Information infrastructure of contemporary humanities and the digital humanities development as a cause of creating new information barriers. A Polish case

Zbigniew Osiński
Maria Curie-Skłodowska University, Faculty of Humanities, Institute of Information and Library Science, Lublin, Poland

Abstract

Purpose—To identify new information barriers created in the process related to the formation and transfer of humanistic scientific knowledge. Methodology—The reanalysis of the data collected during the earlier research projects. Results—Information barriers were identified, which are created by incomplete bibliographic databases, some solutions used in publishing on the Internet, the low level of information competencies of many humanists, their adverse attitude to open access, and by the education system that prefers knowledge transmission in the linear and text form. Conclusion—New information barriers effectively reduce research capabilities of many humanists, thereby creating a communication barrier between traditional and digital humanists.

Correspondence:
Zbigniew Osiński, Marii Curie-Skłodowskiej 4, 20-031 Lublin, Poland.
E-mail:
zbigniew.osinski@gmail.com

1 Introduction

Since the late twentieth century, there has been a belief among some humanists that the incorporation of modern technology into research work, combined with the mass digitization of culture and cultural activities on the Internet, has led to the creation of new opportunities for conducting research and scientific communication. The information space of humanities has increased by the digital sphere which includes bibliographic information, scientific publications and discovery tools, digitized sources for researchers, services that provide new research fields and new tools, and research methods for science (Górny *et al.*, 2015). The Internet has become an area of science communication, the zone of transfer and processing of data, and dissemination of information and knowledge generated by science (Sapa, 2009). There are new trends in the field of humanistic research that assume the use of computer hardware and software to explore the massive humanities databases and look for their relationships and correlations or explore rules and principles. The creation of digital humanities (DH) has significantly increased the role and significance of graphic forms of data presentation, information, and knowledge. It is claimed that visual transmission begins to displace the written word and that a new type of visual narrative appears. Some humanists say that the movies, collections of images and infographics, etc., are equivalent to text as a form of description and storytelling (Brown, 2002; Staley, 2003; Rosenstone, 2008; White, 2008; Paradowski, 2011; Radomski, 2016).

Digital Scholarship in the Humanities, Vol. 34, No. 2, 2019.

doi:10.1093/llc/fqy032 Advance Access published on 6 August 2018

In the recent decade in Poland, research projects have been implemented whose goal was to identify the characteristic features and regularities of the development of both digital information infrastructure and DH. The reanalysis of the data collected in the course of implementing several different research projects (Osiński, 2012, 2015a, 2015b, 2016, 2018; Górny *et al.*, 2017a, 2017b), and now combined into one resource, suggests that it is reasonable to advance a thesis that some aspects of the development of DH and the digital information infrastructure of science generate new information barriers in Polish science and education. As the authors of the studies in question did not directly refer to the thesis, it appears reasonable to reanalyze the data they have collected and the regularities they have identified in order to verify that thesis, important from the standpoint of directions and principles of the development of the humanities on the Internet.

Also in the humanities, the Internet becomes the main space for scientific communication and education. On the Web during the transfer of knowledge, a lot of important processes are performed, which include the publication of new knowledge and its validation and evaluation and providing of access to resources already produced, as well as ensuring cooperation between scientists. However, it can be observed that a multitude of techniques and technologies generates new barriers to research, scientific communication, and education.

In this context, it is important to identify these newest barriers in the process of the humanities development. The main purpose of this article is to identify new information barriers created in the process related to the formation and transfer of humanistic scientific knowledge. The article focuses on the barriers generated by the development of DH infrastructure and by the formation of new fields in research and presentation forms regarding scientific knowledge in the practice of digital humanists.

2 Literature Review—Information Barriers

According to Świgoń (2011) already in the 1960s, the literature mentioned information limits or different obstacles impeding access to information and its use. Technical, financial, and geographical macro-barriers, as well as micro-barriers related to a user's characteristics, were distinguished.

In the 1970s, the term information barriers was popularized. In the Polish 'Dictionary of Information Science Terms [Słownik terminologiczny informacji naukowej]' (1979), information barriers is defined as obstacles that prevent the use of information or information dissemination, e.g. linguistic, economic, technical, or ideological barriers. A broader definition is offered by the German 'Lexikon der Information und Dokumentation' (1984). It identifies information barriers with obstacles disturbing the process of information flow from the originator to the user. They may affect information users regardless of their awareness (objective barrier) or through this awareness (subjective barriers). Information barriers prevent the existing information from being used for two reasons:

(1) The user does not know that it exists or does not use well-known information sources; and
(2) The user's access to information sources, e.g. for financial, technical, legal, or similar reasons, is hindered.

Haag's (1989) definition suggests that information barriers always appear when there are discrepancies between the ideal and actual availability of the published information. In Wilson's conception (1981), in turn, information barriers are related to the characteristics of the information user (personal barriers), the social role s/he performs at work or life (interpersonal barriers), and to external circumstances, e.g. political, economic, technical (environmental barriers). Information barriers hinder the identification and awareness of information needs and search for information.

In the 1980s, the broader term became popular—access to information. After the use of Internet became widespread, three main elements of information access were distinguished (Borgman, 2000):

(1) Connectivity—connection with the network;
(2) Access to the resources of information content and information services; and

(3) Usability—skills enabling the user to utilize both technology and content.

Rice *et al.* (2001) named six elements of access to information:

(1) Access to knowledge contained in books, articles, reports, etc;
(2) Access to technologies and media;
(3) Ability to understand and absorb content;
(4) Control of information flow;
(5) Access to information as a commodity with social and economic value; and
(6) Access to government information resources, enabling participation in democracy.

Each of these elements may become a source of information barriers.

Two main groups of information barriers are distinguished (Engelbert, 1974; Julien, 1999):

(1) Objective and external—independent of the user (resulting from the construction of content and availability of sources of information); and
(2) Subjective and internal—dependent on the user (resulting from the information skills and awareness).

