

Why Google Scholar?

An empirical comparison of the main evaluation tools

Isidro F. Aguillo
isidro.aguillo@csic.es

Research Evaluation

Macro level

**Policies
Monitoring**

Funding: %GDP devoted to R&D

Personnel

Output

Global output: Rank and Evolution

Statistical Office Reports

Micro level

**Career
Related**

Promotion (PhDs, Postdocs, Tenure)

Productivity (Example: Spanish 'Sexenios')

Research funding (projects, contracts)

Prestige

Plenary Speaker at events

Committees Member and/or Chair

Grants, awards, prizes

How it is performed till now?

Using Easy Bibliometrics based on JCR data

Strong reaction against (?)bibliometrics(?)

San Francisco DORA Declaration on Research Assessment

General Recommendation

1. Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions.

What happened? What are the reasons?

Monopoly

Only one bibliometric source (SCI, SSCI)
Paper format, later as CD-ROM for experts
VERY EXPENSIVE

Cumbersome to use, complex to analyze
No automatic computer supporter tools

Journal as Unit

By design, easier to choose the best (core) journals

Journal Citation Report
Dealing with data from thousands of journals instead of millions of
Papers and authors

(ISI) Citation Indexes “Original Sins”

Biased Core of Journals

Disciplinary (poor coverage of many disciplines)

Geographical (most of the journals published in Western countries)

Linguistic: Almost all of them in English

Closed small group

Large impact of Review and Multidisciplinary journals

Citation cartels of “popular” journals guiding future selections

Commercial interests

Other formats excluded

Books, chapters, thesis

Conferences

Reports

Patents with bibliography

Impact Factor “The Big Sins”

Unintended

Garfield introduced “only” for selecting journals

Never intended for evaluating individual papers of authors

Biased algorithm

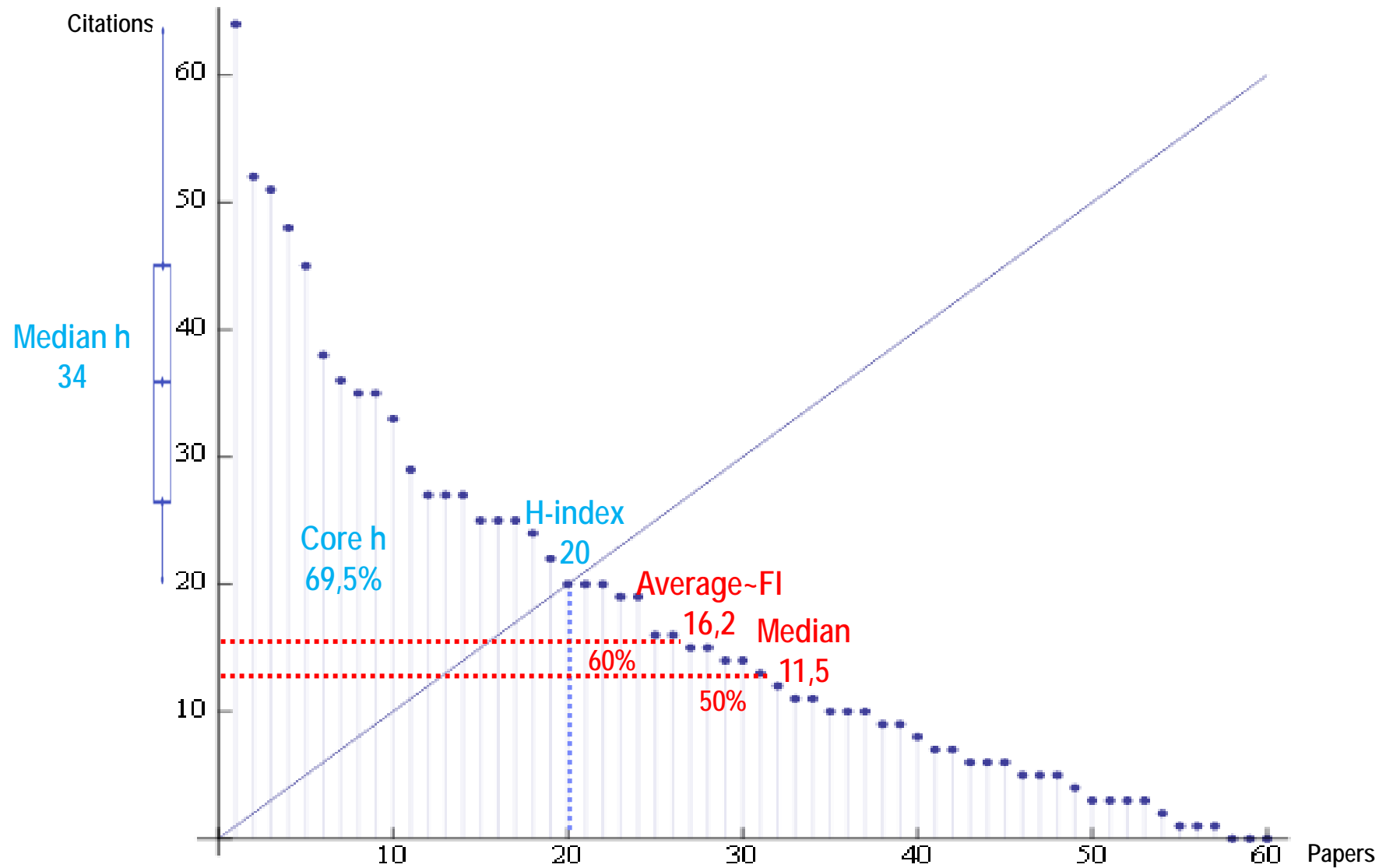
Citation distribution very skewed (“power law”)

Unable to compare different disciplines

Too short (2 years) citation window

False precision (3 decimals!!)

