THE USAGE OF INFORMATION RETRIEVAL TOOLS: DIFFERENCES BETWEEN LITHUANIAN AND PORTUGUESE STUDENTS

Rita Jakutienė
Šiauliai State College
Lithuania

Daiva Civilkienė
Šiauliai State College
Lithuania

Annotation

With the development of information technologies, more and more information retrieval tools, containing a wide variety of information, can be found. Internet resources, universal and specialized search engines, virtual libraries are becoming the main information retrieval tools for students. The article introduces an exploratory (pilot) research aiming to find out the differences of information retrieval between Lithuanian and Portuguese students. Šiauliai State College offers a unique study programme Information Management, having a signed contract of a double diploma with Porto (Portugal) Polytechnic Institute Porto Accounting and Business School (ISCAP). 10 Portuguese and 10 Lithuanian students participated in the research. The online questionnaire was forwarded to the students of ISCAP Porto Accounting and Business School and to the second year students of the study programme of Information Management at Šiauliai State College. The research revealed that the respondents Lithuanian students look for scientific papers and books, whereas Portuguese students search for basic information and scientific papers. All the respondents of the survey identified Google Scholar as the most important search engine for scientific information retrieval.

Key words: information retrieval, information retrieval tools, information retrieval systems, Portuguese students, Lithuanian students.

Introduction

The Relevance of the Topic

It is becoming more complicated to find the way to discover the necessary source of information and not to get lost in abundance of information sources and information retrieval systems due to increasing amount of information. Target information retrieval needs more time, moreover, appropriate information retrieval skills are needed. Information retrieval is generally carried out intuitively; it naturally flows into the process of scientific research and provision of its results. Information can be obtained hardly following any rules, simply trying out different retrieval tools and following personal information retrieval experience. The more information sources and retrieval methods, suitable for science and studies are known, the easier it is to select them, following own needs and find appropriate information. Information retrieval is generally described as an activity that deals with information retrieval systems for obtaining of appropriate information, during which reconciliation of user's information needs to data objects take place. Generally, a user, when solving an information problem, selects familiar and useful information resources, proven in the past. Information retrieval is at the core of our daily lives. Modern search, ranking and indexing systems underpinned by enhanced computing power, fast network speeds and near unlimited data storage capacity mean that we have easy access to all the information we need, when we need it (Harman, 2019).

Information retrieval is a constituent part of user's information behaviour. Theoretical issues of information user behaviour, information retrieval tools, choice of sources, motives, influencing the quality of information search, were researched by T.D. Wilson (2000); R. Fidel, A. M. Pejtersen (2004); P. Vakkari, K. Järvelin (2005); A. Spink, C.B. Cole (2006); K. Järvelin (2011), R. Savolainen (2016, 2018). The dominance of the Internet search engines in students' information search strategy and their influence on the quality of studies was analysed by J.R.Griffiths and P.Brophy (2005), H. Weber, S. Hillmert, K.J. Rott (2018); H. Weber, D. Becker, S. Hillmert (2019).

Lithuanian scientific literature provides only a few researches about students' information behaviour (Janiūnienė, 2012), parliament members' information behaviour (Vernickaitė, 2014);

there are several master degree's theses dealing with information behaviour topic. Information retrieval problems have not been analysed at Šiauliai State College.

Research problem

Students may seek information for various reasons: to understand a specific subject matter or to conduct a research. Knowledge is acquired through the use of various information retrieval tools or sources, so their selection and knowledge of how it works is important for the information user. The topic is relevant to the fact, that the first pilot research has been carried out, trying to single out the differences of information retrieval tools usage, between Portuguese and Lithuanian students, as a unique study programme *Information Management* is being pursued at Šiauliai State College, and the students of the programme, following the contract of double diploma, are provided an opportunity to study both in Lithuania and Portugal during the same study period, and are awarded two diplomas – of Šiauliai State College and Porto (Portugal) Polytechnic Institute Porto Accounting and Business School (ISCAP). This academic year two students from Portugal have come to study a part of their degree in this study programme. During their Information Behaviour traineeship, they analysed the way Portuguese and Lithuanian students retrieve information, what information they look for and what retrieval tools they employ.

