Analysis of the SEO visibility of university libraries and how they impact the web visibility of their universities

Mari Vállez\textsuperscript{a,}\textsuperscript{*}, Anna Ventura\textsuperscript{b}

\textsuperscript{a} Department of Librarianship, Information Science and Audiovisual Communication, University of Barcelona, Molí d'en Palau 140, 08014 Barcelona, Spain
\textsuperscript{b} Department of Communication, Universitat Pompeu Fabra, Rec Boronat 138, 08001 Barcelona, Spain

\textbf{ARTICLE INFO}

Key words: University libraries Web visibility Visibility index Search engine optimization (SEO) Academic search engine optimization (ASEO) Library visibility University visibility

\textbf{ABSTRACT}

This article comparatively analyzes the web visibility of 20 libraries corresponding to the top 10 universities in the Times Higher Education World University Rankings and the 10 largest Spanish universities, and explores whether a correlation exists between the web visibility of libraries and of their respective universities. To study web visibility, a search engine optimization (SEO) tool called Sistrix Toolbox was used. It analyzes a large amount of data, the most notable of which is the visibility index, which combines different data indicators to analyze web visibility. The results are checked with Xavi, another SEO tool that offers its own visibility index. Both tools allow us to observe similar trends in the visibility of library websites.

The results show that university library visibility is generally low and that there is no direct correlation between the visibility index of libraries and that of their universities. Some revealing exceptions were identified, in which libraries have made significant contributions to the web visibility of their universities. The results would suggest that higher education institutions need to implement SEO strategies in order to increase their visibility more effectively.

\textbf{Introduction}

Search engine optimization (SEO) is a set of techniques whose purpose is to position a website in the organic results of search engines and thus achieve visibility (Yalçın & Köse, 2010). SEO use has become a necessity for any institution aspiring to visibility in the digital age. Being visible to search engines means being ranked highly as an information source for some 4.2 billion internet users who formulate around 4 billion daily Google searches.\textsuperscript{1} On average, 51% of a website’s traffic comes from organic searches (BrightEdge Research, 2017), i.e., more than half of the traffic to a website is channelled there by search engines. Therefore, good positioning in search engines means being able to respond to an important niche of user queries as well as standing out as a reference institution. Google rankings are especially important, as Google’s share of the search engine market is 92.42\%.

Universities websites need to adapt to the new information-seeking habits of users. Therefore, good SEO practices should be integrated not just by universities, but especially by their libraries, which need to provide quality information to their users in an efficient and effective manner (Balasubramanian & Sabarish, 2016; Onaifo & Rasmussen, 2013).

It would also be relevant to know if libraries can impact the web visibility of their universities. In this sense, this article answers three research questions. First, what is the visibility index of the selected university libraries in the sample? Second, does this visibility influence the domains of their universities? Finally, how does the visibility of the library subdomains influence the total web visibility of their respective universities?

Although studies of the application of SEO techniques to academic libraries websites exist (Gasparotto, 2014; Lee, Jang, Lee, & Oh, 2016), there are no studies that undertake an international comparison of the visibility index of different libraries. This article describes a comparative analysis of the visibility of 20 university libraries, with a particular emphasis on Spanish libraries.

The general objectives defined for this research were twofold: (1) to analyze and compare the visibility index of academic libraries websites, and (2) to explore whether their libraries enhance the visibility of universities. These objectives, expressed in more specific terms, were as follows:

\textsuperscript{*} Corresponding author.
\textsuperscript{1} E-mail address: marivallez@ub.edu (M. Vállez).
\textsuperscript{1} Data from Internet Live Stats http://www.internetlivestats.com/ [accessed July 15, 2019].
\textsuperscript{2} Data from StatCounter GlobalStats http://gs.statcounter.com/ [accessed July 15, 2019].
a. To analyze the website visibility of academic libraries.
b. To identify possible links between university domain visibility and the corresponding library subdomain.
c. To consider how library subdomains can enhance the visibility of universities.

