Altmetrics of the
Open Access Institutional Repositories
A **Webometrics** Approach

Isidro F. Aguillo
isidro.aguillo@csic.es

SCIMAGO research group Cybermetrics Lab CSIC Consejo Superior de Investigaciones Científicas

Agenda

Objectives	Analyze the mentions of institutional repositories URLs in the main Social Tools and Networks
	Promote to increase the visibility of the contents deposited in Institutional Repositories
Introduction to the sources	Green Open Access: Ranking Web of Repositories
	Social Tools and Networks
	Google: Retrieval web data from its Data Centers
Comparative Analysis	Descriptive statistics: Good performers and gaps
	Results by Regions
Conclusions	Limitations
	Recommendations

Current situation

Open Science

Open Science	EU backed initiative for opening both the research cycle and the access to data, methods and results of the R&D activities
	Other objectives: - Citizen Science (crowdsourcing) - Open Funding (crowdfunding) - Open Metrics - Open educational resources
Open Access	Platinum Open Access (free) limited success, mainly in LATAM
	Gold OA: Authors paying high APCs guided by a metrics' (IF) oriented market. A few editorials controlling top journals and DOIs
	Green OA: Limited success in populating the Repositories, not a primary source for references to full text documents

Ranking Web of Repositories

http://repositories.webometrics.info/

Ranking Web of Repositories

Promote the deposit of full text documents in OA repositories and journals

2008-2017: Composite indicator

Three rankings: Global (Institutional + Subject repositories); Institutional (repositories) and Portals (Regional platforms and portals of journals)

2018: Number of items indexed by Google Scholar

Methods: Institutional Repositories

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

Population

2185 IRs from 102 different countries

Top 10 countries amounts for 58% of IRs

Data collected during July 2017

SCIMAGO research group   Cybermetrics Lab   CSIC Consejo Superior de Investigaciones Científicas

Methods: 28 Social Tools and Networks

GROUP	TOOLS	WEB DOMAIN	RELEASE DATE
ACADEMIC	RESEARCHGATE	researchgate.net	may-08
	ACADEMIA	academia.edu	sep-08
	CITEULIKE	citeulike.org	nov-04
	CROSSREF	crossref.org	2000
	BIBSONOMY	bibsonomy.org	2008
GENERAL	FACEBOOK	facebook.com	feb-04
	LINKEDIN	linkedin.com	may-03
	GOOGLE+	plus.google.com	jun-11
	RENREN	renren.com	dic-05
	VK	vk.com	oct-06
(DATA)DEPOSITS	SCRIBD	scribd.com	mar-07
	SLIDESHARE	slideshare.net	oct-06
	GITHUB	github.com	feb-08
	FIGSHARE	figshare.com	jan-11
	ZENODO	zenodo.org	2013
	DRYAD	datadryad.org	jan-08
WIKIPEDIA	WIKIPEDIA	wikipedia.org	jan-01
	WIKIPEDIA (ENGLISH)	en.wikipedia.org	
	WIKIMEDIA	wikimedia.org	2003
	WIKIA	wikia.com	oct-04
(MICRO/MEDIA)BLOGS	TWITTER	twitter.com	jul-06
	WEIBO	weibo.com	aug-09
	REDDIT	reddit.com	jun-05
	TUMBLR	tumblr.com	feb-07
	YOUTUBE	youtube.com	feb-05
	VIMEO	vimeo.com	nov-04
	PINTEREST	pinterest.com	mar-10
	INSTAGRAM	instagram.com	oct-10

Google: Syntax and Data Centers

Google   site:researchgate.net "digital.csic.es"   🔍

All   Images   News   Maps   Videos   More   Settings   Tools

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

About 21,100 results (0.44 seconds)

[PDF] **Las trayectorias profesionales y académicas de los profesores ... - Csic**
https://www.researchgate.net/.../Las-trayectorias-profesionales-y-aca... ▼ Translate this page
by L Cruz Castro - 2006 - Cited by 25 - Related articles
Junio 2006. Page 2. ÍNDICE. 1. INTRODUCCIÓN ....
.... 1. 1. 1. OBJETO DE ESTUDIO .

