Purdue University

Purdue e-Pubs

Jun 3rd, 12:00 AM

Correlation Between Information Needs and the Library Collection: A Citation Analysis Study of Doctoral Theses at Universidade Federal de Santa Catarina Library

Maria Bernardete Martin Alves
Universidade Federal de Santa Catarina, m.alves@ufsc.br

Karyn Munyk Lehmkuhl
Universidade Federal de Santa Catarina, karyn.lehmkuhl@ufsc.br

Liliane Vieira Pinheiro
Universidade Federal de Santa Catarina, liliane.pinheiro@ufsc.br

Dirce Maris Nunes da Silva
Universidade Federal de Santa Catarina, dirce.maris@ufsc.br

CORRELATION BETWEEN INFORMATION NEEDS AND THE LIBRARY COLLECTION: A CITATION ANALYSIS STUDY OF DOCTORAL THESES AT UNIVERSIDADE FEDERAL DE SANTA CATARINA LIBRARY

Maria Bernardete Martins Alves
Universidade Federal de Santa Catarina, Brazil, m.alves@ufsc.br
Karyn Munyk Lehmkuhl
Universidade Federal de Santa Catarina, Brazil, karyn.lehmkuhl@ufsc.br
Liliane Vieira Pinheiro
Universidade Federal de Santa Catarina, Brazil, liliane.pinheiro@ufsc.br
Dirce Maris Nunes da Silva
Universidade Federal de Santa Catarina, Brazil, dirce.maris@ufsc.br

ABSTRACT

This study aims at measuring the correlation between information needs of patrons and the Universidade Federal de Santa Catarina (UFSC) Library collection. The survey was carried out targeting students from a doctoral course in Scientific and Technological Education especially, analyzing the listing of sources and literature of UFSC doctoral theses submitted in 2012. The postgraduate program was selected for its multidisciplinary nature. Theses were chosen for analysis because they are original studies and represent an innovative contribution. Also, they are available at the library collection and general public can access them both in printed or online format. The goal of this examination was to identify the matching between the literature used at research and the current library collection, to determine the proportion of electronic respectively online resources compared to printed material and their relationship to publication year, resource language and citation patterns. Concerning the used literature, when they were identified as online, was also verified if its content belongs to the library collection or if the access was provided by CAPES Portal. That is a scientific information portal maintained by the Brazilian federal govern. It collects and makes available to institutions of higher education and research in Brazil the best international scientific production. The methods used in this study will be implemented into the library's workflows in order to ensure the continual optimization of the collection development policy at Universidade Federal de Santa Catarina Library.

Keywords: collection development, citation analysis, academic libraries, decision-making process, collection evaluation.

INTRODUCTION

The primary function of a university library is to provide comprehensive resources in order to meet the university community information needs, supporting activities related to learning, teaching and research.
Therefore, libraries must have a coherent and consistent collection, oriented to the profile and information needs of their users. To achieve that goal, librarians can use Bibliometrics which is a useful tool to detect information usage patterns, especially in the context of scientific research.
According to Mattos & Dias (2009) Bibliometrics, through mathematical and statistical models, enables us to analyze published literature and how that information is being used, establishing literature behavior patterns as well as the users' information needs.

Among these models, the analysis of bibliographic citations of publications has been used as an important auxiliary criterion to guide decision making on library collection development. When applied to the set of publications of a particular community or research group, it plays an important role in user studies as, once the patterns of this group are identified, it is possible to establish its specific needs and demands.

Based on the idea that references cited in scientific works express the users information needs, Oliveira (2004) proposes the monitoring of such works to support the decision making process during the development of library collections.

Thus, this study aimed at measuring the correlation between the user's information needs and the University Library collection in the Federal University of Santa Catarina (UFSC).

About UFSC, it is a public university in the state of Santa Catarina, Southern Brazil. The main campus is located in Florianópolis, capital of Santa Catarina, but there are other four campuses in the cities of Joinville, Araranguá, Curitibanos and, from this year on, in Blumenau.

Considering all campuses together, UFSC has over 34,500 students, including undergraduates, graduates, distance learners, nursery and kindergarten children, as well as elementary school, high school and vocational students. The university has a faculty of 1,610 teachers and 2,874 technical and administrative staff. Moreover, UFSC has partnerships and cooperation agreements with institutions and organizations from 46 countries located in all continents (UFSC, 2010).