Central values are wrong for skewed distributions



Evaluation in Spain (and other countries)

A lot of output is excluded

Only WoS-indexed papers are considered
(sometimes –recently- Scopus and other minor sources are also accepted)

JCR rules!

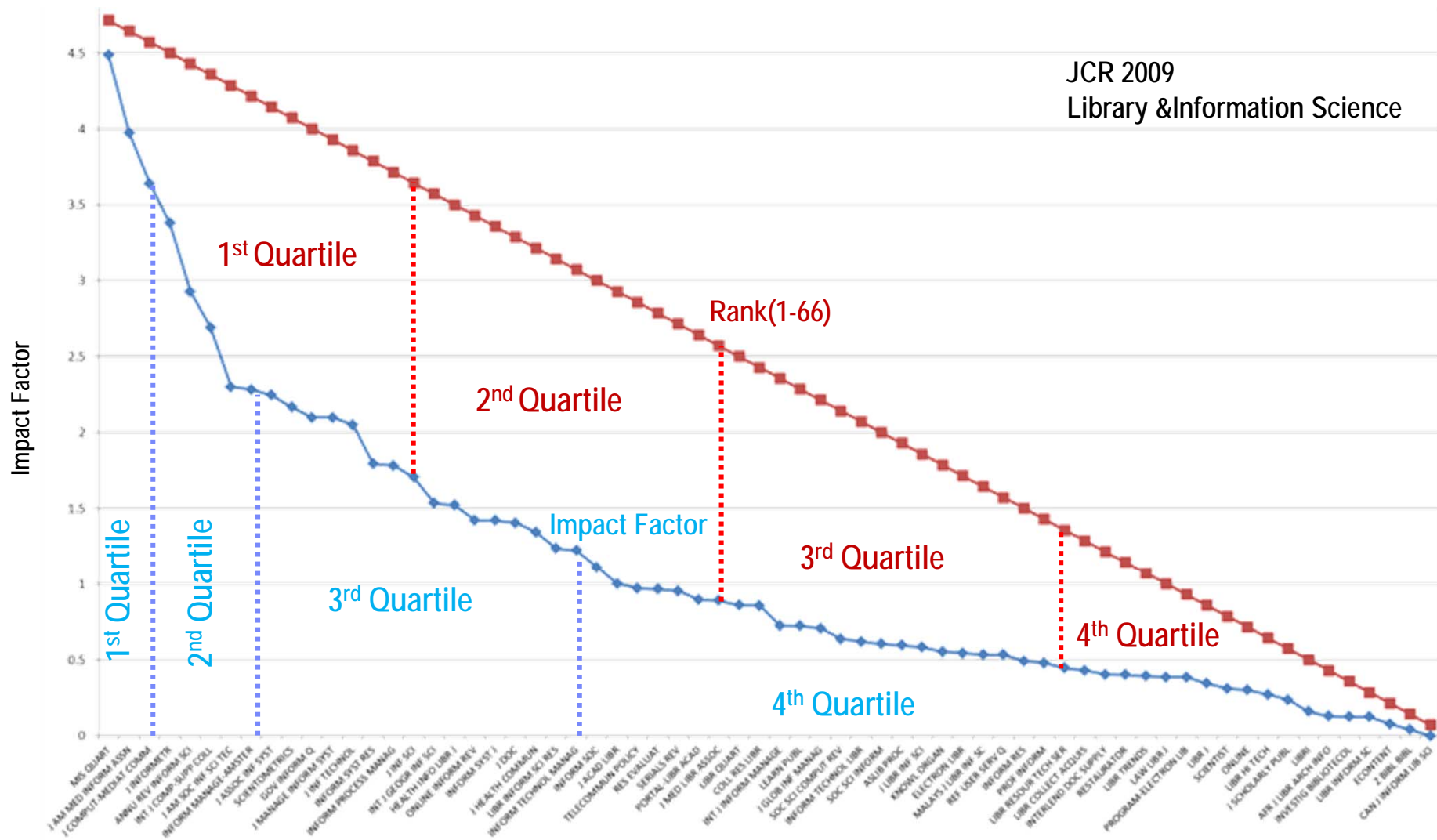
Raw outputs (quantity instead of quality approach)

