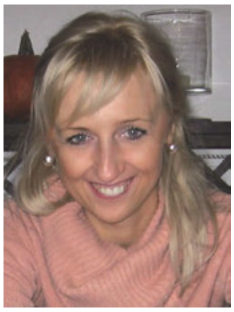
The Role of Visualization in the Shaping and Exploration of the Individual Information Space: Part 1†

Veslava Osinska*, Malgorzata Kowalska**, Zbigniew Osinski***

Nicolaus Copernicus University, Institute of Information Science and Book Studies, Torun, Poland,
*<wieo@umk.pl>, **<koma@umk.pl>
***Maria Curie-Skłodowska University, Institute of Library and Information Science, Lublin, Poland,
<zbigniew.osinski@gmail.com>



Veslava Osinska received an MSc degree in physics and a PhD in information science. She works in the Institute of Information Science and Book Studies at the Nicolaus Copernicus University in Torun. Her research interest concerns information and knowledge visualization, in particular, science domain analysis. She teaches data visualization and design, computer graphics and animations. She has also an experience in ECDL training. Dr. Osinska is collaborating with Places @ Spaces International Project and became its ambassadour in Poland. She is a local coordinator of the Polish Chapter of ISKO and a member of the Polish Information Processing Society.

Malgorzata Kowalska is an assistant professor in library and information science at Nicolaus Copernicus University in Torun, Poland. Her scientific interest focuses on a broad range of LIS topics, such as digital libraries, quality of digital resources, application of ICT and crowdsourcing in GLAM institutions, infobroking, scientific development of LIS as a research field, information management, bibliometrics and scientometrics. Malgorzata Kowalska is a member of the ISKO Polish Chapter.

Zbigniew Osinski is Professor at the Marie Curie Sklodowska University in Lublin. In 2009, he became Chief Department of Information Science in the Institute of Library and Information Science. Since 2017, he has been the editor-in-chief of a scientific journal *Informatio et Scientia. Information Science Research*. His research interests are: information science (educational and scientific internet database), humanists scientific communication, information architecture, bibliometrics and webometrics, digital humanities (information visualization and data analysis of large collections of text and graphics objects, transfer of knowledge through infographics).

Osinska, Veslava, Malgorzata Kowalska and Zbigniew Osinski. 2018. "The Role of Visualization in the Shaping and Exploration of the Individual Information Space: Part 1." *Knowledge Organization* 45(7): 547-558. 30 references. DOI:10.5771/0943-7444-2018-7-547.

Abstract: Studies on the state and structure of digital knowledge concerning science generally relate to macro and meso scales. Supported by visualizations, these studies can deliver knowledge about emerging scientific fields or collaboration between countries, scientific centers, or groups of researchers. Analyses of individual activities or single scientific career paths are rarely presented and discussed. The authors decided to fill this gap and developed a web application for visualizing the scientific output of particular researchers. This free software based on bibliographic data from local databases, provides six layouts for analysis. Researchers can see the dynamic characteristics of their own writing activity, the time and place of publication, and the thematic scope of research problems. They can also identify cooperation networks, and consequently, study the dependencies and regularities in their own scientific activity. The current article presents the results of a study of the application's usability and functionality as well as attempts to define different user groups. A survey about the interface was sent to select researchers employed at Nicolaus Copernicus University. The results were used to answer the question as to whether such a specialized visualization tool can significantly augment the individual information space of the contemporary researcher.

Received: 23 January 2018; Revised: 6 May 2018; Accepted: 10 July 2018

Keywords: information visualization, research, scientific disciplines, data analysis

† The research is sponsored by the Polish National Science Centre (NCN) under grant 2013/11/B/HS2/03048/Information Visualization methods in digital knowledge structure and dynamics study.

1.0 Introduction

The visualization of information, which has been rapidly developing in recent years, provides increasingly functional and attractive forms of knowledge presentation. Professionals mention the reduction of information (brevity) and as a consequence, better adaptation to the level of user perception, as well as better comprehension of a large amount of data, among other advantages. The well-known saying that "one picture is worth more than a thousand words" is often used as a confirmation of this thesis ("Meaning and Origin" 2017). However, the role of information visualization is much more extensive and is not limited to compressing textual content. Its use, even in the business environment, clearly indicates usefulness both in accelerating decision-making processes and in implementing specific solutions.

Visualizations, made in a deliberate way, play an important role in science as well; they create an opportunity to discover mutual connections between authors and identify inter- and multidisciplinary research groups, flows of theory and scientific trends, paradigm shifts, or the development of new research fields. Such images can inform researchers about emerging scientific fields and show where and how they overlap or disappear, giving rise to new areas of knowledge and revealing unpopular research disciplines. Regarding knowledge organization, visualization maps provide valuable information about research theme composition and the delivery methodology to analyse knowledge domains (Smiraglia 2015a; Hjørland 2017). Visualizations represent conceptual relationships between domains that are important for revealing inter- and transdisciplinary areas (López-Huertas 2015). Critical changes in visual patterns on one hand lead to the discovery of scientific frontiers (Chen 2017; Raghavan et al. 2015) and prove the imperfections of knowledge organization; and on the other hand, systems (Osinska and Bala 2010; Buchel and Sedig 2011).

The demand for multithreaded data analyses and their visualizations in science is largely due to the expanded framework of open access to scientific databases. In the last decade, we can note the methodological shift in knowledge organisation into domain analysis (Smiraglia 2015b). This is closely connected with bibliometrics and scientometrics as well as information visualization. By visualizing metadata from bibliographic databases at an appropriate level, we can obtain maps of scientific cooperation on a global, national, institutional, or individual scale (Börner 2015). If the assessment of the scientific condition concerns countries, continents, or more broadly, the whole globe, according to terminology derived from the Social Network Analysis (Newman 2010), it is referred to as "macro" scale research. When such research is conducted at the level of institutions or local communities, it is referred to as a "meso" scale study, and when it measures the activity of individuals, a "micro" scale study (Börner 2015).

