tu był jakiś motyw przewodni, muzyczny właśnie taki... może jakiś Teatru Odwróconego... to, czy on by nie zaburzył tej właśnie takiej hm... zasady czystości przestrzeni na tej stronie, tak bym to nazwał” – użytkownik 2A2 o serwisie Teatru Odwróconego.

7. Wnioski

Opisane w artykule badanie stanowiło próbę zapewnienia luk w piśmiennictwie naukowym z zakresu przekazu symbolicznego serwisów internetowych oraz User Experience serwisów teatrów. Przedstawiony został katalog sześciu funkcji przypisywanych serwisom WWW oraz zarysowano zależności między wartościami pragmatycznymi UX a przekazem symbolicznym. Do czynników wpływających na przekaz symboliczny zaliczono cechy indywidualne użytkownika, takie jak doświadczenie i wiedza oraz nastawienie emocjonalne, a także elementy wnoszone przez interakcję użytkownik-serwis, reprezentowane m.in. przez użyteczność, przy czym pozytywnej ocenie użyteczności serwisu niekoniecznie towarzyszy oczekiwany przekaz symboliczny.

Prowadzone są dalsze badania nad rozszerzoną definicją przekazu symbolicznego i jego wpływem na zachowania informacyjne użytkowników serwisów WWW.

Bibliografia

- AlRoobaea, R. S.; Al-Badi, A. H.; Mayhew, P. J. (2013). Generating a Domain Specific Inspection Evaluation Method through an Adaptive Framework: A Comparative Study on Educational Websites. *International Journal of Human Computer Interaction*, 4(2), 88–116.
- Bargas-Avila, J. A.; Hornbæk, K. (2011). Old Wine in New Bottles or Novel Challenges? A Critical Analysis of Empirical Studies of User Experience. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2689–2698.
- Bhui, K.; Ibrahim, Y. (2013). Marketing the “radical”: Symbolic communication and persuasive technologies in jihadist websites. *Transcultural Psychiatry*, 50(2), 216–234.
- Coursaris, C. K.; Swierenga, S. J.; Pierce, G. L. (2010). Effects of Aesthetics and Playfulness on Web Usability – An empirical investigation. Americas Conference on Information Systems 2010 Proceedings [online], paper 549, AIS Electronic Library [21.06.2017], <http://aisel.aisnet.org/amcis2010/549/>
- Diefenbach, S.; Hassenzahl, M. (2011). The dilemma of the hedonic – Appreciated, but hard to justify. *Interacting with Computers*, 23(5), 461–472.
- Elling, S.; Lentz, L.; Jong, M. de (2012). Combining Concurrent Think-Aloud Protocols and Eye-Tracking Observations: An Analysis of Verbalizations and Silences. *IEEE Transactions On Professional Communication*, 55(3), 206–220.
- Haak, M. J. van den; Jong, M. de; Schellens, P. J. (2009). Evaluating municipal websites: A methodological comparison of three think-aloud variants. *Government Information Quarterly*, 26(1), 193–202.
- Hager, M. A.; Kopczynski-Winkler, M. (2012). Motivational and Demographic Factors for Performing Arts Attendance Across Place and Form. *Nonprofit and Voluntary Sector Quarterly*, 41(3), 474–496.
- Helfenstein, S. (2012). Increasingly emotional design for growingly pragmatic users? A report from Finland. *Behaviour & Information Technology*, 31(2), 185–204.
- Hernik, J. (2012). Morfologia marki organizacji niekomercyjnych na przykładzie polskich teatrów. *Zeszyty Naukowe Uniwersytetu Szczecińskiego. Problemy Zarządzania, Finansów i Marketingu*, 24, 303–316.
- Karapanos, E.; Zimmerman, J.; Forlizzi, J.; Martens J.-B. (2009). User experience over time: an initial framework. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 729–738.
- Kossecki, P.; Świerczyńska-Kaczor, U. (2014). Acquiring a Digital Audience for Theaters – Looking Through The Lenses of Customer Equity and Empirical Research. *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems*, 1277–1284.
- Laskowski, M. (2011). Czynniki zwiększające jakość użytkową interfejsów aplikacji internetowych. *Logistyka*, 6, 2191–2199.

- Nawaz, A. (2013). *Website user experience: A cross-cultural study of the relation between users' cognitive style, context of use, and information architecture of local websites* [online]. Copenhagen Business School Handelshøjskolen [21.06.2017], <http://openarchive.cbs.dk/handle/10398/8871>
- O'Brien, H. L. (2010). The influence of hedonic and utilitarian motivations on user engagement: The case of online shopping experiences. *Interacting with Computers: Special Issue on User Experience*, 22(5), 344–352.
- Pallud, J.; Straub, D. W. (2014). Effective website design for experience-influenced environments: The case of high culture museums. *Information & Management*, 51(3), 359–373.
- Partala, T.; Kallinen, A. (2011). Understanding the most satisfying and unsatisfying user experiences: Emotions, psychological needs, and context. *Interacting with Computers*, 24(1), 25–34.
- Riley-Huff, D. A. (2009). Design Insights and Inspiration from the Tate: What Museum Web Sites Can Offer Us. *Libraries and the Academy*, 9(1), 79–98.
- Sajkiewicz, V. (2010). *Teatr w świecie widowisk*. W: E. Wąchocka (red.) *Widowisko – teatr – dramat: skrypt dla studentów kulturoznawstwa*. Katowice: Wydaw. Uniwersytetu Śląskiego, 21–38.
- Saucken, C. von; Gomez, R. (2014). Unified user experience model enabling a more comprehensive understanding of emotional experience design. *Proceedings of the 9th International Conference on Design and Emotion: The Colors of Care*, 631–640.
- Schuitema, G; Anable, J.; Skippon, S.; Kinnear, N. (2013). The role of instrumental, hedonic and symbolic attributes in the intention to adopt electric vehicles. *Transportation Research Part A - Policy and Practice*, 48, 39–49.
- Siwiak, A; Burszta, W.; Duniec, K.; Guderian-Czaplińska, E.; Wittels, K. (2013). *Badanie publiczności teatrów w stolicy* [online]. Narodowe Centrum Kultury [21.06.2017], <http://www.nck.pl/baza-badan/302562-2013-badanie-publicznosci-teatrow-w-stolicy/>
- Szóstek, A. (2011). Metody ilościowe i jakościowe w dyscyplinie interakcja człowiek–komputer. *Zeszyty Naukowe Politechniki Poznańskiej. Organizacja i Zarządzanie*, 56, 93–107.
- Troilo, G.; Cito, M. C.; Soccia, I. (2014). Repurchase Behavior in the Performing Arts: Do Emotions Matter without Involvement? *Psychology & Marketing*, 31(8), 635–646.
- Trosheva, N. (2013). *The Royal Opera House Website Evaluation Report* [online]. Kingston University [21.06.2017], http://ninatrosheva.com/assets/files/ROH_Evaluation_Report.pdf
- Turrini, A.; Soccia, I.; Maulini, A. (2012). Web communication can help theaters attract and keep younger audiences. *International Journal of Cultural Policy*, 18(4), 474–485.
- Wu, J.; Holsapple, C. W. (2014). Imaginal and emotional experiences in pleasure-oriented IT usage: A hedonic consumption perspective. *Information & Management*, 51(1), 80–92.
- Yamin M.S.; Jaafar E.A. (2013). Web Design for Science Museum towards Engaging User Experience. In: Kotzé P., Marsden G., Lindgaard G., Wesson J., Winckler M. (eds) *Human-Computer Interaction – INTERACT 2013. INTERACT 2013. Lecture Notes in Computer Science*. Berlin: Heidelberg: Springer, 745–754.
- Zaltman, G. (1996). Metaphorically speaking. *Marketing Research*, 8(2), 13–20.

Symbolic Messages and User Experience Approach Exemplified with the Websites of Cracow Theaters

Abstract

Purpose/Thesis: The purpose of the article was to investigate symbolic messages and User Experience approach in the case of the websites of Cracow theaters. The websites were claimed to support users' non-instrumental needs.

Approach/Methods: The study was conducted by means of qualitative research. Data were gathered with TAP and the technique based on ZMET. Ten users participated in the study and all of them had some theater background. Nine theater websites were chosen for the study. All empirical material was examined with thematic analysis.

Results and conclusions: Symbolic messages were depicted based on the Theater Scena STU website. The user's individual characteristics, emotional attitude and elements brought by the user-website interaction were found to be the factors influencing the symbolic messages. It was discovered that the positive assessment of the website usability does not necessarily lead to the expected symbolic message.

Originality/Value: The article is intended as a contribution to further research on the relations between symbolic messages, informative messages and the usability of websites, which have not been studied in the information science so far.

Keywords

Function of website. Symbolic message. Theater website. Usability. User Experience.

MAGDALENA ZYCH jest absolwentką Instytutu Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego. Od roku akad. 2016/17 jest doktorantką informatologii i bibliologii na Wydziale Zarządzania i Komunikacji Społecznej UJ. Zajmuje się problematyką użytkowników serwisów internetowych.

Kontakt z autorką:

magdalena.zych@student.uj.edu.pl

Wydział Zarządzania i Komunikacji Społecznej

Uniwersytet Jagielloński

ul. prof. Stanisława Łojasiewicza 4

30–348 Kraków

Tibor Koltay, Sonja Špiranec, László Z. Karvalics (2016): *Research 2.0 and the future of information literacy*. Amsterdam [etc.]: Elsevier: Chandos Publishing, ss. 180. ISBN 978-0-08-100075-5

Od niemal dziesięciu lat termin *information literacy* (IL) przeżywa renesans i stał się niezwykle modny, o ile oczywiście można mówić o trendach w kontekście badań naukowych. Koncepcja IL jest bardzo intensywnie omawiana i badana. Podejmowane są próby jej rewizji, reorientacji i zbadania zastosowań w różnych sferach nauki i życia. Jednak nadal jądrem definicyjnym IL pozostaje zestaw kompetencji pozwalających na dostęp, ocenę i wykorzystanie informacji. Wyłącznie samemu zagadnieniu edukacji informacyjnej poświęconych jest kilka międzynarodowych konferencji, a w ramach krajowych i międzynarodowych stowarzyszeń i organizacji funkcjonują zespoły ds. IL.

Dla celów niniejszej recenzji teminy *information literacy*, *IL* oraz *edukacja informacyjna* używane będą wymiennie.

Rozwój mediów społecznościowych spowodował chętnie dodawanie rozszerzenia 2.0 do wcześniej funkcjonujących w literaturze terminów, w celu podkreślenia ich funkcjonowania w świecie Web 2.0. Badacze edukacji informacyjnej zauważyli tę tendencję i tak powstało wiele prac, które dyskutują rolę, jaką może odegrać IL w nowo powstałych zjawiskach. Jedną z pierwszych, bardzo dobrze przyjętych w środowisku publikacji była *Information literacy and Library 2.0* z 2008 r.¹

Po ośmiu latach od wydania tej książki dwójce węgierskich badaczy oraz badaczka chorwacka postanowiło omówić rolę, jaką edukacja informacyjna może pełnić na rzecz Nauki 2.0 (ang. *Research 2.0*).

Autorzy to dwóch węgierskich profesorów (Tibor Koltay pracuje na Uniwersytecie Szent István, a László Z. Karvalics na Uniwersytecie w Szegedzie) oraz chorwacka profesor Sonja Špirnec z Uniwersytetu w Zagrzebiu. Wszyscy mają w swoim dorobku publikacje nt. edukacji informacyjnej i tematów pokrewnych. Wybór międzynarodowego wydawnictwa Chandos i jego serii „Advances in Information” uważam za bardzo trafny. Po pierwsze, wydawnictwo to od lat publikuje książki dotyczące edukacji informacyjnej i tematów pokrewnych. Po drugie, jest to wydawca rozpoznawalny i uznany międzynarodowo, a niedawny transfer do grupy wydawniczej Elsevier na pewno tę rozpoznawalność (a co za tym idzie, liczbę czytelników i cytowań) wzmocni. Ponadto serie wydawnicze Chandos Publishing są dobrze zredagowane, przejrzyste i, z technicznego punktu widzenia, przyjazne w lekturze.

Według mojej wiedzy, *Research 2.0 and the future of information literacy* to pierwsza książka opisująca to zagadnienie. Także inne artykuły na ten temat, które ukazały się do tej pory, są autorstwa właśnie wyżej wymienionej trójki badaczy i ukazały się w języku angielskim w międzynarodowych, uznanych czasopismach² lub były publikowane w czasopismach o zasięgu krajowym, w językach ojczystych autorów.

Jak autorzy piszą we wstępie, celem ich pracy jest wyjaśnienie, czym jest Nauka 2.0 i zwrócenie uwagi na czynniki determinujące jej rozwój; pokazanie, że IL w wyniku rozwoju Nauki 2.0 musi ulec w pewnych aspektach modyfikacjom (s. 2). Jednak w kolejnych rozdziałach, a także w podsumowaniu zawartym w rozdziale czwartym, niejednokrotnie ponownie wspominają o głównych założeniach publikacji i w kilku miejscach nie są one tożsame. Na s. 131 można przeczytać, że przedmiotem rozważań jest analiza zależności między Nauką 2.0 a IL w wielu kontekstach; z kolei na s. 143 jako jedno z głównych przesłań tej książki autorzy uznają postulat, iż biblioteki akademickie i bibliotekarze w nich pracujący muszą ciągle się doszkalać, żeby móc wspierać naukowców w procesie komunikacji

¹ P. Godwin & J. Parker (eds.). (2008). *Information literacy meets Library 2.0*. London: Facet.

² Por. np. T. Koltay; S. Špiranec & L. Z. Karvalics (2015). Perspectives On...: The Shift of Information Literacy Towards Research 2.0. *The Journal of Academic Librarianship*, 4187–93.

naukowej; natomiast na s. 153 autorzy dokonują podsumowania głównego przesłania – Nauka 2.0 jako ta, która zmieniła zachowania badaczy, jednak nie wszystkich i w różnym stopniu, w zależności od dyscypliny nauki, środowiska naukowego, czy wieku. Uważny czytelnik może odnieść poczucie pewnego chaosu. Brak bowiem jednoznacznie określonego założenia publikacji.

Jeśli chodzi o dwa tytułowe terminy, *Research 2.0* (Naukę 2.0) autorzy uważają za komplementarną do eNauki (ang. *eScience*) i definiują tę pierwszą jako środki do realizowania założeń drugiej. Zaś co do *information literacy*, uważają ją, oraz szereg pokrewnych *literacies*, za niezbędną w pracy współczesnego badacza i zaliczają do kompetencji praktycznych.

Książka składa się z czterech rozdziałów. W rozdziale pierwszym poruszono zagadnienie zmiany paradygmatów badawczych w stronę Nauki 2.0. Opisując, jakim zmianom podlega współczesna nauka, opisano pojęcie Nauka 2.0. Przedyskutowano z różnych perspektyw zagadnienie danych naukowych: ich wytwarzanie i wykorzystanie, jakość, długoterminowe przechowywanie, ochronę, dzielenie się (udostępnianie) danymi. Zwrócono uwagę na zintensyfikowanie ilości zbieranych danych badawczych, wielość typów danych i operacji na nich oraz na konieczność nowatorskiego zarządzania danymi (ang. *data-intensive paradigm of scientific research*) i ich wpływ na komunikację naukową. Podkreślono rolę, jaką mogą tu odegrać bibliotekarze, zważywszy na długą dotychczasową tradycję współpracy bibliotekarzy z badaczami na rzecz wykazywania znaczenia dorobku tych drugich, głównie w kontekście cytowania danych badawczych. Opisano także, jaką rolę media społecznościowe czy, szerzej, Web 2.0 i Open Access, odgrywają w przekształcaniach komunikacji naukowej i zarządzaniu danymi badawczymi oraz jaki mają wpływ na rzetelność publikowanych treści. Pobieźnie omówiono zagadnienie humanistyki cyfrowej oraz cytowania danych, gdzie poza znanymi powszechnie narzędziami bibliometrycznymi, takimi jak Web of Science czy Google Scholar, wspomniano o bibliometrii alternatywnej oraz semantycznej. W rozdziale pierwszym omówiono także zdolności i umiejętności, jakie powinny cechować współczesnych badaczy (Badaczy 2.0). Ta część wydaje się nieco trywialna, po raz kolejny (i nie ostatni) w tej książce zaprezentowano po prostu zestawienie (skądinąd trafnie dobranego) piśmiennictwa tematu. Ciekawsze i bardziej potrzebne w kontekście dyskusji nad Nauką 2.0 wydaje się, zawarte w tym samym rozdziale, omówienie koncepcji Otwartej Nauki (ang. *Open Science*) i Otwartego Dostępu (ang. *Open Access*), chociaż i to może nie wydawać się nowatorskie dla osób, których zainteresowania w tym zakresie są pogłębione.

Drugi rozdział jest najważniejszą częścią publikacji. Otwiera go zarysowanie historii terminu *information literacy*, jego definicji, modeli zastosowania oraz przyjętych międzynarodowych deklaracji. Opisano także występowanie IL w różnych kontekstach (nauka, praca, uczenie się przez całe życie) oraz terminy pokrewne (np. myślenie krytyczne). Rozdział drugi analizuje również kulturowy kontekst *information literacy*. Przywoływano także termin 'kultura informacyjna' jako jeden z wariantów tłumaczenia terminu *information literacy* na języki narodowe, w tym polski. Autorzy starają się pokazać jak historyczne, językowe, kognitywistyczne, socjologiczne i socjo-poznawcze zmiany w ostatnich dekadach wpłynęły na bibliotekoznawstwo i informację naukową jako dyscyplinę nauki, na zmiany sposobu myślenia, na zachowania i potrzeby użytkowników informacji. Autorzy zdają się opowiadać za postrzeganiem *information literacy* jako osobnej dyscypliny nauki³. Piszą bowiem, że o ile bibliotekoznawstwo i informacja naukowa skupiają się na selekcji, organizowaniu, przechowywaniu, rozpowszechnianiu i wykorzystywaniu informacji, o tyle *information literacy* to połączenie tej dyscypliny z pedagogiką, ponieważ dodatkowo koncentruje się na nauczaniu i uczeniu się. Rozdział drugi przedstawia także omówienie innych *literacies*. Jest ich w literaturze przedmiotu wiele i ciągle proponowane są nowe terminy mające obrazować umiejętne posługiwanie się informacją w różnych kontekstach (np. *digital literacy*, *legal literacy*, *workplace literacy*, *visual literacy* czy *financial literacy*). Ta część książki stanowi ciekawe zestawienie *literacies* 15 lat po pierwszej próbie charakterystyki koncepcji pokrewnych czy

³ Po raz pierwszy pisali o tym S. Webber i B. Johnston. Por. B. Johnston & S. Webber (2006). As we may think: Information literacy as a discipline for the information age. *Research Strategies*, 20, 108–121.

pochodnych *information literacy*, przedstawionej przez D. Bawdena⁴. Z punktu widzenia tytułowej Nauki 2.0 istotne wydają się nowe *literacies*, bezpośrednio powiązane z nauką i szkolnictwem wyższym, tj. *scientific literacy*, *academic literacy* i *data literacy*. Dodałabym jeszcze *methodological literacy*, termin zaproponowany przez J. Steinerovą w 2013 r., przez autorów recenzowanej książki z jakichś powodów pominięty⁵. Większość omówionych przez autorów koncepcji to elementy zestawu umiejętności określanych jako „umiejętności / kompetencje 21. wieku” lub *transferable skills*. Drugi termin jest bardziej atrakcyjny, więc częściej używany w literaturze przedmiotu. W rozdziale drugim zawarto także krótkie zarysowanie współczesnych „problemów” informacyjnych, takich jak: nadmiar informacji, *information anxiety*, zarządzanie własną informacją i wiedzą. Wysoko rozbudowane kompetencje informacyjne wydają się być ułatwieniem w radzeniu sobie z tymi zjawiskami.

Rozdział trzeci, według mnie najmniej nowatorski, poświęcono transformacjom IL spowodowanymi także rozwojem technologii. Pokazano przekształcenie oferty edukacyjnej bibliotek od szkolenia bibliotecznego do edukacji informacyjnej. Opisano włączanie edukacji informacyjnej do programów studiów. *Information literacy* przeanalizowano z perspektywy pedagogiki oraz bibliotekoznawstwa i informacji naukowej. Edukacja informacyjna wyszła poza biblioteki, stając się istotnym elementem w wielu sferach życia. Milowym krokiem było pojawienie się WWW, a później Web 2.0 i mediów społecznościowych, co ściśle łączy się z nadmiarem informacji i koniecznością pracy na dużych zbiorach danych. Kompetencje informacyjne są w takiej pracy niezbędne. Rozdział czwarty stanowi podsumowanie. Przedstawia analizę powiązań między Nauką 2.0 a edukacją informacyjną w wielu kontekstach. Prześledzono charakter IL, podkreślono kluczowe trendy i wyzwania. W rozdziale tym autorzy skupiają się na bibliotekach akademickich jako głównym miejscu, w którym można rozwijać edukację informacyjną w służbie nauki. Jak piszą autorzy, Otwarta Nauka, Nauka 2.0 i Open Access powstały pod wpływem potrzeby ułatwienia dzielenia się i udostępniania wiedzy jako dobra publicznego (s. 132), czyli, podobnie jak w przypadku edukacji informacyjnej, ich celem jest ułatwienie dostępu do informacji. Jeśli nauka zmienia się pod wpływem Web 2.0, również edukacja informacyjna skierowana do badaczy powinna ulegać modyfikacjom. Zasady dostosowywania edukacji informacyjnej do potrzeb badaczy są takie same jak w przypadku potrzeb każdej innej grupy społecznej czy zawodowej, inna jest kontekstualizacja. W rozdziale czwartym omówiono także założenia edukacji informacyjnej skierowanej do badaczy. Podniesiono trudną i wrażliwą kwestię wysokiej samooceny badaczy, którym trudno znieść krytykę i którzy nieprzyzwyczajeni są do odgrywania roli uczących się. Jak słusznie zauważają autorzy książki, „people do not become information literate by osmosis” (s. 137), dlatego tak ważne jest nieustanne dokształcanie się. Edukacja informacyjna łączy się także z umiejętnościami komunikacji naukowej. A biblioteki akademickie powinny również w tej dziedzinie rozwijać swoją ekspertyzę, aby móc pomagać badaczom. Nauka 2.0 wymogła na bibliotekach zaoferowanie innowacyjnych usług i zajęcie się nowymi zadaniami. Autorzy pokazują te zmiany na przykładzie nowego nazewnictwa bibliotekarskich stanowisk (m.in. wymieniane na s. 144–147: *scholarly communication librarian*, *embedded librarian*, *data librarian*, *research informationist*)⁶.