To provide information to all academic community, UFSC has a library system comprising a main library, four branch libraries, which are: health sciences / medicine, agricultural sciences, physical sciences and mathematics, education sciences. There is also a school library in the Laboratory School. All of these are located in Florianopolis. Each campus in the other cities has a library.

The central library has the biggest building, the largest collection and the largest number of librarians and staff compared to the other libraries which compose the system. Moreover, the university library administration as well as the acquisition and cataloging sectors are located there.

The collection of the university library at UFSC comprises 232,158 titles and 476,685 items, including printed and digital material. Among the printed material there are theses, dissertations, books, journals, special collections, proceedings, technical standards and maps. The library also subscribes to a great number of online databases with a vast amount of resources such as e-books, theses and dissertations from all over the world, journals, newspapers, videos, legal information, images, clinical information, etc.

Furthermore, the community at UFSC has full access to CAPES Portal. CAPES means Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, which is a federal agency for support and evaluation of graduate education. Its portal provides access to various electronic information sources, playing a leading role in the Brazilian scientific research.

This research was performed targeting students from the graduate course in Scientific and Technological Education. Specifically, were analyzed the reference lists from doctoral dissertations presented in 2012.

The objective of this research was to verify if the items cited in the dissertations were available at the library collection, identify the different types of publications, if they were printed or online, the publication year, languages and citation patterns.

Currently, UFSC university library offers both electronic and printed information resources and the selection and acquisition of this data follows a collection development policy. However, this policy does not include the citation analysis as a decision-making tool for collection acquisition and evaluation. Therefore, this study is a pilot for future implementation of citation analysis in the routine of UFSC university library.

COLLECTION DEVELOPMENT AND CITATION ANALYSIS

The collections are the central point of the libraries, and the staff and services only exist to maintain and provide access to the collection (Mack, 2003). The wide range of literature in various areas of knowledge as well as the availability of information in multiple media are a challenge to the university libraries when it comes to keeping updated and consistent collections that match their patrons profile.

For this reason, the development of library collections is essential to create and maintain, harmonically, collections in accordance with the institutional goals and capable of meeting the information needs of their users. According to Weitzel (2002), developing collections is, therefore, a technical activity committed with the systematization of a particular area under the institutional approach concerning the interests of those who keep the library.

Nowadays, the collection development demands strategies and techniques that make the selection and acquisition processes more objective, accurate and less laborious. Citation analysis is among such techniques. Haycock (2004) argues that, as a collection development tool, citation analysis can guide and support decisions related to selection, for example, to decide on the acquisition of a journal, subscription maintenance (and cancellation), or even providing electronic access.

Citation analysis is an instrument of Bibliometrics, which is considered the application of statistical and mathematical methods to books and other means of communication, aiming to elucidate the processes of written communication, the nature and development of scientific disciplines, from techniques to count and analyze the various aspects of this communication (Liniers, 1998).

The creation of the Institute for Scientific Information (ISI) and the emergence of citation indexes boosted such studies. The use of citation analysis can help recognize those who are interested in a specific subject and the research fronts, as well as identify the potential of journals and authors (Norton, 2001). It also assists in identifying information use trends in a given area, as it allows the identification of authors that publish as well as the most quoted works on a specific theme, the information sources where the papers were published and the relationships between documents and researchers.

In literature it is possible to identify articles that use the citation analysis as indicators for collection development. As stated by Williams & Fletcher (2006) citation analysis is a non-intrusive method of finding patterns in a specific population's use of research materials. The underlying assumption is that citations reflect the library needs of the citing authors. There are several views on citation analysis and many authors have discussed their uses and limitations.

As an example, Oliveira (2004) proposes the monitoring of scientific papers as a tool to assist in the decision making process of the acquisition of library materials. To corroborate this idea, the author analyzed the references of dissertations from the Department of Informatics of the Federal University of Espírito Santo in a pilot project which outcome was positive for this purpose.

