

# ZIN

Studia Informacyjne  
Information Studies

VOL. 60 2022 NO. 2(120)

p-ISSN 0324-8194

e-ISSN 2392-2648



**WDiB UW**



STOWARZYSZENIE  
BIBLIOTEKARZY  
POLSKICH

## REDAKCJA | EDITORIALS

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

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

Wersja elektroniczna (<http://ojs.sbp.pl/index.php/zin>) jest wersją pierwotną czasopisma.  
Zawartość czasopisma jest indeksowana w CEJSH, CEEOL, CSA, LISTA, Knowledge Organization Literature, PBN, PBB.

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

Pełne teksty roczników 1972–2020 znajdują się w Archiwum SBP: <http://www.ac.sbp.pl/>;  
pełne teksty roczników 2009–2021 znajdują się na Platformie OJS  
Czasopism Naukowych SBP: <http://ojs.sbp.pl/index.php/zin/issue/archive>

**ZAGADNIENIA  
INFORMACJI  
NAUKOWEJ**  
Studia Informacyjne

**ISSUES IN  
INFORMATION  
SCIENCE**  
Information Studies

VOL. 60 2022 NO. 2(121)  
p-ISSN 0324-8194  
e-ISSN 2392-2648



**WDIiB UW**



STOWARZYSZENIE  
BIBLIOTEKARZY  
POLSKICH



Warszawa 2023

## ISSUES IN INFORMATION SCIENCE – INFORMATION STUDIES

The core purpose of *Issues in Information Science – Information Studies* (*Zagadnienia Informatyki Naukowej – Studia Informacyjne*, ZIN – *Studia Informacyjne*) 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.

We publish papers in Polish or English. For each paper a set of metadata is provided: an abstract and keywords in both languages) as well as author's bio and contact information.

The subtitle of the journal – *Information Studies* – 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 – *Information Studies* 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 ZIN – *Information Studies* covers, among else, the issues of:

- information services in institutions of science, culture, business, education and administration,
- 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,
- culture of information,
- information, digital and media 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 the list of journals scored by Polish Ministry of Science and Higher Education and indexed by: Central European Journal in Social Sciences and Humanities (CEJSH), Central and Eastern European Online Library (CEEOL), Cambridge Scientific Abstracts (CSA), Library and Information Science and Technology Abstracts (LISTA), Polish Bibliography of Book Studies (PBB), Knowledge Organization Literature, Worldcat and Polish Scholarly Bibliography (PBN). The journal is registered in the European Reference Index for the Humanities (ERIH Plus).

## ZAGADNIENIA INFORMACJI NAUKOWEJ – STUDIA INFORMACYJNE

Głównym celem *Zagadnień Informatyki Naukowej – Studiów Informatycznych* (ZIN – *Studia Informatyczne*) 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.

Publikujemy artykuły w językach polskim i angielskim. Każdy artykuł posiada zestaw metadanych: abstrakt i słowa kluczowe (w obu językach) oraz nota biograficzna autora i dane do kontaktu z nim.

Czasopismo adresowane jest zarówno do czytelnika polskiego jak i zagranicznego, publikujemy artykuły zarówno w języku polskim jak i angielskim. Podtytuł czasopisma – *Studia Informatyczne* – 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 Informatyczne* obejmuje m.in. problematykę:

- usług informacyjnych w instytucjach nauki, kultury, biznesu, edukacji i administracji,
- 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,
- kultura informacji,
- kompetencji informacyjnych i cyfrowych,
- polityki informacyjnej,
- etyki informacyjnej.

*Zagadnienia Informatyki Naukowej – Studia Informatyczne* 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 różnymi aspektami funkcjonowania informacji we współczesnym świecie.

*Zagadnienia Informatyki Naukowej* znajdują się na liście czasopism punktowanych Ministerstwa Nauki i Szkolnictwa Wyższego. Czasopismo jest indeksowane w bazach: Central European Journal in Social Sciences and Humanities (CEJSH), Central and Eastern European Online Library (CEEOL), Cambridge Scientific Abstracts (CSA), Library and Information Science and Technology Abstracts (LISTA), Polska Bibliografia Bibliologiczna (PBB), Knowledge Organization Literature, Worldcat, Polska Bibliografia Naukowa (PBN). Czasopismo jest zarejestrowane w European Reference Index for the Humanities (ERIH Plus).



# Spis treści | Contents

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

- Arkadiusz Pulikowski  
*Full-Text Search in the Resources of Polish Digital Libraries*  
[Wyszukiwanie pełnotekstowe w zasobach polskich bibliotek cyfrowych] ..... 7
- Łukasz Iwasiński, Weronika Furman  
*Jak być świadomym użytkownikiem algorytmów? O potrzebie rozwijania kompetencji algorytmicznych* [How to Be a Conscious User of Algorithms? On the Necessity of Developing Algorithmic Literacy] ..... 25

## ZAGADNIENIA ARCHITEKTURY INFORMACJI | ISSUES IN INFORMATION ARCHITECTURE

- Paulina Sajna-Kosobucka  
*Architektura informacji w witrynach internetowych Prawa i Sprawiedliwości oraz Platformy Obywatelskiej* [Information Architecture on the Websites of the Law and Justice, and the Civic Platform] ..... 44
- Łukasz Tarnacki  
*Analiza architektury informacji platformy Steam*  
[Analysis of the Information Architecture of the Steam Platform] ..... 72

## RECENZJE I OMÓWIENIA | REVIEWS

- Monika Krakowska (2022). Zachowania informacyjne człowieka w kontekście zjawiska epistemicznej bańki informacyjnej. Propozycja nowej koncepcji* [Human Information Behaviour in the Context of the Epistemic Information Bubble Phenomenon. In Search of a New Concept]. Kraków: Wydaw. Uniwersytetu Jagiellońskiego (Barbara Sosińska-Kalata) ..... 90

- WSKAZÓWKI DLA AUTORÓW | GUIDELINES FOR AUTHORS ..... 97



# Full-Text Search in the Resources of Polish Digital Libraries

Arkadiusz Pulikowski

ORCID 0000-0003-1807-8642

*Institute of Culture Studies, Faculty of Humanities  
University of Silesia in Katowice, Poland*

---

## Abstract

**Purpose/Thesis:** The article aims to analyze the conditions and possibilities of full-text search in Polish digital libraries (DL), taking into account the access to full-text search in individual DLs, the file formats and the software used, as well as the visibility of the DL resources in the Google search engine.

**Approach/Methods:** Forty of the largest Polish DLs, whose resources primarily comprise digitized traditional library resources, were selected for the study. The study examined the type of the software used, the availability and efficiency of the full-text search, and the extent to which the resources were indexed as PDF files in Google and Google Scholar. Finally, the study compared the results of full-text search in ten DLs with those obtained from Google.

**Results and conclusions:** All DLs selected for the study allow for full-text search. There are significant differences between specific interfaces. Each has advantages and disadvantages that require further development. The Google search engine is not currently a viable alternative to accessing content provided in DLs.

**Originality/Value:** The issue of full-text search in DLs is rarely addressed, even though users consider it one of the most important functions of DLs. The result of the study presents a picture of DL's current capacity in this area.

## Keywords

Content Visibility. Digital libraries. Full-text search. Google search engine. Information retrieval.

*Received: 2 December 2022. Reviewed: 11 December 2022. Accepted: 20 January 2023.*

---

## 1. Introduction

1st of October 2022 marked the twentieth anniversary of establishing the Greater Poland Digital Library (Pl. *Wielkopolska Biblioteka Cyfrowa*). It was the first digital library (DL) to employ the dLibra software, developed by the Poznan Supercomputing and Networking Center (Pl. *Poznańskie Centrum Superkomputerowo-Sieciowe* – PSNC). Its dissemination in the following years allowed for a dynamic

growth of similar projects all over the country (Kolasa, 2007). Over the last two decades, the number of institutions hosting or co-hosting DLs grew, as did the number of the objects they made accessible. The user base increased, too. A significant part of the users – c. 50% (Parkoła et al., 2016, 5) – are academics, who use the resources of DLs for their research. These are mostly historians, linguists, and scholars pursuing literary and culture studies. They were quick to appreciate the convenience of remote access ensured by the DLs, allowing them to use the resources at any time and place.

With time, the traditional DLs (based on library resources) hosted by the dLibra platform were joined by the digital resources of museums and multimedia archives, as well as scientific repositories. The repositories were created to share current scientific resources, unlike the digital libraries, museums and archives, which were established in order to protect and promote the premodern artefacts collected by the scientific and cultural institutions (Bednarek-Michalska, 2017, 47). Accordingly, the repositories' collections are dominated by objects which were 'born digital,' while the libraries, museums and archives focus on the digitized images of stored objects.

Digitized collections of various Polish institutions are scattered in more than one hundred and fifty DLs of various types (as per data collected in November 2022)<sup>1</sup>. The resources of these DLs might be browsed simultaneously thanks to the Federation of Digital Libraries aggregator (Pl. *Federacja Bibliotek Cyfrowych* – FBC, [fbc.pionier.net.pl](http://fbc.pionier.net.pl)). Established in 2007, the aggregator allows the user to browse the resources hosted on the dLibra platform, or by the means of another type of software system supporting the OAI-PMH protocol (Open Archives Initiative Protocol for Metadata Harvesting). Using the protocol, the database systematically harvests the metadata describing the objects from the DLs, and creates a single summary nationwide collection, which might be browsed on the FBC website (Lewandowska et al., 2007).

Over the last fifteen years, the FBC search engine has undergone many changes; the most apparent shift occurred in 2015, when the engine's interface was completely reconstructed, allowing for faceted filtering of the results. The current interface is clear and user-friendly, supporting efficient research. The only significant limitation of the database is the absence of the full-text search option. Because the system aggregates only the metadata, rather than the objects' content, full-text search is impossible.

---

<sup>1</sup> Digital libraries, digital museums, digital archives, scientific repositories. All these resources are often referred to as digital libraries, although it is not semantically correct. The practice is partly due to the name of the FBC platform, which was established at a time when the only digital resources were the resources of libraries, and partly due to the lack of a better alternative, which would encompass the digital resources of other types of institutions. This article will continue to use this simplified term, excluding the analysis which will consider only one type of digital libraries.

In order to search the content of the files stored in the scattered digital libraries, it is necessary to conduct separate searches on the websites of specific digital libraries. It is a time-consuming and challenging process. Given that the full-text search of files hosted by DLs is widely recognized as one of the websites' most useful functions (Parkola et al, 2016, 8), we should consider the reality of conducting such searches and potential alternatives.

The aim of the article is to analyze the conditions and possibilities of full-text search in Polish digital libraries (DL), taking into account the availability of full-text search in individual DL, the file formats and the software used, as well as the visibility of the DL resources in the Google search engine.

## 2. Selection of the DLs for the study

For the purposes of this study, it was necessary to make a selection out of over one hundred and fifty DLs registered in the FBC database. The chief criterion was the collection's emphasis on text files (a prerequisite for the full-text search). The FBC's list of sources is dominated by the traditional DLs which digitized their collections. These constitute over a half of the sources available via the FBC. DLs of this type were selected for analysis not only because of their uniform character, allowing for a comparative analysis, but primarily because of the issues with the full-text search that have been consistently obstructing access to the content of the files since the incipience of the DLs (Pulikowski, 2009).

Purely scientific resources, gathered in repositories and on the websites of scientific journals and publishers, are hosted on separate aggregators, such as Google Scholar, which allows its users to search not only the descriptions of the files, but also their contents. The functionality encompasses not only those publications which were created digitally, but also those that have been digitized, with the OCR layer added. The DLs created on the basis of traditional library resources contain files of various theme and origin, which by definition renders impossible the project of indexing them in their entirety on Google Scholar. It is viable only for the scientific sources (see chapter 5).

Taking into account these circumstances, the DLs selected from the FBC register for the purposes of this study do not include repositories, platforms of scientific journals and publishers, as well as digital museums and multimedia archives focusing on non-textual objects<sup>2</sup>. The analysis was conducted on the CSV file downloaded from the FBC website. A further selection was necessary to meet

---

<sup>2</sup> With the exception of the Multimedia Library of the NN Theatre, which contains a rich collection of text publications. The isolated "Library" collection comprising books, journals, articles, studies, etc., amounts to over 26 000 objects.

the time constraints, and to yield reliable data. To that end, the DLs were sorted according to the number of available objects in a descending order. It was assumed that analyzing the largest DLs will be the best solution, as such an analysis will cover the largest summary number of objects. The preliminary cut-off mark was set at 10 000 objects.

The FBC aggregator, collecting data concerning DLs nationwide, seemed like the perfect tool for selecting the DLs. It served its purpose as a vast register, but its data were revealed to be far from accurate. The disparities between the number of objects noted by the FBC and the websites of specific DLs came up to thousands (occasionally – ten thousands), both over and below. The only solution was to collect the data concerning the size of the collections from the DLs' websites over the course of a single day, to ensure that the selection would be valid. The analysis was conducted on November 12th, 2022. All DLs available in the FBC register were examined as otherwise it would not have been possible to determine which DLs meet the accepted criteria regarding the size of the collection. During the process of data collection, it turned out that some of the links in the FBC are expired, or inaccessible. It was doubtful that the register was complete; therefore, the author supplemented his data with those in an extensive DL register developed by Barbara Morawiec, hosted on the *Lustro Biblioteki* (En. *The Library Mirror*) portal (Morawiec, 2021). As a result, the final list encompassed 69 DLs collecting at least 1000 objects, among which 40 DLs collected as many as 10 000 objects, thus meeting the criteria for analysis. The decision to set the cut-off mark at 10 000 proved sound, as the next collection was significantly smaller (7400 objects, Krośniewska DL).

Definite majority of the 40 largest Polish DLs is hosted on the dLibra platform. Only three DLs use their own software (Polona, CRISPA, Polonijna DL). There are three versions of dLibra: 4.x – used by two DLs; 5.x – five DLs; 6.x – thirty-three DLs. The newest version, 6x, is predominant. It was introduced in 2016, and significantly improved on the previous versions (Parkoła et al., 2016, 10–20). Despite its high price, not only the largest DLs in the country chose to purchase the software. When confirming the real number of objects in individual DL, the author also identified the software used in the remaining 29 DLs with collections numbering more than 1000 and less than 10 000 objects. Ten DLs use dLibra 6.x; twelve – 5.x; three – 4.x. Four DLs use alternative software<sup>3</sup>. The information concerning the number of objects and the software used in the DLs selected for further analysis is presented in Table 1.

---

<sup>3</sup> Wolne Lektury (Free Schooltexts) – own software; DL of the Siedlce University of Natural Sciences and Humanities – Dspace; Digital Statistical Library – Aleph; DL of Sieradz Land – Sowa. The last two cases use integrated modules of library systems adjusted for the needs of DLs.

Tab. 1. The number of objects and the software used in the DLs selected for further analysis

| No. | DL name   | Number of objects | Software |
|-----|---|-------------------|----------|
| 1.  | Polona  | 3720884           | own      |
| 2.  | Jagiellonian Digital Library   Jagiellońska Biblioteka Cyfrowa                                      | 854476            | 6.3.14   |
| 3.  | Silesian Digital Library   Śląska Biblioteka Cyfrowa  | 528369            | 6.2.11   |
| 4.  | CRISPA University of Warsaw Digital Library   CRISPA Biblioteka Cyfrowa Uniwersytetu Warszawskiego  | 464576            | own      |
| 5.  | Greater Poland Digital Library   Wielkopolska Biblioteka Cyfrowa                                    | 396783            | 6.3.13   |
| 6.  | Kujawsko-Pomorska Digital Library   Kujawsko-Pomorska Biblioteka Cyfrowa                            | 252073            | 6.3.16   |
| 7.  | Małopolska Digital Library   Małopolska Biblioteka Cyfrowa  | 130535            | 5.8.5    |
| 8.  | NN Theatre Multimedia Library   Biblioteka Multimedialna Teatru NN                                  | 127861            | 6.2.14   |
| 9.  | Digital Library of University of Wrocław   Biblioteka Cyfrowa Uniwersytetu Wrocławskiego            | 119472            | 6.3.15   |
| 10. | Regional Materials of Łódź Land   Regionalia Ziemi Łódzkiej   | 113486            | 6.2.9    |
| 11. | Pomeranian Digital Library   Pomorska Biblioteka Cyfrowa  | 106691            | 6.3.16   |
| 12. | Lower Silesian Digital Library   Dolnośląska Biblioteka Cyfrowa                                     | 92701             | 6.3.15   |
| 13. | Digital Library of the University of Lodz   Biblioteka Cyfrowa Uniwersytetu Łódzkiego               | 90702             | 6.2.14   |
| 14. | Mazovian Digital Library   Mazowiecka Biblioteka Cyfrowa  | 85727             | 5.8.4    |
| 15. | Baltic Digital Library   Bałtycka Biblioteka Cyfrowa  | 72439             | 6.2.14   |
| 16. | Digital Library of the Silesian University of Technology   Biblioteka Cyfrowa Politechniki Śląskiej | 70712             | 6.3.16   |
| 17. | Elbląg Digital Library   Elbląska Biblioteka Cyfrowa  | 67071             | 6.3.15   |
| 18. | Digital Library KUL   Biblioteka Cyfrowa KUL  | 66292             | 6.2.8    |
| 19. | Podlaska Digital Library   Podlaska Biblioteka Cyfrowa  | 58423             | 6.3.13   |
| 20. | KARTA Center Digital Library   Biblioteka Cyfrowa Ośrodka KARTA                                     | 54542             | 6.3.15   |
| 21. | Świętokrzyska Digital Library   Świętokrzyska Biblioteka Cyfrowa                                    | 49816             | 6.0.2    |
| 22. | The West Pomeranian Digital Library 'Pomerania'   Zachodniopomorska Biblioteka Cyfrowa              | 48421             | 6.2.12   |
| 23. | Lower Silesian Digital Library   Cyfrowy Dolny Śląsk  | 47200             | 6.2.14   |
| 24. | Digital Library of Zielona Góra   Zielonogórska Biblioteka Cyfrowa                                  | 47016             | 6.2.12   |
| 25. | UMCS Digital Library   Biblioteka Cyfrowa UMCS  | 44991             | 6.2.13   |
| 26. | Radom Digital Library   Radomska Biblioteka Cyfrowa   | 44659             | 6.2.13   |

| No. | DL name   | Number of objects | Software |
|-----|---|-------------------|----------|
| 27. | Digital Library of Provincial Public Library in Lublin   Biblioteka Cyfrowa Wojewódzkiej Biblioteki Publicznej w Lublinie | 28394             | 5.8.3    |
| 28. | Podkarpacka Digital Library   Podkarpacka Biblioteka Cyfrowa  | 24131             | 6.1.3    |
| 29. | Lodz Regional Digital Library   Łódzka Regionalna Biblioteka Cyfrowa  | 20914             | 6.2.11   |
| 30. | Polonia Digital Library   Polonijna Biblioteka Cyfrowa  | 20795             | own      |
| 31. | Opole Digital Library   Opolska Biblioteka Cyfrowa  | 19754             | 6.2.9    |
| 32. | Bialska Digital Library   Bialska Biblioteka Cyfrowa  | 17996             | 6.2.11   |
| 33. | Digital Library in Leszno   Leszczyńska Biblioteka Cyfrowa  | 17902             | 6.3.13   |
| 34. | Warmia-Mazury Digital Library   Warmińsko-Mazurska Biblioteka Cyfrowa   | 13727             | 6.2.11   |
| 35. | Digital Library of Lublin University of Technology   Biblioteka Cyfrowa Politechniki Lubelskiej                           | 13473             | 6.0.2    |
| 36. | Wejherowo Digital Library   Wejherowska Biblioteka Cyfrowa  | 13440             | 4.0      |
| 37. | Military Digital Library   Wojskowa Biblioteka Cyfrowa  | 13365             | 5.8.4    |
| 38. | Digital Library of Warsaw University of Technology   Biblioteka Cyfrowa Politechniki Warszawskiej                         | 12587             | 6.3.16   |
| 39. | Chelm Digital Library   Chełmska Biblioteka Cyfrowa   | 11865             | 4.0      |
| 40. | Digital Library of Inowroclaw   Inowrocławska Biblioteka Cyfrowa  | 10114             | 5.8.2    |
|     | Total   | 7994375           |          |

### 3. The conditions for full-text search in DLs

Full-text search requires a text. This obvious statement becomes more meaningful when we consider the files created during digitization, which constitute a major part of the traditional DLs' collections, analyzed in this paper. Only the files processed by OCR (Optical Character Recognition) might be subject to a full-text search. The recognized text is usually inscribed as an invisible layer 'beneath' the image of a given page, where the recognized text matches the text visible on the image. This allows for highlighting the searched text and copying it, as it is the case with the documents created digitally.

It is difficult to determine what part of the DL collections is constituted by files with the OCR layer. This information is not a part of the available metadata. Considering that OCR has been employed in large DLs since they were established, we might assume that it features in the majority of their files. This might not be the case for the DLs established by smaller institutions, which might not have the software required for creating the OCR layer. The largest Polish DL – Polona,

which employs its own software for managing its DL, additionally marks the OCR files, which allows the user to narrow the scope of query in advanced search (“Search only objects with a text layer available”), and filtering the results (“text layer available”). Thanks to the marking, it is easy to determine that two thirds of texts available online in Polona have the OCR layer.

Scans/images with a hidden OCR layer are saved as PDF or DjVu files. The PDF is considered to be a standard file for saving information; it is widely recognized and supported by a variety of software and systems. The opposite is the case for DjVu which, despite its many advantages, remains an obscure file format and requires dedicated software. Many users encountered it for the first time in a DL. For many it was a cause for frustration, as they could not open the file and were obliged to install additional extensions for their internet browsers. Furthermore, it was soon impossible to use the extensions, as the browser security updates disabled them. To counteract these problems, in 2015 dLibra introduced a programming library for its version 5, which allowed the users to open the DjVu files without installing any further extensions (Parkoła et al., 2016, 14). The library is also available in the version 6 of dLibra. But the users of the older version 4 still have to face the challenge of opening the DjVu files.

The DjVu file format has not been updated in a long time, and presently is recommended only for very large documents. It has been widely used during the initial stages of digitization in Polish DLs, when its compression capacity was significantly higher than that of a PDF. Even ten years ago, DjVu files constituted almost 80% of all sources in Polish DLs (Szafranski, 2013, 8). It is difficult to determine what part of all sources they constitute now, but according to the author’s calculations, there are at least 1.5 million DjVu files, which would suggest that they will remain in use at Polish DLs<sup>4</sup>.

The full-text search depends not only on the presence of the OCR layer, but also on its efficacy. The accuracy of text recognition in scans/images of documents depends on many factors: the quality of scans/images (resolution, colour depth, dynamic range of the scanner/camera), the condition of the document (stains, discoloration, deformation), the type of paper and ink used (print contrast, bleed-through from the other side of the page), the text (font, language) and the OCR software used (Kotyńska, 2013, 8–9).

The software is a key factor here, as inappropriate selection or inappropriate use might result in errors depicted in Figure 1. Unfortunately, there are many documents in Polish DLs with weak OCR, which means that some part of the text remains unsearchable. The problem affects primarily the files in the DjVu format, created during the digitizing boom in the first decade of the 21st century.

---

<sup>4</sup> Calculations were based on advanced search for 40 DLs studied according to format.

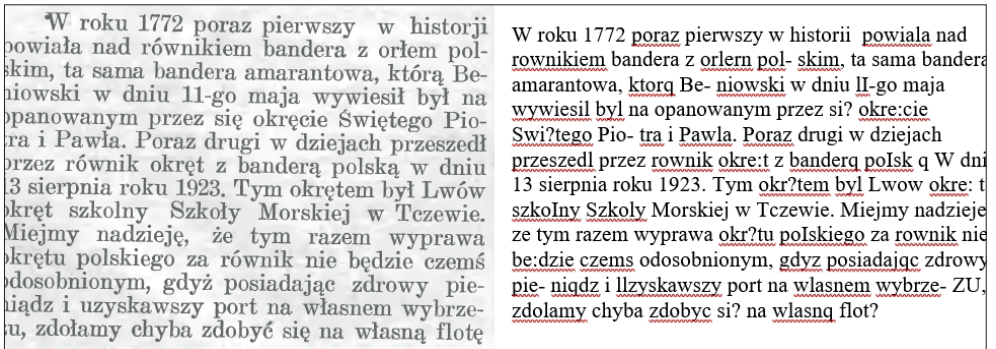


Fig. 1. Problems with text recognition in digitized documents.

Source: own transcription based on: <https://kpbc.umk.pl/dlibra/publication/87307/edition/92270/content>

As it was already mentioned, OCR is not available for all digitized files. While in some cases it is a consequence of lacking OCR software, in other cases text recognition simply does not make sense, particularly for manuscripts, incunabula, maps, and other documents for which OCR accuracy does not reach higher than 50% (Szala, 2015, 267). We might also find files in the PDF or the DjVu format used for storing non-textual objects, primarily photographs. The DjVu file will always have a graphic layer, while the PDF file might not. The PDF is also the format for storing files created digitally. These texts may always be indexed for the purposes of full-text search. Files created digitally might be accessed in DLs in other formats, such as TXT, HTML, EPUB, or MOBI. These might be equally easily searched. Increasingly often, the user finds a single publication in several formats, which is possible thanks to the flexibility of modern DL platforms.

#### 4. The possibilities for full-text search in DLs

As it is impossible to conduct a full-text search of scattered DL resources from a single access point, the users must rely on search engines of individual DLs. It is most likely that they primarily use the largest Polish DL – Polona, the digital library of the National Library in Poland. The regional DLs, and the DLs of the institutions where the users are employed, might be equally useful. Nevertheless, if a query concerning the content of the files is to be exhaustive, then it will be necessary to repeat it in all DLs. Fortunately, majority of DLs is hosted on the dLibra platform, which means that the users will interact with only two interfaces – version 6.x, and the earlier 5.x and 4.x.

All analyzed DLs, presented in Table 1, allow for full-text search. The functionality was tested on single terms and phrases. Regardless of the software used, there were

no serious issues with searching the text of the files. The analysis was conducted between 11th and 14th of November 2022. The remaining part of this chapter discusses the author's experiences using individual searching interfaces.

On the Polona website, the option of full-text search is marked on the right side of the search bar as a box described "Search the content". After selecting this option and entering the query, detailed results load (it usually takes a moment) – see Figure 2. The users might additionally apply filters listed on the left side, and select another method of sorting the results. Under the description of each file, which takes up a large part of the screen, the user will find the fragment of the sentence where the searched phrase occurs, highlighted in blue. The user might see further matches clicking "See more" on the right side. When the query is a full phrase, rather than a single term, one should apply quotation marks; nevertheless, the suggested matches will include sentences where only one element of the phrase will appear. It is particularly bothersome where prepositions are concerned, for example in the phrase "w piaskowej skale" (in Pieskowa Skala Castle), where "w" (in) will be considered as a significant term and generate dozens, if not hundreds of results. What is convenient is the transposition of the information concerning found content to the full text view. The advanced search allows for generating more complex queries, which use Boolean operators. The Polona website does not have a mobile version and requires manual scaling. Since November 2022, mobile users may use the mPolona app, optimized for phones. Unfortunately, for the time being full-text search in the app yields an error page instead of a list of records (state in December 2022).

The screenshot shows the Polona website search results for the query "piaskowa skale". The search bar at the top contains the query and a search icon. The results are displayed in a grid format. The first result is a map titled "Atlas dróg samochodowych w Polsce [1939]". The second result is a book titled "Wierzbowa : Automobilklub Polska, [1935]". The third result is a book titled "Słownik etymologiczny języka polskiego [1927]". The search results are filtered by "Wyniki: 3259". The left sidebar contains a list of filters, including "Filtr wyszukiwania" and "Kategorie". The right sidebar contains a "Znaleziono" section with a "Public Domain" logo and a "Drogi 1918-1939" section.

Fig 2. Search results in Polona.

Source: polona.pl

The 6.x dLibra, predominant in the largest Polish DLs (33 out of 40) provides the full-text search option in the advanced search. The users must select the “Search content” field. After entering the phrase, there appears (after a short while) a list of results, identical to the results yielded by searches of other types, allowing for dynamic filtering, sorting, etc. However, the results do not display the fragments of texts where the searched phrase appear. The user has to open the full text, and search the phrase again. The search is effective both for single terms and for whole phrases. There is no issue with searching the content of DjVu files. Searching further occurrences of a given term, we move to further pages of the file. The search engine automatically searches different grammatical forms of the searched terms and phrases. This should be borne in mind when searching the content of the files – it will be more effective to type in the root of the word. Constructing the query, the user may apply many Boolean operators expanding the possibilities of full-text search. The possible operators are suggested after clicking the question mark next to the search bar. The mobile version of the dLibra 6.x is well-matched to the size of the device and does not hinder the search.

The screenshot displays the CRISPA search interface. At the top, the logo for CRISPA (biblioteka cyfrowa Uniwersytetu Warszawskiego) is visible alongside navigation links: STRONA GŁÓWNA, KOLEKCJE, and POMOC. Utility links for font size (Rozmiar: A A A), contrast (Kontrast: A A A), and language (Język: PL EN) are also present. The search bar contains the text 'Szukaj w treści'. Below the search bar, the search results are shown for the query 'pieskowa skała', indicating 1781 results. The results are organized with a 'Filtrowanie wyników' section on the left, including a 'Twórca' (Author) filter with a list of authors and their counts (e.g., (936) Dmuszewski, ...). A 'Zakres dat' (Date range) filter is also visible, set from 1756 to 2018. The main results area shows two entries. The first entry is from 'GONIEC WIECZORNY' and includes a snippet: '... Pieskowa Skała ... Pokaż źródło: PDF'. The second entry is from 'KURIER WARSZAWSKI' and includes a snippet: '... Pieskowa Skała ... Pokaż źródło: PDF'. The search terms 'Pieskowa Skała' are highlighted in red in the snippets.

Fig 3. CRISPA search results.  
Source: [crispa.uw.edu.pl](http://crispa.uw.edu.pl)

The full-text search option in the former versions of dLibra is not much different. A slightly outdated interface allows for conducting analogous searches (without the option to filter the results). However, there might be problems with searching the content of DjVu files; in the fourth version, these files might fail to open. Curiously enough, the results in the 4.x version include an accuracy coefficient which has not been included in the later versions. Unlike the sixth version of dLibra, the earlier versions are not optimized for mobile use.

When we consider the largest DLs in the country, we should note the DL of the University of Warsaw, CRISPA, which uses its own software. The full-text search is conducted in a separate page, selected above the search bar. It takes time to generate the results, but they are well worth the wait. The results (Fig. 3), next to the cover of each document, show a nested text box, containing the fragments of the text with the search terms highlighted. This form of presenting the results makes it much easier to select files for further analysis. Unfortunately, the searched fragments are not transposed onto the full text view, as in Polona. It is necessary to search the text of each opened file. CRISPA takes into account the grammatical variations of each term, allows the user to use quotation marks to search by phrases, but unfortunately does not recognize other Boolean operators. The mobile version is scaled and fully operational.

## 5. Full-text search in DLs by Google

As conducting a full-text search in DLs is very time-consuming, we might consider the alternative offered by the Google search engine. Thanks to the modifiers *site:* and *filetype:*, it is easy to determine to what extent Google indexes the DL resources. The *site:* modifier limits the search to a given domain; the *filetype:* modifier limits the results to a selected format, here, the PDF. Selecting a single format is necessary to avoid duplicated results – sites describing the files, and the files themselves. It is impossible to conduct an analogous search for the DjVu files as Google does not recognize the format.

The original plan was to compare the results of a Google search with the number of the PDF files in selected DLs. However, it turned out that it is impossible to determine the data credibly for such a high number of DLs. Furthermore, the author was forced to conclude that the results of the Google search might be considered only as estimations, as the numbers change every few days. Despite these limitations, it was possible to establish an approximate number of PDF files that Google “sees” in the DLs. It was also possible to confirm the presence of these documents in Google Scholar (GS).

The search was conducted on November 16th, 2022, and repeated after a week in order to average the results. The search term was as follows: “site:the DL-domain filetype:pdf”.

For the search conducted in GS, the term only included site: as GS groups duplicate results together. The results are presented in Table 2.

