Bibliometric Analysis of the Scientific Production of the Information Architecture Related to Libraries

Vitor Taga
Danielly Oliveira Inomata
Caroline Rodrigues Vaz
Mauricio Uriona Maldonado
Gregório Varvakis
Universidade Federal de Santa Catarina - UFSC, Brasil

Abstract

Objective. Information architecture is a determinant guiding element for content layout and user navigation. This work aims at identifying and characterizing the scientific production about information architecture related to libraries.

Method. In order to achieve the objective, a bibliometric review and a systemic analysis were conducted, the first to identify the most cited articles, the most productive authors in the field, the main journals; and the latter to characterize the debate and results presented in the articles. For the structured literature review, the Proknow-C method was utilized. The search was made in the Web of Science, Scopus and Science Direct databases.

Results. One hundred fifteen articles were identified. Three peaks in the number of publications in 2001 (5), 2002 (6) and 2006 (4) were identified.

Conclusions. The publications were concentrated in the Social Sciences and Computer Science. The main keywords (Digital libraries, Information architecture and World Wide Web) already contextualize and represent the content of the recovered articles. In relation to content analysis of the articles, a possible convergence in papers aimed towards studies of information architecture and digital libraries is discussed, which is very coherent with the current literature. We point out information architecture makes digital information environments better labelled and facilitates information access.

Keywords
Bibliometrics; Information architecture; Libraries

Análise bibliométrica da produção científica de Arquitetura da Informação relacionados às bibliotecas

Resumo

Objetivo. A arquitetura da informação é um elemento norteador determinante para a disposição do conteúdo e facilitador para a navegação do usuário. Este trabalho tem como objetivo identificar e caracterizar a produção científica sobre arquitetura da informação relacionada a bibliotecas.

Método. Realizou-se uma revisão bibliométrica e uma análise sistêmica, a primeira para identificar quais os artigos mais citados, os autores que mais publicam na área, os principais periódicos; a segunda análise para caracterizar as discussões e resultados apresentados nos artigos. Para a revisão de literatura estruturada, utilizou-se o método Proknow-C, a pesquisa foi realizada nas bases de dados Web of Science, Scopus e Science Direct.

Resultados. Foram identificados 115 artigos, desse resultado verifica-se um ápice no número de publicações nos anos de 2001 (5), 2002 (6) e 2006 (4).

Conclusões. As publicações estão concentradas nas subáreas de Ciências Sociais e Ciência da Computação. As principais palavras-chave (Digital libraries, Information architecture e World Wide Web) já contextualizam a representativamente o conteúdo dos artigos recuperados. Com relação à análise do conteúdo dos artigos, discute-se que existe uma convergência de trabalhos voltados para estudos de arquitetura da informação e bibliotecas digitais, o que é muito coerente com o que a literatura sobre o assunto aponta: a arquitetura da informação é um facilitador do acesso à informação, tornando os ambientes informacionais digitais melhor sinalizados.

Palavras-chave
Arquitetura da Informação; Bibliotecas; Bibliometria
1 Introduction

Information architecture (IA) became popular in the beginning of the 1990s due to a growing need for the organization of information resulting from the spreading of the World Wide Web (Web) in commercial activities.

With the rise of the first graphical user interfaces in the Web, firms realized that people visited their websites looking for knowledge and entertainment that could capitalize access to the firms' content, encouraging firms to add more content (Davis, 2012), which resulted in the exponential and uncontrolled growth of disorganized information in the web and the usage of inefficient websites (Bawden & Robinson, 2009).

Firms then realized that the quality of services and products offered in the Web was influenced by the lack of content organization, since if users could not follow up the graphical interface or could not find the information they were looking for, a broad negative effect over the search experience would impact the website and the firm as a whole (Ewing, Magnuson & Schang, 2002; Forlizzi & Battarbee, 2004).