Moreover, according to Araújo (2006), the data obtained from citations allow us to identify the most cited authors, the most productive authors, the research elite, the research front, the impact factor of authors and publications, geographic and/or institutional origins of the most influential authors in a particular field of research, the most used document types, average age of the literature used, literature obsolescence, geographic and/or institutional origin of the used bibliography, the most cited journals - as well as the "core" journals - that make up a field.

Regarding the library collections, citation analysis enables us to check whether users are using the collection as well as if the library has the documents cited in scientific papers.

Mattos (2008) analyses the citations used in theses defended at the Administration Graduate Programme in the Federal University of Rio Grande do Sul Administration School, from 1999 to 2007, aiming at obtaining indicators to support decision making for the development of library collections.

A similar study was held by Fernandes & Cendón (2010) who evaluated the collection of CAPES Portal through articles and journals cited in theses and dissertations from an Information Science Graduate Program. The citations were analyzed to verify their availability throughout the Portal. The bibliometric analysis results showed that the Portal collection contained 45% of the articles cited in the analyzed theses and dissertations, indicating that the collection can be expanded to better meet the needs of its users. The study recommends that the bibliometric analysis becomes part of the criteria adopted for the selection of new titles.

In Brazil the literature on the use of citation analysis as a tool for collection development is not *that* wide, but in the international literature there are many works with various approaches. An example of a research that resembles this one is Kayongo & Helm's (2012) who aimed to determine to what extent the University of Notre Dame collection met the needs of graduate students from that institution. Doctoral dissertations where analyzed and one

of the main objectives was to check if the libraries had the documents that had been cited at the dissertations.

The literature demonstrates that citation analysis is a useful tool for collection development. However, it should be used in conjunction with other indicators, as the quotes reflect usage in a given context and do not necessarily indicate the classic or higher quality works in all fields of knowledge.

METHODOLOGICAL PROCEDURES

This is a quantitative research, it uses bibliometric techniques, more specifically, citation analysis. Quantitative methods aim to obtain the frequency of facts, so that data can be generalized. About the analysis corpus, it consists of eight theses defended in the Scientific and Technological Education Graduate Program (PPGECT) at UFSC in 2012 and deposited in the University Library.

The Graduate Program in Scientific and Technological Education is interdisciplinary and intend to develop teaching and research on the relationship between education and science, specifically on the understanding of teaching-learning processes, that is, the mastery of thinking structures required by science and technology, and its insertion in Brazilian society to, in its light, promote a better education (UFSC, 2014).

This graduate program was selected due to its multidisciplinary nature. The doctoral dissertations were chosen due to the fact that they are original studies and represent an innovative contribution to science since they are results of further investigation on a topic with defined theoretical approach. Moreover, they are available in the library collection to the general public and can be found both printed and online.

Data collection involved the capture of the references listed at the end of each doctoral dissertation and its subsequent insertion into a spreadsheet in Microsoft Excel application. From the references the following data were identified:

a) Type of publication;
b) Format;
c) Year of publication;
d) Language.

About publication type, the defined categories were: book, article, conference paper, book chapter, web site, thesis, dissertation, legislation and others. It is important to explain that the category "article" encompassed articles from any kinds of serial publications, except conference paper and monographic series.

In the following stage, we verified the availability of the works cited at the UFSC University Library collection. We considered 'available in the collection' all material located in the library catalog, in CAPES Portal or in the online databases subscribed by the library.

The access to the local collection is free. It can be accessed at http://150.162.1.90/pergamum/biblioteca/index.php. The online collection is accessible to the entire academic community through IP recognition or through Virtual Private Network (VPN).

It is important to explain that the CAPES Portal is one of the largest collections of scientific publications in the world, containing over 29,000 titles covering all fields of knowledge. A great part of the material is in full text. Moreover, access is offered to over 130 reference databases, 150,000 e-books and patent databases, standards, theses and dissertations (CAPES, 2014). As a public federal institution, UFSC has unrestricted access to the portal content, providing its academic community with a vast source of scientific information.

In the process of reference data collection, problems related to lack of standardization were detected. That fact has raised questions on citation identification. The absence or inconsistency of data often made it hard to identify the type of work. These problems were solved consulting the available databases.

RESULTS AND DISCUSSION

In the analyzed dissertations were collected 1,732 references, with an average of 216.5 references by dissertation.

