Esteban Romero Frías y

María Sánchez González (editores)

Prólogos de Paul Spence y de Nuria Rodríguez Ortega

Esteban Romero Frías; Octavio Rubiera; Belén Álvarez; Carlos Arcila; Mabel Calderín Cruz; Luis Núñez; Ysabel Briceño; Elika Ortega; Silvia Eunice Gutiérrez; María Sánchez González; Daniel Villar-Onrubia; Juan M. García Sierra; Manuel D. Pérez Heredia; Javier Jiménez Cuadros; Patricia Lázaro Pernias; Daniel Tena Parera; David Roca Correa; Josep María Blanco Pont; Alfonso González Quesada; Álvaro Baraibar; Victoria Tur-Viñes; María del Carmen Quiles-Soler; José Manuel de Pablos; Ana María Rodera Bermúdez; Ana María González Ramos; Ernesto Priani Saisó

Ciencias Sociales y Humanidades Digitales

Técnicas, herramientas y experiencias de *e-Research* e investigación en colaboración

CAC, Cuadernos Artesanos de Comunicación / 61

Universidad de La Laguna
Universitat d'Alacant Universidad de Alicante
USC Universidade de Santiago de Compostela
Universidad de Málaga
Sociedad Latina de Comunicación Social

The value of Webmentions as data: mapping attention to the notion of OER in the HE arena

Daniel Villar-Onrubia. Oxford Internet Institute, University of Oxford (Reino Unido) daniel.villaronrubia@oii.ox.ac.uk

1. Introduction

SOCIAL actors increasingly leave traces behind as they share messages, seek information and, more generally, behave in the digital media environment. As Lee *et al.* put it, the social world is "becoming self-documenting and self-archiving" (2008, 8). Thus, the heterogeneous range of practices articulated around the use of information and communication technologies (ICTs) offers unprecedented opportunities to identify and track people's interests, preoccupations, concerns, preferences and even moods. For instance, the queries submitted to a search engine by their users or the messages posted on a micro-blogging service might be approached as data, allowing researchers to draw conclusions on a given population.

This paper focuses on a specific type of online research technique that draws on the analysis of word/s occurrence (i.e. Web mentions) as a way of assessing the level of attention paid to certain concepts on the Web. Apart from referring to relevant literature and reviewing two specific research tools, the paper also illustrates the procedures of data collection and analysis by means of two original studies that

look at the level of attention generated by the notion of Open Educational Resources (OER) across Higher Education (HE) institutions and systems in several countries. The studies were presented at the UNESCO OER World Congress (Villar-Onrubia, 2012) and the OER13 Conference (Villar-Onrubia, 2013) as a work-in-progress. The data on Latin American countries was collected in the context of the OportUnidad Project,[103] with the aim of gaining insight into the overall level of attention paid to OER across universities in such region.

2. Approaching Web mentions as an indicator of relevance

The study of online influence has typically relied on relational approaches, drawing on social network analysis techniques that allow researchers to look at interconnection patterns between different kinds of actors, which are treated as nodes (Ackland, 2013; Maeyer, 2013; Hogan, 2008). However, the procedures examined in this paper are based on the study of influence as inferred from textual content posted on the public Web. More specifically, the occurrence of phrases across webdomains is taken here as an indicator of attention to the issue under investigation.

A Web mention can be defined as a "textual mention in a webpage, typically of a document title or person's name. Nevertheless, a web mention encompasses any non-URL textual description" (Thelwall and Sud, 2011: 1490). The number of webdomains and/or specific URLs mentioning a given element –whether a person, organisation, brand, work or even an idea– may be taken as an indicator of relevance or popularity, which is often conceptualised in the literature as 'impact' or 'resonance.'

There are two main research traditions that draw on Web mentions as an indicator of relevance, namely Webometrics and Digital Methods. Apart from taking Web mentions as a valuable source of evidence,

[103] OportUnidad (www.oportunidadproject.eu) is an action-research project, funded by the European Commission (Programme Alfa III), that aims to foster the adoption of open educational practices across more than 60 HE institutions in Latin America.

researchers working in these two traditions have engaged in the development of research tools that facilitate these kinds of data collection and analysis procedures. Such tools rely on data supplied by commercial search engines, being "capable of automatically submitting queries to search engines and then downloading, saving, and processing the results" (Thelwall, 2009a: 57). Thus, they may be defined as *automatic search engine query submitters* (Thelwall, 2009).