Generally, knowledge domain mapping is limited to macro and meso levels. We can notice the long tradition of mapping science that has developed since the 1960s (Garfield 1994; Börner 2010). The micro level is, however, very rarely presented and discussed. The management staff of scientific institutions is unlikely to be interested in analyzing the paths of individual researchers due to the high costs of generating and aggregating data relating to each individual. In addition, the analysis of individual careers is very time-consuming and biased because of possible subjective assessment. Another limitation is the low representativeness of the research material in terms of statistics: the publication output of a single scientist rarely exceeds 1,000 items (Börner, 2010), which means that statistical significance or reliability decreases dramatically.

Despite the described limitations of a uniform range of knowledge, scientific research conducted in the micro dimension has deep meaning. This article is, on one hand, an attempt to prove this claim and on the other, fill the literature gap in this context. For the purposes of the current study, the authors developed a web application for visualizing the bibliographic data generated for a single researcher. The results were to answer the question whether such a specialized visualization tool can significantly augment the individual information space of the contemporary researcher, and if so, in which scientific disciplines, in what function and scope, and if not, for what reasons.

2.0 Individual information space (IIS)

The origin for the concept of an "individual information space" (IIS) defines the information space. In the science of information, the term is used in parallel to other semantically similar terms, e.g. "information environment" or "infosphere." One of the most comprehensive definitions of what other researchers have called "information space" was developed by Davenport (1997), who, when writing about the information environment in an organization, listed six of its components: information strategy, information policy, information behavior, information workers, information processes, and information architecture. Boisot (1995), in turn, used the concept of information space to allow the study of the evolution and spread of knowledge and information in society and the social process of learning and sharing information within and between social groups. A completely different, very narrow definition of information space was proposed by Dillon (2000), claiming that it is primarily "available content." Similarly, this phenomenon was interpreted by McKnight (2000), perceiving the information space as objects (real or virtual), which a given person uses to obtain information. Polish researchers also undertook the analysis of IIS both in terms of terminology and scope of usefulness in a modern academic environment (Kisilowska 2011; Skórka 2011).

"Individual information space" is sometimes defined in opposition to the public information space considered as a repository of all information resources and channels (Gwizdka 2006). For the realities of science, the term IIS is proposed by Sauermann, van Elst and Dengel (2007), who define "personal information space" as a set of all data needed by the unit to conduct research work, regardless of how they were obtained by the user, and from the source, format, and author. It is recognized that all standard information processes are implemented in IIS. Many tasks related to the performance of professional work, including group tasks, are carried out in this individual information space (Kljun and Dix 2012). Shemberko and Borovik (2012) mention the individual information space in the context of information retrieval. They argue that the activities inherent in the search (verbalization of the information need, division of the problem into smaller topics, selection of sources, and use of search tools) build individual IISs. It should be mentioned that information retrieval through visualization, which plays a role in the interface, belongs to the scope of KO problematics (Raghavan et al. 2015; Batorowska and Kaminska-Czubala 2014).

The most comprehensive descriptions of IIS include the definitions proposed by Jones and Detlor. The Jones IIS concept refers to both the practice and the study of the individuals' activities undertaken to acquire or create, store, organize, conduct, search, use, and disseminate information needed to perform tasks and fulfill different roles. IIS is, in particular, organizing and maintaining individual information collections in which pieces of information, such as paper, electronic documents, e-mails, web pages, and handwritten notes are stored for reuse at a later time. Detlor draws attention to the fact that the individual perspective in information space refers to information needed in everyday life or work. IIS, according to him, is a practice and theory related to the activities undertaken to acquire, organize, archive, and search for information.

The richest (for the state of knowledge in 2017) semantic IIS description brings the latest work by Gorny, Glowacka, Kisilowska and Osinski. They develop an extended concept of IIS, the main assumption being the belief that every person undertaking information activities functions always and exclusively within their IIS. The shape of any individual IIS is not constant. It can change under the influence of many factors. Particular elements of IIS are characterized by a different degree of vulnerability. According to these researchers, IIS is a space of both a physical and mental nature, associated with the acquisition of a specific type of information and its dissemination: information assigned to a given user. In the mental sphere, IIS is a collection of knowledge, experience, and skills in the field of information activities, ways of perceiving the surrounding information infrastructure, along with attitudes, prejudices, and beliefs related to information activities that reveal themselves during the implementation of the information task. In the physical sphere, it is a collection of information resources in various forms, systems, and tools used to present the information available to the user. IIS is a methodological construct that includes information needs, information infrastructure, characteristics, and user information features. However, it does not include information activities of the information the user undertakes. But, because they are the result of a specific shape of an IIS, on their basis one can, to a certain extent, make conclusions about the properties of the IIS, as well as about the mechanisms of processes occurring within the IIS. In terms of methodology, IIS is included in the scheme, which consists of three parts:

a) the actual availability of the information infrastructure for a particular user;
b) the set of skills and information experiences of that user; and,
c) its information awareness team (Gorny et al. 2017).

3.0 Research outline

Accepting from Jones the assumption that IIS refers to both the practice and the study of the individual's activity, the authors designed an application called Scientific Visualizer (http://visualizeme.umk.pl/), which allows the mapping of activity as well as the user's ability to manage information from a long-term perspective. The visualizer, oriented toward individual researchers, gives them an insight into their own scientific productivity and thus creates the opportunity to move forward based on the acquired knowledge from the analysis and their own information needs.

In order to estimate the usability frontiers and potential of the presented tool, it was decided to define potential users and make applicability tests on them. Authors' ICT experience and information literacy indicated that the best results would be gained from the selection of two main groups: 1) researchers interested in the visualization of their own scientific output, and 2) evaluators who, due to Visualizer, can appraise its functionality and effectiviness to evaluate the scientific achievements of others.