Autorzy konkludują stwierdzeniem, że bez wątpienia Nauka 2.0 jest przyszłością edukacji informacyjnej, która ma służyć wsparciu w potrzebach informacyjnych badaczy. Odwołują się też

⁴ Por. D. Bawden (2001). Information and digital literacies: A review of concepts. *Journal of Documentation*, 57(2), 218–259.

⁵ Por. J. Steinerová (2013). Methodological Literacy of Doctoral Students – an Emerging Model. In: S. Kurbanoglu; E. S. Grassian; D. Mizrachi; R. Catts & S. Spiranec (eds.). *Worldwide Commonalities and Challenges in Information Literacy Research and Practice*. European Conference, ECIL 2013, Istanbul, Turkey, October 22–25, 2013. Revised Selected Papers (pp. 148–154). Switzerland: Springer.

⁶ Bardzo dobre źródło obserwacji zmian w nomenklaturze bibliotekarskich stanowisk i potrzeb bibliotek akademickich stanowi także amerykańska lista mailingowa LIBJOBS, rozsyłająca oferty pracy w bibliotekach akademickich głównie Stanów Zjednoczonych i Kanady.

do Czwartego Prawa Ranganathana, w którym mowa o oszczędzaniu czasu czytelnika. Nie od dziś wiadomo, a badania empiryczne to potwierdzają, że rozwinięte kompetencje informacyjne pozwalają zaoszczędzić dużo czasu podczas pracy naukowej.

Zasadnicze rozdziały dopełnia obszerna (21 i pół strony) bibliografia oraz indeks, który jest bardzo przydatny, zważywszy, że większość treści w książce to przytaczanie różnych terminów i konceptów w zmieniających się kontekstach.

Podsumowując, trzeba powiedzieć, że publikacja porządkuje terminy związane z koncepcją Nauki 2.0 i edukacji informacyjnej w sposób przejrzysty, odwołując się do międzynarodowej literatury tematu.

Jest to praca o charakterze przeglądowym, niecałkowicie nowatorska, ale na pewno przydatna. Autorzy przytoczyli wiele definicji czy hipotez, podkreślając te, z którymi się zgadzają i które przyjmują w swojej pracy, jednak trudno doszukać się ich własnego wkładu definicyjnego.

Recepcja tej książki zależy w dużej mierze od tego, kogo uznamy za potencjalnych jej odbiorców. Niestety, autorzy nie zdefiniowali tej grupy. Jeśli mają to być bibliotekarze akademicy (nierazko sami prowadzący badania naukowe) lub akademicy, omówienie tematu jest zbyt ogólne i powierzchowne (ta grupa tę wiedzę już posiada, a przynajmniej powinna posiadać). I nie wnosi wiele więcej niż przegląd literatury i uporządkowanie terminów wcześniej już wprowadzonych w różnorodnych publikacjach. Nie można nie odnieść wrażenia, że wszystko to już zostało omówione wcześniej w literaturze przedmiotu. Fakt, że głównie w pojedynczych artykułach, a niniejsza publikacja całą tę literaturę zgromadziła i usystematyzowała.

Jeśli natomiast odbiorcą mają być studenci czy inne osoby szeroko związane z nauką i szkolnictwem wyższym, które po raz pierwszy stykają się z zagadnieniem edukacji informacyjnej i Nauki 2.0 lub też takie, które chciałyby uporządkować swoją wiedzę, a także poznać międzynarodową literaturę tematu, warto polecić im lekturę tej publikacji.

Książka na pewno ma swój wkład także w rozwój badań. Wydaje się, że zapoczątkowała ona duże, międzynarodowe, porównawcze badania prowadzone od 2014 r., dotyczące umiejętności korzystania z danych badawczych przez pracowników naukowo-dydaktycznych i doktorantów. Ich rezultaty będą ogłoszone na corocznej konferencji ECIL (European Conference on Information Literacy) we wrześniu 2017 r.

Zuzanna Wiorogórska
Biblioteka Uniwersytecka w Warszawie

Nadesłano: 1 kwietnia 2017.

Wśród zagranicznych książek

Jacek Wojciechowski

Przeglądając ofertę publikacyjną, chciałoby się częściej trafić na wypowiedzi możliwie odkrywczę i zarazem napisane frapująco, jednak takie zespolenie przytrafia się rzadko. Interesującym opiniom przeszkadza już to myślowy bałagan lub/ oraz toporny sposób eksplikacji, albo znów sprawna narracja ukrywa (oraz odkrywa) intelektualne niedoróbki.

Żeby więc coś wybrać, kryteria nie mogą być jednolite. Jeśli zatem bodaj jakaś jedna cecha korzystnie wyróżnia tekst, to warto skupić na nim uwagę – udając, że niedostatków nie ma, lub że są mało ważne. No i tak też postępuję.

Uroki indeksacji [****]

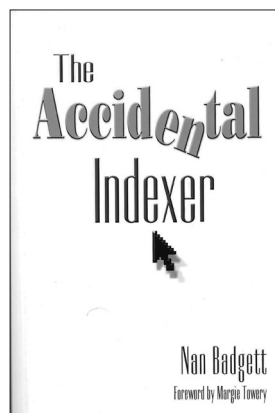
Nan Badgett (2015). *The accidental indexer*. Medford: Information Today Inc., pp. 221. ISBN 978-1-57387-514-1

Indeksowanie, w ujęciu najszerszym (tak w tej książce napisano), polega na penetrowaniu treści przekazów i tworzeniu rejestrów albo map, syntetycznych wyznaczników treściowych, umożliwiających pełne lub wybiórcze, ale nielinearne korzystanie. W tak ogólnym zakresie mieści się więc także tworzenie bibliografii oraz katalogowanie – z tą różnicą, że bibliografia stwierdza istnienie i rejestruje jakiegokolwiek *wszystko*, a katalog lokalizuje to, co jest w danym *miejscu*¹ – ale także indeksowanie w węższym sensie, czyli budowanie wykazów zawartości treściowej dla materiałów pojedynczych albo zgrupowanych razem.

I o takim właśnie, wąsko rozumianym indeksowaniu pisze amerykańska autorka, referując zasady postępowania, często zbieżne bądź wręcz tożsame z katalogowaniem i bibliografowaniem, oraz tworząc charakterystykę takiej specjalności lub wręcz zawodu – bo tak to w amerykańskich warunkach wygląda.

W polskich: niekoniecznie. Nie ma bowiem nawet odpowiedniej nazwy w żadnym rodzimym leksykonie, toteż umownie proponuję termin *indeksator(ka)*, w skojarzeniu z określeniem *inspirator(ka)*. Jest zatem jasne, że o takiej profesji mało kto słyszał i o samej indeksacji wie się niewiele albo jeszcze mniej. Ale ktoś to jednak robi, skoro indeksy istnieją, jakkolwiek wykonawstwo bywa często specjalnością równoległą, albo dodatkową, wobec zawodu prymarnego.

Natomiast ważność nie podlega dyskusji. Trudno dzisiaj zaakceptować monografie lub podręczniki, pozbawione indeksów (niestety są takie), niezależnie od spisów treści. Które też zresztą bywają często rozbudowane do dwóch albo nawet trzech poziomów sygnalizacji szczegółowej².



¹ J. Woźniak-Kasperek (2015). Bibliografia a katalog biblioteczny – dyskusja o pojęciach i terminach. *Przegląd Biblioteczny*, (4), 521–523.

² R. Shimura, S. Yamada, B. Umino, Sh. Toda, K. Kageura (2016). *The structural characteristic of the Japanese paperback book series* [online]. In: J. Sun; Q. Zhu, Ch. Khoo; S. Ou (eds.) *A-LIEP 2016. Proceedings of the 7th*

Autorka sygnalizowanej tu książki, Nan Badgett, uważa że to jest zawód pełnoprawny i wysoce wyspecjalizowany oraz: że to jest jej zawód. Bo też dokonania ma w tej profesji rzeczywiście znaczące i liczne.

Pisze więc o zawodowym indeksatorstwie nie tylko ze znanstwem, ale też z ogromnym zaangażowaniem oraz z sympatią. To fajnie, kiedy ktoś lubi to, co robi! Tę książkę zaś traktuje jako monograficzny podręcznik zawodowy, możliwie kompletny, miejscami zresztą nawet ponadkompletny oraz jako zachęcające wprowadzenie nowych adeptów do tej profesji. Żeby zaś wzmocnić oddziaływanie, oprócz modelowego indeksu książki własnej, dodaje doniesienie z sondażu opinii 74 amerykańskich indeksatorów oraz – na końcu książki – rozmaite załączniki, tak nowym jak i rutynowanym indeksatorom praktycznie przydatne.

Można też pozazdrościć autorce swady i narracyjnego talentu: tekst czyta się znakomicie. Jest przy tym na ogół wysoce kompetentny – chociaż nie obywa się bez trywializmów: niestety w zawyżonej ilości.

Badgett przyznaje się do ukończenia studiów INB, lecz w tekście nie nawiązuje bezpośrednio do bibliotekarstwa, ani do bibliografowania, bądź do katalogowania. Odpowiednie relacje występują wprawdzie, ale czysto przedmiotowe.

Proces tworzenia indeksów (podobnie jak katalogów i bibliografii) określa jako wysoce intelektualny i wieloaspektowy. Oprócz bowiem ustalania rejestru terminów oraz lokalizującej *mapy* w tekście – w materiałach digitalnych często nie ma stron numerowanych, toteż indeksy muszą być nieco inne – trzeba dodatkowo przewidzieć oczekiwania czytelników tekstu i odbiorców indeksu. A to są często różne role i odmienne nastawienia. Poza tym nierzadko bywają konieczne uzgodnienia z autorami, którzy są we współpracy trudni oraz kiepsko sprawdzają się w roli indeksatorów tekstów własnych.

Już choćby dlatego automatyzacja procesów indeksacji nie może być bezgraniczna: udział człowieka jest nieodzowny. Charakterystyczne, że podobną opinię w odniesieniu do katalogów przedstawił Piotr Malak³. To zaś akurat jeden z powodów, dla którego indeksy oraz sieciowe wyszukiwarki nie są wobec siebie tożsame.

Komputer mianowicie wyławia jedynie identyczne terminy i wybrane zbieżne określenia. Nie wnika natomiast w intencje, nie analizuje znaczeń, nie rozróżnia synonimów ani homonimów, a w fazie finalnej często tworzy indeksowe longiery. To nie jest więc prawdziwe i pełne indeksowanie. Za to jest niezastąpiony jako podmiotowe narzędzie pracy i w realizacji kontaktów nie ma alternatywy.

Autorską charakterystykę profesji indeksatorskiej można odnieść częściowo także do bibliografowania lub do katalogerstwa – z tą głównie różnicą, że te dwie specjalności trudno byłoby uprawiać w domu. Natomiast podobnie: to jest zawód słabo przez innych rozpoznawany. Wymaga zaś staranności, uwagi oraz refleksyjności, a także wrodzonej skłonności do porządku oraz do organizowania i kategoryzowania. Guru katalogerstwa, Michael Gorman, powiedziałby: no właśnie. Nie od rzeczy jest też chęć czytania. A poza tym trzeba umieć zachować balans pomiędzy postrzeganiem detali, a zdolnością do syntetyzowania. No i przydaje się dobra pamięć.

Zawodowe przygotowanie indeksatorów jest urozmaicone. Wśród przepytanych, 33% to osoby po studiach INB, 19% po edytorstwie, inni zaś reprezentują specjalności bardzo zróżnicowane. W USA są na uczelniach specjalne kursy, ale np. do indeksowania materiałów online nie przygotowuje nikt. Najważniejsza jest więc praktyka i rozumne samokształcenie. Myślę, że dokładnie to samo potwierdzą doświadczeni katalogerzy i bibliografowie.

Odrębność polega na tym, że indeksatorzy to przeważnie wolni strzelcy – w USA: 90% – i z tego coś musi wynikać. Mianowicie trzeba odpowiednio zorganizować sobie pracę, oddzielając profesję od tego, co prywatne (to duża trudność, kiedy pracuje się w domu), oraz pogodzić się z samotnym przeważnie

Asia-Pacific Conference on Library and Information Education and Practice. Innovation in Library and Information Science in the Age of Big Data. Nanjing: Nanjing University, 178, 182, 187 [18.12.2016], <http://www.vub.ac.be/BIBLIO/nieuwenhuysen/presentations/2016-11-03-aliep-nanjing/A-LIEP%202016%20Proceedings.pdf>

³ P. Malak (2012). *Indeksowanie treści. Porównanie skuteczności metod tradycyjnych i automatycznych.* Warszawa: Wydawnictwo SBP, 168.

i anonimowym charakterem wykonawstwa: to jest zajęcie dla introwertyków. Zapewniające wprawdzie względną swobodę i niezależność, ale powiązane z kolei z brakiem zagwarantowanej ciągłości oraz z nieustannym polowaniem na zlecenia. To jest czynnik mocno stresujący – tak samo, jak narzucane i nierzadko dramatycznie zmieniane terminy wykonawstwa. Ale, powiada Badgett: da się żyć.

Używając nieco innej terminologii, niż przyjęta w obszarze INB, podkreśla różnice pomiędzy zamkniętą formułą indeksowania – jednego dokumentu – i otwartą: wielu różnych materiałów. Przy zamkniętej, poniekąd folksonomicznej, można tworzyć słowniki własne, a indeks da się opracować przed publikacją. W formule otwartej, powiedzmy że taksonomicznej, trzeba posługiwać się słownikami już istniejącymi, zaś kreowanie indeksów następuje dopiero po sfinalizowaniu produkcji dokumentów. No i wtedy wydawcy radykalnie skracają terminy.

W praktyce indeksatorskiej nadal dominują drukowane książki, oraz – w mniejszym zakresie – czasopisma, także drukowane. Przedmiotem dumy zawodowej są indeksy do monografii naukowych, mimo że w wykonaniu trudne i wymagające elementarnej wiedzy przedmiotowej. Z kolei w podręcznikach akademickich i szkolnych potrzebna jest zgodność z przesłaniami dydaktyki, co nie obywa się bez kłopotów. Znacznie trudniej indeksuje się zaś hipertekstowe e-booki – brak bowiem użytecznych standardów – oraz bazy danych online. Przy czym indeksy do nich są o wiele użyteczniejsze, niż oferta automatycznych wyszukiwarek. No więc trzeba, żeby były.

Tymczasem takie zlecenia trafiają się indeksatorom rzadko, a w odniesieniu do stron WWW – prawie nigdy. Trudno zgadnąć, czy to przez odbiorcze preferencje dla sieciowych wyszukiwarek, czy po prostu ze względu na brak zleceńodawców.

Z podręcznikowej i poradnikowej koncepcji tej książki wynikają różne podpowiedzi praktyczne. Za główną wytyczną zawodowego postępowania Badgett uznaje umiarkowanie, a więc zadowolenie z realizacji zleceń już pozyskanych, oraz dystans wobec ambicjonalnego nastawienia na sukces. Sukcesem jest bowiem już sama możliwość wykonawstwa.

O to zaś trzeba zabiegać stale, podtrzymując jednocześnie kontakty dotychczasowe: taka jest specyfika tej pracy. Zlecenia można przyjmować różne, lecz jednak najlepiej wybrać sobie konkretne obszary przedmiotowe, bo przy takiej specjalizacji łatwiej o zlecenia, a i samo indeksowanie toczy się sprawniej. No właśnie: podobnej dyrektywy do bibliografii oraz do katalogowania, odnieść nie można.

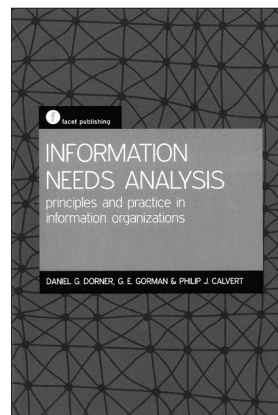
Indywidualny charakter wykonawstwa nie wyklucza współpracy z różnymi partnerami, ewentualnie z podwykonawcami, ale to nie są przypadki częste; natomiast w ogóle jakaś forma kontaktów z innymi indeksatorami, dla wymiany opinii, jest pożądana. A do tych sugestii autorka dorzuca jeszcze rozbudowane wskazówki, jak negocjować racjonalne terminy wykonawstwa oraz możliwie godziwe honoraria – a także wytyczne postępowania autopromocyjnego (marketingowego) i przygotowania osobistych stron WWW w sieci.

Jest poza tym w książce sporo podpowiedzi dla nowych adeptów zawodu, do którego Badgett zachęca, ale z tak detaliczną przesadą, że ta część tekstu razi trywialnością. No i trochę niepokoi opinia, że na dojście do pełnej sprawności indeksacyjnej i zleceniobiorczej potrzeba od 2 do 5 lat praktyki. Ale widocznie tak to jest.

Badanie potrzeb informacyjnych [****]

Daniel G. Dorner; G. E. Gorman; Philip J. Calvert (2015). *Information needs analysis. Principles and practice in information organizations*. London: Facet Publishing, pp. 269. ISBN 978-1-85604-484-4

Niewygodowaną jakością nauk o informacji oraz o bibliotekarstwie, nieomal do granic katastrofy sprowadzają doniesienia o pseudosondażach – wraz



z zerowym autorskim rozeznaniem metodologicznym – które masowo zalewają czasopisma i publikacje zbiorowe z obu zakresów. No pewnie: skoro funkcjonuje u nas aż 20 punktowanych czasopism z tej strefy, oraz drugie tyle imitacyjnych, to ich zawartość czymś trzeba wypełnić. Pojawiają się więc sygnały o rzekomych *badaniach*, gdzie o żadnych technikach badawczych nikt nie ma pojęcia, a sondowane grupy często liczą mniej niż 20 osób. To jest absolutny śmietnik.

Niestety, tak jest nie tylko u nas: to wygląda na śmietnik ogólnobranżowy. Kiedy przeglądam niektóre prezentacje zagraniczne, to zgrzytam zębami. Oto bowiem jakieś osoby próbują robić nam wodę z mózgu i kompromitują te nasze dyscypliny wobec innych specjalistów.

Czasami wdaje się w to wszystko jeszcze i kamuflaż. Oto w gigantycznym doniesieniu z konferencji INB w chińskim Nankinie⁴, referenci z Azji Wschodniej przedstawiali *sondaże* opinii, czasami 15, a niekiedy 9 osób, dokumentując w ten sposób *odkrycia* typu, że koło jest okrągłe. Oszustwo zaś polegało przede wszystkim na wprzęgnięciu w te bzdety całego aparatu przeliczników statystyczny: chi kwadrat, korelacji, wariancji, standardowych odchyień, median itp. Przy dziewięciu respondentach?!

Bezczelność w tym względzie rywalizuje z abnegacją. Porządnych podręczników na temat badawczej metodologii, są wprawdzie setki, a nawet tysiące. Sęk w tym, że należałoby dwa lub trzy przeczytać, ale na to decyduje się mało kto.

Dlatego nie jestem entuzjastą nowych opracowań metodologicznych – wystarczy poczytać to, co już jest – jednak nie twierdzę, że są całkowicie zbędne. Zwłaszcza, jeśli wzbogacają orientację o coś nowego, to warto je zaakceptować. I ostatnio pojawiły się takie: o tyle ciekawe, że nakierowane na analizy obiegu materiałów digitalnych⁵.

Natomiast podręcznik, autorstwa specjalistów z Uniwersytetu w Wellington w Nowej Zelandii, który tu sygnalizuję, już tak dobry nie jest. Rzuca wprawdzie ogólne światło na procedury badawcze, ale bez szczegółów: jedynie z *grubsza*. Za poważny niedostatek uważam zwłaszcza brak odniesienia do konstrukcji zbiorowości badanych, co w praktyce wygląda fatalnie i ta książka tego nie zmienia. Oraz słabością jest również pomylenie metod i technik badawczych, a także całkowite przemilczenie w rejestrze technik – eksperymentu, testu i analizy treści. Poza tym (jak na mój gust) tekst jest niebywale przegadany. Ale może ktoś to lubi.

Autorzy, zapewne świadomi mizerii praktyk rozpoznawczych, zdecydowanie namawiają do prowadzenia rzetelnych sondaży, możliwie systematycznych i metodologicznie poprawnych, zamiast żeby informacyjne nastawienia publiczności szacować według wyrzykowych i pobieżnych obserwacji. Ostrzegają jednak, że nie wolno mylić wyartykułowanych życzeń z rzeczywistym rejestrem potrzeb, które przeważnie są ukryte i niekoniecznie uświadamiane. Ale o sposobach takiego rozróżniania nie ma ani słowa. Jest jeszcze gorzej: w tekście nagminnie ma miejsce mylenie informacyjnego usługodawstwa (zdarzeń zewnętrznych), z (wewnętrzną) istotą odbioru.

Chociaż występuje sugestia, żeby dociekać głębiej i ustalać, co naprawdę wywołuje potrzeby informacyjne – które często bywają nieuświadamiane lub uświadamiane mylnie. Otóż wymaga to analizy szerszych kontekstów towarzyszących procesom informowania. Ale takie objaśnienie jest puste.

W tekście podpowiada się, żeby przedmiotem sondaży byli nie tylko użytkownicy informacji, ale także oferenci i pośrednicy oraz materiały informacyjne kierowane do obiegu. Bardzo to naciągana propozycja i obawiam się, że bezużyteczna. Nie napisano bowiem, jak tego dociekać, a zalecenie żeby

⁴ J. Sun; Q. Zhu, Ch. Khoo; S. Ou, eds. (2016). *A-LIEP 2016. Proceedings of the 7th Asia-Pacific Conference on Library and Information Education and Practice. Innovation in Library and Information Science in the Age of Big Data*. Nanjing: Nanjing University, 178, 182, 187 [18.12.2016], <http://www.vub.ac.be/BIBLIO/nieuwenhuysen/presentations/2016-11-03-aliep-nanjing/A-LIEP%202016%20Proceedings.pdf>.

⁵ S. Hoffman, ed. (2016). *Dynamic research support for academic libraries*. London: Facet Publishing; M. Górska, A. Wandel, red. (2016). *Metody i narzędzia badań piśmiennictwa cyfrowego I jego użytkowników*. Wrocław: Wydaw. Uniwersytetu Wrocławskiego.

analizować środowiskowe normy i wartości, oraz relacje konsumentów informacji z ich otoczeniem, jest tak ogólnikowe, że w praktyce pozbawione treści. Tak więc w konstrukcji badawczych założeń, autorzy nikomu w żadnym stopniu nie pomogą.