2.1. Webometrics

Webometrics is as a field of research devoted to "the study of web-based content with primarily quantitative methods for social science research goals using techniques that are not specific to one field of study" (Thelwall, 2009a: 6). A method-centred definition like this one implies that the overall community of social scientists, beyond definite disciplinary boundaries, might benefit from Webometrics methods and techniques. However, this field has been particularly relevant to researchers working in the area of Information Science, and more specifically to infometricians. Indeed, it "emerged from the realization that methods originally designed for bibliometric analysis of scientific journal article citation patterns could be applied to the Web" (Thelwall, Vaughan and Björneborn, 2005: 81).

Bibliometric research focuses on citation behaviour, usually taking citation counts as an indicator of the level of impact of scholarly worksand prominence of their authors. Those works and authors that are cited most often tend to be seen as being more influential, or having more impact, than those that are quoted less widely. Citation counts work as indicators of scientific performance (Bornmann and Daniel, 2008). They are interpreted as synonymous of quality by most research assessment frameworks and, thus, are key to the accumulation of scientific capital, being often "used to evaluate scholars for hiring, promotion, funding, and other rewards" (Borgman, 2007: 63).

In the late nineties, the authors of a pioneering study (Cronin *et al.*, 1998) decided to use Web mentions retrieved via various commercial search engines in order to assess the influence of five prominent scholars. One of the main goals of the study was to elaborate a typology of the manifold online contexts, alternative to the sources

traditionally used in bibliometric research (i.e. academic publications), in which scholars and their works may be mentioned:

> "While traditional citation analysis can tell us a lot about the formal bases of intellectual influence, it, quite naturally, tells us nothing about the many other modalities of influence which comprise the total impact of an individual's ideas, thinking, and general professional presence."(Cronin *et al.*, 1998: 1326)

Impact assessments based on Web mentions take as a point of departure the assumption "that, other factors being equal, documents or ideas having more impact are likely to be mentioned online more" (Thelwall, 2009a: 9). Like most Webometrics techniques, Web impact assessments usually aim to study the salience of the element/s under investigation on the entire World Wide Web – to be more precise on the publicly indexable Web (Thelwall, Vaughan and Björneborn, 2005) – as represented by a sample of URLs crawled by some of the major commercial search engines. Nevertheless, the scope of webometric analyses may be also restricted to certain types of domains, such as blogs or news sites, or even to a specific set of domains or websites (e.g. Thelwall, Vann and Fairclough, 2006).

Web impact assessments are particularly useful for comparative purposes, as they allow researchers to "compare the influence, spread, or support of competing academic theories, political candidates, or a number of similar books" (Thelwall, 2009a: 9). This type of studies have been particularly concerned with the higher education (HE) arena, whether aiming to assess the impact of leading scholars (see also Thelwall, 2009b), academic organisations (Thelwall and Sud, 2011), journal articles (Vaughan and Shaw, 2004) or scholarly digital resources (Meyer, 2011). Nevertheless, online impact assessments are equally suitable to studies that do not focus on academic contexts and actors (e.g. Wilkinson, Sud and Thelwall, 2014).

At the Statistical Cybermetrics Research Group of the University of Wolverhampton (UK), Mike Thelwall has led the development of several online research tools specifically aimed at social scientist, most notably the Webometrics Analyst[104] –formerly known as LexiURL

[104] http://lexiurl.wlv.ac.uk/

Searcher– and the SocSciBot.[105] The Webometrics Analyst is a free Windows program that relies on data gathered by means of the Application Programming Interface (API) of Bing –a commercial search engine own by Microsoft. It is able to generate various kinds of outputs, including link impact reports, network diagrams and Web impact reports. The latter output displays the number of URLs, domains, sites, second level domains (STLDs) and top level domains (TLDs) matching the query or queries(see Table 1 for a definition of each element).