Expected number of citations (Impact Factor) instead of actual citations

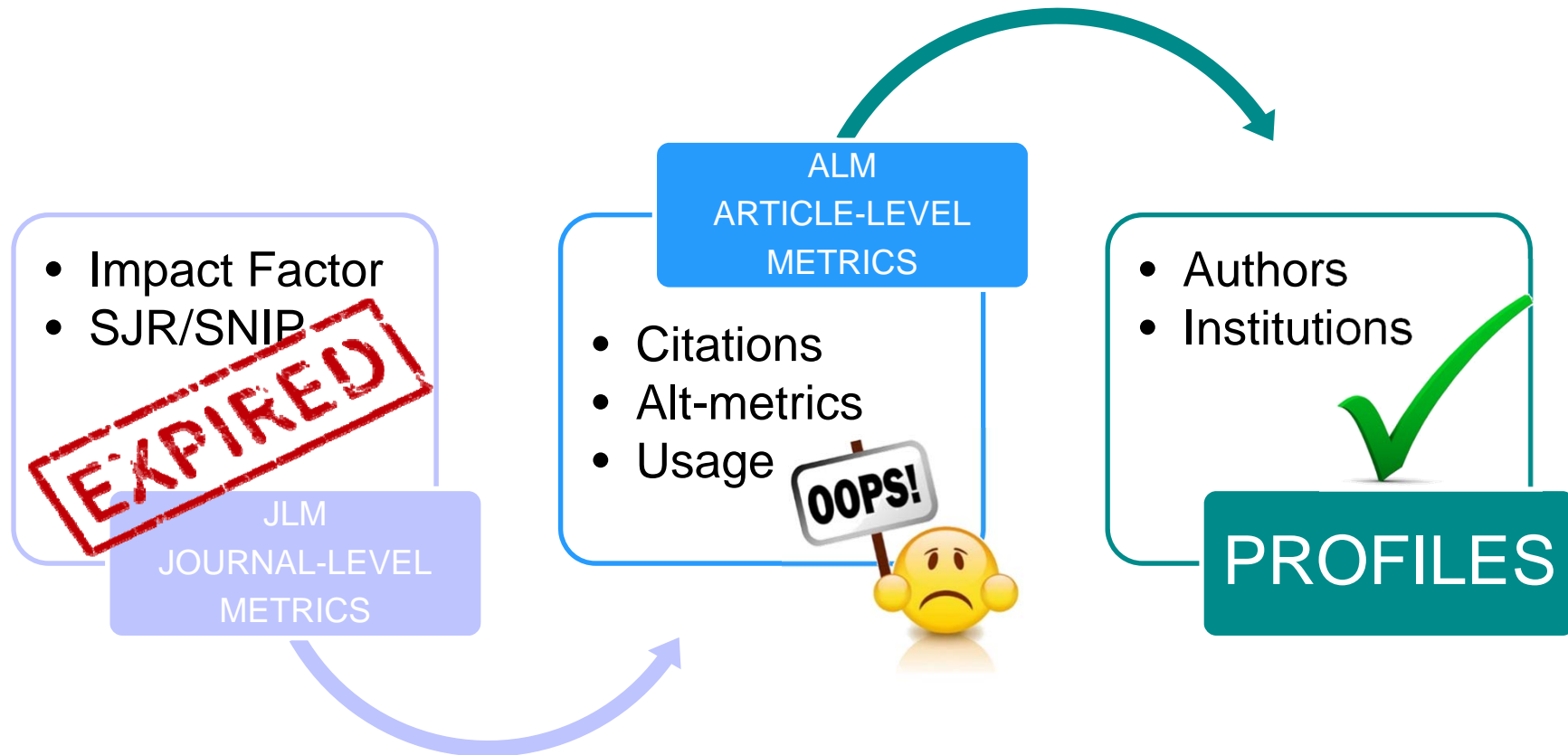
Quartiles: A pseudo solution for solving disciplinary differences

Non existent JCR Arts & Humanities was recommended in the law (!)

Wrongly-built Quartiles (WoS, Scopus)



In the meantime, the Paradigm changed!



Ugly, very ugly!

Lies
Big Lies
Bibliometrics?

Easy Bibliometrics

“says” that ...

but the truth is ...



- Only papers published in international journals need to be counted
- Social Sciences and Humanities are not so frequently cited
- Books have no research impact
- New web formats like blogs, forums, video are useless for evaluation

- Any document should be read and used and it can be cited and monitored
- Authors in these disciplines are heavily cited
- Books are hugely popular and strongly cited too
- New media is highly popular for very large audiences and can be measured

First: Introducing Google Scholar Citations



Anne-Wil Harzing

Follow

Professor of International Management - Middlesex University
 HQ-subsidary relations, International HRM, Language in IB, Quality & Impact of Academic Research, Bibliometrics
 Verified email at mdx.ac.uk - Homepage

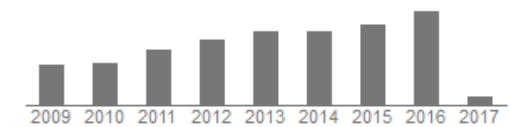
Title	1-20	Cited by	Year
Publish or Perish AW Harzing http://www.harzing.com/pop.htm	584 *	2007	
Acquisitions versus greenfield investments: International strategy and management of entry modes AW Harzing Strategic Management Journal 23 (3), 211-227	575	2002	
When knowledge wins: Transcending the sense and nonsense of academic rankings NJ Adler, AW Harzing The Academy of Management Learning and Education 8 (1), 72-95	557	2009	
Google Scholar as a new source for citation analysis? AW Harzing, R van der Wal Ethics in Science and Environmental Politics 8 (1), 61-73	524	2008	
An empirical analysis and extension of the Bartlett and Ghoshal typology of multinational companies AW Harzing Journal of International Business Studies 31 (1), 101-120	505	2000	
Managing the multinationals: An international study of control mechanisms AW Harzing Edward Elgar	50	1999	

Time required
 <5 min

Google Scholar

Get my own profile

Citation indices	All	Since 2012
Citations	11751	6884
h-index	51	43
i10-index	87	75



Co-authors View all...