Przedstawiona w książce modelowa propozycja postępowania badawczego nie jest przesadnie odkrywczą, ale też do odkrywania pozostaje niewiele. Schemat wygląda tak oto: przygotowanie > zebranie danych > analiza > ocena wyników > raport > wykorzystanie. Trudno mieć do tego zastrzeżenia.

W fazie przygotowań zaleca się staranne przemyślenie celu i formuły badań, oraz dookreślenie potencjalnych zbiorowości (osób) do przebadania, a także wybór metody (w rzeczywistości: techniki) postępowania oraz naszkicowanie przebiegu całego przedsięwzięcia. Natomiast w raporcie autorzy chcą widzieć krótki opis procesu badawczego, jak również charakterystykę rezultatów, sprowadzonych do relacji przyczynowo-skutkowych. To również są sugestie trafne.

Z kolei opis doboru zbiorowości próbnych do badań jest zbyt ogólnikowy. Poza podziałem na dobór losowy oraz celowy, żadnych praktycznych porad nie widać. Nie ma choćby słowa o warunkach reprezentatywności, ani o różnych sposobach losowania, jak też o wariantach (według równych kwot lub statystycznej zależności cech) wyboru celowego. To fatalnie, bo właśnie brak znajomości reguł doboru próby pozbawia sensu większość sondaży.

Sugestia etycznego postępowania w procesie badawczym, jakkolwiek ogólnie słuszna, wypada infantylnie: same frazesy. Natomiast ma sens zalecenie, żeby do oceny przebiegu i efektów badań, zapraszać ekspertów niezależnych, więc spoza grona badaczy. Trafna jest również opinia, że wiarygodność wyników rośnie, jeżeli ktoś inny powtórzy badanie i osiągnie rezultaty zbieżne.

Wśród zalecanych sposobów postępowania jest wskazówka, żeby wykorzystywać już istniejące dane – o zasobach, o użytkownikach oraz o usługach informacyjnych – zwłaszcza zarejestrowane w formie elektronicznej. No dobrze, ale co na to globalne przepisy o tajności danych indywidualnych? W Nowej Zelandii też wszak obowiązują. Brak także szczegółowych wytycznych do takich rozpoznań, mimo że istnieje specjalna technika: analiza dokumentów. Jest tylko ostrzeżenie, że rejestr informacji o udzielonych usługach nie wyczerpuje całego repertuaru transmisji treści informacyjnych.

W książce znalazły się bliższe charakterystyki niektórych innych technik badawczych, jednak wyraźnie dodane do tego, co ewentualny badacz powinien już wiedzieć z innych źródeł. Mam wątpliwości, czy to dobry pomysł. W końcu każdy, kto przystępuje do badań, chciałby mieć porady i wskazówki w jednym materiale, a nie żeby szukać ich w rozproszeniu.

Jako podstawową technikę badań obiegu informacji, nie bez racji, wskazano ankietę. I tak jak napisano: jest tania, łatwa do szerokiego zastosowania oraz nie stresuje respondentów, ponieważ ma przeważnie charakter anonimowy. Jednak nadaje się głównie do analizy postaw i zachowań uświadamianych; jest poza tym trudna do zweryfikowania pod względem wiarygodności, a przy rozsyłaniu, zwrotność bywa mierna. Jest też kłopot z pytaniami, które muszą funkcjonować samodzielnie, bez dodatkowych objaśnień – zamknięte są często dla respondentów zbyt sztywne, mało elastyczne, z kolei zaś odpowiedzi na pytania otwarte kiepsko poddają się grupowaniu i kwantyfikacji. I wprawdzie to o ankietach wiadomo nie od dzisiaj, ale takie przypomnienie ma sens.

Dla poprawy zwrotności i wiarygodności wypowiedzi ankietowych autorzy podpowiadają staranny wstęp wyjaśniający oraz maksymalną klarowność kwestionariusza. Do osiągnięcia w trybie wcześniejszego pilotażu weryfikacyjnego. Rzeczywiście. A już inna sprawa, że nie znam opisu badań ankietowych, który nie zawierałby takich właśnie wskazówek.

Przy charakterystyce wywiadu – w sumie produktywny – również dostrzegam niedobór sugestii szczegółowych. To potwierdza, że jest adresowana do osób już z tą techniką trochę zaprzyjaźnionych.

Oryginalność autorskich podpowiedzi polega na tym, że zaleca się dobieranie do wypowiedzi już to osób, dobrze w temacie zorientowanych, bądź: realizację wywiadów w grupach 8–12 osób, bo współobecność innych respondentów bywa pożyteczna. Otóż takich sugestii jeszcze nie spotkałem.

Przy czym, jak napisano – grupowa zgodność opinii wcale nie jest konieczna. Natomiast w wywiadzie grupowym badacz musi funkcjonować jako moderator oraz motywator wypowiedzi, jakkolwiek oczywiście bez sugerowania poglądów. Zresztą ogólne zalecenie też jest takie, żeby nie ingerować w reakcje respondentów, a nawet nie uprzedzać wcześniej, jaki ma być temat rozmowy. Przy grupowych wywiadach autorzy doradzają, żeby nie tworzyć więcej niż cztery odrębne grupy, za to każdą z nich dobrać ze względu na jakąś jedną wspólną cechę, inną w każdej z grup, a potem uznać je za badane zmienne przyczynowe. To ma sens.

Wywiad zawsze generuje kłopoty rejestracyjne i tak też je zasygnalizowano. Rejestracja przebiegu foniczna albo filmowa peszy respondentów, pisemna zaś – absorbuje badacza. Co można nieco zredukować, wspierając się pytaniami zamkniętymi, ale w tekście takiej podpowiedzi nie ma. Natomiast podzielałam wyartykułowane zastrzeżenia wobec wywiadów telefonicznych lub internetowych: bez kontaktów bezpośrednich to jednak nie są rozmowy pełnowartościowe.

Jest też kilka uwag na temat obserwacji, lecz widać wyraźnie brak autorskiego entuzjazmu, choćby w uwadze, że to jest wyjątkowo czasochłonna technika badań. No to co? Tak jest i już.

Przy obserwacji zewnętrznej (a istnieje też uczestnicząca obserwacja wewnętrzna w grupie), zdaniem autorów trzeba, żeby obserwator był co najmniej tolerowany przez obserwowanych. No pewnie: inaczej mogliby go zbić. Za to o zasadach obserwowania nie ma w tekście nic. Jest tylko powtórzona uwaga, że rejestracja pisemna odrywa obserwatora od przebiegu zdarzeń, a znów mechaniczna – deprymuje wszystkich. Tak jest rzeczywiście, jednak są sposoby minimalizowania tych niedogodności, ale nic o tym nie napisano.

Również powierzchowne są wskazówki, jak analizować rezultaty badań. Jest zalecenie zakodowania danych, numerycznego lub wagowego, a następnie – relatywnych (przyczyny-skutki) zestawień wyników, dla weryfikacji zakładanych pierwotnie hipotez. Ma temu służyć garść przeliczników statystycznych, zasygnalizowanych w książce, lecz tylko ponazywanych. Zabrakło objaśnień, jak je obliczać.

Bardziej użyteczne są porady, jak skonstruować finalny raport z badań. Generalnie: pisemny. Ewentualnie z wariantem w blogu, bądź na stronie WWW. Czasami potrzebna jest jeszcze wersja werbalna, lecz wtedy skrócona i uwolniona od detali.

Autorzy nie bez racji zalecają, żeby wariant lub warianty raportu dopasować do adresatów, którzy bywają rozmaici, a ważne żeby ich takie sprawozdanie przekonało. Najpełniejszej wersji wymagają eksperci. Trzeba więc przedstawić im koncepcje, zastosowane techniki badawcze, dodać rejestr przedmiotowego piśmiennictwa, no i oczywiście zreferować wyniki, a także zasugerować możliwości wykorzystania. Sposób prezentacji ma wpływ na odbiór, toteż warto wprowadzić tabele, szkice i wykresy ilustracyjne, jakkolwiek w racjonalnym wymiarze. Zasugerowano, że prezentacja konkluzji ogólnej nie powinna przekraczać 10% całego raportu. Akurat co do tego miałbym wątpliwości: to wysoce arbitralna opinia. A w dopowiedzi pojawiła się jeszcze uwaga, żeby zreferować też stopień reprezentatywności wyników. Słuszna, tyle że w całej książce nie ma o tym najmniejszej wzmianki.

Nie mam wątpliwości, że osoby słabo zorientowane w badawczej metodologii, nie wyniosą z lektury tej książki istotnych korzyści praktycznych: jest od tego zbyt odległa. Szkoda, bo wydawnictwo ma dobrą renomę, więc szansa na szeroki rezonans była, zatem i na poprawę jakości sondażowych praktyk. Tymczasem to jest raczej próba pobudzenia do staranniejszych przemyśleń metodologicznych tych badaczy, albo kandydatów na badaczy, którzy podstawowe pojęcie o badaniach już mają. A czy pożyteczna? To zależy od nastawienia każdego czytającego. Nowozelandzcy autorzy wyszli z założenia, że może tak być.

Co nam z neuronauki kognitywnej [***]

Richard Passingham (2016). *Cognitive neuroscience. A very short introduction*. Oxford: Oxford University Press, pp. 136. ISBN 978-0-19-878622-1

Ograniczone pole poznawcze nauki o informacji oraz obszerniejsze, ale też limitowane, nauki o bibliotekarstwie – a nawet zwiększone, jednak również *ogrodzone*, nauki o komunikacji, narzuca konieczność zwracania się ku innym dyscyplinom ościennym, jeżeli rozwój ma mieć miejsce. W szczególności trzeba posiłkować się ustaleniami psychologii, bo wszystko co przynależy do tych trzech dyscyplin, finalizuje się w procesach recepcji oraz myślenia. Dlatego niekiedy – i tym razem też – sygnalizuję tu publikację również z obszaru psychologii, jeżeli dostrzegam możliwość odniesień.

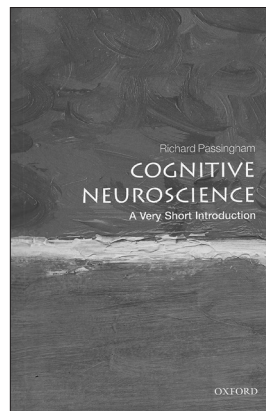
Nie jest to łatwe, nie istnieje bowiem jednolita wiedza psychologiczna. Jej dynamiczny rozwój zmierza w kilku kierunkach, niekoniecznie spójnych, a bywa że sprzecznych. Już dawno temu nurt psychodynamiczny oddzielił się od behawiorystycznego, a potem powstało kolejne ujęcie, mianowicie poznawcze (kognitywistyczne), lecz na tym nie koniec. Od dłuższego czasu istnieje autonomiczna subdyscyplina, nazwana *neuronauką*, która zajmuje się funkcjonowaniem mózgu – dzięki nowym technologiom badawczym: z możliwością obserwowania reakcji neuronalnych na impulsy zewnętrzne oraz wewnętrzne. Przy czym także już w obrębie neuronauki wyodrębnił się nurt kognitywistyczny, traktujący mózg tak, jakby to był komputer, więc mechanistycznie⁶. Ale niezależnie od mnóstwa wątpliwości, można upatrywać w tym nastawieniu nowych, nieznanych dotychczas objaśnień funkcjonowania umysłu, również w trakcie odbioru treści, także informacyjnych. Jak dalece zaś rzeczywiście tak jest, to już inna sprawa.

Właśnie dlatego sygnalizuję tu książkę Richarda Passinghama – to emerytowany profesor psychologii z Uniwersytetu Oksfordzkiego – który zreferował stan wiedzy, dostępnej za sprawą kognitywnej neuronauki. Oczywiście również jest mechanistycznie nastawiony do zjawisk mentalnych, bo to taki nurt, ale bez skrajności, jakie zdarzało mi się już nieraz spotykać. Tak czy inaczej, więcej tu opinii dyskusyjnych, niż bezwarunkowo akceptowalnych, ale możliwych odniesień do procesów komunikowania oraz informowania, da się tu dokonać sporo.

Natomiast ta – mimo wszystko: skrajność – sądów nie całkiem przystaje do znakomitej, zazwyczaj umiarkowanie nastawionej – serii przeglądowej, w ramach której to omówienie wydano. No i określenie *very short* mija się z prawdą: rzeczywiste rozmiary tej publikacji, obficie ilustrowanej, są znaczne.

Autor narzeka na skutki długotrwałego obciążenia psychologii nawisem behawiorystycznym, który forsował wnioskowanie o procesach psychicznych w oparciu o zewnętrzne obserwowanie reakcji na bodźce. Lecz na dobrą sprawę, neuronauka kognitywna, chociaż istnieje od pół wieku, także nie postępuje inaczej, chociaż więcej *widzi*. Polega wszak na obserwowaniu obiegu krwi w mózgu, aktywowanego pod wpływem pobudzeń neuronalnych – i to właśnie można teraz zobaczyć – ale to nie jest nasz spór. Ważniejsze, że jednak nie widać wszystkiego i wobec tego jeszcze nam daleko do wielu fundamentalnych ustaleń.

W kolejnej negacji Passingham nie jest odosobniony. Ma otóż miejsce zakwestionowanie kartezjańskiego oddzielenia myślenia (ducha) od umysłu (ciała), bo to w aktualnym ujęciu ma być taki sam proces organiczny, jak każdy inny. Szczerze mówiąc, trochę robi się nieswojo na myśl, że rozumowanie jest fizjologicznie podobne do trawienia, lub do oddawania moczu – ale oczywiście taka wątpliwość nie jest nauką.



⁶ J. Cacioppo (2016). Social neuroscience. In: R.J. Sternberg, S.T. Fiske, D.J. Foss (eds.) *Scientists making a difference. One Hundred Eminent Behavioral and Brain Scientists Talk about their Most Important Contributions*. New York: Cambridge University Press, 46.

Ważniejsze, że przy takim ujęciu procesów umysłowych nie ma objaśnienia, w jaki sposób tworzy się i uzupełnia świadomość – dla procesów informacyjnych przecież fundamentalna. Że już nie zapytam o generowanie refleksji nowych.

Świadomość jest rezultatem przesyłanych impulsów, ale przetworzonych, powiązana często z relacją bezpośrednią: woda to nie jakieś H₂O, lecz to, co się pije⁷. Wymaga zatem interpretacji odbieranych sygnałów, a komputery (maszyny) nie interpretują⁸. Tak więc neuronaukowy model kognitywistyczny umysłu, traktujący mózg tak jak komputer, nie wystarczy do pełnych objaśnień procesów komunikowania, a zwłaszcza nie potrafi scharakteryzować ich finalizacji, czyli: wzbogacania świadomości. Na takie rozpoznanie trzeba jeszcze poczekać i to niedobrze dla wiedzy o komunikowaniu, o informowaniu, oraz siłą rzeczy: o bibliotekarstwie.

Passingham potwierdza, że neuronauka koncentruje się na obserwacji przebiegu procesów w mózgu i ustala, gdzie następuje oraz przebiega aktywizacja. Określa więc miejsca: obszary reakcji na różne rodzaje bodźców. Każdy taki obszar jest wyspecjalizowany, stosownie do podłączonych sensorów, ma zadania własne, istnieje więc swego rodzaju mapa takich ośrodków, które następnie łączą się z innymi, tworząc razem ujęcia (składniki myślenia) scalone. Tego jednak już nie *widać*.

No i to akurat jest dla nas ważne. Zdaniem autora – a wie, co mówi – neuronaukowe charakterystyki procesów mentalnych nie wkraczają w obszary przyczyn ani następstw, dlatego właśnie nie odnoszą się do świadomości. Otóż tego nam brakuje, ponieważ sfera komunikowania oraz informacji ma charakter świadomościowy.

W książce ma jedynie miejsce nawiązanie, zresztą fragmentaryczne, do podświadomości. Muszę tylko jeszcze dopowiedzieć, że z innych analiz psychologicznych wygenerowała się opinia, że zachowaniami rządzi (inicjują, motywują, intensyfikują) emocje⁹. Otóż Passingham potwierdza, że nieraz przed reakcją decyzyjną (dotyczącą zachowania), widać w mózgu przedprocesową aktywizację, którą określa jako nieświadomą, albo przedświadomą. A to może stanowić dowód, że na przykład dla przyjęcia treści, także informacyjnych, ważne jest poprzedzające nastawienie uczuciowe (właśnie więc aktywizujące): już to na niespodziankę, lub na znaczny pożytek odkrywczy, bądź na potwierdzenie oczekiwań, indywidualnie ważnych. Można z tego wysnuć praktyczne dyrektywy postępowania informacyjnego.

Inny obszar autorskich doniesień, o którym trzeba tu wspomnieć, dotyczy wielosensoryczności, zatem i multisensoryczności, przyjmowanych sygnałów. W komunikacji oraz w informowaniu to jest zagadnienie sposobu transmisji treści: jednorodnego jak przez pismo, bądź wielojęzykowego jak przez hipertekst. A także następstw i rezultatów, czyli przetworzenia tego, co zostało pierwotnie przyjęte do pamięci roboczej, a następnie zachowania części w pamięci trwałej. To nie przebiega automatycznie¹⁰ i jest zależne od sposobu transmisji właśnie, a dokładniej: od rodzaju przejmowanych sygnałów, bo niektóre mogą oddziaływać rozpraszająco. Tymczasem odbiorowi sprzyja natężenie uwagi i skupienie się na przyjmowanych treściach – zresztą nie na wszystkich, ale na wybranych, ze względu na ograniczoną pojemność i krótkotrwałość pamięci roboczej. Potem, z zespolenia tych wybranych treści, już w pamięci trwałej, powstaje przetworzona wiedza indywidualna¹¹.

⁷ J. Zimmermann (2015). *Hermeneutics. A very short introduction*. Oxford: Oxford University Press, 12.

⁸ L. Floridi (2014). *The 4th revolution. How the infosphere is shaping human reality*. Oxford: Oxford University Press, 135; J. Zimmermann (2015). *Hermeneutics. A very short introduction*. Oxford: Oxford University Press, 19, 136.

⁹ A. Damasio (2016). Feelings and decisions. In: R.J. Sternberg, S.T. Fiske, D.J. Foss (eds.) *Scientists making a difference. One Hundred Eminent Behavioral and Brain Scientists Talk about their Most Important Contributions*. New York: Cambridge University Press, 12.

¹⁰ A. Baddeley (2016). Working memory. In: R.J. Sternberg, S.T. Fiske, D.J. Foss (eds.) *Scientists making a difference. One Hundred Eminent Behavioral and Brain Scientists Talk about their Most Important Contributions*. New York: Cambridge University Press, 120.

¹¹ J. Zimmermann (2015). *Hermeneutics. A very short introduction*. Oxford: Oxford University Press, 129.

Dla charakterystyki procesu informowania i jego newralgicznych składników to są objaśnienia ważne. Tymczasem Passingham zgłasza istotną wątpliwość: twierdzi mianowicie, że nie ma dowodów na pojemnościowe oraz czasowe ograniczenia pamięci roboczej, a zatem pośrednio – na wybór treści do trwałego zapamiętania. Z drugiej strony nie twierdzi, że zapamiętuje się absolutnie wszystko i sam pisze o skupieniu, rozproszeniu oraz o skuteczności zapamiętywania. Jak to pogodzić? Powstaje w ten sposób niepotrzebny mętlik.

Prawie całe piśmiennictwo psychologiczne jest zgodne co do tego, że w pamięci roboczej daje się pomieścić niewiele sygnałów naraz i to na czas krótki – dlatego jest też nazywana krótkotrwałą albo nietrwałą. Już tu więc ma miejsce selekcja. Potem następuje zaś kolejny wybór treści, do zapamiętania dłuższego: w pamięci trwałej. Otóż charakter tych wyborów pozostaje w związku z rodzajem przyjętych sygnałów.

Konkretnie od tego, jak dalece umożliwiają skupienie się na przejmowanych treściach, co ułatwia językowa jednorodność sygnałów, np. werbalnych albo pisemnych. Multisemiotyczność, wielojęzyczność – co Passingham potwierdza – wydłuża czas odbioru (zatem pamięć robocza przyjmuje mniej treści), tak jak i nieprzewidywalność lub nieklarowność tych treści, ponieważ w transmisji różnych sygnałów są jednak, częściowo, zaangażowane te same neurony. Tym bardziej zaś dekoncentrują działania uboczne, impulsy rozproszone i niejednorodne. Dlatego rezultaty odbioru przekazów hipertekstowych, w porównaniu do pisemnych, są inne¹². Poza tym skupienie się na jakimś segmencie treści, inspiruje skojarzenia zbieżne, a więc łączy z treściami, które zostały już wcześniej zinternalizowane. Przy rozproszeniu uwagi tak nie jest. Otóż wszystko to trzeba odnieść do odbioru informacji – już to, wobec tego, ograniczonego, lub pogłębionego.

Dostrzegam też jeszcze inny obszar refleksji, sprowokowanych przez ten tekst wprost lub pośrednio, dla informowania i komunikowania moim zdaniem ważny. Mianowicie od dawna wiadomo, że dla różnych reakcji na komunikacyjne sygnały istnieją w mózgu odrębne ośrodki *sterujące*. Passingham sygnalizuje ośrodek mówienia Broca i rozumienia mowy Wernickego, potwierdzając ich autonomię w ten sposób, że przy odbiorze sygnałów z ekranu, w tych akurat ośrodkach żadnych reakcji nie ma.

Podobna specjalizacja ma miejsce w komunikacji piśmienniczej. Ośrodek Exnera steruje pisaniem, zaś ośrodek Dejerine'a czytaniem – jakkolwiek o nich i w ogóle o piśmiennictwie nie ma w tej książce ani słowa. Szkoda, ale trudno. Otóż ta odrębność reagowania na sygnały, niezależnie od późniejszej syntetyzacji przyjętych różnymi sposobami treści, nakazuje ostrożność i wstrzeźliwość w orzekaniu o zlewaniu się różnych form komunikacji, o wypieraniu jednych przez drugie, albo też o ich uniwersalności, czyli o wszechstosowaniu. Tymczasem dostrzegam powszechną nieostrożność.

Z punktu widzenia aktualnej wiedzy o tym, jak pracuje umysł, takie ewentualne scalenie komunikacyjnych form nie znajduje żadnego uzasadnienia. Odrębność jest faktem, więc i współistnienie też musi być faktem.

Osobna grupa ważnych wątpliwości dotyczy w tej książce poglądu na temat ewolucji bądź niezmienności mentalnych reakcji na sygnały komunikacyjne, oraz dostosowania umysłu do nowych form komunikowania. Opinie znawców są bowiem w ogóle rozbieżne.