This paper is organized as follows. First, we review the relevant literature focusing on the website visibility of university institutions and SEO applied to libraries. Next, we describe the methodology used to obtain the visibility index. We then report and discuss the results, and finally, summarize our conclusions, limitations and future lines of research.

Literature review

Web visibility in ACADEMIA

The visibility of universities is a long-standing topic of discussion in the academic literature (Björneborn & Ingwersen, 2004; Thelwall, Vaughan, & Björneborn, 2005). For example, the Webometrics Ranking⁴ which ranks universities according to their presence and impact on the web has been published since 2004 (Aguillo, Ortega, & Fernández, 2008). Several similar international ranking systems exist as a consequence of the globalization of higher education (Aguillo, Bar-Ilan, Levene, & Ortega, 2010). The same kind of analysis has also been applied to libraries (Orduña-Malea & Regazzi, 2013; Verma & Brahma, 2017).

There is a long tradition of measuring both the impact and outcome of libraries (Poll, 2003; Salisbury & Peasley, 2018). However, the impact of library websites on their institutions is poorly studied. For this reason, it is interesting to study whether university libraries provide web visibility to their universities as a whole.

Recent years have seen several studies of SEO in relation to the academic scope, covering topics such as SEO techniques applied to the diffusion of academic production (Codina, 2017; Shahzad et al., 2017), optimization of scholarly articles in Google Scholar (Beel, Gipp, & Wilde, 2009), university visibility in academic social networks (French & Fagan, 2019; González-Díaz, Iglesias-García, & Codina, 2015), and SEO applied to repositories and academic journals (Alhuay-Quipe, Quipe-Riveros, Bautista-Ynofuente, & Pacheco-Mendoza, 2017; Arltitsch & O'Brien, 2015; Shi, Cao, & Zhao, 2010). However, studies dealing specifically with SEO and the visibility of universities in search engines are scarce and existing studies focus on specific cases, whether local (Orduña-Malea, Serrano-Cobos, Ontalba-Ruízérez, & Lloret- Romero, 2010; Yalcın & Kılıç, 2016) or thematic (Gasparotto, 2014; Özkan, Özceylan, Kabak, & Dağdeviren, 2019).

SEO and libraries

Some studies on SEO applied to academic libraries websites have been undertaken. Rushton, Kelehan, and Strong (2008), for instance, described how Binghamton University librarians brought the library closer to its users using web positioning techniques; the authors viewed search engines as allies rather than as competition, given that SEO renders pages and content more visible online. By the end of the pilot project, it was observed that the investment of effort in SEO had improved the quality of the library’s website. However, in evaluating the results, the authors noted that they did not have a control that would have allowed them to compare results before and after the application of SEO strategies.

Another study reported the Open SESMO (Search Engine and Social Media Optimization) project of Montana State University’s library, which aimed to increase the use of subscribed platforms and databases. Different approaches were adopted, including web positioning, structured data definition (using Schema.org vocabularies), the use of linked data and social network diffusion (Clark & Rossmann, 2017). The results pointed to a significant increase in visits and content downloads.

In relation to SEO applied to libraries websites, library content has traditionally faced difficulties in being indexed by search engines, since much content is dynamically generated and forms part of the invisible web (Bergman, 2001; Blandford, 2015). Despite this drawback, users persist in using search engines to locate academic information (Blankstein & Wolf-Eisenberg, 2019; Gauder, 2011). This would suggest that libraries need to adapt to new modes for accessing information. Some proposals to facilitate this adaptation have been made, e.g., facilitating the indexing of online public access catalogues (OPACs) (Lewandowski, 2010) or promoting the semantic identity of libraries (Arlitsch, 2017; Arltitsch, O'Brien, Clark, Young, & Rossmann, 2014).

Contrasting with this openness to search engines, web invisibility has been defended by some researchers, who point to benefits in finding accurate and quality information in the shortest possible time, despite the need for specialist personnel (Alyami & Assiri, 2018; Devine & Egger-Sider, 2004; Su, 2009). This is especially important in sectors like medicine, where quality and up-to-date information is vital.