The problem of the research is described by the problem question – the differences in the use (choice) of information retrieval tools and sources between Lithuanian and Portuguese students

The object of the research is information retrieval tools

The aim of the research is to identify the differences in information retrieval tools usage between Lithuanian and Portuguese students.

The objectives of the research:

1. Describe information retrieval concept and tools.
2. Analyse the differences of information retrieval tools usage between Lithuanian and Portuguese students.

The research methods are the analysis of scientific literature, quantitative research (survey).

The Concept of Information Retrieval

The term 'information retrieval' was coined in 1952 and gained popularity in the research community from 1961 onwards. It describes the process of retrieving information from an organized (indexed) collection of information resources (Chowdhury, 2010). Today this process is most familiarly carried out by search engines, the tools, used to retrieve information from the web. Information retrieval systems originally meant text retrieval systems, since they were designed to deal with textual documents. However, modern information retrieval systems deal with multimedia information, comprising text, audio, images and video. While many features of conventional text retrieval systems are equally applicable to multimedia information retrieval, the specific nature of audio, image and video information has called for the development of new tools and techniques for information retrieval. Broadly speaking, information retrieval encompasses all the activities related to the organization, processing of and access to information resources of all forms and formats (Chowdhury and other, 2007). Information access includes all the typical information retrieval processes and activities ranging from content and data selection, and processing and indexing, to search and retrieval, and use of information and data by a designated user community in order to meet their information requirements (Chowdhury, Foo, 2012). K. R. Chowdhary (2020) states that information retrieval is the identification of documents or other units of information in a collection that are relevant to a particular information need - a set of questions to which someone would like to find an answer. Therefore, information retrieval may be considered, as the quest for content, relevant to a given information need (Goh, 2012).

The Tools of Information Retrieval

Rapid advances in information technologies introduce an increasing number of information retrieval tools, which provide a wide range of information. The enormous growth in the size of scholarly literature makes its retrieval challenging. It is extremely important to choose appropriate information retrieval tools to ensure high quality scientific information sources, meeting the needs of a request for the user information. With the rapid increase of online resources and databases in the modern era, many students read to search and find the required information as fast as possible using these resources. Students now prefer to search the Internet instead of using printed resources in the libraries due to the fast and free access to a large amount of information available online. Internet search engines increasingly serve as the first option for people who want to find information. A diverse range of articles report the results

of studies of the information–seeking and retrieval behaviour observed in search engine environments (Kim, 2009; Thatcher, 2008; Jansen, et al., 2008)

Google is the most popular and the most commonly used web search engine worldwide. It is generally selected by users for its simplicity and the abundance of provided resources. Googling has become synonymous with research (Mostafa, 2005). Recent statistics indicate that Google has become the search interface of choice for many students to address their information needs, far exceeding their use of library catalogues or other online citation databases (Griffiths and Brophy, 2005). Universal search engines are not the most appropriate scientific search measures as they provide millions of sources not related to science alongside with valuable and high quality scientific information, consequently a lot of time is wasted for selection of appropriate resources. A number of important and useful sources of information do not appear among the sources of information, as universal information retrieval systems do not have access to all closed and requiring registration domains (e.g. databases). Universal information retrieval systems are an efficient search technique when looking for popular, entertainment information, but the search for scientific information is much more efficient, when specialized information retrieval tools are employed. Several important aspects have to be taken into consideration when selecting specialized information retrieval tools: majority of these tools are intended for the search of a particular format, type, category or scientific area publications; information retrieval tools are developed by institutions, libraries, international publishers, different organizations and other, therefore, they can be presented of different thematic, linguistic, geographical etc. scope; information sources, provided by information retrieval tools can be available to everyone or only to the users of the subscribing institution (Šarlauskienė, 2014). Different joint retrieval systems or virtual libraries have established a possibility of one–stop principle for the information retrieval. Due to the abundance of information retrieval tools, users take advantage of the easily accessible (open access or subscribed by an institution) sources and the sources, providing the largest choice of necessary literature. The websites of institution libraries provide a sufficient amount of information retrieval tools, intended for science and studies.