With regard to the main goal of the research, and thus the attempt to determine the role of the visualizer in IIS, the two stages of the research procedure are marked out:

Stage 1 (subject of this article):

1. Selecting a group of researchers to test the application on the basis of their publication records;
2. Preparing a questionnaire to find the degree of interest in the designed application;
3. The researchers (respondents) formulating opinions on the usability and functionality of the application; and,
4. The authors classifying opinions and formulating conclusions regarding the necessary changes to be implemented in the application.

Stage 2 (subject of the planned second part of the text):

1. Making the modifications in the application;
2. Selecting (based on the researchers' answers) a group of evaluators who test the applicability of the Visualizer and assess its action;
3. Referring to a group of evaluators, aimed at identifying groups of researchers for whom the Visualizer would be most useful, defining potential areas of its application and irregularities in functioning; and,
4. Delimiting directions for future work to improve the application.

4.0 Data

Research data derive from the Expertus bibliographic database (http://bg.cm.umk.pl/splendor/umk/), which records the scientific output of employees of the Nicolaus Copernicus University in Toruń (NCU). The database collection covers the bibliography of scientific publications in various disciplines. The database platform Splendor provides preparation for detailed bibliometric analyses: the entire institution, individual organizational units, and any author ("Expertus" n.d.).

The Expertus system allows one to register bibliographic descriptions of any type of publication (e.g., article, paper, message, chapter-fragment, monograph, script, patent, report, leaflet, CD-ROM, electronic document, etc.). The basic elements of the bibliographic description are: author, title of publication, title of source, date of publication, page numbers, language, formal type of publication (magazine, journal article, book, folder, electronic document, calendar, PhD/habilitation, map, patent, fragment from a collective position), substantive type of publication (including scientific article, popular science article, scientific book, popular book, volume edition, source text, translation, communication, academic textbook, computer program, report, interview, review, expert opinion, etc.), affiliation of publications, Polish keywords, English keywords, information on the availability of online, assessment measures (like impact factor, ISI Master Journal List, Ministry of Science, and Higher Education score). Searching is carried out in several ways, including the use of simple queries, formulation of queries with Boolean operators, and advanced search. The system allows the display of results in several formats (including "print version") and export of data to Word, OpenOffice editors, etc. (Głębocki and Pietruszewski 2010).

The choice of the Expertus system for testing was not accidental. In addition to the versatility in terms of the hardware platform and operating system, it was also chosen because its implementation in several Polish universities and research centers.[1] In addition to local scientific databases created according to the Expertus system, in Poland there is also a national database of publishing output, the Polish Scientific Bibliography (Polska Biblbiografia Naukowa 2017). Data can be introduced both by the authors (personal scientific bibliography) and by scientific units, thus creating a scientific bibliography of an individual (institutional bibliography). Data added by the authors collected in the open part of the system called the "repository module" allow the researcher to create his or her own research portfolio. Because entering data into the repository module is voluntary, not all scientists use this type of self-presentation. Some scientists who are already present in the database complement their achievements in a selective manner (e.g., they only add books and recent publications, bypassing those from many years ago, or add only their current affiliation) and non-systematic (e.g., once in a year, once every few years). Including this database in research on the visualization of the output, despite the fact that according to the assumption of its owner, The Ministry of Science and Higher Education, that it plays the role of a central database, it would, therefore, be burdened with high risk of including incomplete data in surveys.

As data are cyclically indexed by Expertus (the obligation to provide publications is imposed on the academic staff by the university's authorities), they seem to represent achievements accurately in terms of completeness. Because the creators of the PBN system undertake numerous attempts to cooperate with local centers in sharing bibliographic data, and, therefore, are also based on the Expertus system, conducting the current study seems to be not only reasonable but necessary.

5.0 Research arrangement

In the first stage of research with the Expertus system, sample bibliographic data for a randomly selected researcher were collected. Then, these data were primarily visualized in the prepared application. Depending on type (numeric, textual, or geographical location) the data were presented using the appropriate type of chart or map. Prior to surveying, the interface's errors and glitches were corrected.

The sampling of the respondents was based on the electronic database of employees of Nicolaus Copernicus University in Toruń (https://spis-pracownikow.umk.pl/). Since it was assumed that the most representative results would be obtained by covering the representatives of all fields of science, i.e., social sciences, humanities, and natural and technical sciences, the organisational units of the university that conduct research in the fields of library and information science, sociology, pedagogics, foreign languages, history and archival science, philosophy, physics, mathematics, and computer science were selected from the aforementioned database. Then, on the basis of data published on the websites of particular institutes and chairs (names of employees and lists of their achievements), employees were selected for the research. The sampling of the aforementioned disciplines and their representatives was purposive. The research covered representatives of those disciplines whose research interests focus around various areas of information: its presentation, conceptualization of concepts, reception, quality, structure, and visualization. Therefore, it was assumed that the researchers representing these disciplines would be the most frequent potential future users of the tested application. One assumption of using a nonprobability sample was that their opinions might constitute the most valuable source of information about the studied application (for more information on nonprobability sampling, see Babbie 2007, 205).

In the end, twenty-five researchers—research workers of various disciplines and specialities employed at the NCU—were included in the research. Due to the fact that the present research was to primarily feature an exploratory dimension (collecting qualitative material sufficient to design further extended studies addressed to professionals), it was not necessary to follow the opinions of the entire academic community of the University. On the other hand, due to the pilot nature of the research (the research will be continued on a larger number of people and with a larger data set), at this stage it was not possible—for formal, organizational, and timeframe reasons—to cover the research staff of over 2,500 (as of 2017).