- Markus Pudelko
- B. Sebastian Reiche
- Niels Noorderhaven
- Kraimer, Maria
- Nancy J Adler
- Michelle Brown
- Axèle Giroud
- Isabel Metz
- Ashly Pinnington
- Wilhelm Barner-Rasmussen
- Audra Mockaitis
- Paulo Prochno
- Michael J Morley

Second: An author is an Author, a citation is a Citation

Google Scholar



Emilio Delgado López-Cózar

Professor Research Methods, EC3 Research Group, Universidad de Granada
Scholarly communication, Research Evaluation, Bibliometrics, Scientometrics, Altmetrics

Verified email at ugr.es - Homepage

CITATIONS

REPORT

La edición de revistas científicas: directrices, criterios y modelos de evaluación

E Delgado-López-Cózar, R Ruiz-Pérez, E Jiménez-Contreras
FECYT

130 * 2006

BOOK

La investigación en biblioteconomía y documentación

E Delgado López-Cózar
Ediciones Trea

130 * 2002

REPORT

Manipulating Google Scholar Citations and Google Scholar Metrics: simple, easy and tempting

E Delgado López-Cózar, N Robinson-García, D Torres-Salinas
EC3 Working Papers 6

63 * 2012

PRE-PRINT

Acerca del concepto de documento

E Delgado-Lopez-Cozar
<http://eprints.rclis.org/13917/>

18 * 1992

Citation indices

All

Since 2012

Citations

3557

2397

h-index

32

28

i10-index


88

72



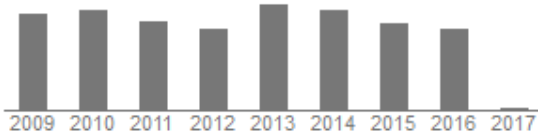
Third: Source is important, source shouldn't be important

Google Scholar



Anurag Acharya
Google Inc
scholarly communication, search engines
Verified email at google.com

Citation indices	All	Since 2012
Citations	4165	925
h-index	31	16
i10-index	53	26



Year	Citations
2009	~1000
2010	~1000
2011	~800
2012	~700
2013	~1000
2014	~1000
2015	~800
2016	~700
2017	~100

Title 1-20 Cited by Year

Information retrieval based on historical data
Anurag Acharya, Matt Cutts, Jeffrey Dean, Paul Haahr, Monika Saha
US Patent 7,346,839,

Active disks: Programming model, algorithms and data structures
A Acharya, M Uysal, J Saltz
ACM SIGPLAN Notices 33 (11), 81-91

Scopus

Author details

Acharya, Anurag
University of California, Santa Barbara, Department of Computer Science, Santa Barbara, United States
Author ID: 7102371436

Documents: 37
Citations: 506 total citations by 476 documents
h-index: 12 ?

Co-authors: 33
Subject area: Computer Science , Engineering [View More](#)

37 Documents | Cited by 476 documents | 33 co-authors

- Analyze author output
- View citation overview
- View h-graph

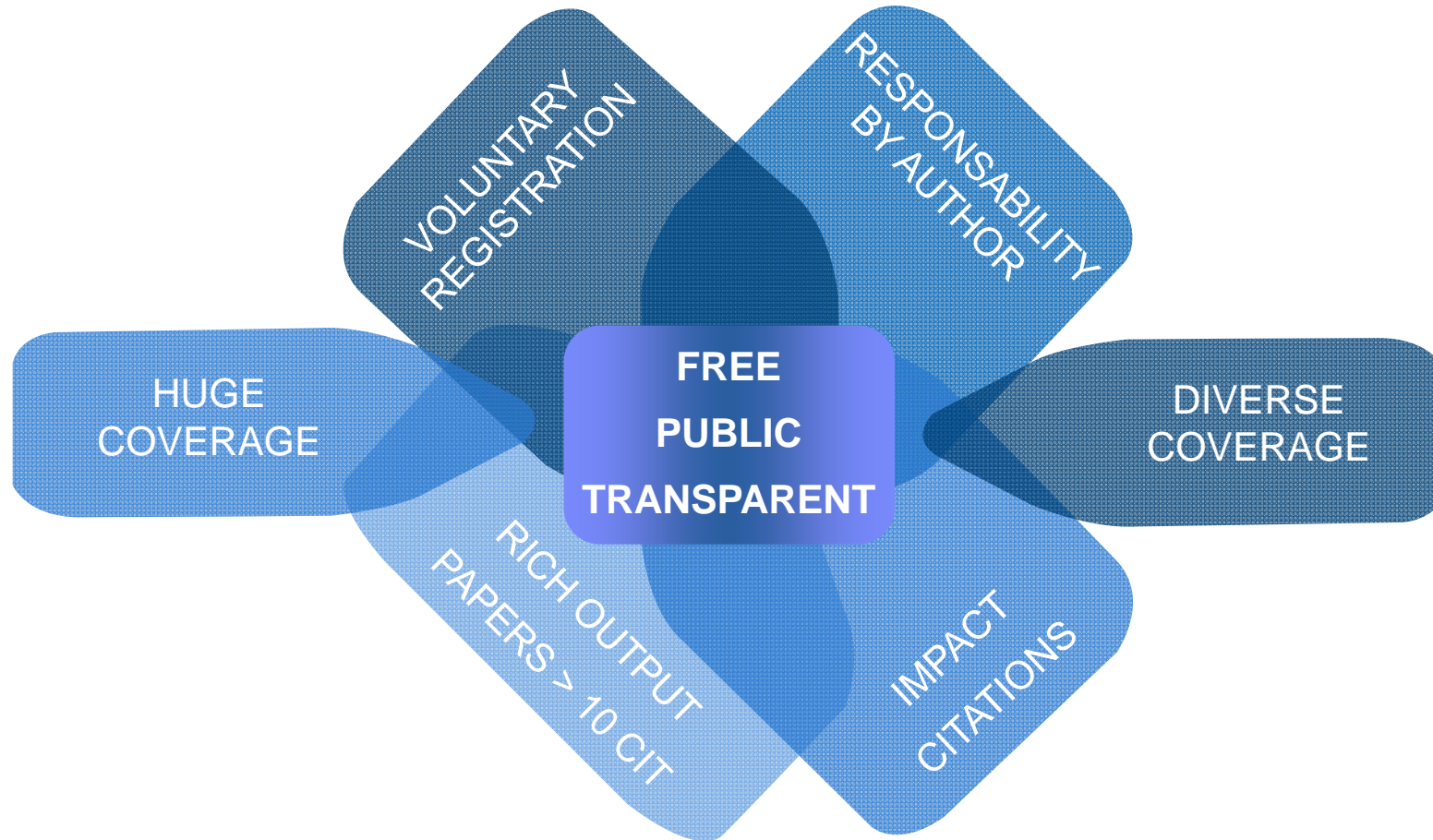
Google Scholar Citations Author Profiles