Świgoń (2006) collected different detailed typologies of information barriers found in scientific literature and systematized them. She distinguishes inter alia:

(1) User-related barriers:
    (a) ignorance barrier—not knowing of the information that can satisfy specific information need;
    (b) barrier of lack of knowledge—education gaps hindering orientation in a specific field of information, as well as hindering understanding information and interpreting it;
    (c) terminological barrier—difficulties caused by the failure to understand specialized vocabulary;
    (d) foreign language barrier;
    (e) ignorance of information sources and effective search strategies;
    (f) inadequate preparation to use modern information technologies, psychological resistance to using digital resources; and
    (g) passive attitude in seeking information, inability to systematically browse scientific literature; and
    (h) lack of time;
(2) Interpersonal barriers:
    (a) fear of asking others for help;
    (b) lack of actual help from those responsible for information provision; and
    (c) problems of communication between persons using different terminology and those who have a different level of knowledge;
(3) Environmental barriers:
    (a) legal barriers;
    (b) financial barriers;
    (c) distance-related barriers; and
    (d) political and ideological barriers;
(4) Barriers related to information sources:
    (a) Library-created barriers:
        (i) lack of purchase of or subscription to specific sources;
        (ii) unfriendly regulations;
        (iii) insufficient information about the holdings; and
        (iv) delays in interlibrary loans;
    (b) Barriers created by authors and publishers:
        (i) long publishing cycles;
        (ii) unavailability of unpublished materials (detailed research data, Master's and PhD theses);
        (iii) excess of published studies;
        (iv) low substantive quality of information;
        (v) discrepancy between bibliographic records and their actual content;
        (vi) unfriendly discovery tools; and
        (vii) dispersion of specific data among many sources.

The scale of the problem of information barriers in science is shown inter alia by the report on

research commissioned by the Research Information Net assessing the nature and scale of key restrictions on access to information resources of importance to researchers, the impact of these restrictions, and the ways in which they might be alleviated or overcome, which appeared in December 2009. (Overcoming barriers. . ., 2009) The report shows that only one in ten researchers in Great Britain did not encounter the situation of being unable to access the materials she/he needed in the institutional library, and two-thirds reported the problem every month or more often. The most frequent reason for lack of access proved to be insufficient institutional budget, electronic unavailability of particular volumes, and failure of the content to meet the adopted acquisition criteria. To seek scholarly content, researchers usually used their institutional library catalog, Google Scholar, and bibliographic and abstract databases. In the next stages of investigation, key barriers to accessing the content material were diagnosed:

(1) Specific content was not digitized and made available online;
(2) Libraries did not buy print materials or licenses for access to online materials although they are aware that the materials are available;
(3) Discovery tools are often complicated and not integrated with library systems;
(4) Licenses for online access are complicated and restrict access for unregistered users or nonmembers of host institutions.

Borovik and Shemberko (2016) argue that the main hindrances that complicate finding information in the humanities and social sciences embrace:

(1) Lack of knowledge about information resources and discovery tools, as well as about methods and techniques for accessing information;
(2) Difficulties in combining information obtained from different sources;
(3) Impossibility of utilizing more advanced data search strategies;
(4) Complicated structure of databases which causes information noise and hampers quick data retrieval;
(5) Lack of information on certain problems or their insufficient relevance;
(6) Language problems that make researchers give up some sources; and
(7) Information overload resulting from the use of many information sources at the same time.

Worth noting are also the studies by Savolainen (2015), who has distinguished six barriers hampering efficient information seeking: unwillingness to see one's needs as information needs, inability to articulate one's information needs, unawareness of information sources, low self-efficacy, poor search skills, and inability to deal with information overload.

3 Literature Review—Digital Humanities

Digital humanities is mentioned in scientific literature in the context of technical and technological transformations, the so-called digital and information revolution, taking place in the sphere of culture and science from the last decade of the twentieth century. It should, however, be admitted that the early attempts to use computer capabilities in the humanities date back to the late 1940s and early 1950s (Index Thomisticus). While we can already say with a large measure of certainty what digital humanities is and what it deals with, it does not appear possible to show an explicit definition of the concept. The outcome of the Day of DH project, scholarly debates conducted at conferences, and at dedicated websites (http://dayofdh2016.linhd.es/; http://dh.fbk.eu/) amounts to over 800 definitions (http://whatisdigitalhumanities.com/). They usually speak of the use of digital technologies to present cultural heritage of humankind and scientifically investigate this heritage, conduct in-depth analysis of and better understand humanistic data, cooperate more efficiently, and present research results. Emphasis is laid on the transdisciplinary nature of research problems and fields, the application of methods derived from other disciplines, and the use of the network and 'the wisdom of the crowd' to perform various tasks.

The programmatic article of this new trend in the humanities, 'Digital Humanities Manifesto 2.0' (2009), reads that print is no longer the exclusive

form of knowledge dissemination. Digital tools, techniques, and media are playing an increasingly important role in the process. Changes have also taken place in the scholarly activity of the humanists. They began to create digital stores of research sources (digitization), full-text databases, and search systems. They started to use digital tools to develop a new methodology of science and new publishing habits. These recent trends are characterized by the combination of qualitative and quantitative analyses, transdisciplinarity, collaboration in multidisciplinary teams, adoption of open-access principles, and publication of studies in multimedia forms, as well as by interest in cyberculture.

The editors of 'A Companion to Digital Humanities' (Schreibman *et al.*, 2004) stress in the Introduction that DH is of interest to researchers representing different research paradigms, methods, and disciplines, who seek new ways of asking questions and seeking answers, and what they have in common is that they use interaction with the computer for this purpose. DH will also interest those who would like to use computers to process large sets of sources, which could provide the data that could not be obtained without this. These researchers maintain that the methods of DH enable the discovery of new horizons in the humanities.