As shown in Figure 1, the book and book chapter categories represent 44.4%, showing that books were the main information source used by the authors of the dissertations published in the program. This result agrees with the results found by (Rosa, 2013) whose study showed the prevalence of books in the references cited by the authors of the theses in the interdisciplinary course at the Federal University of Santa Catarina.
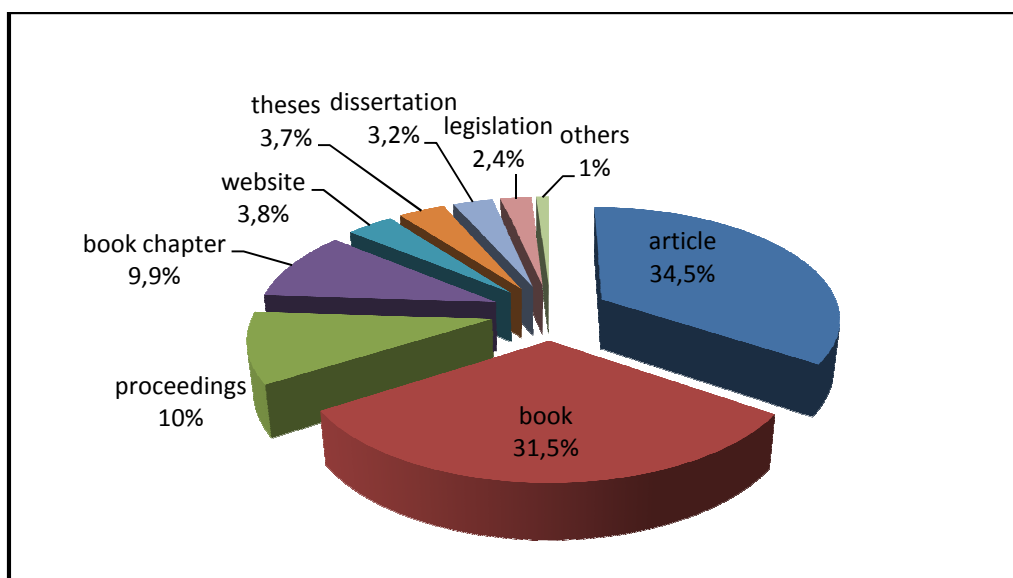

Figure 1: Publication type

On the other hand, it differs from Banateppanvar, Biradar & Kannappanavar's conclusions (2013). According to their research on thesis in the area of botany, the most cited sources were journal articles. It also differs from Williams & Fletcher's conclusions (2006) which also mentioned journal articles as the most used sources - in their case the theses were in the field of engineering. This divergence may be related to the multidisciplinary nature of PPGECT.

In this research, articles (34.5%) were the second type of publication most used by students to support their research, followed by work at events (10%).

The use of articles as a source of information may be related to its characteristics to science, as it is the 'scientific document by excellence' and is characterized as a finished report on a specific aspect of an investigation, as it is a complete piece that is made available to peer consideration (Maltrás Barba, 2003). As books and articles are considered the definitive publication of research results, they are preferably more read and cited (Meadows, 1999).

In the category "others", movies, administrative documents, long players, monographs, interviews, dictionaries and encyclopedias were mentioned.

We observed the prevalence of printed publications, as 1,507 out of 1,732 references collected referred to publications in this format (figure 2). Books and articles were also used in printed version. The low percentage of online material compared to printed material may be due to unfamiliarity with searching information on the internet or lack of training for this purpose (ECHEZONA, OKAFOR & UKWOMA, 2011).