Establishing IA is difficult (Robins, 2002), and there is no general consensus on what exactly an information architect does (Reiss, 2000), known only by their obscure role (Kalbach, 2009). Reiss (2000) argues that part of the problem resides in the variety of activities performed by information architects according to their jobs or education background, as well as a variety of tasks they can perform, depending on the type and size of the firm.

This work recognizes that IA is a determinant guiding element for content layout and a facilitator for user navigation. However, we are specifically interested in answering: Who are the main authors and which are the main articles in the scientific literature, related to IA in libraries? In this sense, this work aims at identifying and characterizing the scientific production about IA under the library field.

2 Information architecture

IA is essentially multidisciplinary. This has lead to many debates regarding its nature (Ewing et al., 2002). The term IA is ambiguous and came to be utilized in a general way (Jacob; Loehrlein, 2009).

Historically, IA was coined by the architect Richard Saul Wurman, and was presented in 1976 in the National Conference of the American Institute of Architects, where Wurman presided as chair and chose the conference theme: “The Information Architect”; even though, this denomination only began to be employed intensively in the 1990s with the popularization of the internet and of the Web (Wyllys, 2000; Kalbach, 2003; Ding; Ling, 2010).

Throughout the years, conflict arose between these fields and subjects dealing with the design of information-rich services, products and environments (Resmini; Rosati, 2011), signaling the beginning of an occupational confusion (Davis, 2011). IA was then confused or defined under distinct names, among them information design, interface design, usability engineering and others. (Robins, 2002).

The demand for information organization on the Web and its massive investments were among the factors that intensified the need for dealing with previous challenges of content organization in more versatile ways, in order to have people being able to navigate the web more naturally and intuitively (Merholz, 2009; Davis, 2012).

In this sense, many specialists from different professional backgrounds focused on information organization, were interested on this issue (Merholz, 2009), such as information designers, information architects, user experience strategists, interaction designers, among others (Cooper, Reimann; Cronin, 2007; Rosenfeld, 2002).

These fields began to deal with a converging problem: discovering how to design new information systems and offer a methodology and a framework to develop them (Rosenfeld, 2002). Evidently, these systems present significant overlap regarding the solution due to a disciplinary and contextual bias from their creators.

However, the organization of information can be interpreted in different ways, especially considering the possible approaches and perspectives to be employed (Dillon; Turnbull, 2005).

IA is hereby stated as emergent from a highly dedicated practice in navigation, problem-solving and of finding content capacity, but not exclusively for websites. With the rise of new Web technologies, such as the Web 2.0, and new applications, these subjects were gradually included in the Web’s scope (Cooper et al., 2007), bringing it closer to interaction design, since both areas concern defining the behavior of the system and of the user (Cooper et al., 2007; Ding & Ling, 2010).
According to Fast (2007), IA was conceived, defined and developed by practitioners from distinct fields (Library Science, Information Science, Human-Computer Interaction (HCI), Information Recovery, etc.). Without a proper theoretical body, IA sustained itself in neighboring fields, hence, being benefited by theories and practices in those fields, contributing to the skills and knowledge involved in IA (EWING et al., 2002; HAVERTY, 2002; KALBACH, 2003; 2009; LATHAM, 2002).

Information architects come from established disciplines (e.g. graphic design) but usually work under the knowledge base of other ones (e.g. usability engineering) (ROSENFELD, 2002). IA is frequently considered unrestricted (V-2 ORGANIZATION, 2001 as cited in Haverty, 2002), which has reflected in the amplitude and practical variety of information architects, consequently making it difficult to characterize them technically and conceptually (REISS, 2000; CAMPBELL, 2007).

IA has become an important emergent subject that is concerned with improving the efficacy of informational environments, particularly websites, being also pushed by academic interest and acceptability from firms (EWING et al., 2002). IA offers a rich professional and academic opportunity, which is claimed by both communities, the professional and the academy (DILLON, 2002; DILLON; TURNBULL, 2005; FAST, 2007).