4 Literature Review—Digital Information Infrastructure for Humanities

Parallel to the development of DH, we are witnessing the buildup of the digital information infrastructure of the humanities. The term information structure entered scientific circulation only in the 1990s. The concept of information infrastructure emerged then but as a political construct—administration and state infrastructure (European Union Commission, 1994). The term became commonly used owing to the programmatic documents of the Organisation for Economic Co-operation and Development (OECD), G-7, European Union, and of the U.S. government, devoted to civilization changes, leading to the development of information society. In scientific reality, the information infrastructure began to be mentioned in the context of the Internet development. Out of the definitions applicable to describing this reality, worth noting is the one formulated by Pironti (2006). In his view, information infrastructure embraces all people, processes, procedures, tools, equipment, and technologies that take part in the process of creation, use, transfer, storage, and destruction of information. The definition offered by Ciborra and Hanseth (1998) says that information infrastructure may mold not only work procedures and methods of action but also the ways of perceiving these practices by people. It also impacts the fact that personnel and citizens treat these practices as 'natural' and gives them the quality of being indispensable and necessary. In this definition, information infrastructure appears as a factor that supports the acceptance and personal implementation of specific organizational solutions and practices.

In the early twenty-first century, the term cyberinfrastructure appeared. (Bowker *et al.*, 2010). It embraces computer hardware and software, information networks, databases, as well as digital sensors, tools, and services, and the appropriately trained personnel. As Hanseth (2002) rightly observed, the foregoing IT solutions differ from traditional information systems in many respects. Admittedly, the new digital infrastructure is being developed by broadening, modifying, and improving the existing solutions, but their digitization and networking creates entirely new opportunities in the sphere of availability and potential possibilities of utilization. In addition, digitalization and networking create a self-perpetuating mechanism that generates next solutions whose form and importance cannot be predicted even today, and which we are certainly going to deal with because of the increasing rate of innovation.

5 Method of Research

As has been said before, the considerations announced in the title of the article will be based on the data from scientific literature—the outcome of several research projects implemented in Poland.

The author reanalyzed the sum of data taken from those projects.

The reader may find it useful to become acquainted with a short presentation of research from which data have been used again. Several years ago, studies were conducted on the efficiency of Internet search tools from the perspective of the needs of the scholar researching Polish history (Osiński, 2012). The starting point was the information about the complete research achievements of selected historians over a period of 5 years. It was checked what portion of these achievements was shown by selected research tools. The following tools were verified (search for names and surnames as well as keywords): scientific search engines, library online catalogs, bibliographical databases, scientific social networks, and bibliography management programs. It was found that all the studied Internet search tools do not provide, for various reasons, possibly complete bibliographical information in the field of Polish history; consequently, they should not be treated as a fully reliable source of information because they are misleading as far as the true state of research on Polish history is concerned.

The next studies taken into account were meant to determine the current state of the construction of open access to research results in Central European Historiography (Osiński, 2015b). The method of study assumed that at the beginning, there would be a search and analysis of the information resources in Central European scientific journals that include historical subjects and provide free Internet access of articles contained at least in one complete year's issue. The investigations utilized the websites of journals and their publishers, bibliographical databases, journal–indexing bases, and the Google and Google Scholar search engines. Another step, after the list of journals meeting the foregoing criteria was made, was the analysis of the contribution of individual countries and languages to building the scientific Internet resources available on an open-access basis. The investigation also established the subjects of articles preferred by individual journals, the way and place of making the content available, and diagnosed the availability of articles to those potentially interested, as well as the impact of individual journals on the development of science measured by citations indexed by the Publish or Perish program. The research showed that the databases available to Central European historians could not be regarded as fairly complete and reliable sources of information about scientific journals offering articles on the Internet on an open-access basis because they contain significant gaps and inaccuracies. The widely used search engines Google and Google Scholar are not, however, dependable providers of data about the presence in the Web of complete texts of articles in historiography, which is caused by errors committed mainly by the editors of periodicals.

The goal of the next inquiry was to obtain the answer to the questions: can those interested in Polish history effectively use the Scopus database? Will they obtain the actual data on the state of their scientific discipline? (Osiński, 2016). The method was to conduct a comparative analysis between the image of foreign historiography of Poland created based on the Scopus database resources and the picture emerging from the information about the 2nd Congress of Foreign Scholars of Polish History. Foreign historiography of Poland was chosen as the comparative field because bibliographical data on the subject are contained in the Scopus database, and at the same time, there are other data about this group of scholars and their publications that enable comparison. Considerable inconsistency was found between the list of the surnames and affiliations of scholars recognized as significant by the Scopus database mechanisms and the one created by the decisions of the Congress organizers. The latter should be regarded as more reliable because those foreign scholars were invited to deliver reports at the Congress who, in the opinion of Polish historians, were the best known and the most important. Consequently, the bibliography of any problem concerning Polish history constructed from the data obtained from the Scopus database must be regarded as fragmentary and limited, particularly in respect of the studies by historians highly regarded in their own circles.

The data were also used from studies aimed to determine whether data and information visualizations can be an effective way of transmitting

specifically humanistic information and knowledge (Osiński, 2018). The following problem was studied: Whether a visual communication, containing by far more graphic elements than text, could be something more for students than a complement to the essential text (as in school handbooks)? Can it be a prevailing form of transmission? To obtain the research data, a series of didactic tests was conducted and administered to an intentionally selected several-dozen-strong group of students. They were the persons with predominantly humanistic competencies, but to whom modern engineering and technologies did not constitute a barrier to the cognitive process. The test results showed the usefulness of infographics as a form of organization and presentation, but only of detailed data and information. The visualization of information did not turn out to be a good basis for creating generalizations and deductions.