Tab. 2. The presence of Polish DL resources in Google and Google Scholar

| No. | DL name   | Google PDF 1 | Google PDF 2 | Google PDF avg. | GS 1 | GS 2 |
|-----|---|--------------|--------------|-----------------|------|------|
| 1.  | Polona  | 8            | 8            | 8               | 0    | 0    |
| 2.  | Jagiellonian Digital Library   Jagiellońska Biblioteka Cyfrowa                                      | 17300        | 16400        | 16850           | 749  | 751  |
| 3.  | Silesian Digital Library   Śląska Biblioteka Cyfrowa  | 16200        | 16300        | 16250           | 0    | 0    |
| 4.  | CRISPA University of Warsaw Digital Library   CRISPA Biblioteka Cyfrowa Uniwersytetu Warszawskiego  | 18           | 17           | 18              | 0    | 0    |
| 5.  | Greater Poland Digital Library   Wielkopolska Biblioteka Cyfrowa                                    | 10700        | 11800        | 11250           | 2290 | 2290 |
| 6.  | Kujawsko-Pomorska Digital Library   Kujawsko-Pomorska Biblioteka Cyfrowa                            | 17100        | 18100        | 17600           | 391  | 390  |
| 7.  | Małopolska Digital Library   Małopolska Biblioteka Cyfrowa  | 4870         | 4010         | 4440            | 0    | 0    |
| 8.  | NN Theatre Multimedia Library   Biblioteka Multimedialna Teatru NN                                  | 7110         | 6770         | 6940            | 185  | 186  |
| 9.  | Digital Library of University of Wrocław   Biblioteka Cyfrowa Uniwersytetu Wrocławskiego            | 15600        | 14900        | 15250           | 584  | 593  |
| 10. | Regional Materials of Łódź Land   Regionalia Ziemi Łódzkiej   | 4850         | 4360         | 4605            | 17   | 17   |
| 11. | Pomeranian Digital Library   Pomorska Biblioteka Cyfrowa  | 7920         | 7910         | 7915            | 434  | 434  |
| 12. | Lower Silesian Digital Library   Dolnośląska Biblioteka Cyfrowa                                     | 11900        | 12800        | 12350           | 3270 | 3310 |
| 13. | Digital Library of the University of Lodz   Biblioteka Cyfrowa Uniwersytetu Łódzkiego               | 32           | 47           | 40              | 14   | 14   |
| 14. | Mazovian Digital Library   Mazowiecka Biblioteka Cyfrowa  | 9250         | 9750         | 9500            | 38   | 35   |
| 15. | Baltic Digital Library   Bałtycka Biblioteka Cyfrowa  | 8330         | 7830         | 8080            | 35   | 35   |
| 16. | Digital Library of the Silesian University of Technology   Biblioteka Cyfrowa Politechniki Śląskiej | 19400        | 19300        | 19350           | 3530 | 3520 |

| No. | DL name   | Google PDF 1 | Google PDF 2 | Google PDF avg. | GS 1 | GS 2 |
|-----|---|--------------|--------------|-----------------|------|------|
| 17. | Elbląg Digital Library  <br>Elbląska Biblioteka Cyfrowa   | 1240         | 1320         | 1280            | 60   | 60   |
| 18. | Digital Library KUL  <br>Biblioteka Cyfrowa KUL   | 8350         | 8610         | 8480            | 441  | 441  |
| 19. | Podlaska Digital Library  <br>Podlaska Biblioteka Cyfrowa   | 8540         | 9080         | 8810            | 539  | 537  |
| 20. | KARTA Center Digital Library  <br>Biblioteka Cyfrowa Ośrodka KARTA  | 23           | 24           | 24              | 0    | 0    |
| 21. | Świętokrzyska Digital Library  <br>Świętokrzyska Biblioteka Cyfrowa   | 3560         | 3900         | 3730            | 4    | 4    |
| 22. | The West Pomeranian Digital Library<br>'Pomerania'  <br>Zachodniopomorska Biblioteka Cyfrowa                                      | 865          | 718          | 792             | 123  | 123  |
| 23. | Lower Silesian Digital Library  <br>Cyfrowy Dolny Śląsk   | 3800         | 3900         | 3850            | 33   | 33   |
| 24. | Digital Library of Zielona Góra  <br>Zielonogórska Biblioteka Cyfrowa   | 4590         | 4580         | 4585            | 33   | 2820 |
| 25. | UMCS Digital Library  <br>Biblioteka Cyfrowa UMCS   | 5110         | 5080         | 5095            | 1290 | 1290 |
| 26. | Radom Digital Library  <br>Radomska Biblioteka Cyfrowa  | 774          | 750          | 762             | 15   | 15   |
| 27. | Digital Library of Provincial Public Library<br>in Lublin   Biblioteka Cyfrowa Wojewódz-<br>kiej Biblioteki Publicznej w Lublinie | 768          | 877          | 823             | 0    | 0    |
| 28. | Podkarpacka Digital Library  <br>Podkarpacka Biblioteka Cyfrowa   | 688          | 817          | 753             | 9    | 9    |
| 29. | Lodz Regional Digital Library  <br>Łódzka Regionalna Biblioteka Cyfrowa   | 5120         | 4940         | 5030            | 546  | 546  |
| 30. | Polonia Digital Library  <br>Polonijna Biblioteka Cyfrowa   | 4140         | 3980         | 4060            | 8    | 8    |
| 31. | Opole Digital Library  <br>Opolska Biblioteka Cyfrowa   | 1610         | 1690         | 1650            | 47   | 47   |
| 32. | Bialska Digital Library  <br>Bialska Biblioteka Cyfrowa   | 792          | 834          | 813             | 12   | 12   |
| 33. | Digital Library in Leszno  <br>Leszczyńska Biblioteka Cyfrowa   | 25           | 25           | 25              | 0    | 0    |
| 34. | Warmia-Mazury Digital Library  <br>Warmińsko-Mazurska Biblioteka Cyfrowa  | 2360         | 2300         | 2330            | 407  | 407  |

| No. | DL name   | Google PDF 1 | Google PDF 2 | Google PDF avg. | GS 1 | GS 2 |
|-----|---|--------------|--------------|-----------------|------|------|
| 35. | Digital Library of Lublin University of Technology   Biblioteka Cyfrowa Politechniki Lubelskiej   | 10500        | 10900        | 10700           | 752  | 752  |
| 36. | Wejherowo Digital Library   Wejherowska Biblioteka Cyfrowa  | 0            | 0            | 0               | 0    | 0    |
| 37. | Military Digital Library   Wojskowa Biblioteka Cyfrowa  | 6030         | 6010         | 6020            | 0    | 0    |
| 38. | Digital Library of Warsaw University of Technology   Biblioteka Cyfrowa Politechniki Warszawskiej | 10500        | 10500        | 10500           | 351  | 351  |
| 39. | Chelm Digital Library   Chełmska Biblioteka Cyfrowa   | 82           | 98           | 90              | 9    | 9    |
| 40. | Digital Library of Inowroclaw   Inowrocławska Biblioteka Cyfrowa                                  | 634          | 678          | 656             | 1    | 1    |

The column “Google PDF 1” contains the results of the first search; “Google PDF 2” – the search repeated after a week. The average values are presented in the third column. The disparities between the results in the first and second column reach thousands, both over and below. Nevertheless, it is clear which DLs are less visible in Google. It is certainly surprising that both Polona and CRISPA are practically invisible. The DLs of the technical universities are very well represented – despite the relatively small size of their collections, they have over 10 thousand documents filed in Google; the library of the Silesian University of Technology has the highest result of all DLs.

The results of a search conducted in Google Scholar are much more precise. The disparities between the results presented in columns “GS 1” and “GS 2” are minimal. The visibility of DL resources in GS is more varied. More DLs are less visible. It should not come as a surprise as GS prioritizes repositories over DLs. Nevertheless, some DL resources do appear in GS. The libraries with the highest number of files in GS were the Greater Poland Digital Library, the Lower Silesian Digital Library, the Digital Library of the Silesian University of Technology and the UMCS Digital Library.

The data in Table 2 concerning the visibility of the DL resources in Google shows that they are somewhat visible. But this is not enough to consider Google as a tool for browsing the DL resources, particularly where full-text search is concerned. A point of reference is required. According to the author’s calculations, the number of PDF files in all DLs studied is not lower than a million.<sup>5</sup> For reference, the

<sup>5</sup> As in the case of the DjVu files, the calculations were based on the advanced search conducted for 40 DLs considered in the article, using the format field.

average number of results of a Google search (presented in Table 2) comes up to c. 230 thousand total. This would suggest that Google indexes no more than 23% of all DL resources. These are obviously mere estimations, but they indicate the scale of the disparity between the number of available files and the results of the Google search.

To achieve a more precise assessment of Google's capacity to enable a full-text search of the DL resources, additional analysis was conducted. It involved a direct comparison of the results yielded by searching the phrase "pieskowa skała" in DLs (content search) and in Google (limiting the results to the DLs' domains). The analysis focused on the ten largest DLs, not counting Polona and CRISPA, whose resources are unindexed. The results of the comparison are presented in Table 3.

Tab. 3. The comparison of the results of full-text search conducted in DLs and in Google

| No. | DL name  | DL   | Google | % Google |
|-----|--|------|--------|----------|
| 1.  | Jagiellonian Digital Library   Jagiellońska Biblioteka Cyfrowa                           | 1074 | 22     | 2        |
| 2.  | Silesian Digital Library   Śląska Biblioteka Cyfrowa                                     | 905  | 64     | 7        |
| 3.  | Greater Poland Digital Library   Wielkopolska Biblioteka Cyfrowa                         | 373  | 12     | 3        |
| 4.  | Kujawsko-Pomorska Digital Library   Kujawsko-Pomorska Biblioteka Cyfrowa                 | 189  | 16     | 8        |
| 5.  | Małopolska Digital Library   Małopolska Biblioteka Cyfrowa                               | 291  | 27     | 9        |
| 6.  | NN Theatre Multimedia Library   Biblioteka Multimedialna Teatru NN                       | 13   | 5      | 38       |
| 7.  | Digital Library of University of Wrocław   Biblioteka Cyfrowa Uniwersytetu Wrocławskiego | 26   | 2      | 8        |
| 8.  | Regional Materials of Łódź Land   Regionalia Ziemi Łódzkiej                              | 143  | 5      | 3        |
| 9.  | Pomeranian Digital Library   Pomorska Biblioteka Cyfrowa                                 | 85   | 8      | 9        |
| 10. | Lower Silesian Digital Library   Dolnośląska Biblioteka Cyfrowa                          | 27   | 5      | 19       |
|     | Total  | 3126 | 166    | 5        |

There is a vast difference between the results yielded by DLs and by Google. It is partially explained by the fact that Google does not index the DjVu files (Lewandowski, 2014). Google retrieved only 5% of the documents found in the DL resources. To confirm these results, other queries were tested. The disparity is so

large that it rules out Google as a viable tool for accessing the DL resources. At least for the time being, the tools offered by particular DLs are the best method of accessing their resources.

## 6. Conclusions

More than anything, this analysis of the conditions and realities of full-text search in Polish DLs shows how much remains to be done in order to improve the users' access to the content of the files collected by the DLs. The most serious issue is the lack of a single point of access which would allow for conducting a full-text search of files scattered in various libraries using a unified index. Google cannot fill this gap – Google Custom Search might have served this purpose if Google were capable of indexing a larger part of the DL resources. The situation might be improved if the software used by DLs is optimized for the visibility of the files in Google. The disparities in visibility of resources collected by particular DLs are apparent when we compare Polona and CRISPA with the digital libraries of Polish technical universities. Equally crucial for indexing the DL files in Google will be replacing the DjVu files with PDF files, or uploading the PDF files alongside DjVus. The quality of OCR layer in the documents created digitally will be important, too, as it will allow for accurate indexing by Google and by the DL software. Many older files, particularly those in the DjVu format, should have their OCR layer corrected, or added, so that the users may access their content. We should not neglect that convenience is another important factor in browsing the DL resources. It depends on the efficacy of the system, and the operability of the software employed. We might observe significant changes for the better in this regard, but there is room for improvement.

## References

- Bednarek-Michalska, B. (2017). Polish Digital Libraries and Repositories. Origins, Operation and Usage. *Przegląd Biblioteczny*, 85 (spec. iss.), 46–69. <https://doi.org/10.36702/pb.854>
- Kolasa, W. M. (2007). dLibra Digital Library Framework – platforma do budowy bibliotek cyfrowych. In: J. Woźniak-Kasperek & J. Franke (eds.), *Biblioteki cyfrowe : projekty, realizacje, technologie* (67–88). Warszawa: Wydaw. SBP. <http://eprints.rclis.org/16562>
- Kotyńska, E. (2013). *Korekta OCR – problemy i rozwiązania* [online]. Biblioteka cyfrowa dziś a wyzwania jutra. Międzynarodowa konferencja naukowa. Biblioteka Jagiellońska, Kraków, 24–25 stycznia 2013 [19.01.2023], <https://jbc.bj.uj.edu.pl/Content/218090>
- Lewandowska, A., Mazurek, C., Werla, M. (2007). *Federacja Bibliotek Cyfrowych w sieci PIONIER – Dostęp do otwartych bibliotek cyfrowych i repozytoriów* [online]. IV Ogólnopolska Konferencja EBIB Internet w bibliotekach. Open Access. Toruń, 7–8 grudnia

- 2007 [19.01.2023], [http://www.ebib.pl/publikacje/matkonf/mat18/lewandowska\\_mazurek\\_werla.php](http://www.ebib.pl/publikacje/matkonf/mat18/lewandowska_mazurek_werla.php)
- Lewandowski, T. (2014). *Google Scholar a repozytoria i biblioteki cyfrowe w Polsce* [online]. *Otwarta nauka.pl* [19.01.2023], <https://otwartanauka.pl/analysis/case-studies/google-scholar-a-repozytoria-i-biblioteki-cyfrowe-w-polsce>
- Morawiec, B., M. (2021). *Biblioteki cyfrowe w Polsce* [online]. *Lustro Biblioteki* [20.12.2022], <https://lustrbiblioteki.pl/biblioteki-cyfrowe-polsce>
- Parkoła, T., Bohdanowicz, K., Werla, M. (2016). Realizacja potrzeb użytkowników bibliotek cyfrowych na przykładzie systemu dLibra 6. *Biuletyn EBIB* [online], 8 (170), <http://ebibojs.pl/index.php/ebib/article/view/168>
- Pulikowski, A. (2009). *Wyszukiwanie pełnotekstowe w zasobach bibliotek cyfrowych* [online]. X Forum Informacji Naukowej i Technicznej. Zakopane, 22–25 września 2009 [19.01.2023], <https://sbc.org.pl/dlibra/publication/15686/edition/13873/content>
- Szala, M. (2015). Cyfrowe oblicza zbiorów. In: G. Piotrowicz (ed.) *Wykorzystanie nowoczesnych technologii i mediów cyfrowych w Bibliotece Uniwersyteckiej we Wrocławiu. Stan na rok 2015* (259–276). Wrocław: Biblioteka Uniwersytecka we Wrocławiu. <https://www.bibliotekacyfrowa.pl/publication/75789>
- Szafański, L. (2013). *Dokumenty cyfrowe w JBC – próba charakterystyki* [online]. Biblioteka cyfrowa dziś a wyzwania jutra. Międzynarodowa konferencja naukowa. Biblioteka Jagiellońska, Kraków, 24–25 stycznia 2013 [19.01.2023], <https://www.jbc.bj.uj.edu.pl/dlibra/publication/229579/edition/218080>

---

## Wyszukiwanie pełnotekstowe w zasobach polskich bibliotek cyfrowych

### Abstrakt

**Cel/teza:** Celem artykułu jest analiza uwarunkowań i możliwości wyszukiwania pełnotekstowego w zasobach polskich bibliotek cyfrowych (BC), uwzględniająca dostępność wyszukiwania w treści dokumentów w poszczególnych BC, wykorzystywane formaty plików i oprogramowanie, a także widoczność zasobów BC w wyszukiwarce Google.

**Koncepcja/Metody badań:** Do badań wytypowano 40 największych polskich BC, których zasoby stanowią głównie digitalizowane zbiory tradycyjnych bibliotek. Sprawdzone rodzaj wykorzystywanego oprogramowania, dostępność i możliwości wyszukiwania pełnotekstowego oraz stopień indeksowania zasobów w formacie PDF w wyszukiwarce Google i Google Scholar. Na koniec porównano wyniki wyszukiwania pełnotekstowego w dziesięciu BC z uzyskanymi w Google.

**Wyniki/Wnioski:** Wszystkie wybrane do badań BC posiadają możliwość wyszukiwania pełnotekstowego. Wyodrębnione interfejsy wyszukiwawcze różnią się znacznie między sobą. Każdy ma zalety i wady wymagające dalszych prac rozwojowych. Wyszukiwarka Google nie nadaje się obecnie do wykorzystania jako alternatywny sposób dostępu do treści udostępnianych w polskich BC.

**Oryginalność/Wartość poznawcza:** Problematyka wyszukiwania pełnotekstowego w BC jest rzadko podejmowana, mimo iż użytkownicy uznają wyszukiwanie w treści dokumentów za jedną z najbardziej przydatnych funkcji BC. Wyniki przeprowadzonych badań tworzą obraz aktualnych możliwości BC w tym zakresie.

**Słowa kluczowe**

Biblioteki cyfrowe. Widoczność zasobów. Wyszukiwanie informacji. Wyszukiwanie pełnotekstowe. Wyszukiwarka Google.

---

*ARKADIUSZ PULIKOWSKI, Associate Professor at the Institute of Culture Studies at the Faculty of Humanities of University of Silesia in Katowice. Research interests: information search; information behavior; infometrics; digitization of information. Major recent publications: Modelowanie procesu wyszukiwania informacji naukowej. Strategie i interakcje (2018), Searching for LIS scholarly publications: a comparison of search results from Google, Google Scholar, EDS, and LISA (Journal of Academic Librarianship, 2021, co-author: A.Matysek), The Relation Between the Structure of Abstracts in LIS and Anthropology Journals and Their Rank (Zagadnienia Informatyki Naukowej, 2020).*

*Contact to the Author*

*arkadiusz.pulikowski@us.edu.pl*

*Institute of Culture Studies,*

*Faculty of Humanities,*

*University of Silesia in Katowice*

*ul. Bankowa 11*

*40-007 Katowice, Poland*

# Jak być świadomym użytkownikiem algorytmów? O potrzebie rozwijania kompetencji algorytmicznych

Łukasz Iwasiński

ORCID 0000-0003-2126-7735

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

Weronika Furman

---

## Abstrakt

**Cel/Teza:** Wszegobecność algorytmów i ich oddziaływanie na niemal każdy aspekt życia jednostek oraz społeczeństwa stanowi istotne wyzwanie dla współczesnego świata. Odpowiedzią na nie powinno być kształtowanie szeroko pojętych kompetencji algorytmicznych. Podstawowym celem artykułu jest analiza tego pojęcia. Drugim celem jest wskazanie, dlaczego owe kompetencje są ważne dla efektywnego i świadomego funkcjonowania we współczesnym świecie.

**Koncepcja/Metody:** W artykule posłużono się metodą narracyjnego przeglądu literatury i elementami krytycznej analizy dyskursu. Analiza opiera się na literaturze teoretycznej i empirycznej. Podjęte w artykule rozważania zilustrowano przykładami zaczerpniętymi ze źródeł zastanych oraz z własnych obserwacji.

**Wyniki i wnioski:** Kompetencje algorytmiczne należy ujmować wielowymiarowo, pamiętając o aspekcie praktycznym i teoretycznym. Niezbędna wydaje się powszechna edukacja obejmująca przynajmniej elementy kompetencji algorytmicznych. Poza wiedzą techniczną wpływ na kompetencje algorytmiczne ma także wiedza społeczno-kulturowa. Algorytmy często wzmacniają władzę hegemonicznych struktur społecznych i reprodukuja istniejące w kulturze uprzedzenia. Dlatego dla jakości kompetencji algorytmicznych, zwłaszcza ich etycznego wymiaru, znaczenie może mieć znajomość struktur i sił społecznych, które za tymi algorytmami stoją. W ujęciu krytycznym na kompetencje algorytmiczne można spojrzeć jak na budowanie świadomości istnienia władzy algorytmicznej i zdolność do stawiania jej oporu.

**Oryginalność/Wartość poznawcza:** Refleksja nad kompetencjami algorytmicznymi, a tym bardziej próby empirycznej analizy tej kategorii, są w początkowym stadium. Artykuł krytycznie omawia aktualny stan badań w tym – właściwie nierozpoznanym w polskim piśmiennictwie – obszarze.

## Słowa kluczowe

Kompetencje algorytmiczne. Przejrzystość algorytmów. Świadomość algorytmiczna. Władza algorytmów.

*Otrzymano: 13 października 2022. Zrecenzowano: 3 grudnia 2022.*

*Poprawiono: 12 grudnia 2022. Zaakceptowano: 12 stycznia 2023.*

---

## 1. Wprowadzenie

Zagadnienie kompetencji informacyjnych (ang. *information literacy*) oraz kompetencji cyfrowych (ang. *digital literacy*) jest od lat szeroko komentowane w literaturze z obszaru informatologii, nauk społecznych i medioznawstwa (Dobson & Willinsky, 2009; Doyle, 1994; Koltay, 2011; Reddy et al. 2020; Snavely & Cooper, 1997). Obok tych pojęć w piśmiennictwie spotkać można wiele, mniej lub bardziej zbliżonych znaczeniowo kategorii, jak m.in.<sup>1</sup>: *information technology literacy* (Ezziane, 2007), *media literacy* (Livingstone, 2004; Potter, 2013), *social media literacy* (Festl, 2021; Livingstone, 2014), *new media literacy* (Chen et al., 2011; Luke 2007), *computer literacy* (Hoffman & Blake, 2003), *internet literacy* (Kim & Yang, 2016; Livingstone, 2008), *Web literacy* (Kuiper et al., 2008), *network literacy* (McClure, 2018), *cyber-literacy* (Stiller & LeBlanc, 2006), *software literacy* (Khoo et al., 2017), *data literacy* (Koltay, 2017), *big data literacy* (D’Ignazio & Bhargava, 2015), *coding literacy* (Hutchison et al., 2016), *platform literacy* (Žagar et al., 2021), *data infrastructure literacy* (Gray et al., 2018), *privacy literacy* (odnoszące się do prywatności w internecie) (Bartsch & Dienlin, 2016), w końcu – *algorithmic literacy*. Namnażanie tych terminów stanowi odpowiedź na dynamicznie zmieniającą się rzeczywistość – zarówno w obszarze technologii, jak i wiedzy oraz świadomości jej użytkowników.

Nie ma ogólnie akceptowanych definicji przytoczonych wyżej pojęć. Są one ujmowane na różne sposoby przez różnych badaczy i na gruncie różnych dyscyplin. Tym samym nie sposób jednoznacznie rozstrzygnąć, jaka jest między rzeczonymi pojęciami relacja. Niewątpliwie jednak zakresy znaczeniowe wielu z nich w jakiejś części się pokrywają i być może niektóre – zgodnie z zasadą brzytwy Ockhama – należałoby uznać za redundantne. Sądzymy jednak, że kategoria *algorithmic literacy* jest warta zachowania. Stanowi ona dobry klucz do opisu ważnego dla efektywnego i świadomego funkcjonowania we współczesnym świecie zestawu kompetencji. Celem niniejszego artykułu jest analiza tego pojęcia, a także wskazanie, dlaczego kompetencje algorytmiczne są ważne dla efektywnego i świadomego funkcjonowania we współczesnym świecie. Zastosowaliśmy metodę narracyjnego przeglądu literatury z elementami krytycznej analizy dyskursu.

## 2. Pojęcie *literacy*

We wskazanych wyżej kolokacjach pojęcie wyrażane angielskim terminem *literacy* zwykle w języku polskim określane jest jako kompetencje. Dosłownie termin ten oznacza jednak *alfabetyzm* – i w niektórych kontekstach pojawia się również taki właśnie przykład. Pierwotnie angielski termin *literacy*, podobnie jak polski

<sup>1</sup> Podajemy je, za przytoczonymi źródłami, w wersji angielskiej.

*alfabetyzm*, oznaczał zdolność czytania i pisania. Początkowo te słowa, zarówno w języku angielskim, jak i polskim, odnosiły się do czysto technicznego aspektu, czyli znajomości liter i umiejętność ich używania do zapisania/odczytania tekstu. Nabywanie tych umiejętności miało odbywać się w toku sformalizowanego procesu edukacji szkolnej. Alfabetyzm postrzegany był – w wymiarze jednostkowym – jako niezbędny i samoistny warunek produktywnego uczestnictwa w życiu społecznym i – w wymiarze społecznym – wehikuł rozwoju. W drugiej połowie XX w. zwrócono uwagę na fakt, że profity z posługiwania się pismem i rozumienia tekstu pisanego są zawsze uzależnione od wielu czynników o charakterze osobistym, strukturalnym czy kulturowym (Graff, 2010). Z czasem sens tego pojęcia ulegał poszerzeniu, obejmując zróżnicowaną wiedzę i umiejętności determinujące efektywne funkcjonowanie w społeczeństwie, bądź – po odpowiednim dookreśleniu – wiedzę i umiejętności w jakiejś konkretnej dziedzinie. Tak oto mówi się np. o *health literacy* (na język polski tłumaczonym zarówno jako *alfabetyzm zdrowotny*, jak i *kompetencje zdrowotne*) i wszystkich wymienionych wyżej rodzajach *literacies* związanych z kontekstem medialnym i informacyjnym (tłumaczonych zwykle jako kompetencje)<sup>2</sup>. Czasem podejmuje się próby operacjonalizacji i proponuje się zamknięty katalog kompetencji w ramach określonej dziedziny, tak by stopień danej *literacy* uczynić mierzalnym. Osiągnięcie wskazanego poziomu ma gwarantować funkcjonalny alfabetyzm w tej dziedzinie.

Jakkolwiek jak najbardziej poprawnym przekładem terminu *algorithmic literacy* jest *algorytmiczny alfabetyzm*, jak zaznaczyliśmy, w odniesieniu do problematyki medialnej i informacyjnej termin *literacy* w przeważającej liczbie źródeł tłumaczony jest jako kompetencje. Przyjmujemy więc taką konwencję i w dalszej części tekstu używać będziemy terminu *kompetencje algorytmiczne*.

Istnieje także pojęcie *critical literacy*. Nie akcentuje się w nim, w każdym razie nie w pierwszym rzędzie, wiedzy i umiejętności pozwalających na adaptację do systemu. Kładzie się natomiast nacisk na rozwój kompetencji pozwalających na modyfikowanie zastanej rzeczywistości, przekształcanie istniejących relacji, w szczególności – stawianie oporu siłom zniewalającym, uprzedmiotawiającym. *Critical literacy* można widzieć jako narzędzie krytycznej interpretacji różnorodnych tekstów kultury i, zgodnie ze stanowiskiem pragmatycznym, używania ich do realizacji szeroko rozumianych celów emancypacyjnych, przeciwstawiania się opresji czy manipulacji (Coffey, 2008). Dla praktykowania tak pojętej *critical literacy* nieodzowna jest umiejętność rozpoznawania struktur władzy i mechanizmów wyzysku. Nie dąży się tu do wyliczania konkretnych kompetencji, gdyż uznaje się, że krytyczna wiedza i umiejętności są kontekstualne, zrelatywizowane

<sup>2</sup> W piśmiennictwie znaleźć można tak rozmaite rodzaje *literacies*, jak choćby: *health literacy*, *death literacy*, *financial literacy*, *physical literacy*, *ocean literacy*, *religious literacy*, *visual literacy*, *dancing literacy*, *spatial literacy*, *screen literacy* czy *porn literacy* (Ridley & Pawlick-Potts, 2021).

do środowiska i sytuacji społecznej. Z takiego ujęcia czerpie koncepcja *critical algorithmic literacy* (Cotter, 2020).

### 3. Pojęcie algorytmu

Termin *algorytm* pochodzi od nazwiska perskiego matematyka żyjącego na przełomie VIII i IX w., Abu Abdullaha Muhammada ibn Musy al-Chuwarizmiego. Zniekształcona, zlatynizowana wersja jego nazwiska brzmiała Algoritmi. W średniowieczu mianem *algorithmus* określano wszelkie operacje wykonywane na liczbach naturalnych (Zawojski, 2018, 18). Współcześnie pojęcie algorytmu kojarzone jest przede wszystkim z informatyką, w której oznacza skończoną sekwencję instrukcji, zgodnie z którymi przetwarzane są dane. Jest to powszechne, choć potoczne, a nie ściśle naukowe, rozumienie algorytmu. Utrwaliło się ono na gruncie informatyki w latach 60. XX w. (Miyazaki, 2012). Rygorystyczna definicja algorytmu nastrocza wielu problemów i nie brak głosów, że nie da się takiej wypracować (Gurevich, 2012). Często jednakże wskazuje się – za Donaldem Knuthem (Knuth, 1973, 4–6) – na pięć kluczowych cech algorytmów: skończoność (po skończonej liczbie kroków algorytm się zatrzyma), dobre zdefiniowanie (każdy krok algorytmu musi być jednoznacznie zdefiniowany), dane wejściowe, dane wyjściowe, efektywność (każdy krok możliwy jest do zrealizowania w praktyce, w skończonym czasie). Algorytmy nie muszą być deterministyczne – innymi słowy wynik działania danego algorytmu nie musi być jednoznacznie określony przez dane wejściowe. Istnieją algorytmy probabilistyczne, korzystające z losowości. Większość algorytmów realizujących procesy uczenia maszynowego ma taki właśnie charakter.

Dla nauk społecznych i humanistycznych kluczowy jest nie tyle techniczny wymiar algorytmów, ile ich społeczna konstrukcja, zawarte w nich znaczenia i ich performatywny charakter (Bucher, 2018, 29; Kreft, 2019, 29–31; Szpunar, 2019, 23). W perspektywie społeczno-kulturowej algorytm jest zatem czymś więcej niż zbiorem następujących po sobie operacji. Można na niego spojrzeć jako na społeczno-technologiczny system tworzenia wiedzy, dla którego nie mniej istotna od zawierającego listę instrukcji kodu jest kultura, obowiązujące w niej normy i wartości. Tak rozumiany algorytm jest instytucją o potencjalnie ogromnym wpływie na wszelkie wymiary współczesnego życia.

Pojęcie algorytmu w dyskursie opisującym współczesny świat odmieniane jest przez wszystkie przypadki – w literaturze mnożą się takie kategorie, jak: zwrot algorytmiczny (Uricchio, 2011), ideologia algorytmiczna (Mager, 2012), algorytmiczne życie (Amoore & Poitukh, 2015), tożsamość algorytmiczna (Cheney-Lippold, 2011), algorytmiczne zarządzanie (Issar & Aneesh, 2022). Badacze społeczni piszą, że żyjemy w wieku algorytmów (Mai, 2019), kulturze algorytmów (Szpunar, 2019) czy algokracji (Danaher, 2016).

## 4. Władza algorytmów

Postępująca digitalizacja kolejnych obszarów rzeczywistości warunkuje ich algorytmizację, ponieważ cyfrowe dane łatwo poddają się algorytmicznemu przetwarzaniu. Poprzez znajdowanie relacji między zmiennymi, klasyfikowanie czy dokonywanie predykcji algorytmy regulują przepływy treści w sieci oraz stają się ważnym, jeśli nie podstawowym, mechanizmem zarządzania informacją i wiedzą w środowisku społecznym w ogóle. Dopasowują do nas różnorodne przekazy medialne, wspierają podejmowanie decyzji przez nas samych i decyzji nas dotyczących, regulują stosunki społeczne, wpływają na nadawanie znaczeń i kategoryzowanie różnych elementów świata – w tym ludzi, co David Lyon (2003) nazywa „sortowaniem społecznym”.

Wielu autorów popularnonaukowych publikacji, chcąc sugestywnie uzmysłowić czytelnikowi wszechobecność algorytmów i ich wpływ na życie jednostki, opisuje rutynowe czynności i decyzje, jakie każdy z nas podejmuje, pokazując, jak kształtowane są one przez algorytmy. Tak oto informatyk i ekspert uczenia maszynowego Pedro Domingos (2016, 11–13) w swej opowieści rekonstruuje hipotetyczny dzień z życia współczesnego człowieka. Budzi się on rano, słysząc piosenkę wybraną przez algorytmy (a może i stworzoną z ich użyciem), mieszka w nadzorowanym przez algorytmy inteligentnym domu, w mieście, w którym dzięki analizom big data dokonywanym przez algorytmy spadła przestępczość. Porusza się kontrolowanym przez algorytmy samochodem, wyznaczoną przez algorytmy drogą. Jest wspomagany przez algorytmy w pracy, w czasie wolnym, w codziennych zakupach, w zarządzaniu finansami, w planowaniu właściwie wszelkich aspektów życia. Odbiera list, który trafił pod właściwy adres dzięki opartej na algorytmach uczenia maszynowego technologii inteligentnego rozpoznawanie znaków (ang. *Intelligent Character Recognition* – ICR) (Ptucha et al., 2019). Przyjmuje leki, które zaprojektowane zostały przy użyciu uczących się algorytmów, jego lekarz stawia diagnozę, korzystając z interpretacji wyników badań dokonanej przez algorytmy itd.

Filozofka i specjalistka od sztucznej inteligencji Aleksandra Przegalińska wybiega w przyszłość, kreśląc wizję jednego dnia z życia w świecie wszechobecnych algorytmów sztucznej inteligencji (Przegalińska & Oksanowicz, 2020, 11–34). Opisywana przez Przegalińską postać budzi się, mając na czole opaskę z wbudowanymi sensorami do detekcji fazy snu – algorytmy decydują, kiedy przebudzenie będzie najłagodniejsze. Jej inteligentny dom stara się rozpoznać potrzeby właścicielki i adaptuje się do nich, regulując najróżniejsze parametry, od temperatury powietrza po ilość światła, czy podpowiadając ubranie adekwatne do pogody i okazji – w ich doborze pomagają algorytmy przetwarzania obrazu. Następnie jedzie do pracy autonomicznym samochodem, który wykorzystuje wiele typów algorytmów, w tym algorytmy widzenia maszynowego. Technologie ubieralne (Iwasiński, 2017) nieustannie monitorują jej organizm i w przypadku wykrycia nieprawidłowości proponują działania zaradcze, np. kontakt z lekarzem. W pracy ma za zadanie

przeprowadzić wywiad. Jeśli ma zamkniętą listę pytań i rozmowa ma względnie zestandaryzowaną formę, może wyręczyć ją bot, bazujący na algorytmach do prowadzenia rozmowy, a potem algorytmy do tworzenia treści mogą skonstruować na podstawie tej konwersacji artykuł. Po powrocie do domu inteligentny asystent, analizując liczne dane, m.in. fizjologiczne (zbierane za pomocą urządzeń ubieralnych) oraz behawioralne, na przykład liczbę wypowiedzianych w danym dniu słów i ich tzw. wydźwięk (ang. *sentiment analysis*) (Medhat et al., 2014), bada nastroj bohaterki i jej potrzebę interakcji. Po analizie wyników algorytm podejmuje decyzję, czy asystent ma podjąć rozmowę i jak ma ona przebiegać. Ewentualna konwersacja – jej treść i forma – staje się źródłem nowych danych, zasilających różnorodne algorytmy. Wieczorem algorytm poleci nie tylko film z platformy streamingowej, ale także hybrydową grę wirtualną, w której bohaterka snutej tu historii bierze udział, i której scenariusz tworzony jest w czasie rzeczywistym przez algorytmy analizujące wypowiedziane przez nią słowa, jej zachowanie, stan organizmu (mimikę, parametry fizjologiczne) oraz cały pakiet wcześniej zebranych danych. W takiej rzeczywistości niemal każda interakcja ze światem, a nawet interakcja z samym sobą – dzięki tzw. self-trackingowi (Iwasiński, 2017) – jest zapośredniczona przez algorytmy.

Poza wpływem na życie jednostek algorytmy kształtują i modyfikują system społeczny, gospodarczy i polityczny. Początkowo algorytmy jedynie automatyzowały różne procesy. Wraz z rozwojem systemów opartych na sztucznej inteligencji, w szczególności uczenia maszynowego, nadzorowane przez algorytmy procesy coraz bardziej się autonomizują, a więc ich zależność od decyzji człowieka i jego ingerencja zmniejsza się. Dotychczas to człowiek wyznacza nadrzędne cele, a algorytm je realizuje, ewentualnie samodzielnie określając cele cząstkowe. W przyszłości algorytmy być może same będą wskazywać i realizować także cele strategiczne (Wodecki, 2021, 47). Autonomia dotyczy nie tylko obszaru i zakresu działania algorytmu, ale również metody. W przypadku meta learning<sup>3</sup> algorytmy same uczą się najlepszych dla danego problemu i sytuacji metod uczenia się. Można przyjąć, że algorytmy zyskują sprawczość; tym samym pojawia się ryzyko tego, że wymkną się ludzkiej kontroli i ulegną reifikacji (Szpunar, 2019, 11–35).

Wspomniany wyżej Domingos pisze:

żaden człowiek nie ma większego wpływu niż algorytm Amazonu na to, jakie książki czyta się obecnie na świecie. Algorytmy NSA uznają, kto zostanie uznany za potencjalnego terrorystę. Modele klimatu decydują, jakie są bezpieczne poziomy dwutlenku węgla w atmosferze. Modele wybierające akcje na giełdzie oddziałują na gospodarkę w większym stopniu, niż większość z nas (Domingos, 2016, 15).

