STI 2018 Leiden

23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at https://hdl.handle.net/1887/64521

ISBN: 978-90-9031204-0

23rd International Conference on Science and Technology Indicators (STI 2018)
"Science, Technology and Innovation indicators in transition"
12 - 14 September 2018 | Leiden, The Netherlands
#STI18LDN

Altmetrics of the Open Access Institutional Repositories: A Webometrics Approach

Isidro F. Aguillo*

*isidro.aguillo@csic.es
Cybermetrics Lab (IPP-CSIC), Albasanz, 26-28, Of. 3E14, Madrid 28037 (Spain)

Abstract
Self-archiving in Institutional Repositories (IRs) is playing a central role in the success of the Open Access initiatives. Deposited documents are more visible and probably they get more downloads and citations, but making them freely available in a local repository is not enough. Social tools, both public and academic targeting, networking or silo oriented, should be taken into account for reaching larger audiences and increase not only the scholarly but also the social impact. This communication explores the presence of IRs contents in 28 social tools (Academia, Bibsonomy, CiteUlike, CrossRef, Datadryad, Facebook, Figshare, Google+, GitHub, Instagram, LinkedIn, Pinterest, Reddit, RenRen, ResearchGate, Scribd, SlideShare, Tumblr, Twitter, Vimeo, VKontakte, Weibo, Wikipedia All Languages, Wikipedia English, Wikia, Wikimedia, YouTube and Zenodo) using a webometric approach. The link mentions of 2185 IRs in the cited tools were collected during July 2017 from Google selected data centers. The results showed that most of the IRs have no strong presence in the most specializes tools and even for the most popular services the figures are not high enough too. Lack of strategies and bad practices are suggested as possible explanations for the low altmetrics figures.

Keywords
Institutional repositories, Open access, Webometrics, Altmetrics, Social tools, Academic networks

Introduction

Since mid-nineties webometrics is slowly placing a role in the description and evaluation of the scholarly communication (Thelwall, Vaughan & Björneborn, 2005; Orduña-Malea & Aguillo, 2014),). The lack of reliable data sources for link analysis is still one of the main barriers for its fully acceptance by the metric community (Thelwall, 2010). But links are not the only web indicators that can be used for measuring science impact and several alternatives has been proposed like mentions of the names of institutions/authors or of paper/monographs titles (Cronin et al., 1998; Kretschmer & Aguillo, 2004, 2005). Problems with name variants, incomplete texts or non-ASCII characters are a formidable obstacle, so it become popular to use link mentions (Ortega, Orduña-Malea & Aguillo, 2014), i.e. the strings of the target URL of links not being necessarily an active link (for example the domain of a mail server present in the after-@ email addresses).

The emergence of social tools in recent years provides new opportunities for metric analysis of the scholarly impact. The most successfully proposal for profiting from that opportunity, the Jason Priem 'altmetrics' (Priem et al., 2010) that stands for alternative metrics, consists of a large and very heterogeneous set of measures with a very unfortunate umbrella term.

Social web-based metrics (=altmetrics) exploit a wide range of platforms, from almost bibliographic services like Mendeley, ResearchGate or Academia to general social networks like Facebook or Google +. The number of social tools are probably exceeding the thousand, but not all of them are equally popular and services like Twitter, Wikipedia or YouTube are plenty of academic users and contents. (Holmberg, 2015; Kousha, Thelwall & Abdoli, 2012; Haustein et al., 2014)

From a webometric point of view, the altmetrics are not only useful as article level metrics, but they also can be applied to research-related units like individuals (Mas-Bleda et al., 2014). In the same way the old webometrics issues regarding the inconsistency of results, understanding the meaning of the results or the vulnerability to manipulation are equally valid regarding altmetrics (Shema, Bar-Ilan & Thelwall, 2014; Kousha & Thelwall, 2015).

In spite of these limitations, altmetrics is being explored as a potential tool for measuring research impact beyond the scientific communities, the so-called societal impact. Several authors have found correlation between altmetrics and citation measurements (Bornman, 2014 Erdt et al. 2016; Eysenbach, 2011; Hammarfelt, 2014;; Sud & Thelwall, 2014; Mohammadi & Thelwall, 2014; Zahedi, Costas & Wouters, 2014), but the choosing of sources can influence that result. There are also composite indicators like the Altmetric Attention Score developed by the company Altmetric.com (https://www.altmetric.com/about-our-data/the-donut-and-score/) that are becoming popular providing article level metrics to repositories and journals.