[PDF] **Internet invisible o Infranet - ResearchGate**
https://www.researchgate.net/.../La-informacion-especializada-en-Internet-Directorio-d... ▼
by Á Maldonado Martínez - 2001 - Related articles
Jan 1, 2001 - Capítulo 1. Herramientas de localización de recursos de información en Internet. Angeles
Maldonado Martínez. 1.1 Introducción .

[PDF] **La denominación y el contenido de los Repositorios ... - E-LIS reposit...**
https://www.researchgate.net/.../La-denominacion-y-el-contenido-d... ▼ Translate this page
by S Sánchez - 2006 - Cited by 24 - Related articles
Resumen. Se analizan los fenómenos que han dado lugar al debate sobre la correcta denominación y
los contenidos más apropiados para incorporar a los " ...

Methodology

A common request is made to all the Google Data centers using their IPs

The most common result is chosen as well as all the IPs giving it

The proposed syntax is used in the selected IPs

SCIMAGO research group   Cybermetrics Lab   CSIC Consejo Superior de Investigaciones Científicas

Results: Descriptive statistics

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

Comments (Coverage)

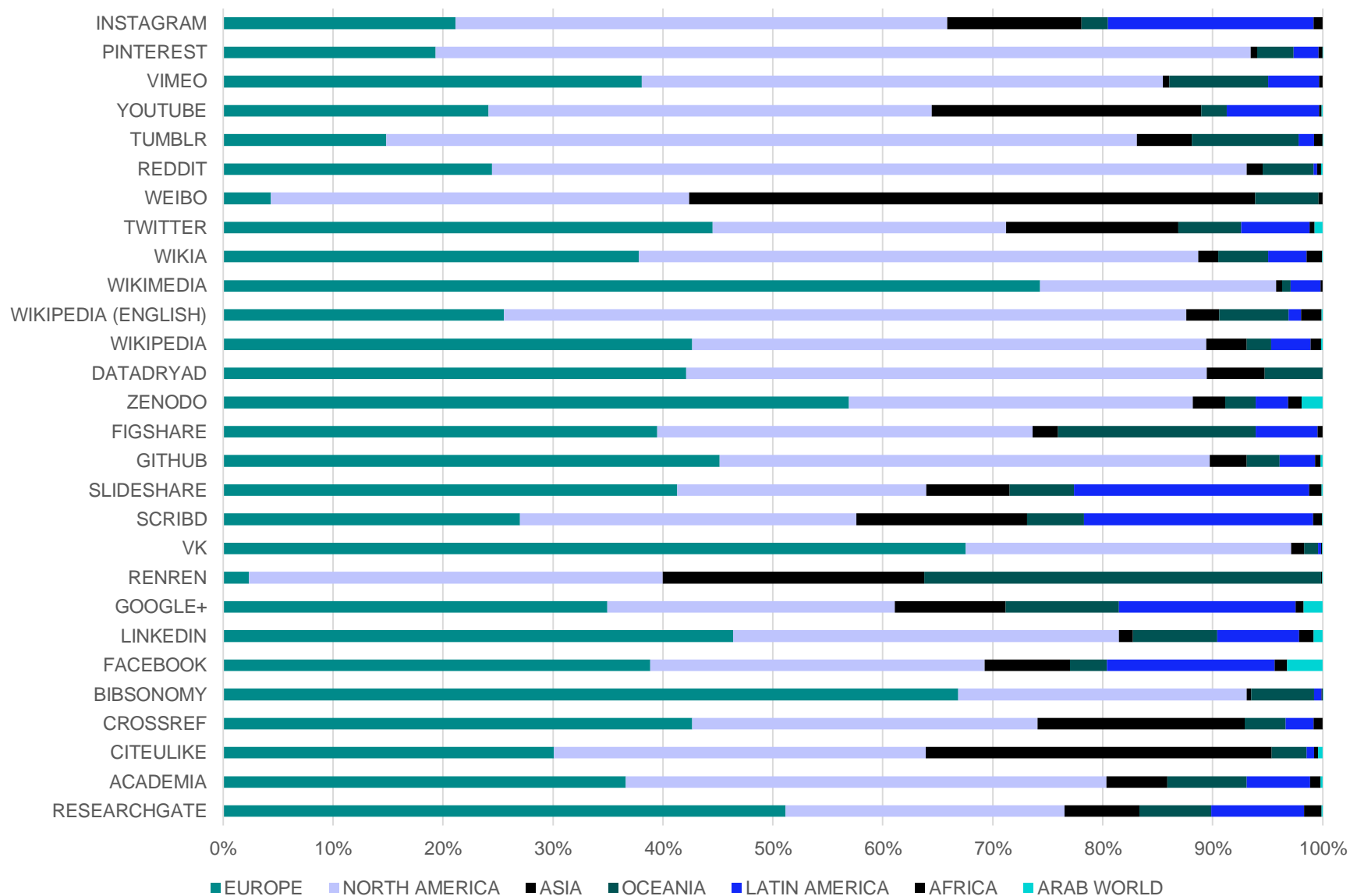|	
Usage gaps	For 19 **tools** (68% of 28) more than 1000 **IRs** (of 2185) has zero values
	For 7 **tools** (25%) there are 2000 **IRs** without any mention recorded
Social networks	The largest and the most popular academic networks (RG and Academia) showed an average of less than 300 mentions. CrossRef uses DOI
	General networks are even less popular, specially the huge local Russian (Vk) and Chinese (Ren Ren) ones
Deposits	ScriBD (e-books) is relatively the most used tool in this analysis
	Average numbers looks low for SlideShare (presentations) and GitHub (software), but this is probably due to their (current) low coverage by IRs
Wikipedia	Wikipedia is a lost opportunity. It should be mandatory to exploit the advantages of having open access to full text documents in this tool
	Other green OA sources can be cited: Regional portals in the local language versions (Scielo, Redalyc) or items in RG or Academia
Blogs	There are many problems related to Twitter: Only a fraction of contents are indexed by Google and it makes extensive use of URL shorteners
	YouTube and other media-supported blogging platforms are very specialized and perhaps can play a major role in the future