According to Davis (2010), communication is inefficient among a broad community of interest that surrounds the IA. Dillon (2001) considers information systems design as a complex activity that requires multiple abilities that go beyond any field. For Reiss (2000) and Resmini and Rosati (2011), IA seems to be accepted by the broad Web design community and considers interaction design, usability engineering, information visualization, content strategy, graphic design, content management as complementary subjects that composing the main field of user experience.

In the libraries’ context, IA is supported by the entrance of digital technologies for the access and organization of traditional services (e.g. the book search in the archive/collection), and also by the modernization of digital services, such as the case of digital libraries (NONATO et al., 2008).

3 Methodology

This research is classified as exploratory and descriptive regarding its objectives. We used the bibliometric review method for identifying the most cited articles, the most productive authors in the field, the main journals; and a systemic analysis, in order to characterize the discussions and results presented in the articles.

The instrument used was the Knowledge Development Process – Constructivist (Proknow-C), which was proposed by Ensslin et al. (2010). This instrument helps in the selection of a bibliographic portfolio through a process subdivided in four phases (Figure 1): 1) selection of articles database – defining the keywords and the databases; 2) initial filtering – filtering of the bank of raw articles, title reading, and exclusion of repeated articles; 3) main filtering articles database – determining the scientific recognition of the articles and the alignment of the articles from the abstract; 4) filtering for alignment of the full article – full reading of the articles.

Figure 1 – Distribuição de frequência dos descritores.
The search was conducted in Web of Science, Scopus and Science Direct, with the keywords: "Information architecture" and library. It is worth mentioning that the bibliometric analysis was conducted on the bibliographic portfolio (the final list of articles included for analysis), which resulted from phase four.

The bibliometric review made possible to retrieve the word count on the following fields (title, abstract, keywords and subject). The search was made in November, 2014.

4 Bibliographic portfolio

From the database searches, 168 articles were recovered and 35 duplicates were removed. After the use of the Proknow-C protocol, 49 documents were selected (Table 1). This was the number of documents made available for download.

<table>
<thead>
<tr>
<th>DATABASE</th>
<th>KEYWORDS</th>
<th>FILTER*</th>
<th>RESULT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scopus</td>
<td>&quot;Information architecture&quot; AND library</td>
<td>Tit.Abs.Key</td>
<td>115</td>
</tr>
<tr>
<td>WoS</td>
<td>&quot;Information architecture&quot; AND library</td>
<td>Topic</td>
<td>50</td>
</tr>
<tr>
<td>Science Direct</td>
<td>Tit.Abs.Key</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td><strong>Total of articles</strong></td>
<td></td>
<td></td>
<td><strong>168</strong></td>
</tr>
<tr>
<td><strong>Duplicated articles</strong></td>
<td></td>
<td></td>
<td><strong>35</strong></td>
</tr>
<tr>
<td><strong>Articles available and in line with the research</strong></td>
<td></td>
<td></td>
<td><strong>49</strong></td>
</tr>
</tbody>
</table>

* Characteristic from each base

Table 1 – Distribuição de frequência dos descritores.

Table 1 shows that the higher share of recovered items was in Scopus, which might mean that the journals indexed in this database concentrate a larger amount of studies of interest for our research.

4.1 Scientific publications

The following graph presents the bibliographic portfolio per publication type. The scientific publications that composed the bibliographic portfolio present the highest amounts of publications in scientific journals (34), followed scientific conferences (9), and with a smaller representation, bulletins (2) (Bulletin of the Association for Information Science & Technology), book chapters (2) (Research and Advanced Technology for Digital Libraries, and Usability for the Web), and review article (1) (Annual Review Information Science & Technology – ARIST).
The main scientific journals were: Journal of the Association for Information Science and Technology (JASIST), OCLC Systems & Services, and The Electronic Library, both the OCLC and the JASIST with four articles published about the subject (Figure 3). The publications’ details can be seen in Appendix A.

Data shows the predominance of the subject in publications of scientific journals aimed towards research with a technological focus linked to libraries’ environment and to the science field of information.