Władzę algorytmów legitymizuje typowa dla nowoczesności – narastająca od Oświecenia – wiara w bezosobowe, abstrakcyjne, zmatematyzowane systemy

<sup>3</sup> Meta learning to poddziedzina uczenia maszynowego, w ramach której algorytmy, korzystając z metaanaliz dotyczących uczenia maszynowego, dobierają optymalne metody uczenia dla określonych problemów (Vanschoren, 2018).

analityczne; do przyznawania takim właśnie systemom legitymacji do ostatecznego rozstrzygnięcia, co jest prawdą i jaka jest rzeczywistość (Porter, 1995). Tarleton Gillespie (2014, 179) twierdzi, że algorytmy postrzegane są jako narzędzia wolne od subiektywnych ocen, sprawiedliwe, rzetelne, bezbłędne i jako takie stały się stabilizatorami społecznego zaufania. Bezstronność algorytmów jest jednak w oczywisty sposób pozorna. W każdym algorytmie zakodowane są określone wartości i/bądź interesy, nie istnieje algorytm obiektywny, aksjologicznie neutralny. Stronniczość algorytmicznych rozstrzygnięć stała się w ostatnich latach przedmiotem rozlicznych analiz, prowadzi się coraz więcej badań dotyczących tej problematyki (Iwasiński, 2020). Algorytmy bazujące na błędnych, zniekształconych danych i nieprawidłowych sposobach ich przetwarzania mogą deformować różnorodne systemy i instytucje. Nawet algorytmy oparte na prawidłowych danych – to znaczy takich, w przypadku których zbiór treningowy jest reprezentatywny dla populacji, na której algorytm uczenia maszynowego będzie stosowany – mogą powodować utrwalanie istniejących uprzedzeń czy nieprawd.

A zatem algorytmy ingerują w niemal każdą dziedzinę życia indywidualnego, a także społecznego. Kształtują nasze zachowanie i osobiste wybory (w obszarze polityki, tworzenia relacji, rozrywki, finansów, zdrowia, edukacji, turystyki i in. – sukcesywnie obejmując kolejne obszary), a zarazem wpływają na funkcjonowanie najróżniejszych organizacji, stając się medium społecznej reprodukcji (niekiedy odtwarzają oraz wzmacniają istniejące struktury i porządek, a czasem je przekształcają)<sup>4</sup>.

## 5. Kompetencje algorytmiczne

W tym kontekście zrozumiały wydaje się postulat edukacji algorytmicznej i kształtowania kompetencji algorytmicznych – obejmujących wiedzę na temat funkcjonowania algorytmów, ich roli i konsekwencji ich działania, a także uprawnień, jakie przysługują osobom i grupom będącym przedmiotem realizowanych czy wspomaganych przez algorytmy operacji. Taka edukacja, zorientowana zarówno teoretycznie, jak i praktycznie, powinna przyczyniać się do efektywnego i świadomego korzystania z bazujących na tej technologii narzędzi oraz dawać podstawy do ich krytycznej oceny.

Pojęcie kompetencji algorytmicznych w naukowym dyskursie pojawiło się w połowie pierwszej dekady XXI w.<sup>5</sup> Funkcjonowało wtedy jako określenie stricte

---

<sup>4</sup> Swoistym społecznym efektem ubocznym funkcjonowania algorytmów jest fakt, że ujawniają istniejące w kulturze uprzedzenia czy różne ideologiczne nachylenia – multiplikują je, przez co stają się one bardziej widoczne i ewidentne.

<sup>5</sup> Czas określamy orientacyjnie na podstawie kwerendy bazy Google Scholar. Seel i Casey (2003, 39) wskazują, że pojęcie *algorithmic literacy* pojawiło się w książce Johna Culberstona z 1986 r. Nie jest to jednak wskazanie precyzyjne. W istocie Culberston (1986, 128) wymienia rodzaje *computer literacy*, wśród których znajduje się *algorithmic reasoning*.

technicznych kompetencji związanych z analizą danych, niezbędnych do prowadzenia zaawansowanych badań naukowych i było traktowane jako element szerszego pojęcia *inferential literacy*, obejmującego znajomość ogólnej metodologii badań eksperymentalnych, statystyki i komputerowych technik obliczeniowych (Miron & Nadon, 2006; Walker & Hughes, 2008). W kolejnych latach utrwaliło się jako określenie umiejętności tworzenia algorytmów obliczeniowych. Z tak rozumianymi kompetencjami algorytmicznymi powiązane są pojęcia umiejętności algorytmicznych (ang. *algorithmic skills*) i myślenia algorytmicznego (ang. *algorithmic thinking*). Oba te terminy obecne są w literaturze od lat 70 XX w. Dziś występują przede wszystkim w kontekście edukacji, a także badań i oznaczają wiedzę pozwalającą projektować, tworzyć, ewentualnie implementować algorytmy oraz określony typ myślenia, który sprzyja efektywnemu wykonywaniu tych czynności.

A zatem, początkowo kompetencje algorytmiczne odnoszono wyłącznie do osób tworzących algorytmy. Niespełna dekadę temu zaczęło upowszechniać się odmienne rozumienie kompetencji algorytmicznych, skupiające się na świadomości mechanizmów działania algorytmów wśród ludzi niepracujących w branży technicznej, a będących użytkownikami algorytmów i/lub przedmiotem ich oddziaływania. Tak pojęte kompetencje algorytmiczne proponowano traktować jako składową *information literacy* (Bakke, 2020), *data literacy*, bądź *digital literacy* (Kampa & Balzer, 2021). W literaturze funkcjonuje również pojęcie świadomości algorytmicznej (ang. *algorithmic awareness*) występujące w podobnym znaczeniu, co kompetencje algorytmiczne.

Dla określenia sposobu, w jaki użytkownicy wyobrażają sobie, czym są algorytmy, czym być powinny i jak powinny działać, zaproponowano termin *imaginarium algorytmiczne* (ang. *algorithmic imaginary*) (Bucher, 2017). Zaczęto analizować, jak postrzeganie algorytmów i emocje, jakie wywołują, przekładają się na zadowolenie z użytkowania serwisów i podejmowane z nimi interakcje. Jedną z zaproponowanych koncepcji mówi, że, rekonstruując działanie złożonych technologii, ludzie posługują się teoriami potocznymi (ang. *folk theories*). Są to odwołujące się do potocznej wiedzy narracje, które pomagają zrozumieć działanie nierozpoznanych, skomplikowanych mechanizmów. Zidentyfikowano stosowane przez użytkowników Twittera i Facebooka intuicyjne, nieformalne próby wyjaśniania ich pracy, wyników i skutków, wpisujące się w koncepcję teorii potocznych (DeVito et al., 2017; Eslami et al., 2016). Teorie te wpływają na reakcje i zachowanie wobec systemów bazujących na algorytmach. W przypadku różnego rodzaju serwisów internetowych wyobrażenia na temat sposobu profilowania czy też rekomendacji kształtują oczekiwania użytkowników w stosunku do otrzymywanych treści. W ostatnich latach podjęto próby wypracowania bardziej precyzyjnych konceptualizacji i narzędzi badania kompetencji algorytmicznych.

Propozycja Leyli Dogruel (2021) dotyczy algorytmów funkcjonujących w Internecie. Badaczka ujmuje kompetencje algorytmiczne w dwóch wymiarach – poznawczym

oraz behawioralnym. Wymiar poznawczy składa się z dwóch komponentów: świadomości i wiedzy (ang. *awareness and knowledge*) oraz zdolności do krytycznej oceny (ang. *critical evaluation*). Świadomość dotyczy możliwości zidentyfikowania algorytmów w konkretnym serwisie i określenia zakresu oraz sposobu ich działania (sortowanie, filtrowanie, profilowanie). Wiedza to ogólne rozumienie zasad działania internetowych algorytmów i regulacji prawnych ich dotyczących. Krytyczna ocena skupia się na potencjalnych szansach i zagrożeniach, które mogą towarzyszyć użytkownikowi platform wykorzystujących algorytmy, związanych m.in. z: manipulacją, dyskryminacją, istnieniem baniek informacyjnych, naruszeniami prywatności. Wymiar behawioralny obejmuje również dwa komponenty: sposoby radzenia sobie z algorytmami (ang. *coping behaviors*) oraz zdolności projektowania i tworzenia algorytmów (ang. *creation and design*). Pierwszy dotyczy kompetentnego korzystania z algorytmów, związany jest m.in. z: umiejętnością modyfikowania domyślnych ustawień aplikacji i platform, wpływaniem na wyniki otrzymywane w wyszukiwarkach i platformach, na selekcję i ekspozycję treści publikowanych na platformach, czy też porównywaniem wyników różnych algorytmów. Dogruel wyróżnia trzy główne rodzaje strategii przyjmowanych przez użytkowników: środki służące ochronie prywatności (m.in. stosowanie odpowiednich ustawień, usuwanie ciasteczek), działania wpływające na treści (m.in. porównywanie wyników różnych wyszukiwarek, personalizowanie ustawień w wyszukiwarkach i platformach, własne formy interakcji z algorytmem) oraz aktywizm (różne formy sprzeciwu wobec działań algorytmów, czy ich modyfikacji – od angażowania się w dyskusje, petycje, po aktywności hakerskie).

Według badań Daniela Kluga i współpracowników (2021) użytkownicy TikToka zakładają, że istotny wpływ na ekspozycję publikowanych przez nich treści (materiałów video) ma zaangażowanie użytkowników (liczba wyświetleń, komentarzy, polubień), godziny zamieszczenia materiału i dodanie odpowiednich, popularnych w danym momencie hashtagów. Wyobrażenia te często determinują strategię publikowania treści. Analiza badaczy wykazała, że założenie dotyczące skuteczności dodawania hashtagów jako sposobu na zwiększenie popularności filmu są błędne. Brak transparentności algorytmu zmusza jednak użytkowników do eksperymentowania i poszukiwania sposobów pomagających w promowaniu zamieszczanych przez nich filmów. Wśród mniej popularnych prób wpłynięcia na wyniki algorytmu znalazło się wykorzystywanie utworów muzycznych zdobywających popularność, skracanie długości filmu oraz bezpośredni zwrot twórcy do publiczności. Praktykowana przez użytkowników gra, stanowiąca działanie wpływające na zarządzane przez algorytm treści, wpisuje się w behawioralny wymiar kompetencji algorytmicznych.

Na przełomie lipca i sierpnia 2022 r. użytkownicy Instagrama wyrazili swój sprzeciw wobec zmiany dotychczas działającego algorytmu na nowy, który faworyzuje filmy video zwane rolkami względem publikowanych zdjęć. Inicjatywa opierała się na udostępnianiu grafiki o treści „MAKE INSTAGRAM INSTAGRAM AGAIN”, która szybko stała się wiralem. Użytkownicy platformy zarzucali twórcom aplikacji chęć

upodobnienia jej do platformy TikTok oraz niekorzystny ich zdaniem sposób filtrowania i prezentowania treści (posty wyświetlane niechronologicznie, obecność reklam). Jest to przykład aktywizmu również wpisujący się w behawioralny wymiar kompetencji algorytmicznych. W obu przedstawionych przypadkach działania użytkowników wymagały posiadania odpowiedniego poziomu świadomości algorytmicznej.

Drugi komponent wymiaru behawioralnego kompetencji algorytmicznych dotyczy projektowania i tworzenia algorytmów. Wiąże się z przynajmniej podstawową znajomością programowania, pisania i rozumienia kodu. Może obejmować również modyfikację istniejącego już algorytmu, różne formy dostosowywania aplikacji i platform do własnych potrzeb, także wbrew intencjom ich twórców. Tworzenie i projektowanie algorytmów jest związane z kompetencjami określonymi wcześniej, jako umiejętności algorytmiczne i myślenie algorytmiczne – ich rola w społeczeństwie informacyjnym będzie zapewne rosła.

W 2021 r. Brahim Zarouali, Sophie C. Boerman i Claes H. de Vreese skonstruowali i poddali walidacji skalę AMCA (ang. *Algorithmic Media Content Awareness*) mierzącą świadomość algorytmiczną użytkowników Internetu (Zarouali et al., 2021). Narzędzie do jej badania oparte jest na kwestionariuszu składającym się z trzynastu pytań. Autorzy zaproponowali cztery wymiary świadomości algorytmicznej. Są to:

- filtrowanie treści – bada świadomość tego, że algorytmy dopasowują treści do użytkowników (4 pytania);
- automatyczne podejmowanie decyzji – bada świadomość tego, że algorytmy automatycznie decydują w zakresie dostarczania użytkownikowi treści (3 pytania);
- interakcja z algorytmem – bada świadomość tego, że działanie użytkownika wpływa na dostarczane treści (3 pytania);
- świadomość konsekwencji etycznych – bada świadomość etycznych konsekwencji działania algorytmów (3 pytania).

Badani za pomocą skali Likerta oceniają poziom świadomości (1 – zupełnie nieświadomy, 5 – w pełni świadomy). Autorzy zaznaczają jednak, że wyodrębnione w wyniku analizy literatury wymiary świadomości algorytmicznej mogą być niekompletne i prawdopodobnie nie obejmują wszystkich istotnych obszarów. Odpowiedzi respondentów oparte są na deklarowanym poziomie świadomości, w związku z czym nie sposób je zobiektywizować – dlatego stosowanie takiego narzędzia do badań porównawczych jest dyskusyjne.

W 2022 r. Dogruel wraz ze współpracownikami zaproponowali dwuwymiarową skalę pomiaru kompetencji algorytmicznych i stworzyli oraz zwalidowali narzędzie badawcze w postaci kwestionariusza zawierającego 22 pytania (Dogruel et al., 2022). Wspomniane wymiary to:

- świadomość na temat wykorzystywanych algorytmów (11 pytań);
- wiedza o algorytmach (11 pytań).

Pierwszy dotyczy świadomości na temat celów, w jakich można używać algorytmów oraz świadomości obszarów (jak: wyszukiwarki, serwisy społecznościowe, spersonalizowane wiadomości, targetowane reklamy) i aplikacji lub urządzeń (jak: smartfony, przeglądarki internetowe, media społecznościowe, serwisy bankowości internetowej), w których algorytmy są rzeczywiście wykorzystywane. Wymiar ten jednak nie odnosi się do wiedzy na temat tego, jak algorytmy faktycznie działają. O tym informuje drugi z wspomnianych wymiarów. Koncentruje się on na bardziej zaawansowanym rozumieniu mechanizmów działania systemów algorytmicznych i ich konsekwencjach (jak personalizacja informacji i inne formy filtrowania treści). Badania potwierdziły korelację między wynikami w obu wymiarach a praktyczną umiejętnością interakcji z algorytmami.

## 6. Przejrzystość algorytmów

Szczególne miejsce w dyskusji nad problemem świadomego korzystania z algorytmów zajmuje problem ich przejrzystości. Ze względu na swą złożoność i dynamiczną naturę procesu obliczeniowego wiele algorytmów jest nietransparentnych, kompletny obraz ich działania jest nieprzenikniony nawet dla ich autorów – dlatego mówi się, że są one rodzajem czarnej skrzynki. Metafora ta odnosi się do sytuacji, w której znane są informacje wejściowe i wyjściowe, ale nie ma (w każdym razie pełnej) wiedzy o sposobie ich przetwarzania. Problem ten dotyczy m.in. powszechnych dziś algorytmów stosowanych w procesie głębokiego uczenia, opartych na sieciach neuronowych. Są one czarną skrzynką w tym sensie, że nie jesteśmy w stanie odtworzyć wszystkich kroków algorytmu i sposobu przetwarzania konkretnych zmiennych; nie wiemy, które spośród ogromnej liczby zmiennych zostały uwzględnione w ukrytych warstwach sieci i tym samym nie potrafimy nadać sensu działaniom tam zachodzącym. A zatem, jakkolwiek mamy dostęp do matematycznej reprezentacji realizowanych przez algorytm procesów, nie potrafimy ich zinterpretować – odpowiedź na pytanie, dlaczego algorytm wskazał określony wynik, jakie były powody tego rozstrzygnięcia, jakie zmienne o tym przesądziły, nie jest możliwa. Brak przejrzystości wynikać może nie tylko ze zjawiska „technicznej” czarnej skrzynki – tak, jak zdefiniowaliśmy ją wyżej – ale także celowego ukrywania mechanizmu działania algorytmu przez podmiot będący jego właścicielem i beneficjentem jego działania.

Problem transparentności algorytmów jest coraz szerzej dyskutowany; podejmuje się próby wypracowania najróżniejszych rozwiązań technicznych i regulacji prawnych mających na celu, jeśli nie zapewnienie przejrzystości, to przynajmniej zredukowanie nieprzejrzystości i ochronę obywateli przed konsekwencjami braku przejrzystości. W ograniczeniu nieprzejrzystości algorytmów, w szczególności wykazaniu ich stronniczości, zakodowanych w nich uprzedzeń, pomocne bywają: inżynieria wsteczna (Diakopoulos, 2014), różne formy audytu algorytmów (Sandvig

et al., 2014), krytyczna analiza kulturowa, analizy etnograficzne, czy interdyscyplinarne badania prowadzone w ramach nurtu studiów nad nauką i technologią, czerpiącego zresztą z wcześniej wymienionych dziedzin (Christin, 2020).

W odpowiedzi na postulaty dotyczące przejrzystości algorytmów rozwija się koncepcja, a także nowy obszar badań, określaną jako: Explainable AI (XAI), Interpretable AI bądź Explainable Machine Learning (XML) (Mohseni et al., 2021). W dyskursie, jaki wypracowano w ramach tego obszaru zwykle przyjmuje się, że przejrzystość (ang. *transparency*) algorytmów jest warunkiem koniecznym ich interpretowalności (ang. *interpretability*), wytłumaczalności (ang. *explainability*), bądź rozliczalności (ang. *accountability*)<sup>6</sup> i stanowi najważniejszą podstawę uzasadnienia podjętej przez algorytm decyzji (ang. *justified decision*). Mówi się również o nowym rodzaju uprawnień – określanym jako Right to Explanation (Kaminski, 2019), czyli o prawie do wyjaśnienia wyniku działania algorytmu. Sądzymy, że nie będzie przesady w stwierdzeniu, że we współczesnym świecie owo uprawnienie urasta do rangi jednego z fundamentalnych praw obywatelskich.

Z drugiej strony pojawiają się stanowiska argumentujące, że algorytmy nie muszą być przejrzyste. Ich orędownicy odwołują się m.in. do koncepcji reliabilizmu obliczeniowego<sup>7</sup>. Pojęcie to wprowadzone zostało przez Juana M. Durána i Nico Formankę (2018), w ich ujęciu odnosiło się do symulacji komputerowych, a do refleksji nad algorytmami zaadaptowane zostało przez Juana M. Durána i Karin Rolandę Jongsmę (2021). Źródłem tej idei jest reliabilizm Alvina Goldmana (1979), który uznaje, że rzetelny proces poznawczy może gwarantować uzasadnione, godne zaufania przekonania. Rzetelne procesy to takie, które zwykle wytwarzają prawdziwe rezultaty. Reliabilści stoją na stanowisku, że przekonanie może być uzasadnione i godne zaufania, nawet jeśli wyznająca je osoba nie rozumie procesu, który czyni to przekonanie wiarygodnym – podobnie jak postrzeganie czegoś za pomocą zmysłu wzroku prowadzi nas do uzyskania przekonania o istnieniu tego czegoś, mimo że nie rozumiemy procesów poznawczych, które odpowiadają za udany akt percepcji<sup>8</sup>. Dodatkowo zwolennicy reliabilizmu obliczeniowego proponują testy pomocne w ocenie rzetelności algorytmu, jak m.in.: weryfikacja, walidacja, ocena ekspertów. Kluczowym sprawdzianem są jednak faktyczne efekty działania

<sup>6</sup> Pojęcia te często bywają używane zamiennie i tu też traktujemy je synonimicznie. Bardziej zniuansowana dyskusja nad ich sensem przekracza ramy niniejszego artykułu.

<sup>7</sup> Przywołujemy to stanowisko jako najlepiej opracowane i najklarowniejsze, ale głosów i argumentów na rzecz odstąpienia od wymogu transparentności algorytmów jest więcej (Ananny & Crawford, 2018).

<sup>8</sup> Ta argumentacja rodzi jednak wiele wątpliwości, wynikających z zastrzeżeń zgłaszanych do reliabilizmu jako takiego (Mordka, 2017; Ziemińska, 2001), w tym tej jego szczególnej formy. Pamiętajmy także, że zdaniem Alvina Goldmana, do którego koncepcji reliabilizmu odwołują się autorzy pojęcia reliabilizmu obliczeniowego, uzasadnienie jest stopniowalne i nigdy ostateczne – zawsze może okazać się fałszywe.

algorytmu (Price, 2018). W tej perspektywie należy przyjąć, że algorytm może być uzasadniony, pomimo braku przejrzystości.

Świadomi tych argumentów stoimy wszakże na stanowisku, że transparentność algorytmów jest bezpośrednio związana z ich interpretowalnością czy też rozliczalnością i stanowi ważny warunek uzasadnialności (ang. *justification*) ich rozstrzygnięć. W pełni świadome używanie algorytmów, w pełni świadome podejmowanie z nimi interakcji, w pełni świadoma zgoda na bycie przedmiotem ich rozstrzygnięć, możliwe są tylko wtedy, gdy w pełni rozumiemy ich działanie (jakkolwiek taką sytuację należy traktować jako idealną). Rozumiejąc, że transparentność algorytmów nie jest zawsze możliwa do osiągnięcia, a uzasadnienie wyników algorytmicznych operacji i zaufanie do nich wynika niekoniecznie z ich przejrzystości, uważamy, że należy dążyć do zawężania obszaru nieprzejrzystości. Zgadza się, że trudno przeciwdziałać nadużyciom i dyskryminacji realizowanym przez algorytmy, jeśli stanowią one czarną skrzynkę. Wyrażamy także przekonanie, że algorytmy przejrzyste i dostępne kontroli społecznej mają większe szanse służyć dobru publicznemu i wartościom demokratycznym.

## 7. Zakończenie

Wszegobecność algorytmów i ich oddziaływanie na niemal każdy aspekt życia indywidualnego oraz społecznego stanowi istotne wyzwanie dla współczesnego świata. Odpowiedzią na nie, formą przygotowania do bardziej świadomego funkcjonowania w algorytmicznej rzeczywistości, powinno być kształtowanie szeroko pojętych kompetencji algorytmicznych. Należy ujmować je wielowymiarowo, pamiętając o aspekcie praktycznym i teoretycznym. Niezbędna wydaje się powszechna edukacja obejmująca przynajmniej elementy umiejętności algorytmicznych i myślenia algorytmicznego (tak, jak zostały one wyżej zdefiniowane) oraz kształtowania świadomości algorytmicznej. Poza wiedzą techniczną, wpływ na kompetencje algorytmiczne ma także wiedza społeczno-kulturowa. Algorytmy często wzmocniają władzę hegemonicznych struktur społecznych i reprodukują istniejące w kulturze uprzedzenia. Dlatego dla jakości kompetencji algorytmicznych, zwłaszcza ich etycznego wymiaru, znaczenie może mieć znajomość struktur i sił społecznych, które za tymi algorytmami stoją. W ujęciu krytycznym na kompetencje algorytmiczne można spojrzeć jak na budowanie świadomości istnienia władzy algorytmicznej i zdolność do stawiania jej oporu.

Przedstawione w tekście wstępne próby budowania narzędzi pomiaru kompetencji algorytmicznych ograniczają się do algorytmów zarządzających treścią w Internecie, pomijają także szeroki społeczny kontekst funkcjonowania algorytmów i zaniedbują wymiar krytyczny. Uważamy, że wymagają one dalszych, pogłębionych prac.

## Bibliografia

- Amoore, L., Poitukh, V., eds. (2015). *Algorithmic Life. Calculative Devices in the Age of Big Data*. London: Routledge.
- Ananny, M., Crawford, K. (2018). Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability. *New Media & Society*, 20(3), 973–989. <https://doi.org/10.1177/14614448166766>
- Bakke, A. (2020). Everyday Googling: Results of an Observational Study and Applications for Teaching Algorithmic Literacy. *Computers and Composition*, 57. <https://doi.org/10.1177/14614448166766>
- Bartsch, M., Dienlin, T. (2016). Control Your Facebook: An Analysis of Online Privacy Literacy. *Computers in Human Behavior*, 56, 147–154. <https://doi.org/10.1016/j.chb.2015.11.022>
- Bucher, T. (2017). The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms. *Information, Communication & Society*, 20(1), 30–44. <https://doi.org/10.1080/1369118X.2016.1154086>
- Bucher, T. (2018). *If... Then: Algorithmic Power and Politics*. New York, NY: Oxford University Press.
- Chen, D. T., Wu, J., Wang, Y. M. (2011). Unpacking New Media Literacy. *Journal of Systems, Cybernetics and Informatics*, 9(2), 84–88.
- Cheney-Lippold, J. (2011). A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control. *Theory, Culture & Society*, 28(6), 164–181. <https://doi.org/10.1177/0263276411424420>
- Christin, A. (2020). The Ethnographer and the Algorithm: Beyond the Black Box. *Theory and Society*, 49(5), 897–918. <https://doi.org/10.1007/s11186-020-09411-3>
- Coffey, H. (2008). *Critical Literacy* [online]. Learn NC, [08.10.2022] [https://teachingaround.com/uploads/1/2/2/8/122845797/critical\\_literacy\\_coffey.pdf](https://teachingaround.com/uploads/1/2/2/8/122845797/critical_literacy_coffey.pdf)
- Cotter, K. (2020). *Critical Algorithmic Literacy: Power, Epistemology, and Platforms*. East Lansing, MI: Michigan State University.
- Culbertson, J. (1986). Chapter VI: Whither Computer Literacy? *Teachers College Record*, 87(5), 109–131.
- D’Ignazio, C., Bhargava, R. (2015). *Approaches to Building Big Data Literacy* [online]. Proceedings of the Bloomberg Data for Good Exchange Conference, [08.10.2022] <https://www.media.mit.edu/publications/approaches-to-building-big-data-literacy/>
- Danaher, J. (2016). The Threat of Algocracy: Reality, Resistance and Accommodation. *Philosophy & Technology*, 29(3), 245–268. <https://doi.org/10.1007/s13347-015-0211--1>
- DeVito, M. A., Gergle, D., Birnholtz, J. (2017). “Algorithms Ruin Everything” #RIPTwitter, Folk Theories, and Resistance to Algorithmic Change in Social Media [online]. CHI ‘17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (3163–3174) [03.01.2023], <https://doi.org/10.1145/3025453.3025659>
- Diakopoulos, N. (2014). *Algorithmic Accountability Reporting: On the Investigation of Black Boxes* [online]. Tow Center for Digital Journalism, Columbia University, [08.10.2022], <https://doi.org/10.7916/D8ZK5TW2>

- Dobson, T.M., Willinsky, J. (2009). Digital Literacy. In: D. Olson, N. Torrance (eds.) *Cambridge Handbook of Literacy*. Cambridge: Cambridge University Press, 286–312.
- Dogrue, L. (2021). What is Algorithm Literacy? A Conceptualization and Challenges Regarding Its Empirical Measurement. In: M. Taddicken, C. Schumann (eds.) *Algorithms and Communication*, 67–93. <https://doi.org/10.48541/dcr.v9.3>
- Dogrue, L., Masur, P., Joeckel, S. (2022). Development and Validation of an Algorithm Literacy Scale for Internet Users. *Communication Methods and Measures*, 16(2), 115–133. <https://doi.org/10.1080/19312458.2021.1968361>
- Domingos, P. (2016). *Naczelny algorytm: Jak jego odkrycie zmieni nasz świat*. Gliwice: Wydawnictwo Helion.
- Doyle, C. S. (1994). *Information Literacy in an Information Society: A Concept for the Information Age*. Syracuse, NY: Syracuse University.
- Durán, J. M., Formanek, N. (2018). Grounds for Trust: Essential Epistemic Opacity and Computational Reliabilism. *Minds and Machines*, 28(4), 645–666. <https://doi.org/10.1007/s11023-018-9481-6>
- Durán, J. M., Jongsma, K. R. (2021). Who is Afraid of Black Box Algorithms? On the Epistemological and Ethical Basis of Trust in Medical AI. *Journal of Medical Ethics*, 47(5), 329–335. <https://doi.org/10.1136/medethics-2020-106820>
- Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., Kirlik, A. (2016). *First I “like” It, Then I Hide It: Folk Theories of Social Feeds* [online]. CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (2371–2382). <https://doi.org/10.1145/2858036.2858494>
- Ezziane, Z. (2007). Information Technology Literacy: Implications on Teaching and Learning. *Journal of Educational Technology & Society*, 10(3), 175–191.
- Festl, R. (2021). Social Media Literacy & Adolescent Social Online Behavior in Germany. *Journal of Children and Media*, 15(2), 249–271. <https://doi.org/10.1080/17482798.2020.1770110>
- Gillespie, T. (2014). The Relevance of Algorithms. In: T. Gillespie, P.J. Boczkowski & K. A. Foot (eds.) *Media technologies: Essays on communication, materiality, and society* (167–193). Cambridge, MA: MIT Press.
- Goldman, A.I. (1979). What is Justified Belief?. In: G.S. Pappas (ed.) *Justification and Knowledge*. Philosophical Studies Series in Philosophy, vol 17. Dordrecht: Springer, 1–25.
- Graff, H. J. (2010). The Literacy Myth at Thirty. *Journal of Social History*, 43(3), 635–661.
- Gray, J., Gerlitz, C., Bounegru, L. (2018). Data Infrastructure Literacy. *Big Data & Society* [online], 5(2), [08.10.2022], <https://journals.sagepub.com/doi/10.1177/2053951718786316>
- Gurevich, Y. (2012). What Is an Algorithm? In: M. Bieliková, G. Friedrich, G. Gottlob, S. Katzenbeisser & G. Turán (eds.) *SOFSEM 2012: Theory and Practice of Computer Science* (31–42). Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-642-27660-6\\_3](https://doi.org/10.1007/978-3-642-27660-6_3)
- Hoffman, M., Blake, J. (2003). Computer Literacy: Today and Tomorrow. *Journal of Computing Sciences in Colleges*, 18(5), 221–233.
- Hutchison, A., Nadolny, L., Estapa, A. (2016). Using Coding Apps to Support Literacy Instruction and Develop Coding Literacy. *The Reading Teacher*, 69(5), 493–503. <https://doi.org/10.1002/trtr.1440>

- Issar, S., Aneesh, A. (2022). What is Algorithmic Governance? *Sociology Compass* [online], 16(1), [08.10.2022], <https://doi.org/10.1111/soc4.12955>
- Iwasiński, Ł. (2017). Quantified Self. Self-tracking a problem tożsamości. *Zagadnienia Informatyki Naukowej*, 55(2), 126–136. <https://doi.org/10.36702/zin.369>
- Iwasiński, Ł. (2020). Social Implications of Algorithmic Bias. In: B. Sosińska-Kalata, M. Roszkowski & Z. Wiorogórska (eds.) *Nauka o informacji w okresie zmian. Rewolucja cyfrowa – dziś i jutro: Infrastruktura, usługi, użytkownicy* (25–35). Warszawa: Wydaw. Naukowe i Edukacyjne SBP.
- Kaminski, M. E. (2019). The Right to Explanation, Explained. *Berkeley Technology Law Journal* [online], 34, [08.10.2022], <https://doi.org/10.15779/Z38TD9N83H>
- Kampa, P., Balzer, F. (2021). Algorithmic Literacy in Medical Students – Results of a Knowledge Test Conducted in Germany. *Health Information & Libraries Journal*, 38(3), 224–230. <https://doi.org/10.1111/hir.12392>
- Khoo, E., Hight, C., Torrens, R., Cowie, B., eds. (2017). *Software Literacy: Education and Beyond*. Singapore: Springer.
- Kim, E. M., Yang, S. (2016). Internet Literacy and Digital Natives' Civic Engagement: Internet Skill Literacy or Internet Information Literacy? *Journal of Youth Studies*, 19(4), 438–456. <https://doi.org/10.1080/13676261.2015.1083961>
- Klug, D., Qin, Y., Evans, M., Kaufman, G. (2021). *Trick and Please. A Mixed-Method Study on User Assumptions About the TikTok Algorithm*. In: WebSci '21: 13th ACM Web Science Conference 2021 (84–92). <https://doi.org/10.1145/3447535.3462512>
- Knuth, D. E. (1973). *The Art of Computer Programming* (Vol. 3). Reading, MA: Addison-Wesley
- Koltay, T. (2011). The Media and the Literacies: Media Literacy, Information Literacy, Digital Literacy. *Media, Culture & Society*, 33(2), 211–221. <https://doi.org/10.1177/01634437103933>
- Koltay, T. (2017). Data Literacy for Researchers and Data Librarians. *Journal of Librarianship and Information Science*, 49(1), 3–14. <https://doi.org/10.1177/0961000615616450>
- Kreft, J. (2019). *Władza algorytmów: u źródeł potęgi Google i Facebooka*. Kraków: Wydaw. Uniwersytetu Jagiellońskiego.
- Kuiper, E., Volman, M., Terwel, J. (2008). Students' Use of Web Literacy Skills and Strategies: Searching, Reading and Evaluating Web Information. *Information Research: An International Electronic Journal* [online], 13(3), [08.10.2022], <https://eric.ed.gov/?id=EJ837273>
- Livingstone, S. (2004). What Is Media Literacy? *Intermedia*, 32(3), 18–20.
- Livingstone, S. (2008). Internet Literacy: Young People's Negotiation of New Online Opportunities. In: T. McPherson (eds.) *Digital Youth, Innovation, and the Unexpected* (101–121). Cambridge, MA: MIT Press.
- Livingstone, S. (2014). Developing Social Media Literacy: How Children Learn to Interpret Risky Opportunities on Social Network Sites. *Communications*, 39(3), 283–303. <https://doi.org/10.1515/commun-2014-0113>
- Luke, C. (2007). As Seen on TV or Was That My Phone? New Media Literacy. *Policy Futures in Education*, 5(1), 50–58. <https://doi.org/10.2304/pfie.2007.5>
- Lyon, D. (2003). *Surveillance As Social Sorting: Privacy, Risk, and Digital Discrimination*. London: Routledge.