Nie tak dawno temu, w dwóch światowych bestsellerach eseistycznych, lansowano pogląd, że nowe technologie komunikacyjne podporządkowują sobie umysły, wywołując w mózgach nieodwracalne zmiany fizjologiczne¹³. Tyle tylko, że obaj autorzy nie spotkali się z psychologią nawet w tramwaju, a wypowiedzi oparli na zasłyszanych sądach obiegowych, bądź na własnych wyobrażeniach.

Ale Passingham jest uznanym ekspertem, a jednak też opowiada się za przypuszczeniem, że to otoczenie – także komunikacyjne – wpływało i wpływa na ukształtowanie i funkcjonowanie umysłu.

¹² J. Nielsen (2010). *iPad and Kindle reading speeds* [online]. NN/g Nielsen Norman Group [18.12.2016], *iPad and Kindle reading speeds*.

¹³ N. Carr (2010). *The shallows. What the Internet is doing to our brain*. New York: W. W. Norton & Company, 31, 34, 44, 50; D. Tapscott (2010). *Cyfrowa dorosłość. Jak pokolenie sieci zmienia nasz świat*. Warszawa: Wydawnictwa Akademickie i Profesjonalne, 77, 176.

I to nie tylko w trybie ewolucyjnego przystosowania, ale wręcz przez wygenerowanie w mózgu odpowiednich struktur neuronalnych i funkcji fizjologicznych.

Z tego punktu widzenia, neguje ideę gramatyki generatywnej Noama Chomskiego, który uważał, że neuronalne struktury w mózgu, zaangażowane w porozumiewanie się ludzi, nie są nabyte, ale wrodzone z natury¹⁴. Jednak ani Passingham, ani cała neuronauka, nie dysponują żadnymi argumentami, że tak jest lub nie jest. Nie ma bowiem świadectw z przeszłości, ewentualnie określających dawniejszy kształt oraz formy przeobrażania się mózgu. Możliwa adaptacyjna elastyczność nie oznacza wszak rekonfiguracji całego mózgu od nowa. Więc dopiero przyszłe obserwacje neuronaukowe będą mogły ustalić, jak to rzeczywiście jest, lecz na to potrzeba dziesięcioleci.

Tymczasem, tak jak Chomsky, wielu znanych psychologów twierdzi, że to jednak wrodzone predyspozycje umysłu stały i stoją u podstaw tworzenia się różnych form komunikacji. Czyli, że zanim te formy powstały – i może dlatego powstały oraz powstają właśnie takie – neuronalne predyspozycje już były. Wybitny znawca psychologicznych aspektów komunikowania, Stanislas Dehaene, sformułował jednoznaczna opinię na temat genealogii komunikacji pisemnej. Mianowicie, że to pismo dostosowało się do neurofizjologii mózgu, a nie odwrotnie¹⁵. Podobnie sądził Janusz Dunin sugerując, że pismo zostało odkryte, a nie wynalezione¹⁶.

Co zaś miałyby z tego wynikać? Ano to, że formy komunikowania, więc również informowania, mogą zmieniać się i wzbogacać oraz rozwijać, ale w ogólnych granicach, wyznaczonych przez podstawowe i w *rdzeniu* niezmiennie struktury ludzkiego umysłu. Co wykluczałoby eliminację jednych komunikacyjnych form podstawowych przez drugie, a nawet redukowałoby ich identyczne zastosowanie do tego samego celu – skoro mają w umyśle inną neuronalną matrycę konfiguracyjną.

Eliminacja byłaby prawdopodobna tylko w wypadku przebudowy fizjologii umysłu, a niczego takiego nie zaobserwowano. Przeciwnie: w powiązanej strukturze DNA nie ma śladu zmian w okresie natężonych przeobrażeń komunikacyjnych¹⁷. Dlatego trzeba z ogromnym dystansem przyjmować łatwowierne a bezdowodowe deklaracje o nadchodzących repulsa-cjach form komunikacji, oraz o dominacji określonych wariantów informowania. O czym, kiedy, w jakim wymiarze, dla jakich celów i jak głęboko? Bez dobrze uzasadnionej argumentacji, lepiej nie dawać temu wiary. Bo zamiast nauki, rozprzestrzenia się ble ble.

Poza tym zaś raz jeszcze odniosę się do zjawiska świadomości, z którym nie tylko neuronauka radzi sobie marnie, bo to jest abstrakt, a nie *wytwór* neurofizjologiczny. Jednak przecież istnieje i co najważniejsze: wieńczy procesy komunikacyjne, w tym oczywiście również informacyjne. Oraz *jakieś* powiązania w neurofizjologią myślenia niewątpliwi ma.

Świadczy o tym choćby odrębność, w ogólnej pamięci trwałej, przechowywania przetworzonych sygnałów z rzeczywistości realnej, w pamięci epizodycznej, a sygnałów treściowych (komunikacyjnych) w pamięci semantycznej. Późniejsze scalenia nie mają nic do rzeczy. To jest (umownie) taka różnica, jak pomiędzy indywidualną wiedzą, a indywidualnymi umiejętnościami.

Nie byłoby takiego rozróżnienia bez strukturalnych różnic w pamięci, a zatem i w świadomości. Ale wiarygodnych objaśnień (na razie?) nie ma. Dlatego wiedza o komunikowaniu oraz o informowaniu jest wciąż daleka od kompletności.

Nadesłano: 18 grudnia 2016.

¹⁴ N. Chomsky (1982). *Zagadnienia teorii składni*. Wrocław: Ossolineum, 14–15, 33–34.

¹⁵ S. Dehaene (2009). *Reading In the brain*. New York: Penguin/Viking, 3.

¹⁶ J. Dunin (1998). *Pismo zmienia świat*. Warszawa, Łódź: Wydaw. Nauk. PWN, 18.

¹⁷ L. Floridi (2014). *The 4th revolution. How the infosphere is reshaping human reality*. Oxford: Oxford University Press, 95.

Przegląd polskich nowości wydawniczych

Anna Stanis

Biblioteka Uniwersytecka w Warszawie

Metody i narzędzia badań piśmiennictwa cyfrowego i jego użytkowników. Pod red. Małgorzaty Góralskiej i Agnieszki Wandel. Wrocław: Wydaw. Uniwersytetu Wrocławskiego, 2016, ss. 220. Seria „Książka–Dokument–Informacja”. Acta Universitatis Wratislaviensis; 3698. ISBN 987-83-229-3514-9

Szybki rozwój zasobów cyfrowych i technologii związanych z ich stosowaniem był inspiracją do dyskusji nad metodologią badań zasobów sieciowych i sposobami ich wykorzystania. Efektem tej dyskusji jest publikacja zawierająca 15 tekstów podzielonych na trzy części tematyczne:

- (1) Cyfrowy wymiar bibliologii i informatologii
- (2) Zasoby elektroniczne – procesy i perspektywy
- (3) Użytkowanie i użytkownicy piśmiennictwa cyfrowego.

W części pierwszej zawarto teksty dotyczące: gatunków tekstu piśmiennictwa cyfrowego, narzędzi analizy zawartości dokumentów cyfrowych, wykorzystania technik komputerowej wizualizacji w badaniach naukometrycznych, badań z zakresu typografii kinetycznej.

W części drugiej znajdują się teksty dotyczące praktycznych aspektów organizowania i wykorzystywania zasobów cyfrowych. Jednym z narzędzi badań bibliotek cyfrowych jest bibliomining czyli „korzystanie z eksploracji danych w badaniach bibliotecznych zasobów danych”. Omówiono model bibliomingu zwracając uwagę na ograniczenia jego wykorzystania przez bariery techniczne, prawne i organizacyjne. Praktyczny aspekt wykorzystania zasobów cyfrowych zawiera również tekst poświęcony ocenie działań bibliotek państwowych wyższych szkół zawodowych w zakresie digitalizacji od 1998 do 2014 r. Innym przykładem praktycznego ujęcia tematu jest tekst poświęcony zarządzaniu zbiorami hybrydowymi na przykładzie Biblioteki Głównej Politechniki Warszawskiej. Ciekawym projektem opisanym w tej części publikacji jest utworzenie w 2014 r. na Politechnice Wrocławskiej Centrum Wiedzy i Informacji Naukowo-Technicznej, jednostki ogólnouczelnianej realizującej zadania naukowe, badawcze i usługowe. W strukturze Centrum znajduje się Biblioteka Klasyczna i Biblioteka Elektroniczna stanowiąca podstawę systemu biblioteczno-informacyjnego.

W części trzeciej zawarto badania potrzeb i zachowań użytkowników piśmiennictwa cyfrowego a w szczególności badania wykorzystania e-booków w bibliotekach, badanie czytelnictwa płatnych treści przez wydawców prasowych, wykorzystanie naukowych zasobów cyfrowych przez studentów w kontekście badań zagranicznych. Większość badań nad e-bookami w bibliotekach to badania prowadzone w USA. Kompleksowe badania w tym zakresie realizowane są od 2010 r., osobno dla bibliotek szkolnych, publicznych i akademickich. Temat wykorzystania naukowych zasobów cyfrowych przez studentów podejmowano w wielu krajach ale największe zainteresowanie wzbudził w Indiach (zrealizowano tam połowę wszystkich badań). Stosowaną techniką badawczą najczęściej była ankieta. W tej części pracy znajdują się także tematy dotyczące recepcji polskich czasopism humanistycznych w literaturze światowej oraz badania nad charakterem komunikacji naukowej w sieci (bibliometria, webometria, nowe narzędzia oceny badaczy i ich prac naukowych) na przykładzie komunikacji naukowej pracowników Politechniki Wrocławskiej.

Fedorowicz-Kruszewska Małgorzata, Cyrklaff Magdalena J.: *Media w środowisku osób zagrożonych wykluczeniem społecznym*. Toruń, Wyższa Szkoła Bankowa, 2016, ss. 179. ISBN 978-83-943316-1-0

Celem pracy jest opisanie zjawiska wykluczenia społecznego ze szczególnym uwzględnieniem mediów, które dzięki informacji pozyskiwanej za ich pośrednictwem, mogą przyczynić się do zapobiegania zjawisku wykluczenia społecznego. Pojęcie mediów zostało określone „jako nośniki informacji (środki transmisji), do których zalicza się m. in. gazety, książki, filmy, radio, sieci informatyczne”. Autorki osobne rozdziały pracy poświęciły mediom analogowym i cyfrowym. W tej pierwszej grupie znalazły się opisy publikacji brajlowskich, książek dotykowych, tyflografiki („grafiki dla niewidomych i słabo widzących”). W drugiej grupie znajdują się opisy audiobooków, filmów i audycji telewizyjnych dla niesłyszących oraz technologii wspomagających, definiowanych przez autorki za M. Fedorowiczem jako „różnorodne urządzenia elektroniczne, które pozwalają ludziom niepełnosprawnym na samodzielność lub tę samodzielność zwiększają”.

Charakterystykę współcześnie funkcjonujących mediów poprzedzają rozdziały poświęcone definicjom pojęcia wykluczenia społecznego (ekskluzji społecznej), przyczynom tego zjawiska, kryteriom kategoryzacji grup zagrożonych wykluczeniem, sposobom zapobiegania wykluczeniu. Omówiono także odzwierciedlenie przeciwdziałania wykluczeniu cyfrowemu w europejskich i polskich dokumentach. Pracę uzupełnia opis organizacji i inicjatyw wspierających rozwój i wykorzystanie mediów, a także obszerna bibliografia, wykaz polecanych zasobów elektronicznych, indeks przedmiotowy.

Włodarczyk Bartłomiej, Woźniak-Kasperek Jadwiga: *Język Haseł Przedmiotowych Biblioteki Narodowej – od analizy dokumentu do opisu przedmiotowego* Warszawa, Wydaw. SBP, 2017, ss. 184. Seria „Nauka, Dydaktyka, Praktyka”; 173. ISBN: 978-83-64203-76-3

W ostatnich latach języki haseł przedmiotowych, w tym język haseł przedmiotowych Biblioteki Narodowej (JHP BN), ulegały licznym przeobrażeniom. Przyczyną tych zmian są m. in. szybko zmieniające się technologie informacyjne, a co za tym idzie – nowe oczekiwania użytkowników bibliotek. JHP BN jest jednym z najbardziej znaczących języków haseł przedmiotowych w Polsce. Tworzony od lat 50. XX w., stosowany był przez wiele lat w bibliografii narodowej, a także do opracowania zbiorów w większości bibliotek publicznych i niektórych bibliotekach naukowych. Prace nad JHP BN w dotychczasowym kształcie zostały zawieszona pod koniec 2015 r. Od stycznia 2016 r. trwają prace nad przekształceniem JHP BN w język deskryptorowy (Deskryptory Biblioteki Narodowej, DBN). W ciągu kilkudziesięciu lat tworzenia JHP BN ukazało się wiele prac o tym języku głównie o charakterze warsztatowym i poradnikowym. Omawiana praca jest – jak piszą we wprowadzeniu autorzy – „dokumentem o charakterze źródłowym, który jest zapisem stanu JHP BN na dzień 30 kwietnia 2014 r.”. Ten jedyny w swoim rodzaju dokument jest cenną pomocą zarówno dla bibliotekarzy praktyków, którzy opracowują jeszcze dokumenty w JHP BN jak i dla teoretyków JIW oraz studentów kierunków związanych z bibliotekoznawstwem i informacją naukową.

Monografia składa się z trzech części: Część pierwsza w opracowaniu Jadwigi Woźniak-Kasperek omawia niezwykle istotną problematykę analizy dokumentu poprzedzając powstanie opisu przedmiotowego. Autorka zwraca uwagę na błędy popełniane przy analizowaniu dokumentów, które identyfikowana są z mankamentami JHP. Wyjaśnia często mylone pojęcia katalogowania przedmiotowego i katalogowania rzeczowego, różnice pomiędzy pojęciem treści a pojęciem przedmiotu i omawia podstawowe etapy opracowania przedmiotowego. W tej części znalazły się też podstawowe wiadomości o rodzajach przedmiotu, jego ujęciu i formie.

Część druga w opracowaniu Bartłomieja Włodarczyka poświęcona została charakterystyce JHP BN, również w ujęciu historycznym. Autor tej części omówił szczegółowo różne rodzaje tematów

w JHP BN, określniki i hasła rozwinięte. Zwrócił uwagę na niejednorodność form jednostek leksykalnych w JHP BN, szczególnie w przypadku tematów ogólnych a także na nie zawsze konsekwentnie ustalane relacje pomiędzy tematami. Wskazując na istotną wartość już istniejących opisów przedmiotowych autor podkreśla, że znajomość zasad tworzenia i stosowania JHP BN jest konieczna zarówno w opracowaniu dokumentów znajdujących się w bibliotekach, jak również w przypadku podejmowania decyzji o jego przekształcaniu.

Praca zawiera także obszerną część trzecią z przykładami dokumentów opracowanych z wykorzystaniem JHP BN. Całość uzupełnia bogata bibliografia załącznikowa i wykaz skrótów.

Nauka o informacji w okresie zmian. Informatologia i humanistyka cyfrowa. Pod red. Barbary Sosińskiej-Kalaty przy udziale Marii Przystek-Samokowej i Zuzanny Wiorogórskiej. Warszawa: Wydaw. SBP, 2016, ss. 276. *Miscellanea Informatologica Varsoviensia*; vol. 8. Seria „Nauka-Dydaktyka-Praktyka”; 169. ISBN 978-83-64203-67-1

Omawiana pozycja jest pokłosiem trzeciej konferencji z cyklu „Nauka o informacji (informacja naukowa) w okresie zmian”, która odbyła się w dniach 11–12 maja 2015 r.¹, zorganizowanej przez Zakład Systemów Informacyjnych Instytutu Informacji Naukowej i Studiów Bibliologicznych Uniwersytetu Warszawskiego pod patronatem ISKO PL (International Society for Knowledge Organization, The Polish Chapter). Tematem przewodnim konferencji, a zarazem prezentowanej publikacji, było określenie relacji między nauką o informacji (informatologią) a dynamicznie rozwijającą się dziedziną humanistyką cyfrową (e-humanistyką). W książce przedstawiono 17 artykułów powstałych na bazie referatów wygłoszonych na konferencji, pogrupowanych w czterech częściach tematycznych: część I *Problematyka, źródła informacji i narzędzia badawcze humanistyki cyfrowej z perspektywy informatologii*; część II *Badacze nauk humanistycznych w cyfrowym świecie*; część III *Projekty badawcze, wyniki badań humanistyki cyfrowej oraz dostęp do nich w przestrzeni cyfrowej*; część IV *Biblioteki akademickie wobec wyzwań humanistyki cyfrowej*.

W najobszerniejszej części pierwszej znalazło się osiem artykułów poświęconych: analizie wielkich zasobów danych (*big data*), z wykorzystaniem zaawansowanych technologii komputerowych (Piotr Gawrysiak), problematyce humanistyki cyfrowej w piśmiennictwie informatologicznym (Barbara Sosińska-Kalata), archiwizacji e-informacji (Małgorzata Góralska), bibliotekom cyfrowym (Widad Mustafa El Hadi), repozytorium danych (Bożena Bednarek-Michalska), prawu autorskiemu (Teresa Świątkowska), społeczności programistycznym (Sebastian Grabowski) i denatyzacji rzeczywistości (Łukasz Iwasiński).

Część druga zawiera trzy artykuły: Remigiusza Sapy, który omówił zarządzanie indywidualnymi kolekcjami informacji naukowej w humanistyce w świetle piśmiennictwa anglojęzycznego; Marzeny Świogoń na temat technologii w zarządzaniu i dzieleniu się wiedzą i informacją; Małgorzaty Jaskowskiej o wskaźnikach altrimetrycznych w systemie wartościowania prac naukowych.

W części trzeciej umieszczono trzy artykuły: dwa pierwsze – Anny Kamińskiej i Darii Rzeplieli – poświęcono różnym prezentacjom informacji. Trzeci tekst, Marka Nahotki, dotyczy nowych gatunków (cybergatunków) tekstów użytkowych (nieliterackich).

Publikację zamyka część czwarta z trzema artykułami, które poświęcono bibliotekarzom i bibliotekom akademickim, stawianym wobec wyzwań humanistyki cyfrowej (Katarzyna Materska, Lilianna Nalewajska, Anna Myśliwska).

¹ Sprawozdanie z tej konferencji ukazało się w *Zagadnieniach Informacji Naukowej* 2015 nr 1, 145–151.

Nauka o informacji. Pod red. Wiesława Babika. Warszawa, Wydaw. SBP, 2016, ss. 712. Seria „Nauka-Dydaktyka-Praktyka”; 174. ISBN 978-83-64203-82-4

Niezwykle obszerna i szczegółowa monografia obecnego stanu nauki o informacji (informacji naukowej, informatologii), ze szczególnym uwzględnieniem aktualnego rozwoju i przyszłości tej dziedziny nauki. Przeznaczona dla osób zainteresowanych problematyką szeroko rozumianej informacji i działalności informacyjnej. W 24 tekstach, które można uznać za osobne rozdziały, grono pracowników naukowych większości naszych uczelni wyższych, przedstawiło szereg zagadnień z zakresu działalności informacyjnej w Polsce i za granicą. Publikacja jest najobszerniejszym w Polsce kompendium wiedzy o informacji traktowanej z punktu widzenia zarówno teoretycznego, jak i praktycznego. Zminimalizowano kontekst historyczny na rzecz niezbędnych aktualizacji i przedstawienia bieżących problemów informatologii.

W poszczególnych rozdziałach omówiono zagadnienia informacji jako dyscypliny naukowej, jej terminologię, problemy lingwistyczne, obszary badań, metodologię. Zwrócono uwagę na naukę o informacji jako przedmiotu w kontekście kształcenia akademickiego pracowników informacji. Poruszono również zagadnienia organizacji i zarządzania informacją i wiedzą, architektury i prezentacji informacji.

Osobne rozdziały poświęcono środowisku informacyjnemu człowieka, działalności informacyjnej, źródłom informacji, komputerowemu przetwarzaniu informacji, usługom informacyjnym, wyszukiwaniu informacji i zachowaniom informacyjnym. Autorzy mają nadzieję, że przedstawiony w tekstach obraz informacji będzie inspiracją do utworzenia wspólnej badawczej płaszczyzny nad informacją naukową do badań zespołowych o charakterze interdyscyplinarnym.

Monografia pełni również funkcję podręcznika akademickiego informatologii, każdy rozdział uzupełniono bogatą i wyczerpującą bibliografią tematu. Na uwagę zasługuje imponująco rozbudowany aparat pomocniczy składający się z: informacji o autorach; Wykazu ważniejszych skrótów i akronimów; wyboru bibliografii, encyklopedii, informatorów i opracowań; słownika angielsko-polskiego; Indeksu osobowego; Indeksu przedmiotowego.

Morawiec Barbara, Maria: *Biblioteki cyfrowe. Tworzenie, zarządzanie, odbiór*. Gliwice: Helion, 2016, ss. 176. ISBN 978-83283-2658-3

W książce przedstawiono wyniki porównania koncepcji architektów bibliotek cyfrowych z potrzebami i oczekiwaniami użytkowników tych bibliotek. Porównanie to poprzedzone zostało omówieniem problemów terminologicznych i wskazaniem celów i funkcji bibliotek cyfrowych. Pracę oparto na przykładach dwóch bibliotek cyfrowych: Małopolskiej Biblioteki Cyfrowej i Jagiellońskiej Biblioteki Cyfrowej, omawiając genezę i rozwój każdej z nich, a także realizację projektów digitalizacyjnych. Obie biblioteki współpracują ze sobą, wykorzystują oprogramowanie dLibra i udostępniają podobne obiekty cyfrowe.

Autorka przedstawiła również ewaluację obu bibliotek, opierając się na analizie zawartości ich stron internetowych w oparciu o kryteria: identyfikacji wizualnej, podstawowych informacji, struktury strony głównej i architektury informacji. Osobny rozdział poświęciła koncepcji architektów bibliotek cyfrowych określając podstawowe funkcje przez nich pełnione w pracy w bibliotece cyfrowej, organizację pracy z uwzględnieniem kontaktu użytkownika biblioteki z architektem.

Ostatni rozdział poświęcono użytkownikom obu bibliotek. W omówionym tu badaniu zastosowano technikę ankiety na podstawie, której scharakteryzowano profile użytkowników i opracowano ich typologię. Wśród oczekiwań użytkowników obu poddanych analizie bibliotek dominują: nieograniczony dostęp do wszystkich publikacji cyfrowych, wiarygodność informacji, możliwość pobierania i kopiowania całości publikacji, sprawne funkcjonowanie i organizacja biblioteki.