General search engines are not always able to search the content of databases as efficiently as specialized information search engines. Specialized information retrieval systems search for a certain type of documents, therefore, the results better meet users' needs. The most popular specialized scholarly literature retrieval system is Google Scholar. Google Scholar provides a simple way to broadly search for scholarly literature. From one place, you can search across many disciplines and sources: articles, theses, books, abstracts and court opinions, from academic publishers, professional societies, online repositories, universities and other websites. Google Scholar helps to find relevant work across the world of scholarly research. Google Scholar aims to rank documents the way researchers do, weighing the full text of each document, where it was published, who it was written by, as well as how often and how recently it has been cited in other scholarly literature (Google Scholar, 2021). Today Google Scholar provides its own version of citation information, with a 'ranking' technology that reports how often the item has been cited in other scholarly literature. This certainly tells the reader that authors have found this item valuable enough to cite, and tells the authors how valuable their work has been to other authors (Friend, 2006). Google Scholar has selected the most scientific information resources and is freely available to all Internet users. Google Scholar automatically identifies, based on certain criteria, which information resources are attributable to the production of scientific information and selects all available resources. Google Scholar selects web content and, when it finds the right information resource for the user, directs it to a database, archive, institutional cache, i.e. to the actual content owners (Grigas et.al. 2016).

Libraries across the globe started massively digitizing their collections, either alone or with the active involvement of commercial partners. In 2004, Google announced partnerships with several libraries and started digitizing books, with the ultimate vision of digitizing and making available the entire human knowledge online (Matulionyte, 2016). Keeping up with its innovative approach to large–scale problems, Google has developed the necessary set of enabling technologies, proving quite spectacularly that access to digitization time has indeed arrived. Google Books is a specialized book retrieval system. Books are provided to the system by publishers and authors or libraries, that is why books of different publishing years, types and topics can be found. Information about books, excerpts from the books or the entire text of a book can be provided.

WorldWideScience.org is a global science gateway comprised of national and international scientific databases and portals. WorldWideScience.org accelerates scientific discovery and progress by providing one–stop searching of databases from around the world. Multilingual WorldWideScience.org provides real time searching and translation of globally dispersed multilingual scientific literature. By using existing federated search and complex

translations technologies, WorldWideScience.org allows the user to search multiple scientific sources around the world with a single query. The user can review his result list and proceed to the host site for a particular result for more detailed information such as the bibliographic citation, abstract, and in many cases, full text of the document. Ten languages currently supported are Arabic, Chinese, English, French, German, Japanese, Korean, Portuguese, Russian, and Spanish. Native speakers of these languages now have unprecedented access to the English content. Multilingual translations are powered by Microsoft® Translator (WorldWideScience.org, 2020).

ScienceResearch.com portal aims "to unify the World Wide Web's dispersed science to become the world's most comprehensive portal for science." Additionally, the portal seeks to make "long tail science," the very specialized science that may appear to be of limited interest available to a larger audience through which applications may be found. ScienceResearch.com provides a single point of access to more than 400 high quality, publicly searchable science and technology collections with a new, robust user interface specifically designed for advanced scientific research. ScienceResearch.com also searches Science Conferences, a portal, providing access to some of the best conference proceedings. These results are aggregated with the results returned by individual sources. ScienceResearch.com is divided into 15 categories. Categories have also been created for Science News and Patents. Users can search any number of categories, searching all collections within the categories selected, or choose specific collections within a category to narrow their search. ScienceResearch.com's categories are managed by volunteer moderators who help the ScienceResarch.com team select the best, most authoritative collections to include in each category (Hane, 2009).

The Online Knowledge Library (B–on) provides unlimited and permanent access to research and higher education institutions to the full texts of thousands of scientific journals and online eBooks from some of the most important content providers, through subscriptions negotiated at national level. It started operating in March 2004 giving access to thousands of scientific publications and is today a reference in access to international scientific information. It brings together institutions of different types: higher education, scientific research and technological development, hospitals, public and private non-profit administration. This information retrieval system provides access only to subscribed users.

The abundance of information sources and resources on the Internet has encouraged developing specialized information portals and virtual libraries (Šarlauskienė, 2014).Virtual libraries are characteristic of being intended for meeting the needs of an actual institution, a region or a specific scientific area. They provide information retrieval from different resources (e-catalogues of libraries, institutional e-repositories and other open access resources, subscribed databases and other).