The aim of this contribution is to use a webometric approach, using link mentions (URLs appearing in third party websites), to analyse the presence of the contents of Open Access Institutional Repositories (IRs) in a wide range of social networks and tools.

Methodology

The general search engines are not commonly used in bibliometric papers due to the limitations and shortcomings of these tools for quantitative analysis of the scholarly communication (Vaughan, 2004; Vaughan & Thelwall, 2004; Bar-Ilan, 2004; Thelwall, 2016; van den Bosch et al., 2016).). Since 2004 with the introduction of the Rankings web (Aguillo et al., 2006), our team has developed a series of strategies for reducing the impact of the sometimes-weird behaviour of the major search engines. Usually one of the key issues is the level of noise when the terms searched are very common or they have many variants (including different languages versions), so trying to use web domains instead of names of individuals or institutions was the preferred option. However, even in these cases short domains (as for example the domain of the University of Seville in Spain is "us.es") tended to over-estimate the results.

However, the institutional repositories usually use not only the institutional web domain, but they add their own subdomain, so using at least 3 "words" instead of 2 clearly reduces the mentioned noise.

The obvious choice for extracting data was Google, the largest and most popular search engine. However when checking the number of records answered by Google for the same request using computers at the same location or repeating the search after a few moments in the same computer, the figures can be very different. The reason is that requests sent to

Google can be answered by different data centres located around the globe for avoiding punctual saturation of the servers, so if the closest centre (in a sense that does not mean necessarily geographical proximity) is not available at that moment, then another one will fulfil the request. Unfortunately, the data centres are not updated at the same time and it could be that their databases can be (greatly) different during large periods, a fact that can be totally unnoticed to the users.

In order to face this problem, we identified the IPs of several of the Google data centres (Table 1) and made a simultaneous common request to all of them. When comparing the numeric results we realised that most commonly obtained count can be used for choosing a set of IPs with exactly the same results that means they share the same Google database. These IPs addresses are those we used later for extracting the IRs altmetrics. The Table 1 includes only the IPs that gave the same results during the experiment, the total number of addresses tested was far larger (150). Most of them were unreachable at that moment, but this is a volatile situation as in previous tests several of them were active, so we recommend to check in advance the availability and results of the candidate IPs.

Table 1. List of IPs of Google Data Centres

GOOGLE DATA CENTERS	
173.194.44.6	46.108.1.182
74.125.230.193	46.28.247.25
201.191.202.178	58.27.108.187
173.194.38.128	64.233.161.99
173.194.69.102	64.233.183.93
173.194.70.113	74.125.226.65
209.85.225.103	74.125.227.38
212.188.7.12	74.125.24.139

For extracting the altmetrics figures, we used a webometric approach based on the advanced operators of Google. The syntax includes two parts. First part is used for filtering for the webdomain of the social network through the operator “site:”, while the second one consists of the URL of the host of the repository between quotes that forces exactly that sequence of characters. The number obtained is referred as “url mention” or “link mention” So if we wish to obtain the link-mentions of the items deposited in the CSIC institutional repository (http://digital.csic.es/) in the Researchgate (http://www.researchgate.net) portal the syntax will be:

site:researchgate.net “digital.csic.es”

Other search engines can be used, but Google has a far larger coverage. However, even Google does not index the whole contents of the most important social tools, especially those highly dynamic and volatile like Facebook or Twitter.

The list of IRs was extracted from the Ranking Web of Repositories (http://repositories.webometrics.info), excluding portals of journals, disciplinary repositories and faculty, school or groups’ repositories when the main organization (mostly a university) has its own repository. The master list includes 2296 IRs from all over the world.

We applied the webometric method described for obtaining altmetrics indicators for those IRs according the following 28 sources: Academia, Bibsonomy, CiteUlike, CrossRef, Datadryad, Facebook, Figshare, Google+, GitHub, Instagram, LinkedIn, Pinterest, Reddit, RenRen, ResearchGate, Scribd, SlideShare, Tumblr, Twitter, Vimeo, VKontakte, Weibo, Wikipedia (all languages), Wikipedia English, Wikia, Wikimedia, YouTube and Zenodo. Unfortunately, Elsevier's Mendeley is not included, as Google is not indexing it. In this paper, we prepare a grouping of the social tools (see Table 3) only for descriptive purposes, but without any taxonomic intention.