from



Why GSC?

New paradigm



Fourth: The Coverage



TLD & cTLD	GOOGLE	SCHOLAR	%
COM	25,270,000,000	88,500,000	0.4%
ORG	4,590,000,000	31,800,000	0.7%
CHINA PR	210,000,000	23,700,000	11.3%
EDU	640,000,000	6,910,000	1.1%
JAPAN	1,090,000,000	6,590,000	0.6%
NET	1,360,000,000	3,780,000	0.3%
GOV	1,230,000,000	3,260,000	0.3%
RUSSIA	735,000,000	3,170,000	0.4%
BRAZIL	585,000,000	2,600,000	0.4%
FRANCE	1,520,000,000	2,330,000	0.2%
KOREA	145,000,000	1,660,000	1.1%
UKRAINE	83,300,000	1,280,000	1.5%
SPAIN	513,000,000	1,220,000	0.2%
INDONESIA	70,800,000	1,220,000	1.7%
POLAND	396,000,000	1,030,000	0.3%
GERMANY	2,090,000,000	985,000	0.0%
AUSTRALIA	709,000,000	848,000	0.1%
UNITED KINGDOM	1,690,000,000	838,000	0.0%
ITALY	737,000,000	698,000	0.1%
CANADA	1,120,000,000	661,000	0.1%
TOTAL	57,203,318,749	193,824,176	0.3%

January 2017

Scopus

Journals

More than 66 million records in Scopus, which includes:

- Over 22,748 peer-reviewed journals, of which more than 4,470 are full open access (see the [Scopus Source List](#))
- Over 360 trade publications

Books

Over 558 book serials are covered in Scopus, accounting for 34,000 individual book volumes and 1.3 million items

More than 138,000 non-serial books (see the [book title list](#)), available in Scopus and 20,000 added each year.

WEB OF SCIENCE™

(from Web of Science Core Collection)

62,973,782 in the data limits

(from Russian Science Citation Index)

388,623 in the data limits

(from KCI-Korean Journal Database)