However, it is important to emphasize - as stated before - that problems related to the lack of standardization were detected. That leads us to suspect that, although the citations did not present the data that characterized them as online sources (eg: URL and date of access), researchers may have used the online source - but they omitted these data. Kayongo & Helm (2012) also had difficulty in defining the impact of the use of online works because the researchers did not necessarily inform, at the reference, that it was an online material.
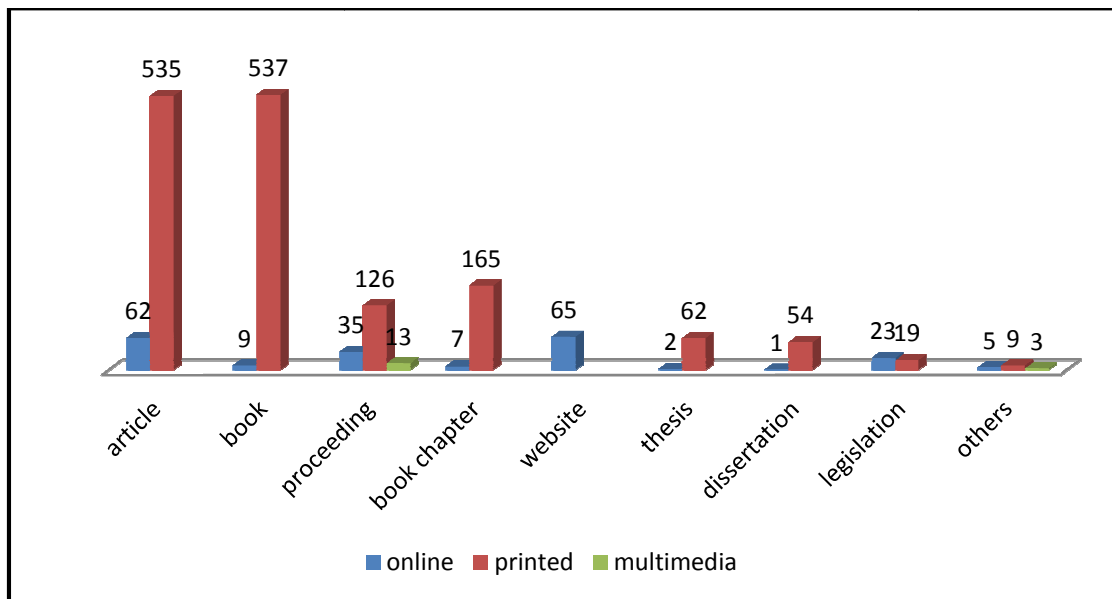
Figure 2: Publication type x form

About the publications language, Portuguese prevailed (figure 3); 72.7% of the information sources used, that is, 1,259 of them, were in that language. The other languages significantly used are English (17.2%) and Spanish (8.9%).

The fact that Portuguese prevailed may be related to the topics discussed, as they refer to the national reality and require theoretical support of national authors.
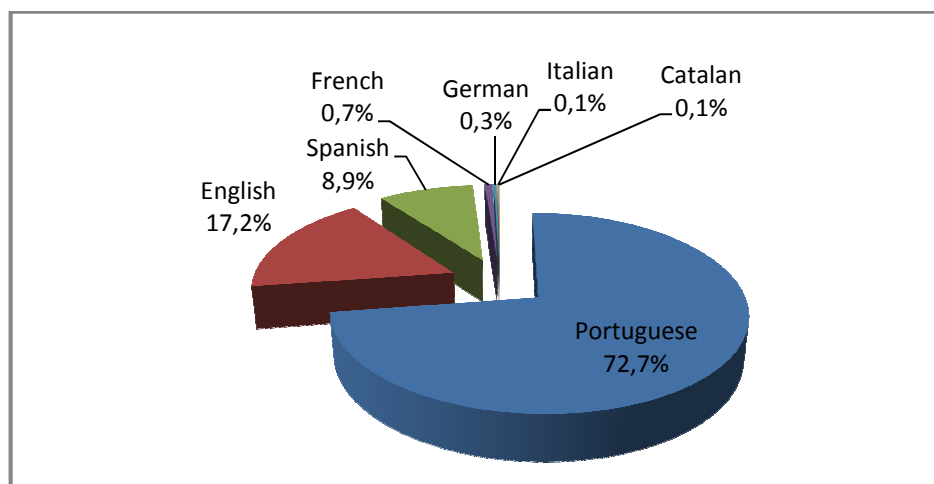