Information retrieval is a creative process, the success of which depends on information retrieval knowledge and skills, the character and type of the required information, properly selected information retrieval tools.

Methodology

Exploratory (pilot) research type was selected in order to find out what information retrieval tools students use to search for information for their studies, what the differences are between Lithuanian and Portuguese students in choosing information sources. Questionnaire survey is the most popular research method used to find out the opinion of the majority of respondents (Kardelis, 2016). The respondents were provided with a questionnaire with closed-ended questions to which they could choose from the options provided. The survey was carried out in December 2020. An online questionnaire was forwarded to the students of ISCAP Porto Accounting and Business School and to the second year students of the study programme of Information Management at Šiauliai State College. Only 11 students study in the second year of the Information Management study program at Šiauliai State College, therefore it was decided to interview an equal number of both Lithuanian and Portuguese students. Twenty replies to the survey were analysed: 10 of Lithuanian students and 10 of Portuguese students. We have to acknowledge that our sample size is quite small, which possibly might affect the validity and generalization of our findings.

Results

As mentioned above, this survey was provided to the students of ISCAP Porto Accounting and Business School and the second year students of the study programme of Information Management at Šiauliai State College. 20 replies were submitted – 10 from Lithuanian students and 10 from Portuguese students. 75 percent of the respondents were female and 25 percent were male. Age distribution of the survey respondents was 19 – 23 years of age.

The results of the research revealed that the most common searched type of information is scientific papers, accounting for 85 percent, followed by basic information, accounting for 70 percent, academic papers and books, accounting for 35 percent and scientific research reports, accounting for 25 percent. The data is provided in the diagram below.

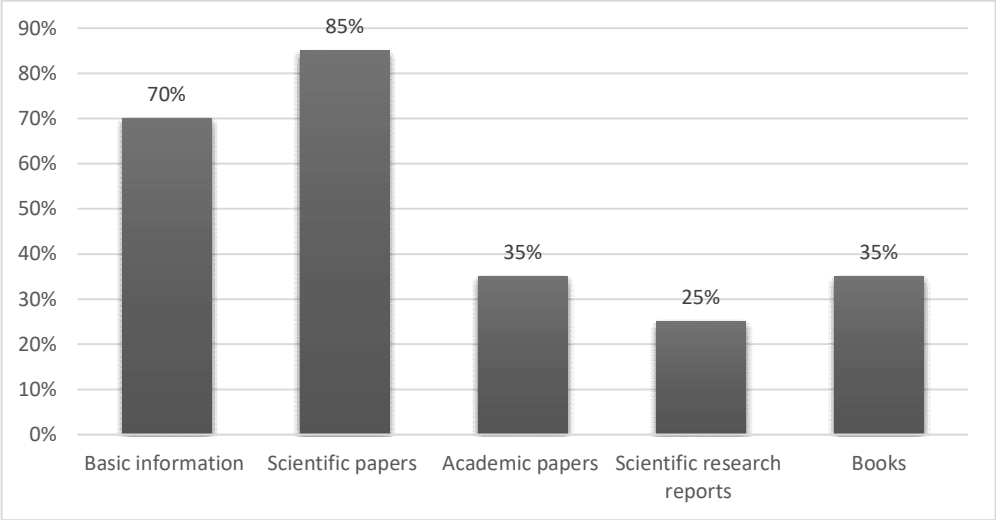

Fig. 1. Information sources necessary for the respondents

The results of the research revealed that the respondents Lithuanian students most often look for scientific papers (90 percent) and books (60 percent), while Portuguese students focus on basic information (90 percent) and scientific papers (80 percent). The research stated that only 10 percent of the respondents Portuguese students tend to look for books. 50 percent of the respondents Portuguese students and 20 percent of Lithuanian students, who participated in the research, look for academic papers. The data is provided in the diagram below.