Table 1. Glossary of elements included in the Web impact reports generated by the Webometrics Analyst

Elements	Definitions
URLs	The number of URLs returned by the search engine (NOT the estimated number of URLs it reports).
Domains	The domain names of the URLs matching the query.[106]
Sites	The distinguishing end of the domain names of the URLs matching the query (e.g., microsoft.com, ox.ac.uk, w3.org, yahoo.co.uk - it is always the SLD plus one extra section on the left).
STLD	The second level domain (when existing, otherwise the top level domain) of the URLs matching the query (e.g., .com, .ac.uk, .edu, .co.nz).
TLD	The top level domain of the URLs matching the query (e.g., .com, .uk, .edu, .nz).

Source: definitions of terms as included in the overview of search results supplied by the Webometrics Analyst.

Commercial search engines only return a limited number of webpages per query, usually under 1,000. In order to overcome such limitation when a certain query matches a higher number of URLs, the professional version of the Webometrics Analyst offers an advanced

[105] http://socscibot.wlv.ac.uk/

[106] Domain names are equated with domain name segments that are allocated to particular micro-sites (e.g. oii.ox.ac.uk) within wider sites (e.g. ox.ac.uk).

technique called query-splitting, which increase the number of URLs that can be retrieved. This functionality involves the automatic selection of a term mentioned in the titles and snippets of 10% of the first results returned by Bing. After that, the software performs two new queries, one adding such term to the original query and the other subtracting it. The researcher must specify the number of modified versions of the original query to be submitted (up to 10) and, after completing the process a report combining the results of all the queries is generated (for further information see Thelwall, 2009: 93-94).

2.2. Digital Methods

Digital Methods is an epistemological programme concerned with "how to diagnose cultural change and societal conditions with the Internet"(Rogers, 2009: 8). In this regard, it takes online artifacts (e.g. hyperlinks, websites, search engines, hits, likes, etc.) as sources of evidence. As Rogers puts it:

> "… it is a proposal to reorient the field of Internet-related research by studying and repurposing what I term the methods of the medium, or perhaps more straightforwardly methods embedded in online devices. For example, crawling, scraping, crowd sourcing, and folksonomy, while of different genus and species, are all web techniques for data collection and sorting." (Rogers, 2013: 1)

The Digital Methods Initiative[107] (DMI), a collaboration between the University of Amsterdam and the Govcom.org Foundation led by Richard Rogers, has released a comprehensive collection of online research tools specifically designed to:

> "… extend the research into the blogosphere, online news sphere, discussion lists and forums, folksonomies as well as search engine behavior. These tools include scripts to scrape web, blog, news, image and social bookmarking search engines, as well as simple analytical machines that output data sets as

[107] https://www.digitalmethods.net

well as graphical visualizations." (Digital Methods Initiative, n.d.)

Web mentions are also taken as a valuable source of data under the Digital Methods programme. The so-called Google Scraper,[108] afterwards transformed and renamed as the Lippmannian Device,[109] is an online research tool specifically devised to enable the analysis of "partisanship and issue resonance" (Rogers, 2013: 113), relying on data retrieved by means of Google's API. In particular, the Lippmannian Device facilitates the submission of batch queries to Google, making it possible to query one or various URLs or domains, named as 'sources,' for one or more keywords, named as 'issues.'

Inspired by the work of Walter Lippmann on the study of public opinion (1922; 1927), the goal of this tool is to offer a "coarse view of the partisanship of an actor" (Rogers, 2010: 253). In other words, it aims to help researchers to gain insight into the level of commitment of various actors with one or more issues. For instance, one of the first studies making use of this tool looked at the extent to which a number of well-known climate change sceptics were mentioned across the first 100 sites returned by Google after submitting the query "climate change" (Rogers, 2010; Rogers, 2013).

The first steps involved compiling the list of sceptics and extracting the URLs of the most relevant sites –according to Google's PageRank algorithm– in relation to climate change. Subsequently, the Lippmannian Device was used in order to query each of the sources for the names of all the sceptics.

> "There were skeptic-friendly sites where the names of the skeptics resonate, such as Marshall.org (Marshall Institute), and there were watchdog sites, such as Sourcewatch.org, where the skeptics too received a great deal of scrutiny." (Rogers, 2013: 114).