- Mager, A. (2012). Algorithmic Ideology: How Capitalist Society Shapes Search Engines. *Information, Communication & Society*, 15(5), 769–787. <https://doi.org/10.1080/1369118X.2012.676056>
- Mai, J. E. (2019). Situating Personal Information: Privacy in the Algorithmic Age. In: R. F. Jørgensen (ed.) *Human rights in the age of platforms* (95–116). Cambridge, MA: MIT Press.
- McClure, C. R. (2018). Network Literacy in an Electronic Society: An Educational Disconnect? In: R. Kubey (ed.) *Media Literacy in an Information Age: Current Perspectives* (403–439). New Brunswick: Transaction Publisher.
- Medhat, W., Hassan, A., Korashy, H. (2014). Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Miron, M., Nadon, R. (2006). Inferential Literacy for Experimental High-Throughput Biology. *Trends in Genetics*, 22(2), 84–89. <https://doi.org/10.1016/j.tig.2005.12.001>
- Miyazaki, S. (2012). Algorhythmic: Understanding Micro-Temporality in Computational Cultures. *Computational Culture*, 2, 1–16.
- Mohseni, S., Zarei, N., Ragan, E. D. (2021). A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 11(3/4), 1–45. <https://doi.org/10.1145/3387166>
- Mordka, C. (2017). Zawodność reliabilizmu A. Goldmana w kontekście sporu o uzasadnienie. *Annales Universitatis Mariae Curie-Skłodowska, Sectio I–Philosophia-Sociologia*, 42(2), 89–101. <http://dx.doi.org/10.17951/i.2017.42.2.89>
- Porter, T.M. (1995). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
- Potter, W. J. (2013). Review of Literature on Media Literacy. *Sociology Compass*, 7(6), 417–435. <https://doi.org/10.1111/soc4.12041>
- Price, W. N. (2018). Big Data and Black-Box Medical Algorithms. *Science Translational Medicine*, 10(471). <https://doi.org/10.1126/scitranslmed.aao5333>
- Przegalińska, A., Oksanowicz, P. (2020). *Sztuczna inteligencja: nieludzka, arcyłudzka*. Kraków: Wydaw. Znak.
- Ptucha, R., Such, F. P., Pillai, S., Brockler, F., Singh, V., Hutkowski, P. (2019). Intelligent Character Recognition Using Fully Convolutional Neural Networks. *Pattern Recognition*, 88, 604–613. <https://doi.org/10.1016/j.patcog.2018.12.017>
- Reddy, P., Sharma, B., Chaudhary, K. (2020). Digital Literacy: A Review of Literature. *International Journal of Technoethics (IJT)*, 11(2), 65–94. <https://doi.org/10.4018/IJT.20200701.0a1>
- Ridley, M., Pawlick-Potts, D. (2021). Algorithmic Literacy and the Role for Libraries. *Information Technology and Libraries* [online], 40(2), [08.10.2022], <https://doi.org/10.6017/ital.v40i2.12963>
- Sandvig, C., Hamilton, K., Karahalios, K., Langbort, C. (2014). Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms. *Data and Discrimination: Converting Critical Concerns Into Productive Inquiry*, 22, 4349–4357.
- Seel, N. M., Casey, N. C. (2003). Changing Conceptions of Technological Literacy. *Disadvantaged Teens and Computer Technologies*, 35–55.

- Snavely, L., Cooper, N. (1997). The Information Literacy Debate. *The Journal of Academic Librarianship*, 23(1), 9–14.
- Stiller, E., LeBlanc, C. (2006). From Computer Literacy to Cyber-Literacy. *Journal of Computing Sciences in Colleges*, 21(6), 4–13.
- Szpunar, M. (2019). *Kultura algorytmów*. Kraków: Instytut Dziennikarstwa, Mediów i Komunikacji Społecznej Uniwersytetu Jagiellońskiego.
- Uricchio, W. (2011). The Algorithmic Turn: Photosynth, Augmented Reality and the Changing Implications of the Image. *Visual Studies*, 26(1), 25–35. <https://doi.org/10.1080/1472586X.2011.548486>
- Vanschoren, J. (2018). Meta-learning: A survey [online]. Arxiv, [08.10.2022], <https://doi.org/10.48550/arXiv.1810.03548>
- Walker, M. S., Hughes, T. A. (2008). Messenger RNA Expression Profiling Using DNA Microarray Technology: Diagnostic Tool, Scientific Analysis or Un-Interpretable Data? *International Journal of Molecular Medicine*, 21(1), 13–17. <https://doi.org/10.3892/ijmm.21.1.13>
- Wodecki A. (2021). *Sztuczna inteligencja we współczesnych organizacjach*. Warszawa: Wydaw. PWN
- Zarouali, B., Boerman, S. C., de Vreese, C. H. (2021). Is This Recommended By an Algorithm? The Development and Validation of the Algorithmic Media Content Awareness Scale (AMCA-scale). *Telematics and Informatics* [online], 62, [08.10.2022], <https://doi.org/10.1016/j.tele.2021.101607>
- Zawojski, P. (2018). Rewolucja algorytmiczna. We władzy cyfr i liczb. *Opcje*, 4, 14–21.
- Ziemińska, R. (2001). Alvina I. Goldmana teoria wiedzy i uzasadnienia. *Roczniki Filozoficzne*, 77–107.
- Žagar, M., Samardžija, J., Mutka, A., Mestrovic, A. H., Mihaljevic, B., Vejzagic, V., Labas, I. (2021). Digital Platform Education For Stem Professions. In: *EDULEARN21 Proceedings*, 3074–3078. <https://dx.doi.org/10.21125/edulearn.2021.0657>
- 

## How to Be a Conscious User of Algorithms? On the Necessity of Developing Algorithmic Literacy

### Abstract

**Purpose/Thesis:** The ubiquity of algorithms and their impact on almost every aspect of the individual's life and society pose a significant challenge to the modern world. We should respond by developing algorithmic literacy. The main purpose of this article is to analyze this concept. We also indicate why this type of literacy is crucial for effective and informed functioning in the modern world.

**Approach/Methods:** The article employs a narrative literature review and critical discourse analysis. The analysis is based on theoretical and empirical literature, and illustrated with the authors' own observations.

**Results and conclusions:** Algorithmic literacy should be considered in all its practical and theoretical aspects. It seems necessary to include at least elements of algorithmic literacy in general education. Algorithmic literacy is influenced not only by technical knowledge, but also by cultural literacy. Algorithms often reinforce hegemonic social structures and reproduce existing cultural biases. Therefore, it may be vital to know social structures and forces behind algorithms to possess algorithmic literacy, particularly when it concerns the ethics of algorithms. From a critical perspective, algorithmic literacy can be seen as raising the awareness of the algorithms' power and acquiring the ability to resist it.

**Originality/Value:** Discussions of algorithmic literacy and empirical studies of the concept are not far-advanced. This article critically discusses the current state of research in this area, thus far unrecognized in Polish literature.

### Keywords

Algorithmic literacy. Algorithmic transparency. Algorithmic awareness. Algorithmic power.

---

*Dr ŁUKASZ IWASIŃSKI jest adiunktem w Katedrze Informatologii na Wydziale Dziennikarstwa, Informacji i Bibliologii Uniwersytetu Warszawskiego. Absolwent Wydziału Organizacji i Zarządzania Politechniki Łódzkiej (tytuł magistra inżyniera – 2006 r.) oraz socjologii na Uniwersytecie Łódzkim (tytuł magistra – 2007 r., stopień doktora – 2013 r.). Jego zainteresowania badawcze obejmują socjologię kultury, wiedzy i nauki, w szczególności wpływ technologii informacyjnych i komunikacyjnych na społeczeństwo i społeczne konsekwencje danetyzacji, Najważniejsze publikacje: Społeczne zagrożenia danetyzacji rzeczywistości (B. Sosińska-Kalata i in., red., Nauka o informacji w okresie zmian. Informatologia i humanistyka cyfrowa, 2016), Theoretical bases of critical data studies (Zagadnienia Informacji Naukowej, 2020), Big data a problem reprezentacji poznawczej (Człowiek i Społeczeństwo, 2022).*

*Kontakt z autorem:*

*[liwasinski@uw.edu.pl](mailto:liwasinski@uw.edu.pl)*

*Katedra Informatologii*

*Wydział Dziennikarstwa, Informacji i Bibliologii*

*Uniwersytet Warszawski*

*ul. Nowy Świat 69*

*00-046 Warszawa*

*WERONIKA FURMAN jest absolwentką studiów I stopnia na kierunku Kognitywistyka na Uniwersytecie Jagiellońskim. Obecnie jest studentką studiów II stopnia na kierunkach Kognitywistyka oraz Architektura przestrzeni informacyjnych na Uniwersytecie Warszawskim.*

*Kontakt z autorką:*

*[w.furman3@student.uw.edu.pl](mailto:w.furman3@student.uw.edu.pl)*

*Wydział Dziennikarstwa, Informacji i Bibliologii*

*Uniwersytet Warszawski*

*ul. Nowy Świat 69*

*00-046 Warszawa*

# Architektura informacji w witrynach internetowych Prawa i Sprawiedliwości oraz Platformy Obywatelskiej

Paulina Sajna-Kosobucka

ORCID: 0000-0002-8587-1280

*Interdyscyplinarna Szkoła Doktorska Nauk Społecznych – Academia Rerum Socialium  
Uniwersytet Mikołaja Kopernika w Toruniu*

---

## Abstrakt

**Cel/Teza:** Głównym celem badań jest ocena jakości architektury informacji w witrynach internetowych dwóch partii dominujących obecnie na polskiej scenie politycznej: Prawa i Sprawiedliwości (PiS) oraz Platformy Obywatelskiej (PO).

**Koncepcja/Metody badań:** W badaniach wykorzystano triangulację kilku metod badawczych: ocenę jakościowo-heurystyczną architektury informacji na stronach WWW partii PiS i PO, ocenę ekspercką, analizę porównawczą witryn oraz obserwacje własne.

**Wyniki i wnioski:** Choć obie witryny charakteryzują się standardową jakością, zostały zaprojektowane z uwzględnieniem zasad funkcjonalności. W obrębie badanych stron WWW widoczny jest brak niektórych istotnych składowych architektury informacji. Na podstawie przeprowadzonych analiz stwierdzono, że witryna partii PO jest jakościowo nieco lepsza od witryny partii PiS.

**Zastosowania praktyczne:** Wyniki przeprowadzonych badań mogą być podstawą dla doskonalenia witryn poddanych analizie. Mogą mieć one również wpływ nie tylko na ocenę zaawansowania technologicznego partii PiS i PO, ale i także na ich wizerunek oraz odbiór społeczny, a w związku z tym też skuteczność w dotarciu do wyborców danej partii.

**Oryginalność/Wartość poznawcza:** Przygotowane opracowanie ukazuje nową interpretację metody badawczej, którą jest ocena jakościowo-heurystyczna oparta na 74 kryteriach skupionych na potrzebach i kompetencjach użytkownika danego serwisu. Owo podejście wspomaga badania stron WWW w celu poprawy jakości doświadczeń użytkowników i bezproblemowej realizacji procesów informacyjnych.

## Słowa kluczowe

Algotymy Google. Analiza jakościowo-heurystyczna. Architektura informacji. Partie polityczne. Platforma Obywatelska. Prawo i Sprawiedliwość. Strony WWW. Systemy etykietowania. Systemy nawigacyjne. Systemy organizacyjne. Systemy wyszukiwawcze.

*Otrzymany: 4 kwietnia 2022. Zrecenzowany: 26 czerwca 2022. Poprawiony: 13 września 2022. Zrecenzowany ponownie: 18 grudnia 2022. Poprawiony ponownie: 23 grudnia 2022. Zaakceptowany: 23 grudnia 2022.*

---

## 1. Wstęp

Architektura informacji wywodzi się z nauk bibliologicznych i jest młodą meta-dyscypliną naukową, której celem jest zadbanie o bezproblemowy dostęp do informacji w przestrzeni informacyjnej. Według Stanisława Skórki: „Z punktu widzenia architektury informacji (AI) przestrzeń informacyjna definiowana jest jako przestrzeń składająca się z trzech oddziałujących na siebie elementów: treści, kontekstu i użytkownika” (Skórka, 2011, 47). Pewien rodzaj przestrzeni informacyjnej stanowią zatem witryny internetowe. Z kolei architekturę informacji w sensie praktycznym pojmuje się jako proces organizowania przestrzeni informacyjnej z wykorzystaniem odpowiednich narzędzi, na które składają się: systemy organizacyjne, systemy etykietowania, systemy nawigacyjne i systemy wyszukiwawcze (zob. Rosenfeld et al., 2017). Wreszcie, przez architekturę informacji konkretnej przestrzeni informacyjnej rozumie się określoną jej strukturę organizacyjną, wyznaczoną przez zastosowanie tych systemów i determinującą możliwości dostępu użytkowników do zawartych w niej treści. Obszar zainteresowań nauki o architekturze informacji obejmuje środowiska informacyjne badane z perspektywy: „projektowania, tj. nadawania im kształtu, struktury, organizowania treści, łączenia nawigacją; komunikowania – np. zastosowania odpowiedniego kodu i formy do utrwalenia i przekazu treści; ewaluacji – oceny np. zastosowanych kategorii zawartości, metadanych, składowych ułatwiających zrozumienie, efektywności systemów nawigacji itp.” (Skórka, 2021, 11). Kryteria, względem których dokonuje się oceny jakościowo-heurystycznej architektury informacji wybranych stron WWW, obejmują aspekty projektowe i komunikacyjne, które, oprócz wymienionych systemów, mogą odnosić się także do dodatkowych elementów czy własności witryn, ustalonych przez badacza. W procesie ewaluacji warto uwzględnić np. to, czy dana strona WWW spełnia wymogi stawiane przez algorytmy wyszukiwarki Google, jako najpopularniejszej wyszukiwarki na świecie. Algorytmy te indeksują treści i hierarchizują wyniki wyszukiwania. Niewątpliwie dla każdego właściciela serwisu internetowego, a w szczególności dla partii politycznych, pozycja w wynikach wyszukiwania odgrywa ważną rolę. Prowadząc zatem badania nad architekturą informacji, nie powinno się pomijać oceny stopnia, w jakim przyjęte rozwiązania spełniają wymogi narzucone przez wspomniane algorytmy.

Internet jest najważniejszym narzędziem politycznym w krajach rozwiniętych, w których istnieją gwarancje wolności słowa i powszechnego dostępu do nowych technologii komunikacyjnych. Siła Internetu była widoczna już w połowie lat 90. XX w., np. gdy internauci wspierali ruch zapatystowski w Meksyku (Darling, 2008; Muñoz Ramírez, 2008). Ponadto już na początku XXI w. specjaliści od mediatyzacji polityki zauważyli, że politycy coraz bardziej interesowali się wówczas poszukiwaniem sposobów dotarcia do odbiorców przez Internet (Street, 2001, 159), przy czym zakładano, że polityka, zarządzanie i tożsamość ulegną zmianie

z powodu wdrożenia nowych technologii medialnych (Louw, 2005, 118). Według Philippe'a J. Maareka istnieją trzy grupy zastosowań Internetu w komunikacji politycznej: instytucjonalne zastosowania polityczne, zastosowania w kampaniach wyborczych oraz zastosowania przez grupy nacisku i grupy „marginalne” (Maarek, 2016). Z perspektywy niniejszego artykułu kluczowe są pierwsze dwie kategorie, gdyż obejmują one oficjalne strony internetowe partii politycznych.

Serwisy WWW partii stanowią narzędzie w walce politycznej, służące pozyskaniu nowych wyborców, kreowaniu wizerunku i utrzymaniu właściwych relacji ze zwolennikami partii. Ponadto serwisy WWW są elektroniczną wersją dawnych partyjnych gazet i ulotek informacyjnych. Witryny internetowe partii politycznych prowadzone są jednak nie tylko dla zwolenników partii i nowych wyborców, ale też dla członków partii (na co wskazuje m.in. zakładka „Archiwum” na oficjalnej stronie WWW partii PiS, chroniona hasłem), konkurentów, analityków, naukowców oraz innych – heterogenicznych – grup odbiorców. Partie polityczne, zwłaszcza te największe, nie chcą ograniczać się do specyficznych grup docelowych – ich serwisy są otwarte dla wszystkich. Wobec tego każdy odwiedzający te witryny powinien mieć możliwość sprawnego poruszania się po nich – nawet odbiorcy najmniej kompetentni, których potrzeby i możliwości też muszą zostać wzięte pod uwagę. Pod tym właśnie kątem została przeprowadzona ocena jakościowo-heurystyczna stron WWW dwóch wybranych polskich partii politycznych. Z tego typu serwisów korzystają ludzie należący do różnych kategorii wiekowych i o różnym poziomie wykształcenia – nawet osoby niepełnoletnie (nastoletnie, zaczynające interesować się polityką) czy bez wykształcenia. Wobec tego interfejs serwisów partyjnych powinien charakteryzować się intuicyjnością i nie powinien zawierać zbyt specjalistycznego słownictwa ani mieć zbyt skomplikowanej budowy.

W niniejszym artykule przedstawione są wyniki badania, którego przedmiot stanowi architektura informacji w witrynach internetowych dwóch partii dominujących w ostatnich latach na polskiej scenie politycznej: Prawa i Sprawiedliwości (PiS) oraz Platformy Obywatelskiej (PO).

## 2. Polityka i nowe technologie a architektura informacji

W przypadku partii PiS i PO bardzo wyraźnie sprawdza się stwierdzenie, że polityczność można traktować jako walkę (Haratyk, 2011, 41); we wszystkich kanałach komunikowania wykorzystywanych przez wspomniane partie można dostrzec bezpośrednie celowe działania, zmierzające do ukształtowania konkretnych postaw i poglądów, przy jednoczesnym krytykowaniu przeciwnika. Obserwacja treści na stronach internetowych tych partii pozwala stwierdzić, że polityczna „batalia” toczona jest przez nie m.in. w Internecie. Ponieważ serwisy WWW partii pełnią rolę narzędzia w walce politycznej, przekłada się to na stosowane zasady funkcjonalności

architektury informacji: im lepiej zaprojektowane systemy architektury informacji, tym lepsza korelacja z potrzebami, oczekiwaniami oraz kompetencjami zwolenników partii, nowych potencjalnych wyborców oraz członków partii.

Zarówno architektura informacji, jak i rola Internetu oraz nowych technologii w polityce są przedmiotami licznych badań. Problematyka architektury informacji pojawiała się m.in. w publikacjach: Martyna Dade-Robertsona (2011), Nathaniela Davisa (2010; 2018), Petera van Dijcka (2003), Ewy Głowackiej (2016), Amira Hosseina Seddighi (2021), Briana Arbogasta de Huberta-Millera (2006), Katarzyny Lange-Sadzińskiej (2011), Veslavy Osińskiej (2014), Zbigniewa Osińskiego (2014), Marcina Roszkowskiego (2019), Stanisława Skórki (2011; 2021), Barbary Sosińskiej-Kalaty (2013), Piotra Tafiłowskiego (2016) czy Doroty Utrackiej (2017). Badania nad rolą nowych technologii i Internetu w polityce podejmowali m.in.: Dominika Borowska i Nikola Rozworska (2022), Adam Drosik (2010), Tomasz Gajowniczek (2019), Jan Garlicki i Daniel Mider (2012), Anna Leszczuk-Fiedziukiewicz (2011), Johanna Neuman (1996), Zbigniew Oniszczyk (2018), Eugene B. Skolnikoff (1992), David Trend (1997) oraz Christopher Weare (2002). Interdyscyplinarne badania łączące architekturę informacji i nauki o polityce w kontekście e-demokracji i partyjnych stron internetowych przeprowadziła Andrea Ricci (2013). Natomiast Małgorzata Cieślak-Florczyk (2014) zbadała strony internetowe pięciu polskich partii politycznych, jednak wykorzystując inne metody badawcze niż zastosowano w tym artykule. Dotychczas nie powstała publikacja skoncentrowana na architekturze informacji w witrynach internetowych polskich partii politycznych.

### 3. Organizacja i metody badań

Głównym celem badań jest ocena jakości architektury informacji w witrynach internetowych partii PiS (<https://pis.org.pl>) oraz PO (<https://platforma.org>) jako dominujących partii na polskiej scenie politycznej. W badaniach wykorzystano triangulację kilku metod badawczych: ocenę jakościowo-heurystyczną architektury informacji w wybranych witrynach, ocenę ekspercką, analizę porównawczą wybranych stron WWW oraz obserwacje własne.

Kryteria oceny jakościowo-heurystycznej architektury informacji zostały opracowane na podstawie książki Louisa Rosenfelda, Petera Morville'a i Jorge Arango (2017). Podzielono je na cztery grupy odpowiadające wspomnianym wcześniej systemom architektury informacji oraz dwie grupy kryteriów dodatkowych i kryteriów oceny realizacji wymogów stawianych przez algorytmy Google. Dokonano oceny eksperckiej porównywalnych elementów składowych architektury informacji w badanych witrynach internetowych bez wykorzystywania w tym celu specjalistycznego narzędzia, gdyż nie jest ono konieczne do tego typu badań. W załączniku do tego artykułu zawarto pełny katalog kryteriów oceny jakościowo-heurystycznej,

którym łącznie przypisano 76 punktów. Punkty przypisywane były za jakość ocenianych elementów na podstawie trójelementowej skali (0, 0.5, 1): jeśli dany element był obecny na stronie, ale nie został zaprojektowany prawidłowo, wówczas nie przyznano punktu; gdy dany element funkcjonował częściowo poprawnie i tym samym tylko częściowo spełniał oczekiwania użytkownika, przyznano pół punktu; za prawidłowo opracowany element witryna mogła uzyskać maksymalnie jeden punkt. W przypadku dwóch specyficznych kryteriów (prawidłowego funkcjonowania okna wyszukiwarki na stronie głównej oraz prawidłowego funkcjonowania mechanizmu wyszukiwania zaawansowanego) przyznano maksymalnie po dwa punkty. Za brakujące składowe systemów przyznano zero punktów. Oczywiście meritum nie stanowi wymóg, aby każda podstrona witryny korzystała ze wszystkich możliwych schematów i opcji, gdyż przeładowana strona traci na przejrzystości i prostocie korzystania; nie na każdej podstronie muszą znajdować się takie elementy, jak np. chmura tagów czy indeks alfabetyczny, ale z perspektywy całego serwisu, jeśli dane elementy rozmieszczone są prawidłowo na różnych podstronach, mogą poprawić doświadczenia użytkowników i usprawnić dotarcie do informacji. Bardzo dobrym przykładem jest aktualna wersja strony internetowej CIA.gov (stan na dzień 1.09.2022), która zawiera prawie wszystkie możliwe elementy architektury informacji w podziale na różne podstrony, a jednocześnie nie jest przeładowana i charakteryzuje się nowoczesnym layoutem, przyjaznym dla użytkownika. Taki efekt uzyskano m.in. dzięki licznym, odpowiednio rozlokowanym na podstronach, dynamicznym, rozwijającym menu.

Jakość strony WWW powinna być oceniana z punktu widzenia potrzeb i kompetencji grup docelowych, dlatego przyjęte kryteria oceny uwzględniają perspektywę użytkownika badanych serwisów. Badania przeprowadzono w sposób metodyczny, dążąc do zapewnienia obiektywności oceny zgodnie z wyznaczonymi kryteriami oraz przestrzegając bezstronności politycznej.

## 4. Analiza jakościowo-heurystyczna architektury informacji w witrynach internetowych

### 4.1. Witryna internetowa partii Prawo i Sprawiedliwość

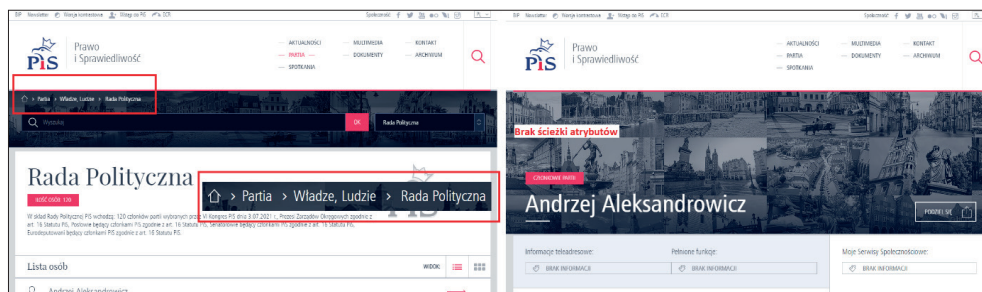
Już po dokonaniu wstępnych obserwacji strony WWW partii PiS można zauważyć, że witryna ogólnie robi dobre wrażenie, a pierwszym elementem, który przykuwa wzrok użytkownika, jest baner ze zdjęciem prezesa partii PiS Jarosława Kaczyńskiego. Ponadto jego postać widnieje na każdym slajdzie w menu dynamicznym (wg stanu na dzień: 01.09.2022). Szczegółowe badania wykazały, że w witrynie partii PiS wykorzystano wiele schematów organizowania informacji, tj.: schemat chronologiczny w aktualnościach; zadaniowy – m.in. opcje typu „podziel się”

i „zobacz multimedia” oraz „zapisz się” itd.; tematyczny jako podział na sekcje tematyczne – np. „historia partii”; otwarty – jako że nie trzeba logować się poprzez konto użytkownika, aby mieć pełen dostęp do serwisu, z wyjątkiem „archiwum”, do którego dostęp mają tylko upoważnieni użytkownicy; audytorium – kierowany do konkretnej grupy ludzi: obywatele Polski; i metaforę – w architekturze informacji schemat ten oznacza odniesienia do sytuacji z życia, nawiązujący m.in. do wartości. W ramach struktur organizacyjnych można tu wyróżnić model hierarchiczny – szeroki i płytki, widoczny w menu i hipertekstowy – oparty na hiperłączach.

Systemy etykietowania funkcjonują bez zarzutu. Jedno kryterium oceny tych systemów wydaje się być zależne od składowej systemu nawigacyjnego, tj. „etykiety pozwalające na nawigację powrotną w formie ścieżek atrybutów”. W przypadku witryny internetowej partii PiS, nawigacja powrotna do kategorii nadrzędnych w formie ścieżki atrybutów nie na wszystkich podstronach jest dostępna, jednak wszędzie tam, gdzie się pojawia, etykiety są bezbłędne. Mechanizm ścieżki atrybutów zaprojektowany został tylko dla niektórych sekcji. Brak tej ścieżki stanowi zdecydowane utrudnienie poruszania się po stronie WWW<sup>1</sup>. Analizując system nawigowania i docierania do informacji, zauważono, że w obrębie witryny zachodzi specyficzne zjawisko: jedenastu członków partii przedstawiono na stronie głównej w formie slajdera, jednak nie sposób trafić do miejsca, w którym wymienieni zostali wszyscy członkowie. Najpierw należy udać się do podstrony pod nazwą „Partia”, a następnie: „Władze, Ludzie” – co nie jest zbyt oczywiste, gdy serwis nie jest znany użytkownikowi. Jednak to jeszcze nie koniec procesu, jeśli użytkownik chce zdobyć informacje o osobie spoza prezydium partii. Na podstronie „Władze, Ludzie” zaprezentowani są: prezes i wiceprezesi partii, a pod spodem znajdują się kategorie opatrzone nagłówkiem „Ludzie Prawa i Sprawiedliwości”. Kategorie te wyświetlane są w formie kafelkowej i jest ich osiem, a każda stanowi osobną podstronę: są to np. „Komitet Polityczny”, „Komisja Etyki” itd. Wybrana podstrona zawiera listę członków partii, a każdy wiersz z listy stanowi kolejną podstronę. Po przejściu do wybranej osoby, np. Marzeny Machałek, widnieją nieuzupełnione pola typu „informacje teledadresowe” i „pełnione funkcje”; na podstronie znajduje się tylko odnośnik do profilu na Facebooku. W niektórych przypadkach brakuje nawet odnośników do mediów społecznościowych (zob. Rys. 1). Gdy użytkownik dokonuje przeglądu członków partii, nie może cofnąć się do kategorii nadrzędnej z poziomu interfejsu i właśnie m.in. w tym przypadku brak ścieżek atrybutów bardzo utrudnia przeglądanie witryny. Przez to łatwo jest się zgubić, co może wpłynąć na zniechęcenie użytkownika – tym bardziej, że po przeszukiwaniach kolejnych

<sup>1</sup> Dla przykładu, ścieżka dostępu wyświetla się dla podstrony „Rada Polityczna” i innych równorzędnych podstron (ikona domu [*home*] – „Partia” – „Władza, Ludzie” – „Rada Polityczna”), ale po przejściu do konkretnego członka partii (tzn. do kolejnej podstrony) ścieżka znika i nie można powrócić do kategorii nadrzędnej z perspektywy interfejsu.

podstron nie znajduje on informacji, które go interesują (tu: informacji o członkach partii). Atutem maskującym tę niedogodność jest wyszukiwarka członków partii (zob. Rys. 1.), dająca możliwość wyszukiwania w konkretnych sekcjach. Systemy nawigowania są ogólnie dobre jakościowo i oprócz standardowej nawigacji wbudowanej zawierają też nawigację pomocniczą, choć niestety ubogą.



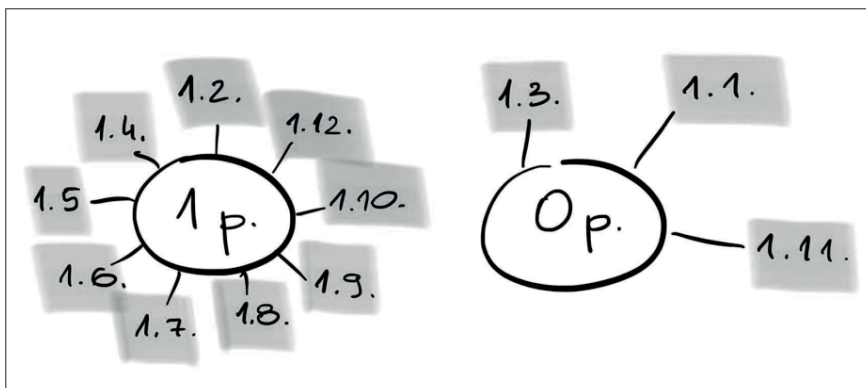
Rys. 1. Dwa zestawione zrzuty ekranu podstrony witryny internetowej partii PiS – ze ścieżką atrybutów i bez tej ścieżki.

Opracowanie własne na podstawie: Prawo i Sprawiedliwość [21.12.2022],  
[https://pis.org.pl/partia/wladze-ludzie/rada\\_polityczna](https://pis.org.pl/partia/wladze-ludzie/rada_polityczna)  
 oraz [https://pis.org.pl/osoba/andrzej\\_aleksandrowicz](https://pis.org.pl/osoba/andrzej_aleksandrowicz)

W systemach wyszukiwawczych wyszukiwarka ma charakter podstawowy z elementem zaawansowania, którym jest możliwość sortowania elementów od najstarszych do najnowszych i na odwrót lub wyszukiwania w konkretnych sekcjach witryny (zob. Rys. 1), a system podpowiada jedynie wyszukiwane wcześniej frazy, bez podpowiedzi ze strony podsystemów samej witryny. Wyszukiwarka mobilna, tak samo jak webowa, funkcjonuje poprawnie. Na uwagę zasługuje duży rozmiar okna wyszukiwarki w tej witrynie, co przyciąga uwagę użytkownika.

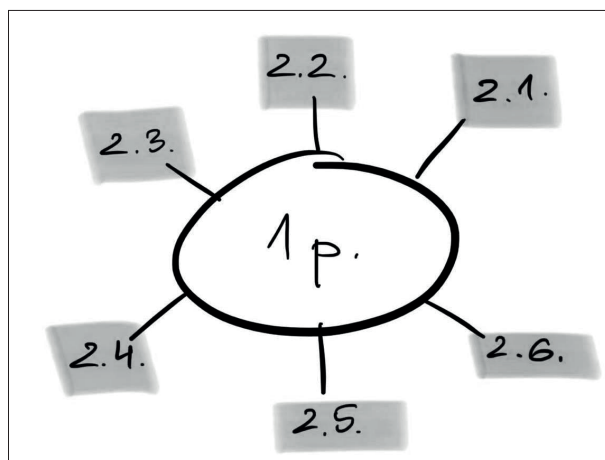
Odnośnie kryteriów dodatkowych należy zwrócić uwagę na niezaktualizowaną stopkę, w której widnieje rok 2018 (stan na dzień 1.09.2022). Inne aspekty wymagające uwagi to: ilość tekstu na stronie głównej – np. przytłaczający dla użytkownika, zbyt długi i słabo wyróżniony tekst opisu partii; bezpieczeństwo łącza – w listopadzie 2021 r. łącze było niezabezpieczone, ale w 2022 r. sytuacja uległa zmianie; brak możliwości ustawienia większej czcionki dla osób niedowidzących, jest za to możliwość ustawienia wersji kontrastowej. Mocną zaletą są opisy alternatywne plików graficznych. Nagłówki witryny aktualizowane są średnio co kilka miesięcy.

Poniżej, w formie grafów (Rys. 2–7) przedstawiono liczbę punktów przyznanych stronie WWW partii PiS w ocenie poszczególnych kryteriów (opis kryteriów szczegółowych – zob. Załącznik, Rys. 17–22).



Rys. 2. Grafy przedstawiające wynik oceny strony WWW partii PiS na podstawie kryteriów dotyczących systemów organizacyjnych

W wyniku analizy jakościowo-heurystycznej, na podstawie kryteriów odnoszących się do systemów organizacyjnych, strona WWW partii PiS uzyskała 9 na możliwych 12 punktów (rys. 2).

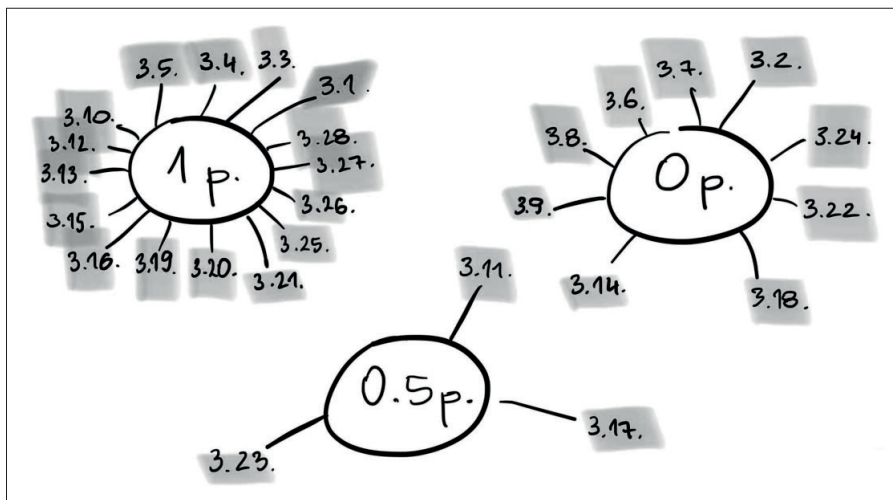


Rys. 3. Graf przedstawiający wynik oceny strony WWW partii PiS na podstawie kryteriów dotyczących systemów etykietowania

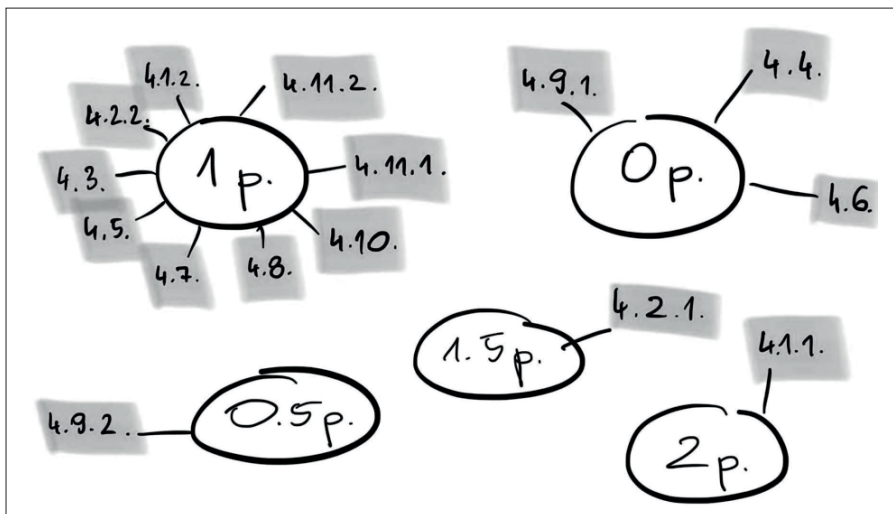
Na podstawie oceny systemów etykietowania (rys. 3) stronie WWW partii PiS przyznano 6, tj. pełną pulę punktów możliwych do uzyskania za tę grupę kryteriów.