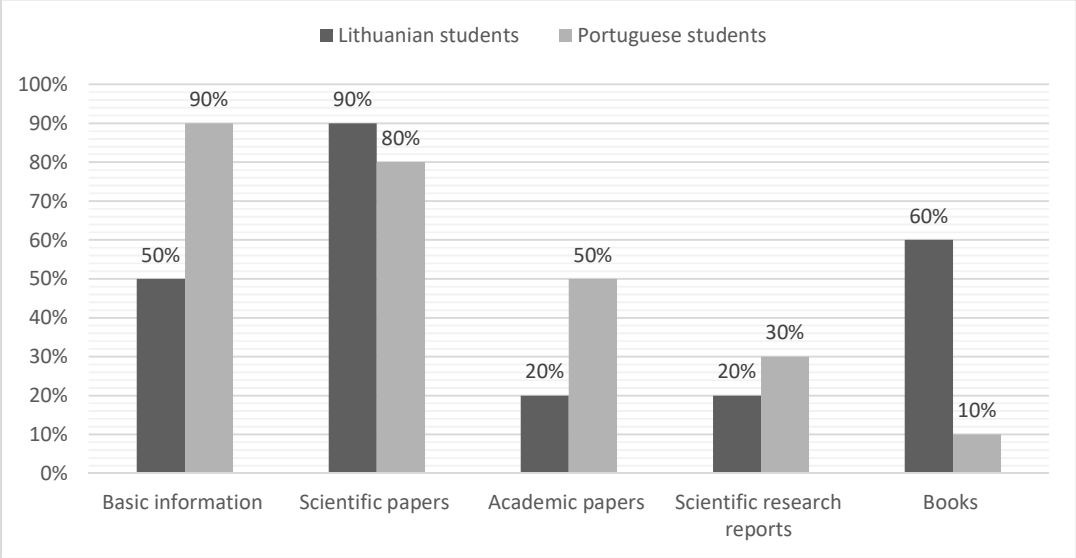

Fig. 2. Information sources necessary for Lithuanian and Portuguese students

It can be implied that information users, when solving information problems, select such sources of information, which would probably provide the solution to the problem task.
We can presume that the need of information source is determined by the topic of the studied subject, the tasks, provided by teachers, accessibility of the source of information, or possibility to borrow it from the library.

As for the question "What browsers/platforms do you usually use to carry out your research?" it was found out, that all the respondents take advantage of Google Scholar, this one being the most commonly used browser to search for the necessary information, followed by Google. The data is provided in the diagram below.

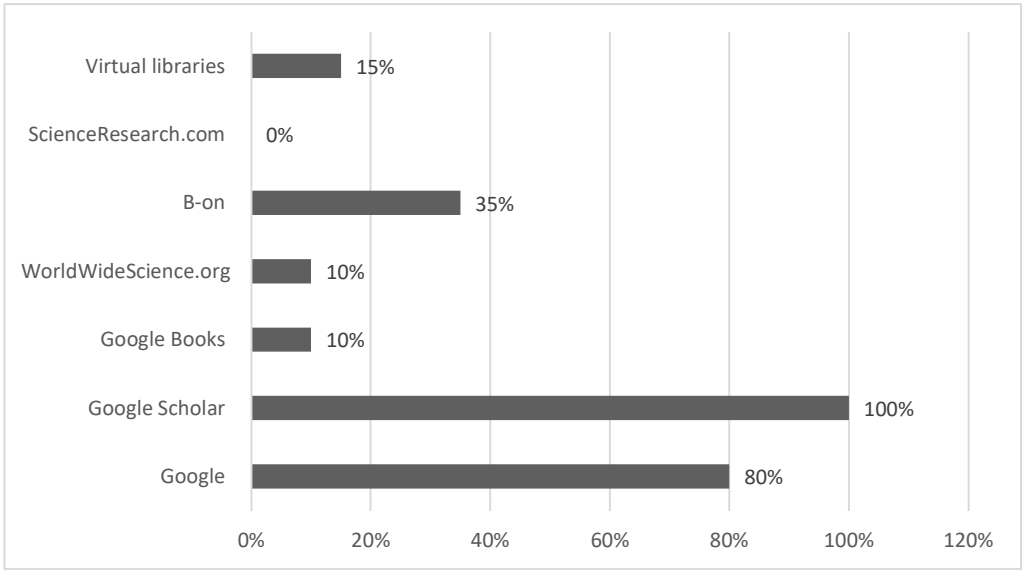

Fig. 3. Browsers/platforms used by the respondents

The results of the research allow stating, that information search engine Google Scholar is always used for the retrieval of the scholarly information or literature. It is very convenient for an information user to look for different documents in one place: articles, theses, books, summaries from different sources of information.