In the publications of scientific conferences, a similar scenario is perceived, where there was a predominance of the conferences with a technological and information focus. In total, there were nine publications in conferences. In only two cases two publications in the same event occurred, though in distinct years, which was the case of the Annual Meeting of the Association for Information Science and Technology (ASIS&T), in 2002 and 2012, and the Join conference on Digital (JCDL), in 2001 and 2002.

The remaining events have only one publication, they are: International Conference on Management Science and Engineering (ICMSEM), in 2006; In 2011 there were two conferences: SouthesatCon and International Conference on Internet Technology and Applications (iTAP); in 2012, the International Conference on Information Management, Innovation Management and Industrial Engineering (ICIII); and in 2013 in the conference from the Special Interest Group on Design Communication (SIGDOC).
The other means of scientific communication (book chapters, bulletins, the review article and the panel abstract article) also converge in subjects aimed towards technology and information.

4.2 Scientific publications by year

Publications within our bibliographic portfolio began in 1998, with a publication of an article in a scientific journal (Program), following to the year 2014. In 1999, three articles in scientific journals were published (Library Philosophy and Practice, Journal of Documentation, and The Electronic Library). In the following year, the number of publications was reduced to one article publication in a journal (The Electronic Library). From 1998 to 2000, the publications are concentrated in scientific journals only. The years with the least numbers of publications were 1998, 2000 and 2012.

![Figure 4 – Distribution by Publication Date](image)

From the total sum of works published by year, 2002 was the year with the highest number of publications, with a total of six publications, composed by a bulletin article (ASIS&T), a conference article (ASIS&T Annual Meeting), and four journals articles, three of them in JASIST, and one in the Information Technology and Libraries.

From the period (1998 to 2014) analyzed, the year 2001 was the one that presented the second highest number of publications - five publications, a conference article, a discussion panel abstract (CHI), a book chapter (Usability for the Web), and two scientific journals, being one from the Library Hi-Tech and the other from College & Research Libraries.

4.3 Keywords

From the 49 publications analyzed, 22 did not have keywords next to the Abstract. Thus, the analysis of the keywords of the works was restricted to those that presented keywords appearing next to the Abstract. Altogether, 123 keywords were registered, 28 of which were repeated. Furthermore, after excluding repeated keywords, 95 keywords were obtained. An average of five keywords per work was registered. The keywords were listed and sorted on a descending order according to the frequency of appearance, down to two occurrences, as shown in Table 2.
<table>
<thead>
<tr>
<th>KEYWORDS</th>
<th>AMOUNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital libraries</td>
<td>5</td>
</tr>
<tr>
<td>Information Architecture</td>
<td>5</td>
</tr>
<tr>
<td>World Wide Web</td>
<td>5</td>
</tr>
<tr>
<td>Libraries</td>
<td>3</td>
</tr>
<tr>
<td>Classification</td>
<td>3</td>
</tr>
<tr>
<td>Semantics</td>
<td>3</td>
</tr>
<tr>
<td>Design</td>
<td>3</td>
</tr>
<tr>
<td>Evaluation</td>
<td>3</td>
</tr>
<tr>
<td>Usability Testing</td>
<td>2</td>
</tr>
<tr>
<td>Internet</td>
<td>2</td>
</tr>
<tr>
<td>Metadata</td>
<td>2</td>
</tr>
<tr>
<td>Information Technology</td>
<td>2</td>
</tr>
<tr>
<td>Information Organization</td>
<td>2</td>
</tr>
<tr>
<td>Education</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2 – Main Keywords

As presented in the previous table, the keywords “digital libraries”, “information architecture” and “world wide web” were the most used, with five occurrences each. Then, the keywords “libraries”, “classification”, “semantics”, “design” and “evaluation” appear, with three occurrences each. Six keywords presented two occurrences: “usability testing”, “internet”, “metadata”, “information technology”, “information organization” and “education”.