Leaving aside this dichotomy, librarians are accustomed to using metadata to facilitate access to library collections. So the use of SEO is not an approach that would be alien to their work philosophy. Dickinson and Smit (2016), in response to the question why, if SEO techniques are useful, they are not being applied on a large scale in libraries, reported that libraries generally do not have a policy regarding their visibility in search engines. According to (Arlitsch, O'Brien, Rossmann, 2013; Pérez-Montoro & Codina, 2016), the training and attitudes of librarians are fundamental to successful SEO implementation. Although such training may be lengthy, if SEO-related policies and strategies were established for universities, libraries could contribute with their full potential in the digital world.

Methodology

A quantitative design was used to evaluate the web visibility of libraries and how they impact on the visibility of their respective universities. Selected for analysis were 20 universities: the top 10 international universities listed in the Times Higher Education 2019 World University Rankings and the 10 largest Spanish universities in terms of number of students⁵ (see Table A1 in the Appendix).

Data on website visibility of the sample of university libraries captured using Sistrix Toolbox,⁴ resulting in a visibility index reflecting performance in Google (not SEO-traffic index to the website). The library subdomains were SEO-analyzed during the last week of April 2019. The Sistrix Toolbox crawls the web, calculates a visibility index for domains and shows data for comparable sites. The calculation is based on checking the top 100 Google weekly search results for one million keywords or combination of keywords (search phrases) reflecting search behaviours in different countries. Obtained each week are 100 million data items (one million keywords X 100 top search results) attributed to domains. The visibility index is weighted according to the volume of searches and the ranking of search results where it appears. The Sistrix Toolbox visibility index is not a website traffic index; instead, it reflects the position of a domain or subdomain in Google while excluding external factors that impinge on traffic, such


⁵ Accessible through https://www.sistrix.com/.

---

as seasonality (Härtling, Mohl, Steinhauser, & Möhring, 2016).

There are other SEO indicators which calculate web visibility. However, the visibility index of Sistrix Toolbox was chosen because it incorporates different indicators (number of keywords positioned, volume of searches, and inbound links) that facilitates the comparison of different domains in a more comprehensive way.

To check that the results obtained with Sistrix Toolbox were not biased, a control test was implemented with a tool with similar features, Xovi, which returned results that confirmed those obtained with Sistrix Toolbox (see Figs. A2–A4 in the Appendix). Keep in mind that since they are different tools, each one has its own way of displaying data. It is not possible to compare the same information, it only allows for the observation of similar trends.

The web visibility of the 20 university libraries websites was thus based on a performance indicator that considered the number of keywords visible in Google’s search engine, and that depicted how a website had evolved and how it compared with those of competitors.

Results

All of the data obtained from the Sistrix Toolbox can be found in Fig. A1 of the Appendix. These data are the source for the figures used in this section.

University Library visibility

Fig. 1 shows the visibility index for the 20 libraries analyzed. The collected data included the number of backlinks, the number of keywords in the top 100 search results and the number of URLs in the top 100 results. It can be observed that the visibility index of the sampled libraries is generally low, with many libraries websites having zero or near zero visibility. However, two libraries stand out for their significantly higher visibility; namely, the libraries of the Open University of Catalonia (UOC) and the Complutense University of Madrid (UCM), both in Spain. These institutions are followed at some distance by the University of Granada (UGR) and the University of Oxford.

Fig. 2, using data obtained directly from Sistrix Toolbox, shows how the visibility indexes have evolved since 2013 for the four libraries with the highest visibility index. The UCM library has maintained reasonably high visibility over time, although a decline since around 2018 coincided with a notable increase in UOC library visibility. Visibility of the UGR and University of Oxford libraries has remained largely steady, although with some notable peaks, especially for the UGR library in early 2017.

Fig. 3 is a scatter plot depicting the general and mobile device visibility indexes. The correlation trend is generally positive, with mobile visibility increasing with general visibility. The bubble size is based on the libraries’ visibility index. Notable exceptions, however, are the University of Oxford and University of Seville (USE) libraries, both of which have greater mobile visibility; in contrast, mobile device visibility for the Massachusetts Institute of Technology (MIT) and UGR libraries is low.