Pracę uzupełnia obszerna bibliografia, indeks osobowy, kwestionariusz przeprowadzonej ankiety, tabele z przedstawieniem dokładnych wyników badań użytkowników obu bibliotek. Publikacja stanowi pracę magisterską autorki powstałą w Instytucie Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego pt. *Architekci i użytkownicy bibliotek cyfrowych – konfrontacja koncepcji i oczekiwań*.

Nadesłano: 17 maja 2017.

Międzynarodowa Konferencja „Wizualizacja informacji w humanistyce”

Toruń, 23–24 marca 2017

W dniach 23–24 marca 2017 r. w Toruniu odbyła się Międzynarodowa Konferencja „Wizualizacja informacji w humanistyce”, zorganizowana przez Instytut Informacji Naukowej i Bibliologii Uniwersytetu Mikołaja Kopernika, przy udziale Biblioteki Uniwersyteckiej UMK (<http://wizualizacja.strikingly.com/>). Konferencja była znaczącym wydarzeniem promującym nowoczesne metody pracy oraz technologie w nauce, a adresowanym przede wszystkim do przedstawicieli nauk społecznych i humanistów.

Wizualizacja informacji i danych jest nowym polem badań, wkraczającym obecnie w każdą sferę ludzkiej działalności: od badawczo-rozwojowej po biznesową, społeczną i artystyczną. Obszar ten wymaga praktycznej wiedzy o tym, jak graficznie „opanować” ogromne zbiory danych, opisujących dany aspekt rzeczywistości i jak wykorzystać tę wiedzę w dzisiejszym społeczeństwie. Tam gdzie jest dużo danych i informacji, powstaje zapotrzebowanie na ich przedstawienie w formie skondensowanej i zrozumiałej dla odbiorcy. Poza naukami przyrodniczymi, ściśle opartymi na empirii, zbieranie danych w wyniku procesów obserwacji czy symulacji coraz częściej dotyczy również nauk społecznych (np. ekonomia, socjologia, politologia) i humanistycznych (takich jak historia, bibliologia i informatologia, językoznawstwo itp.). Z takimi wyzwaniem jak rzetelna analiza dużych zbiorów danych w oparciu o wizualizację borykają się specjaliści bibliometrii i ekonometrii, socjologowie, psychologowie czy administracja uczelni. Postać cyfrowa zasobów naukowych wymusza również na humanistach stosowanie metod wizualizacji, a rozwój Web 2.0 wpływa na zwiększenie zainteresowania tymi metodami. Z myślą o specjalistach z różnych dyscyplin szeroko zdefiniowano zakres tematyczny konferencji: od mapowania wiedzy i big data poprzez historię wizualną i kognitywistykę po estetyczny aspekt prezentacji wizualnych.

Tytuł konferencji symbolicznie został zawężony do humanistyki po to, żeby podkreślić znaczenie humanistyki cyfrowej we współczesnej nauce, gdzie umiejętności informacyjno-komunikacyjne często decydują o zakresie, rozwoju i promocji badań. W związku z tym, patronat merytoryczny nad konferencją objęło konsorcjum DARIAH-PL (www.dariah.pl). W przyszłości planuje się przygotować podobne konferencje, kierując problematykę na zastosowania wizualizacji w mediach lub biznesie.

W konferencji wzięło udział przeszło 90 osób, z czego połowa – czynnych uczestników (z referatami bądź warsztatami), w tym 16 prelegentów z zagranicy. Pierwszego dnia w dwóch angielskojęzycznych sesjach plenarnych wystąpili zaproszeni goście – krajowi i światowi specjaliści w zakresie problematyki szeroko rozumianej wizualizacji informacji. Dr Raphael Ball, dyrektor Biblioteki Uczelnianej Politechniki w Zurichu, przedstawił historię rozwoju komunikacji naukowej, zaznaczając, że powstanie nowych, nieliniowych form tekstu wymaga od naukowców połączenia tradycyjnej aktywności z intensyfikacją prac nad zarządzaniem danymi, ich strukturalizacją i publikowaniem (*The end of linear text in scholarly communication?*). Prof. Yuriy Holovatch z Ukraińskiej Akademii Nauk wprowadził słuchaczy do rozpowszechnionej obecnie metody badawczej – analiza sieci społecznych (*Social Networks Analysis, SNA*), przywołując ciekawe przykłady symulacji międzynarodowego ruchu lotniczego (*Visualizing and understanding complex systems: from transport networks to online role-playing games*). Trzeci prelegent – znany w kraju i na świecie kognitywista i fizyk, prof. Włodzisław Duch, pokazał jakie procesy i struktury w mózgu są odpowiedzialne za przetwarzanie obrazu i czytanie wizualizacji (*Visualization of mind inside the brain*). Dr Jan Kozłowski z Ministerstwa Nauki i Szkolnictwa Wyższego przekazał słuchaczom unikalne informacje o wizualizacji wiedzy

w systemach naukometrycznych, które działają obecnie w różnych krajach (*Visualisation for science policy*). Dr Suzana Pasanec Preprotić z Uniwersytetu w Zagrzebiu (Wydział Sztuk Artystycznych) opowiedziała, jak wizualizacja pomaga w efektywnym nauczaniu studentów sztuki oprawy ksiąg i tegumentologii (*Visualization framework in teaching bookbinding*). Prof. Panayiota Polydoratou natomiast argumentowała, że wizualizacja informacji jest ważnym elementem systemów Open Data (*An introduction to data journals, linking data and visualizations in publications in an integrated way*). Dr Monika Szetela, teolog z wykształcenia, wykładowca w Wyższej Szkole Kultury Społecznej i Medialnej w Toruniu, za pomocą oryginalnej analizy wizualnej zaprezentowała interpretację postaci oblubieńców z biblijnej Pieśni nad Pieśniami (*An attempt to a new reading of the Bible – the visualization of the dynamics of the relationship, between the Bridegroom and the Bride in the Song of Songs*). Abdelaziz Blilid z Uniwersytetu w Lille (Francja) pokazał serię wizualizacji sieciowych odnoszących się do geograficznej analizy społeczności Berberów rozproszonej w różnych krajach Azji (*Visualization of Hypertext Links of Transborder Cultural Identity: A Case of Berber Diaspora*).

W czterech sesjach równoległych przedstawiono 15-minutowe referaty. Na krótsze, 7-minutowe wystąpienia przydzielono osobną sesję Ignite Talk, w której przedstawiono dziesięć takich prezentacji. Wszystkie prezentacje, włącznie z relacją filmową za zgodą autorów zostały opublikowane na stronie internetowej¹: <http://www.wizualizacja.nauki.umk.pl/konferencja/>.

Sesja „Wizualizacja źródeł i danych” obejmowała referaty wprowadzające w problematykę wizualizacji danych wielkoskalowych pod względem interpretacji (*Wizualizacja danych z różnych źródeł: interpretacyjne pułapki normalizacji na przykładzie narodowych bibliograficznych baz danych*, Emanuel Kulczycki), dostosowania narzędzi do badanego zagadnienia (*Sieciowe narzędzia do wizualizacji wiedzy w popularyzacji humanistyki cyfrowej*, Radosław Bomba) oraz w dużej mierze w zakresie studiów nad książką (*Wizualizacja w opracowaniu książki dawnej*, Jolanta Gwioździk; *Wizualizacje cyfrowe w badaniu utworów literackich*, Katarzyna Staniuk; *Współczesna polska bibliologia – Słownik Pracowników Książki Polskiej w liczbach i w perspektywie sieciowej*, Magdalena Paul).

Sesja „Piękno a funkcjonalność w prezentacji informacji”, ukierunkowana na aspekt estetyczny z jednej strony, a funkcjonalność – z drugiej, cieszyła się dużym zainteresowaniem słuchaczy. Dyskutowano na niej o infografikach i zawartym w nich pierwiastku manipulacji (*Prezentacja danych na wykresach i infografikach – pomiędzy informacją a manipulacją*, Anna Seweryn), szumie informacyjnym (*Infografiki w humanistyce – przekaz informacji czy szum informacyjny?*, Zbigniew Osiński) czy ekologii informacji (*Prezentacja informacji z punktu widzenia ekologii informacji*, Wiesław Babik) oraz modelach estetyki z perspektywy naukowej (*Ze spostrzeżenia w wyobrażenie – kod ikoniczny komunikatów public relations organizacji*, Anna Waszkiewicz-Raviv) i empirycznej (*Czy informacja może być piękna? Wizualizacja informacji a sztuka*, Małgorzata Lebda i Barbara Krasieńska).

Nowe możliwości przed humanistą cyfrowym otwiera analiza obrazów i/lub ich sekwencji w procesie wizualizacji – temat ten omawiany był w trzeciej sesji równoległej „Obrazy statyczne i ruchome w wizualizacji”. Andrzej Radomski zaprezentował obiecującą metodę analizy obrazów przy wykorzystaniu charakterystyk spektralnych okładek czasopism polskich (*Analiza i wizualizacja obrazów na przykładzie badania okładek polskich tygodników społeczno-politycznych*). Stanisław Skórka pokazał techniczne rozwiązania wizualizacji nawigacji w serwisach typu *discovery* (*Wizualizacja nawigacji w serwisach typu discovery*). Z technicznymi aspektami wizualizacji obiektów architektonicznych zapoznał słuchaczy zespół informatyków z Wyższej Szkoły Komunikacji Społecznej i Mediów (*Znaczenie i wpływ doboru parametrów oświetlenia na rzeczywisty odbiór modelowanych obiektów sakralnych*, Zbigniew Chaniecki, Błażej Świątek, Grzegorz Osiński).

W sesji „Wizualizacja wiedzy – aspekty kognitywne i metodologiczne” zgrupowano zagadnienia metodologiczne wizualizacji informacji (*Perspektywy i dobre praktyki w wizualizacji wiedzy*

¹ Strona internetowa projektu NCN 2013/11/B/HS2/03048/ Badanie dynamiki i struktury cyfrowych zasobów wiedzy przy pomocy metod wizualizacji, w ramach którego zorganizowano konferencję.

o jednostkach terytorialnych, Jan Fazlagic), terminologiczne (*Wizualizacja 'wizualizacji' w procesie dydaktycznym: case study przedstawienia terminu dokonane przez studentów Zarządzania informacją INIB UJ*, Małgorzata Janik), jak i aspekty kognitywne w odniesieniu do badań interdyscyplinarnych (*Kognitywistyczny i kognitywny wymiary wizualizacji dla badań interdyscyplinarnych w zakresie science of team science oraz science of science communication*, Tomasz Komendziński; *Struktury wiedzy osobistej jako technika wspierająca uczenie się głębokie, Dorota Ciechanowska; Rola graficznych korelatów przedpojęciowych schematów wyobrażeniowych w wizualizacji wiedzy, Albert Leśniak*).

Dwujęzyczna sesja Ignite Talk łączyła różnorodne tematy, jak i sposób prelekcji: za pomocą prezentacji lub posteru. W sesji wzięła także udział grupa studentów z Uniwersytetu w Thessalonikach (ATEI). Równoległe do Ignite Talk odbyła się ważka debata na temat humanistyki cyfrowej, adresowana do wszystkich pracowników Wydziału Nauk Historycznych UMK, z udziałem kierowników grup roboczych konsorcjum DARIAH-PL. Dzięki niej – mamy nadzieję – technologie wizualizacyjne jeszcze mocniej ugruntowały swoją pozycję w tym środowisku.

Ważnym elementem konferencji były warsztaty w drugim dniu konferencji, ukierunkowane na praktyczne przekazanie wiedzy o współczesnych narzędziach wizualizacyjnych i praktykach graficznej prezentacji informacji. Osiem warsztatów w dwóch turach, przy równoległym wykorzystaniu czterech laboratoriów komputerowych z dwunastoma stanowiskami pozwoliło na zaspokojenie potrzeb większości uczestników. Podczas warsztatów prezentowano narzędzia do analiz bibliometrycznych i sieci społecznościowych: Gephi (Radosław Bomba), Pajek (Ewa Rozkosz), do analiz historycznych – Palladio (Maria Skóra, Dawid Stanny, Veslava Osińska), do eksploracji korpusów tekstowych – CLARIN-PL (Piotr Malak), do wizualizacji streamów Twittera (Marcin Wilkowski), profesjonalnych analiz zbioru obrazów (Andrzej Radomski). Zorganizowano również szkolenie w zakresie tworzenia czytelnych infografik (Anna Sobko i Agnieszka Dura).

Cenną wiedzę na temat tego, jak zaprojektować profesjonalną prezentację multimedialną przekazali przedstawiciele firmy Bain & Company: Kinga Dąbrowska i Rafał Wodzis. Przeprowadzenie tego warsztatu oraz warsztatu naukometrycznego, poprowadzonego przez głównego sponsora konferencji – firmę Clarivate Analytics – w zakresie użytkowania komercyjnego narzędzia InCites, nadały części szkoleniowej konferencji częściowo biznesowy charakter.

Pomimo szeroko zakrojonego zakresu tematycznego audytorium zdominowali polscy bibliolodzy i informatolodzy. Po raz kolejny pokazuje to, jak chłonne na nowe technologie i jak ambitne jest środowisko specjalistów informacji naukowej. Organizatorzy świadomi tego, że tylko niewielka część społeczności bibliologów i bibliotekarzy skorzystała z wykładów i warsztatów postanowili opublikować materiały pokonferencyjne w modelu otwartym „open access” (publikacja jest w trakcie opracowywania) oraz udostępnić w czasopiśmie bibliotekarskim *Biuletyn EBIB*² szczegółowy tutorial, poświęcony współczesnym analizom bibliometrycznym, w których wykorzystywane są narzędzia wizualizacyjne.

Wydaje się, że ta nowoczesna tematyka była dla bibliotekarzy i naukowców inspirująca, wskazywała na nowe trendy pojawiające się na świecie w zakresie przystępnego i zrozumiałego przekazywania informacji i zbiorów danych. Wykłady poszerzyły świadomość społeczności akademickiej na temat prezentowania informacji, a warsztaty wskazywały na innowacyjne narzędzia i metody, jakie można przy tym stosować. Narzędzia i metody bowiem, które do tej pory stosowało środowisko naukowe, mogą teraz być wzbogacone o nowsze, bardziej wyrafinowane, strukturalnie bogatsze, efektywniejsze i plastycznie zróżnicowane. Bibliotekarze mają świadomość, że dzięki tej wiedzy mogą lepiej przygotowywać informacje, wspierać naukowców i administrację uczelni. Uczelnie zaś powinny docenić te

² V. Osińska (2017). *Gephi – nowoczesnym, niezwykle przydatnym narzędziem dla bibliometrów. Instrukcja obsługi (tutorial)*. bibliotekach [online]. *Biuletyn EBIB*, 2017, 173. [10.08.2017], <http://open.ebib.pl/ojs/index.php/ebib/issue/view/128>.

zdolności, ponieważ coraz częściej w zarządzaniu uczelnią wykorzystuje się skomplikowane struktury danych, które trzeba prosto i przekonująco przedstawić różnym słuchaczom, także decydentom i donatorom. Humanistyka cyfrowa ma wiele wymiarów, a wizualizacje są jednym z ciekawszych.

Veslava Osińska

Institut Informacji Naukowej i Bibliologii, Uniwersytet Mikołaja Kopernika w Toruniu

Bożena Bednarek-Michalska

Biblioteka Uniwersytecka, Uniwersytet Mikołaja Kopernika w Toruniu

IV Międzynarodowa Konferencja Naukowa „Nauka o informacji (informacja naukowa) w okresie zmian” Innowacyjne usługi informacyjne (Warszawa, 15–16 maja 2017 r.)

W dniach 15–16 maja 2017 r. w salach Biblioteki Uniwersyteckiej w Warszawie odbyła się czwarta już edycja międzynarodowej konferencji naukowej „Nauka o informacji (informacja naukowa) w okresie zmian”. Jest to konferencja organizowana cyklicznie co dwa lata przez dawny Zakład Systemów Informacyjnych Instytutu Informacji Naukowej i Studiów Bibliologicznych Uniwersytetu Warszawskiego. W tym roku po raz pierwszy zorganizowana została pod szyldem nowej jednostki uniwersyteckiej: Katedry Informatologii Wydziału Dziennikarstwa, Informacji i Bibliologii UW.

Tematem przewodnim tegorocznego spotkania były innowacyjne usługi informacyjne. Wszyscy na co dzień doświadczamy zmian, jakie zachodzą w naszym otoczeniu informacyjnym. Przemianom ulegają kanały komunikacyjne, media, sposoby przekazywania, ale także konsumowania czy wykorzystywania informacji; zmieniają się technologie, kontekst społeczny, oczekiwania i role użytkowników informacji. Usługi informacyjne są jednym z tych fragmentów pola zainteresowań informatologii, o których najbardziej zasadnie można powiedzieć, że znajdują się w okresie zmian.

Zakres tematyczny konferencji, określony przez Komitet Programowy, objął szeroką gamę teoretycznych i praktycznych aspektów użytkowania nowoczesnych narzędzi i metod działalności informacyjnej, organizacji i zarządzania informacją i wiedzą, analizy i oceny nowych technologii wykorzystywanych w społecznym transferze informacji i wiedzy oraz ich recepcji wśród użytkowników. Przyjęliśmy, że debata toczyć się będzie wokół innowacyjnych usług informacyjnych w nauce i edukacji, w biznesie, mediach i administracji publicznej czy w różnego rodzaju instytucjach kultury. W polu naszych zainteresowań znalazły się społeczne i filozoficzne aspekty innowacyjnych technologii i usług informacyjnych, technologie semantyczne w usługach informacyjnych, mobilne usługi informacyjne, media społecznościowe, wizualizacja informacji, innowacje w organizacji wiedzy, architekturze informacji oraz w zarządzaniu informacją, a wreszcie użytkownicy informacji oraz edukacja informacyjna.

W konferencji udział wzięło niemal 100 osób, zarówno prelegentów, jak i słuchaczy. Wygłoszono 56 referatów. Mieliśmy przyjemność gościć wielu specjalistów badających problematykę informatologiczną, wywodzących się z wielu ośrodków zarówno krajowych, jak i zagranicznych, reprezentujących różne pokolenia, różne spojrzenia na dyscyplinę, różne ujęcia badawcze i kierunki jej rozwoju. Dzięki

temu wydarzenie to, jedno z najważniejszych na informatologicznej mapie naszego kraju, odegrało rolę nie tylko forum wymiany wiedzy i doświadczeń, prezentacji nowych koncepcji czy też rozwiązań, lecz także pozwoliło na obserwację dynamiki zmian zachodzących w nauce o informacji, a wreszcie umożliwiło dokonywanie pewnych porównań pomiędzy zainteresowaniami badawczymi naszych koleżanek i kolegów z różnych ośrodków polskich i zagranicznych.

Konferencję rozpoczęło uroczyste otwarcie w Sali Konferencyjnej Biblioteki Uniwersyteckiej w Warszawie. Słowo wstępne wygłosili: prof. dr hab. Barbara Sosińska-Kalata – przewodnicząca Komitetu Programowego konferencji, prof. dr hab. Janusz W. Adamowski – dziekan Wydziału Dziennikarstwa, Informacji i Bibliologii UW oraz prof. dr hab. Wiesław Babik – z Instytutu Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego, reprezentujący stałego honorowego patrona konferencji – International Society for Knowledge Organization, ISKO PL. W otwarciu konferencji wziął też udział prodziekan WDIB UW, prof. dr hab. Dariusz Kuźmina oraz przedstawiciele sponsorów konferencji.

Największa liczba gości zagranicznych przyjechała z francuskiego Uniwersytetu Charlesa de Gaulle'a – Lille 3, współpracującego z Katedrą Informatologii Laboratorium GERiiCO (Groupe d'Études et de Recherche Interdisciplinaire en Information et Communication). Dwa z wystąpień wprowadzających w problematykę konferencji wygłoszone zostały przez reprezentujące tenże ośrodek prof. Widad Mustafę El Hadi, która nakreśliła historyczny kontekst nauki o informacji i dokumentacji, oraz prof. Laurence Favier, która mówiła o ewaluacji informacji w kontekście serwisów społecznościowych. Oba paniom towarzyszyli ich doktoranci i współpracownicy, którzy również brali czynny udział w obradach, wygłaszając w czasie późniejszych sesji równoległych referaty dotyczące zachowań historyków w środowisku cyfrowym (Samia Takhtoukh), etycznych i międzykulturowych wymiarów systemów organizacji wiedzy (Tan Tran), zastosowania crowdsourcingu w humanistyce cyfrowej (dr Amel Fraisse) oraz wpływu zachowań innych użytkowników portali społecznościowych na podejmowane przez nas decyzje (Zhenfei Feng).

Ponadto gośćmi specjalnymi konferencji byli profesorowie: Gerhard Budin z Uniwersytetu Wiedeńskiego (Zentrum für Translationswissenschaft), Mieczysław Muraszkiewicz z Politechniki Warszawskiej, Tibor Koltay (Uniwersytet Eszterházy Károly w Egerze), a także przedstawiciele firm prowadzących komercyjne usługi informacyjne – Maciej Dziubecki (ALEPH Polska) i Christopher Biedermann (Emitel Ltd).

Gerhard Budin mówił o potrzebie rozwoju infrastruktury usług informacyjnych dla nauki i edukacji przez wdrażanie systemów organizacji wiedzy umożliwiających wieloaspektowy i spersonalizowany dostęp do danych badawczych w wirtualnym środowisku badawczym na różnych etapach prowadzenia badań i wykorzystywania ich wyników w edukacji. Mieczysław Muraszkiewicz zachęcał do myślenia o sztucznej inteligencji jako motorze innowacyjności, która implikuje dziś głębokie zmiany w niemal wszystkich obszarach życia społecznego. Tibor Koltay zaproponował spojrzenie na biblioteki akademickie jako ośrodki wspierające rozwój Nauki 2.0. Maciej Dziubecki poświęcił swoje wystąpienie zastosowaniu badań użytkowników w procesie projektowania bibliotecznych systemów informacyjnych, Christopher Biedermann natomiast cyberbezpieczeństwu w Internecie Rzeczy.

Pierwszy dzień konferencji wypełniony był w całości wystąpieniami w języku angielskim. Po zakończeniu pierwszej części, na którą składały się dwie sesje plenarne, moderowane przez prof. B. Sosińską-Kalotę i prof. W. Babikę i obejmujące wymienione powyżej keynote'y, obradowano w czterech międzynarodowych sesjach równoległych: „e-Administracja i e-Biznes”, „e-Kultura”, „e-Nauka” oraz „Systemy informacyjne i ich użytkownicy”.