Figure 3: language

In scientific research the data topicality is a relevant factor, especially because it is a part of the work to collect previous studies in order to avoid duplicity and to know the research theme. About the publication date, the analyzed dissertations mentioned works published on different dates, including titles from the nineteenth century. However, most publications are recent; 1,163 of them have been published from 2000 on (figure 4). In six references it wasn't possible to identify the date.
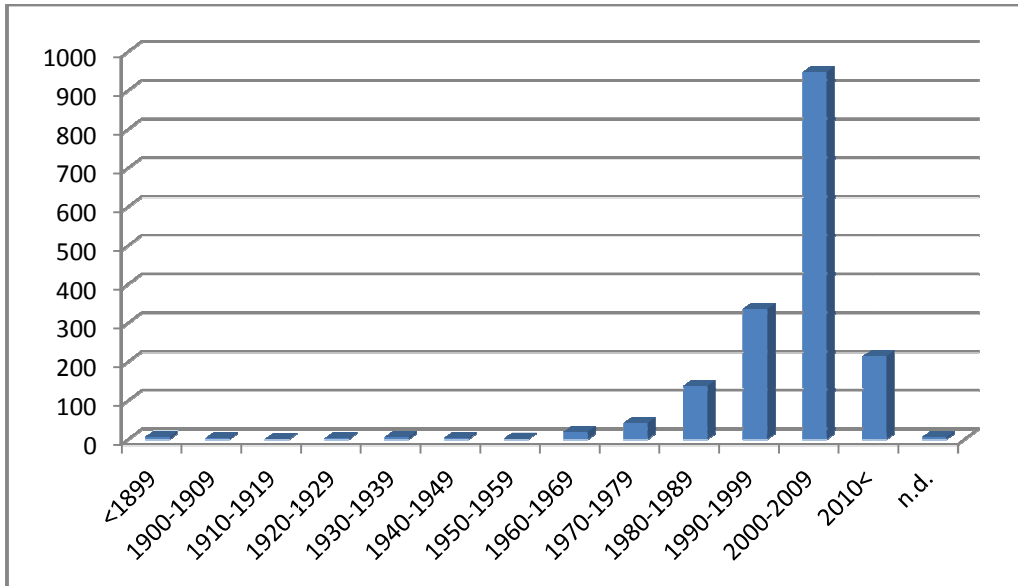
Figure 4: Publication date.

Regarding the availability of cited materials in UFSC University Library collection, it was found that the library provides or partially provides 49.42% of the cited information sources (figure 5). Recalling, the library collection includes all content available in the catalog, as well as e-books collections, online databases and CAPES Portal.

Nagel & Löbler (2007) got to a similar conclusion when analyzing bibliographic citations of thirteen Master's Degree in Business Administration thesis from the Social and Human Sciences Centre of Federal University of Santa Maria in 2005. The analysis results showed that 57.63% of the cited books are provided by the Library.

According to this research, 499 sources (28.81%) were found in its collection, 276 of them (15.94%) are available at CAPES Portal and one (0.06%) was retrieved from the e-book collections. It was considered that 81 of them (4.7%) are partially provided, as the library offers the information source cited, but in a different edition from the one indicated by the researcher.
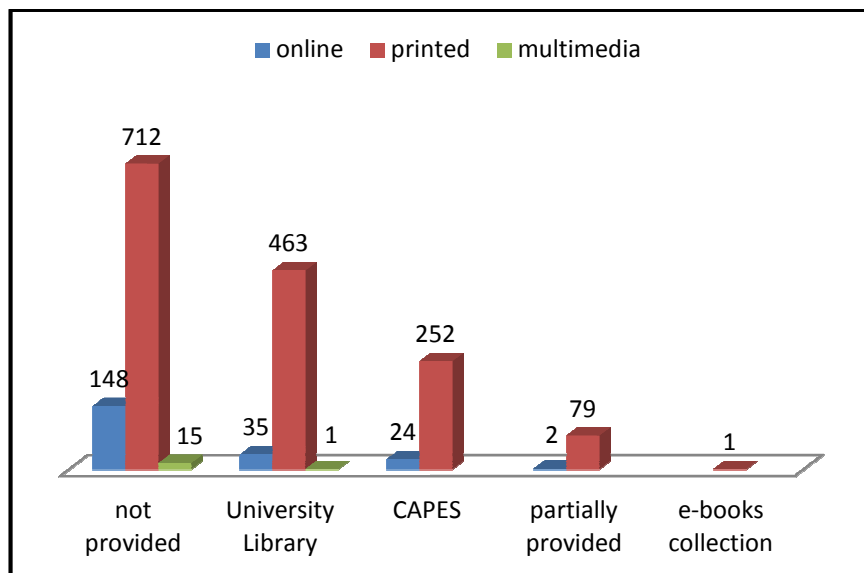

Figure 5: Provided information sources.