Regarding a possible correlation between the number of backlinks received by libraries and the number of URLs in the top 100 keyword results, the scatter plot in Fig. 4 points to no relationship between the two variables. In general, university libraries in the USA receive more backlinks compared to the number of URLs they position, whereas the opposite happens with Spanish university libraries. The UK universities as a whole do not follow either pattern, as indicated, for instance, by the University of Oxford and the University of Cambridge reflecting the two extremes.

Library visibility correlation with university visibility

The scatter plot in Fig. 5, depicting relationships between the visibility indexes of libraries and of their respective universities, points to no relationship in most cases and identifies some critical outliers. The UOC library is again notable, for a highly visible library and a largely invisible university, while the reverse is true of Harvard University and USE libraries.

The scatter plot in Fig. 6 shows the relationship between the number of the universities’ and libraries’ URLs that rank in the top 100 results, with the bubble size reflecting the number of the libraries’ URLs. Of the university libraries standing out from the rest, the case of the UOC is especially interesting: compared to other universities, the library manages to position far more URLs than the university does.

Library contribution to their university’s visibility

Subdomains for the universities with the highest number of keywords in the Google top 100 results were analyzed using the Sistrix Toolbox to assess how libraries might raise the visibility of their respective universities. An analysis of the subdomains pointed to how, in some cases, libraries and other subdomains linked to libraries feature, e.g., institutional repositories or academic journals (even though these do not always have an explicit link to the library). To determine their impact, the percentage of keywords for each subdomain linked to a library in relation to the keywords for the university domain as a whole was calculated. Fig. 7 shows the results, indicating how libraries impact on their university’s online visibility. It can be observed that the UOC and USE libraries contribute very substantially to their university’s visibility.

Discussion

The findings indicate that the visibility of university libraries websites is generally very low or zero, with a handful of exceptions, namely Spain’s UOC, USE and UCM libraries and the UK’s University of Oxford library. Some of the possible reasons for low library visibility are listed:

1. The library does not have its own separate subdomain, for example, if its URL is a directory of the university’s website address. This is the case of the library of the Autonomous University of Barcelona (UAB), whose web address is https://www.uab.cat/web/servicio-de-bibliotecas-1345733231312.html.
2. External tools that do not maintain the library subdomain are used, e.g., for thematic guides, meta-search engines, blogs, digital collections, etc. A specific example is the MIT’s library research guides at https://libguides.mit.edu/history.
3. The library subdomain is decentralized, which happens when faculties have their own libraries. An example is the University of Cambridge library network, with libraries depending on the corresponding faculty subdomain; for instance, the URL of the Faculty of Architecture and History of Art is https://www.aha.cam.ac.uk/Library. Greater visibility is achieved with models such as that used by the University of Oxford’s libraries; for example, its Medical Sciences library, for which URL is https://www.bodleian.ox.ac.uk/medicine.
4. Content is timeless and not linked to current issues, which is what generate more searches. A model of good practice is the UOC library, as, despite being one of the youngest universities in the sample, its library has the highest visibility index. One reason for this visibility is that the UOC library prepares dossiers on topical subjects; an example is its fake news dossier at http://biblioteca.uoc.edu/en/resources/fake-news.

Although libraries may take little interest in topical issues, doing so could bring them certain advantages over their competitors in terms of

*A Accessible through https://www.xovi.com/*
positioning their content in search engines. Successive modifications to Google’s algorithms\(^7\) can have favoured the type of content that libraries publish. Libraries, by their very nature, use white-hat SEO techniques\(^8\) to position themselves (Roslin & Shahirah, 2019; Scott, 2015) according to Google rules. Elements that add weight to visibility include using reputable backlinks in a natural way, using enriched formats, updating content regularly, generating content using synonyms and elements from related fields of meaning, avoiding duplicated content, promoting original quality resources, using visual features (images, graphs and videos) and optimizing websites for mobile devices.