The results of the research showed that all (100 percent) the respondents Portuguese students conduct research in universal search system Google, whereas only 60 percent of the respondents Lithuanian students state to be searching for necessary information in this source of information. The data is provided in the diagram below.

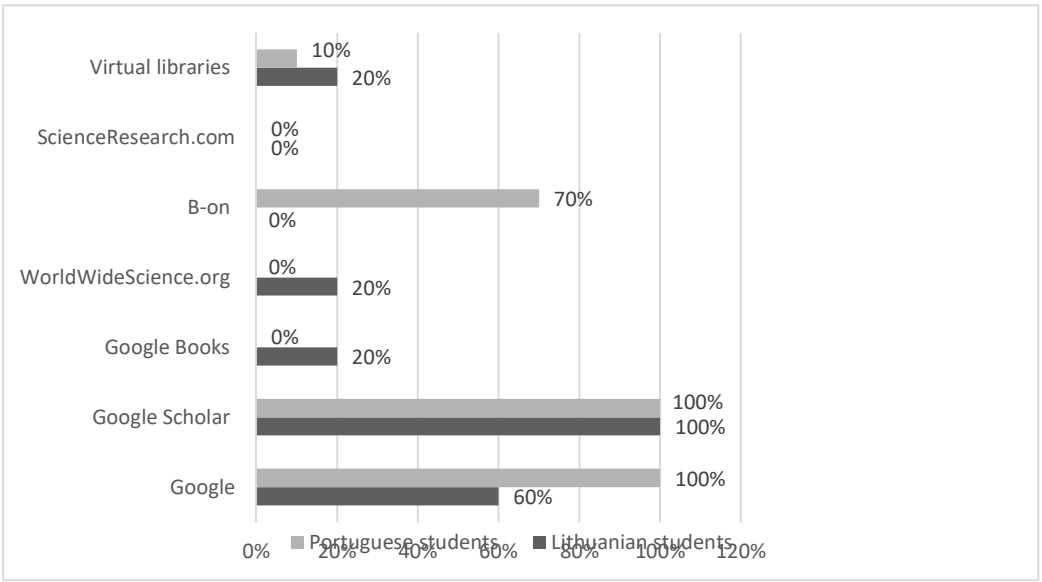

Fig. 4. Browsers/platforms used by Lithuanian and Portuguese students

20 percent of the respondents Lithuanian students carry out information retrieval in specialized book retrieval system Google Books, equally, 20 percent of the respondents

Lithuanian students look for information in WorldWideScience.org – global science gateway, where users can find scientific information in national and international scientific databases and portals from around the world. The respondents Portuguese students claim not to be using such search engines as Google Books and WorldWideScience.org. The given results revealed that neither Lithuanian nor Portuguese students take advantage of ScienceResearch.com information retrieval tool, providing access to more than 400 highest quality science and technology collections, equipped with new and reliable user's interface, specially designed for advanced scientific research. The research found out that 70 percent of the respondents Portuguese students use B–on information search tool for the information search. It should be noted that B–on information search tool provides the Portuguese students access to thousands of different sources of information – periodicals, electronic books and databases, and the possibility to carry out the information retrieval taking advantage of one–stop principle.

The respondents Lithuanian students have never used information search tool B–on, as this source of information is intended for subscribed users only in Portugal. It can be assumed, that in the future this great information search tool will be accessible and used by Šiauliai State College students, as the study programme *Information Management* has a signed double diploma contract with Porto (Portugal) Polytechnic Institute Porto Accounting and Business School (ISCAP), and the students will have an opportunity to join the users of B–on information retrieval tool.

20 percent of the respondents Lithuanian students of the research survey and 10 percent of the respondents Portuguese students stated that they look for information in virtual libraries as well. It can be noted that virtual libraries is not a very popular information search tool, as a big part of the documents introduced there provides only bibliographic description, but does not allow the user to get access to the entire text of the documents.