In order to gather the most diverse group of respondents possible, the sample also included people holding the degrees of master, doctor, doctor *Habilitatus* (*doktor habilitowany* in Polish—the highest academic degree in Poland, confirming the scientific independence of a researcher, after which one can apply for the title of professor), as well as professors. The sampling of the respondents was not based on the biological age of the researchers, but on their productivity in science. It is worth stressing that it was not the aim of the research to determine the behaviours and opinions related to the tested visualizer by age groups. It became much more important to capture the differences in the assessment of the usability of this tool for persons who can demonstrate significant scientific achievements and those just starting their scientific careers: persons representing soft sciences and those representing hard sciences. It was assumed that the approach to the application of people with higher scientific achievements would be different from that of those with lower number of publications. Other application evaluations were also expected from the representatives of the humanities and technical sciences. Since age of the researcher was not a study criterion, in the future, after collecting a larger amount of empirical material, it seems justified to undertake research aimed at learning the opinions and behaviours of the researchers also with this criterion applied.

For the people selected in the Expertus database, ASCII files covering publication records were generated (scholar's portfolio). The files together with a questionnaire (see Appendix) were sent to respondents by e-mail with a request to upload them to the interface and to evaluate the various forms of visualization. Due to the low response, after a month the request to test the Visualizer and fill the questionnaire was repeated.

The questionnaire was composed of seven closed and two open questions. The first question concerned the assessment of the completeness of the records downloaded from the Expertus database compared with the current lists of publications owned by the authors. As it turned out, the statement of completeness of records or its lack determined the researchers' attitudes (bigger or smaller involvement) for further data analysis.

Because the most important goal was to determine the functions that the application could perform in an academic environment, the second question asked respondents to identify possible application areas. To help researchers formulate their opinions, a few suggestions are included, but it also leaves the possibility of giving one's own answer. It is not a coincidence that the question about the functionality of the application was placed at the beginning of the questionnaire. In the authors' assumption, it was intended to show the researchers the scale of possible applications and to encourage them to look for new, non-trivial ones. Another question, the knowledge that can be gained from publishing activity visualization, was to encourage researchers to discover irregularities in the Visualizer's operation, and thus contribute to identifying new, potential areas of its application.

Two more questions opened the door to direct reflection on respondents' own scientific achievements. The fourth question concerned possible inspirations for introducing changes in publishing activity. The authors of the questionnaire wanted the respondents to consider whether the visualization of the output could be an impetus to undertake publishing activity in new or different sources of publishing than hitherto. Again, to help respondents formulate opinions, the question was accompanied by examples of manifestations of such activities, and respondents were asked to choose those most appropriate. In answering this question, a helpful tool was a map of the geographical distribution of publishing places. The fifth question was the same, but this time it was interested in whether the visualized output could become an inspiration for expanding research interests. Once again, the answer to the question was indicated: the dimensions of inspiration, taking up related subjects in the area of research already being carried out, looking for new research topics, establishing a cooperation network, acquiring foreign contacts, and acquiring new sources of financing. The set of expected answers to this question was to provide an idea of how NCU scientists approach individual knowledge management of their scientific achievements: whether they have developed systematics or a strategy in this area and whether they intend to change them under the influence of visualized information.

The sixth and seventh questions were to verify whether the researchers are interested in the achievements of other scientists, whether they conduct comparisons of their own scientific achievements with others, and who the object of their interests is and for what reasons. In this case, the idea was to encourage researchers in some way to reflect on possible directions of changes in the application's operation, and more strictly on the sense of making it available to a wider audience and creating a universal opportunity to visualize not only personal achievements but also other people's achievements.

The last two questions (questions 8 and 9) concerned the usability of the application interface. The first one asked which chart, according to the respondents, was the most useful, while the second one asked to help complete the list of comments to improve the operation of the Visualizer (detection of errors and gaps, indicating complementary elements and modifications).

The responses and opinions collected based on the questionnaire were to indicate not only the directions of changes necessary to implement the application but also to help to select an appropriate group of evaluators that is necessary for the second stage of research. After collecting the survey data from the researchers, a list of future modifications of the application was prepared by submitting it for testing and assesment by the evaluators.

6.0 Interface

At the foundation of the Scientific Visualizer was an assumption that research visualization helps scientists in seeking answers to questions about the cooperation of academic groups, the development of a chosen research discipline, and its place in relation to other similar disciplines. Finally, it encourages, especially for those interested in the broadly understood history of science and technology, the analyzed research fields to be placed on the global science map (Borner 2015; Osinska 2016).

For these reasons, the Visualizer's interface (Figure 1) was built primarily from six visualization layouts. Additionally, it is equipped with the main menu and the "Upload" button used to load the RTF text file with bibliographic data.[2]

To visualize data that are loaded into memory from the MongoDB database system, the application uses an exten-

Figure 1. Screenshot of Visualizer's interface
Source: Scientific Output Visualizer (online), accessed December 12, 2017. http://visualizeme.umk.pl/en/.

sive library of D3.js scripts,[3] and to indicate the geographical location it uses the Google Maps API application interface. Displaying the total number of works published by a given author is possible using a bar chart (layout 1). All documents (articles, reviews, books) are ordered on a timeline according to the year of publication. Such a solution allows a holistic view of the dynamics of a scientist's own productivity and captures the changes taking place within it.

Comprehensive knowledge about cooperation is provided by graphs and maps, which are included in layouts 2 and 3. While the graph shows cooperation networks, or more precisely co-authors of the publications and co-editors of collective monographs and periodicals, the map indicates the geographical location of publications. The display area is automatically scaled to the geographical scope of recorded locations. The national and international cooperation was revealed through the network visualizations, which reflected the prevailing tendencies in science toward group scholarly activity and a greater appreciation of both global and local evaluation systems by authors from various scientific centers (internationalization).