The list is not exhaustive and not all the tools are similarly relevant for the researchers, even for the members of the metrics community (Haustein et al., 2014). But taken into account that most of the IRs are managed by librarians we assumed that overall visibility is their main aim and so we do not restrict the analysis only to the "mainstream" media.

The data was collected during the first two weeks of July 2017. Requests were made twice (two times the same day) for avoiding collection errors. Then the maximum value of these two attempts for each IRs/tool request is chosen.

The list of IRs was cleaned by excluding the lowest values of the duplicate entries (19) and the 92 repositories with zero counts for every one of the sources. The final list includes 2185 entries from 102 different countries. The top 30 countries represented according to the number of IRs is shown in Table 2.

Table 2. Countries with the largest number of Institutional Repositories (IRs) analysed in this study (July 2017)

COUNTRY	IRs	COUNTRY	IRs	COUNTRY	IRs
USA	371	Australia	50	Ecuador	24
Japan	240	Brazil	49	Russia	24
United Kingdom	131	Canada	46	Turkey	24
France	109	Colombia	45	Belarus	21
Germany	106	India	38	Peru	21
Indonesia	71	Sweden	37	South Africa	20
Spain	69	Portugal	35	Argentina	19
Ukraine	58	China	30	Croatia	19
Taiwan	54	Poland	30	Hungary	19
Italy	51	Malaysia	26	South Korea	15

The top 10 countries includes the 58% of the IRs analysed. Indonesia, Ukraine or Taiwan are also present in that group that probably are more related to local universities initiatives than a true national open access policy. Among those countries lacking in the list perhaps Mexico, Belgium, Netherlands and Switzerland are the most surprising absences, although regarding the last three, their relative small number of universities need to be taken into account.

Results

Descriptive summary statistics for the coverage of the IRs are provided for the 28 tools in the Table 3. The population consists of 2185 IRs from which their web address (domain or subdomain) converted into strings of characters are checked for being mentioned in the cited social tools according to Google. As previously described, the proposed link counting method provides not exact numbers, but the counts may be accurate relative to each other.

The last column is very relevant as it shows that for 19 (68%) of the social tools, more than 1000 IRs have zero values, including 7 (25%) with more 2000 IRs (of 2185) in that situation.

Considering that the number of items deposited in the global IRs is in the order of several millions and that even considering overlaps and duplicates, the number of link mentions for all of the social tools is low, even for the academic networks (RG and Academia with averages below 300 mentions). In that sense, Scribd (e-books deposit) could relatively the most successful of the analysed tools. Low results from CrossRef are probably due to the fact the sources are mainly journals themselves so the links used are surely DOIs.

Table 3. Descriptive statistics of the link mentions in the 28 social tools for the web domains/subdomains of all the Institutional repositories (n=2185; Google, July 2017)

GROUP	TOOLS	MEAN	STD ERROR	MAXIMUM	SUM	NON ZERO
ACADEMIC	RESEARCHGATE	282.9	25.4	40400	618073	1918
	ACADEMIA	249.2	58.5	125000	544570	1833
	CITEULIKE	0.8	0.3	540	1833	287
	CROSSREF	0.1	0.0	32	239	117
	BIBSONOMY	2.9	0.7	1070	6416	314
GENERAL	FACEBOOK	128.2	7.5	5610	280221	1706
	LINKEDIN	35.0	2.5	2090	76439	1306
	GOOGLE+	3.8	0.4	347	8298	1004
	RENREN	0.6	0.2	297	1396	59
	VK	6.6	1.0	1350	14359	621
(DATA)DEPOSITS	SCRIBD	146.8	11.9	11700	320728	1488
	SLIDESHARE	18.5	1.5	1360	40520	1358
	GITHUB	8.1	1.8	3560	17657	1447
	FIGSHARE	0.2	0.1	99	429	130
	ZENODO	1.1	0.2	299	2368	666
	DATADRYAD	0.0	0.0	2	19	17
WIKIPEDIA	WIKIPEDIA	15.5	2.7	4980	33866	1262
	WIKIPEDIA (ENGLISH)	3.2	0.4	625	6943	781
	WIKIMEDIA	3.7	1.2	2390	8088	382
	WIKIA	0.7	0.1	210	1434	359
(MICRO/MEDIA)BLOGS	TWITTER	33.5	2.4	1590	73089	1500
	WEIBO	0.1	0.0	45	278	104
	REDDIT	7.3	1.2	2200	15921	800
	TUMBLR	6.3	1.4	2470	13726	696
	YOUTUBE	7.0	1.7	3380	15201	891
	VIMEO	0.1	0.0	20	323	126
	PINTEREST	216.1	82.6	178000	472127	884
	INSTAGRAM	0.1	0.0	13	125	86