The objective of the next studies taken into consideration was to investigate transformations occurring in the digital information infrastructure of the humanities, prepare its model, and determine the already visible and foreseen impact of these changes and of the developing model on the research practices and publications of scholars (Górny *et al.*, 2017b). The method of Internet exploration was used to obtain data concerning individual constituents of the digital information infrastructure and exploration of library servers to gain data on the use of this infrastructure by selected groups of humanities scholars. As a result, the model of digital information infrastructure of the humanities was constructed; parameters for investigating it were distinguished; a connection between the use of infrastructure components, and the level of information competencies and the state of information awareness was demonstrated; the current state of the use of digital infrastructure by humanities scholars was described, as well as the mechanisms of its potential impact on the future research practices of humanities scholars were developed.

The data from studies on the development of individual information spaces of Polish humanities scholars were also utilized (Górny *et al.*, 2017a). The scholars devised the model of this space and determined a set of indices that characterize it together with their values. To obtain proper data, research scholars were surveyed, and individual questions were specifically constructed to investigate the knowledge and use of the information infrastructure, information and IT competencies, information awareness and undertaken information activities while performing professional duties. Additional data were obtained from bibliographic analyses of a group of periodicals and library servers. The outcome was the description of the model of how the individual information spaces of humanities scholars transformed themselves from traditional into digital ones, and the factors restricting this process were indicated.

The foregoing studies had in fact different goals and methodology and were not associated with identification of information barriers, but the data obtained in the course of investigation can be used to achieve a new goal, i.e. to identify new information barriers in the humanities. For they point out the state of the information infrastructure in the humanities, its use, and competencies of its users. They also identify new development trends in the humanities: the emergence of DH and the problems that the process encounters.

Summarizing:

The subjects of the research projects mentioned earlier were:

(1) Elements of the digital infrastructure of humanities such as databases, digital resources of scientific texts, and other tools available on the Internet;
(2) Available online examples concerning the scientific activities of DH such as Web pages created while implementing scientific projects; and
(3) Humanists (scholars and students) such as historians, linguists, and representatives of library and information science from four Polish scientific centers (Warsaw, Lublin, Poznan, and Torun).

The following research methods were used:

(1) Internet exploration to find components of digital infrastructure useful for humanists and digital scientific studies they have created;

(2) Surveys among humanists concerning the level of knowledge about the digital world and the use of digital scientific infrastructure;
(3) Analysis of data coming from university library servers regarding the use of elements of digital information infrastructure by selected groups of humanists; and
(4) Tests among the students on reading visual and nonlinear forms of presentation of information and knowledge.

6 Results—Three Groups of Problems

The analysis of publications resulting from the abovementioned projects and the data contained in them shows that here are three groups of problems:

1. Online databases and Internet search tools do not show the true condition of humanities research. The problems arise from the principles of data selection for bibliographic databases and the formats of scientific papers and journals available online. This observation is the outcome of analysis of the data from which those concerning Polish historiography were selected for the purpose of the present article (Osiński 2012, 2015a, 2015b, 2016).

The most complete database of *Polish scientific journals Arianta* (http://www.arianta.pl/) shows that there are about 400 periodicals in the field of history, and the European database ERIH Plus lists about seventy. In contrast, in the most prestigious, international, and multidiscipline databases, like the Web of Science and Scopus, there is only one such journal at present. Similar discrepancies are observable in the case of Polish history journals that offer open access to at least some articles on the Internet. The Arianta database lists ca. ninety such periodicals, the *Electronic Journals Library*—forty-three, *Central European Journal of Social Sciences and Humanities*—twenty-seven, Slavic Humanities Index—twelve, Central and Eastern European Online Library—ten, European Reference Index for the Humanities (ERIH) Plus—three, Web of Science and Scopus—one each, Directory of Open Access Journals, JURN, and the *Journal for Free*—zero. Moreover, a large portion of humanities journals is interdisciplinary, and the editors of databases sometimes classify the journals that do not contain historical articles into the History category. There are also cases when periodicals containing historical subjects are not classified into the History section.

A highly incomplete and at the same time random selection of the data on the achievements of Polish historiography can be shown using the example of the Scopus database. It contains information about 423 titles classified into the subject area: History, published in Europe and North America (studies about the history of Poland are written mainly in these regions of the world). In the past dozen years, (since 2004) only seventy-eight periodicals have published any studies devoted to Polish history, sixty-nine of which—only incidentally. It turns out that to 248 such articles there are 234 names of the authors (few articles have two or three authors), of which forty-eight authors are affiliated with Polish universities (apart from them, the dominant researchers are those working in Germany—fifty-seven, the USA—forty-six, and Great Britain—twenty-two). What is surprising is the vestigial presence in the Scopus database of historians from many countries with which Poland had diverse historical relationships. That is why they can be expected to conduct large-scale studies on mutual relations (e.g. the Czech Republic, Hungary, Slovakia, Sweden, Ukraine, and Russia). It would be absurd to assume that apart from several countries, there are no historians studying Poland's history, who would be good enough scholars to publish an article in a prestigious journal (and the Scopus editors suggest they only take high-quality papers into consideration).

A partial explanation of the problem can be obtained using the information about the participants in the II Congress of International Researchers of Polish History, which took place in September 2012 in Krakow. The organizers invited the international scholars who, according to Polish historians, were best known and most significant to present their reports. The problems discussed in their lectures were recognized as identical with the research subjects concerning Polish history that are the most popular outside Poland. When confronting the

picture of the foreign historiography of Poland's history presented by the Scopus database with the picture created by the decisions of the II Congress organizers, we can see a great discrepancy between the two sets of names and affiliations of the researchers regarded as eminent. First of all, in the list of the speakers at the II Congress, there is no prevalence of historians from Germany (Scopus—fifty-seven, and reporters at the Congress—twenty-four), the USA (forty-six, six) and from Great Britain (twenty-two, five). In contrast, there is a large contingent of scholars from Central-Eastern European countries: Austria (Scopus five, and speakers at the Congress nine), Belarus (zero, five), Czechia (two, thirteen), Hungary (two, ten), Lithuania (two, eleven), Russia (two, nine), Romania (two, three), Slovakia (five, three), and Ukraine (one, nineteen). If we compare the names of 186 historians and their non-Polish affiliations contained in the Scopus with 152 speakers at the II Congress, it will turn out that the two sets have little in common. The data from the Scopus database show that outside Poland the interest in studying Poland's history is negligible as evidenced by an extremely modest number of articles on the subject published over the past dozen years or so in the journals regarded by the Scopus editors as valuable. This conclusion is also confirmed by other data from this database: only one in six indexed journals published any article about Polish history during the investigated period: a vast majority of these periodicals published such studies incidentally, while the overwhelming majority of the authors, historians with non-Polish affiliations, were incidentally concerned with Polish issues, usually in the context of the history of other countries and societies. This picture is highly incompatible with the conclusions resulting from the sessions of the II Congress of International Researchers of Polish History because it turned out that over a thousand researchers are concerned with many aspects of Polish history.