The next ring chart presents a map of domains (layout 4). It shows in which disciplines the particular researcher publishes. To perform this task, it is necessary to verify (e.g., on the basis of Web of Science indexes) the lists of scientific journals assigned to particular disciplines, and in the case of monographs, to determine their main subject and subsequently assign them to the areas of knowledge. However, the difficulty in categorizing the scientific activity of a single researcher lies in the fact that increasingly, publications can be attributed to not one but a few equivalent disciplines. The problem in this context is, therefore, the proper geometry of arranging objects (articles) on the domain map (for more on the geometry of placement of objects, see Osinska 2016).

Because in Polish bibliometric systems, the number of characters (rather than the number of words) has an evaluation value, it is converted directly into publishing sheets, and those convert to Ministerial points. The next chart, the bubble chart (layout 5), illustrates the total number of pages of all papers published in a given calendar year. Below the bubble chart, a line chart shows the cumulative number of Ministerial points awarded in a given year. Such a combination is not accidental. In the face of the "disease" that affects Polish academics known as "point disease" (Kulczycki 2017), these two charts clearly show that neither the number of pages, nor even more predetermined points, can be a measure of the quality of a scientific achievement.

The last graph proposed in the Visualizer is a word cloud (layout 6). This type of visualization allows a separate combination of Polish and English keywords, and thus a comparison of the subjects taken by a single researcher in the pages of local and international journals.

In addition to visualizing, the interface allows users to generate a PDF file for each graph separately. In addition, it also gives users the opportunity to display statistics next to the charts, which relate to the numerical values of the aggregated data (e.g., summary number of papers, number of co-authors, average number of publications, number of monographs, etc.). These statistics may be more comprehendable for some users than graphical representations.

In addition to the built-in layouts and visualization, the Scientific Visualizer is armed with a main menu. Individual tabs inform users of the purpose of the application, team characteristics, and accepted types of data being processed. To make the interface user friendly, a sample file is also included for visual analysis and a description of the data structure, along with some tips on how to prepare the input data. The users' tutorial was attached to the application; it includes the steps to prepare a bibliographic file from the Expertus system and application functions focused on data analysis (http://visualizeme.umk.pl/examples/tutorial.pdf).

7.0 Results

As already mentioned, this research was based on a nonprobability sample. Despite previous declarations or expressions of willingness to participate in research, ultimately, of the twenty-five people selected, twenty-two sent their opinions. Ten women and twelve men were surveyed. Five persons held the title of professor, five held postdoctoral degrees, eight doctoral degrees, and four master's degrees. The final group comprised representatives of library and information science, archival science, history, German philology, philosophy, pedagogics, physics, mathematics, and computer science.

The e-mailed responses were aggregated for each question. As the questions were multiple-choice, in many cases the number of responses did not add up to twenty-two (100% of the respondents who replied to the questionnaires). The statistical data were supplemented with verbal descriptions and additional observations and statements of the respondents. Due to the small sample, the results were interpreted collectively, without referring to partial results for each of the groups (differentiated by gender, academic rank, and discipline). As the age of the researcher was not a study criterion, it was not subject to analysis in the evaluation of the results. Due to the small amount of empirical material, the following discussion of the study results is not illustrated by tables or diagrams. In this case there is no need to calculate statistical significance.

Answers provided allowed the authors to formulate the following observations (according to the order in which the questions were asked in the survey: see Appendix):

1. Most people (sixteen) confirmed completeness of data downloaded from the Expertus database. Six people reported shortcomings, pointing out that the list contained few recent works and contained errors in the bibliographic descriptions. This may indicate that either the Expertus system correctly included all scientific publications for majority of authors (which seems unlikely after random verification), or that these people did not make any effort to verify the original data source.
2. Opinions formulated by researchers prove that a visualizer can be used in several functions: serve as a tool for assessing one's own publication activity (ten indications), analyzing its directions (seven indications), discovering some regularities in it (three indications), and discovering delays in records reporting as well as bug detection in the Expertus system (three indications). These results are shown in Figure 2. Only two persons proposed items other than what was suggested in the answers. The physics representative said that the application of this kind of tool can be compared with the ways researchers in various disciplines operate in science. In turn, the representative of archival science raised the psychological dimension of data visualization, i.e., its impact on raising or lowering self-esteem.
3. The collected answers indicate that the most useful format was the map of areas (layout 4); nine people pointed it out (Figure 2). The others rank accordingly: a word cloud (layout 6), a graph of publication activity (layout 1) and a bubble chart (layout 5). Such results prove that the Visualizer allowed the respondents to first and foremost determine the place of their own achievements on the predefined science map and outline the periods of highest and lowest scientific productivity.
4. Three of the respondents, thanks to the visualization of data, have discovered some regularities of which they had not previously been aware. The representative of archival science noticed that as a result of visualization, the publications showed the dominance of problems different from what he expected. A physicist realized that he publishes within the common area for both natural sciences and humanities, and a computer scientist found that he publishes little in traditional printed format.
5. The potential of the Visualizer in forecasting the directions of research interests was noticed only by researchers positioned at the interface between different sciences (interdisciplinary fields). These researchers most often indicated that the results of data visualization inspire them to initiate new research topics (seven indications, Figure 2), apply for research grants for current research (six indications), and make decisions on the need to extend research cooperation on a national or international scale (four indications). Three people mentioned taking up topics related to those in the area of research they have already conducted. Researchers focused on narrow specializations (for example, bibliology and informatics) were not interested in changing or broadening their interests towards other fields. For one person, visualization has become an inspiration to undertake research on communicativeness of keywords for publications in Polish and English. Two other people, however, stated that the application could not play any role in the analysis or planning of scientific development.
6. Regarding the impact of data visualization on increasing publication activity, the most frequent answers were the attempts to publish articles in national journals (seven indications) and international (six indications), the desire to increase conference activity (four indications), the attempt to publish Open Access (OA journals, digital repositories, first online, etc.) (four indications), and increase in the number of monographs (three indications). None of the people indicated that they would try to develop collective publishing.
7. Studies have shown that if the application offered the possibility of comparing one's own scientific achievements with other researchers, the majority of respondents would be interested in such a possibility towards verification of colleagues from other academic centers and other countries (nine indications), e.g., in terms of the total number, types and places of publishing. Only two people would be interested in the possibility of compiling their own publication activity with colleagues from their alma mater. One person would like to have the chance to compare their achievements with representatives of other completely different fields of science, particularly, the time spent in the development of a monograph. Surprisingly, none of the researchers would be interested in the visual compilation of their own achievements to compare with the achievements of direct superiors or university authorities. This proves that scientists prefer to make comparisons with peers in the hierarchy, that is, colleagues, which may be caused by the willingness to assess opportunities in the field of gaining new career levels or simple curiosity and rivalry.
8. Most researchers (fourteen people) agreed that in the future a Visualizer could be a helpful tool in supporting the process of selecting reviewers in promotion proceedings, applying for subsequent academic degrees, and selecting experts for evaluating grants, as well as supporting superiors in their periodic evaluations of employees.