W pierwszej sesji, prowadzonej przez prof. Remigiusza Sapę (Uniwersytet Jagielloński), problemami administracji publicznej w środowisku cyfrowym zajął się Mariusz Luterek z Wydziału Dziennikarstwa, Informacji i Bibliologii UW, przedstawiając wyniki badań bibliometrycznych, których celem było ustalenie udziału informatologów w rozwoju usług informacyjnych w tym zakresie. O usługach informacyjnych dla biznesu mówili w swych wystąpieniach Erika Janiūnienė z Uniwersytetu w Wilnie i Radosław Lipiński (niezależny ekspert zarządzania wiedzą). E. Janiūnienė opowiadała o ocenie

jakości systemów informacyjnych dla biznesu, R. Lipiński zaś zaprezentował zaprojektowany przez siebie model przepływu informacji, który stanowić powinien narzędzie wykorzystywane do wpięrania procesów zmian w organizacjach.

W sesji poświęconej zjawiskom e-kultury, moderowanej przez prof. Ewę Głowacką (UMK w Toruniu), najwięcej uwagi poświęcono bibliotekom. Emmanuelle Chevry Pébayle (Uniwersytet w Strasburgu) analizowała wpływ obecności bibliotek akademickich w mediach społecznościowych na powstawanie i implementację innowacyjnych usług informacyjnych. Alicja Waszkiewicz-Raviv (WDIB UW) mówiła o wykorzystaniu elementów wizualnych w promocji bibliotek publicznych, Karolina Zawada (Biblioteka UMK w Toruniu) – o dostępnym online repozytorium surowych danych naukowych z zakresu astronomii. Projekt realizowany jest na UMK we współpracy z biblioteką uniwersytecką. Czwartym prelegentem tego panelu był Zhenfei Feng, o którego wystąpieniu była już mowa.

W moderowanej przez prof. Marię Próchnicką (UJ) trzeciej i moderowanej przez prof. Piotra Tafiłowskiego (WDIB UW) czwartej sesji tego dnia mieliśmy okazję wysłuchać, oprócz wystąpień wspomnianych już gości z GERiiCO, także referatu Silvii Cardoso i Patricii de Almeida (Uniwersytet w Coimbrze), które zastanawiały się, czy tagowanie społecznościowe może wzbogacić katalog biblioteki oraz czy może być ono sposobem na podniesienie jakości usług w zakresie indeksowania w porównaniu do aktualnie używanego języka kontrolowanego. Odnośne badania wykonane zostały w bibliotece Escola Superior de Saúde de Santa Maria w Porto, gdzie pracuje jedna z prelegentek. W ostatnim referacie w tej sesji Pablo Benalcazar (Instytut Gospodarki Surowcami Mineralnymi i Energią PAN) omawiał zagadnienia związane z metodami pomiaru i identyfikacji gwałtownie wzrastającej ilości informacji patentowej.

W drugim dniu konferencji, w dwunastu sesjach tematycznych, zorganizowanych w trzech równoległych ciągach, obradowano w języku polskim. Sesje, podobnie jak w pierwszym dniu, moderowane były przez członków Komitetów Programowego i Organizacyjnego konferencji: prof. Mirosława Górnego (Uniwersytet Adama Mickiewicza w Poznaniu), prof. Jacka Tomaszczyka (Uniwersytet Śląski), prof. Barbarę Sosińską-Kalotę (WDIB UW), prof. Katarzynę Materską (Uniwersytet Kardynała Stefana Wyszyńskiego), dr. Marcina Roszkowskiego (WDIB UW), prof. Jadwigę Woźniak-Kasperek (WDIB UW), prof. Piotra Tafiłowskiego (WDIB UW), prof. Dianę Pietruch-Reizes (UJ), prof. Wiesława Babika (UJ), dr. Mariusza Luterka (WDIB UW) i dr. Teresę Święckowską (WDIB UW). Większość sesji składała się z trzech wystąpień i dyskusji.

Pierwszą sesję, pod hasłem „Innowacje w informacji naukowej”, otworzyła Diana Pietruch-Reizes (UJ) referatem poświęconym innowacyjnym usługom informacyjnym w nauce. Marzena Świ-goń (Uniwersytet Warmińsko-Mazurski w Olsztynie) sygnalizowała kierunki zmian zachodzących w komunikacji naukowej, wzywając młode pokolenie badaczy do przeprowadzenia rewolucji w tym zakresie, zaś Teresa Święckowska (WDIB UW) umieściła innowacje w usługach wyszukiwania informacji w kontekście prawa autorskiego. Tematyka ta kontynuowana była w sesji czwartej, w ramach której referaty wygłosili dwaj goście z Poznańskiego Centrum Superkomputerowo-Sieciowego oraz pracownik WDIB UW. Marcin Werla, kierownik Działu Bibliotek Cyfrowych i Platform Wiedzy PCSS, zaprezentował kluczowe zagadnienia związane z wykorzystaniem analityki webowej (ang. *web analytics*) w kontekście systemów informacyjnych, takich jak repozytoria, biblioteki cyfrowe, bazy bibliograficzne, czy systemy informacji naukowej. Michał Kozak mówił o jakości danych w systemie informacji naukowej, zaś Piotr Tafiłowski (WDIB UW) omówił projekt kartografowania wiedzy, który chciałby zrealizować w przyszłości, po uzyskaniu środków finansowych na ten cel.

Sesja druga poświęcona została użytkownikom usług informacyjnych. Remigiusz Sapa (UJ) poddał reinterpretacji koncepcję użytkownika usług informacyjnych. Monika Krakowska i Paloma Korycińska z tej samej uczelni zaprezentowały zastosowanie metody osoby w badaniach informatologicznych oraz projektowaniu nowych rozwiązań zorientowanych na użytkownika. Julita Niedźwiecka-Ambroziak (UMK) mówiła o badaniach eyetrackingowych (percepcja graficznych elementów okładek książek), prowadzonych w bibliotece Wyższej Szkoły Bankowej w Toruniu.

Sesje trzecia i szósta poświęcone były innowacjom w informacyjnych usługach bibliotecznych i bibliograficznych. Maria Próchnicka (UJ) zaprezentowała mobilne usługi biblioteczne jako przedmiot badań informatologii, dokonując przeglądu publikacji w tym zakresie. Bożena Łazowska (Centralna Biblioteka Statystyczna w Warszawie) mówiła o innowacyjnych usługach informacyjnych w kierowanej przez nią bibliotece. Dorota Siwecka (Uniwersytet Wrocławski) przedstawiła korzyści i zagrożenia, jakie niesie ze sobą wykorzystanie koncepcji danych powiązanych (Linked Data) w działalności bibliotecznej. Kontynuację tematyki innowacji w informacyjnych usługach bibliotecznych i bibliograficznych w sesji szóstej rozpoczęło wystąpienie Justyny Antosik i Anny Kamińskiej (Biblioteka Politechniki Warszawskiej), które zademonstrowały blog „Niezbędnik Chemika” jako narzędzie edukacji informacyjnej. Następnie Grzegorz Kłębek (Biblioteka Uniwersytecka w Warszawie) zaprezentował nowe oprogramowanie dla elektronicznej biblioteki Uniwersytetu Warszawskiego. Wreszcie Karolina Jaworska (Instytut Badań Literackich PAN) zapoznała słuchaczy z nową bazą danych *Polskiej Bibliografii Literackiej*.

W sesji piątej, poświęconej innowacjom na rynku wydawniczym, wygłoszono tylko dwa referaty: Kamila Augustyn (UWr) mówiła o cyfryzacji oferty wydawniczej i perspektywach globalnego rozwoju rynku wydawniczego w odniesieniu do książki naukowej, zaś Maria Przystek-Samokowa (WDIB UW) zaprezentowała projekt pod nazwą Obserwatorium Językowe 2.0.

Dwie popołudniowe sekwencje sesji równoległych obejmowały problematykę innowacji w kolejnych obszarach wykorzystywania usług informacyjnych. Sesja siódma, poświęcona humanistyce cyfrowej, rozpoczęła się wystąpieniem Zbigniewa Osińskiego z Uniwersytetu Marii Curie-Skłodowskiej w Lublinie, który przedstawił wyniki zespołowego projektu badawczego (poza prelegentem realizowanego przez Ewę Głowacką z UMK w Toruniu, Mirosława Górnego z UAM i Małgorzatę Kisilowską z UW), poświęconego przekształceniom zachodzącym w infrastrukturze informacyjnej humanistyki oraz widocznym i przewidywanym wpływom tych zmian na praktyki badawcze i publikacyjne badaczy. W drugim referacie w tej sesji Jacek Tomaszczyk (UŚ w Katowicach) zademonstrował możliwości wykorzystania bezpłatnego programu Microsoft OneNote w pracy naukowej do gromadzenia, organizowania i opracowywania różnego typu materiałów – tekstowych, graficznych i audiowizualnych.

Sesja ósma dotyczyła innowacyjnych usług informacyjnych w archiwach. Problematyka ta omówiona została w trzech referatach. Justyna Soliwoda (Narodowe Archiwum Cyfrowe) zapoznała słuchaczy z systemem ZoSIA – budowanym i rozwijanym przez NAC narzędziem informatycznym, przeznaczonym dla archiwów, umożliwiającym opracowywanie zbiorów materiałów archiwalnych każdego typu, w oparciu o uporządkowaną hierarchiczną strukturę opisu i przepisy archiwalne Naczelnego Dyrektora Archiwów Państwowych, a także międzynarodowy standard opisu archiwalnego ISAD(G). Monika Cołbecka (UMK) mówiła o schematach wyszukiwania informacji w zasobach archiwalnych w dobie elektronicznych pomocy archiwalnych, zaś Kamila Biernat (Archiwum Akt Nowych w Warszawie, UMK) zaprezentowała pomoce wyszukiwawcze dla użytkowników archiwów.

Interesująca była także sesja dziewiąta, trwająca równolegle pod hasłem „Nowe trendy w dyscyplinach informacyjnych”. Katarzyna Materska (Biblioteka Uniwersytetu Kardynała Stefana Wyszyńskiego w Warszawie) zaproponowała umieszczenie innowacji i procesu innowacyjności w humanistycznej, filozoficznej, a także psychologicznej perspektywie pytając, czy innowacje powinny być dla człowieka, czy też człowiek jest dla innowacji? Nowatorskie spojrzenie na dyscyplinę zaproponował Wiesław Babik (UJ), mówiący o logistyce informacji, której zadaniem staje się organizacja i koordynacja procesów przepływu informacji i wiedzy w kanałach i łańcuchach logistycznych. Strukturalistyczna koncepcja architektury informacji była przedmiotem rozważań Stanisława Skórki (Uniwersytet Pedagogiczny w Krakowie).

Ostatnią grupę sesji równoległych poprzedziło wystąpienie przedstawicieli firmy Newspoint, jednego z patronów medialnych konferencji. Firma zajmuje się monitoringiem mediów w czasie

rzeczywistym. Reprezentujący ją Jakub Kozłowski i Kamil Niedbała omówili możliwości oferowanego przez nich komercyjnego narzędzia do monitorowania mediów. Prezentacja ta spotkała się z żywym zainteresowaniem słuchaczy.

Sesję dziesiątą, pt. „Innowacje w organizacji wiedzy” otworzył referat Barbary Sosińskiej-Kalaty poświęcony optymalizacji systemów organizacji wiedzy. Następnie Marcin Roszkowski z tej samej jednostki mówił o adnotacjach semantycznych w kolekcjach cyfrowych. Prelegent rozpoczął przemówienie od przedstawienia wniosków, które następnie, w trakcie wykładu, uzasadniał. Autorkami ostatniej prezentacji tej sesji były Jadwiga Woźniak-Kasperek (WDIB UW) i Anna Kamińska (Biblioteka Główna Politechniki Śląskiej w Katowicach), które omówiły możliwości wykorzystania języka UML w celu eliminowania wieloznaczności komunikatów, przede wszystkim komunikatów naukowych.

W sesji jedenastej, „Usługi informacyjne w gospodarce, ochronie zdrowia i kulturze”, Agnieszka Młodzka-Stybel (Centralny Instytut Ochrony Pracy – Państwowy Instytut Badawczy) przedstawiła wyniki własnego badania ankietowego, dotyczącego zastosowania technologii informacyjnych w wybranych sektorach gospodarki narodowej. Jan Kozakowski (Narodowy Instytut Zdrowia Publicznego – Państwowy Zakład Higieny, Biblioteka Naukowa, Warszawa) mówił o zasadach informowania o zdrowiu publicznym. Sesję zakończyło wystąpienie Małgorzaty Kisilowskiej, która w imieniu własnym oraz swoich współpracownic, Justyny Jasiewicz i Anny Mierzeckiej, opowiedziała o badaniach dotyczących oferty cyfrowej Narodowego Instytutu Audiowizualnego, których celem było wskazanie przyczyn niskiego zainteresowania materiałami cyfrowymi oraz zaproponowanie działań, które mogłyby spowodować pozytywne zmiany w tym zakresie.

Ostatnia, dwunasta sesja, poświęcona została nowym mediom w biznesie i kulturze. W pierwszym wygłoszonym w jej ramach referacie Łukasz Iwasiński (WDIB UW) skupił uwagę na problemie utowarowienia informacji medialnej dostępnej w sieci Internet. Następnie Anna Miotk (UKSW) mówiła o badaniach audytorium mediowego, które dostarczają informacji stanowiących podstawę dla podejmowania decyzji biznesowych. W trzecim referacie Grzegorz Gmiterek (WDIB UW) zaprezentował analizę funkcjonalności polskich aplikacji mobilnych wykorzystywanych w muzeach.

Obrazy konferencji zakończyła dyskusja podsumowująca omawiane na niej zagadnienia. Podkreślano ich bogactwo, potwierdzające dynamiczny rozwój nauki o informacji. Podnoszono też potrzebę kontynuacji tej cyklicznej konferencji, która umożliwi konfrontację wyników szerokiego spektrum badań prowadzonych w różnych obszarach współczesnej informatologii przez badaczy z różnych ośrodków polskich i zagranicznych.

Wydarzenie z pewnością osiągnęło cel zamierzony przez organizatorów, zapewniając forum wymiany wiedzy, doświadczeń i przemyśleń jego uczestników. Zgodnie z intencjami organizatorów gośćmi konferencji byli zarówno akademicy, badający teoretyczne zagadnienia realizacji usług informacyjnych, jak i praktycy wdrażający je w różnych dziedzinach życia. Szczególnie cenna była możliwość wymiany doświadczeń między środowiskiem naukowym i środowiskiem różnych gałęzi biznesu informacyjnego. Spotkanie przyniosło wiele interesujących (i mamy nadzieję, że również owocnych) dyskusji kulturalnych.

Piotr Tafiłowski

Katedra Informatologii

Wydział Dziennikarstwa, Informacji i Bibliologii

Uniwersytet Warszawski

Nadesłano: 17 czerwca 2017

I Ogólnopolska Konferencja „Organizacja i recepcja treści w środowisku cyfrowym”

(Katowice, 23 maja 2017 r.)

W dniu 23 maja 2017 r. w Katowicach odbyła się I Ogólnopolska Konferencja „Organizacja i recepcja treści w środowisku cyfrowym”, zorganizowana przez Instytut Bibliotekoznawstwa i Informatyki Naukowej Uniwersytetu Śląskiego w Katowicach oraz Centrum Informatyki Naukowej i Bibliotekę Akademicką Uniwersytetu Śląskiego i Uniwersytetu Ekonomicznego w Katowicach. Głównym celem konferencji była dyskusja nad różnymi aspektami organizowania, udostępniania, publikowania i odbioru tekstu w przestrzeni cyfrowej. Uroczystym dopełnieniem konferencji była promocja książki *Kultura książki i informacji* (Wydaw. Uniwersytetu Śląskiego, 2017, red. A. Pulikowski) dedykowanej Profesor Elżbiecie Gondek w związku z jubileuszem 70-lecia jej urodzin.

Obrady rozpoczął prof. Jacek Tomaszczyk, dyrektor Instytutu Bibliotekoznawstwa i Informatyki Naukowej Uniwersytetu Śląskiego w Katowicach, który powitał uczestników spotkania. Pierwsza sesja, prowadzona przez dr. Karola Maklesa z Uniwersytetu Śląskiego, rozpoczęła się wystąpieniem zatytułowanym *Percepcja tekstu w formie drukowanej i elektronicznej*. Katarzyna Tałuc i Jacek Tomaszczyk z Uniwersytetu Śląskiego zaprezentowali wyniki badań empirycznych dotyczących odbioru tekstu udostępnianego w wersjach drukowanej oraz cyfrowej. Eksperyment, przeprowadzony w grupie studentów kierunku humanistycznego, pokazał, że niezależnie od urządzenia, interfejs cyfrowy nie wpływa na rozumienie i interpretację tekstu. Kolejny referat pt. *Organizacja i recepcja informacji w środowisku cyfrowym w świetle ekologii informacji* wygłosił Wiesław Babik z Uniwersytetu Jagiellońskiego. Poruszył kwestie organizacji i odbioru informacji oraz produktów informacyjnych, szczególną uwagę poświęcając problemom nadmiaru informacji, przeciążenia informacyjnego i chorób z nimi związanych. Justyna Adamus-Kowalska z Uniwersytetu Śląskiego zaprezentowała referat pt. *Organizacja informacji archiwalnej w środowisku cyfrowym*. Swoją uwagę skupiła na zasobach informacji archiwalnej tworzonej w administracji publicznej, narzędziach wykorzystywanych do udostępniania tej informacji w wersji elektronicznej oraz wskazała konieczności opracowania jednolitych metadanych niezbędnych do efektywnego przetwarzania informacji archiwalnej.

Drugą sesję, również prowadzoną przez dr. Karola Maklesa, rozpoczęła Natalia Pamuła-Cieślak z Uniwersytetu Mikołaja Kopernika w Toruniu, przedstawiając referat pt. *Przyczyny niepowodzeń. Dlaczego wielu polskich czasopism otwartych nie ma w Directory of Open Access Journal?* Autorka wymieniła problemy uniemożliwiające rejestrację polskich czasopism w katalogu. Dokonała zestawienia częstych błędów popełnianych przez redakcje czasopism i ich wydawców i porównała je z kryteriami niezbędnymi do rejestracji czasopisma w DOAJ. Przybliżyła także dobre praktyki publikowania naukowego wdrożone przez redakcję bazy. Kolejny referat, zatytułowany *Cytowania „Nowej Biblioteki. Usług, Technologii Informacyjnych i Mediów” jako element komunikacji naukowej w świecie cyfrowym*, wygłosiły Marta Gawlik i Edyta Kosik z Uniwersytetu Śląskiego. W swoich badaniach uwagę skupiły na analizie rzeczowej cytowań, w celu określenia stopnia wykorzystania literatury naukowej w periodyku w latach 2010–2015, a także ustalenia klasyfikacji tematycznej tych dzieł. Anna Małgorzata Kamińska z Politechniki Śląskiej wygłosiła referat pt. *Potencjał bibliometryczny bibliograficznej bazy danych CYTBIN w świetle prostych i złożonych wskaźników analitycznych*. Przedstawiła w nim wyniki analiz bibliometrycznych, zrealizowanych na danych udostępnianych z poziomu przeglądarki aplikacji systemu CYTBIN, m.in.: sieci współautorstwa i cytowań, miary Erdősa, Hirscha i PageRank. Prezentację wzbogaciły wizualizacje danych dotyczące najaktywniej cytujących autorów, współcytowań oraz powiązań bibliograficznych. Przedstawione wyniki potwierdziły tezę o dużym potencjale

bibliometrycznym bazy CYTBIN. Następnie głos zabrała Małgorzata Gajos-Grzetic z Uniwersytetu Śląskiego, która zaprezentowała referat zatytułowany *Semantyczne aspekty cyberprzestrzeni i cyberprzestępczości*. Wyjaśniła w nim pojęcie cyberprzestrzeni, przedstawiła jej cechy i powiązania z cyberbezpieczeństwem, by następnie przejść do rodzajów cyberzagrożeń oraz definicji i form cyberprzestępczości. Sesję zakończył referat Anny Radoszewskiej pt. *IBUK libra – sprawdzona technologia dla edukacji i rozwoju*, w którym autorka zaprezentowała zasoby edukacyjne i możliwości wyszukiwawcze wirtualnej czytelnicy.

Następnie odbyła się promocja książki dedykowanej prof. dr hab. Elżbiecie Gondek, szerzej przedstawiona w drugiej części sprawozdania. Po uroczystości odbyła się ostatnia sesja konferencji prowadzona przez dr Annę Matyszek z Uniwersytetu Śląskiego. Pierwszy z prelegentów, Artur Ogurek z Uniwersytetu Wrocławskiego, wygłosił referat zatytułowany *Problem udostępniania zasobów cyfrowych w świetle najnowszych przepisów prawa oraz orzecznictwa sądów*. Poruszył w nim zagadnienie udostępniania zbiorów bibliotecznych za pośrednictwem terminali w bibliotekach publicznych. Szczegółowo omówił wyrok C-117/13 Trybunału Sprawiedliwości Unii Europejskiej, dotyczący problemu udostępniania zbiorów bibliotecznych w wersji cyfrowej oraz przybliżył planowane zmiany w prawie autorskim. Kolejny referat pt. *Dokumentacja codzienności w wybranych zasobach internetowych (z perspektywy językoznawcy o sąsiedzkich konfliktach)* przedstawiła Aneta Banaszek-Szapowalowa z Uniwersytetu Śląskiego. W wystąpieniu podjęła próbę odpowiedzi na pytanie, jak zdarzenia realne, w tym przypadku konflikty sąsiedzkie, są prezentowane w przestrzeni wirtualnej przez zwykłych użytkowników Internetu. Rozpatrując przykłady polsko – i rosyjskojęzycznych źródeł internetowych, wśród sposobów dokumentacji konfliktów sąsiedzkich zaprezentowała werbalne i wizualne treści. Następny referat pt. *Wyszukiwanie informacji o właściwościach szablonów wyglądu dla najpopularniejszych systemów CMS* wygłosił Mariusz Jarocki z Uniwersytetu Mikołaja Kopernika. Przedstawił w nim wyniki badań, których przedmiotem były możliwości systemów wyszukiwawczych, znajdujących się na witrynach domowych najpopularniejszych na świecie systemów CMS. Wyszukiwarki testowano pod kątem stwierdzenia poziomu trudności odnajdywania informacji o pożądanym przez użytkownika właściwościach szablonu, a kierowane do systemu zapytania odnosiły się m.in. do takich zagadnień, jak responsywność szablonów, czy też spełnienie norm dostępności dla osób z niepełnosprawnościami. Sesję zakończył referat Agnieszki Jezierskiej i Andrzeja Koziara z Centrum Informacji Naukowej i Biblioteki Akademickiej pt. *Analiza ryzyka jako narzędzie planowania i wspomagania zarządzania bezpieczeństwem informacji*. Autorzy zaprezentowali w nim podstawy prawne i normatywne analizy ryzyka jako narzędzia do planowania i wdrażania zarządzania bezpieczeństwem informacji. Wskazali szanse i zagrożenia dla instytucji na różnych poziomach organizacji pracy, a także dylematy zarządzających przy korygowaniu działań, na podstawie wskazań wynikających z przeprowadzonych audytów bezpieczeństwa.