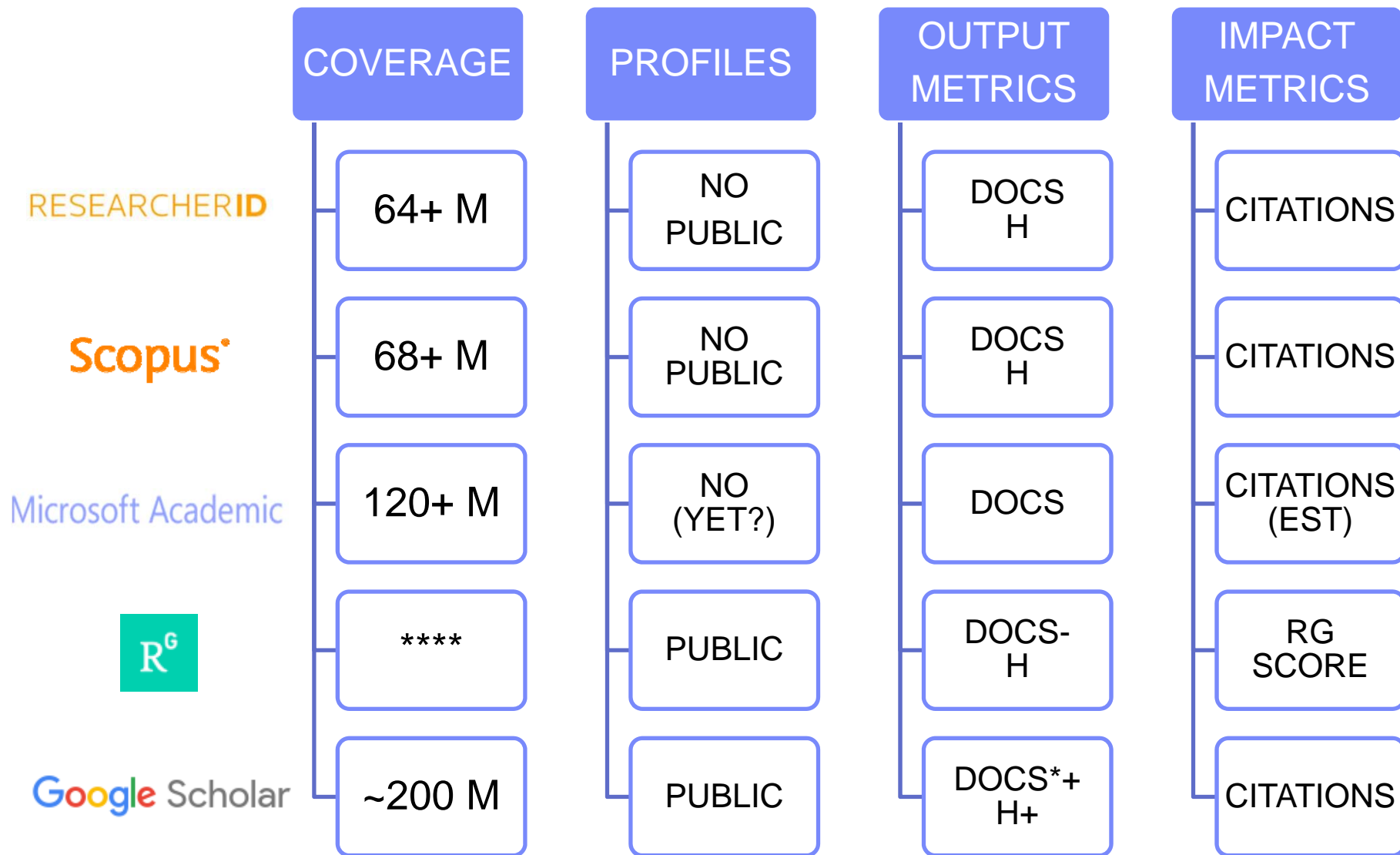
1,169,446 in the data limits

(from SciELO Citation Index)

530,654 in the data limits

Book Citation Index
60,000 books

Comparative analysis



Fifth: Supporting Open Access

RANKING WEB OF REPOSITORIES

COUNTRY	GOOGLE SCHOLAR		
	cTLD	REPOSITORIES	TOP INSTITUTIONAL ITEMS
CHINA PR	23,700,000	18,013,522	Xiamen University Institutional Repository 66,600
JAPAN	6,590,000	571,118	University of Tsukuba Repository 27,800
RUSSIA	3,170,000	1,185,136	Kazan Federal University Repository 16,900
BRAZIL	2,600,000	923,907	Universidade de São Paulo Teses e Dissertações 62,000
FRANCE	2,330,000	908,112	Université de Toulouse Archive Ouverte 3,620
KOREA	1,660,000	120,552	Seoul National University Repository S-Space 16,600
UKRAINE	1,280,000	270,666	Sumy State University Institutional Repository 22,900
SPAIN	1,220,000	* 1,224,239	Universitat Autònoma de Barcelona Dipòsit Digital 69,300
INDONESIA	1,220,000	537,152	Gadjah Mada State University Repository 52,300
POLAND	1,030,000	62,214	University of Lodz Repository 6,830
GERMANY	985,000	* 2,047,455	Universitat Heidelberg Dokumentenserver 23,400
AUSTRALIA	848,000	219,819	University of Queensland UQ eSpace 36,700
UNITED KINGDOM	838,000	* 2,826,668	UCL Discovery University College London 19,300
ITALY	698,000	134,444	Università di Pisa Electronic Theses 22,800
CANADA	661,000	370,079	University of British Columbia Open Collections 47,900
IRAN	658,000	369,360	Ardabil University of Medical Sciences Digital Repos 2,910
TAIWAN	569,000	314,915	National Chiao Tung University Institutional Reposit 52,700
NETHERLANDS	590,000	251,592	Universiteit van Amsterdam Digital Academic Repos 107,000
GRAN TOTAL	193,824,176	57,966,020	

January 2017

Sixth: YES, Books are (highly) cited too!

Book	Author	Date*	Discipline	Citations
The Structure of Scientific Revolutions	Thomas Kuhn	1962	Philosophy	81,311
Diffusion of Innovations	Everett Rogers	1962	Sociology	72,780
Pedagogy of the Oppressed	Paulo Freire	1968/1970	Education	72,359
Competitive Strategy	Michael E Porter	1980	Economics	65,406
Imagined Communities	Benedict Anderson	1983	Political Science	64,167
Mind in Society	LS Vygotsky	1978	Psychology	63,809
Discipline and Punish	Michel Foucault	1976/1977	Philosophy	60,700
A Theory of Justice	John Rawls	1971	Political Science	58,594
Social Foundations of Thought and Action	Albert Bandura	1986	Psychology	55,324
The Interpretation of Cultures	Clifford Geertz	1973	Anthropology	48,984
The History of Sexuality (3 Volumes)	Michel Foucault	1978-1986	Philosophy	47,955
Situated Learning: Legitimate Peripheral Participation	Jean Lave and Etienne Wenger	1991	Education	47,627
The Fifth Discipline	Peter M Senge	1992	Management	43,876
Institutions, Institutional Change and Economic Performance	Douglass North	1990	Economics	43,411
Culture's Consequences	Geert Hofstede	1980	Management	42,144
The Presentation of the Self in Everyday Life	Erving Goffman	1959	Sociology	40,573
Das Kapital	Karl Marx	1867-1894	Economics	40,237
Distinction: A Social Critique of the Judgement of Taste	Pierre Bourdieu	1984	Sociology	39,729
The Social Construction of Reality	Peter Berger and Thomas Luckmann	1966	Sociology	38,845
Metaphors We Live By	George Lakoff and Mark Johnson	1980	Linguistics	38,723
Stress, Appraisal and Coping	Richard Lazarus and Susan Folkman	1984	Psychology	38,665
Communities of Practice	Etienne Wenger	1999	Psychology	37,775
The Economic Institutions of Capitalism	Oliver Williamson	1985	Economics	37,651
Motivation and Personality	Abraham Maslow	1954	Psychology	37,614
Attachment	John Bowlby	1969	Psychology	37,318

What are the most-cited publications in the social sciences (according to Google Scholar)?