In addition to assessing the usefulness of the application to visualize one's own achievements, the respondents

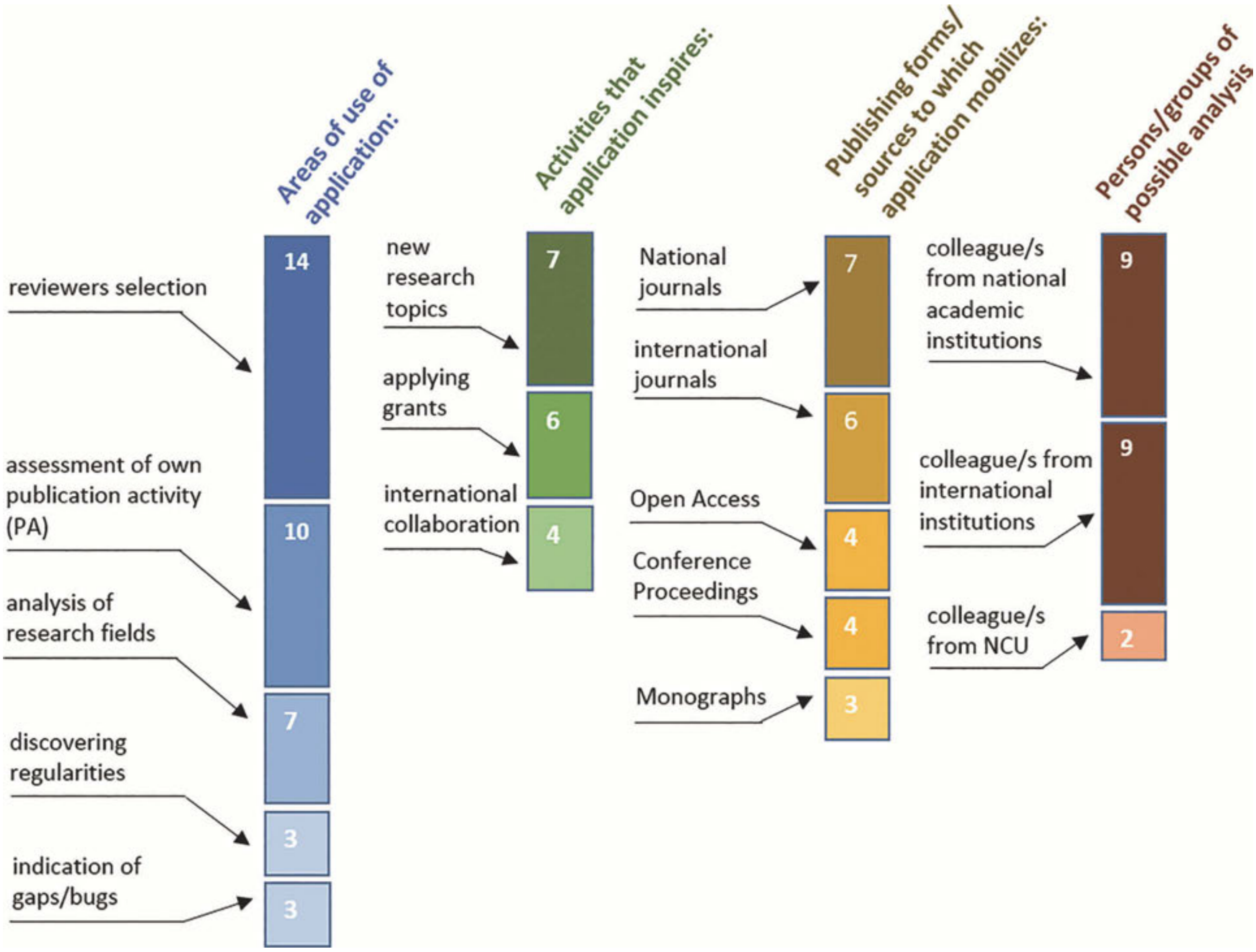


Figure 2. Infographic presenting some surveys results.

should also evaluate its functionality. In terms of improving user interface, respondents suggested the following improvements: to display the publication titles on a bar chart, to implement a function to upload a file within a permanent link with the university database instead of with a button. Other ideas concerned the magnification of the Ministerial score chart as an important element in assessement, better readability of the domain maps and word cloud, changing the counting algorithm to the total number of pages, and separation of document types. It was also proposed to supplement the visualization with data on conference activity and to present participation in research projects.