Po wygłoszeniu wszystkich referatów odbyła się dyskusja, w której głos zabrali m.in. Elżbieta Gondek, Wiesław Babik, Andrzej Koziara. Podsumowania i oficjalnego zakończenia konferencji dokonał dr hab. Jacek Tomaszczyk, który dziękując uczestnikom za udział w konferencji zaprosił na kolejną edycję.

Uświetnieniem konferencji była promocja książki dedykowanej Profesor Elżbiecie Gondek. Irena Socha przedstawiła sylwetkę Jubilatki, jako naukowca oraz osoby sumiennej, odpowiedzialnej i rzetelnej, a jednocześnie życzliwej dla innych i spieszącej z pomocą. Elżbieta Gondek już we wczesnej młodości interesowała się nauczaniem i pedagogiką³. Studia, na kierunku filologia polska podjęła w Studium Nauczycielskim w Cieszynie, które kontynuowała na poziomie magisterskim na Uniwersytecie Śląskim w Katowicach, pracując jednocześnie jako nauczyciel w szkole podstawowej. Talentem

³ Sylwetkę prof. dr hab. Elżbiety Gondek przedstawiono na podstawie: Socha, I. (2017) Profesor Elżbieta Gondek – w służbie nauki o książce oraz akademickiej dydaktyki bibliotekoznawczej i informatologicznej. W: A. Pulikowski (red.). *Kultura książki i informacji*. Katowice, Wydaw. Uniwersytetu Śląskiego, 7–18.

naukowym wykazała się już podczas seminarium magisterskiego dr Janiny Saratowicz-Stolarzewiczowej, gdzie przygotowywała pracę *Monografia „Katolika” 1890–1894*. Pracę na Uniwersytecie Śląskim rozpoczęła w 1974 r., gdy powołano kierunek bibliotekoznawstwo i informacja naukowa oraz Pracownię do Badań nad Życiem Literackim Śląska i Zagłębia Dąbrowskiego, kierowaną przez docenta Adama Jarosza. Początkowo Elżbieta Gondek jako asystent naukowo-badawczy przyczyniła się do stworzenia bogatej i dobrze opracowanej kartoteki zawartości zagłębiowskiej i śląskiej prasy w ramach Pracowni zagłębiowsko-śląskiej. Pod kierunkiem doc. dr. Adama Jarosza przygotowywała rozprawę doktorską na temat *Rola prasy Zagłębia Dąbrowskiego w kształtowaniu kultury literackiej regionu (1897–1918)*, którą obroniła w 1989 r., a opublikowała trzy lata później pod tytułem *Prasa i literatura w Zagłębiu Dąbrowskim 1897–1918*. Podstawą habilitacji była, oprócz licznych artykułów naukowych, monografia pt. *Polska książka literacka na Śląsku pod panowaniem pruskim 1795–1863* opublikowana w Wydawnictwie Uniwersytetu Śląskiego (1995). Praca zatytułowana *Rynek książki na Śląsku pod panowaniem pruskim w XIX wieku. Drukarnie, księgarnie, płatne wypożyczalnie publikacji* (2001), stała się z kolei podstawą procedury uwieńczoną nadaniem Jubilatce tytułu profesora nauk humanistycznych w 2003 r. Obszerny dorobek naukowy, dokumentacyjny i dydaktyczny Elżbiety Gondek, liczący łącznie ponad 130 pozycji, lokuje się w kilku głównych nurtach: rozległe i wielotematyczne badania regionalne dotyczące kultury książki i literatury Zagłębia Dąbrowskiego i Śląska; historia książki polskiej w XIX wieku; teoria i historia bibliografii polskiej oraz jej znaczenie w rozwoju bibliologii.

W trakcie ponad czterdziestu lat działalności naukowej i akademickiej prof. Elżbieta Gondek podejmowała wiele różnych działań: sprawowała funkcje wicedyrektora ds. dydaktycznych (1996–2005), a potem dyrektora (2005–2012), jest także kierownikiem Zakładu Zarządzania Informacją (od 2011). Na Wydziale Filologicznym pełniła funkcje prodziekana ds. dydaktycznych (2000–2002), członka Rady Dydaktycznej (2008–2012), a także funkcję przewodniczącej Rady Programowej Biblioteki Uniwersytetu Śląskiego. Była członkiem Centralnej Komisji ds. Stopni i Tytułów (2011–2012) oraz wiceprzewodniczącą Centralnej Komisji Egzaminacyjnej dla Kandydatów na Bibliotekarzy Dyplomowanych i Dokumentalistów Dyplomowanych (2008–2011). Uczestniczyła także w pracach nad doskonaleniem jakości kształcenia w zakresie bibliologii, bibliotekoznawstwa i informacji naukowej w Uniwersyteckiej, a potem Państwowej Komisji Akredytacyjnej, przygotowując jednocześnie projekty dydaktyczne i programy nauczania. Prof. Elżbieta Gondek wypromowała trzech doktorów, recenzowała liczne prace doktorskie, kilkakrotnie dorobek habilitacyjny, opiniowała wnioski w postępowaniu o nadanie tytułu profesora. Ma w swoim dorobku także liczne recenzje wydawnicze prac naukowych, grantów dla Komitetu Badań Naukowych oraz projektów naukowych MNiSW. Współpracowała z ogólnopolskimi recenzowanymi czasopismami naukowymi jako członek Rady Naukowej oraz Komitetu Redakcyjnego. Prof. Elżbieta Gondek wniosła trwałe i znaczący wkład do rozwoju śląskiego bibliotekoznawstwa oraz wiedzy o kulturze książki na Śląsku i w Zagłębiu, a także do bibliologicznych badań naukowych.

Po wysłuchaniu biografii Jubilatki, gratulacje składali: dziekan Wydziału Filologicznego, prof. dr hab. Krzysztof Jarosz, a następnie dyrektorzy i przedstawiciele zaprzyjaźnionych instytutów bibliotekoznawstwa, profesorowie: Ewa Głowacka, Maria Juda, Michał Rogoż i Jacek Wojciechowski. Następnie głos zabrala wzruszona Jubilatka, która kierując słowa podziękowania do obecnych gości, dodała kilka osobistych wspomnień z początków pracy na Uniwersytecie Śląskim oraz podkreśliła wpływ profesora Jana Zaremby na jej zainteresowania śląskoznawcze.

Doniosłym punktem uroczystości było wręczenie przez redaktora Arkadiusza Pulikowskiego księgi jubileuszowej zatytułowanej *Kultura książki i informacji*, opublikowanej przez Wydawnictwo Uniwersytetu Śląskiego. Arkadiusz Pulikowski opowiedział o przygotowaniach książki, wyborze autorów oraz samej treści. Publikacja rozpoczyna się biografią Jubilatki napisaną przez Irenę Sochę oraz bibliografią publikacji za lata 1980–2015, przygotowaną przez Agnieszkę Gołdę. Pozostałe prace 27 autorów zebrano w trzech częściach: *Kultura książki*, *Ludzie książki*, *Kultura informacji*.

Na zakończenie głos zabrał jeszcze Dariusz Pawelec, wspominając współpracę prof. Elżbiety Gondek z Biblioteką Uniwersytetu Śląskiego, zaś oficjalne życzenia w imieniu macierzystego Instytutu złożyli Jacek Tomaszczyk i Karol Makles. Następnie składano Jubilatce indywidualne gratulacje, a uroczystość zakończyła się miłym poczęstunkiem i nieformalnymi rozmowami.

Anna Matysek

*Instytut Bibliotekoznawstwa i Informatyki Naukowej
Uniwersytet Śląski w Katowicach*

Nadesłano: 1 lipca 2017.

IV Konferencja Naukowa „Biblioteka w przestrzeni edukacyjnej” Książki w plikach – publikowanie, udostępnianie i użytkowanie (Kraków, 24–25 maja 2017 r.)

Biblioteka Główna oraz ówczesny Instytut Informatyki Naukowej i Bibliotekoznawstwa, a obecnie Instytut Nauk o Informatyce Uniwersytetu Pedagogicznego im. Komisji Edukacji Narodowej w Krakowie poczynawszy od 2011 r., co dwa lata organizuje konferencję naukową, której hasło przewodnie brzmi: Biblioteka w przestrzeni edukacyjnej.

Pierwsza konferencja, zorganizowana w 2011 r., ukazała potrzebę wymiany poglądów i doświadczeń na temat m.in. zastosowania nowych technologii komunikacyjnych w pracy bibliotekarza. Podczas kolejnej, w 2013 r., skupiono się na zagadnieniach związanych z kierunkami rozwoju profesji bibliotekarza 2.0: jego wiedzy, umiejętności i kompetencji. W 2015 r. podjęto temat projektowania informacji.

W 2017 r., w dniach 24–25 maja na Uniwersytecie Pedagogicznym w Krakowie odbyła się już czwarta edycja konferencji „Biblioteka w przestrzeni edukacyjnej”, tym razem poświęcona tematu: *Książki w plikach – publikowanie, udostępnianie i użytkowanie*. Głównymi organizatorami wydarzenia byli: Biblioteka Główna oraz Instytut Nauk o Informatyce Uniwersytetu Pedagogicznego w Krakowie. Honorowy patronat nad konferencją objęli Prezydent Miasta Krakowa, prof. dr hab. Jacek Majchrowski oraz JM Rektor Uniwersytetu Pedagogicznego, prof. dr hab. Kazimierz Karolczak.

Obrady uświetniło wystąpienie Prorektora ds. Nauki, prof. dr hab. Mariusza Wołosa oraz Prodziekana Wydziału Filologicznego, dr hab. prof. UP. Marka Władysława Kolasy. Ponadto głos zabrali: dyrektor Biblioteki Głównej UP im. KEN, dr Stanisław Skórka i dyrektor Instytut Nauk o Informatyce, dr hab. Michał Rogoź.

Konferencja dotyczyła kwestii związanych z audiobookami, e-bookami i innymi dokumentami niedrukowanymi w praktyce biblioteczno-informacyjnej. Miała na celu wymianę doświadczeń z zakresu wykorzystania, opracowywania, udostępniania, gromadzenia, czytelności, percepcji i funkcjonalności książek w postaci cyfrowej.

Podczas dwudniowego spotkania odbyło się pięć sesji tematycznych prowadzonych przez: prof. dr hab. Barbarę Sosińską-Kalotę, prof. dr hab. Jadwigę Woźniak-Kasperek, dr hab. Michała Rogoźa, prof. dr hab. Wiesława Babikę oraz dr hab. Małgorzatę Kowalską.

W pierwszej sesji – „Funkcjonalność i dostępność mediów cyfrowych” wykład wprowadzający *Człowiek w świecie mediów cyfrowych* wygłosił prof. dr hab. inż. Janusz Morbitzer. Prelekcja odnosiła się do oddziaływania mediów na człowieka i społeczeństwo, istoty edukacji medialnej, zagrożeń ze

strony mediów oraz dominacji ikonosfery nad logosferą. Prof. Morbitzer w referacie nawiązał do podziału społeczeństwa Marca Prensky'ego na dwie kategorie: cyfrowych tubylców i cyfrowych imigrantów. Dla pierwszej grupy, do której zalicza się osoby urodzone po 1983 r. świat mediów i Internet jest naturalnym środowiskiem życia. W tym przypadku pierwszeństwo ma obraz, myślenie hipertekstowe, a zamiast książek – Google. Druga grupa posługuje się językiem ery przedmedialnej, przedkładając tekst nad obraz i dźwięk. Współcześni ludzie żyją w warunkach znacznego nadmiaru informacji oraz bodźców z telewizji, gier komputerowych, Internetu. Pod wpływem nadmiaru informacji ludzki mózg przechodzi w tzw. stan awaryjny i skraca obieg informacji. Człowiek obojętnie na wszystko, co go nie dotyczy, a obniżeniu ulega poziom empatii, altruizmu i tolerancji. Media działają destrukcyjnie na wyobraźnię, a to ma kluczowe znaczenie w kwestii odpowiedzialności, a więc myślenia w kategoriach przyszłości. Internet powoduje zmiany w budowie mózgu, a co za tym idzie niezdolność do głębszej refleksji (G. Smell, G. Vorgan). Powrót do pisma obrazkowego oznaczałby regres cywilizacyjny. Prof. Morbitzer zacytował Marshalla McLuhana: „My kształtujemy narzędzia, a potem one kształtują nas”. Nadmienił, że nie wszystko, co można technicznie zrealizować, należy moralnie zaakceptować, a istotnym zadaniem edukacji medialnej jest kształtowanie odpowiedzialności u uczących się osób. Wspomniał o informacyjnym zalewie, wielości opinii i poglądów, dominacji informacji nad wiedzą i mądrością. Hipermediatyżacja pozbawia nas własnych doświadczeń, oddając we władzę rzeczywistości zapośredniczonej. Konsekwencją tej sytuacji jest to, że każdy z nas rozumie jedynie to, czego bezpośrednio doświadczył. Prelegent omówił także zagrożenia wynikające ze stałego oglądania telewizji przez dzieci, prowadzące do trwałych zmian osobowości oraz hamujące rozwój (R. Patzleff). Telewizja osłabia zdolność myślenia abstrakcyjnego, a wraz z nią – rozumienia. Postawił również pytanie: „Co my robimy z mediami, co media robią z nami?”. Podkreślił, że zagrożeniem jest brak umiejętności właściwego odnoszenia się do mediów i racjonalnego korzystania z nich. Istotna jest wiedza dotycząca mechanizmów oddziaływania mediów na jednostkę i społeczeństwo. W sesji tej głos zabrali także Małgorzata Kowalska oraz Katarzyna Jarczeńska-Walendziak, Karolina Bohdanowicz, Tomasz Parkoła i Marcin Werla, jak również Stanisław Skórka. Referat *Platformy e-booków w polskich bibliotekach akademickich – ocena funkcjonalności i dostępności* Małgorzaty Kowalskiej oraz Katarzyny Jarczeńskiej-Walendziak odnosił się do wyniku testów 10 wybranych platform e-książek używanych w bibliotekach akademickich, w tym Elsevier eBooks, Ibuk Libra, NASBI, Science Direct. Badania zostały przeprowadzone pod kątem możliwości zmian układu graficznego, dostosowywania tekstu, wyszukiwania oraz nawigacji, konwersji tekstu na mowę, a także druku. W prezentacji *W kierunku zwiększania dostępności zasobów udostępnianych przez polskie biblioteki cyfrowe – nowoczesne rozwiązania w systemie dLibra 6* Karolina Bohdanowicz, Tomasz Parkoła i Marcin Werla z Poznańskiego Centrum Superkomputerowo-Sieciowego zaprezentowali rozwój szóstej generacji systemu dLibra, jednego z popularnych oprogramowań do tworzenia bibliotek cyfrowych w Polsce. Przedstawione zostały kluczowe właściwości systemu dLibra 6 w zakresie dostępności interfejsu użytkownika portalu WWW w związku z badaniem przeprowadzonym w 2016 r. Referat *Użyteczność czytelników książek elektronicznych w świetle norm i standardów* Stanisława Skórki przybliżył zasady oraz normy używane przy projektowaniu aplikacji dla urządzeń mobilnych oraz elementy interfejsu użytkowników.

W drugiej sesji pt. „Integracja zasobów cyfrowych w bibliotekach” prezentacje mieli: Wojciech Kowalewski, Agnieszka Gałęcka-Golec i Iwona Łojewska oraz Bożena Winiarska.

Referat *Integracja książek elektronicznych w katalogach OPAC a wzrost czytelnictwa* Wojciecha Kowalewskiego dotyczył zespelenia metadanych subskrybowanych zasobów elektronicznych w katalogach wybranych bibliotek publicznych. Agnieszka Gałęcka-Golec i Iwona Łojewska w prelekcji *Repozytorium Uniwersytetu w Białymstoku – jako narzędzie gromadzenia, upowszechniania i promowania dorobku uczelni* przedstawiły pierwsze repozytorium instytucjonalne archiwizujące i upowszechniające dorobek naukowy i dydaktyczny pracowników, doktorantów i studentów Uniwersytetu w Białymstoku. Wystąpienie Bożeny Winiarskiej *Książnica w pliku – pliki w Książnicy*

poświęcone było różnego rodzaju plikom elektronicznym używanym w Książnicy Pomorskiej wraz z porównaniem z danymi statystycznymi.

Podczas sesji trzeciej zatytułowanej „Udostępnianie dokumentów elektronicznych” wystąpili: Bartłomiej Siedlarz, Mariusz Jarocki, Michał Górski, Łukasz Tomkiewicz i Grzegorz Wajda oraz Joanna Hoły. Referat *Dokumenty ciągłe w polskich bibliotekach cyfrowych – podejścia do użycia zasobu i sposoby prezentacji treści* Bartłomieja Siedlarza przybliżył przykładowe rozwiązania ukazujące wykorzystywane sposoby prezentacji dokumentów ciągłych. Prelekcja Mariusza Jarockiego *Aplikacje mobilne dla systemu Android a zasoby cyfrowe bibliotek polskich* dotyczyła odpowiedzi na pytanie, czy istnieje idealna aplikacja mobilna do odczytu zasobów pochodzących z polskich bibliotek cyfrowych i repozytoriów. Ocenie zostały poddane darmowe aplikacje dostępne za pośrednictwem usługi Google Play. W wystąpieniu *Piractwo czy nieformalny obieg książek elektronicznych w Polsce? Próba charakterystyki* Michała Górskiego, Łukasza Tomkiewicza i Grzegorza Wajdy przedstawiono słuchaczom pojęcia związane z poruszaną problematyką głównie w kontekście książek elektronicznych. Podjęto także próbę wyszczególnienia istotnych czynników odpowiedzialnych za rozwój nielegalnego obiegu dzieł kultury oraz przykłady metod zwalczających piractwo. Odczyt *Copyright w bibliotekach cyfrowych* Joanny Hoły odnosił się do przestrzegania ustawy o prawie autorskim, prawach pokrewnych i związanych z tym problemach na przykładzie Jagiellońskiej Biblioteki Cyfrowej.

W drugim dniu konferencji podczas czwartej sesji zatytułowanej *Użyteczność i recepcja książki elektronicznej* głos oddano: Barbarze Kamińskiej-Czubale, Beacie Langer, Monice Curyło, Małgorzacie Lebdzie i Barbarze Krasieńskiej, a także Katarzynie Staniuk. Prelekcja *E-podręczniki w perspektywie architektury informacji* Barbary Kamińskiej-Czubały dotyczyła oceny jakości e-podręczników przez wykorzystanie kryteriów i metod wypracowanych przez architekturę informacji i używanych w ulepszaniu internetowych serwisów informacyjnych. W referacie *Czytelność i gęstość informacyjna e-podręczników do historii dla kształcenia podstawowego w klasie VI* Beata Langer poruszała kwestie związane z pomiarami zmierzającymi do określenia czytelności i gęstości informacyjnej wybranych e-podręczników. Monika Curyło w wystąpieniu *Czytelnictwo e-booków w Polsce: stan obecny i perspektywy* zaprezentowała zagadnienie e-czytelnictwa, głównie wśród młodych ludzi, nieoznaczające jednak wyparcia literatury drukowanej, gdyż literatura elektroniczna i drukowana będą uzupełniać się i współistnieć. Małgorzata Lebda i Barbara Krasieńska w referacie *Czy oceniać po okładce? O szacie graficznej współczesnych publikacji elektronicznych* omówiły analizę szaty graficznej, w szczególności okładki, a także tendencje ich projektowania w gronie największych wydawnictw, publikujących dzieła typu „digital born”. Odczyt *Literatura w kulturze cyfrowej* Katarzyna Staniuk poświęciła zagadnieniom dotyczącym przejścia od kultury pisma do kultury obrazu, wiążącym się z wizualizacją literatury.

Podczas piątej sesji „Rynek dokumentów elektronicznych” głos zabrali: Agnieszka Łakomy, Piotr Andrusiewicz, Michał Górski, Łukasz Tomkiewicz i Grzegorz Wajda, Natalia Pamuła-Cieślak, Marzena Błach i Michał Rogoż. Agnieszka Łakomy w prelekcji *Rynek polskich czasopism przeznaczonych na tablety* omówiła wyniki analizy oferty rynkowej polskich periodyków przeznaczonych na tablety, ponadto określiła typy dostępnych tytułów oraz stosunek ich zawartości do odpowiedników drukowanych i internetowych portali prasowych. Piotr Andrusiewicz w wystąpieniu *Multiscreening i shifting w procesie lektury za pomocą czytników ebooków* omawiał metody sprzedaży i udostępniania e-książek na rynku polskich księgarni internetowych, jak również rozwiązania wspierające coraz popularniejszy trend korzystania z wielu ekranów jednocześnie, w tym rozpoczynania lektury na jednym i kontynuowania jej na innym urządzeniu. Referat *Dokumenty polskojęzyczne w kolekcji eBooks and Texts serwisu Internet Archive* Michała Górskiego, Łukasza Tomkiewicza i Grzegorza Wajdy miał na celu przybliżenie idei rozpowszechniania wiedzy w ramach projektu Internet Archive oraz przedstawienie polskojęzycznych dokumentów z kolekcji eBooks and Texts i obsługi Open Library z platformą Bookserver, umożliwiających wypożyczanie elektronicznych wersji tekstów objętych prawami autorskimi. W wystąpieniu *Otwarte książki w sieci na przykładzie Directory of Open Access*

Books Natalia Pamuła-Cieślak omówiła platformę udostępniającą i wyszukującą książki wydawane w modelu otwartego dostępu, skupiając uwagę na modelu funkcjonowania serwisu, upowszechnianiu wiedzy o otwartych książkach i wydawcach oraz porównaniu działania platformy DOABooks z jej podobnymi. Odczyt *Alternatywne wydania przewodników po Krakowie: e-przewodniki* Marzeny Błach miał za zadanie próbę przedstawienia metodologii badań nad e-przewodnikami na przykładzie publikacji dotyczących Krakowa. W prelekcji *Nie tylko biblioteki. Cyfrowe materiały informacyjne na stronach polskich parków narodowych* Michał Rogoż omówił ofertę Polskich Parków Narodowych w zakresie cyfrowych zasobów informacyjnych związaną głównie z badaniami naukowymi, edukacją ekologiczną oraz turystyką stanowiącą zazwyczaj uzupełnienie istniejących zasobów drukowanych.