Another problem is that from the standpoint of a researcher seeking articles on a specific subject on the Internet, what is essential is the visibility of these materials in the search results. The achievements of Polish historians are barely visible on the Internet. The Google Scholar shows the presence on the Web of only some papers by these researchers, contained in various digital resources (depending on the researcher, from 3% to 33% of all publications available on the Internet). Only single items of information about some of the publications by abovementioned researchers can be found by means of such search tools as Bielefeld Academic Search Engine, Microsoft Academic Search, and OpenDOAR. The reasons for this state of affairs are diverse. In the case of a journal in the form of a WWW site with articles in the html format, it is possible to reach their content through the general Google search engine because the specialist Google Scholar search engine does not show htlm articles, even from scientific journals, on the first three result pages. None of the search engines will show papers that are part of journals or collective publications placed on the Web as one, overall file. A similar barrier is created by sharing a scientific article in the graphic file formats (scans and photos) without Optical character recognition (OCR) layer. Moreover, articles are shared on the Internet even with a several-year delay as compared to the print version, sometimes only some issues and articles are available on an open-access basis. Managers of some digital repositories and libraries did not see to implementing the The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) protocol, owing to which it is possible to exchange metadata with aggregators like Agregator CeON and Open DOAR. Not all institutions of this type have been registered in the OpenAIRE project combining the European open-access infrastructure.

2. Most humanists do not use the opportunities offered by digital infrastructure. There are major problems with information skills and information awareness among Polish humanists. Conducting humanities research with the use of traditional tools and methods does not allow the development of information skills in the digital world. Many humanists are only satisfied with the tools and resources that support their traditional research activities. Digital publications do not have the same prestige as printed works. The principles of open access are not trusted.

Studies (Górny *et al.*, 2017a) conducted among Polish humanists (historians, linguists, literary

scholars, and specialists in biology and information science) have demonstrated that when looking for needed scientific publications, they mainly use library catalogs, subject bibliographies, and the Google Scholar search. A great majority of specialist search tools, bibliographic databases, and lists of periodicals are not utilized. Lack of knowledge that these tools and databases exist has been found in two thirds of the investigated humanists. In the survey, this group of respondents did not give any name of a tool or database, or they named only the Google Scholar. Few of them (only 7%) realized that they had too low competencies in searching and using scientific publications available on the Web, and the lack of access to specific tools and resources—declared by half of the subjects—was actually a manifestation of ignorance of their existence. A good example is the statistics of accessing various databases through the server of one university (the server records accesses to paid databases, for which the university purchased license, both from the university's computers and from private hardware of the faculty members through the homepage of the university library). It turns out that in 2015 such databases as the International Directory of Medievalists and International Medieval Bibliography were not accessed despite the fact that the university employs several dozen researchers interested in the Middle Ages (philosophers, historians, literary scholars, linguists, culture experts, and archeologists). The statistics of accessing multidisciplinary databases (Ebsco, Ebrary, eJournals, iBuk, Science Direct, Scopus, and WoS) show that they were used on average several times a year per one faculty member.

The attitude of the studied groups toward digital publishing is characteristic. Almost three quarters do not decide to publish their articles in digital versions only. A similar group of humanists has greater confidence in print sources than in online sources, and 62% of them do not consent to the publication of their studies in the open-access format although the idea itself is supported by more than half of the subjects. A digital text is still perceived as something potentially transitory (the problem of disappearing pages on the Internet), and first of all as something scientifically less valuable. Open-access publishing, in turn, appears to be mysterious and potentially dangerous (easy to plagiarize). Similar results (two thirds of the subjects answered—no) were obtained from questions about depositing one's own publications in scientific repositories, about reading blogs conducted by scientists or using scientific social networking.

The respondents usually learn about information search tools and scientific databases by themselves, from their own search. Training in the use of digital information search tools is not particularly popular among humanities researchers. More than half of them do not use the tools at all; a large group—definitely seldom. The respondents' information on archiving tools and managing bibliographies and materials collected from search queries comes, first of all, from their own search, informal contacts, and publications. Most of the respondents never took part in training in this field or do not find instructions of this kind in blogs. The question about text analysis programs showed that almost all humanists did not know them. Despite a significant number of respondents representing disciplines for which text is the fundamental research material, the majority of respondents do not know computer text analysis tools.

A special case is the question of the ability to conduct analyses of large sets of humanistic data (big data and data mining), like collections of texts, photos, pictures, or films, containing several thousand or even several hundred thousand items. On the one hand, some scholars are aware of the fact that analyses of big data sets may yield results that would revolutionize many humanistic theories; we have computer software enabling the automation of such analyses, and the community of DH scholars in Poland are implementing research projects based on big data and data mining, and organizing conferences on these problems. On the other hand, however, almost no humanities scholar outside this community understands those problems and cannot use appropriate software, which the author found when conducting trainings for academic personnel. Observations from these courses allow us to advance a thesis that a vast majority of Polish humanities scholars are not even aware of the possibilities that big data and data mining can open before them.