In spite their huge popularity, even among researchers, none of the main global platforms (Facebook, Linkedin or Google+) not the large local Chinese (RenRen) or Russian (Vkontakte) are being use to disseminate papers deposited in the IRs. However this method only identify links shared in public pages, so links exchanged in private groups are not considered.

In the case of Twitter there are evidence that the Google index only about a mere 5% of its contents, a figure that decreases year after year (Enge, 2018)

Regarding YouTube, the specific characteristics of the video media can explain their low usage.

However, perhaps the most surprising result refers to the Wikipedia. As it is not allowed to publish original research, it can be expected that all the scholarly items will include several academic references to recent papers, preferably open access full text versions, those usually deposited in IRs. The low numbers can be due to several reasons: (1) OA papers referred by its DOI or other pURL different to the IR address. (2) Many citations can refer to canonical sources not OA or not included yet in repositories. (3) Perhaps many documents are referred to other OA sources, global portals like Researchgate, Academia or regional portals like Scielo or Redalyc.

The poor indexing of seminal papers by top institutions in their own repository is a serious concern for the OA community. The situation is clearly illustrated by the low number of records of the Oxbridge universities repositories included in the Google Scholar database (http://repositories.webometrics.info/en/transparent).

Even although English is by far the main language in academic papers and Wikipedia edition in that language is also the largest one, it looks that independently of the language of the paper, the IRs use the entries in their local Wikipedia for adding the link mentions to their assets.

In order to check specific situations, we identified the most popular IRs for each one of the social tools (Table 4). For 12 (43%) of them, the SAO/NASA Astrophysics Data System is the main contributor, although its huge size (about 13 million documents) and the fact that it cannot be properly tagged as a true IR should advice against its inclusion in the analysis.

For the rest of the tools the institutions represented are very diverse. The very large network of institutions of the University of California heads Facebook, while CiteUlike is especially liked by a small Christian Indonesian University. Slideshare is very popular in Latin America and the local Russian and Chinese social tools are represented by “local” institutions (Belorussian and Hong Kong ones).

Table 4. Repositories with the maximum number of link mentions for each one of the social tools (Google, July 2017).

GROUP	TOOLS	INSTITUTION	WEB ADDRESS	MENTIONS
ACADEMIC	RESEARCHGATE	University of Michigan	deepblue.lib.umich.edu	40400
	ACADEMIA	Smithsonian/NASA	adsabs.harvard.edu	125000
	CITEULIKE	Petra Christian University	repository.petra.ac.id	540
	CROSSREF	Smithsonian/NASA	adsabs.harvard.edu	32
	BIBSONOMY	Smithsonian/NASA	adsabs.harvard.edu	1070
GENERAL	FACEBOOK	University of California	escholarship.org	5610
	LINKEDIN	The World Bank	openknowledge.worldbank.org	2090
	GOOGLE+	Smithsonian/NASA	adsabs.harvard.edu	347
	RENREN	Hong Kong University of Science and Technology	repository.ust.hk	297
	VK	Belarusian State University	elib.bsu.by	1350
(DATA)DEPOSITS	SCRIBD	University of Sumatera Utara	repository.usu.ac.id	11700
	SLIDESHARE	Universidad de los Andes (Venezuela)	saber.ula.ve	1360
	GITHUB	Smithsonian/NASA	adsabs.harvard.edu	3560
	FIGSHARE	Smithsonian/NASA	adsabs.harvard.edu	99
	ZENODO	Smithsonian/NASA	adsabs.harvard.edu	299
	DATADRYAD	Purdue University	docs.lib.purdue.edu	2
WIKIPEDIA	WIKIPEDIA	Smithsonian/NASA	adsabs.harvard.edu	4980
	WIKIPEDIA (ENGLISH)	Smithsonian/NASA	adsabs.harvard.edu	625
	WIKIMEDIA	Sverdlovsk Regional Universal Scientific Library	elib.uraic.ru	2390
	WIKIA	Smithsonian/NASA	adsabs.harvard.edu	210
(MICRO/MEDIA)BLOGS	TWITTER	National Chiao Tung University	ir.nctu.edu.tw	1590
	WEIBO	University of Michigan Deep Blue	deepblue.lib.umich.edu	45
	REDDIT	NASA Technical Reports Server	ntrs.nasa.gov	2200
	TUMBLR	Smithsonian/NASA	adsabs.harvard.edu	2470
	YOUTUBE	National Chiao Tung University	ir.nctu.edu.tw	3380
	VIMEO	Smithsonian/NASA	adsabs.harvard.edu	20
	PINTEREST	University of Florida	ufdc.ufl.edu	178000
	INSTAGRAM	Instituto Federal do Rio Grande do Norte	memoria.ifrn.edu.br	13