The research revealed that both Portuguese and Lithuanian students give priority to the sources of information to their native languages.

Conclusions

1. Information retrieval is a creative process, the success of which depends on information retrieval knowledge and skills, the character and type of the required information, properly selected information retrieval tools. Universal information retrieval systems are very convenient when looking for basic information. Specialized information retrieval systems look for a certain type of documents; therefore, the search results better meet the needs of the users. The most popular specialized information retrieval system is Google Scholar.

2. Exploratory (pilot) research was carried out for the first time, in order to find out the differences in information retrieval and information retrieval tool usage between Lithuanian students and Portuguese students. The online questionnaire was forwarded to the students of ISCAP Porto Accounting and Business School and to the second year students of the study programme of Information Management at Šiauliai State College. The research revealed that the respondents Lithuanian students look for scientific papers and books, whereas Portuguese students search for basic information and scientific papers. All the respondents of the survey identified Google Scholar as the most important search engine for scientific information retrieval. The Portuguese students most commonly took advantage of Google and B–on information search tools, whereas Lithuanian students used Google, Google Books, WorldWideScience.org information search tools. However, all the respondents of the questionnaire noted, that they started information retrieval from the sources of information in their native language first. Our findings suggest that students increasingly rely on search engines to support their study tasks. The retrieval tools are increasingly becoming an integral part of their lives, turning them into taken–for–granted background tools.

References

1. Chowdhury, G. (2010*). Introduction to Modern Information Retrieval*. (3rd edition). London: Facet publishing.

2. Chowdhury, G., & Foo, S. (2012). Information access. In G. Chowdhury & F. Schubert (Eds.). *Digital Libraries and Information Access: Research Perspectives* (p.47-68). doi:10.29085/9781856049764.005.

3. Chowdhury, G., Burton, P., McMenemy, D., & Poulter, A. (2007). Information retrieval. In *Librarianship: An Introduction* (p.123-140). doi:10.29085/9781856049146.017.

4. Chowdhary, K.R. (2020). *Information retrieval.* Retrieved from https://www.researchgate.net/publication/340433372_Information_Retrieval.

5. Griffiths, J.R., & Brophy, P. (2005). Student searching behavior and the Web: Use of academic resources and Google. *Library Trends*, *53*(4), 539–554. Retrieved from:

https://www.ideals.illinois.edu/bitstream/handle/2142/1749/Griffiths539554.pdf?sequence=2&isAllowed=y.

6. Fidel, R., & Pejtersen, A.M. (2004). From information behaviour research to the design of

information systems: the cognitive work analysis framework. *Information Research*, 10(1). Retrieved from: http://InformationR.net/ir/10-1/paper210.html.

7. Friend, F.J. (2006). Google Scholar: Potentially Good for Users of Academic Information. *Journal of Electronic Publishing, 9*(1). https://doi.org/10.3998/3336451.0009.105.

8. Grigas, V., Šuminas, A., Gudinavičius, A., & Grigaliūnas, M. (2016). *Šiuolaikinės medijos ir informacija: požeminiai garažai Google karalystėje.* Vilnius: Akademinė leidyba.

9. Goh, D. (2012). Collaborative search and retrieval in digital libraries. In G. Chowdhury & F. Schubert (Eds.). *Digital Libraries and Information Access: Research Perspectives* (p. 69-82). doi:10.29085/9781856049764.006.

10. Google Scholar. (2021). *About*. Retrieved from: https://scholar.google.ca/intl/en/scholar/about.html.

11. Hane, P.J. (2009). *Deep Web Tech Relaunches ScienceResearch.com.* Retrieved from: http://newsbreaks.infotoday.com/NewsBreaks/Deep-Web-Tech-Relaunches-ScienceResearchcom-54675.asp.

12. Harman, D. (2019). Information Retrieval: The Early Years. *Foundations and Trends® in Information Retrieval,13*(5), 425-577. http://dx.doi.org/10.1561/1500000065.

13. Janiūnienė, E. (2012). Studentų informacinę elgseną motyvuojantys veiksniai. *Informacijos mokslai*, *61*, 83-92. https://doi.org/10.15388/Im.2012.0.1067.