While the content creation process of libraries is naturally favourable in terms of facilitating web visibility, it is also useful for libraries to establish a formal SEO policy, as the definition of a strategy about SEO content is key to visibility (Blakiston, 2013; Buchanan, 2017; Datig, 2018). An SEO policy should reconcile the library’s strategic plan and objectives, the institution’s idiosyncrasies, usage metrics, and user feedback (Kettunen, 2007). Training for, and the involvement of, librarians are especially essential to implementing an SEO strategy.

As evidenced by the findings of this research, improving library

---

\(^7\) Google Algorithm Update History. [https://moz.com/google-algorithm-change](https://moz.com/google-algorithm-change) [accessed July 15, 2019].

\(^8\) White-hat SEO refers to the usage of optimization strategies, techniques and tactics that focus on a human audience opposed to follow search engine rules and policies completely.
Conclusions

This research provides a general overview of the web visibility of leading university libraries websites and demonstrates its importance in positioning the respective universities in search engines. The analysis focuses on the world’s top 10 universities (located in the UK and USA) and the 10 largest universities in Spain. The results show that university libraries generally have little visibility in search engine results rankings, with some exceptions, namely, the UOC, UCM and University of Oxford libraries.

As for the relationship between the web visibility of universities and of their libraries, there is no direct correlation. Most universities have high visibility relative to their libraries, with a notable exception being the UOC library, where the reverse occurs. Interestingly, libraries whose visibility index is high have other services associated with them – such as institutional repositories or academic journals – that provide outstanding visibility to their university domains.

Web positioning can be enhanced by tackling the reasons for low visibility listed in the discussion, not least, the critical issue of libraries having their own subdomain. More generally, libraries above all need to adapt to new information access modes and deploy effective SEO strategies. In addition, it would be interesting to monitor the SEO strategies implemented to measure their impact.
Limitations and future lines of research

This research has two main limitations. First, the fact that much library content is beyond the reach of search engines in the invisible web or is only accessible through tools external to the library has a direct impact on the visibility of libraries. Second, use of the commercial Sistrix Toolbox could have introduced an element of bias, although this was minimized by implementing control tests using Xovi, which ratified the results obtained with the Sistrix Toolbox.

To further explore the visibility of higher education institutions, an SEO audit is planned for the most visible libraries as identified in this research. This audit will allow an in-depth study of the value of the contributions of libraries and make recommendations aimed at improving the visibility of academic institutions and university library websites.

Acknowledgements

Our thanks to Carlos Gonzalo for helping us to evaluate SEO for libraries and Lluís Codina for his contributions and for reviewing the text. Special thanks to the anonymous reviewers who have contributed to improving the work.

Funding

There was no outside source of funding.

Availability of data and materials

The datasets used and/or analyzed are available from the corresponding author on reasonable request.

Authors’ contributions

MV conceived the research idea. MV and AV designed the research framework. MV collected and analyzed the data and extracted the results. MV and AV drafted and reviewed the manuscript. The two authors read and approved the final version of the manuscript.

Declaration of competing interest

The authors declare that they have no competing interests.