As the institutional patterns are not evident and it looks strongly dependent of local initiatives or projects, we decided to focus on geographical aggregations: Regions (Table 5) and (selected) countries (Table 6).

Table 5. Global number of link mentions by Region of the social tools (Google, July 2017)

GROUP	TOOLS	EUROPE	NORTHAMER	ASIA	OCEANIA	LATAM	AFRICA	ARABWORLD
ACADEMIC	RESEARCHGATE	315111	156286	41806	40260	52194	9590	739
	ACADEMIA	198646	237277	29907	39316	31126	5268	987
	CITEULIKE	551	619	576	58	13	6	8
	CROSSREF	102	75	45	9	6	2	0
	BIBSONOMY	4290	1682	26	368	43	6	1
GENERAL	FACEBOOK	106487	83316	21149	9207	41826	3097	8873
	LINKEDIN	34329	25947	948	5664	5522	979	602
	GOOGLE+	2877	2154	830	849	1322	61	141
	RENREN	33	525	332	504	0	1	1
	VK	9553	4184	168	180	37	15	5
(DATA)DEPOSITS	SCRIBD	85616	96998	49139	16453	66023	2584	192
	SLIDESHARE	16539	9062	3031	2353	8558	444	45
	GITHUB	7954	7843	591	531	566	88	33
	FIGSHARE	169	146	10	77	24	2	0
	ZENODO	1345	739	70	65	70	29	45
	DATADRYAD	8	9	1	1	0	0	0
WIKIPEDIA	WIKIPEDIA	14406	15808	1237	749	1217	320	48
	WIKIPEDIA (ENGLISH)	1763	4285	207	433	81	128	6
	WIKIMEDIA	6009	1734	45	61	222	15	0
	WIKIA	542	730	26	65	50	20	1
(MICRO/MEDIA)BLOGS	TWITTER	31761	19013	11188	4067	4438	320	523
	WEIBO	12	105	142	16	0	1	0
	REDDIT	3850	10813	230	729	53	60	20
	TUMBLR	2034	9346	681	1336	188	105	4
	YOUTUBE	3649	6102	3705	352	1271	31	16
	VIMEO	123	153	2	29	15	1	0
	PINTEREST	90769	347960	2727	15560	10650	1727	66
	INSTAGRAM	26	55	15	3	23	1	0

European and North American IRs are virtually tied for every tool, although it should be noted that RG is far more used in Europe while the North Americans prefer Academia. The rest of the regions, including Asia, are far from the figures of the first two. Even for a few tools, Latin America, home of the large journal portals Redalyc and Scielo that topped the development of local IRs, is performing better than the whole Asia. As according to Table 2 the Asian countries are well represented by the number of IRs perhaps the gap can be explained by the size (number of records) and/or the visibility policies of these IRs.

We choose six countries (USA, UK, Japan, Australia, Brazil and South Africa) from different regions for a comparative analysis. Academia mentions less items for every country than RG, except for the USA, although Australian figures are very similar. Brazilian IRs are very active in Facebook and Scribd (here with figures close to Australia).