<http://www.monabaker.org/?p=4779>

Seventh: And the “other” disciplines too!



Karl Popper

London School of Economics, University of London
[Philosophy, Epistemology, Logic, Philosophy of science](#)
 Verified email at zoho.com - [Homepage](#)

Google Scholar

Citation indices	All	Since 2012
Citations	140305	36361
h-index	118	67
i10-index	553	247



Noam Chomsky

Professor of Linguistics (Emeritus), MIT
[Linguistic Theory, Syntax, Semantics, Philosophy of Language](#)
 No verified email - [Homepage](#)

Google Scholar

Citation indices	All	Since 2012
Citations	318239	83065
h-index	158	97
i10-index	865	414

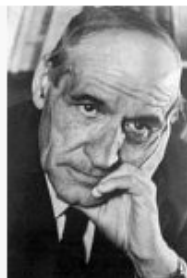


Hannah Arendt

Stanford, Princeton, Yale, University of Chicago, etc.
[political theory, philosophy, political philosophy](#)
 No verified email - [Homepage](#)

Google Scholar

Citation indices	All	Since 2012
Citations	143253	58113
h-index	118	79
i10-index	541	329



Jose Ortega y Gasset

Universidad Central
[Filosofía](#)
 No verified email

Google Scholar

Citation indices	All	Since 2012
Citations	20545	6190
h-index	54	34
i10-index	166	76

Eighth: Focus on Authors' own responsibility

Search results for 'colombia' on the R6 platform:

- National University of Colombia**: 15439 Members
- Colombia Aprende**
- Fundación Cardiovascular de Colombia**: 53 Members
- Fundación instituto de Inmunología de Colombia**: 15 Members
- Centro de Investigación de la Caña de Azúcar de Colombia**: 30 Members
- Universidad de Córdoba (Colombia)**: 261 Members

+70,000 profiles

RANKING WEB OF RESEARCHERS

COLOMBIA

RANK	NAME	INSTITUTION	H-INDEX	CITATIONS
1	Marta Losada Falk	Universidad Antonio Nariño	101	49714
2	Gabriela Navarro	Universidad Antonio Nariño	95	48647
3	Carlos Sandoval	Universidad Antonio Nariño	89	41460
4	Diego A Milanes	Universidad Nacional de Colombia	83	28095
5	Manuel Elkin Patarroyo Murillo	Fundación Instituto de Inmunología de Colombia	64	17234
6	Juan Manuel Anaya	Universidad del Rosario	57	10446
8	José Posada Villa	Universidad El Bosque	51	15627
9	Patricio López Jaramillo	Fundación Oftalmológica de Santander	50	9085
10	Roberto Martínez	Universidad Nacional de Colombia	43	9876
11	Felipe Guhl	Universidad de los Andes Colombia	42	4985
12	Francisco Lopera	Universidad de Antioquia	40	5756
13	Yovani Marrero Ponce	Universidad Tecnológica de Bolívar, Universidad Central de Las Villas	39	3540
14	Christophe Béné	International Center for Tropical Agriculture CIAT Colombia	38	4568
15	Aldo Pardo García	Universidad de Pamplona Colombia	35	9088
16	Omar D Cardona	Universidad Nacional de Colombia	34	5568
17	Esther de Vries	Pontificia Universidad Javeriana	34	4955
18	Olga L Sarmiento	Universidad de los Andes Colombia	32	8839
19	Juan Camilo Cárdenas	Universidad de Los Andes Colombia	32	5780
20	Iván D Vélez	Universidad de Antioquia	32	5617

~10,000 profiles

Advantages Google Scholar ...

offers free universal service

covers most of the scientific production

provides transparent indicators

promotes Open Access (repositories green OA)

requires user involvement

mandates author individual / institutional responsibility

Desiderata

FRIENDLY INTERFACE

API

Current data
Historical archive

EXTRA FIELDS

Country
Discipline

EXTRA INDICATORS

Academic Age
Co-author Index

EXTRA SOURCES

Usage-Metrics
Alt-metrics

BETTER PROCESSING

Parsing Improvement
Duplicates merging

From
promoting Quantity
to
measuring Quality

Ranking by countries

RANKING WEB OF RESEARCHERS

SPAIN

RANK	NAME	INSTITUTION	H INDEX	CITATIONS
1	Valentín Fuster	Centro Nacional de Investigaciones Cardiovasculares	166	164096
2	Santiago González de La Hoz	Instituto de Física Corpuscular UV CSIC	160	159028
3	Andrés Pacheco Pages	Institut de Física d' Altes Energies	151	125104
4	Alberto Ruíz Jimeno	Instituto de Física de Cantabria UNICAN CSIC	138	100754
5	Francisco Javier Rojas Ruiz ##	Universidad de Granada	127	71601
6	Ignacio Cirac	Max Planck Institute of Quantum Optics	123	68113
7	Fernando Barreiro	Universidad Autónoma de Madrid	121	81038
8	Antoni Torres	Hospital Clinic Universidad de Barcelona; CIBERes; Institut d'Investigacions Biomèdiques August Pi i Sunyer	119	57984
9	Jaime Bosch #	Hospital Clínic Universitat de Barcelona	115	43365
10	Francisco Herrera	Universidad de Granada	112	48259
11	Jesús San Miquel	Clinica Universidad de Navarra Universidad de Navarra	107	50616
12	Teun Avan Dijk	Universitat Pompeu Fabra	106	75670
13	Pablo García Abia	Centro de Investigaciones Energéticas Medioambientales y Tecnológicas	106	60595
14	Mariano Barbacid	Centro Nacional de Investigaciones Oncológicas	105	51676
15	Marcel Vos	Instituto de Física Corpuscular UV CSIC	104	62143