Z kolei ocena systemów nawigowania (rys. 4) pozwoliła przypisać stronie WWW partii PiS 17.5 na 28 punktów możliwych do uzyskania za tę grupę kryteriów. Po pół punktu przyznano za kryteria 3.11. (możliwość eksportu wyników i metadanych jako system nawigacji pomocniczej), 3.17. (ścieżki atrybutów z bezpośrednim dostępem do kategorii nadrzędnych) i 3.23. (menu rozwijane), ponieważ można

udostępnić konkretne nagłówki do mediów społecznościowych, ale nie można zapisać wszystkich wyników, ścieżka atrybutów jest dostępna tylko w wybranych sekcjach strony WWW, a menu rozwijane jest dostępne tylko w wersji mobilnej. Kontrowersyjną kwestią jest przyznanie 0 punktów za mapę witryny (kryterium 3.6.), jednak jest ona zdecydowanie za słabo rozbudowana.

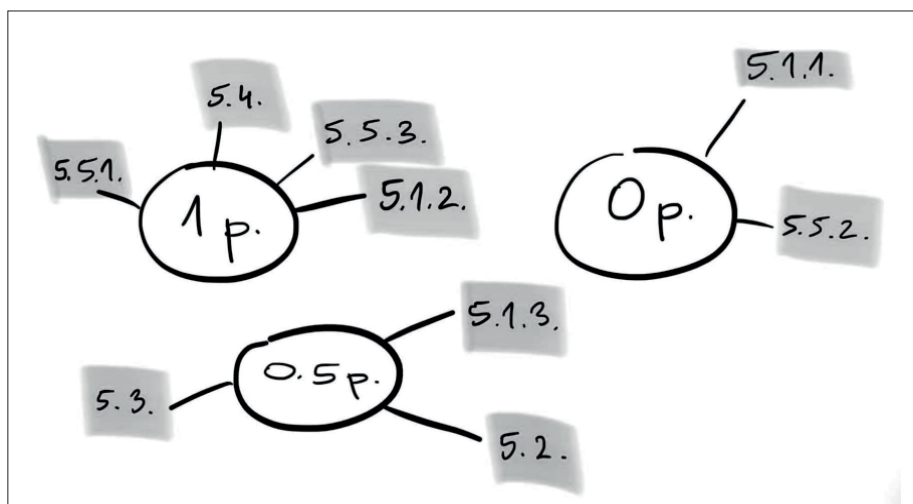


Rys. 4. Grafy przedstawiające wynik oceny strony WWW partii PiS na podstawie kryteriów dotyczących systemów nawigacyjnych



Rys. 5. Grafy przedstawiające wynik oceny strony WWW partii PiS na podstawie kryteriów dotyczących systemów wyszukiwawczych

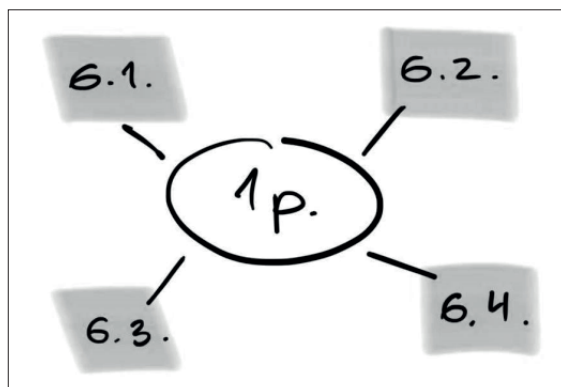
Ocena na podstawie kryteriów odnoszących się do systemów wyszukiwania pozwoliła przyznać stronie WWW partii PiS 13 na 17 punktów możliwych do uzyskania za tę grupę kryteriów (Rys. 5). Pół punktu przyznano za kryterium 4.9.2. (rozsadne i poprawnie działające opcje sortowania wyników wyszukiwania), ponieważ narzędzia sortowania zapewniają tylko porządkowania wyników od najnowszych do najstarszych wiadomości i na odwrót. Natomiast 1.5 punktu na 2 przyznano za kryterium 4.2.1. (prawidłowo działający mechanizm wyszukiwania zaawansowanego), gdyż zaawansowane mechanizmy wyszukiwawcze są ograniczone w swojej budowie.



Rys. 6. Grafy przedstawiające wynik oceny strony WWW partii PiS na podstawie grupy dodatkowych kryteriów

Analiza jakościowo-heurystyczna na podstawie grupy kryteriów dodatkowych pozwoliła przyznać stronie WWW partii PiS 5.5 na 9 punktów możliwych do uzyskania (Rys. 6). Po pół punktu przyznano za kryteria 5.1.3. (tekst w stopce ograniczony do minimum), 5.2. (aktualność wyróżnionych nagłówków w witrynie) oraz 5.3. (odpowiednia ilość tekstu na stronie głównej), ponieważ: nie zaktualizowano daty w stopce, nagłówki aktualizowane są raz na kilka miesięcy<sup>2</sup>, a tekst w witrynie internetowej jest przeciętnie wyróżniony, co może wpłynąć negatywnie na jego odbiór.

<sup>2</sup> Wynikać to może jednak z faktu, że niektóre nagłówki przez swoją specyfikę są wciąż aktualne – nawet przez kilka miesięcy. Mimo to brak aktualizacji nie wygląda najlepiej z perspektywy użytkownika.



Rys. 7. Graf przedstawiający wynik oceny strony WWW partii PiS na podstawie kryteriów dotyczących spełniania wymogów algorytmów Google

Analiza na podstawie ostatniej grupy kryteriów, związanych ze spełnianiem wymagań algorytmów wyszukiwarki Google pozwoliła przyznać stronie WWW partii PiS maksimum punktów możliwych do uzyskania za tę grupę kryteriów, tj. 4 punkty (Rys. 7).

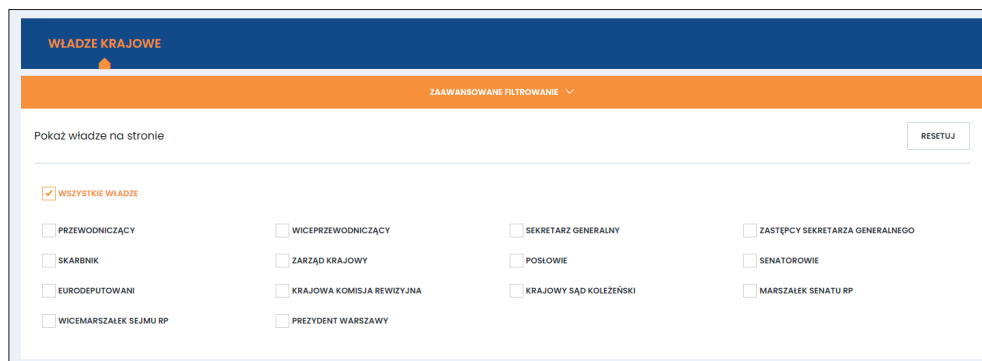
#### 4.2. Witryna internetowa partii Platforma Obywatelska

Podobnie jak na stronie WWW partii PiS, także na stronie głównej witryny internetowej PO widnieje baner ze zdjęciem przewodniczącego partii (tu: Donalda Tuska), jego postać widnieje też na każdym slajdzie w menu dynamicznym (stan na dzień: 01.09.2022). W witrynie partii PO jakość systemów organizacyjnych jest wysoka, brakuje jedynie schematu geograficznego. Schemat alfabetyczny pojawia się tylko w jednym fragmencie witryny: na podstronie z etykietą „Ludzie”, w sekcji „Wiceprzewodniczący”. W pozostałych sekcjach kolejność nazwisk nie jest uporządkowana alfabetycznie. Na tej stronie WWW wykorzystano też wszystkie modele struktur organizowania: model hierarchiczny – hierarchia wąska i głęboka oraz szeroka i płytka, model bazodanowy – na co wskazuje m.in. możliwość filtrowania wyników wyszukiwania względem miejscowości, oraz model hipertekstowy.

Systemy etykietowania funkcjonują bezbłędnie. W witrynie znajdują się liczne etykiety tekstowe i graficzne, a każda z nich sprawnie pełni swoją rolę.

W systemach nawigowania widoczny jest błąd na mapie strony, która nie zawiera dwóch zakładek zawartych w menu, tj. „Sekretariat samorządowy” i „Dokumenty”, ale za to widnieje dodatkowa zakładka: „Nasz Program” (stan na dzień: 01.09.2022). Za namiastkę przewodników można uznać komunikaty typu „przewiń stronę” i informacje nad oknem wyszukiwawczym: „Może zasugerujemy?” (tu: hasło wyszukiwawcze) oraz (sic!) „Kliknij ENTER aby wyszukać albo ESC aby wyjść”. Większość opcji systemów nawigacyjnych funkcjonuje bez zarzutu.

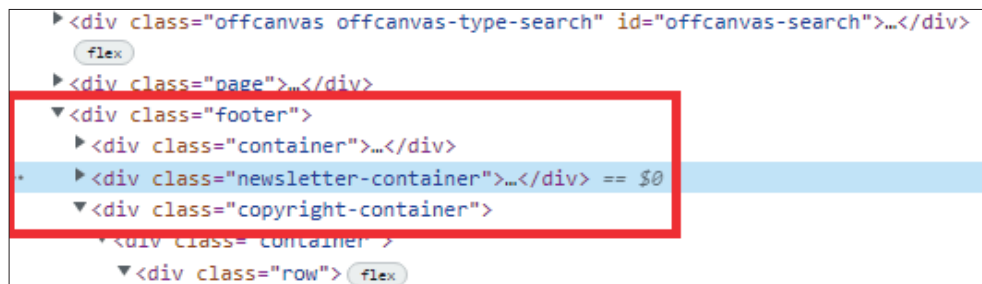
Systemy wyszukiwawcze pełnią swoją funkcję prawidłowo, a zaletą jest główna wyszukiwarka wewnętrzna, oferująca gotowe zapytania. Wyszukiwarka ta nie jest jednak tak zaawansowana, jak wyszukiwarka pomocnicza zawierająca „zaawansowane filtrowanie” w zakładce „Ludzie” (Rys. 8).



Rys. 8. Fragment podstrony WWW Platformy Obywatelskiej: „Ludzie”.

Źródło: Ludzie – Platforma Obywatelska [21.12.2022],

<https://platforma.org/ludzie>



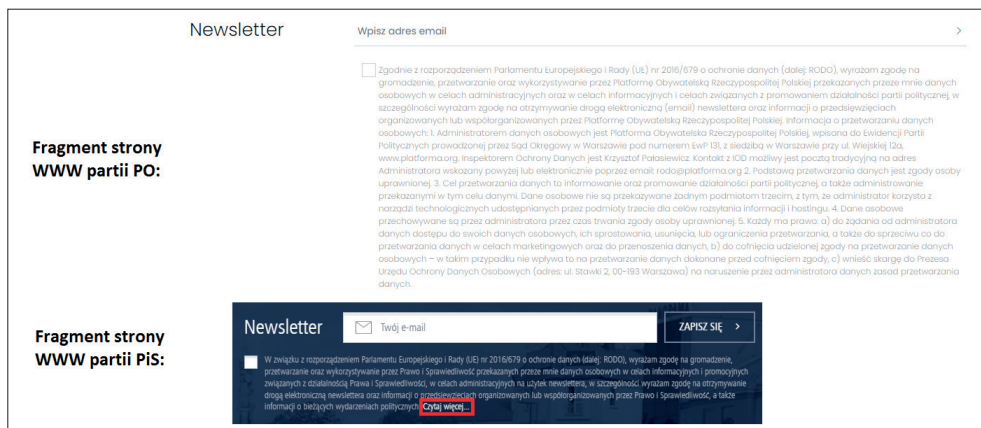
Rys. 9. Fragment kodu źródłowego stopki strony WWW Platformy Obywatelskiej.

Opracowanie własne na podstawie: Platforma Obywatelska [21.12.2022],

<https://platforma.org/>

W wyszukiwarce głównej, po wyszukiwaniu, brakuje podpowiedzi w postaci propozycji dodatkowych typu: „zobacz też...”. Przechodząc do dodatkowej puli kryteriów, można zauważyć, że nie wszędzie ulokowano opisy alternatywne plików graficznych. Dużym atutem są jednak: możliwość ustawienia kontrastu i większych liter, zabezpieczone łącze, regularnie aktualizowane nagłówki, zaktualizowana stopka oraz odpowiednia ilość tekstu na stronie głównej, czego jednak nie można stwierdzić w przypadku tekstu zawartego – według kodu źródłowego (Rys. 9) – w sekcji *footer*; nadmierną ilość tekstu przy formularzu do newslettera można było zwinąć i oznaczyć opcją „Czytaj dalej”, tak jak w przypadku witryny

partii PiS. Porównanie organizacji tej sekcji na stronach WWW obu partii przedstawiono na Rys. 10.



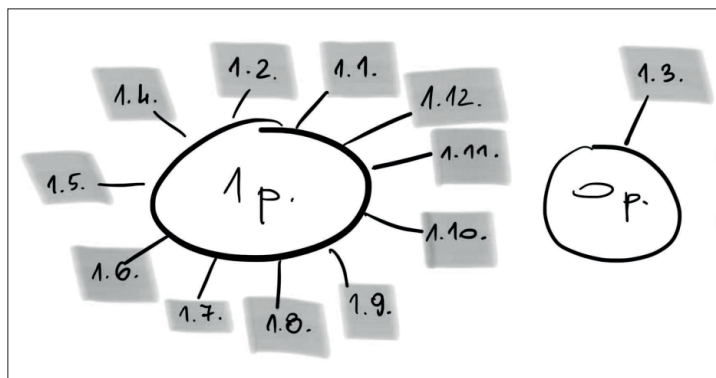
Rys. 10. Fragmenty sekcji *footer* w witrynach partii PiS oraz PO z możliwością zapisu do newslettera.

Opracowanie własne na podstawie:

Platforma Obywatelska [21.12.2022], <https://platforma.org/>  
 oraz Prawo i Sprawiedliwość [21.12.2022], <https://pis.org.pl/>

Zdecydowane atuty interfejsu w serwisie partii PO to profesjonalnie wykonana oś czasu w opowiadaniu historii partii i wszelkie inne elementy interaktywne, np. ruchome wykresy czy przejrzysta prezentacja klubu parlamentarnego. Członkowie partii podzieleni są na kategorie etykietami i separatorami, co utrzymuje porządek na stronie, przy czym niektórzy członkowie partii nie zostali opisani (podobnie jak niektórzy członkowie PiS na stronie WWW tej partii). Jednak w witrynie partii PO proces przeglądania członków partii jest dużo prostszy niż na stronie WWW partii PiS. Ciekawym rozwiązaniem zastosowanym w interfejsie serwisu partii PO jest prezentacja wyników wyszukiwania, które wyświetlają się w różnych kategoriach, przy czym widnieje m.in. podział na listę z artykułami i dokumentami. Ponadto media prezentują się w formie kafelkowej w postaci kafli średniej wielkości, a „ludzie” – również w formie kafelkowej, jednak są to kafle mniejsze. Pod wynikami wyszukiwania wyświetla się hasło wyszukiwawcze w dużym formacie – tak samo, jak podczas wpisywania tego hasła. Wyniki wyszukiwawcze dla różnych, wypróbowanych w trakcie badania haseł, są jednak dość ograniczone. Nawet po wpisaniu haseł podpowiadanych przez sam mechanizm wyszukiwawczy wyświetla się zaledwie kilka wyników. Z tego powodu nie można określić rodzaju nawigacji służącej do przechodzenia przez strony z wynikami, co znacząco obniżyło łączną punktację za systemy nawigacyjne.

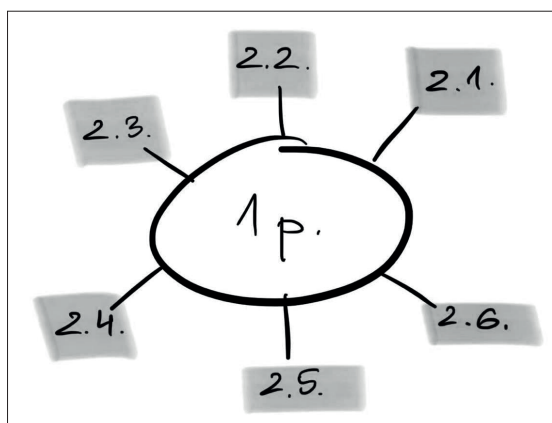
Poniżej za pomocą grafów (Rys. 11–16) przedstawiono wyniki oceny strony WWW partii PO na podstawie poszczególnych kryteriów, których szczegółowy opis znajduje się w Załączniku.



Rys. 11. Grafy przedstawiające wynik oceny strony WWW partii PO na podstawie kryteriów dotyczących systemów organizacyjnych

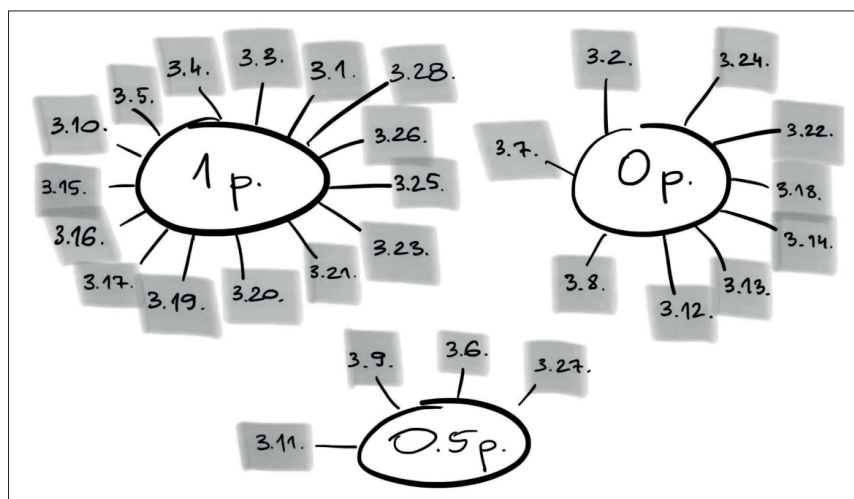
Analiza strony WWW partii PO pod względem jakości i różnorodności systemów organizacyjnych pozwoliła przypisać jej 11 punktów na 12 możliwych do uzyskania w tej grupie kryteriów.

Graf na rysunku 12 demonstruje wynik oceny systemów etykietowania zastosowanych na stronie WWW partii PO. W odniesieniu do tej grupy kryteriów stronie przyznano maksymalną liczbę punktów, tj. 6.



Rys. 12. Graf przedstawiający wynik oceny strony WWW partii PO na podstawie kryteriów dotyczących systemów etykietowania

Jakość i różnorodność systemów nawigacyjnych zastosowanych na stronie WWW partii PO oceniona została na 17 punktów (na 28 możliwych do uzyskania). Po pół punktu przyznano za: ograniczoną możliwość eksportu wyników i metadanych, podobnie jak w przypadku witryny partii PiS (kryterium 3.11.), *quasi*-przewodnik (kryterium 3.9.) oraz mapę serwisu (kryterium 3.6.) – czytelną, jednak niekompletną.

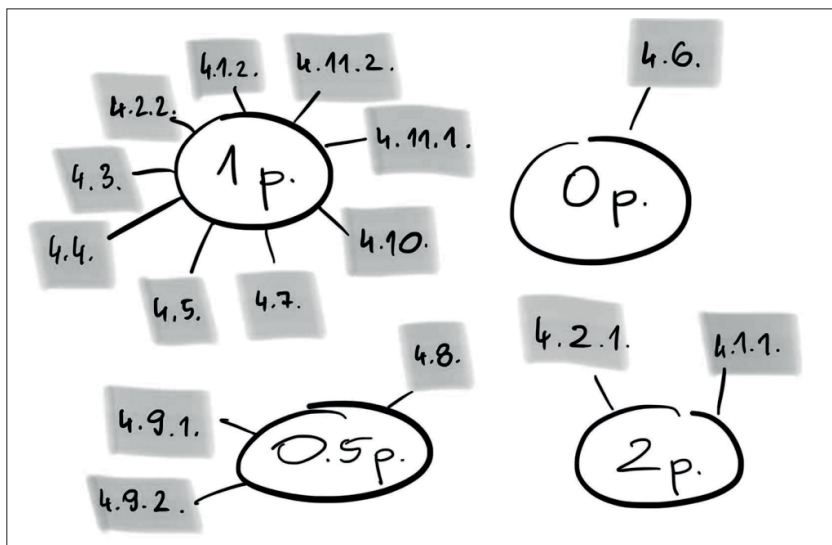


Rys. 13. Grafy przedstawiające wynik oceny strony WWW partii PO na podstawie kryteriów dotyczących systemów nawigacyjnych

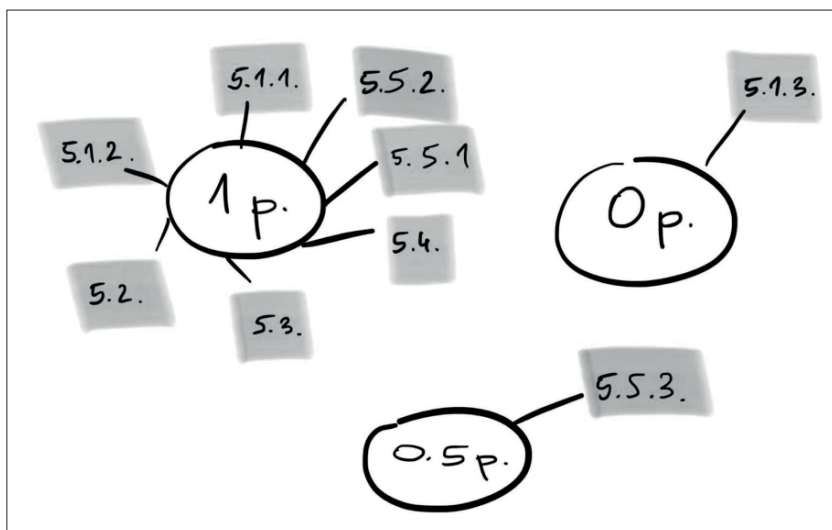
Jakość i różnorodność systemów wyszukiwania zastosowanych na stronie WWW partii PO została oceniona na 14.5 punktu na 17 możliwych do uzyskania (Rys. 14). Pół punktu przyznano za sposoby filtrowania zaawansowanego (kryterium 4.9.1.), gdyż to kryterium skierowane jest w stronę głównej wyszukiwarki wewnętrznej, a nie wyszukiwarek pomocniczych. Gdyby wyszukiwarka główna zawierała przynajmniej takie samo „zaawansowane filtrowanie” jak wyszukiwarka pomocnicza na podstronie „Ludzie” (Rys. 8), punktacja wzrosłaby. Kolejne pół punktu przyznano za sposoby sortowania wyników (kryterium 4.9.2.), a ostatnie pół punktu – za system pozwalający na oszacowanie liczby wyników wyszukiwania (kryterium 4.8.).

Analiza jakościowo-heurystyczna na podstawie grupy kryteriów dodatkowych pozwoliła przyznać stronie WWW partii PO 7.5 na 9 punktów możliwych do uzyskania w tej części oceny (Rys. 15). Pół punktu przyznano za opisy alternatywne plików graficznych (kryterium 5.7.). Kwestią dyskusyjną jest ocena kryterium 5.1.3. (tekst w stopce ograniczony do minimum), gdyż wizualnie stopka wydaje się być odseparowana od newslettera opatrzonego nadmierną ilością tekstu i zawierać wyłącznie logo partii, rok, informacje o właścicielu i prawach autorskich oraz

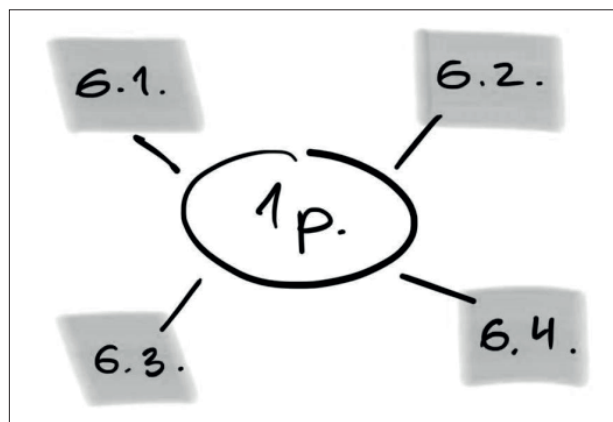
logo grupy EPL. W kodzie źródłowym witryny widać niestety, iż sekcja newsletteru znajduje się w sekcji stopki (Rys. 9), która z kolei obejmuje też mapę strony, znajdującą się nad newsletterem.



Rys. 14. Grafy przedstawiające wynik oceny strony WWW partii PO na podstawie kryteriów dotyczących systemów wyszukiwawczych



Rys. 15. Grafy przedstawiające wynik oceny strony WWW partii PO na podstawie dodatkowych kryteriów



Rys. 16. Graf przedstawiający wynik oceny strony WWW partii PO w oparciu na podstawie kryteriów dotyczących spełniania wymogów algorytmów Google

Analiza strony WWW partii PO na podstawie ostatniej grupy kryteriów pozwoliła stwierdzić, że spełnia ona wszystkie wymogi stawiane przez wybrane algorytmy Google (Rys. 16).

## 5. Analiza porównawcza i ocena ekspercka

Ocena ekspercka opiera się na uśrednieniu wyników oceny przeprowadzonej na podstawie różnych kryteriów zastosowanych w analizie jakościowo-heurystycznej. Uśrednienie takie uzyskano, przypisując ocenie witryny na podstawie każdej z wyodrębnionych grup kryteriów taką samą skalę punktów, która odwzorowuje odsetek punktów uzyskanych przez witrynę za poszczególne grupy kryteriów w analizie jakościowo-heurystycznej. Zabieg ten dla laika może wydawać się zabiegiem enigmatycznym, nie mającym uzasadnienia i nie dostarczającym precyzyjnej informacji zwrotnej. Jednak jest wprost odwrotnie, gdyż z obserwacji własnych wynika, że jeśli dane grupy systemów architektury informacji mieszczą się w punktacji w przedziale od 0% do 20%, ich podsystemy nie funkcjonują wystarczająco dobrze. Ta sama tendencja zauważalna jest w przypadku holistycznej oceny całego serwisu pod kątem architektury informacji. W przypadku przedziału od 0.00% do 20.99% przydzielono 0 punktów, gdyż tak niski wynik oznacza niedostateczne funkcjonowanie podsystemów. Według tej samej zasady, od 21.00% do 39.99% przyznano 0.5 punktu; od 40.00% do 59.99% – 1 punkt; od 60.00% do 79.99% – 1.5 punktu; od 80.00% do 100% – maksymalną liczbę punktów, tj. 2 na 2.

Jak pokazują dane zebrane w tabeli 1, strona WWW partii PiS spełnia bez zarzutu wymogi wybranych algorytmów Google. Bezbłędnie funkcjonują systemy etykietowania w tej witrynie. Nieco słabiej, ale dość dobrze prezentuje się ocena

systemów wyszukiwawczych, organizacyjnych i nawigacyjnych. Najslabiej w witrynie internetowej PiS wypadła ocena elementów analizowanych na podstawie kryteriów dodatkowych, co ostatecznie obniżyło całościową ocenę jakości witryny. Mimo pewnych niedoskonałości, strona WWW partii PiS jest zaprojektowana zgodnie z zasadami zapewniania funkcjonalności.

Tab. 1. Zestawienie wyników oceny architektury informacji w witrynie internetowej partii PiS na podstawie analizy jakościowo-heurystycznej z wynikami oceny eksperckiej

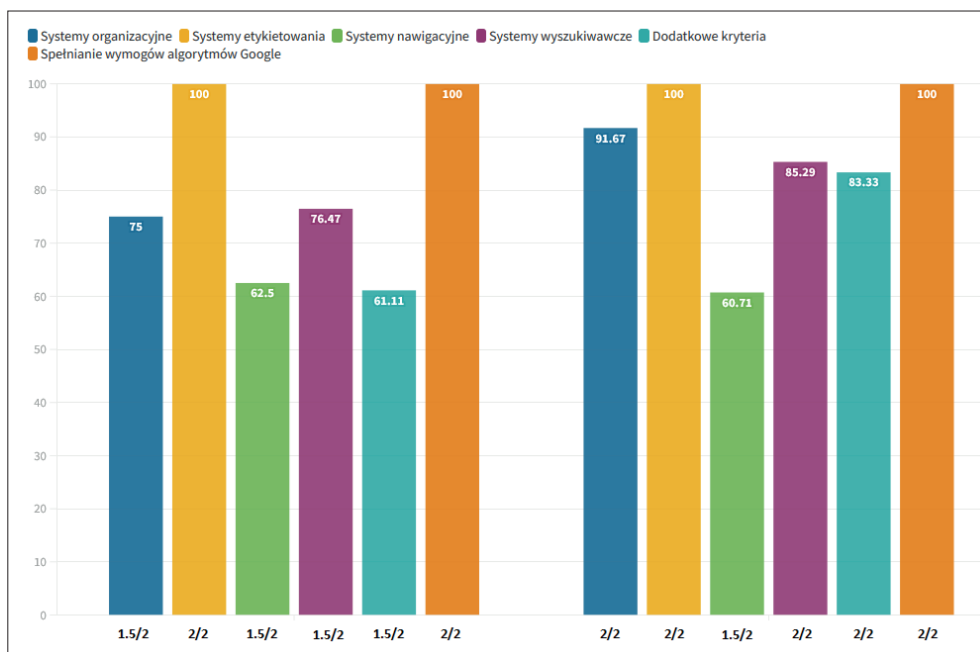
|  | Ocena jakościowo-heurystyczna architektury informacji<br>liczba punktów uzyskanych /<br>liczba punktów możliwych do<br>uzyskania [odsetek punktów<br>możliwych do uzyskania] | Ocena ekspercka<br>liczba punktów uzyskanych /<br>liczba punktów możliwych<br>do uzyskania |
|--|--|--|
| Systemy organizacyjne                                | 9 / 12 [75.00%]  | 1.5 / 2  |
| Systemy etykietowania                                | 6 / 6 [100%]   | 2 / 2  |
| Systemy nawigacyjne                                  | 17.5 / 28 [62.5%]  | 1.5 / 2  |
| Systemy wyszukiwawcze                                | 13 / 17 [76.47%]   | 1.5 / 2  |
| Dodatkowe kryteria                                   | 5.5 / 9 [61.11%]   | 1.5 / 2  |
| Spełnienie wymogów<br>wybranych algorytmów<br>Google | 4 / 4 [100.00%]  | 2 / 2  |

Tab. 2. Zestawienie wyników oceny architektury informacji w witrynie internetowej partii PO na podstawie analizy jakościowo-heurystycznej z wynikami oceny eksperckiej

|  | Ocena jakościowo-heurystyczna architektury informacji<br>liczba punktów uzyskanych /<br>liczba punktów możliwych do<br>uzyskania [odsetek punktów<br>możliwych do uzyskania] | Ocena ekspercka<br>liczba punktów uzyskanych /<br>liczba punktów możliwych<br>do uzyskania |
|--|--|--|
| Systemy organizacyjne                                | 11 / 12 [91.67%]   | 2 / 2  |
| Systemy etykietowania                                | 6 / 6 [100.00%]  | 2 / 2  |
| Systemy nawigacyjne                                  | 17 / 28 [60.71%]   | 1.5 / 2  |
| Systemy wyszukiwawcze                                | 14.5 / 17 [85.29%]   | 2 / 2  |
| Dodatkowe kryteria                                   | 7.5 / 9 [83.33%]   | 2 / 2  |
| Spełnienie wymogów<br>wybranych algorytmów<br>Google | 4 / 4 [100.00%]  | 2 / 2  |

Dane zaprezentowane w tabeli 2 pokazują, że witrynie internetowej partii PO przyznano wysoką punktację za jakość składowych architektury informacji. Najlepiej prezentują się systemy etykietowania, a wymogi wybranych algorytmów Google spełnione są bez zastrzeżeń. Architektura informacji tej strony WWW przejawia też wysoką jakość systemów organizacyjnych, wyszukiwawczych oraz elementów ocenianych na podstawie dodatkowych kryteriów. Najsłabiej w tej witrynie internetowej zostały ocenione systemy nawigacyjne. Mimo to także w odniesieniu do witryny PO można stwierdzić, że wszelkie zasady funkcjonalności zostały w niej wdrożone.

Wykres 1 prezentuje porównanie oceny badanych stron WWW na podstawie wyników jakościowo-heurystycznej analizy ich architektury informacji oraz uśrednionych wyników oceny eksperckiej.



Wykr. 1. Analiza porównawcza wyników oceny architektury informacji witryn PiS i PO (ocena jakościowo-heurystyczna wyrażona w odsetkach spełniania wyznaczonych kryteriów i ocena ekspercka wyrażona w punktach).

Opracowanie własne na podstawie: PiS vs. PO | Flourish [22.12.2022],  
<https://public.flourish.studio/visualisation/10804424/>

Z analizy porównawczej przedstawionej na wykresie 1 wynika, że systemy organizacyjne witryny partii PiS są słabsze jakościowo od systemów witryny PO o 16.67%, systemy wyszukiwawcze – o 8.82%, zaś elementy oceniane na podstawie

dotychczasowych kryteriów prezentują się słabiej aż o 22.22%. Systemy nawigacyjne są lepsze jakościowo na stronie WWW partii PiS o 1.79% (o przewadze zdecydował brak nawigacji krokowej i stronicowej oraz przewijania w witrynie partii PO, w systemie prezentacji wyników wyszukiwania z perspektywy interfejsu głównej wyszukiwarki wewnętrznej). Systemy etykietowania funkcjonują tak samo bezbłędnie w przypadku witryn obu partii, podobnie – obie strony WWW bez zarzutu spełniają wymogi algorytmów Google.

Ocena jakościowo-heurystyczna oraz ocena ekspercka architektury informacji badanych witryn w ujęciu holistycznym zaprezentowana została w tabeli 3.

Tab. 3. Ocena jakościowo-heurystyczna oraz ocena ekspercka architektury informacji witryn PiS i PO w ujęciu holistycznym

| Witryna partii | Ocena jakościowo-heurystyczna architektury informacji             |  | Ocena ekspercka<br>liczba punktów uzyskanych / liczba punktów możliwych do uzyskania |
|----------------|---|--|--|
|                | liczba punktów uzyskanych / liczba punktów możliwych do uzyskania | odsetek punktów możliwych do uzyskania |  |
| PiS            | 55 / 76   | 72.37%                                 | 1.5 / 2  |
| PO             | 60 / 76   | 78.95%                                 | 1.5 / 2  |

Z danych przedstawionych w tabeli 3 wynika, że witryna partii PO zdobyła przewagę wysokości 6.58% (5 punktów) nad witryną partii PiS pod względem jakości architektury informacji. Generalnie należy ponownie podkreślić, że obie strony WWW zaprojektowano z uwzględnieniem zasad funkcjonalności badanych składowych, gdyż wynik końcowy dla stron obu partii wynosi 1.5 na 2 punkty. Stronę WWW partii PO, od uzyskania maksymalnej punktacji (2 / 2) dzieli jednak tylko 1.05%.