Still about the sources provided by the library (figure 5), it was observed that 15.99% of them are provided in a different format from the one specified by the researcher. Among the

printed sources, 252 (14.5%) were available online at CAPES Portal and one (0.06%) in the e-books collections. Among those cited as 'online', 35 sources (2%) were available in the library.
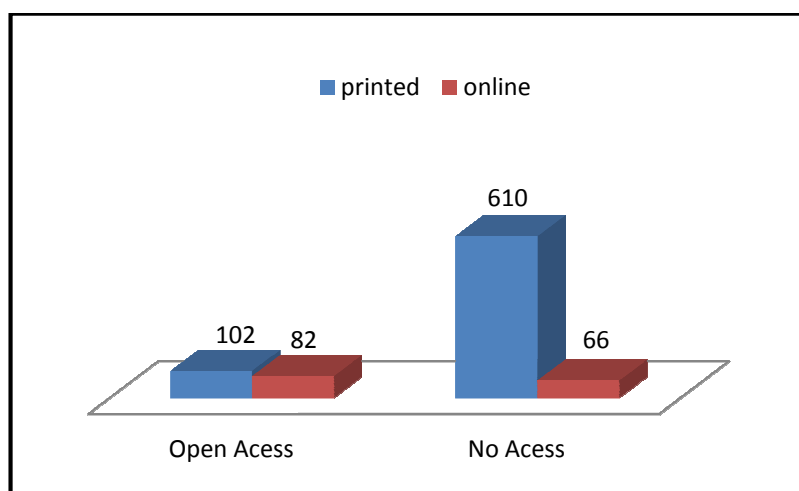

Figure 6: Internet availability x format

By comparing the information sources format identified as 'not provided by the library', it was found that 184 were freely available on the Internet (Figure 6). Also, 82 of all sources cited as 'online' (55.4%) and 102 of all sources cited as 'printed' (14.12%) were available on the Internet for free.

CONCLUSION

This research was intended to measure the correlation between UFSC University Library collection and the information needs of its users. Thus, we conclude that, among the sources used, 28.81% are in the physical library collection and 16% are in collections accessed or provided by the library. These data demonstrate that the library does not fully meet the information needs of researchers, who were the focus of this work. However, it is worth noting that the results demonstrate the use of information sources by a specific group of researchers.

Concerning the other goals, about the proportion of electronic resources in comparison with the use of printed material, although the library offers a wide range of online works, this type of format is not expressively used. However, it should be noted that the lack of reference standardization does not make it clear whether the document was actually consulted in printed or digital format. Furthermore, through the citation analysis it is not possible to know what led the researcher to cite a specific work, what was the context of that choice. And yet, even if the library has that specific work, it does not mean that the researcher used the copy available in the collection.

Regarding the publication year, the works cited are recent, a fact that is expected in a graduate program that addresses science and technology, even on the educational aspect. In these areas, topicality is a relevant factor as they undergo changes more constantly than the humanities, for example.

Concerning the language, Portuguese prevailed. This prevalence is common in the Brazilian research, even at the graduate level. It may be due to the need for theoretical support from national authors due to the choice of local subjects.

Regarding citation patterns, this study pointed out problems, both concerning reference standardization and information use. Based on this finding and because it is a multidisciplinary area where an information use pattern in not so clear, future studies are recommended to examine how researchers use information, the context of information source choice, their education field and the focus of the research.

This study was a pilot so that, in the future, citation analysis is adopted as a part of indicators for decision making in collection development of UFSC University Library. After all, it is important to bear in mind that, even citation being a good indicator of informational needs, it does not mean that the works cited are the most renowned in a particular area or the ones with the highest quality.

The citation analysis allows knowing the needs of users, what they are reading and using in their work. However, the decision making for collection development requires multiple and judicious approaches.