The keywords above, were used to plot a cloud graph, which shows the variation of occurrence as a variation of font size, and also presents keywords clusters according to their relation (Figure 5).
Three clusters or main groups are observed. The first presents the word “web” in display, the second the word “digital” and the third “technology”. Around these terms, smaller terms appear; in the first case the words “information”, “internet”, “management” and “strategic”; in the second case the words “user”, “worldwide”, “web”, “organization”, “classification”, and “libraries”; and in the third case the words “semantics”, “design”, “libraries” and “architecture”.

This data reinforces the indications that the main themes or subjects are related to the thematic, likewise, how these themes are interlinked or associated to themselves or to the other themes, which are highlighted in relation to the others.

### 4.4 Authors

One hundred twenty two authors were counted as a whole. From the 49 publications, 14 articles accounted a single author, 15 accounted two authors, 11 presented three, five displayed four authors, one presented five authors, two presented six authors, and one, a book chapter, displayed 12 authors. Table 3 lists the eight authors that most published in a descending order.

<table>
<thead>
<tr>
<th>PLACE</th>
<th>AUTHORS</th>
<th>AMOUNT OF PUBLICATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ROBINS, David</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>BEHESHTI, Jamshid</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>DEMPESEY, Lorcan</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>FOX, Robert</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>LARGE, Andrew</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>MARCHIONINI, Gary</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>MURRAY, Robin</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>TOMS, Elaine G.</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3 – Main Authors

Approximately 80 of the 122 authors either come from the fields of Library Science and Information Science or work in specialized libraries and universities. This data shows that approximately 65% of the publications resulted from developed research in these knowledge/work fields. In the case of the other researchers, the majority comes from the computing field, the remaining come from the fields of management, health, agriculture, among others.

The specific case of the eight authors that presented two publications is presented below:

1. David Robins is Assistant Professor in School of Information Sciences at the University of Pittsburgh, Pennsylvania, United States of America.
2. Robert Fox is the manager for the Web and Software Engineering Unit for the Hesburgh Libraries, University Libraries of Notre Dame, Indiana, United States of America.
3. Jamshid Beheshti is Associated Professor in School of Information Studies at McGill University, Montreal, Canada.
4. Elaine G. Toms nowadays is Professor of Information Science at The University of Sheffield, England.
5. Andrew Large is Professor Emeritus in Information Studies at McGill University, Montreal, Canada.
6. Robin Murray is Technical Director, Fretwell-Downing Informatics, United Kingdom.
7. Gary Marchionini is Professor in School of Information and Library Science, University of North Carolina at Chapel Hill, United States of America.


Andrew Large and Jamshid Beheshti are co-authors of the paper “Design Criteria for Children’s Web Portals: The Users Speak Out” and “Information Architecture for the Web: The IA Matrix Approach to Designing Children’s Portals”. It is noteworthy that Elaine G. Toms is the co-author of the articles “Usability of the Academic Library Web site” and “The impact of information architecture on academic web site usability” which are among the most cited articles in Scopus and WoS, as shown onward.

4.5 Citations

Although there is certain variation in the number of citations, the most cited publications display symmetry, that is, the papers with the highest number of citations in a database were also the most cited papers in the others, as shown in Figures 6 and 7. In this work, due to variation in the number of citations pointed out between Scopus and WoS, the academic relevance was analyzed through Google Scholar due to the fact that this database presents a compilation of citations of all other databases.

In Figure 6, publications from 1998 to 2003 were listed, and in Figure 7 publications from 2003 to 2013. The oldest publications can be seen as the ones that received the highest amount of citations.
Figure 7 – Publication Date Distribution II


4.6 Systemic analysis

With a systemic analysis of content of each of the 49 publications, it was possible to identify five core subjects that the researches (Figure 8) aimed at directly or indirectly:


Teaching and Academic Guidance (TAG): studies that discussed information architecture in teaching, required abilities, academic curriculum, teaching institutions and research directions were considered. In all, 6 studies were classified in this topic: Rosenfeld & Instone (2001), Robins (2002), Zhang, Strand, Fisher, Kneip, & Ayoub (2002), Jacob & Loehrlein (2009), Pattuelli (2010), Gradmann (2014).