3. The visual transmission of knowledge used by digital humanists (for example, infographics) is difficult to understand by students and by a large part of the population. Problems were created by the proportions between text and graphics when they were more advantageous for graphics. Next problems arose in the course of interpreting and generalizing the details of infographics and reading qualitative data. The education system did not provide us with appropriate skills.

A survey (Osiński, 2018) was conducted among students at the Faculty of Humanities at one Polish university, intended to show to what extent they were able to read the content presented in visual form and create their own messages in this form. The survey was carried out based on different types of infographics available on the Internet, containing explanations and information in visual form and integrating pictures and words. To obtain research data enabling determination of the efficacy of search for information in infographics and the degree of understanding their message, a series of didactic tests was conducted. Verification of the way and extent of understanding these forms of transmission was also carried out through tasks that required the student's own creation of infographics for a set topic. Tests were administered to a specifically selected group of students, which can be identified with a typical graduate of Polish education system in terms of his/her information competencies. The group performed test tasks with diverse infographics which required: reading detailed information given directly or guessable from the context and symbols, inference based on the analysis of a set of charts or materials forming a project in DH, reading the general message of an infographic and detailed problems contained in it, and creating a narrative based on infographics or materials forming a project in DH. Another group of tasks required composing infographics on one's own according to specific criteria. All the questions and tasks required giving unassisted answers rather than selecting them from among ready-made distractors. The point was to make the form of the test maximally close to the practice of daily use of information sources. The infographics were selected in such a way as to have the main information types in them: numbers, text, location in time, and location in space. The information in the infographics was visualized in the form of graphs, diagrams, schematics, maps, and pictures.

The infographics did not turn out to be a source of data, information, and knowledge that would create an impenetrable information barrier. A clear relationship was observable; however, the more advantageous the proportions between text and graphics were for the text, the fewer problems the students had with correctly reading the content. Another relationship was also noticeable—the more information the analyzed material contained, the more data, especially the peripheral ones, were omitted while performing a task. The respondents did not have difficulties reading detailed information, but they had problems with generalizing it. They found quantitative data without problems, but they found it difficult to discern qualitative data. A considerable challenge was to correctly name the analyzed phenomenon unless the viewed material showed it directly. It was also a problem to read the temporal and spatial variability of the phenomena presented in an infographic form, particularly when this involved comparing several different graphic and text elements. It was a great difficulty to interpret details read from an infographic and translate their set into a consistent message. This proved especially difficult in the case of subjects that the students did not know well. A frequently observable behavior of the respondents was attempts to find additional information although the information in the infographic was entirely sufficient to perform the task. Infographics proved best in the role of the resource that organized and ordered detailed information. The infographics created by the investigated students were in turn characterized by the use of visual elements mainly as embellishments rather than an information source on an equal footing with text.

7 Results—New Information Barriers in Humanities

The conclusions from the analysis of the present data show that the development of the digital information structure of science and DH generates new

information barriers both to the researchers themselves and to recipients of scientific knowledge. Several barriers of this kind can be shown, which the literature on information barriers has so far not mentioned:

1. International, online bibliographic databases are incomplete, especially in the case of humanities. The basic conclusion that arises after the analysis of information on Polish historical periodicals, visible in international bibliographic databases and on the lists of journals, is pessimistic. To the researcher of Polish history (and also to researchers of the history and culture of other countries and societies who use sources other than in English), the value of these databases and registers is negligible because all too often they present incomplete or even misleading data. Essentially, they omit the vast majority of the achievements of humanities that were published outside large publishing companies and the subject matter based on sources other than in English and on national problems. A researcher who would like, by means of the foregoing bibliographic databases and lists of periodicals, to find journals and articles devoted to Polish history (and more broadly, in the sphere of culture and history outside English) runs the risk of time-consuming search which can be fragmentary and incomplete. This risk pertains, first of all, to researchers who do not speak Polish (or Czech, Russian, Romanian, Hungarian, etc.) and who use bibliographic databases and lists of journals in English. The picture of the culture and history of many nations that emerges from these resources has little in common with the actual state.

2. Some of full-text databases, digital editions of journals, and scientific books are placed on the Internet in formats that prevent indexing by the most popular search engines. Due to problems with data formats, the studies published in more than one journal or deposited in many a repository or digital library are not shown in the search results of many specialist search engines, thereby not being visible to the academic community. The processes of digitization of archival issues of journals were often based on wrong technologies (scanning without OCR, the whole issue of a journal in the form of one file). Cooperation with indexing databases was not properly seen to, while the authors themselves were not used to complementing their texts with abstracts in English, with keywords, and with reference lists (bibliographical footnotes being popular, instead). Self-archiving of scientific papers in different repositories took place without due care of the file format and its metadata. As a result, it is impossible to effectively reach a considerable portion of papers by Polish historians placed in network resources if the complete access address of a database is not known. A researcher who is looking for articles she/he does not yet know encounters a barrier created by the founders of many a journal and the full-text database, and by the authors of papers themselves—a barrier that hides a portion of the scholarly achievements.

3. A large part of digital information infrastructure remains beyond the awareness and skills of many humanists. They do not know of its existence and they are not aware of its usefulness. The personal information space of these scholars does not include a large part of digital space, thereby creating a barrier to the use of research and publishing opportunities available on the Internet. Polish humanists acquire too few competencies in digital information infrastructure and in communication on the Internet. They are not aware of the existence of many resources and tools and cannot keep up with learning new developments in this field. Nor do they know new research opportunities emerging in the digital reality. They prefer a traditional approach to research and publishing. They block their opportunities of accessing a large portion of the achievements of humanities not known to them and of utilizing research approaches other than traditional. A growing division of humanist scholars into two groups is observable, who are able to collaborate only to a negligible extent: traditional humanists and digital humanists.