Table 6. Global number of link mentions by Country (selected) of the social tools (Google, July 2017)

GROUP	TOOLS	USA	UK	JAPAN	AUSTRALIA	BRAZIL	SOUTH AFRICA
ACADEMIC	RESEARCHGATE	142111	67939	8939	36443	35340	7278
	ACADEMIA	227572	51155	3385	36172	14100	4418
	CITEULIKE	601	140	5	56	5	6
	CROSSREF	71	36	2	9	4	2
	BIBSONOMY	1598	1178	14	132	12	5
GENERAL	FACEBOOK	74793	13456	3036	7938	19004	1597
	LINKEDIN	22513	5623	94	5104	3646	863
	GOOGLE+	2044	399	108	807	340	37
	RENREN	522	5	23	504	0	0
	VK	3988	612	67	161	25	14
(DATA)DEPOSITS	SCRIBD	92741	18916	742	15201	15479	2069
	SLIDESHARE	8602	4329	158	2212	1789	366
	GITHUB	7364	1475	268	453	159	70
	FIGSHARE	134	77	1	66	6	1
	ZENODO	716	281	5	54	26	14
	DATADRYAD	8	4	0	1	0	0
WIKIPEDIA	WIKIPEDIA	15248	1148	616	641	447	296
	WIKIPEDIA (ENGLISH)	4062	519	57	370	25	109
	WIKIMEDIA	1655	44	17	54	17	15
	WIKIA	694	80	12	58	29	18
(MICRO/MEDIA)BLOGS	TWITTER	17247	9322	8688	2968	540	269
	WEIBO	105	0	8	16	0	1
	REDDIT	10133	1405	114	615	30	52
	TUMBLR	9118	564	465	1028	76	105
	YOUTUBE	5821	474	49	318	569	27
	VIMEO	137	41	1	26	11	1
	PINTEREST	338889	21907	1396	13800	2354	1660
	INSTAGRAM	54	0	0	2	16	1

Discussion and Conclusions

There are relevant limitations regarding the results obtained. The documents deposited should be referred in the social tools using the domain/subdomain of the repository. In fact, the opposite is truer, as many of them are recommending the use of handles (like for example http://hdl.handle.net/10261/148387). The handles are a type of pURL (permanent URL) that hide not only the name of the repository and its hosting institution, but even basic information about the paper like the names of author(s) or source, the publishing year or title keywords. This is relevant as the URL mention in the social tool can be the only piece of information the reader has to decide if clicking for reading or downloading the paper. Obviously, a prestigious university web domain can be a relevant hint in a tweet for both scientists and non-scholars users. In the case of Twitter, the use of shortening tools is very popular; so many mentions are lost when the URL does not include the host of the original web address. In fact, handles commonly are long strings of meaningless characters. A similar situation refers to the use of DOIs (another pURL) that it looks a good option with papers published in gold open access journals as both versions (published and deposited ones) are cited at the same time.

The Ranking Web of Repositories explicitly stated that its aims was not only to promote the green model of the open access initiatives (increasing the deposit in repositories), but to support good practices towards the authors depositing and their institutions. Therefore, the variables were designed for considering only those backlinks or social link mentions relevant to the repository that explicitly used the institution web domain. So since 2016 when altmetrics-based indicators were introduced, the use of pURLs in the IRs penalized their positions in the cited Ranking.

The results showed that most of the IRs managers (librarians in most of the cases) are not actively posting their contents to the social tools. It is possible that many items are really mentioned in the academic networks, but according to the data, they are not cited by the URL of the repository that offers information about the institutional authorship, a hint regarding the quality of the documents. However, there are other possibilities as far more authors than librarians are present on social media and they are surely interested in promoting their research, so when mentioning their papers probably they do not provide the digital location in their IR.

Excluding the most popular tools, local (or even individual) strategies and policies can explain the results for the most specialised tools. The Russian and Chinese services are virtually ignored outside their regional reach, although they have indeed very large audiences. GitHub, Figshare, Zenodo and Datadryad has scarce impact outside Europe and North America.

The recommendation of using pURLs for citing IRs items is probably sound, but using neutral or non-institutional web addresses decreases the informative value derived from the identification of the hosting institution. We suggest that this can penalize the usage of the involved OA papers as prestigious names can attract more visits. It can be also considered a bad practice regarding the institutional (moral) rights, as its authorship is explicitly excluded.

For scholarly communication purposes, researchers themselves are more and more active in both large general and academic tools, like RG, Facebook or Twitter, but from the results obtained it looks that IRs contents play a minor role. Regarding the most specialised tools the results suggests mostly local or individual initiatives.