<http://www.webometrics.info/en/node/167>

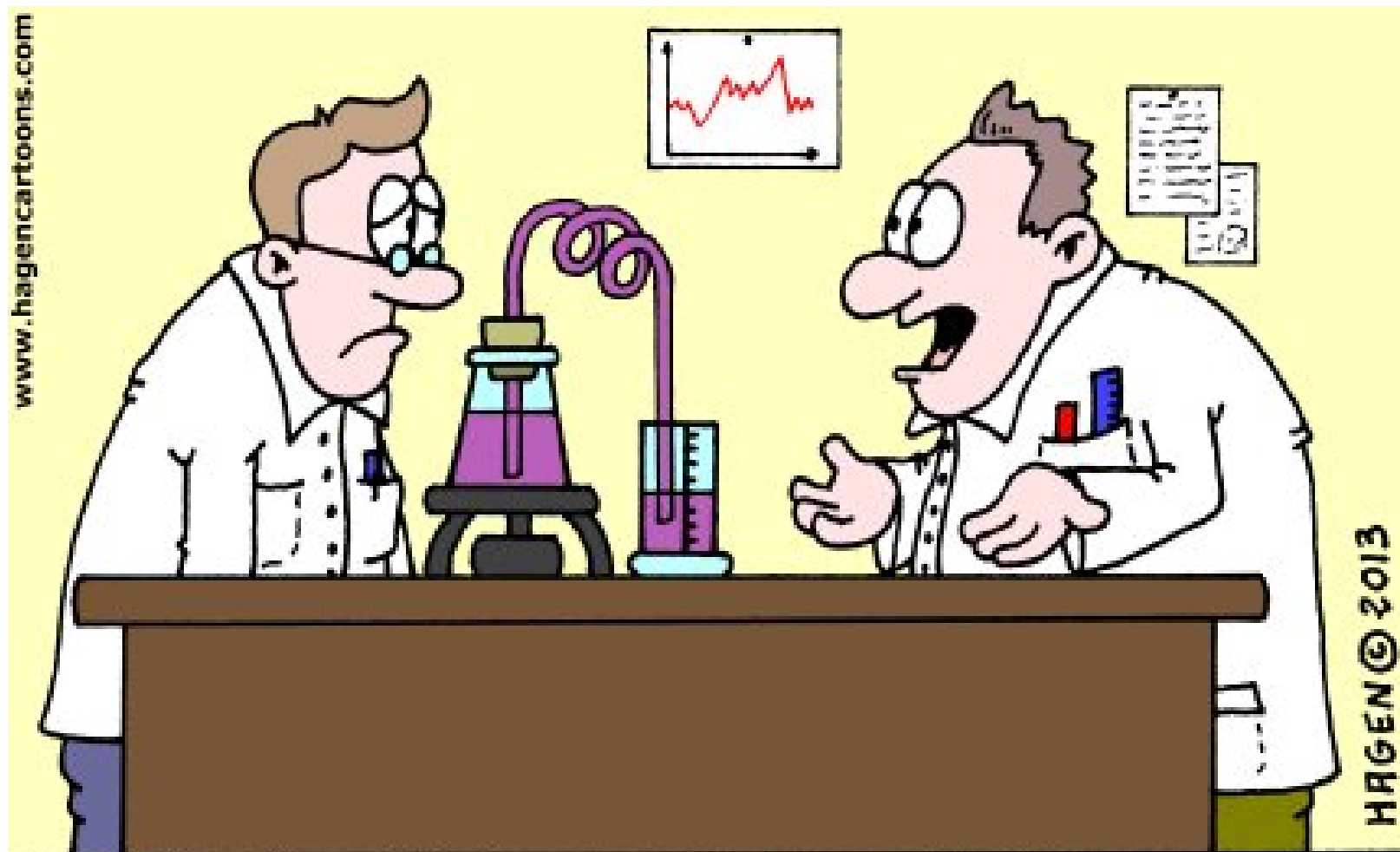
RANKING WEB OF RESEARCHERS

H>100

RANK	NAME	ORGANIZATION	H-INDEX	CITATIONS
1	Sigmund Freud	University of Vienna	269	488396
2	Graham Colditz	Washington University in St Louis	264	256415
3	Eugene Braunwald	Brigham and Women's Hospital; Harvard Medical School	246	290831
4	Ronald C Kessler	Harvard University	245	263006
5	Pierre Bourdieu	Centre de Sociologie Européenne; Collège de France	242	528228
7	Solomon H Snyder	Johns Hopkins University	240	216313
6	Michel Foucault	Collège de France	237	690001
8	Robert Langer	Massachusetts Institute of Technology MIT	232	216122
9	Bert Vogelstein	Johns Hopkins University	230	315600
10	Eric Lander	Broad Institute Harvard MIT	225	294683
11	Michael Karin	University of California San Diego	223	210430
12	Gordon Guyatt	McMaster University	217	187432
13	Michael Graetzel	Ecole Polytechnique Fédérale de Lausanne	216	235390
14	Salim Yusuf	McMaster University	214	248236
15	Richard A Flavell	Yale University; HHMI	214	171241

<http://www.webometrics.info/en/node/58>

Evaluation 2.0



**She didn't want to go out with you
even though you told her you had a h-index of 37!
Boy I don't understand women!**

Questions?



Thank you!