8.0 Discussion and conclusions

The proposed approach demonstrates that the visualization of scientific achievements can have a significant impact on the formation of IIS of individual researchers and thus on the performance of various tasks and fulfilment of different roles. Visualization of writing activity dynamics, the time and place of publication, the thematic scope of research problems and their importance in the whole of scientific output, and, finally, the establishment of cooperation networks can undoubtedly influence the perception of various relationships, dependencies, and regularities in one's own scientific activity, and consequently, a better reflection in the researcher's mind about his or her scholarly achievements. Only proper consideration of the scientific reality, and more broadly, building information awareness, can be the bases for making rational decisions in various areas of professional activity, especially those parts of it that directly relate to the area of information (acquiring, collecting, processing, verifying, publishing, presenting, and sharing). It can also lead to broadening one's own information competencies, to use a wider part of the science information infrastructure and to create new information needs. Confronting one's own scientific achievements with the achievements of other researchers, thanks to the visualizations, may lead to extensive study, not only on the problem of the quality of output, but also on the structure, size, and properties of one's own IIS in comparison with IIS of other researchers to help spot deficiencies within it.

According to the specialists on problematics in information visualization (Chen 2017; Tufte 2001; Ware 2004), visualization, thanks to the synthesis, integration, and concentration of information, serves the condensed transmission of knowledge, multiplying its interpretation potential in comparison with the textual message.

Visualizations in certain cases more effectively than text affect our brain, make complex data sets easier to understand, and allow for their effective interpretation. The results of the research confirm that presenting information in a striking and innovative form attracts the user's attention, sustains the engagement of the recipient, allows him to notice something that he would not otherwise notice, and facilitates memorization of information.

Designed by the authors, the Visualizer can certainly significantly enrich the individual information space of the contemporary researcher. It can facilitate the evaluation of the publication activity and raise awareness of the need to monitor the indexation of bibliographic data in selected databases. The proposed application can also help discover the relationship between research issues, inspire researchers to broaden horizons and cooperation networks, and also be a tool supporting administrative activities (employee evaluation, promotion proceedings, accreditation, appointment of experts, and separation of funds).

According to the respondents' suggestions, the application requires some functional improvements, such as a counting algorithm of total number of pages or improving the readability of the domain map. From the designers' perspectives, it will also be necessary to prepare tutorials for users who have little or no knowledge about visual language. The development of tips and samples of data will certainly facilitate the reading of complicated and non-traditional graphs.

The users suggested changing the "Upload" button into a permanent link to the university database. They consider the function of loading bibliographic records to be a psychological barrier in initiating extended analysis. But in the authors' opinions, it is worth keeping unchanged as an interaction mechanism. It seems that it compels the user to take the trouble to filter Expertus data and quickly review them, and thus stimulates curiosity through a series of questions that naturally arise in the space of the mind. Although the automation of this process would certainly facilitate access to data, on the other hand, as research conducted on software usability shows (Tufte 1977), it would definitely reduce cognitive curiosity.

It is worth noting that some of the scholars have reluctantly proceeded with the visual analysis of their own output. One can only presume that it was related either to the general reluctance to participate in this type of research or the unrepresentative nature of scientific achievements. Unsatisfied responses to this research may suggest that this type of research should be carried out in direct contact with scientists (e.g., in a laboratory group). It seems, however, that in this case the effect would not be satisfactory. The general comparison of scientific achievements of various researchers could be too controversial for some, if not painful, and risks making their effects too public.

Because the database created in the Expertus system is not devoid of errors and gaps in the bibliographic descriptions (which was discovered thanks to the prepared Visualizer), and due to the fact that it does not give the possibility of graphical presentation of the output taking into account criteria other than the basic formal criteria (total number of works, types of publications), the prepared application is a good way to supplement or extend the solutions proposed there. In addition to its cognitive and utilitarian values, it can perform integrative functions in the scientific community, i.e., initiate discussions and encourage the exchange of views or encourage closer and further cooperation. Is this not the core of visualization?

Endnotes

1. Unfortunately, Expertus's interface has been designed for Polish researchers, although it allows manipulating by English keywords and terms from publication titles. Expertus is implemented by Splendor (Expertus n.d.). There is also the possiblility to test a demo database and see how the filters function (http://www.splendor.net.pl/demo/).
2. By clicking on the "data" option in the main menu (http://visualizeme.umk.pl/info/data), we are able to find a link to sample files packed into on archive file and test them through uploading the current application. We also improved information systems about the application; we published a detailed tutorial in Polish and added more options to manipulate the data as well as provided a detailed description of each layout.
3. D3.js (Data-Driven Documents) is a library in JavaScript designed for creating advanced data visualizations in various internet applications. It allows not only generation of charts, but also allows users to create maps, interactive diagrams, and control panels for data, reports, etc.

References

Babbie, Earl. 2007. *The Practice of Social Research.* 11th ed. Belmont, CA Wadsworth Cangage Learning.

Batorowska, Hanna, and Barbara Kaminska-Czubala. 2014. "Information Retrieval Support: Visualisation of the Information Space of a Document." In *Knowledge*

Organization in the 21st Century: Between Historical Patterns and Future Prospects. Proceedings of the Thirteenth International ISKO Conference 19-22 May 2014, Kraków, Poland, ed. Wieslaw Babik. Advances in Knowledge Organization 14. Würzburg: Ergon, 343-50.

Boisot, Max H. 1995. *Information Space. A Framework for Learning in Organizations, Institutions and Culture*. Organizations: Theory & Behaviour 2. London: Routledge.

Börner, Katy. 2010. *Atlas of Science: Visualizing What We Know*. Cambridge, MA: MIT Press.

Börner, Katy. 2015. *Atlas of Knowledge: ANyone Can Map*. MA, Cambridge: MIT Press.

Buchel, Olha and Kamran Sedig. 2011. "Extending Map-Based Visualizations to Support Visual Tasks: The Role of Ontological Properties." *Knowledge Organization* 38: 204–29.

Chen, Chaomei. 2017. *Mapping Scientific Frontiers: The Quest for Knowledge*. 2nd ed. softcover reprint. London: Springer.

Davenport, Thomas H. 1997. *Information ecology: Mastering the Information and Knowledge Environment*. New York: Oxford University Press.