4. Many researchers are afraid to share their works in the open-access model. They are mainly afraid of plagiarism. What is responsible for this state of affairs is the ignorance of the technological side of the Internet and open-access principles, comparatively often noticeable among Polish humanists. They perceive the Internet from the angle of frequent disappearance of WWW pages and widespread copying of what others have made

available there. The traditional approach to science tells them that traditional publication in print form is more valuable. They, thus, block the possibility of presenting their accomplishments to a wide circle of recipients. They seem to disregard the need to make their achievements available beyond their narrowly understood academic community, especially to the younger part of society who uses mainly the texts that can be found on the Web and utilized free of charge.

5. Visual forms of presentation of humanistic knowledge create a cognitive barrier to graduates because of the current model of education based on the text and linear knowledge transmission. It should be emphasized that a typical graduate of the Polish educational system is not suitably prepared to receive the visual content concerning abstract phenomena and processes. Hence, the hope of digital humanists that writing and linear narrative may be superseded by visual forms of transmission in describing the world is premature in the case of a significant part of Polish society. It will be a long time before publications based on visualizations will be a remedy for the perceptible decline of traditional reading in Poland. Visualizations do not, admittedly, increase information noise, but, at present, they are not becoming a significant form of transmission of abstract knowledge from scientists and experts to society. An indisputable barrier is the model of education based on the verbal and textual, as well as linear transmission of knowledge. The information competencies of a graduate of this system do not allow the efficient use of information visualization. Consequently, the development of visual forms of information and knowledge transmission used by digital humanists creates another information barrier.

8 Conclusions

It can be said that the development of digital infrastructure for humanities and the development of DH overtakes the development of information skills and information awareness among many humanists. Their personal information spaces represent an ever-smaller part of the information space of science. New information barriers effectively reduce their research capabilities, which creates a communication barrier between traditional and digital humanists. Therefore, traditional humanists do not cope well with their educational role—they do not prepare young people for the challenges of the digital world.

The sources of these problems should be sought in the model, still recently prevailing in Poland, of training the academic personnel of universities/colleges and in the rules of scholarly assessment and promotion. A young humanities scholar developed his/her abilities under the guidance of his/her master, who taught him/her the patterns of scientific behavior that s/he had her/himself acquired many years ago. Their important component was the conservative cult of printed sources: consequently, research was conducted using printed materials collected in libraries and archives, and the principal research activity was to carefully read and analyze these materials. That practice was strengthened by the extremely weak development of digital resources for study. Research results were published exclusively in the form of printed books and articles because digital forms were treated as inferior and insufficient to obtain an academic promotion. Accustomed to such realities, humanities scholars do not seem to be interested in professionally moving publications to the Internet; they cannot ensure the implementation of standards that allow the indexing of journals in bibliographical databases and articles in Internet search engines. They have a great mistrust of the open-access rules, Creative Commons licenses, and computer-assisted analysis of research sources. In recent years, this situation has changed, its indication being the development of DH. However, without legal and organizational changes supporting the publication of scientific studies on the Internet and permitting of scientific communications in which text is not dominant, as well as without the systematic training of research scholars, the Polish humanities scholars may, for many years to come, have problems with the effective use of the rapidly growing information infrastructure and computer programs that support the research process.

Funding

This work was supported by the Maria Curie Sklodowska University as part of funds for statutory research.

References

Borgman, C. L. (2000). *From Gutenberg to the Global Information Infrastructure. Access to Information in the Networked World*. Cambridge: MIT Press, pp. 56–7.

Borovik, M. A. and Shemberko, L. V. (2016). The challenges of information retrieval in social sciences and humanities and ways to overcome information barriers. *Scientific and Technical Information Processing*, **43**(2): 99–105.

Bowker, G. C., Baker, K., Millerand, F., and Ribes D. (2010). Toward information infrastructure studies: ways of knowing in a networked environment. In Hunsinger, J., Klastrup, L., Allen, M., (eds), *International Handbook of Internet Research*. Dordrecht: Springer Link.

Brown, J. S. (2002). Learning in the digital age. In: *Forum Futures 2002*. New York, NY: Educause, pp. 20–3.

Ciborra, C. U. and Hanseth O. (1998). From tools to gestell. Agendas for managing information infrastructures. *Information Technology and People*, **11**(4): 305–27.

Digital Humanities Manifesto 2.0. (2009). http://manifesto.humanities.ucla.edu/2009/05/29/the-digital-humanities-manifesto-20/ (accessed 10 November 2017).

Engelbert, H. (1974). Informationsbarrieren. *Informatik*, **6**: 47–51.

European Union Commission. (1994). Report on Europe and the Global Information Society: Recommendations of the High-level Group on the Information Society to the Corfu European Council. Bulletin of the European Union, Supplement No. 2. http://aei.pitt.edu/1199/ (accessed 10 November 2017).

Górny, M., Głowacka, E., Kisilowska, M., and Osiński, Z. (2015). Model Indywidualnej Przestrzeni Informacyjnej jako narzędzie badania procesów informacyjnych w humanistyce cyfrowej [The model of personal information space as a tool of investigating information processes in digital humanities]. *Zagadnienia Informacji Naukowej [Problems of Information Science]*, **2**: 18–30.

Górny, M., Głowacka, E., Kisilowska, M., and Osiński, Z. (2017a). *Mechanisms of the Formation and Evolution of Personal Information spaces in the Humanities*. Poznan: Wydawnictwo Rys.

Górny, M., Głowacka, E., Kisilowska, M., and Osiński, Z. (2017b). Wpływ cyfryzacji infrastruktury informacyjnej na procesy badawcze w humanistyce. Wybrane aspekty zjawiska [The impact of digitization of information infrastructure on research processes in the humanities. Selected aspects of the phenomenon]. *Zagadnienia Informacji Naukowej [Problems of Information Science]*, **2**: 137–55.

Haag, D. E. (1989). *Information Barriers, Pretoria: Human Science Research Council.* Pretoria: Human Science Research Council, p. 5.

Hanseth, O. (2002). From systems and tools to networks and infrastructures - from design to cultivation. *Towards a theory of ICT solutions and its design methodology implications*. http://heim.ifi.uio.no/~oleha/Publications/ib_ISR_3rd_resubm2.html (accessed 10 November 2017).