The current version of the Lippmannian Device requires off-loading the requests to a client (i.e. the researchers' computer) by means of a

[108] https://wiki.digitalmethods.net/Dmi/ToolGoogleScraper
[109] https://wiki.digitalmethods.net/Dmi/ToolLippmannianDevice

Firefox extension.[110] That is, researchers submit the queries to Google using their own IP. Therefore, results might be exposed to potential biased resulting from search histories and other mechanisms that contribute to making search results personalised. A few steps must be followed in order to avoid that Google's customisation of results biases the scrapes. One of the potential solutions entails creating a specific Firefox profile that researchers should reserve for use in combination with the Lippmannian Device. In addition, before submitting the queries all Google services should be logged out and the personalisation based on "signed-out search activity" must be also disabled.[111]

Both the Webometrics and the Digital Methods traditions imply that Web mentions might be observed "either to directly assess web impact or to indirectly assess predominantly offline impact through measurement of the online component of that impact" (Thelwall, 2009a: 26). In this regard, Rogers introduces the term 'online groundness' in relation to "research that follows the medium [i.e. the Internet], captures its dynamics, and makes grounded claims about cultural and societal change" (2013: 23) that go beyond the online realm. While online data can sometimes be helpful in order to identify patterns and trends that are representative of entire organisations and other populations, it is not always the case (e.g. Mitchell and Hitlin, 2013). Indeed, the usefulness of data and credibility of findings depends on both the issues and populations under analysis.

3. Mapping attention to OER via Web mentions

The phrase Open Educational Resources (OER) refers to "teaching, learning, and research resources that reside in the public domain or have been released under an intellectual property license that permits their free use or re-purposing by others" (Atkins, Brown and Hammond, 2007: 4). Over the last few years, the idea of fostering the release of OER has captured the imagination of governments, organisations and educational institutions all over the world. For instance, the UNESCO, the Commonwealth of Learning or the

[110] https://wiki.digitalmethods.net/Dmi/FirefoxToolBar
[111] https://www.google.com/history/optout

European Commission are acting as some of the key advocates at an international level (UNESCO, 2012; UNESCO and COL, 2011; European Commission, 2013). This section discusses the value of Web mentions as an indicator of the level of attention paid to the notion of OER across universities and HE systems at several levels of aggregation (i.e. organisation, country and region). In particular, it focuses on Spanish-speaking countries.

In those studies that take Web mentions as a source of data, the way researchers construct their queries is crucial for the reliability and overall quality of the analysis. Queries might produce a considerable number of spurious matches, so certain criteria ought to be followed in order to maximise the meaningfulness and quality of data, for example avoiding the use of general or polysemic phrases and checking afterwards that "the overwhelming majority (e.g. 90%) of the pages returned correctly mention the desired document or idea" (Thelwall, 2009a: 11). The queries used in the analyses discussed here are specific enough to ensure that the chances of producing spurious matches are extremely low. The original term in English (i.e. "open educational resources") and the two equivalent expressions in Spanish (i.e. "recursos educativos abiertos" and "recursos educativos en abierto") are used in the procedures discussed below.

Before looking at particular sectors and countries, the Webometrics Analyst might be used in order to gain insight into the overall impact of the concepts under examination on the entire Web, as indexed by Bing.

Table 2 shows an overview of the results reported by the Webometrics Analyst after submitting the two phrases in Spanish.

Table 2. Overview of search results supplied by the Webometrics AnalystBase query	URLs	Domains	Sites	STLDs	TLDs
"recursoseducativosabiertos"	601	469	250	31	26
"recursoseducativos en abierto"	133	91	60	10	10