REFERENCES

Araújo, C. A. (2006). Bibliometria: evolução histórica e questões atuais. *Em Questão, 12*(1), 11-32.

Banateppanvar, K., & Biradar, B. S., & Kannappanavar, B. U (2013). Citation analysis of doctoral theses in botany submitted to Kuvempu University, India: a case study. *Collection Building*,32(1), 12-21. DOI: 10.1108/01604951311295058.

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (2014). *Portal CAPES*. Avaiable at http://periodicos.capes.gov.br/.

Echezona, R. I., Okafor, V. N., & Ukwoma, S. C. (2011). Information sources used by postgraduate students in library and information science: a citation analysis of dissertations. *Library Philosophy and Practice*. Retrieved from http://go.galegroup.com/ps/i.do?id=GALE%7CA267422510&v=2.1&u=capes58&it=r&p=AONE&sw=w&asid=f4c130a7f90b65f86307a393e18b0953.

Fernandes, W. R., & Cendón, B. V. (2010). Portal de Periódicos da Capes: proposta de um modelo de estudo bibliométrico para avaliação da coleção. (Portuguese). [Article]. Capes Portal of Periodicals: proposal for a bibliometric model for evaluation of the collection. (English). *Revista Brasileira de Pós-Graduação: RBPG, 7*(13), 318-349. Retrieved from http://ojs.rbpg.capes.gov.br/index.php/rbpg/article/view/196/190.

Haycock, L. A. (2004). Citation analysis of education dissertation for collection development. *Library Services and Technical Services*, 48 (2), 102-106. Retrieved from http://yunus.hacettepe.edu.tr/~tonta/courses/spring2011/bby704/citation-analysis-Haycock.pdf.

Kayongo, J. & Helm, C. (2012). Relevance of Library Collections for Graduate Student Research: A Citation Analysis Study of Doctoral Dissertations at Notre Dame. *College & Research Libraries*, 73(1), 47-67. Retrieved from http://crl.acrl.org/content/73/1/47.full.pdf+html.

Liniers, M. C. R.(1998). Bibliometría y ciencias sociales. *Clío*, 7, p.1-7. Retrieved from http://clio.rediris.es/ numero007.html

Mack, D. C. (2003).*Collection development policies*: new directions for changing collections. New York: Haworth Information.

Maltrás Barba, B. (2003). *Los indicadores bibliométricos*: fundamentos y aplicación al análisis de la ciencia. Gijón: Trea.

Mattos, A. M. (2008). *Análise de citações de teses como apoio a tomada de decisão do desenvolvimento de coleções de bibliotecas universitária.* (Dissertation). Universidade de Minas Gerais, Brasil.

Mattos, A. M., & Dias, E. J. W. (2009). Desenvolvimento de coleções em bibliotecas universitárias: uma abordagem quantitativa. *Perspectivas em Ciência da Informação, 14*, 38-60.

Meadows, A. J. (1999). *A comunicação científica*. Brasília: Briquet de Lemos.

Nagel, M. A. d. B., & Löbler, M. L. (2007). Aplicação da técnica de análise de citações como um instrumento de avaliação da Biblioteca Setorial do CCSH/UFSM. *Revista Sociais e Humanas, 20*, 239-250.

Norton, M. J. (2001). Bibliometrics. In: Norton, M. J. *Introductory concepts in information science*. Medford: ASIS.

Oliveira, E. (2004). Identificando necessidades de acervo de usuários de um departamento acadêmico: uma abordagem bibliométrica. *Ciência da Informação, 33*, 126-131.

Universidade Federal de Santa Catarina (2014). *Pós-Graduação em Educação Científica e Tecnológica*. Retrieved from http://ppgect.ufsc.br/apresentacao/.

Universidade Federal de Santa Catarina (2010). *UFSC's information in English*. Retrieved from http://en.ufsc.br/.

Weitzel, S. R. (2002). O desenvolvimento de coleções e a organização do conhecimento: suas origens e desafios. *Perspectivas em Ciência da Informação*, 7(1), 61-67.

Williams, V. K. & Fletcher, C. L. (2006). *Materials used by master's students in engineering and implications for collection development*: a citation analysis. Retrieved from http://www.istl.org/06-winter/refereed1.html.