Julien, H. (1999). Barriers to adolescents' information seeking for career decision making. *Journal of the Association for Information Science and Technology*, **50**(1): 38–48.

Osiński, Z. (2012). Sprawność internetowych narzędzi wyszukiwawczych z punktu widzenia badacza dziejów Polski [The efficiency of internet search tools from the standpoint of researchers of polish history]. In: Turyn, J. (ed.), *Rola informatyki w naukach ekonomicznych i społecznych. Innowacje i implikacje interdyscyplinarne [The Role of Information Science in Economic Sciences. Interdisciplinary Innovations and Implications]*, vol. **1**. Kielce: Wydawnictwo Wyższej Szkoły Handlowej, pp. 452–61.

Osiński, Z. (2015a). Współczesna humanistyka w Internecie—główne tendencje rozwojowe z perspektywy informatologicznej [The contemporary humanities on the internet—main development tendencies from the information science perspective]. In: Osiński, Z., Malesa, R., and Kotuła S. (eds), *Biblioteka, książka, informacja, Internet 2014 [Library, Book, Information, Internet 2014]*. Lublin: Wydawnictwo UMCS, pp. 11–32.

Osiński, Z. (2015b). Open access w środkowoeuropejskiej historiografii—perspektywa informatologiczna. [Open access in central-european historiography—the information science perspective]. *Zagadnienia Informacji Naukowej [Problems of Information Science]*, **2**: 31–64.

Osiński, Z. (2016). Historia Polski w czasopismach indeksowanych w bazie Scopus [Polish history in indexed

journals in the scopus database]. In: Głowacka, E., Jarocki, M., and Pamuła-Cieślak, N. (eds), *Współczesne oblicza komunikacji i informacji. Przestrzeń informacyjna nauki [The Contemporary Faces of Communication and Information. Information Space of Science]*. Toruń: Wydawnictwo Naukowe UMK, pp. 191–212.

Osiński, Z. (2018). Infographics in humanities: transmission of information or information noise? The polish case. In: Osinska, V. and Osiński, G. (eds), *Information Visualization Techniques in the Social Sciences and Humanities*. Hershey: IGI Global, pp. 50–67.

Overcoming Barriers: Access to Research Information Content. A Research Information Network report. (2009), http://www.rin.ac.uk/system/files/attachments/overcoming_barriers_report.pdf (accessed 10 November 2017).

Paradowski, M. (2011). Wizualizacja danych—dużo więcej niż prezentacja [Visualization of data—much more than presentation]. In: Kluza, M. (ed.), *Materiały konferencji "Wizualizacja wiedzy. Od Biblii Pauperum do hipertekstu" [Conference Proceedings "Visualization of Knowledge. From the Bible Pauperum to Hypertext"]*. Lublin: Wiedza i Edukacja, pp. 37–60.

Pironti, J. P. (2006). Key elements of a threat and vulnerability management program. http://iparchitects.com/wp-content/uploads/2016/07/Key-Elements-of-a-Threat-and-Vulnerability-Management-Program-ISACA-Member-Journal-May-2006.pdf (accessed 10 November 2017).

Radomski, A. (2016). Wizualne analizy, wizualne narracje [Visual analyses, visual narratives]. In: Bomba, R., Radomski, A., Solska, E. (eds), *Humanistyka cyfrowa. Badanie tekstów, obrazów i dźwięku [Digital humanities. Examination of Texts, Images and Sound]*. Lublin: Wydawnictwo E-Naukowiec, pp. 147–58.

Rice, R. E., McCreadie, M., and Chang, S-J. L. (2001). *Accessing and Browsing Information and Communication*. London: MIT Press, pp. 45–55.

Rosenstone, R. (2008). Historia w obrazach/historia w słowach: rozważania nad możliwością przedstawienia historii na taśmie filmowej [History in images/history in words: reflections on the possibility of presenting a story on film]. In: Kurz, I. (ed.), *Film i historia. Antologia [Film and History. An Anthology]*. Warszawa: Wyd. Uniwersytetu Warszawskiego, pp. 93–116.

Rückl, S. and Schmoll, G. (eds). (1984). *Lexikon der Information und Dokumentation*. Leipzig: Bibliographisches Institut, p. 187.

Sapa, R. (2009). *Metodologia badań obszaru pośredniczenia w komunikacji naukowej z perspektywy nauki o informacji [Methodology of Research on the Area of Mediation in Scholarly Communication from the Perspective of Information Science]*. Krakow: Wydawnictwo Uniwersytetu Jagiellońskiego.

Savolainen, R. (2015). Cognitive barriers to information seeking: a conceptual analysis. *Journal of Information Scvience*, **44**(5): 1–11.

Schreibman, S., Siemens, R., and Unsworth, J. (eds) (2004). *A Companion to Digital Humanities*. Oxford: Blackwell. http://www.digitalhumanities.org/companion/ (accessed 10 November 2017).

Słownik terminologiczny informacji naukowej [Dictionary of Information Science Terms], (1979). Warszawa: Zakład Narodowy im. Ossolińskich, p. 28.

Staley, D. J. (2003). *Computers, Visualization, and History: How New Technology Will Transform Our Understanding of the Past*. Armonk, NY: M. E. Sharpe.

Świgoń, M. (2006). *Bariery informacyjne [Information Barriers]*. Warszawa: Wydawnictwo SBP. pp. 24–36.

Świgoń, M. (2011). Information limits: definition, typology and types. *Aslib Proceedings*, **63**(4): 364–79.

White, H. (2008). Historiografia i historiofotia [Historiography and historiophotics]. In: Kurz, I. (ed.), *Film i historia. Antologia [Film and History. An Anthology]*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego, pp. 117–30.

Wilson, T. D. (1981). On user studies and information needs. *Journal of Documentation*, **37**(1): 8.