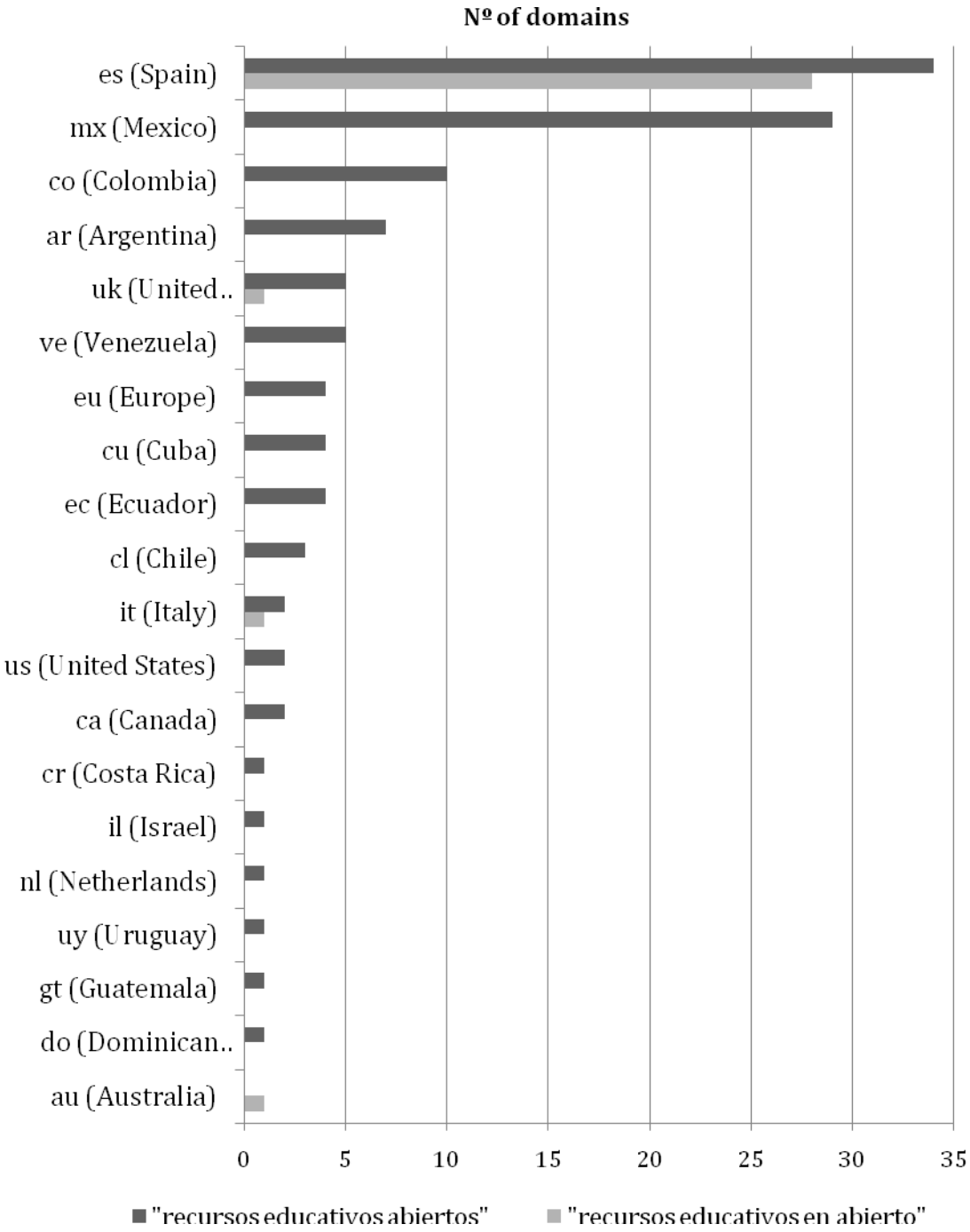
Source: Data gathered from Bing by means of the Webometrics Analyst, in August 2013.

As noted by the Webometrics Analyst, it should be borne in mind that search engines do not index all webpages. Likewise, search engines do not necessarily report all the URLs matching a particular query. In addition, as already mentioned, search engines return a maximum of 1,000 results per query. Therefore, the information supplied by the Webometrics Analyst may be incomplete.

In spite of those limitations and potential flaws, the results supplied by this tool are valuable in relative terms, that is, for comparative purposes. For instance, we can be quite confident that the first phrase is much more usual than the second one. Whereas such conclusion might be inferred from both URLs and domains count, the "most reliable impact indicator is normally the number of domains rather than the number of URLs because of the possibility that text or links are copied across multiple pages within a web site" (Thelwall, 2009a: 62). As already noted, domains are equated with the domain name segments that may identify specific micro-sites within wider websites (see Table 1).

Since many top-level domains (TLDs) are associated with specific regions or countries (i.e. country code top-level domains), the data gathered by means of this tool can be also useful in order to gain insight into the geographic contexts in which the queried phrases are used most often. Figure 1 shows the number of domains matching the queries that are associated with a country code TLD. In any case, it should be taken into account that most of the domains returned by both queries are under generic TLDs, namely .com, .org and .edu.