References

Aguillo, I. F.; Granadino, B.; Ortega, J. L. & Prieto, J. A. (2006). Scientific Research Activity and Communication Measured With Cybermetrics Indicators. *Journal of the American Society for Information Science*, 57(10), 1296–1302.

Aguillo, I. F.; Ortega, J. L., Fernández, M. & Utrilla, A. M. (2010). Indicators for a webometric ranking of open access repositories. *Scientometrics*, 82(3), 477-486.

Bar-Ilan, J. (2004). The use of web search engines in information science research. *Annual Review of Information Science and Technology*, 38: 231-288. doi:10.1002/aris.1440380106

Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, 8 (4), 895-903.

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.

Enge, E. (2018). How Much Does Google Index Twitter in 2018?. Retrieved July 18, 2018 from https://www.stonetemple.com/twitter-indexing-study-2018

Erdt, M., Nagarajan, A., Sin, S.C.J. & Theng, Y.L. (2016). Altmetrics: an analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics*, 109(2), 1117-1166.

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with tradition al metrics of scientific impact. *Journal of Medical Internet Research*, 13(4), e123.

Hammarfelt, B. (2014). Using altmetrics for assessing research impact in the humanities. *Scientometrics*, 101(2), 1419-1430.

Haustein, S., Peters, I., Bar-Ilan; J., Priem, J., Shema, H. & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, 101 (2), 1145–1163.

Haustein, S., Peters, I., Sugimoto, C. R., Thelwall, M. & Larivière, V. (2014). Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *Journal of the Association for Information Science & Technology*, 65(4), 656-669.

Holmberg, K. J. (2015). *Altmetrics for information professionals: Past, present and future*. Chandos Publishing.

Kousha, K., Thelwall, M. & Abdoli, M. (2012). The role of online videos in research communication: A content analysis of YouTube videos cited in academic publications. *Journal of the Association for Information Science & Technology*, 63(9), 1710-1727.

Kretschmer, H. & Aguillo, I. F. (2004). Visibility of collaboration on the Web. *Scientometrics*, 61(3), 405-426.

Kretschmer, H. & Aguillo, I. F. (2005): New Indicators for Gender Studies in Web Networks. *Information Processing & Management*, 41 (6), 1481- 1494.

Mas-Bleda, A., Thelwall, M., Kousha, K. & Aguillo, I. F. (2014). Do highly cited researchers successfully use the social web?. *Scientometrics*, 101(1), 337-356.

Mohammadi, E. & Thelwall, M. (2014). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science & Technology*, 65(8), 1627-1638.

Orduña-Malea, E. & Aguillo, I. F. (2014). Cibermetría: midiendo el espacio red. Barcelona: Editorial UOC.

Ortega, J. L.; Orduña-Malea, E. & Aguillo, I. F. (2014). Are web mentions accurate substitutes for inlinks for Spanish universities?. *Online Information Review*, 38 (1), 59-77.

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). Altmetrics: A manifesto. Retrieved March 18, 2018 from http://altmetrics.org/manifesto/

Shema, H., Bar-Ilan, J. & Thelwall, M. (2014). Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics. *Journal of the Association for Information Science & Technology*, 65(5), 1018-1027.

Sud, P. & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics*, 98(2), 1131-1143.

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

Thelwall, M. (2010). Webometrics: emergent or doomed?. *Information Research*, 15(4), colis713. Retrieved March 10, 2018 from http://InformationR.net/ir/15-4/colis713.html.

Thelwall, M. (2016). *Web Indicators for Research Evaluation: A Practical Guide*. San Rafael, CA: Morgan & Claypool. DOI: 10.2200/S00733ED1V01Y201609ICR052)

van den Bosch, A., Bogers, T. & de Kunder, M. (2016). Estimating search engine index size variability: a 9-year longitudinal study. *Scientometrics*, 107(2), 839–856

Vaughan, L. (2004). New measurements for search engine evaluation proposed and tested. *Information Processing & Management*, 40(4), 677–691.

Vaughan, L. & Thelwall, M. (2004). Search engine coverage bias: Evidence and possible causes. *Information Processing & Management*, 40(4), 693–707.

Zahedi, Z., Costas, R. & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of ‘alternative metrics’ in scientific publications. *Scientometrics*, 101 (2), 1491-1513