## 6. Wnioski

Badania wykazały, że obie strony WWW – partii PiS i PO – charakteryzują się standardową, porównywalną jakością, przy czym witrynę partii PO dzieli niewiele punktów procentowych od najlepszego wyniku oceny eksperckiej. Dostęp do informacji w obrębie tych witryn nie jest specjalnie ułatwiony, ale po zapoznaniu się z serwisami stwierdzono, że nie ma w nich znaczących problemów. Niemniej interfejs witryny powinien być intuicyjny, a architektura informacji powinna wspierać użytkownika i pozwolić zaoszczędzić mu czas. W obrębie badanych stron WWW brak niektórych istotnych składowych architektury informacji, które zostały wskazane w artykule. Obie strony WWW zaprojektowano zgodnie z zasadami zapewnienia funkcjonalności.

Należy zwrócić uwagę, że twórcy witryn partii mieli różne pomysły projektowe, stąd zidentyfikowane w nich błędy i/lub braki w architekturze informacji mają różny charakter. We współczesnej praktyce projektowania witryn internetowych widoczne jest podążanie za pewnymi trendami globalnymi, co prowadzi do wykorzystywania podobnych rozwiązań. W przypadku badanych stron partii politycznych nie można mówić o radykalnych różnicach, choć na tle innych witryn często one występują. Niewątpliwie członkom partii politycznych zależy na pozytywnym odbiorze, dlatego witryny obu partii generalnie robią dobre wrażenie i są poprawnie skonstruowane.

W dalszych badaniach można dodatkowo przeanalizować witryny pod kątem lewostronnego tekstu głównego w witrynach, gdyż w architekturze informacji zakłada się, że tekst lewostronny jest łatwiejszy w odczycie od pozostałych układów treści. Warto też dokonać badań doświadczeń konkretnych użytkowników (ang. *UX research*) na Skali Poczucia Alienacji „Z” (zob. Turska-Kawa, 2011) i w ich ramach bazować m.in. na dyferencjale semantycznym. Wówczas, na podstawie wyników ankiety, można uzupełnić wnioski i dokonać jeszcze bardziej rozbudowanej oceny eksperckiej. Wyniki przeprowadzonych badań mogą mieć wpływ nie tylko na ocenę zaawansowania technologicznego partii, ale i na ich wizerunek oraz odbiór społeczny, a w związku z tym też skuteczność w dotarciu do wyborców danej partii.

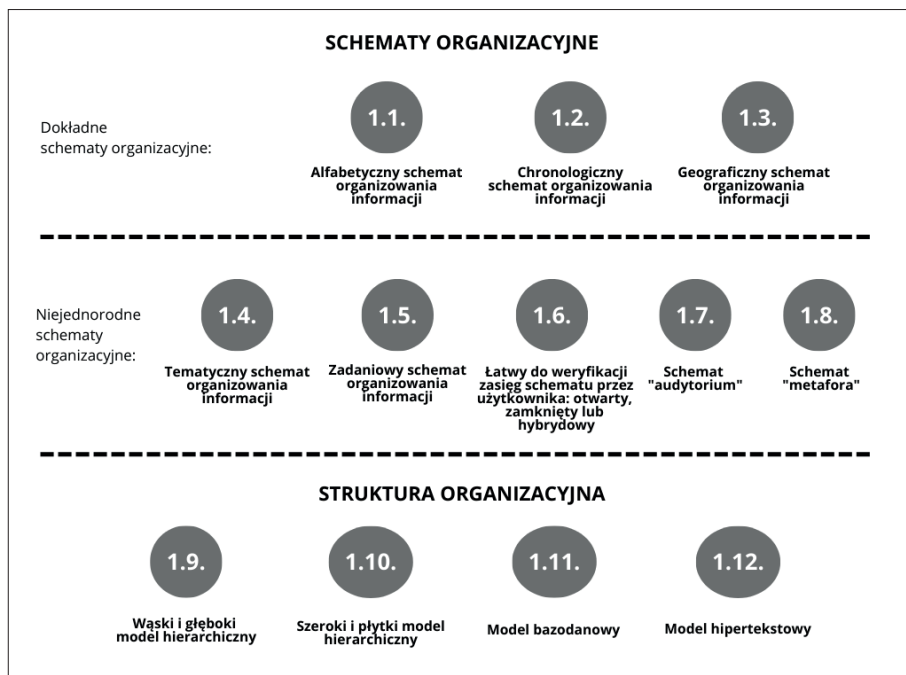
Ze względu na rosnącą popularność korzystania z serwisów internetowych na urządzeniach mobilnych potrzebne jest przeprowadzenie analizy mobilnych wersji stron WWW. Projektanci powinni dbać o to, by witryny wyświetlały się bez zakłóceń na wszystkich urządzeniach. Wstępne badania wersji mobilnych stron WWW partii PiS i PO wykazały, że zakłócenia w wyświetlaniu treści niestety występują na różnych ekranach – pewne elementy nakrywają się, inne nie wyświetlają się w całości itd. Jednak w badaniach omówionych w niniejszym artykule responsywność nie była przedmiotem oceny. W przyszłości warto byłoby poszerzyć analizę o ten aspekt.

W kryteriach oceny uwzględniono spełnianie przez badane witryny wymagań algorytmów wyszukiwarki Google, gdyż jest najbardziej popularną wyszukiwarką na świecie. Niemniej warto uwzględnić też inne wyszukiwarki oraz ich algorytmy w swojej strategii optymalizacyjnej. Do tych wyszukiwarek, często wykorzystywanych przez użytkowników, należą obecnie również: Bing, Yahoo, Baidu (Grupa iCEA, b.d.) czy Yandex (Kurzak, 2022). W dalszych badaniach można poszerzyć analizę stron WWW o stopień realizacji wymogów stawianych przez algorytmy wymienionych wyszukiwarek.

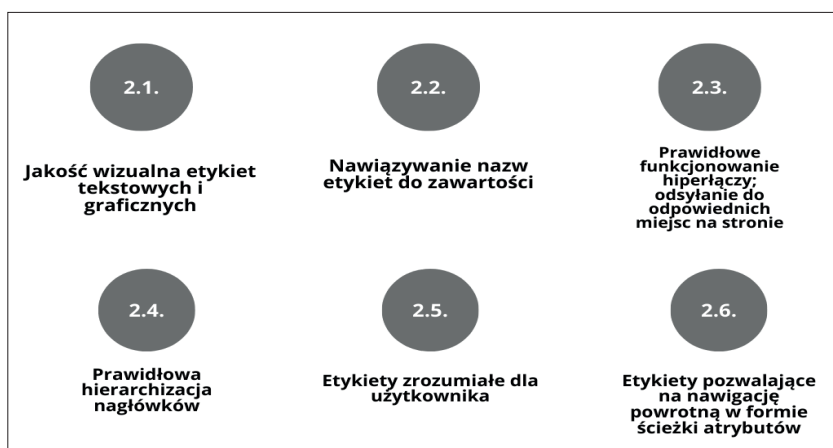
## Załącznik

### Katalog kryteriów oceny jakościowo-heurystycznej architektury informacji

Sposób przyznawania punktów opisano w sekcji: 3. Organizacja i metody badań.



Rys. 17. Kryteria oceny funkcjonalności systemów organizacyjnych



Rys. 18. Kryteria oceny funkcjonalności systemów etykietowania

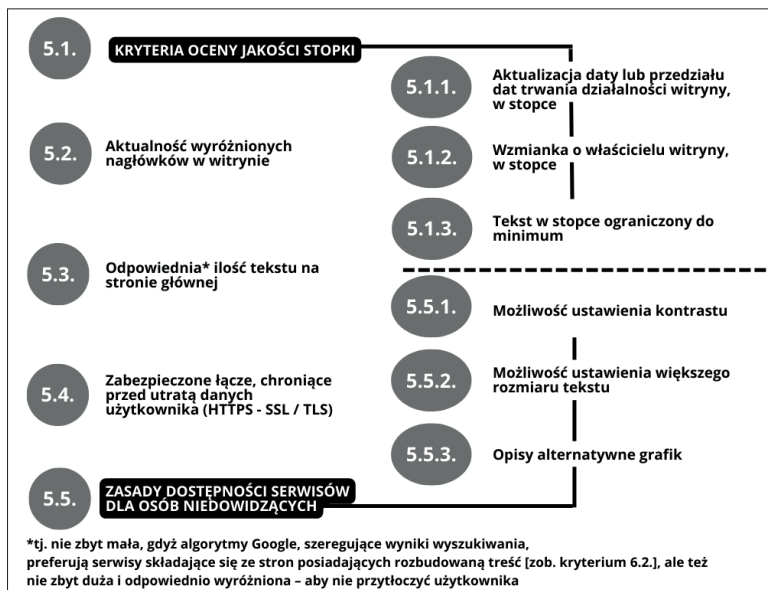


Rys. 19. Kryteria oceny funkcjonalności systemów nawigacyjnych



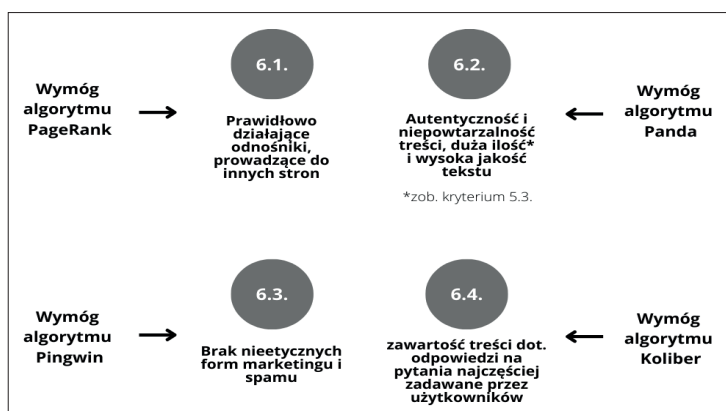
Rys. 20. Kryteria oceny funkcjonalności systemów wyszukiwawczych

Kryteria 4.1., 4.2., 4.9. i 4.11. są wyłącznie kryteriami wyjściowymi i nie podlegają punktacji, natomiast kolejne ich podpunkty, tj.: 4.1.1. i 4.1.2., 4.2.1. i 4.2.2., 4.9.1. i 4.9.2., 4.11.1. oraz 4.11.2. – są punktowane.



Rys. 21. Dodatkowe kryteria oceny

Kryteria 5.1. i 5.5. są wyłącznie kryteriami wyjściowymi i nie podlegają punktacji, natomiast kolejne ich podpunkty, tj.: 5.1.1., 5.1.2. i 5.1.3. oraz 5.5.1., 5.5.2. i 5.5.3. – są punktowane.



Rys. 22. Kryteria oceny stopnia spełniania wymogów narzuconych przez najważniejsze algorytmy Google w celu hierarchizacji wyników wyszukiwania

## Bibliografia

- Borowska, D., Rozworska, N. (2022). Wpływ internetu na sposób komunikowania polityków z elektoratem. *Acta Politica Polonica* 53(1), 75–87. <https://doi.org/10.18276/ap.2022.53.06>
- Cieślak-Florczyk, M. (2014). Analiza strukturalna stron internetowych głównych polskich partii politycznych. *Świat Idei i Polityki*, 13, 23–35.
- Dade-Robertson, M. (2011). *The Architecture of Information: Architecture, Interaction Design and the Patterning of Digital Information*. London: Routledge.
- Darling, J. (2008). *Latin America, Media, and Revolution. Communication in Modern Mesoamerica*, New York, NY: Palgrave Macmillan.
- Davis, N. (2010). Information Architecture, Black Holes and Discipline: On Developing a Framework for a Practice of Information Architecture. *Bulletin of the American Society for Information Science and Technology*, 36 (6), 25–29. <https://doi.org/10.1002/bult.2010.1720360608>
- Davis, N. (2018). *What is Information Architecture?* [online]. Methodbrain [18.08.2022], <https://methodbrain.com/2018/10/08/what-is-information-architecture/#information-architecture-science>
- Dijk Van, P. (2003). *Information Architecture for Designers. Structuring Websites for Business Success*. Mies: RotoVision SA.
- Drosik, A. (2010). Wykorzystanie Internetu jako podstawowej formy komunikacji w kampanii wyborczej. Studium kampanii do PE Sergiusza Najara (165–178). W: M. Jeliński, W. Peszyński, A. Seklecka (red.), *Wybory do Parlamentu Europejskiego. Media i marketing polityczny*. Toruń: Wydaw. „Dom Organizatora”.
- Gajowniczek, T. (2019). Internet w komunikowaniu politycznym. *Media – Kultura – Komunikacja Społeczna*, 2(15), 53–67. <https://doi.org/10.31648/mkks.5136>
- Garlicki, J., Mider, D. (2012). Wykorzystanie Internetu w komunikowaniu politycznym elit z wyborcami. *Studia Politologiczne*, 25, 177–209.
- Głowacka, E. (2016). Propozycja modelu badań funkcjonalności serwisów WWW bibliotek cyfrowych opartego na zasadach architektury informacji. *Biblioteka Nostra. Śląski Kwartalnik Naukowy*, (2), 23–32. <http://www.sbc.org.pl/dlibra/publication?id=267003&tab=3>
- Grupa iCEA (b.d.). *Alternatywne wyszukiwarki – na które warto zwrócić uwagę?* [online]. Grupa iCEA [2.08.2022], <https://www.grupa-icea.pl/alternatywne-wyszukiwarki-na-ktore-warto-zwrocic-uwage/>
- Haratyk, K. (2011). Open source = open democracy? Perspektywy przemian polityczności pod wpływem nowych mediów. W: K. Kopecka-Piech, A. Woźny (red.), *New Media (Studies)* (37–53). Wrocław: Wydaw. UWŕ.
- Hubert-Miller de, B.A. (2006). The IA of Potentiality: Toward a Grounded Theory of Information Architecture. Philosophy, Theory and Research. *Bulletin of the American Society for Information Science and Technology*; 32 (6), 10–12. <https://doi.org/10.1002/bult.2006.1720320605>
- Kurzak, T. (2022). *Yandex, czyli rosyjskie Google – czym jest, co oferuje i komu chce jeszcze zagrozić* [online]. Komputer Świat – Komputery, Testy Sprzętu, Newsy [2.08.2022], <https://www.komputerswiat.pl/artykuly/redakcyjne/historia-yandexu-czy-rosyjski-gigant-jest-za-duzy-zeby-upasc/1h3k8sk>

- Lange-Sadzińska, K. (2011). Architektura informacji w praktyce [online]. *Studies & Proceedings of Polish Association for Knowledge Management*, (53), 93–103 [27.05.2021], [http://www.pszw.edu.pl/images/publikacje/t053\\_pszw\\_2011\\_lange-sadzinska\\_-\\_architektura\\_informacji\\_w\\_praktyce.pdf](http://www.pszw.edu.pl/images/publikacje/t053_pszw_2011_lange-sadzinska_-_architektura_informacji_w_praktyce.pdf)
- Leszczuk-Fiedziukiewicz, A. (2011). Internet jako narzędzie kreowania wizerunku polityka. *Nowe Media*, 2 31–54. <https://doi.org/10.12775/NM.2011.002>
- Louw, E. (2005). *The Media and Political Process*. London: SAGE.
- Maarek, P.J. (2016). *Marketing político y comunicación. Claves para una buena información política*. Barcelona: Paidós.
- Muñoz Ramírez, G. (2008). *El Fuego y La Palabra: Una Historia del Movimiento Zapatista*. San Francisco, CA: City Lights Books.
- Neuman, J. (1996). *Lights, Camera, War: Is Media Technology Driving International Politics?*, Manhattan, NY: St Martin's Press.
- Oniszczyk, Z. (2018). Internet jako ważna przestrzeń komunikacji politycznej [online]. W: A. Czyż, S. Kubas (red.), *Na drodze do wyjaśnienia problemów politycznych: księga jubileuszowa ofiarowana profesorowi Markowi Barańskiemu (664–674)*. Katowice: Wydaw. UŚ [17.08.2022], [https://rebus.us.edu.pl/bitstream/20.500.12128/21557/1/Oniszczyk\\_Internet\\_jako\\_wazna\\_przestrzen\\_komunikacji\\_politycznej.pdf](https://rebus.us.edu.pl/bitstream/20.500.12128/21557/1/Oniszczyk_Internet_jako_wazna_przestrzen_komunikacji_politycznej.pdf)
- Osińska, V. (2014). Rola mechanizmów wizualizacyjnych w architekturze informacji. *Toruńskie Studia Bibliologiczne*, 6 (2), 81–96. <https://doi.org/10.12775/TSB.2013.023>
- Osiński, Z. (2014). Architektura informacji polskich internetowych serwisów edukacyjnych. W: B. Sosińska-Kalata, E. Chuchro (red.), *Nauka o informacji w okresie zmian (569–596)*. Warszawa: Wydaw. SBP.
- Ricci A. (2013). *The Early Political Web, 1995–2005: A Ten-year Observational Research Seeking Evidence of eDemocracy in the Information Architecture of Political Parties Web Sites Worldwide* [online]. Université Libre de Bruxelles, DI-fusion [17.08.2022], <https://difusion.ulb.ac.be/vufind/Record/ULB-DIPOT:oai:dipot.ulb.ac.be:2013/209496/Holdings>
- Rosenfeld, L., Morville, P., Arango, J. (2017). *Architektura informacji w serwisach internetowych i nie tylko. Wydanie IV*, Gliwice: Helion.
- Roszkowski M. (2019). Każda reprezentacja jest interpretacją – w stronę hermeneutycznej koncepcji architektury informacji. *Zagadnienia Informacji Naukowej – Studia Informacyjne*, 57, 2(114), 61–79. <https://doi.org/10.36702/zin.455>
- Seddighi, A.H. (2021). An Information Architecture Evaluation Framework for Websites. *Iranian Journal of Information Processing and Management*, 36 (4), 995–1030. <https://doi.org/10.52547/jipm.36.4.995>
- Skolnikoff, E.B. (1992). *The Elusive Transformation: Science, Technology, and the Evolution of International Politics*. Princeton, NJ: Princeton University Press.
- Skórka, S. (2011). Architekt informacji – kreator przestrzeni informacyjnych [online]. *Przegląd Biblioteczny*, 79(1), 47–61. <https://doi.org/10.36702/pb.42>
- Skórka, S. (2021). Nauka o architekturze informacji: koncepcja dyscypliny naukowej. *Zagadnienia Informacji Naukowej – Studia Informacyjne*, 59(1), 7–28. <https://doi.org/10.36702/zin.720>
- Smith, A.W. (2019). User Experience Design for Older Adults: Experience Architecture and Methodology for Users Aged 60+. *SIGDOC '19: Proceedings of the 37th*

- ACM International Conference on the Design of Communication*, 1–9. <https://doi.org/10.1145/3328020.3353952>.
- Sosińska-Kalata, B. (2013). Obszary badań współczesnej informatologii (nauki o informacji). *Zagadnienia Informatologii – Studia Informacyjne*, 51 (2), 9–41. <https://doi.org/10.36702/zin.600>
- Street, J. (2001). *Mass Media, Politics and Democracy*. New York, NY: Palgrave Macmillan.
- Tafiłowski, P. (2016). Architektura informacji jako problem badawczy informatologii, *Zagadnienia Informatologii – Studia Informacyjne*, 54(2), 37–47. <https://doi.org/10.36702/zin.302>
- Trend, D. (1997). *Cultural Democracy: Politics, Media, New Technology*. New York, NY: State University of New York Press.
- Turska-Kawa, A. (2011). *Poczucie alienacji a użytkowanie mediów. W poszukiwaniu nowych obszarów zastosowania teorii użytkowania i gratyfikacji*. Katowice: Wydaw. Nauk. „Śląsk”.
- Utracka, D. (2017). Humanistyczne aspekty architektury informacji. *Rekonosans. Zagadnienia Rodzajów Literackich*, (3), 171–189.
- Weare, C. (2002). The Internet and Democracy: The Causal Links Between Technology and Politics. *International Journal of Public Administration*, 25 (5), 659–691. <https://doi.org/10.1081/PAD-120003294>
- 

## Information Architecture on the Websites of the Law and Justice, and the Civic Platform

### Abstract

**Purpose/Thesis:** The main objective of the presented study is to assess the quality of information architecture on the websites of Law and Justice (PiS) and Civic Platform (PO), the dominant parties in Polish politics.

**Approach/Methods:** The research methodology comprises a qualitative-heuristic assessment of the information architecture on the websites of PiS and PO, expert assessment, a comparative analysis of selected websites and the author’s own observations.

**Results and conclusions:** Although both sites are of standard quality, they have been designed with the principles of functionality in mind. They lack some important components of information architecture. The website of PO is of marginally higher quality than the website of PiS.

**Practical implications:** The results of the conducted research may have an impact not only on the assessment of the parties’ technological advancement, but also on their image and public perception, and therefore also their effectiveness in reaching their voters.

**Originality/Cognitive value:** The study shows a new interpretation of a qualitative and heuristic assessment based on 74 criteria concerning the needs and competences of the user of a given website. This approach will advance the research into websites and thus improve the quality of user experience, allowing for a seamless implementation of information processes.

---

## Keywords

Civic Platform. Google algorithms. Information architecture. Labeling systems. Law and Justice. Navigation systems. Qualitative-heuristic analysis. Organization systems. Political parties. Search systems. Websites.

---

*PAULINA SAJNA-KOSOBUCKA jest studentką Interdyscyplinarnej Szkoły Doktorskiej Nauk Społecznych Academia Rerum Socialium w dyscyplinie nauk o komunikacji społecznej i mediach. Najważniejsze publikacje: Architektura informacji a synestezja – rozważania teoretyczne i przykłady zastosowania w praktyce (Toruńskie Studia Bibliologiczne 2020, 13 (2), 9–28); Data Journalism and Politics: Election DataBot, European Data Journalism Network and Media 3.0 Foundation (Świat Idei i Polityki, 2019, 18, 393–408); The Future of Media: Data Journalism and Synesthetic Communication (R. Sajna-Kunowsky, A. Garczewska, eds., Future of Media, Changing Journalism and New Communication (53–67). Bydgoszcz 2020).*

*Kontakt z autorką:*

*psk@doktorant.umk.pl*

*Interdyscyplinarna Szkoła Doktorska Nauk Społecznych – Academia Rerum Socialium*

*Uniwersytet Mikołaja Kopernika w Toruniu*

*Collegium Humanisticum*

*ul. Bojarskiego 1*

*87-100 Toruń*

# Analiza architektury informacji platformy Steam

Łukasz Tarnacki<sup>1</sup>

---

## Abstrakt

**Cel/Teza:** Celem artykułu jest analiza architektury informacji w platformie cyfrowej dystrybucji gier komputerowych Steam. Przybliżono zarówno cele strategiczne platformy Steam, jak i praktyczne zastosowanie różnych systemów organizacji informacji w samej usłudze.

**Koncepcja/Metody badań:** Analizy przeprowadzono na podstawie modelu architektury informacji autorstwa Rosenfelda, Morville'a i Arango, opublikowanego w książce *Architektura informacji w serwisach internetowych i nie tylko*. W artykule scharakteryzowano sposoby implementacji w usłudze Steam czterech elementów składowych architektury informacji. Dla każdego z nich sformułowano wnioski w odniesieniu do działania ogółu platformy. Część wniosków została także poparta materiałami i treściami od producentów platformy.

**Wyniki i wnioski:** Przeprowadzone badania pozwoliły na identyfikację i charakterystykę narzędzi organizacji informacji w usłudze Steam w perspektywie elementów składowych architektury informacji.

**Oryginalność/Wartość poznawcza:** Steam jest jedną z najpopularniejszych platform cyfrowej dystrybucji gier komputerowych. Stanowi ona jednak stosunkowo nowe i niezbadane środowisko z punktu widzenia architektury informacji. Biorąc pod uwagę sukces finansowy platformy, studia nad jej architekturą informacji mają wartość poznawczą w odniesieniu do tego gatunku usług informacyjnych.

**Ograniczenia badań:** Do głównych ograniczeń w przeprowadzonych badaniach należy ograniczony dostęp do oficjalnych informacji związanych z mechanizmami architektury informacji platformy Steam. Część wniosków jest zatem oparta na wynikach inspekcji interfejsu graficznego użytkowników oraz interakcji z narzędziami dostępu.

## Słowa kluczowe

Architektura informacji. Gry komputerowe. Steam.

Otrzymany: 16 sierpnia 2022. Zrecenzowany: 30 września 2022.

Poprawiony: 31 października 2022. Zaakceptowany: 30 grudnia 2022.

---

<sup>1</sup> Artykuł jest oparty na wynikach badań przeprowadzonych przez autora w ramach pracy dyplomowej pt. *Architektura informacji platformy dystrybucyjnej Steam jako wielofunkcyjnej usługi dla graczy i producentów gier*, przygotowanej pod opieką dr. hab. Piotra Tańkowskiego, prof. ucz. na studiach I stopnia na kierunku Architektura przestrzeni informacyjnych na Wydziale Dziennikarstwa, Informacji i Bibliologii Uniwersytetu Warszawskiego.

## 1. Wstęp

Gry komputerowe stanowią z jednej strony istotny element współczesnej kultury masowej, z drugiej – reprezentują dynamicznie rozwijającą się gałąź gospodarki. Dystrybucja gier komputerowych ma miejsce zarówno w formie tradycyjnej, jak i cyfrowej. Jednak to środowisko cyfrowe World Wide Web w postaci sklepów i usług sieciowych jest obecnie głównym kanałem rozpowszechniania informacji o grach komputerowych oraz samych gier jako produktów komercyjnych. Pojęcie cyfrowej dystrybucji gier komputerowych w niniejszym artykule rozumiane jest w sposób zaproponowany przez Chmielarza i Szumskiego (2019, 44):

Cyfrowa dystrybucja gier komputerowych jest tu traktowana jako metoda rozpowszechniania pomiędzy graczami produktów (odpowiednich plików instalacyjnych lub kodu niezbędnego do uruchomienia aplikacji za pośrednictwem Internetu) przez wydawcę (dystrybutora). Rozpowszechnianie odbywa się za pośrednictwem odpowiedniej platformy, która zapewnia wsparcie w procesie sprzedaży i posprzedażowej obsłudze klienta.

Cyfrowe platformy dystrybucji gier komputerowych są więc pewnym gatunkiem sieciowych usług informacyjnych, które realizują określone cele strategiczne, są skierowane do określonej grupy użytkowników, a ich zasadnicze kolekcje tworzą opisy produktów komercyjnych w postaci gier komputerowych. Jedną z najpopularniejszych platform tego typu jest Steam firmy Valve Corporation (USA).

Celem artykułu jest charakterystyka architektury informacji (AI) platformy dystrybucji gier komputerowych Steam (<https://store.steampowered.com/>). Przedmiotem badań jest aplikacja Steam przeznaczona na komputery stacjonarne. Badania przeprowadzono z wykorzystaniem analizy heurystycznej w oparciu o model architektury informacji Rosenfelda, Morville'a i Arango (2017) poprzez inspekcję interfejsu graficznego użytkownika. Przedmiotem analiz były podstawowe elementy składowe AI: systemy organizacji, systemy nawigacji, systemy etykietowania i systemy wyszukiwania. Głównym celem badania było zastosowanie koncepcji architektury informacji jako sposobu interpretacji organizacji informacji i dostępu do zasobów platformy Steam, która z jednej strony realizuje określone cele biznesowe, a z drugiej – jest oparta na zaangażowaniu społeczności w jej rozwój.

## 2. Platforma Steam

Historia platformy Steam sięga początków XXI w., a jej oficjalna premiera miała miejsce w 2003 r. Właścicielem Steam jest Valve – amerykański producent gier komputerowych. Steam jest platformą cyfrowej dystrybucji gier komputerowych. Początkowo była to usługa przeznaczona na komputery z systemem operacyjnym Windows. Obecnie Steam obsługuje również inne systemy operacyjne zarówno dla urządzeń stacjonarnych (np. Linux, macOS), jak i mobilnych (np. Android, iOS).

Usługa Steam dostępna jest również jako serwis internetowy. Strona internetowa Steam posiada funkcje oraz budowę interfejsu graficznego zbliżoną do aplikacji na urządzenia stacjonarne lecz ma ograniczone funkcjonalności w zakresie obsługi gier.

Podstawowe cele strategiczne platformy Steam to:

- (1) Udostępnianie i sprzedaż gier komputerowych oraz treści z nimi związanych.
- (2) Optymalizacja rozgrywki i zapewnianie wsparcia dla obsługiwanych gier.
- (3) Modyfikacja gier i treści oraz rozwój potrzebnych ku temu narzędzi.
- (4) Zapewnienie środków i przestrzeni do interakcji dla społeczności.

Steam nie jest platformą przeznaczoną jedynie dla gier sieciowych czy opartych wyłącznie na trybie wieloosobowym. W aplikacji dostępne jest przejście na wersję offline, co pozwala grać w każdą pobraną i w pełni zaktualizowaną grę, która nie wymaga podłączenia do sieci.

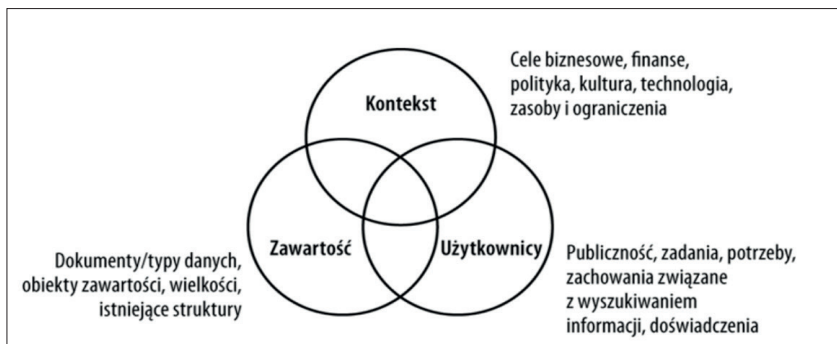
### 3. Główne założenia architektury informacji platformy Steam

Platforma Steam jest usługą sieciową dającą dostęp do dużej ilości treści i oferującą znaczny zakres funkcjonalności, co wymaga odpowiedniej organizacji zarówno jej kolekcji, jak i sposobów dostępu do jej zawartości. W badaniu tych zagadnień można wykorzystać m.in. koncepcję architektury informacji, która koncentruje się na projektowaniu cyfrowych środowisk informacyjnych w celu optymalizacji ich użyteczności i funkcjonalności. AI wykorzystywana jest przy tworzeniu serwisów internetowych i aplikacji mobilnych niezależnie od ich stopnia zaawansowania. Ramy AI są dosyć płynne, jednak główny zakres jej działalności można wyznaczyć, odwołując się do jednej z najważniejszych dla AI publikacji: szeroko wykorzystywanej w edukacji AI książki *Architektura informacji w serwisach internetowych i nie tylko* (Rosenfeld et al., 2017). Autorzy dla lepszego zrozumienia roli architektury informacji w środowiskach informacyjnych proponują jej model, na który składa się kontekst, zawartość oraz użytkownicy cyfrowych przestrzeni informacyjnych (Rys.1) (Rosenfeld et al., 2017, 37–44). Warto przyjrzeć się, jak te elementy można zinterpretować w kontekście platformy Steam.

Element kontekstu w odniesieniu do platformy Steam można interpretować na dwa sposoby. Po pierwsze, jej cele strategiczne są związane z realizacją założeń komercyjnych poprzez osiąganie przychodów z dystrybucji gier oraz budowę identyfikacji marki i zaufania wśród klientów. Po drugie, ma to być przestrzeń wzajemnej interakcji, zapewniająca dużą swobodę użytkownika w zgodzie z polityką organizacji. Innymi słowy, AI platformy Steam powinna łączyć funkcję sklepu internetowego oraz przestrzeni do komunikacji interpersonalnej i interakcji dla graczy.

Element zawartości odnosi się do specyfiki kolekcji platformy Steam. Mowa tu bowiem nie tylko o grach i programach dostępnych w sklepie, ale także o „dokumentach, aplikacjach, usługach, schematach i metadanych, z których użytkownicy

korzystają lub które odnajdują w systemach” (Rosenfeld et al, 2017, 42). AI platformy Steam ma za zadanie nie tylko informować gracza o produktach będących w ofercie z wykorzystaniem adekwatnych elementów metadanych (np. producent, data wydania czy wielkości pliku), ale także zapewniać szybki dostęp do podstron społeczności związanych z produktem (Rys. 2).



Rys. 1. Elementy architektury informacji [Rosenfeld, et al., 2017]



Rys. 2. Strona biblioteki Steam na przykładzie gry CS:GO

Interpretacja użytkownika jako trzeciego elementu modelu AI wiąże się z identyfikacją grupy użytkowników docelowych, czyli w tym przypadku graczy, oraz charakterystyką ich potrzeb i zachowań informacyjnych. Valve skupiło się przede wszystkim na zaspokojeniu dwóch podstawowych potrzeb użytkowników: zakupu gier oraz umożliwieniu rozgrywki za pośrednictwem platformy Steam. Jest

to szczególnie widoczne w sposobie prezentacji informacji, gdzie podczas przeglądania produktów produkty w sklepie czy bibliotece Steam jedynym wyraźnie przyciągającym wzrok elementem jest przycisk „Graj” lub „Dodaj do koszyka” (Rys. 2). Jaskrawy zielony kolor kontrastuje tu z pozostałymi elementami interfejsu, utrzymanymi w ciemniejszych barwach.

#### 4. Charakterystyka systemów architektury informacji na platformie Steam

Przed przejściem do szczegółowego omówienia elementów AI, warto nakreślić ogólny schemat struktury platformy Steam. W publikacji *The Information Architecture of Behavior Change Websites* (Danaher et al., 2005) wyodrębniono cztery modele projektowania przestrzeni informacyjnych odnoszące się do sposobu ich eksploracji przez użytkowników. Są to:

- projekt macierzowy (ang. *matrix design*) – nieustrukturyzowany, oparty na hipertekście i licznej nawigacji kontekstowej, pozwalający na dużą swobodę wyboru sposobów eksploracji;
- projekt hierarchiczny (ang. *hierarchical design*) – odgórnie zorganizowany system ścieżek eksploracji ze stroną początkową jako punktem wyjścia;
- projekt tunelowy (ang. *tunnel design*) – zamknięty i zadaniowy, prowadzący użytkownika krok po kroku do celu;
- projekt hybrydowy (ang. *hybrid design*) – łączący cechy powyższych modeli.