Figure 1. Country code Top-Level Domains of pages matching the following queries: "recursos educativos abiertos" and "recursos educativos en abierto".



Source: Data gathered from Bing by means of the Webometrics Analyst, in August 2013.

As noted by Thelwall, content analysis "is normally an important component of web impact analyses because of the need to interpret

the figures produced" (2009a: 26). In order to get a more nuanced understanding of the contexts in which the queried phrases were mentioned, it is advisable to analyse either a random sample or all the URLs matching the queries –depending on the overall size of the study and resources available. A typology of online mentions can be created in an inductive way, for instance, focusing on the nature of the organisations (e.g. universities, NGOs, governmental agencies, personal sites, etc.) that mention the phrases or the specific types of sites or documents were mentions are included (e.g. blog-posts, academic syllabuses, mission statements, etc.). The results of this analysis will be reported in forthcoming works.

While the Webometrics Analyst was devised to assess the impact of issues on the entire Web, at least as indexed by Bing, the Lippmannian Device is only suitable for analysis restricted to a predefined set of domains or sites. In this regard, it is particularly helpful in studies aiming to cover sectors with well-defined boundaries. When trying to gain insight into the level of attention paid to OER across the HE sector of a particular country or group of countries, the first step involves creating a list, as complete as possible, of institutions and their domain names. Covering more than 21,000 HE institutions from all over the world, the *Ranking Web of World Universities* (CSIC - Cybermetrics Lab, 2013) is probably the most comprehensive register of university domains available. Besides universities, this ranking also encompasses the sites of other types of HE institutions and centres, such as research institutes and also units (e.g. schools or departments) with their own domain names.

After extracting the domain names of all the relevant HE institutions (i.e. from Spain and the Spanish-speaking countries in the Latin American region) included in the above-mentioned ranking, the Lippmannian Devices may be used to submit the queries. In order to carry out a comparative analysis between countries, it is necessary to submit the queries to the domains of each country separately.

Using the Boolean operator OR, one single query (Query 1 = "recursoseducativosabiertos" OR "recursoseducativos en abierto") can retrieve all the URLs that mention at least one of the phrases in Spanish. Taking into account that the OER concept was originally coined in English, it is worth it to submit a second query aimed at

assessing the extent to which the phrase "open educational resources" (Query 2) is used across HE institutions in Spanish-speaking countries.

The output supplied by the Lippmannian Device makes it possible to tell apart those institutions that seem to be paying some level of attention to the notion of OER in each country from those that do not seem to do so.

Of course, this is not to deny that people and organisations that do not write about OER on the Web might still be aware of the notion and associated practices. Nevertheless, the fact that someone within a particular organisation is writing online about a given issue can be construed as a sign of stronger attention or commitment, either at a personal or organisational level. Moreover, in the case of issues that are directly related to digital practices, such as OER, it seems reasonable to assume that certain level of interest or attention is likely to prompt the occurrence of Web mentions.

By aggregating the data, it is also possible to get an overview of the level of attention paid to OER at a country level and, therefore, to carry out a cross-country analysis. Table 3 shows the overall number of HE institutions' sites per country, as indexed in the *Ranking Web of World Universities*, along with the number of sites matching any of the queries in each country.

The data allow us to assess the level of attention at a country level not only in absolute terms (i.e. those countries with a high number of institutions matching the queries), but also in relative terms (i.e. taking into account the proportion of institutions matching the queries in relation to the overall number of institutions).

Table 3. HE institutions sites per country and number of sites matching at least one of the queries

Country	HE Institutions' sites	Sites matching any of the queries
Spain	229	63
Mexico	898	47
Colombia	285	42
Argentina	114	24
Ecuador	59	17
Uruguay	34	17
Chile	78	15
Venezuela	72	14
Peru	92	12
Cuba	26	6
Dominican Rep.	32	6
Puerto Rico	39	5
Costa Rica	64	4
Salvador	37	4
Guatemala	19	3
Nicaragua	41	3
Paraguay	42	3
Bolivia	44	2
Panama	29	2
Honduras	12	1
Total	2246	290

Source: HE institutions as indexed by the *Ranking Web of World Universities*, edition of July 2013. Data gathered from Google by means of the Lippmannian Device in September 2013.