SCIMAGO research group Cybermetrics Lab CSIC

IRs MENTIONS BY SOCIAL TOOLS GROUPED BY COUNTRY

Legend: EUROPE, NORTH AMERICA, ASIA, OCEANIA, LATIN AMERICA, AFRICA, ARAB WORLD

Comments (Regions)

North America and Europe	There are few differences between the two developed regions regarding most of the main tools, although …
	RG is more used in Europe, while the North Americans prefer Academia
	But Facebook & LinkedIn are more popular in USA
Asia	The best represented country in the region is Japan (not China)
	Considering the large number of Asian IRs, their global contribution is relatively lower, perhaps due to their small size or visibility policies
Latin America	Successful LATAM portals (Scielo, Redalyc) collect local production, so high RG/Academia figures are reflecting papers in international journals
	ScriBD (e-books) probably includes a high number of textbooks
Others	Given its size Oceania IRs are well represented in academic tools
	Contributions of Africa and Arab World (North Africa and Middle East) to the global picture is limited

SCIMAGO research group · Cybermetrics Lab · CSIC Consejo Superior de Investigaciones Científicas

Limitations

Webometrics	Google (and other search engines) does not give exact results and figures can be very volatile changing wildly even in short periods
	Syntax can provide false positives especially for "short" URLs
Altmetrics (URL mentions)	Tools and authors are reluctant to mention URLs, specially if they are very long and complex
	Extensive (and recommended) usage of pURLS (handle, DOIs) hide the web domain of the IRs
Social Networks	Search engine robots do not index completely the social databases
	Large general networks in Russia (VK) or China (Ren Ren, Weibo) use local language, papers in English are not so commonly mentioned
Who make the mention?	Most of IRs have limited resources so probably most of the mentions are made by the author themselves
	Promotion of papers by authors in social networks probably uses journal bibliographic record (incl. DOI), but not the deposited version