Stwierdzenie Danahera i in. (2005), że treść mająca na celu zmianę zachowania musi być prezentowana w sposób atrakcyjny i użyteczny, by mogła mieć korzystny wpływ na decyzje użytkownika odnosi się nie tylko do zagadnień behawioralnych, ale jest ważna także z perspektywy biznesowej. Wykorzystanie dobrych praktyk opracowanych dzięki AI bezpośrednio przekłada się więc na efektywność projektowanych usług. Platforma Steam zbudowana jest według modelu hybrydowego, umożliwiającego realizację różnorodnych celów i użytkowników. O ile klasyczny schemat hierarchii góra-dół jest obecny we wszystkich sekcjach platformy, o tyle w niektórych miejscach wzbogacony jest on o inne modele. W dziale „Sklep” wykorzystano w znacznej mierze projektowanie macierzowe, oferując wiele opcji przeglądania i przemieszczania się pomiędzy produktami oraz ich pochodnymi. Sama procedura zakupu utrzymana jest natomiast w konwencji tunelu, aby nie dać użytkownikowi pretekstu ani przestrzeni do rozproszenia się lub zmiany zdania.

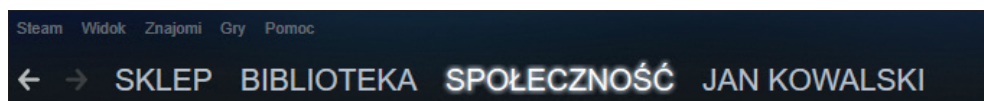
W badaniu AI platformy Steam przeanalizowano zarówno elementy serwisu widoczne w interfejsie graficznym użytkownika, jak i procesy zachodzące wewnątrz usługi. AI w ujęciu Rosenfelda, Morville’a i Arango to cztery komponenty składowe architektury informacji, będące tutaj punktem wyjścia do szczegółowej analizy (Rosenfeld et al., 2017, 86):

- (1) systemy organizacji – odpowiedzialne za grupowanie i klasyfikowanie informacji;
- (2) systemy etykietowania – dbające o odpowiednie nazewnictwo i reprezentację treści;
- (3) systemy nawigowania – określające sposób poruszania się po serwisie;
- (4) systemy wyszukiwania – optymalizujące efektywność wyszukiwarek zawartych w serwisach.

#### 4.1. Systemy organizacji

Steam to platforma wielozadaniowa z zawartością o bardzo zróżnicowanym charakterze. W samym sklepie znajduje się ponad 59 tys. gier wideo (stan na lipiec 2022) (Galyonkin, 2022). Dla zapewnienia dobrego poziomu funkcjonalności aplikacji wymagane jest więc zastosowanie zaawansowanego systemu organizacji. Systemy te mają na celu wyodrębnienie i pogrupowanie dostępnych na platformie treści, aby zapewnić do nich możliwie jak najlepszy dostęp. Jest to o tyle ważne, że sposób organizacji determinuje w znacznym stopniu efektywność systemów wyszukiwania i nawigowania po serwisie.

Wśród schematów organizacyjnych wykorzystywanych na platformie przeważają rozwiązania o charakterze zadaniowym i bazującym na metaforach (Rosenfeld et al., 2017, 97–105). Jest to widoczne w sposobie organizacji nawigacji globalnej, gdzie wyraźnie wydzielone są cztery przestrzenie spełniające różne funkcje. *Sklep* i *Biblioteka* odpowiadają za komercyjne cele usługi, natomiast *Społeczność* oraz profil użytkownika (reprezentowany za pomocą nazwy profilu, np. *Jan Kowalski*) to opcje skierowane na indywidualne i grupowe aktywności graczy (Rys. 3). Cyfrowe odpowiedniki fizycznych miejsc i przestrzeni prywatnej stanowią, dzięki swojej uniwersalności, zrozumiałe metafory dla użytkowników.



Rys. 3. Organizacja zawartości w nawigacji globalnej Steam

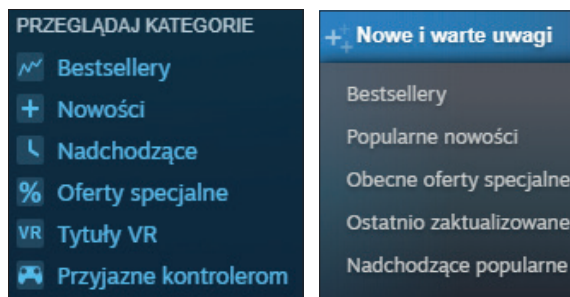
Na rysunku 3 widoczna jest również nawigacja narzędziowa. Pomimo odmiennego przeznaczenia, ten sposób organizacji dostępu cechuje podobna konstrukcja. Mowa tu o zastosowaniu płytkiego schematu hierarchicznego (tylko jeden stopień głębokości) pozwalającego na organizację i dostęp do zawartości platformy. Nie brak także elementów świadczących o polihierarchii wykorzystanej w tym schemacie – do sekcji *Biblioteka* prowadzą aż trzy różne miejsca w systemie organizacji. Są to rozwiązania bardzo podobne do tych stosowanych także w sklepie Steam (zob. sekcja 4.3).

## 4.2. Systemy etykietowania

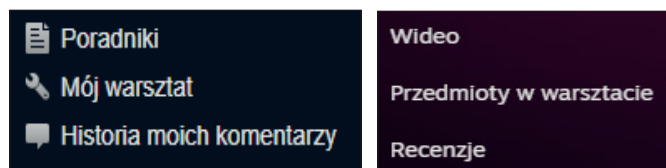
Posegregowane już treści wymagają czytelnej i zrozumiałej reprezentacji. Za to właśnie odpowiedzialny jest system etykietowania, czyli nazywania tego, co zgrupowano w danych obszarach przestrzeni informacyjnej. Etykiety na platformie Steam przyjmują dwojaką formę – tekstu oraz ikon, pełnią także różne funkcje w zależności od sposobu ich wykorzystania (Rosenfeld et al., 2017, 124–135). W przypadku Steam warto rozpatrzyć system etykietowania w systemach nawigacji oraz w systemie tagowania wykorzystywanego do organizacji informacji.

Treści platformy w znacznej większości opisane są przez etykiety tekstowe. Ikony wykorzystano zazwyczaj jako ich uzupełnienie – np. znak plusa przy opcji „dodaj grę” lub też w formie samodzielnej reprezentacji powszechnie znanych rozwiązań – koło zębate jako ustawienia czy też koperta jako sekcja z wiadomościami.

Etykiety nawigacyjne pozwalają użytkownikowi na określenie swojego miejsca w danej przestrzeni informacyjnej. Rosenfeld et al. (2017, 131) piszą, że „skutecznie stosowane etykiety są integralną częścią budowania poczucia znajomości [danego serwisu]”, stąd ważne jest, aby były one utrzymane w jednolitym stylu (językowym i wizualnym) oraz na podobnym poziomie szczegółowości. Dobrze zaprojektowane etykiety nie powinny też zmieniać swojej formy w zależności od miejsca, w którym znajduje się użytkownik. Jest to niestety jedna z wad platformy, w której elementy nawigacji często są reprezentowane niespójnie. Np. w różnych miejscach sklepu natrafić można na „oferty specjalne” i „obecne oferty specjalne”, prowadzące ostatecznie do tej samej podstrony (Rys. 4. i 5.).



Rys. 4. Niespójne etykiety w różnych miejscach usługi Steam

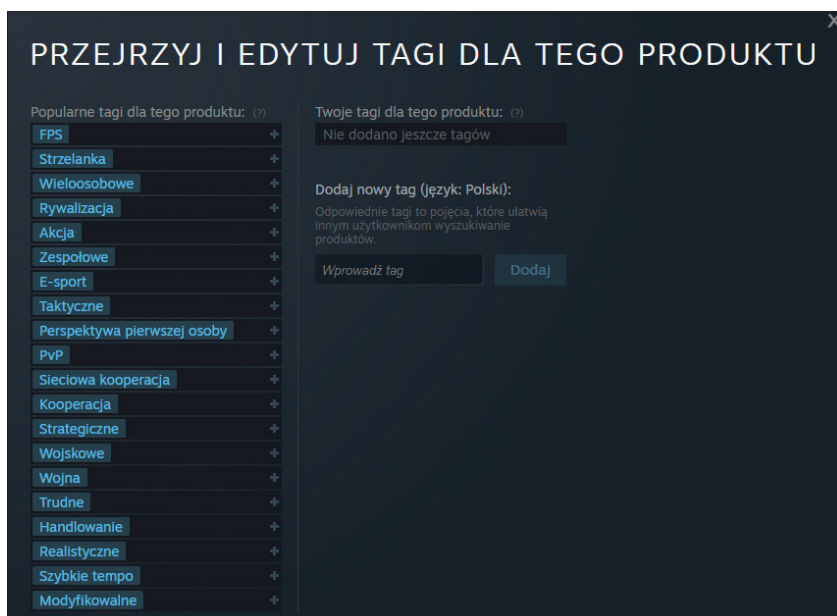


Rys. 5. Niespójna etykieta (warsztat) w różnych miejscach usługi Steam

#### 4.2.1. Systemy tagowania

Podczas projektowania sklepowego systemu etykietowania w Steam, projektanci Valve skorzystali z trzech źródeł nazewnictwa: własnej analizy zawartości, sugestii autorów oraz wyborów użytkowników (zob. Rosenfeld et al., 2017, 138–152). Warto zaznaczyć, że w obliczu globalnego zasięgu platformy Steam jej systemy etykietowania są unikalne dla każdej wersji językowej usługi. Oznacza to, że tagi przypisane do danej gry w wersji polskiej oraz angielskiej mogą się różnić (Valve Corporation, 2022).

System *Tagi Steam* ma charakter dwuetapowy (Valve Corporation, 2022). Pierwszy z nich to zaangażowanie producentów gier do ich opisu. Przed opublikowaniem swojej gry na platformie Steam muszą oni zaproponować adekwatne ich zdaniem tagi. Odbywa się to za pomocą przeznaczonego dla autorów produkcji kreatora tagów – narzędzia, które zawiera setki tagów zatwierdzonych przez moderatorów Steam. Kreator ten pozwala na wybranie nielimitowanej liczby tagów dla danej gry. Tagi następnie pozycjonowane są według wagi nadawanej im przez producentów. W interfejsie graficznym Steam widoczne jest dwadzieścia tagów o największej wadze, jednak tylko pięć ma kluczowy wpływ na pozycjonowanie gry w sklepie. W odpowiednim pozycjonowaniu tagów pomaga m.in. automatyczny system sugerowania priorytetów oraz możliwość przeglądania innych podobnie otagowanych tytułów.



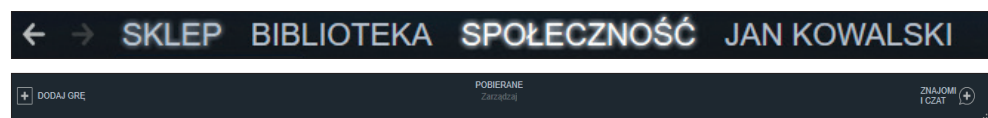
Rys. 6. Aktualnie wyświetlane Tagi Steam dla gry CS:GO

Gry opisane wstępnie przez producentów publikowane są w sklepie Steam. Wtedy to rozpoczyna się drugi etap procesu tagowania, tym razem realizowany przez użytkowników platformy. Odbywa się on na zasadzie połączenia metod swobodnej listy (samodzielne opisywanie elementu) z sortowaniem kart (dopasowywanie istniejących tagów do elementów) (zob. Rosenfeld et al., 2017, 146–149). W ramach każdego sklepowego produktu gracze zatwierdzają zaproponowane uprzednio tagi lub dodają własne, ich zdaniem bardziej trafne (Rys. 6). Oczywiście postrzeganie gry z perspektywy autora oraz gracza może znacząco od siebie odbiegać, stąd też popularne tagi stosowane przez społeczność mogą być odmienne od tych ustalonych pierwotnie, a co za tym idzie, zająć ich miejsce w sklepowej witrynie. Valve (2022) zaznacza jednak, że nowy tag „może nie być widoczny dla pozostałych użytkowników, dopóki wystarczająca liczba osób nie użyje go do otagowania tego produktu”. Ma to na celu zniwelowanie znaczenia nietrafionych opisów lub tagów-zartów, a także zachowanie opisu producenta.

### 4.3. Systemy nawigacji

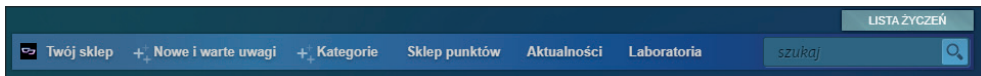
Mechanizmy nawigacyjne determinują sposób poruszania się użytkownika po serwisie. Rosenfeld, Morville i Arango (2017, 157) określają funkcje nawigacji jako „tworzenie poczucia miejsca za pomocą języka i udostępnianie czytelnych ścieżek poznawania środowiska”. Wyodrębnia się trzy główne systemy nawigacji: globalną, lokalną i kontekstową (Rosenfeld et al., 2017, 154–168). Steam jako usługa rozbudowana i wielofunkcyjna implementuje wszystkie typy systemów nawigacji.

Nawigacja globalna jest nieodłącznym systemem każdego serwisu sieciowego. Najczęściej przybiera ona formę modułu z kategoriami tematycznymi, który pozostaje niezmienny niezależnie od miejsca serwisu, w którym przebywa użytkownik (Rosenfeld et al., 2017, 160–161). Podobnie rzecz ma się w przypadku platformy Steam. Idąc za platformowym konceptem usługi w usłudze, czyli sklepu będącego integralną częścią aplikacji, można tu mówić o zastosowaniu podwójnej nawigacji globalnej.



Rys. 7. Nawigacja globalna platformy Steam

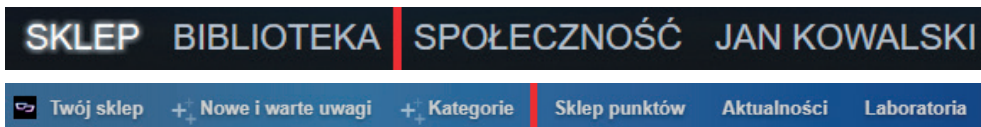
Pierwszą z nich jest nawigacja globalna Steam jako usługi (Rys. 7). Występuje ona w formie górnego paska menu z rozwijanymi opcjami oraz stopki wyposażonej w trzy elementy funkcjonalne. Warto zauważyć, że Steam nie zawiera jednej strony głównej, posiada natomiast strzałki „wstecz” i „dalej” wzorem przeglądarek internetowych. Miejsce użytkownika w aplikacji określone jest za pomocą podświetlenia nazwy wybranej sekcji (w tym przypadku „Społeczności”).



Rys. 8. Nawigacja globalna sklepu Steam

Druga nawigacja globalna odnosi się do sklepu Steam zintegrowanego z aplikacją (Rys. 8). Globalny układ tej sekcji w pewnym stopniu przypomina inne popularne sklepy internetowe. Na elementy podobne składają się tu rozwijane menu, w tym polihierarchiczne megamenu z zakładki „Kategorie”, a także moduł wyszukiwania. Specyficzny charakter nawigacji ujawnia się po przeanalizowaniu etykiet proponowanych sekcji. Tylko połowa z nich bezpośrednio dotyczy treści prezentowanych w sklepie. „Sklep punktów” (system nagród Steama), „Aktualności” oraz „Laboratoria” to odnośniki do osobnych miejsc w usłudze, które nie są bezpośrednio związane z procesami kupna i sprzedaży gier, a mimo to znalazły się na głównym szczeblu sklepowej nawigacji. Jest to efektem ingerencji użytkowników w procesy projektowe.

Warto zauważyć pewną symetryczność platformy Steam w kontekście nawigacyjnym i organizacyjnym. W tym przypadku warto wrócić do tezy o podwójnej zadaniowości AI Steama, nastawionej na jednoczesne generowanie przychodów oraz pełnienie funkcji społecznościowych. Zostało to zrealizowane poprzez rozdzielanie funkcjonalności i obszarów zawartości. Jest to widoczne już w nawigacji globalnej. Na rysunku 9 czerwona linia odgradza funkcjonalności o charakterze komercyjnym od społecznościowych.

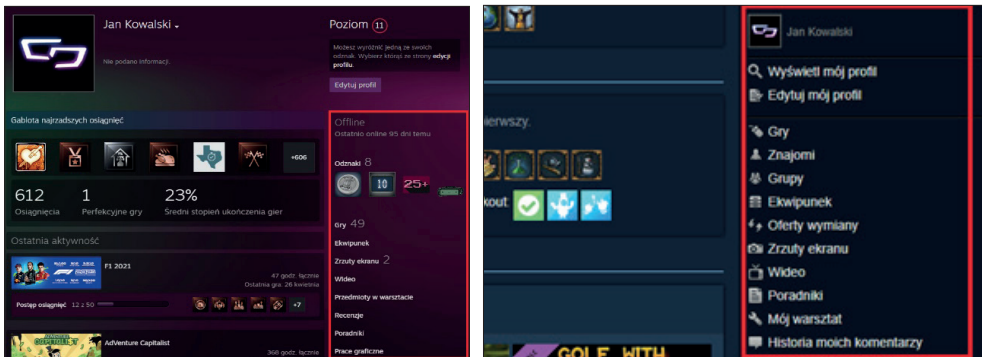


Rys. 9. Organizacja wybranych elementów nawigacji w serwisie Steam

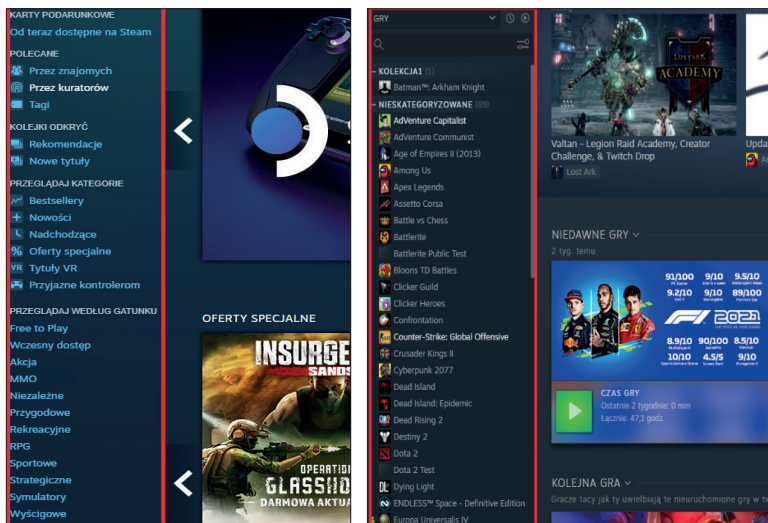
Prawidłowość ta przechodzi w zmienionej formie także na szczebel nawigacji lokalnej, która odpowiada za poruszanie się po obszarach serwisu. O ile prawie każda z nich różni się pełnionymi przez siebie funkcjami, o tyle punktem wspólnym pozostaje tu przemyślane położenie względem zawartości sekcji. Dla przykładu, systemy nawigacji lokalnej w zakładkach aktywności znajomych oraz profilu znajdują się po prawej stronie, natomiast sklep oraz biblioteka posiadają je po lewej (Rys. 10. i 11.).

Odnieść się tu można do praktyki zwanej tworzeniem miejsc (ang. *place-making*), omówionej w książce *Pervasive Information Architecture* (Resmini & Rosati, 2011). Jej autorzy odróżniają w cyfrowych środowiskach pojęcia miejsca (ang. *place*), skonkretyzowanego i powiązanego z użytkownikiem oraz przestrzeni (ang. *space*) pojmowanej jako ogół serwisu. *Place-making* ma za zadanie kreować miejsca o w pełni zrozumiałym kontekście, które pozwolą użytkownikom na uzyskanie komfortu i odpowiedniej orientacji względem reszty usługi. Istotną rolę odgrywają w tym hiperłącza, które stanowią narzędzia pozwalające poruszać się po tej

cyfrowej przestrzeni. Hiperłącza, operując relacjami semantycznymi, „plotą” sieć powiązań i przejść, tworząc zrozumiałe dla danych użytkowników miejsca. Remini i Rosati (2011, 73–77) twierdzą, że *place-making* jest tak samo potrzebne jak właściwa nawigacja i sposób odnajdywania drogi, aby sprawić, żeby każdy projekt mógł stanowić „dom” dla swoich użytkowników. Zarówno w mediach społecznościowych, jak i na platformie Steam przykładem tworzenia miejsc może być profil użytkownika. Popularny schemat wykorzystujący duże zdjęcie oraz przywołujący najważniejsze informacje na górze strony daje poczucie indywidualnego charakteru takiego miejsca, natomiast zmniejszenie rozmiaru fotografii oraz sprowadzenie elementów jej towarzyszących do nawigacji lokalnej świadczy o tym, że użytkownik znajduje się już w sekcji o bardziej publicznym charakterze (Rys. 10).

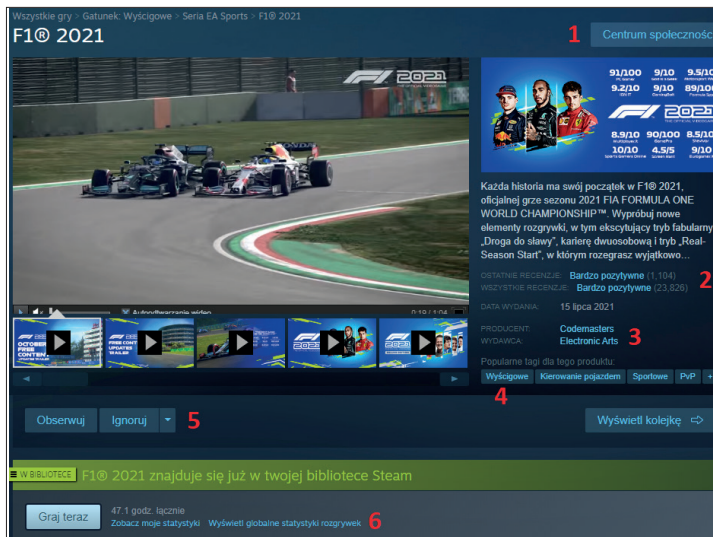


Rys. 10. Nawigacja lokalna sekcjach społecznościowych



Rys. 11. Nawigacja lokalna sekcji komercyjnej

Do elementów zawartości występującej zazwyczaj w centralnej części witryn odnosi się nawigacja kontekstowa. Jest to nawigacja o największym stopniu szczegółowości. Działa ona na zasadzie skojarzeń niejako obudowując konkretną zawartość elementami w mniejszym lub większym stopniu z nią związanymi (zob. Rosenfeld et al., 2017, 164–166). Steam wykorzystuje obszerną nawigację kontekstową, która najbardziej widoczna jest w części sklepowej. Każdy produkt zawiera metadane oraz dodatkowe elementy opisowe (Rys. 12). Niektóre występują w formie hiperłączy prowadzących do innych miejsc w usłudze związanych z produktem. Pozwala to np. z pozycji karty produktu przenieść się do obszaru społeczności związanej z grą (1), zapoznać się z jej recenzjami (2) czy przejść do profilu producenta i wydawcy gry (3). Widać też wspomniane już tagi (4) i opcje związane z funkcją tzw. kolejki odkryć (5). W przypadku posiadania danej gry w bibliotece pojawia się również możliwość dostępu do wybranych statystyk związanych z użytkowaniem gry przez społeczność (6).



Rys. 12. Witryna produktu w sklepie Steam na przykładzie gry F1 2021

Szczególnie istotną rolę na platformie Steam odgrywają zagadnienia personalizacji i dostosowywania treści. Według Rosenfelda et al. (2017, 177) „zarówno personalizacja, jak i dostosowywanie mogą być wykorzystywane do udoskonalania lub uzupełniania istniejących systemów nawigacji”. Personalizacja, czyli selekcjonowanie treści w oparciu o zachowania i potrzeby użytkownika przejawia w postaci funkcji systemów rekomendacyjnych. Jest to szczególnie widoczne w przypadku szczegółowych kategorii produktów prezentowanych w sklepie. Generowane kategorie odnoszą się bezpośrednio do preferencji użytkownika na podstawie jego poprzednich interakcji (Rys. 13).

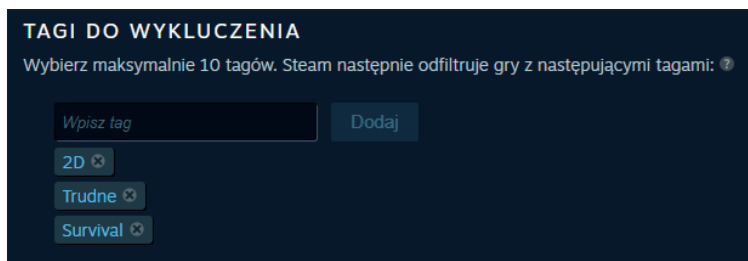


Rys. 13. Wybrane kategorie generowane w sklepie Steam.

Źródło: Valve Corporation, aplikacja Steam, 2022

Dostosowywanie odnosi się do możliwości zarządzania przez użytkownika daną zawartością według własnych preferencji. W przypadku platformy Steam jest ono wykorzystywane np. w systemie tagowania poprzez możliwość organizacji zawartości według popularności wśród użytkowników. Główny nacisk położony jest tutaj jednak na systemy nawigacyjne i indywidualny dobór wyświetlanych treści. Dla dostosowywania treści w sklepie powstała funkcja listy życzeń. Gracze mogą dodawać do niej wszystkie interesujące i nieposiadane jeszcze produkty.

Dla użytkowników chcących głębiej ingerować w proponowaną im zawartość stworzony został rozbudowany system określania preferencji. Obejmuje on niemal wszystkie treści dostępne w sklepie i sekcji społeczności, z których najciekawsza jest funkcja wykluczania tagów, odfiltrująca ze sklepu gry o niepożądanych przez użytkownika atrybutach (Rys. 14).



Rys. 14. System wykluczania treści za pomocą tagów

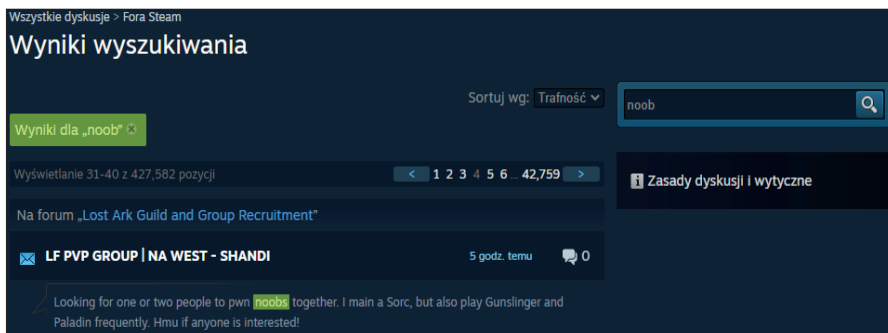
Jednym z celów dostosowywania preferencji jest też eliminacja toksycznych zachowań przejawianych przez innych graczy. Z tego względu dostosowanie sekcji społecznościowych opiera się głównie na eliminacji wyrażen wulgarnych oraz wprowadzających w błąd. Ważnym czynnikiem dla całokształtu wyświetlanej zawartości jest też opcja zmiany języka usługi. Steam pozwala na jednoczesną prezentację treści zawartych we wszystkich wybranych wersjach językowych platformy.

#### 4.4. Systemy wyszukiwania

Systemy wyszukiwania stanowią czwarty element architektury informacji. Ich celem jest zapewnienie użytkownikowi możliwości efektywnego odnalezienia wybranego fragmentu zawartości z wykorzystaniem wyszukiwarki internetowej (zob. Rosenfeld et al., 2017, 183–188). Proces wyszukiwania odbywa się poprzez interakcję z obecną w serwisie wyszukiwarką.

Na platformie Steam wyodrębniono kilka osobnych stref wyszukiwania, tworzących zgodnie z definicją Rosenfelda et al. (2017, 189) „podzbiory środowiska informacyjnego, które zostały zaindeksowane oddzielnie od reszty zawartości”. Stosowane na platformie strefy wyszukiwania odnoszą się do głównych obszarów zawartości. Każdy z nich (m.in. sklep, biblioteka, społeczność) zawiera przynajmniej jeden mechanizm pozwalający na przeszukiwanie jego zawartości. Narzędzia te różnią się od siebie z uwagi na odmienny charakter informacji obecny w danym obszarze.

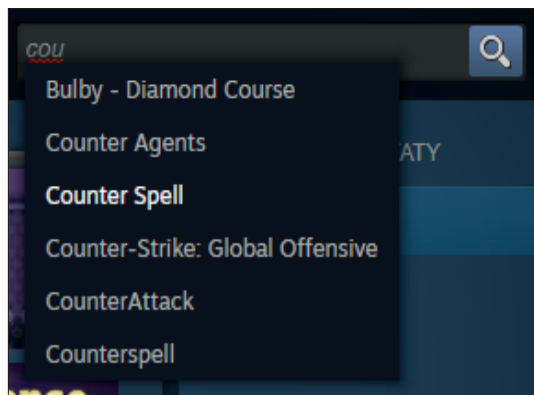
System wyszukiwania w sekcji zawierającej fora społecznościowe funkcjonujące na platformie Steam pozwala na odnalezienie wątków w dyskusjach, w których wystąpiło dane wyrażenie użyte przez użytkownika w zapytaniu (Rys. 15.). Nieco bardziej złożone jest wyszukiwanie treści sekcji społecznościowej *Warsztat*, gdzie znajduje się zawartość stworzona przez użytkowników. Odbywa się ono dwuetapowo – należy zacząć od wyboru określonego tytułu gry, a następnie sformułować zapytanie dotyczące poszukiwanych modyfikacji z nim związanych (Rys. 16). Pozwala to na uniknięcie przeciążenia informacyjnego, gdyż cały warsztat zawiera tysiące projektów.



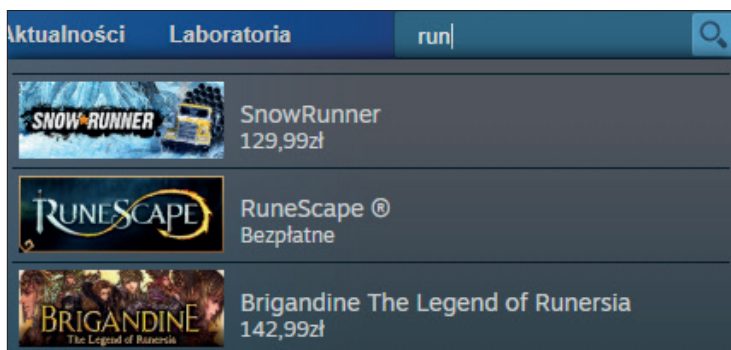
Rys. 15. System wyszukiwania dyskusji

System wyszukiwania cechujący się największym stopniem zaawansowania jest wykorzystywany w sekcji sklepu. Dzięki funkcji autosugerowania, po wpisaniu danej frazy mechanizm wyszukiwania proponuje graczom gry zawierające ją w swoim tytule. Rezultaty wyszukiwania są prezentowane według trafności, ale

z eksponowaniem gier cieszących się aktualnie dużym zainteresowaniem wśród użytkowników. Z tej pozycji bezpośrednio przejść można do witryny wybranej gry (Rys. 17).



Rys. 16. System wyszukiwania w warsztacie Steam

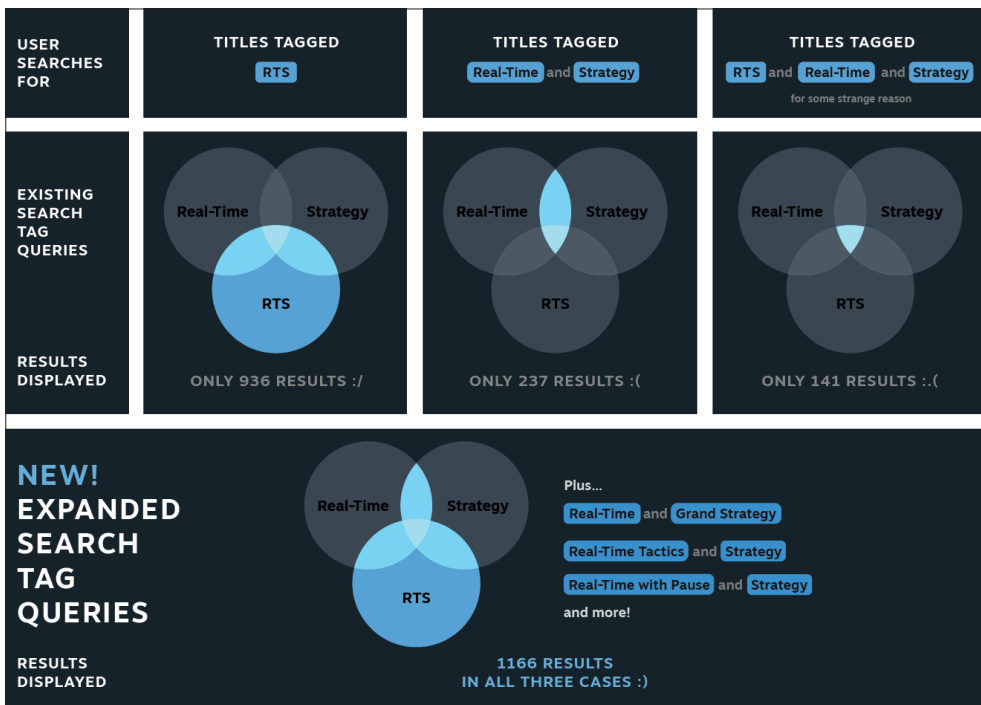


Rys. 17. System autosugestii dla frazy „run” w sklepie Steam

W celu zawężenia i sprecyzowania rezultatów wyszukiwania w wyszukiwarce zainstalowanej na platformie Steam zaimplementowano szereg filtrów. Filtry sklepowe odnoszą się np. do wybranych funkcji platformy Steam, stosowanego systemu operacyjnego, preferowanej liczby graczy, typu gry, czy tagów zastosowanych w opisie. W przypadku systemu tagów mechanizm wyszukiwawczy Steam stosuje sieć relacji semantycznych występujących między tagami, pozwalając na lepszą efektywność wyszukiwania informacji.

We współpracy z graczami wprowadzony został nowy system rozszerzania zapytań wyszukiwarki. Opiera się on na manualnie wprowadzonych mapach relacji pomiędzy tagami. Według Valve (2020) „wliczają się w to tagi rozbijające się na inne (np. FPS – Perspektywa pierwszej osoby + Strzelanka) oraz tagi konstruowane

z innych (np. Strategiczne + RPG – Strategiczne RPG)”. Zastosowanie oznaczeń relacji synonimicznych między wybranymi tagami Valve uzasadnia ich użytecznością w procesie wyszukiwania. Zdecydowanie większy nacisk postawiono tu na równoznaczność tagów, odsuwając na bok ich asocjacje, np. „[tag] fantasy nie sugeruje tagu magia” (Valve Corporation, 2020), mimo iż często występują one wspólnie przy danych produktach. Wprowadzone zmiany poskutkowały znacznie bardziej trafnymi wynikami wyszukiwania. System interpretuje zależności pomiędzy zakresami znaczeniowymi tagów, budując dla nich wspólne pola semantyczne (Rys. 18).