14. Jansen, B.J., Booth, D.L., & Spink, A. (2008). Determining the informational, navigational, and transactional intent of Web queries. *Information Processing & Management*, *44*(3), 1251–1266. Retrieved from: https://www.researchgate.net/publication/222824696_Determining_the_Informational_Navigational_and_Transactional_Intent_of_Web_Queries/link/59e62b5245851598a13a5b30/download.

15. Järvelin, K. (2011). IR Research: Systems, Interaction, Evaluation and Theories. *ACM SIGIR Forum*, 45(2), 17-31. Retrieved from: https://sigir.org/files/forum/2011D/conferences/2011d_sigirforum_jarvelin.pdf.

16. Kardelis, K. (2016). *Mokslinių tyrimų metodologija ir metodai: vadovėlis*. Vilnius: Mokslo ir enciklopedijų leidybos centras.

17. Kim, J. (2009). Describing and predicting information-seeking behavior on the Web. *Journal of the American Society for Information Science and Technology*, *60*(4), 679–693. http://dx.doi.org/10.1002/asi.21035.

18. Matulionyte, R. (2016). 10 years for Google Books and Europeana: copyright law lessons that the EU could learn from the USA. *International Journal of Law and Information Technology, 24*, 44–71. doi: 10.1093/ijlit/eav018.

19. Mostafa, J. (2005). Seeking better Web searches. *Scientific American*, *292*(2), 67–73. http://dx.doi.org/10.1038/scientificamerican0205-66.

20. Savolainen, R. (2016). Information seeking and searching strategies as plans and patterns of action: A conceptual analysis. *Journal of Documentation*, *72*(6),1154-1180. https://doi.org/10.1108/JD-03-2016-0033.

21. Savolainen, R. (2018). Pioneering models for information interaction in the context of information seeking and retrieval. *Journal of Documentation*, *74*(5), 966-986. https://doi.org/10.1108/JD-11-2017-0154.

22. Spink, A., & Cole, C.B. (2006). Human information behavior: integrating diverse approaches and information use. *Journal of the American Society for Information Science and Technology, 57*(1), 25-35. https://doi.org/10.1002/asi.20249.

23. Šarlauskienė, L. (2014). *Informacijos paieškos startegija.* Retrieved from: https://www.lmba.lt/sites/default/files/2-modulis.pdf.

24. Thatcher, A. (2008). Web search strategies: The influence of Web experience and task type. *Information Processing and Management*, *44*(3), 1308–1329. Retrieved from: https://www.sciencedirect.com/science/article/abs/pii/S0306457307001719.

25. Vakkari, P., & Järvelin, K. (2005). Explanation in Information Seeking and Retrieval. In: Spink A., Cole C. (eds). *New Directions in Cognitive Information Retrieval. The Information Retrieval Series*, *19*. https://doi.org/10.1007/1-4020-4014-8_7.

26. Vernickaitė, A. (2014). Parlamentarų informacinės elgsenos ypatumai*:* Lietuvos Respublikos Seimo narių atvejis*. Parlamento studijos*, *15*, 66-97. Retrieved from: *http://www.parlamentostudijos.lt/Nr15/15_informacija_1.htm*

27. Weber, H., Becker, D., & Hillmert, S. (2019). Information-seeking behaviour and academic success in higher education: Which search strategies matter for grade differences among university students and how does this relevance differ by field of study? *Higher Education,77*, 657–678. https://doi.org/10.1007/s10734-018-0296-4.

28. Weber, H., Hillmert, S., & Rott, K.J. (2018). Can digital information literacy among undergraduates be improved? Evidence from an experimental study. *Teaching in Higher Education, 23*(8), 909-926. doi:10.1080/13562517.2018.1449740.

29. Wilson, T.D. (2000). Human information behavior. *Informing Science, 3*(2), 49-56. Retrieved from: http://inform.nu/Articles/Vol3/v3n2p49-56.pdf.

30. WorldWideScience.org. (2020). *About.* Retrieved from: https://worldwidescience.org/about.html.

Received: 8 March 2021
Accepted: 31 May 2021