SCIMAGO research group Cybermetrics Lab CSIC Consejo Superior de Investigaciones Científicas

Conclusions

Webometrics' role

IRs are not only deposits of documents, but tools to disseminate the institution's output. Usage (downloads) is a limited and biased indicator, but altmetrics can provide a better and larger picture of that impact

Webometrics (use of search engines as a proxy) can provide valuable information, but it is badly needed that URLs should preserve the identity of the source

IRs' role

Active promoting of new deposited items in social networks can be a very demanding task for IRs with limited resources, but it is a key strategy for their success

There are far more tools than those included in this analysis. Promotion of items should taking care of opportunities provided by large diversity of local and specialized tools

The authors have also important roles increasing the diversity of deposits (data, software, slides, video), giving added value to their social contributions and referencing (green) open access full text versions correctly (besides the DOI)

SCIMAGO research group Cybermetrics Lab CSIC Consejo Superior de Investigaciones Científicas

Recommendations

IR Managers should be more active mentioning their items in social networks by themselves or encouraging authors to include URLs of the deposited documents when promoting their work in those networks

pURLs have an important role in scholarly communication, but they are associated with the journal version, so it is recommended, that besides DOIs, the URL of the full text version (pdf suffixed, not handle) will be used too

Social tools full range should be considered. It will be unfeasible to use all of them, but diversity is relevant for achieving better impact:
- Immediacy of Twitter plus permanence in Wikipedia
- Academic (RG) plus general public (Linkedin) audiences
- Global (Facebook) plus local (Vk) coverage
- Paper (Academia) based plus extra additional (GitHub) contents
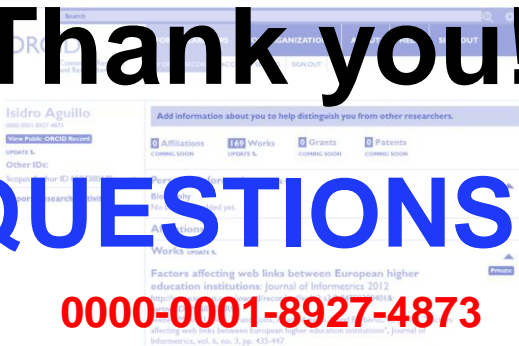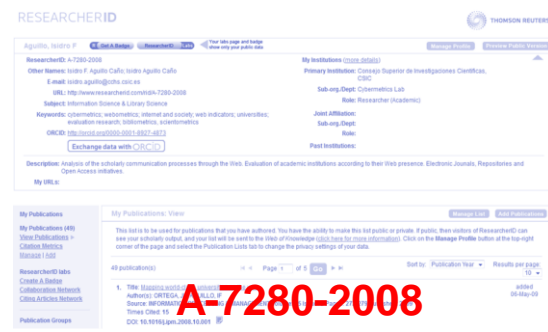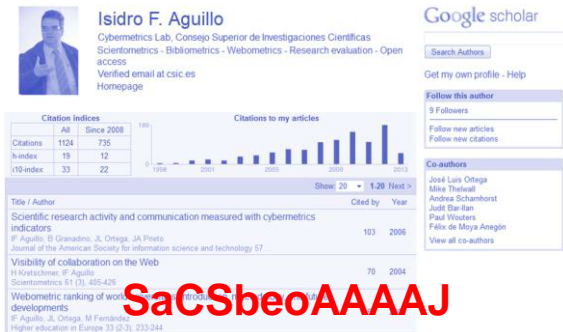
Thank you!

QUESTIONS?

SaCSbeoAAAAJ

124106

A-7280-2008

3481717

0000-0001-8927-4873

6507380175

@isidroaguillo

IsidroAguillo

isidro-f-aguillo

SCIMAGO research group · Cybermetrics Lab · CSIC Consejo Superior de Investigaciones Científicas