Dillon, Andrew. 2000. "Spatial Semantics and Individual Differences in the Perception of Shape in Information Space." *Journal of the American Society for Information Science* 51: 521-8.

Expertus – "Review of Application." (n.d.) Splendor® Information Systems. http://www.splendor.net.pl/english.html

Garfield, Eugene. 1994. "Scientography: Mapping the Tracks of Science." *Current Contents: Social & Behavioural Sciences* 7, no. 45: 5-10.

Glębocki, Jacek, and Grzegorz Pietruszewski. 2010. "Przeglad zastosowan systemu Expertus w Polsce i za granica." *Forum bibliotek medycznych* 3, no. 1: 352-80.

Gorny, Miroslaw, Ewa Glowacka, Malgorzata Kisilowska, and Zbigniew Osinski. 2017. *Mechanisms of the Formation and Evolution of Personal Information Spaces in the Humanities*, trans. John Catlow. Poznań: Wydawnictwo Rys.

Gwizdka, Jacek. 2006. "Finding to Keep and Organize: Personal Information Collections as Context." Paper presented at Personal Information Management. A SIGIR 2006 Workshop. August 10-11, 2006, Seattle, Washington. http://pim.ischool.washington.edu/pim06/files/gwizdka-paper.pdf

Hjørland, Birger. 2017. "Domain Analysis." *Knowledge Organization* 44: 436-64.

Kisilowska, Małgorzata. 2011. "Przestrzen informacyjna jako termin informatologiczny." *Zagadnienia informacji naukowej* no. 2: 35-52.

Kulczycki, Emanuel. 2017. "Punktoza jako strategia w grze parametrycznej w Polsce." *Nauka i szkolnictwo wyzsze* 1, no. 49: 63-78. doi:10.14746/nisw.2017.1.4

López-Huertas. María J. 2015. "Domain Analysis for Interdisciplinary Knowledge Domains." *Knowledge Organization* 42: 570-80.

McKnight, Cliff. 2000. "The Personal Construction of Information Space." *Journal of the American Society for Information Science* 51: 730-3.

"The Meaning and Origin of the Expression: A Picture is Worth a Thousand Words." 2017. *The Phrase Finder*. http://www.phrases.org.uk/meanings/a-picture-is-worth-a-thousand-words.html

Newman, M. E. J. 2010. *Networks: An Introduction*. Oxford: Oxford University Press.

Osinska, Veslava and Piotr Bala. 2010. "New Methods for Visualization and Improvement of Classification Schemes: The Case of Computer Science." *Knowledge Organization* 37: 157-72.

Polska Bibliografia Naukowa – Pomoc Systemu. 2017. https://polon.nauka.gov.pl/help_pbn/doku.php

Raghavan, K. S., K. H. Apoorva, and Jivrajani Aarti. 2015. "Information Retrieval as a Domain: Visualizations Based on Two Data Sets." *Knowledge Organization* 42: 591-601.

Sauermann, Leo, Ludger van Elst, and Andreas Dengel. 2007. "PIMO – a Framework for Representing Personal Information Models." In *Proceedings of I-Semantics' 07, JUCS*, ed. Tassilo Pellegrini. Graz: Verl. der Techn. Univ. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.162.84

Shemberko, L. V., and M. A. Borovik. 2012. "Information Resources for Social Sciences and Humanities: Support for Research on Innovative Development in Russia." *Scientific and Technical Information Processing* 39: 153–63.

Skórka, Stanislaw. 2011. "Architekt informacji – Tworca Przestrzeni informacyjnej." *Przeglad biblioteczny* 79, no. 1: 47-61.

Smiraglia, Richard P. 2015a. *Domain Analysis for Knowledge Organization: Tools for Ontology Extraction*. Amsterdam: Chandos/Elsevier.

Smiraglia. Richard P. 2015b. "Domain Analysis of Domain Analysis for Knowledge Organization: Observations on an Emergent Methodological Cluster." *Knowledge Organization* 42: 602-12.

Tufte, Edward R. 2001. *The Visual Display of Quantitative Information*. 2nd ed. Cheshire, CT: Graphics Press.

Ware, Colin. 2004. *Information Visualization: Perception for Design*. 2nd ed. San Francisco: Morgan Kaufman.

Appendix

Questionary for Authors

1. **Is the list of author's publications obtained from Expertus database complete?**
2. **In your opinion, what is the use of presented application:**
   a) assessment of own publication activity,
   b) analysis of research fields,
   c) discovering regularities and gaps in the publication activity,
   d) indication of irregularities (i.e. delays and gaps) in Expertus system,
   e) other ....
3. **Have you discovered any regularity, that you have not previously been aware of? If so, please specify?**
   a) articles in national magazines,
   b) publishing articles in foreign journals,
   c) increasing in conference activity,
   d) Open Access publishing (OA journals, digital repositories, first online etc.),
   e) developing a collective publishing,
   f) increasing the number of monographs,
   g) other ....
4. **Could visualized output be an inspiration to expand research interests? If so, please mark the possible dimensions:**
   a) undertaking the topics related to the area of already carried out research,
   b) undertaking new research topics,
   c) applying for research grants
   d) making a decision on the need to extend research cooperation in a national or international dimension,
   e) other: ....
5. **Knowing that the application gave you the opportunity to view the scientific achievements of other researchers, would you use this function? If so, please specify who could be the target of this study:**
   a) the Rector /Dean
   b) the Director of the Institute /Department
   c) a colleague,
   d) a colleague from another national academic institution,
   e) researchers from other countries,
   f) other ....
6. **What would be the purpose of this study:**
   a) willingness to compare own achievements with the achievements of other researchers,
   b) willingness to check the fields of research carried out by other researchers,
   c) willingness to establish research cooperation,
   d) willingness to verify/support reviewers appointed in promotion,
   e) other ....
7. **Which graph turned out to be the most useful?**
8. **What would you change or add in the tested application?**