Appendix A

Table A1

<table>
<thead>
<tr>
<th>University</th>
<th>Acronym</th>
<th>Country</th>
<th>URL Library</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanford University</td>
<td>SU</td>
<td>US</td>
<td><a href="http://library.stanford.edu/">http://library.stanford.edu/</a></td>
</tr>
<tr>
<td>Massachusetts Institute of Technology</td>
<td>MIT</td>
<td>US</td>
<td><a href="https://libraries.mit.edu/">https://libraries.mit.edu/</a></td>
</tr>
<tr>
<td>California Institute of Technology</td>
<td>Caltech</td>
<td>US</td>
<td><a href="https://library.caltech.edu/">https://library.caltech.edu/</a></td>
</tr>
<tr>
<td>Harvard University</td>
<td>HU</td>
<td>US</td>
<td><a href="https://library.harvard.edu/">https://library.harvard.edu/</a></td>
</tr>
<tr>
<td>Princeton University</td>
<td>PU</td>
<td>US</td>
<td><a href="http://library.princeton.edu/">http://library.princeton.edu/</a></td>
</tr>
<tr>
<td>Yale University</td>
<td>YU</td>
<td>US</td>
<td><a href="https://web.library.yale.edu/">https://web.library.yale.edu/</a></td>
</tr>
<tr>
<td>University of Chicago</td>
<td>Uch</td>
<td>US</td>
<td><a href="https://www.lib.uchicago.edu/">https://www.lib.uchicago.edu/</a></td>
</tr>
<tr>
<td>University of Oxford</td>
<td>OX</td>
<td>UK</td>
<td><a href="http://www.bodleian.ox.ac.uk/">http://www.bodleian.ox.ac.uk/</a></td>
</tr>
<tr>
<td>University of Cambridge</td>
<td>UC</td>
<td>UK</td>
<td><a href="http://www.lib.cam.ac.uk/">http://www.lib.cam.ac.uk/</a></td>
</tr>
<tr>
<td>Imperial College London</td>
<td>ICL</td>
<td>UK</td>
<td><a href="https://www.imperial.ac.uk/admin-services/library/">https://www.imperial.ac.uk/admin-services/library/</a></td>
</tr>
<tr>
<td>The National Distance Education University</td>
<td>UNED</td>
<td>ES</td>
<td><a href="http://portal.uned.es/portal/page?_pageid=93,50543">http://portal.uned.es/portal/page?_pageid=93,50543</a></td>
</tr>
<tr>
<td>Complutense University of Madrid</td>
<td>UCM</td>
<td>ES</td>
<td><a href="https://biblioteca.ucm.es/">https://biblioteca.ucm.es/</a></td>
</tr>
<tr>
<td>University of Barcelona</td>
<td>UB</td>
<td>ES</td>
<td><a href="http://crai.ub.edu/">http://crai.ub.edu/</a></td>
</tr>
<tr>
<td>University of Seville</td>
<td>USE</td>
<td>ES</td>
<td><a href="https://bib.us.es/">https://bib.us.es/</a></td>
</tr>
<tr>
<td>University of Granada</td>
<td>UGR</td>
<td>ES</td>
<td><a href="https://biblioteca.ugr.es/">https://biblioteca.ugr.es/</a></td>
</tr>
<tr>
<td>University of Valencia</td>
<td>UV</td>
<td>ES</td>
<td><a href="https://www.uv.es/uvweb/servicio-bibliotecas-docu">https://www.uv.es/uvweb/servicio-bibliotecas-docu</a></td>
</tr>
<tr>
<td>Open University of Catalonia</td>
<td>UOC</td>
<td>ES</td>
<td><a href="http://biblioteca.uoc.edu/">http://biblioteca.uoc.edu/</a></td>
</tr>
<tr>
<td>King Juan Carlos University</td>
<td>URJ</td>
<td>ES</td>
<td><a href="https://www.urjc.es/estudiar-en-la-urjc/biblioteca">https://www.urjc.es/estudiar-en-la-urjc/biblioteca</a></td>
</tr>
<tr>
<td>University of the Basque Country</td>
<td>UPV</td>
<td>ES</td>
<td><a href="https://www.ehu.eus/es/web/biblioteca">https://www.ehu.eus/es/web/biblioteca</a></td>
</tr>
<tr>
<td>Autonomous University of Barcelona</td>
<td>UAB</td>
<td>ES</td>
<td><a href="https://www.uab.cat/web/service-de-bibliotecas-134">https://www.uab.cat/web/service-de-bibliotecas-134</a></td>
</tr>
</tbody>
</table>
Fig. A1. All data recollected from Sistrix Toolbox.

Fig. A2. Control test with Xovi for Spanish university libraries.

Fig. A3. Control test with Xovi for UK university libraries.
Fig. A4. Control test with Xovi for USA university libraries.

References