Reasumując, na konferencji w ciągu dwóch dni, wygłoszono 22 referaty w pięciu sesjach tematycznych mających związek z funkcjonalnością i dostępnością mediów cyfrowych, integracją zasobów cyfrowych w bibliotekach, udostępnianiem dokumentów elektronicznych, użytecznością i recepcją książki elektronicznej a rynkiem dokumentów elektronicznych.

Sponsorami konferencji były firmy: Elibron, Międzynarodowe Centrum Budownictwa, Ebsco oraz Ibuk Libra.

W konferencji udział wzięło 67 uczestników, przedstawicieli wielu ośrodków, nie tylko akademickich, ale także bibliotek pedagogicznych, publicznych oraz szkolnych. Organizatorzy planują wydać publikację pokonferencyjną, a za dwa lata zorganizować piątą Konferencję Naukową z cyklu „Biblioteka w przestrzeni edukacyjnej”.

Agnieszka Folga

Joanna Kołakowska

Biblioteka Główna Uniwersytetu Pedagogicznego w Krakowie

Nadesłano: 30 czerwca 2017.

III Konferencja Naukowa Konsorcjum BazTech

Bibliograficzne bazy danych: perspektywy i problemy rozwoju

Kraków, 26–27 czerwca 2017

Bibliografio! Ty nad poziomy wylatuj!

W dniach 26 i 27 czerwca 2017 w Krakowie odbyła się III Konferencja Naukowa Konsorcjum BazTech zatytułowana Bibliograficzne bazy danych: perspektywy i problemy rozwoju. W tym roku zorganizowały ją Biblioteka Politechniki Krakowskiej i Biblioteka Główna Uniwersytetu Technologiczno-Przyrodniczego w Bydgoszczy. Honorowy patronat objęli JM Rektor Politechniki Krakowskiej prof. dr hab. inż. Jan Kazior oraz JM Rektor Uniwersytetu Technologiczno-Przyrodniczego prof. dr hab. inż. Tomasz Topoliński. Poprzednie konferencje Konsorcjum BazTech miały miejsce w 2013 r. (Bibliograficzne bazy danych i ich rola w rozwoju nauki, Poznań) i 2009 r. (Bibliograficzne bazy danych: kierunki rozwoju i możliwości współpracy, Bydgoszcz).

Zakres omawianych problemów szczegółowo został określony w zaproszeniu i obejmował:

- tworzenie dziedzinowych baz danych rejestrujących artykuły z czasopism,

- znaczenie baz bibliograficznych dla upowszechniania nauki,
- wsparcie procesu oceny czasopism naukowych przez bazy bibliograficzne,
- rola i miejsce baz danych w rozwoju otwartej nauki, w tym otwartego dostępu do treści naukowych,
- programy do edycji, udostępniania i agregowania baz,
- analizy bibliometryczne i webometryczne,
- zasoby bibliograficznych baz danych a inne serwisy, np. repozytoria (w tym też repozytoria surowych danych), wyszukiwarki naukowe, bazy cytowań, platformy czasopism,
- współpraca z wydawcami i redakcjami czasopism,
- najlepsze praktyki i standardy wydawnicze,
- sposoby skutecznego rozpowszechniania informacji o bazach, głównie w środowisku naukowców, studentów oraz wydawców i redakcji czasopism.

Organizatorzy pogrupowali zgłoszone referaty w pięć sesji tematycznych. Pierwsza – wprowadzająca – składała się z trzech wystąpień. Jan Kozłowski, radca w Ministerstwie Nauki i Szkolnictwa Wyższego, zacytował swoją wypowiedź *Bibliografia – gatunek z przyszłością czy gatunek zagrożony?* Na panoramicznym tle historii bibliografii, jako gatunku piśmienniczego oraz zmian, które obecnie dokonują się w jej otoczeniu społecznym, naukowym, gospodarczym, politycznym, technologicznym, uwzględniając koncepcję Kevina Kelly’ego strategii dla świata połączonego z siecią, prelegent zadał kilka ważnych pytań o przyszłość bibliografii, pozostawiając otwartą kwestię odpowiedzi na nie. Moją uwagę zwróciły zwłaszcza trzy następujące pytania: (1) Jeśli bibliografia jest zbiorem opisów publikacji dokonanych za pomocą takich standardowych charakterystyk, jak m.in. autor, tytuł, słowa kluczowe, streszczenie, cytowania, czyli innym słowy „mapą”, jaka może być jej rola, gdy mamy dziś bezpośredni dostęp do „terytorium”, czyli samych publikacji (otwarty dostęp)? (2) Czy bibliograf nie powinien posunąć się o ogniwo (lub ogniwa) dalej w „łańcuchu wartości” tworzenia wiedzy i przeobrazić się w „meta-analityka”, „eksploratora tekstu”, „syntetyka”? (3) Czy bibliografie nie chwytają tylko pewnych elementów publikacji, przeocząc przy tym inne ważne składniki (styl, dane, podejścia itd.)? W zacytowanych pytaniach są zawarte supozycje odpowiedzi – nie ze wszystkimi mogę się zgodzić. Metafora bibliografii jako mapy przemawia do wyobraźni. Dziś już pewnie niewiele osób korzysta z papierowych map, planów miast czy przewodników. Zastąpiły je nawigacje, aplikacje w telefonach usprawniające poruszanie się na przykład po nieznanym mieście, przewodniki cyfrowe, audio (ciągle powszechne w muzeach i galeriach) itp. Ale czy mimo wszelkich różnic sprzętowo-technologicznych nadal nie służą realizacji tego samego zadania, jakim jest identyfikowanie i informowanie poprzez przeprowadzenie mnie z punktu A do punktu B, czasami z uwzględnieniem jakiś dodatkowych informacji? Czy wiele jest osób, które podjęłyby ryzyko przejechania samochodem „tylko” z Suwałk do Faro bez wsparcia się jakąś „mapą”, choć terytorium jest i jest otwarte? Oczywiście, bez udziału bibliografii też się da coś znaleźć, ale czy istotnie zawsze chodzi nam o „coś” zamiast „to, co optymalne” (cokolwiek by to miało znaczyć w konkretnym przypadku)? Poza tym nie możemy zapominać, że każda bibliografia pełni jeszcze jedną istotną rolę – jest depozytariuszem obrazu dorobku jakiejś dyscypliny, narodu, państwa, regionu itd. Podlega zasadom archiwizacji i zachowania, które nie dotyczą rejestrowanych obiektów. Całkowicie się zgadzam, że należy zintegrować rejestrację bibliograficzną z rejestracją samych dokumentów, obiektów, ale nie rezygnować z tej pierwszej, bo ona wnosi porządek – wartość naddaną – do repozytorium, archiwum tekstów. Zresztą, autor referatu sam sugeruje, że możliwe i potrzebne byłoby rozszerzenie owego „łańcucha wartości”. Wnikliwego rozważenia, wykraczającego poza ograniczenia tej wypowiedzi, wymagają pytania drugie i trzecie. Bezpośrednia odpowiedź, której udzieliłabym na nie, brzmi tak. Ale diabeł tkwi w szczegółach, zwłaszcza, że jak słusznie zauważył prelegent, przeobrażenia badań naukowych już zaczynają chwiać tradycyjnie rozumianą publikacją jako podstawową jednostką opisu bibliograficznego.

W sesji wprowadzającej były jeszcze dwa inne referaty: Jadwigi Sadowskiej *Bibliografia dziedzinowa dzisiaj – jej rola i perspektywy* oraz Henryka Hollendra *Bibliografia dla nauki a nauka na krawędzi*.

Co robimy? Pierwszy z wymienionych referatów był rekapitulacją wiedzy na temat bibliografii dziedzin i poglądów J. Sadowskiej na ich przyszłość. W drugim H. Hollender zwrócił uwagę na, z jednej strony zjawisko stygmatyzowania pewnych typów wypowiedzi naukowych (np. recenzje), z drugiej zaś „zrytualizowane naśladownictwo”, prowadzące do „stopniowego psucia jakości tekstów” i poziomu czasopism („Czasopisma drapieżne i tandetne jako przepustka do świata wpływu [impact]). Wprowadził do wywodu pojęcie nauki aparadygmatycznej (post-normalnej) jako zacierającej granice między wiedzą naukową a jej zastosowaniami i wdrożeniami, „w obrębie której funkcjonują nierozwiązywalne sprzeczności pomiędzy «niepewnymi faktami, niezgodnionymi wartościami, wysokim ryzykiem i pilną potrzebą decyzji»”. Celnie zwrócił uwagę na uwikłania nauki „normalnej” stymulowane przez biurokrację, której nauka podlega, m.in. takie jak „błędy w wykorzystywaniu statystyki; nadmierna ufność, pokładana w data mining; bałamutne zachęty (publish or perish); przeciwnie skuteczne wskaźniki i pomiary”. W referacie znalazły też miejsce refleksje dotyczące rozpoznawania „jakości i doniosłości” tekstu (z udziałem bibliografa i algorytmów sztucznej inteligencji) oraz potencjalnego udziału bibliografii w dokumentowaniu tych zjawisk i ich efektów. Szeroko pojmowana racjonalność nauki, jak i jej udział w praktycznym opanowaniu i przekształcaniu świata moim zdaniem tworzą większą rewolucję niż jakakolwiek zmiana paradygmatu. Jeśli od nauki oczekuje się użyteczności, praktyczności, zastosowności, to prawdziwość wiedzy naukowej przestaje być konieczna (żeby spełnić warunek użyteczności). A to już na skrót prowadzi do tego, nad czym ubolewał H. Hollender. Osobiście relacje nauka normalna ÷ nauka post-normalna widzę jako komplementarną, a nie jako członki alternatywy. Produkty jednej i drugiej powinny być w sferze zainteresowań bibliografów, co nie oznacza, automatyzmu rejestracji bibliograficznej i jednych, i drugich.

Kolejną sesję wypełnił przegląd zasobów, użytkowników i użytkownika krajowych baz bibliograficznych: AGRO (Zofia Kasprzak, Mariusz Polarczyk, Krzysztof Gmerek *Przyrodnicze i rolnicze zasoby nauki udostępniane cyfrowo w bazie AGRO*), BazHum (Klaudia Grabowska *BazHum w działaniu. Badania użytkowników bazy czasopism*), BazEkon (Urszula Cieraszewska *BazEkon – bogactwo wiedzy i możliwości*), BazTech (Katarzyna Bikowska *Marketing społecznościowy czy społeczności? Promocja bibliograficznej bazy danych BazTech*), BEP (Inga B. Kuźma, Małgorzata Wilbik, Anna Deredas, Alicja Piotrowska, *Bibliografia Etnografii Polskiej (BEP) jako narzędzie pomocne w informacji naukowej i dydaktyce*), PEDAGOG (Dorota Wierzbička-Próchnika, Danuta Szewczyk *Baza PEDAGOG – czy dziedzinowe bazy danych nadal są potrzebne użytkownikom informacji naukowej?*), PBLekarska (Dorota Ubysz, Jacek Głębocki, *Polska Bibliografia Lekarska – perspektywy rozwoju*), PBLiteracka (Karolina Jaworska, Beata Koper, Tomasz Umerle, Piotr Wyciślik, *Polska Bibliografia Literacka – nowe wyzwania*).

Pierwszy dzień obrad zamknęła sesja zatytułowana „Bibliograficzne bazy danych – perspektywy” oraz dyskusja panelowa. W sesji tematycznej wystąpili Emanuel Kulczycki (*Rola baz danych w ocenie czasopism naukowych w Polsce – aktualne cele i wyzwania polityki naukowej*) oraz Marek Niezgodka (*Perspektywy rozwoju modeli otwartej nauki w Polsce*). Kulczycki w swoim referacie zwrócił uwagę m.in. na ważny aspekt udziału baz danych (bibliografii zawartości) nie tylko w ewaluacji periodyków, ale również w realnym kształtowaniu polityki naukowej w Polsce. Sformułował tezę, że żadne polskie bazy nie powinny być w żaden sposób brane pod uwagę w instytucjonalnej ewaluacji nauki w Polsce, natomiast powinny być finansowane oraz wspierane jako element systemu informacji i komunikacji naukowej. Końcowe pytania referatu, a wśród nich: Komu są potrzebne bibliograficzno-abstraktowe bazy danych w czasach Web of Science, Scopus, Google Scholar i Microsoft Academic? Kto używa baz dziedzinowych? Skoro klasyfikacje dziedzin i dyscyplin tracą znaczenie, to po co robić bazy dziedzinowe? Jak finansowanie baz przekłada się na realizację celów polityki naukowej? W jaki sposób bazy mogą wspierać inną sferę naukowej praktyki niż ewaluacja czasopism na potrzeby oceny instytucjonalnej? są niewątpliwie ważne i koniecznie należy im się namysł, choć nie można też nie zauważyć w nich pewnej dozy politykierstwa. Jak mantrę powtórzę, że reagowanie na aktualne (najlepiej rozumiane) zapotrzebowanie polityczne nie może być naczelną regułą przestrzeganą

w nauce, a także w bibliotekach. Wysoko wnosimy sztandar racjonalizacji wydatków, na przykład wstrzymując finansowanie wielkich i ważnych przedsięwzięć bibliograficznych, jak Bibliografia Historii Polskiej, a hojną ręką finansujemy kolejne tworzone bez składu i ładu biblioteki cyfrowe i inne paranaukowe „projekty”.

Punktem wyjścia wystąpienia Marka Niezgódki było stwierdzenie, że nauka (nie tylko) w Polsce jest w stanie kryzysu powodowanego m. in. jej systemową zamkniętością. Otwartość nadal jest w większym rozmiarze ideą niż praktyką. Zdaniem M. Niezgódki najważniejsze dylematy i wyzwania stojące przed otwartą nauką w Polsce to przede wszystkim: brak decyzji (niesprzecznych z działaniami oddolnymi) o krajowej infrastrukturze danych i publikacji, brak spójności decyzyjnej skutkującej wielokrotnym finansowaniem zbliżonych działań, brak świadomości znaczenia trwałości e-infrastruktur. W referacie zwrócono również uwagę na znaczenie spójnej koncepcji nowego krajowego modelu publikowania naukowego i dokumentowania wyników naukowych (otwarte platformy, zróżnicowane i otwarte modele recenzowania, uznanie integralnej przynależności dokumentacji badawczej do procesu naukowego i co za tym idzie konieczność jej publikacji itp.). Według M. Niezgódki w Polsce obserwujemy obecnie pogłębiające się cofnięcie w stosunku do międzynarodowych standardów ewaluacji instytucjonalnej i oceny indywidualnej aktywności naukowej. Potrzebujemy „zdrowych relacji krajowej e-infrastruktury informacji naukowej (publikacje, dane badawcze, dokumentacja itp.) i dostępu do zasobów zagranicznych w modelu gwarantującym trwałość”.

Pierwszy dzień prac konferencji zamknęła dyskusja panelowa „Gramy do jednej bramki” z udziałem Marka Niezgódki (ICM UW), Emanuela Kulczyckiego (UAM), Jana Kozłowskiego (MNiSZW), Małgorzaty Paszkowskiej (OPI PIB), Jolanty Przyłuskiej (Instytut Medycyny Pracy) i Danuty Szyszki (redakcja czasopisma „Mining Science”), moderowana przez Dorotę Buzdygan i Lidie Derfert-Wolf. Osiami problemowymi dyskusji, zaproponowanymi przez organizatorów, były następujące zagadnienia: polityka naukowa państwa, parametryzacja jednostek naukowych i ocena czasopism naukowych, współdzielenie krajowych zasobów informacyjnych, kierunki rozwoju baz bibliograficznych.

Drugiego dnia referaty dotyczyły roli baz w komunikacji naukowej oraz praktyk i standardów wydawniczych czasopism naukowych w Polsce. W pierwszej grupie znalazły się wypowiedzi Anety Drabek *Czasopisma Open Access w polskich bazach danych*, Edyty Strzelczyk *Otwarte dane badawcze – kolejny krok do otwierania nauki*, Małgorzaty Rychlik *Upowszechnianie dorobku naukowego w repozytoriach i bazach danych – działania komplementarne czy konkurencyjne?*, Marzeny Błaszczczyńskiej, Michała Kozaka, Cezarego Mazurka, Marcina Szymczaka i Marcina Werli *System SINUS – otwarte narzędzie do budowy bibliograficznych baz danych*. Z badań A. Drabek wynika m.in., iż bazy danych stanowią trzeci co do popularności kanał dostępu do informacji i tekstów publikacji naukowych. Jest to również odpowiedź na pytanie E. Kulczyckiego „Kto używa baz dziedzinowych?”. Badając dostępność do pełnych tekstów, referentka zauważyła, że informacja o możliwościach wykorzystania pełnych tekstów jest słabo widoczna, brak jest takiej informacji (lub jest niedostateczna) na stronach czasopism (266 z 885), w artykułach brak z kolei danych o licencji (uwaga ta nie dotyczy bazy BazHum). Referaty Edyty Strzelczyk i Małgorzaty Rychlik zawierały wiele cennych, niekiedy szczegółowych informacji, których nie sposób streścić w sprawozdaniu. Zachęcam do ich lektury (jak i wszystkich pozostałych tekstów), gdy zostaną opublikowane⁴.

W grupie referatów omawiających praktyki i standardy wydawnicze czasopism naukowych w Polsce znalazły się wystąpienia: Małgorzaty Kowalskiej *Platformy czasopism naukowych a bibliograficzne bazy danych: obszary przenikania, narzędzia, usługi*, Danuty Szyszki *Budowanie i wdrażanie strategii rozwoju czasopisma naukowego „Mining Science”*, Ewy Rozkosz *Wykorzystanie diagramów COPE w zasadach etyki publikacyjnej czasopisma naukowego*, Natalii Pamuły-Cieślak *Polskie czasopisma otwarte w Directory of Open Access Journal – aplikowanie, indeksowanie i dobre praktyki*. Powinny

⁴ Organizatorzy zapowiadają opublikowanie pełnych tekstów wystąpień online w serii „Materiały Konferencyjne EBIB” oraz w repozytorium e-LIS.

się z nimi zapoznać zwłaszcza osoby zaangażowane w wydawanie czasopism naukowych, chcące rozwijać ich jakość, doskonałość, otwartość. Warto zaznaczyć, że organizatorzy konferencji umożliwili udział w obradach licznemu gronu wydawców i redaktorów czasopism.

Z mojego punktu widzenia konferencja „Bibliograficzne bazy danych: perspektywy i problemy rozwoju” była niezwykle wartościowa i udana. Na sukces złożyło się kilka czynników, a wśród nich: staranne i szczegółowe opracowanie koncepcji konferencji i jej realizacja przez zespoły z bibliotek Politechniki Krakowskiej i Uniwersytetu Technologiczno-Przyrodniczego w Bydgoszczy, którym przewodzili Dorota Buzdygan, Marek Górski i Lidia Derfert-Wolf; znakomite grono prelegentów, którzy podjęli trud podzielenia się z licznie zebranymi uczestnikami (ponad 160 osób!) swoją wiedzą, doświadczeniem, poglądami; atmosfera sprzyjająca otwartej wymianie myśli i poglądów. Tylko poprzez dialog i wzajemne wsłuchiwanie się w swoje racje możemy zbliżyć się do prawdy, w tym również i tej dotyczącej obecnego i przyszłego statusu bibliografii, także bibliografii dziedzinowych. Na zakończenie dodam, że w moim odczuciu zaobserwowane różnice poglądów częściowo biorą się z pewnej nieprecyzyjności wypowiedzi. Czasami jest to efekt „nieznośnej lekkości” wypowiedziania się, a czasami braku szerszego spojrzenia na poruszane zagadnienia. Tak czy owak, szkoda, że na następną konferencję z tego cyklu trzeba będzie czekać cztery lata.

Jadwiga Woźniak-Kasperek

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

Nadesłano: 24 sierpnia 2017

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.iinsb@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. wykonywanych 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).

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 Informatycznych* 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 6th (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 6th. 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 zapisany kursywą, trzy kropki 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*. Medford, NJ: Information Today, 65–80.
- Gawrysiak, P. (2000). W stronę inteligentnych systemów wyszukiwawczych. W: Cz. Daniłowicz (red.) *Multimedialne i sieciowe systemy informacyjne*. Wrocław: Oficyna Politechniki Wrocławskiej, 59–69.

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
- 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). In: *Encyclopaedia Britannica*, Vol. 1, 13th ed., London and New York, NY: Encyclopaedia Britannica, 765–771.
- Iluminatorstwo (1971). W: *Encyklopedia Wiedzy o Książce*. Wrocław – Warszawa – Kraków: Zakł. Narod. im. Ossolińskich, 911–952.
- Big data. (2013, November 12). In: *Wikipedia, The Free Encyclopedia* [online] [12.11.2013], 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 164 a 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
- 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.iinsb@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 of the editor followed by abbreviation *Ed.*: (Name, Ed., Year of publication), eg. (Kocojowa, Ed., 2005). If there is no author or editor information, put the first word from the title in italics followed by ellipsis (...) and the year of publication : (Word..., Year of publication), eg. (*Libraries...*, 1995). Use the following pattern when referring to specific pages in the cited publications: (Dembowska, 1991, 15) or (Cisek & Sapa, 2007, 40–42) or (Kocojowa, Ed., 2005, 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*. Medford, NJ: Information Today, 65–80.

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

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). In: *Encyclopaedia Britannica*, Vol. 1, 13th ed., London and New York, NY: Encyclopaedia Britannica, 765–771.

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

Big data. (2013, November 12). In: *Wikipedia, The Free Encyclopedia* [online] [12.11.2013], 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 2017 r. – 118 zł

Wydawnictwo SBP – Warszawa 2017. Nakład 350 egz.

Ark. wyd. 15,72. Ark. druk. 12,75.

Skład i łamanie: Justyna Grzymała-Łuszcz

Druk i oprawa: Mazowieckie Centrum Poligrafii

ul. Piłsudskiego 2A, 05-270 Marki, www.c-p.com.pl

e-mail: biuro@c-p.com.pl, tel. 22 497 66 55