Rys. 18. Sposób działania systemu wyszukiwania sklepu Steam

## 5. Zakończenie

Steam to jedna z pierwszych platform o globalnym zasięgu zajmująca się dystrybucją gier wideo. Z tego względu musiała ona samodzielnie wytyczać ścieżki dla swojego rozwoju, w tym także budować strategię architektury informacji. Duży udział miała w tym społeczność graczy, którzy często dzielili się wrażeniami na temat nowych rozwiązań. Steam jest doskonałym przykładem wieloaspektowego postrzegania organizacji informacji przez pryzmat architektury informacji. Można

sądzić, że dzięki wysoce funkcjonalnej AI platforma jest w stanie nieustannie zwiększać liczbę swoich użytkowników, tym samym maksymalizując zyski.

Steam wyodrębnia w swojej usłudze miejsca realizujące cele komercyjne i społecznościowe. Przejawem tego jest choćby dwojaki podział nawigacji globalnej i lokalnej czy też utworzenie unikalnych stref wyszukiwania dla osobnych sekcji usługi. Przy projektowaniu elementów architektury informacji Valve nie tylko posiłkuje się opinią graczy, ale też w niektórych obszarach daje im wolną rękę.

## Bibliografia

- Chalk, A. (2021). *Only one of Epic's 'first wave' of exclusives made money, and it wasn't Metro* [online]. PC Gamer, [20.12.2022], <https://www.pcgamer.com/only-one-of-epics-first-wave-of-exclusives-made-money-and-it-wasnt-metro/>
- Chmielarz, W., Szumski, O. (2019). Analiza porównawcza serwisów dystrybutorów gier komputerowych. *Przegląd Organizacji*, 43–47. <https://doi.org/10.33141/po.2019.02.06>
- Clemnt, J. (2021). *Number of games available in the Epic Games Store from 2019 to 2021* [online]. Statista, [20.12.2022], <https://www.statista.com/statistics/1234037/epic-games-store-games-available/>
- Danaher, B. G., McKay, H. G., Seeley, J. R. (2005). The Information Architecture of Behavior Change Websites. *Journal of Medical Internet Research* [online], 7(2), <https://doi.org/10.2196/jmir.7.2.e12>
- Galyonkin, S. (2022). *Monthly summaries* [online]. Steamspy, [20.12.2022], <https://steamspy.com/year/>
- Resmini, A., Rosati, L. (2011). *Pervasive Information Architecture: Designing Cross-Channel User Experiences*. Burlington: Morgan Kaufmann.
- Rosenfeld, L., Morville, P., Arango, J. (2017). *Architektura informacji w serwisach internetowych i nie tylko*. Gliwice: Helion.
- Valve Corporation. (2020). *Eksperyment laboratoriów Steam 004.1: rozszerzenie zapytania wyszukiwarki* [online]. Steam, [20.12.2022], <https://steamcommunity.com/games/593110/announcements/detail/1792935884902997963>
- Valve Corporation. (2022). *Przeglądaj Steam w nowy sposób* [online], Steam, [20.12.2022], <https://store.steampowered.com/tag/?l=polish>
- Wańtuchowicz, P. (2019). Epic Games ma bardzo złą opinię wśród graczy. Przepaść w stosunku do Steama [online]. Eurogamer, [20–12-2022], <https://www.eurogamer.pl/epic-games-ma-bardzo-zla-opinie-wsrod-graczy-przepasc-w-stosunku-do-steama>

## Analysis of the information architecture of the Steam platform

### Abstract

**Purpose/Thesis:** The purpose of this article is to present and analyze the information architecture used in the digital distribution service Steam. It discusses both the strategy informing the design of the Steam and the practical application of the service's systems.

**Approach/Methods:** A significant part of the analysis was carried out on the basis of the book entitled Information architecture. For the web and beyond. Four main systems identified in the book are considered. For each of them, conclusions were drawn with regard to the operation of the entire platform. The analysis was supported by content provided by the platform's producers.

**Results and conclusions:** The conducted research allowed the author to identify and characterize the methods for organizing information employed by the Steam platform in the light of the basic elements of information architecture.

**Originality/Value:** Steam is one of the most popular digital video-game distribution platforms. Nevertheless, it is a relatively new and underexplored field in the context of information architecture studies. The platform's enormous financial success warrants research into its information architecture as it yields insight into the operations of this type of information services.

**Research limitations:** The main challenge encountered during research was the limited access to official information concerning the mechanics of Steam's information architecture. Therefore, the article partly relies on the analysis of the platform's interface as accessed by its users.

### Keywords

Information architecture. Steam. Video games.

---

*ŁUKASZ TARNACKI jest absolwentem studiów I stopnia na kierunku Architektura przestrzeni informacyjnych na Uniwersytecie Warszawskim. Obecnie jest studentem studiów II stopnia na kierunku Logistyka i administrowanie w mediach na Uniwersytecie Warszawskim. Niniejszy artykuł napisany został na podstawie pracy licencjackiej pod tytułem Architektura informacji platformy dystrybucyjnej Steam jako wielofunkcyjnej usługi dla graczy i producentów gier.*

*Kontakt z autorem:*

*tarnacki.lukasz@gmail.com*

*Wydział Dziennikarstwa, Informacji i Bibliologii  
Uniwersytet Warszawski*

**Monika Krakowska (2022).** *Zachowania informacyjne człowieka w kontekście zjawiska epistemicznej bańki informacyjnej. Propozycja nowej koncepcji.* Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, ss. 336. ISBN: 978-83-233-5134-4

Nigel Ford w znanej książce *Introduction to Information Behaviour* zauważył, że zachowania informacyjne to aktywność, która przenika całe nasze życie (Ford, 2015). Bezustannie przetwarzamy informację, która stanowi materiał wyjściowy, surowiec do tworzenia naszej wiedzy. Jeśli ten surowiec jest trafnie wybrany i efektywnie przetworzony, może stać się fundamentem ludzkiej mądrości. Informacja, którą pozyskujemy, podejmując różnorakie zachowania informacyjne jest naszą codzienną „intelektualną strawą”, od której jakości i umiejętności jej przetworzenia zależy jakość i efektywność naszej intelektualnej kondycji. Jak pisze Ford – naszego „intelektualnego zdrowia” (Ford, 2015, 1). Warto przypomnieć te słowa, zwracając uwagę na najnowszą publikację Moniki Krakowskiej, badaczki z Instytutu Studiów Informacyjnych na Uniwersytecie Jagiellońskim, która od lat zajmuje się studioowaniem ludzkich zachowań związanych z poszukiwaniem i wykorzystywaniem informacji.

Analiza i modelowanie zachowań informacyjnych człowieka od ponad 30 lat stanowią jeden z głównych nurtów badań w nauce o informacji (por. np. Bates, 2010; Saracevic, 2010). Nawiązując do trzech głównych pytań badawczych nauki o informacji, które – jak przyjmuje się za Marcją Bates – wyznaczają intelektualną strukturę tej dyscypliny<sup>1</sup>, głównym celem badania zachowań informacyjnych ludzi jest wyjaśnienie, jak wchodzą oni w powiązania ze światem informacji, jak jej szukają i jak ją wykorzystują. Na rozwój tego nurtu decydujący wpływ miało upowszechnienie technologii informacyjnej i komunikacyjnej, a wraz z nim – różnorodnych serwisów informacyjnych adresowanych do użytkowników nieprofesjonalnych, wymagających dostosowania tych usług do ich potrzeb, oczekiwań, nawyków i umiejętności w zakresie poszukiwania i/lub korzystania z różnego typu treści. Badania prowadzone na różnych grupach odbiorców informacji, w różnych środowiskach społecznych i technologicznych oraz w różnych dziedzinach wiedzy i życia społecznego ujawniły nie tylko duże zróżnicowanie przebiegu interakcji między człowiekiem i informacją, działań podejmowanych przez ludzi w celu pozyskania potrzebnej im informacji oraz potrzeb i nawyków komunikacyjnych, ale też liczne i różnorodne czynniki kształtujące zachowania informacyjne człowieka i decydujące zarówno o wyborze składających się na nie działań, jak i o tym, jak poszukujący informacji oceniają te, które uzyskują w wyniku tych poszukiwań.

Rozwój cyfrowych technologii sieciowych, w szczególności rozwój mediów społecznościowych i inteligentnych technologii filtrowania informacji, który nastąpił w ostatnich

---

<sup>1</sup> “Three Big Questions can be identified within the above framework: (1) the physical question: What are the features and laws of the recorded-information universe? (2) The social question: How do people relate to, seek, and use information? (3) The design question: How can access to recorded information be made most rapid and effective?” (Bates, 1999, 1048).

kilkunastu latach, zwrócił uwagę badaczy na problemy związane z wieloma negatywnymi zjawiskami informacyjnymi, takimi jak: przeciążenie informacyjne, tworzenie się baniek informacyjnych czy negatywne praktyki informacyjne, np. manipulacja informacją, stosowanie tzw. mrocznych wzorców (ang. *black patterns*) w projektowaniu usług i aplikacji cyfrowych, celowe rozesyłanie informacji nieprawdziwej, kradzież tożsamości i inne działania nieetyczne czy cyberprzestępstwa. Problemy te wyraźnie ujawniły z jednej strony niewystarczające kompetencje współczesnych odbiorców informacji zarówno w zakresie oceny wiarygodności otrzymywanych treści, jak i radzenia sobie z ich nadmiarem oraz ze złymi praktykami innych uczestników cyfrowej komunikacji, a z drugiej strony – negatywne konsekwencje rozwiązań tych problemów oferowanych przez technologię cyfrową (np. algorytmy personalizacji miały zapewnić ograniczenie nadmiaru informacji, a stały się mechanizmem generowania baniek informacyjnych czy filtrujących). Najnowsza książka Moniki Krakowskiej wpisuje się w tę bardzo obecnie aktualną problematykę, prezentując dotychczasowe ustalenia badaczy zachowań informacyjnych oraz oferując nową koncepcję badań coraz bardziej złożonych problemów dotyczących interakcji współczesnego człowieka z wielkimi i różnorodnymi zasobami informacyjnymi w technologicznie nasyconym środowisku informacyjnym.

Na strukturę książki składa się pięć obszernych rozdziałów, które poprzedza kilkunastostronicowy wstęp, a zamyka zakończenie podsumowujące całość rozważań oraz streszczenie w języku angielskim i starannie opracowany zestaw spisów pomocniczych, w tym bogata bibliografia wykorzystanej literatury (obejmująca ponad 600 pozycji) i indeks nazwisk badaczy, których pomysły i prace są omówione w książce. Treści każdego rozdziału zostały uporządkowane za pomocą śródtytułów drugiego i trzeciego stopnia, czytelnie wyodrębniających kolejno omawiane przez Autorkę badania, teorie, problemy i pomysły.

We wstępie Autorka napisała, że:

Zasadniczym celem rozważań przedstawionych w tej książce jest zaproponowanie nowego konceptu – epistemicznej bańki informacyjnej – stworzonego na potrzeby teoretycznych i empirycznych dociekań w obszarze zachowań informacyjnych. Koncepcja ta, charakteryzująca określone zjawisko w kontekście aktywności informacyjnych, poprzez wnikliwą analizę różnorodnych modeli, teorii i założeń ugruntowanych w interdyscyplinarnym ujęciu nauki o informacji, powinna otworzyć nową i szerszą perspektywę dla eksplorowania zachowań informacyjnych. Ma być pewnego rodzaju metateoretycznym spoiwem łączącym różnorodne podejścia, m.in. społeczne, afektywne, kognitywne, a także stanowić wielowymiarowy element empirycznie rozwijanych badań dotyczących reakcji i interakcji człowieka z informacją (s. 16–17).

Tak sformułowanemu głównemu celowi książki została podporządkowana jej konstrukcja. W czterech pierwszych rozdziałach są omówione wybrane problemy, pojęcia i teorie, służące naświetleniu różnych aspektów kształtowania się i funkcjonowania specyficznych, zamkniętych lub częściowo zamkniętych struktur przestrzeni informacyjnej człowieka, które Autorka objęła pojęciem epistemicznej bańki informacyjnej. Każdy rozdział otwiera wprowadzenie zawierające wyjaśnienie kryteriów wyboru zagadnień w nim omówionych, a zamyka podsumowanie, w syntetycznej formie eksponujące najważniejsze koncepcje, tezy i spostrzeżenia poczynione przez innych badaczy, które tworzą fundament własnej koncepcji teoretycznej Autorki. Koncepcja ta została szczegółowo przedstawiona i przedyskutowana w rozdziale piątym.

We wstępie do książki czytelnik znajdzie klarowne wprowadzenie do podjętej w niej problematyki. Autorka wyjaśniła tu główny przedmiot swoich rozważań, tj. pojęcie epistemicznej bańki informacyjnej, osadzając je w kontekście zarówno takich pojęć jak bańka filtrująca, bańka informacyjna czy komory echa (pogłosowe), które funkcjonują już od pewnego czasu w piśmiennictwie innych dyscyplin (np. medioznawstwa, nauk politycznych czy socjologii kultury), jak i w kontekście pojęć związanych z dotychczasową agendą badawczą obszaru zachowań informacyjnych, w szczególności ewolucyjnego ujęcia zachowań informacyjnych zaproponowanego przez Amandę Spink (2010). Monika Krakowska podkreśliła też ważny aspekt swoich badań, tj. ich interdyscyplinarność czy multidyscyplinarność: w poszukiwaniu wyjaśnień i uzasadnień swojej koncepcji badania specyficznych zachowań informacyjnych w warunkach ograniczania dostępu do informacji sięgała bowiem po teorie i ustalenia wielu innych nauk, od socjologii i psychologii, przez kognitywistykę i neurologię, aż do nauk biologicznych. Wysoki stopień złożoności problemów związanych z kształtowaniem się zachowań informacyjnych człowieka i różnorodnością czynników je determinujących w sposób oczywisty uzasadnia wybór takiej postawy badawczej. Jako główną metodę zastosowaną w badaniach, na podstawie których powstała książka, Autorka wskazała krytyczną analizę treści interdyscyplinarnego piśmiennictwa oraz metodę przeglądu i krytyki literatury z wybranego zakresu (ang. *scoping review*), wywodzącą się z nauk medycznych i wykorzystaną do problemowej syntezy zgromadzonego materiału. Zwięzłe wskazanie najważniejszych publikacji wśród bogatego piśmiennictwa badawczego – zagranicznego i polskiego – poświęconego problematyce użytkowników i użytkowania informacji oraz zachowań informacyjnych, posłużyło Autorce do zwrócenia uwagi na brak publikacji poddających kompleksowej analizie, wyjaśnieniu i naukowej refleksji najnowsze zjawiska związane z coraz powszechniejszym funkcjonowaniem użytkowników mediów cyfrowych w specyficznej infosferze zamkniętych lub częściowo zamkniętych przestrzeni informacyjnych, generowanych przez algorytmy komputerowe. Dotychczasowe badania zachowań informacyjnych ludzi dostarczyły natomiast licznych inspiracji i pomysłów do rozwiązania różnych cząstkowych problemów, objętych holistyczną koncepcją teoretyczną epistemicznej bańki informacyjnej. Swoim analizom Autorka poddała piśmiennictwo opublikowane w językach angielskim i polskim, które wyodrębniła na podstawie starannego przeszukania międzynarodowej dziedziny bazy bibliograficznej (Library and Information Science and Technology Abstracts firmy EBSCO), przeglądów badań w zakresie nauki o informacji publikowanych w *Annual Review of Information Science & Technology*<sup>2</sup> oraz przeglądów zawartości najważniejszych międzynarodowych i polskich czasopism specjalizujących się w problematyce nauki o informacji. Ze względu na interdyscyplinarny charakter problematyki stanowiącej obszar zainteresowań Autorki, relewantnego piśmiennictwa poszukiwała ona również w bazach bibliograficznych z zakresu nauk pokrewnych, tj.: psychologii, nauk społecznych, nauk o życiu oraz w bazach multidyscyplinarnych. Zasięg chronologiczny zgromadzonej literatury zasadniczo obejmuje dwie ostatnie dekady (okres od 2001 r. do 2021 r.), jakkolwiek słusznie skupiając uwagę przede wszystkim na najnowszej literaturze, na podstawie zawartych w niej cytowań Autorka włączyła do swoich analiz także pewną liczbę

<sup>2</sup> Rocznic ten był publikowany w latach 1966–2011, a po dziesięcioletniej przerwie jego wydawanie zostało wznowione w 2021 r.

wcześniejszych publikacji o szczególnym znaczeniu dla kształtowania podstaw teoretycznych badań nad zachowaniami informacyjnymi człowieka. Szczegółowa i systematyczna analiza zgromadzonego piśmiennictwa zapewniła jego pełne i wnikliwe wykorzystanie.

Pierwszy rozdział książki, zatytułowany *Zachowania informacyjne człowieka – wybrane konteksty*, jest poświęcony omówieniu pojęcia zachowania informacyjnego oraz cech i wymiarów ludzkiej aktywności, której ono dotyczy. Ta podstawowa kategoria badawcza odniesiona została do innych pojęć z nią powiązanych, takich jak praktyki informacyjne, doświadczanie informacji, kultura informacyjna oraz do pojęć bardziej szczegółowych, wskazujących różne typy zachowań informacyjnych lub formy działań składających się na zachowania informacyjne (pozyskiwanie informacji, poszukiwanie informacji, przypadkowe i nieintencjonalne pozyskiwanie informacji, wyszukiwanie informacji, wykorzystywanie informacji). Rozdział ten bardzo dobrze zbiera i systematyzuje wiedzę o różnorodnych formach ludzkiej aktywności, które w badaniach informatologicznych obejmuje się uogólnionym pojęciem zachowań informacyjnych. Autorka zamknęła rozdział prezentacją wybranych kategoryzacji zachowań informacyjnych zaproponowanych przez innych badaczy oraz własnego schematu stanowiącego próbę uporządkowania tego istotnie dość chaotycznego zestawu pojęć i terminów.

Rozdział drugi nosi tytuł *Interdyscyplinarne perspektywy zachowań informacyjnych człowieka* i poświęcony jest omówieniu badań dotyczących zachowań informacyjnych pogrupowanych według czterech perspektyw ich analizy i eksplanacji: ewolucyjnej, poznawczej, afektywnej i społecznej. Te różne ujęcia pozwoliły Autorce ukazać wielowymiarowość badań nad ludzkimi zachowaniami informacyjnymi, różnorodność czynników na nie oddziałujących i kontekstów, w których są osadzone. Każde z tych ujęć wiąże się z wykorzystaniem przez badaczy koncepcji i teorii wywodzących się z innych nauk: biologii (perspektywa ewolucyjna), psychologii (perspektywy kognitywna i afektywna) i socjologii (perspektywa społeczna). Badania przedstawione w tej części książki dotyczą wyjaśniania wielu kategorii badawczych wprowadzonych w celu konceptualizacji czynników wewnętrznych, związanych ze specyfiką psychiki podmiotów zachowań informacyjnych i kształtujących ich postawy, motywacje i działania poznawcze. Należą do nich takie kategorie jak: potrzeby informacyjne, intencje informacyjne, odkrywanie informacji czy nadawanie sensu. Rozdział ten demonstruje i porządkuje kierunki coraz bardziej pogłębionych analiz różnorodnych uwarunkowań interakcji człowieka z informacją.

Rozdział trzeci został zatytułowany *Kontekst oraz przestrzeń informacyjna człowieka*. Poświęcony jest badaniom skupionym na analizie czynników zewnętrznych oddziałujących na kształtowanie zachowań informacyjnych, a zatem czynników charakterystycznych dla szeroko rozumianego środowiska, w którym te zachowania są podejmowane. Rozpoznanie takich czynników zewnętrznych, które tworzą kontekst podejmowania przez ludzi interakcji z informacją, odgrywa kluczową rolę w badaniu specyfiki zachowań informacyjnych. Autorka przedstawiła i omówiła teorie związane z określaniem różnych rodzajów kontekstu zachowań informacyjnych (kontekstu wyznaczonego przez czas, środowisko, otoczenie społeczne i technologiczne) i różnych modeli przestrzeni informacyjnej oraz uporządkowała wprowadzone w nich pojęcia (środowisko informacyjne, informacyjne światy, horyzonty informacyjne, krajobrazy informacyjne, pola informacyjne, małe światy, osobista czy indywidualna przestrzeń informacyjna, tymczasowe miejsca informacyjne).

W umieszczonym w podsumowaniu rozdziału schemacie Autorka zestawiała te pojęcia, ukazując zachodzące między nimi powiązania.

Czwarty rozdział książki nosi tytuł *Bańki: filtrująca, epistemiczna i informacyjna oraz komory echa – charakterystyka zjawisk*, a jego celem jest wprowadzenie pojęć odnoszących się do zamkniętych przestrzeni informacyjnych i krytyczna analiza teorii, w których zostały ukształtowane. Kolejne podrozdziały zawierają zatem omówienie pojęć bańki filtrującej, bańki epistemicznej, wysp poznawczych, komór echa, bańki informacyjnej. Podjęte zostały tu także problemy związane z funkcjonowaniem zamkniętych przestrzeni informacyjnych, takie jak: domykanie przestrzeni informacyjnych, manipulowanie informacją, gatekeeping w zamkniętych przestrzeniach informacyjnych. Warto wspomnieć, że poznanie mechanizmów powstawania i funkcjonowania tego typu zamkniętych przestrzeni informacyjnych w ostatnich latach budzi wielkie zainteresowanie środowiska badawczego różnych nauk, stając się doskonałym przykładem interdyscyplinarnego wysiłku podjętego w celu zmierzenia się z trudnym i poważnym problemem współczesnego świata. O wadze zagrożeń wynikających z wykorzystywania tego rodzaju zjawisk w celu dezinformacji i polaryzacji opinii społecznej w dobitny sposób przekonały nas zarówno doświadczenia związane z politycznymi kampaniami wyborczymi (np. przebieg kampanii w wyborach prezydenckich w Stanach Zjednoczonych i Francji czy referendum w Wielkiej Brytanii dotyczące Brexitu), jak i obserwacja aktywności ruchów antyszczepionkowych w czasie pandemii COVID-19 i ich wpływu na postawy społeczeństw. Z niepokojem obserwujemy obecnie próby wykorzystania tych mechanizmów w sterowaniu nastrojami społecznymi w krajach zachodnich wobec wojny w Ukrainie i wspierania uchodźców. W takich okolicznościach media społecznościowe i profilowanie ich użytkowników bez wątpienia stały się nowym rodzajem efektywnej i trudnej do zwalczenia broni, poznanie mechanizmów której staje się pilną potrzebą. Jako ciekawe przykłady prób wyjaśniania takich mechanizmów warto tu wskazać niedawno opublikowane prace fizyków dotyczące komór echa, których działanie przeanalizowane zostało w kategoriach wielowarstwowych przestrzeni topologicznych i zjawiska polaryzacji (Bauman et al., 2020; Gajewski et al., 2022).

Ostatni rozdział książki Moniki Krakowskiej, zatytułowany *Epistemiczna bańka informacyjna i jej interdyscyplinarne determinanty w kontekście zachowań informacyjnych* zawiera szczegółową i systematyczną prezentację koncepcji epistemicznej bańki informacyjnej, stanowiącej konstrukt teoretyczny, który Autorka zaproponowała jako narzędzie analizy zamkniętych lub częściowo zamkniętych przestrzeni informacyjnych i dokonujących się w nich różnorodnych procesów interakcji człowieka i informacji, a także interakcji całych ludzkich społeczności i świata informacji. Koncepcja ta oparta jest na syntezie i integracji elementów wcześniejszych teorii, obserwacji i ustaleń, które Autorka zebrała, połączyła i rozszerzyła tak, aby uzyskać ogólny, holistyczny i wewnętrznie spójny model opisu i analizy całkowicie lub częściowo zamkniętego środowiska informacyjnego coraz częściej generowanego przez narzędzia cyfrowe i coraz bardziej charakterystycznego dla współczesnego człowieka i współczesnych społeczeństw intensywnie korzystających z komunikacji za pośrednictwem mediów cyfrowych i rzadko świadomych działania czynników modulujących zakres i przebieg tej komunikacji. Pojęcie epistemicznej bańki informacyjnej bezpośrednio nawiązuje do pojęć bańki informacyjnej czy bańki filtrującej, jest jednak od nich szersze, odnosi się bowiem do wszelkich zamkniętych lub niedomkniętych

przestrzeni informacyjnych, które ograniczają dostęp do informacji i w których funkcjonuje jednostka ludzka lub pewna ludzka zbiorowość. Epistemiczna bańka informacyjna jest metaforycznym, wielowymiarowym opisem takich przestrzeni, traktowanych jako specyficzna sytuacja, okoliczności i uwarunkowania oddziałujące na podejmowane w nich zachowania informacyjne. Może być ona rozpatrywana w trzech podstawowych wymiarach: w wymiarze jednostkowym jako przestrzeń osobista, w wymiarze jednostka – grupa, jako przestrzeń, w której dana jednostka funkcjonuje jako członek pewnej grupy czy wspólnoty zainteresowań, oraz w wymiarze społecznym, jako przestrzeń, w której podejmowane są zachowania informacyjne w pewnym społeczeństwie, determinowane przez jego kulturę, normy i przekonania. Autorka przedstawiła zastosowanie swojej koncepcji epistemicznej bańki informacyjnej do opisu i wyjaśniania zachowań informacyjnych, odwołując się do wybranych modeli i teorii, wskazując ewolucyjne, poznawcze, afektywne i społeczne czynniki oddziałujące na ich tworzenie się i kształtowanie. Ten wielowymiarowy opis, niezwykle bogaty pod względem wskazanych w nim aspektów i elementów zachowań informacyjnych i ich uwarunkowań, przekonuje o dużym potencjale eksplanacyjnym zaproponowanej koncepcji do precyzyjnego wyjaśniania właściwości i procesów rządzących zamkniętymi lub częściowo zamkniętymi przestrzeniami informacyjnymi zarówno generowanymi przez media cyfrowe jak i powstającymi w środowisku fizycznym. Warto pamiętać, że zamkniętość czy hermetyzm przestrzeni informacyjnych, w których funkcjonują ludzie, kształtując swój obraz świata oraz stosowanie w nich mechanizmów nadzorczych i manipulacyjnych może mieć miejsce i często ma miejsce także w świecie analogowym.

Traktując przedstawioną w książce koncepcję epistemicznej bańki informacyjnej jako „prolegomena do nowego, bardziej komplementarnego i zharmonizowanego rozpatrywania obszaru zachowań informacyjnych człowieka” (s. 277), w zakończeniu książki Autorka wskazała listę siedmiu grup zagadnień, które można poddać holistycznej i wieloaspektowej analizie wykorzystując ten model. Zaproponowała m.in., aby użyć tej koncepcji do zbadania relacji między zachowaniami informacyjnymi jednostek funkcjonujących w subiektywnych bańkach informacyjnych oraz wzajemnego oddziaływania różnych epistemicznych zamkniętych przestrzeni informacyjnych w kontekście procesów informacyjnych. Zdaniem Autorki model mógłby także pozwolić na bardziej wszechstronną eksplorację sposobów rozszczelnienia zamkniętych przestrzeni informacyjnych, ich niwelowania czy zapobiegania ich powstawaniu. Takie badania bez wątpienia miałyby nie tylko wartość poznawczą, ale też wielką wartość praktyczną, bo może przybliżyłyby nas do rozwiązania coraz bardziej niepokojących problemów związanych z wykorzystywaniem mediów cyfrowych do dezinformacji i polaryzacji społeczeństw.

Książka Moniki Krakowskiej jest ciekawym i pod wieloma względami nowatorskim opracowaniem aktualnej i ważnej problematyki, która w literaturze polskiej reprezentowana jest skromnie. Trzeba też zaznaczyć, że mimo dużej liczby badań dotyczących zachowań informacyjnych, które w ostatnich dekadach prowadzono na świecie, złożoność badanych zjawisk i procesów oraz ich zmienność w czasie, zwłaszcza pod wpływem rozwoju technologii informacyjnej, powoduje, iż nadal wiele kwestii nie doczekało się jeszcze wyjaśnienia. Badania zachowań informacyjnych najczęściej mają charakter fragmentaryczny, stosunkowo niewiele jest natomiast koncepcji holistycznych, które umożliwiają ujęcie interakcji człowieka z informacją w szerokiej perspektywie różnych jej form i różnych czynników na

nią oddziałujących. Taką koncepcję Monika Krakowska zaproponowała w swojej książce, co czyni tę książkę szczególnie cenną.

Interdyscyplinarne piśmiennictwo badawcze zachowań informacyjnych, w którym operuje się różnymi systemami pojęciowymi i różnorodną terminologią, stanowi materiał trudny do analizowania i systematyzowania wyodrębnionych z niego koncepcji i teorii. Autorka recenzowanej książki wielokrotnie zwracała w niej uwagę na potrzebę uporządkowania pojęć i terminów w tym obszarze, to trudne zadanie wyznaczyła sobie też jako jeden z dodatkowych celów swojego badania. Niewątpliwie wysiłki te przyniosły dobry rezultat, jednak złożoność analizowanej i omawianej w książce materii powoduje, że dążenie do precyzyjnego formułowania myśli niekiedy skutkuje budowaniem trudnych w odbiorze złożonych zdań. W konsekwencji książka wymaga uważnej i wnikliwej lektury, z pewnością jednak warto ten wysiłek podjąć.

## Bibliografia

- Bates, M. (1999). The Invisible Substrate of Information Science. *Journal of the American Society for Information Science*, 50 (12), 1043–1050.
- Bates, M. (2010). Information Behavior. In: M.J. Bates, M.N. Maack (eds.). *Encyclopedia of Library and Information Sciences* (2381–2391). Boca Raton: CRC Press.
- Baumann, F., Lorenz-Spreen, P., Sokolov, I.M., Starnini, M. (2020). Modeling Echo Chambers and Polarization Dynamics in Social Networks, *Physical Review Letters*, 124 (4), 048301. <https://doi.org/10.1103/PhysRevLett.124.048301>
- Ford, N. (2015). *Introduction to Information Behaviour*. London: Facet Publishing.
- Gajewski, Ł.G., Sienkiewicz, J., Hołyst, J.A. (2022). Transitions between polarization and radicalization in a temporal bilayer echo-chamber model. *Physical Review E*, 105 (2), 024125. <https://doi.org/10.1103/PhysRevE.105.024125>
- Saracevic, T. (2010). Information Science. In: M. J. Bates, M.N. Maack (eds.). *Encyclopedia of Library and Information Sciences* (2570–2585). Boca Raton: CRC Press.
- Spink, A. (2010). *Information Behavior. An Evolutionary Instinct*. Berlin, New York: Springer.

Barbara Sosińska-Kalata

ORCID: 0000-0002-4511-4701

Katedra Informatologii

Wydział Dziennikarstwa, Informacji i Bibliologii

Uniwersytet Warszawski

Nadesłano: 28 listopada 2022.

## 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 dobieranych 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 zgłaszać za pośrednictwem platformy OJS: [www.ojs.sbp.pl/index.php/zin](http://www.ojs.sbp.pl/index.php/zin)

### 1. Zasady ogólne

#### 1.1. Format

Wszystkie pliki (tekst artykułu, materiały ilustracyjne) należy przesyłać 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, identyfikator ORCID)
- 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). Opisy publikacji powinny być sporządzone zgodnie z zasadami APA Style 6th.

### 1.5. *Abstrakt ustrukturyzowany*

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

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

### 1.6. *Słowa kluczowe*

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

### 1.7. *Oświadczenie o oryginalności tekstu*

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

## 2. *Zasady opracowania artykułu*

### 2.1. *Organizacja i podział tekstu*

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

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

#### 1. **Pierwszy poziom podziału**

##### 1.1. **Drugi poziom podziału**

###### 1.1.1 **Trzeci poziom podziału**

### 2.2. *Przypisy*

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

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

### 2.3. *Pisownia tytułów w tekście artykułu*

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

### 2.4. *Wyróżnienia w tekście*

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

## 2.5. Materiały ilustracyjne i ich oznaczanie w tekście

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

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

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

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

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

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

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

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

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

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

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

### 2.6.1. Przykłady redagowania opisów bibliograficznych

#### KSIĄŻKA

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

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

#### PRACA ZBIOROWA

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

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

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

#### ARTYKUŁ W CZASOPIŚMIE

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

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

## ARTYKUŁ W PRACY ZBIOROWEJ

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

## ARTYKUŁ W CZASOPISIMIE ELEKTRONICZNYM

- Berners-Lee, T., Hendler, J., Lassila, O. (2001). The Semantic Web. *Scientific American* [online], May, [30.06.2013], <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>
- Bartalesi, V., Meghini, C. (2016). Using an Ontology for Representing the Knowledge on Literary Texts: The Dante Alighieri Case Study. *Semantic Web* [online], 8(3), 385–394. <http://doi.org/10.3233/SW-150198>
- Miller, H. (2013). Big-Data in Cloud Computing: A Taxonomy of Risks. *Information Research* [online], 18(1), [15.07.2013], <http://informationr.net/ir/18-1/paper571.html>

## HASŁA ENCYKLOPEDYCZNE

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

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

## DOKUMENT Z WITRYNY INSTYTUCJI, ORGANIZACJI LUB OSOBY PRYWATNEJ

- Aristotle (2009). *Organon*. From 1a to 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](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 the OJS platform: [ojs.sbp.pl/index.php/zin](http://ojs.sbp.pl/index.php/zin)

### 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 and the ORCID numbers,
- the e-mail address of the corresponding author,
- address for correspondence,
- biographic note (see below),
- structured abstract (see below),
- keywords (see below),
- statement of originality (see below).

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

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

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

#### 1.5. Structured abstract

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

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

### **1.6. Keywords**

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

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

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

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

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

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

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

- 1. First-level heading**
- 1.1. Second-level heading**
- 1.1.1 Third-level heading**

### **2.2. References**

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

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

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

### **2.4. Emphasis**

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

### **2.5. Illustrative content**

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

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

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

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

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

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

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

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

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

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

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

### 2.6.1. References List Examples

#### BOOK

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

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

#### BOOK (EDITED)

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

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

#### JOURNAL ARTICLE

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

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

#### BOOK CHAPTER

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

#### ELECTRONIC JOURNAL ARTICLE

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

Bartalesi, V., Meghini, C. (2016). Using an Ontology for Representing the Knowledge on Literary Texts: The Dante Alighieri Case Study. *Semantic Web* [online], 8(3), 385–394. <http://doi.org/10.3233/SW-150198>

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

#### ARTICLE IN ENCYCLOPEDIA

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

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

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

Article in encyclopedia with author information describe as book chapter.

ELECTRONIC DOCUMENT FROM WEBSITE

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

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

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

---

**Adres Wydawnictwa**

ul. Konopczyńskiego 5/7  
00-335 Warszawa, tel. 22 827 52 96

**Prenumerata i sprzedaż**

Dział Promocji i Kolportażu SBP  
Al. Niepodległości 213, 02-086 Warszawa, tel. 22 608 28 26  
**Cena prenumeraty na 2022 r. – 118 zł**

Wydawnictwo Naukowe i Edukacyjne SBP – Warszawa 2022.

Ark. wyd. 5,72. Ark. druk. 6,5.

Skład i łamanie: Justyna Grzymała-Łuszcz