Other Themes (OT): group composed by studies that presented different themes than the majority, two studies enrolled in this group: Twining (1999) and Clyde (2000).

Figure 8 – Central Themes by Publication Date

All 49 publications approach the central topics above directly or indirectly and correlativey. The occurrence of four core themes related to Library was observed in these papers: Library Websites (LB), Digital Libraries (DL), Library Science and Information Science (LSIC), and Other Topics (OTT). Figure 9 below shows the distribution of publications amidst the topics related to Library separately and organized by year of publication.
Figure 9 – Central Topics Related to Library


And the fourth subject was Library Science and Information Science (LSIC), with 5 publications: Twining (1999), Gardner (2001), Robins (2002), Zhang et al. (2002), Garshol (2004).

The data of Figures 8 and 9 combined with the publications made it possible to verify the relationship between the five core themes with the four core topics related to Library (Figure 10).
The popularity between certain core themes with core topics related to Library is observed. From the point of view of the central topics related to Library, the most researched theme related to Library Websites is User Test. From the point of view of the central theme, Technological Interoperability and Infrastructure is mostly related to other topics related to Library, for example, there is no publication regarding this subject in Library Websites.

5 Conclusions

Our bibliometric and systemic analysis of the IA field evidenced the need for unifying research in the scientific community. As pointed out by Robins (2002) and Reiss (2000), there is no unique definition for IA.

The analysis exposed some issues, shown below, that might help the understanding of how IA has been studied in previous literature, indexed in three of the most recognized databases by the scientific community.

1. The thematic focus on fields such as Information Science and Library Science, since the main publication outlets and the main authors are from these knowledge fields.

2. There is a high number of papers in the digital context, especially regarding websites. This can be reinforced by the given convergence between the main keywords (Digital libraries, Information architecture and World Wide Web) used for representing the publications. The keywords, when grouped individually, cluster around three main terms – “web”, “digital” and “technology” –, which is a result that is very coherent to what most of the articles’ systemic analysis showed: there is a representation of articles aimed towards practices, uses and elaboration methodologies in IA.

3. The thematic that is most related to “Library Websites” is “User Test”; With “Digital Libraries” there seems to be a relation between themes regarding “Practices, Uses, Elaboration Methodologies”; “Library Science and Information Science” is most related to “Teaching and Academic Directions”; and “Other Topics” is related to “Technological Interoperability and Infrastructure”.

4. Regarding the collaboration between authors, single, double and triple authorships are more common among the analyzed publications.

5. The most productive authors on the subject come from the fields of Information Science and Library Science, act as librarians, develop research or employ work related to the environment of libraries and of informational nature.

In conclusion, it can be stated that there is a convergence of works aimed towards IA and digital libraries studies, which is very coherent to what the literature about the subject points out: information architecture is a helper in the access to information, which makes digital information environment better signalized.
References


DAVIS, N. Framing the Practice of Information Architecture. UX Matters, 2007.


Dados dos autores

**Vitor Taga**  
Mestre em Ciência da Informação, Universidade Federal de Santa Catarina. Departamento de Ciência da Informação.  
vitor.taga@gmail.com

**Danielly Oliveira Inomata**  
inomata.danielly@gmail.com

**Caroline Rodrigues Vaz**  
Pós-Doutoranda na Universidade Federal de Santa Catarina. Departamento de Pós-Graduação em Engenharia de Produção.  
caroline.vaz@posgrad.ufsc.br

**Mauricio Uriona Maldonado**  
Professor na Universidade Federal de Santa Catarina. Departamento de Engenharia de Produção e Sistemas.  
m.uriona@ufsc.br

**Gregório Varvakis**  
Professor na Universidade Federal de Santa Catarina. Departamento de Engenharia e Gestão do Conhecimento, e do Departamento de Ciência da Informação.  
g.varvakis@ufsc.br

**Recibido - Received:** 2016-04-27  
**Aceito - Accepted:** 2017-05-14