By separating the queries by language it is also possible to spot difference across countries. Although, in this population, Web mentions of the notion of OER are more usual in Spanish than in English, there are some countries in which the difference is very small

or the proportions are even inverted (i.e. Bolivia, Guatemala and Spain).

Since the Lippmannian Device counts the URLs matching each query (providing there are less than 1,000 matches) across all the domains under examination, it is also possible to identify those institutions that are particularly concerned with the idea of OER[112]. For instance, there are four universities in Spain that stand out in this regard: the UniversitatOberta de Catalunya, the Universidad Politécnica de Madrid, the Universidad de Alicante and the Universidad Nacional de Educación a Distancia. Moreover, by taking into account the characteristics of each institution (e.g. age, private vs. public nature, location, etc.) it is possible to identify relevant patterns. For instance, private universities in Spain seem to be less likely to pay attention to OER than public ones.[113]

As already noted, domain counts are more reliable than URL counts. Thus, it is of utmost importance to cleanse the data in order to make any interpretation based on URLs counts more robust, minimising the presence of duplicates (i.e. same content under different URLs) that could bias the conclusions. Moreover, it is important to carry out also a qualitative content analysis at this level, as the "variety of reasons why a web page could be created […] make it difficult to give a simple explanation of what a count of online mentions really means (Thelwall, 2009a: 17).

University domains tend to consist of a highly diverse range of sites and pages that are not only published and maintained by different kinds of actors, but also targeted at different types of audiences. In this regard, the implications of finding the term OER embedded into a mission statement are quite different from the implications of finding the very same term in a blog-post, a reading list or maybe a brief comment at an online forum.

---

[112]See Villar-Onrubia (2012) for a preliminary analysis of the level of attention paid to the notion of OER across HE organisations in Spanish-speaking countries in Latin America.

[113] The findings of an analysis focusing on the attention to OER across Spanish universities will be presented in a forthcoming paper (for preliminary results see Villar-Onrubia, 2013).

4. Conclusion

This paper focuses on the value of Web mentions as a source of data in social research. After 1) looking at certain research traditions that make use of this approach and 2) reviewing two online research tools that facilitate the collection of this type of data, it illustrates the potential of Web mention analyses with examples. In particular, it focused on the opportunities for assessing the level of attention paid to the notion of OER across the HE systems of Spanish-speaking countries.

Despite the limitations of data supplied by commercial search engines, and provided that some measures are taken in order to minimise potential flaws, Web mentions can be a very useful source of evidence in relation to the level of interest generated by certain issues (e.g. people, ideas, documents, etc.). This type of data may be especially relevant in the context of studies on issues that are somehow to do with digital practices (e.g. OER), as attention might be more likely to lead to the occurrence of online mentions.

However, findings exclusively based on Web mentions analyses should be seen just as providing a coarse perspective (Rogers, 2010), in other words, as indicative instead of definitive (Thelwall, 2009a). Therefore, these kinds of exploratory analyses can help researchers to chart emerging areas or serve as the basis for subsequent explanatory studies, drawing on a wider range of methods and techniques.

References

Ackland, R.(2013). *Web Social Science: Concepts, Data and Tools for Social Scientists in the Digital Age*. Los Angeles, Calif.: SAGE.

Atkins, D.E., Brown,J.S. &Hammond, A.L. (2007). "A Review of the Open Educational Resources (OER) Movement: Achievements, Challenges, and New Opportunities". Report to The William and Flora Hewlett Foundation.

Borgman, C.(2007). *Scholarship in the Digital Age: Information, Infrastructure*. Cambridge, Mass: MIT.

Bornmann, L. & Hans-Dieter, D. (2008). "What Do Citation Counts Measure? A Review of Studies on Citing Behavior".*Journal of Documentation* 64(1) (January 18): 45–80.

Cronin, B., Snyder, H.W., Rosenbaum, H., Martinson, A.& Callahan, E. (1998). "Invoked on the Web". *Journal of the American Society for Information Science,* 49(14): 1319-1328.

CSIC - Cybermetrics Lab (2013). "Ranking Web of Universities July 2013: Spain".Disponible en:http://webometrics.info/(consultado el 13 de octubre de 2013).

European Commission (2013). "Open Education Europa". Disponible en: http://www.openeducationeuropa.eu/ (consultado el 13 de octubre de 2013).

Hogan, B.(2008). "Analyzing Social Networks via the Internet". In *The SAGE Handbook of Online Research Methods*, edited by Nigel Fielding, Raymond M Lee, and Grant Blank, 141-160. Los Angeles: SAGE.

Lee, R.M., Fielding, N.G. & Blank, G.(2008). "The Internet as a Research Medium". In *The SAGE Handbook of Online Research Methods*, edited by Nigel Fielding, Raymond M Lee, and Grant Blank, 3-20. Los Angeles: SAGE.

Lippmann, W. (1922). *Public Opinion*. New York: Macmillan.

Lippmann, W. (1927). *The Phantom Public: a Sequel to 'Public Opinion'*. New York: Macmillan.

Maeyer, J. (2013). "Towards a Hyperlinked Society: A Critical Review of Link Studies". *New Media & Society* 15(5) (August 1): 737–751.

Meyer, E.T. (2011). *Splashes and Ripples: Synthesizing the Evidence on the Impacts of Digital Resources*. London: JISC.

Mitchell, A.&Hitlin, P.(2013). "Twitter Reaction to Events Often at Odds with Overall Public Opinion". Pew Research Center. Disponible en: http://www.pewresearch.org/2013/03/04/twitter-reaction-to-events-often-at-odds-with-overall-public-opinion/ (consultado el 13 de octubre de 2013).

Rogers, R. (2009). *The End of the Virtual: Digital Methods*. Amsterdam: Amsterdam University Press.

Rogers, R. (2010). "Internet Research: The Question of Method—A Keynote Address from the YouTube and the 2008 Election Cycle in the United States Conference". *Journal of Information Technology & Politics* 7 (2-3): 241–260.

Rogers, R. (2013). *Digital Methods*. Cambridge, Massachusetts□; London, England: The MIT Press.

Thelwall, M.(2009a). *Introduction to Webometrics: Quantitative Web Research for the Social Sciences*. Morgan & Claypool.

Thelwall, M.(2009b). "A Webometric Analysis of Olle Persson". *Celebrating Scholarly Communication Studies Communication Studies*: 61.

Thelwall, M. &Sud, P. (2011). "A Comparison of Methods for Collecting Web Citation Data for Academic Organizations". *Journal of the American Society for Information Science and Technology,* 62(8): 1488–1497.

Thelwall, M., Vann, K.,& Fairclough, R. (2006). "Web Issue Analysis: An Integrated Water Resource Management Case Study". *Journal of the American Society for Information Science and Technology,* 57 (10): 1303-1314.

Thelwall, M.,Vaughan, L. & Björneborn, L. (2005). "Webometrics". *Annual Review of Information Science and Technology*, 39 (1): 81-135.

UNESCO(2012). "Paris OER Declaration". Disponible en: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/Events/Paris%20OER%20Declaration_01.pdf

UNESCO& COL. (2011). "Guidelines for Open Educational Resources (OER) in Higher Education". United Nations Educational, Scientific and Cultural Organisation; Commonwealth of Learning. Disponible en: http://unesdoc.unesco.org/images/0021/002136/213605E.pdf (consultado el 13 de octubre de 2013).

Vaughan, L.& Shaw, D. (2004). "Can Web Citations Be a Measure of Impact? An Investigation of Journals in the Life Sciences". *Proceedings of the American Society for Information Science and Technology,* 41(1): 516–526.

Villar-Onrubia, D.(2012). "Assessing Awareness on Open Education by Means of Online Research Tools (Working Paper)". Work presented at the UNESCO Open Educational Resources (OER)

World Congress, June 20, Paris. http://papers.ssrn.com/abstract=2085585 (consultado el 13 de octubre de 2013).

Villar-Onrubia, D.(2013). "Mapping Attention to OER Across Spain's HE Sector: a Web Impact Assessment". In *Proceedings of OER13: Creating a Virtuous Circle*. Nottingham, England. http://www.oer13.org (consultado el 13 de octubre de 2013).

Wilkinson, D., Sud, P. & Thelwall, M. (2014). "Substance Without Citation: Evaluating the Online Impact of Grey Literature". *Scientometrics*, 98(2): 797-806.