

Which citation database: Web of Science, Scopus, or Google Scholar?


Special Libraries Association 2009 Annual Conference

Lokman I. Meho

School of Library and Information Science
Indiana University, Bloomington
meho@indiana.edu

June 16, 2009

End of a monopoly?

- Until 2004, most large-scale citation analysis and bibliometric studies have relied on data exclusively from three ISI (currently Thomson Reuters) databases:
 - Arts & Humanities Citation Index
 - Science Citation Index
 - Social Sciences Citation Index

Web of Science
- The emergence of **Scopus** and **Google Scholar** in late 2004 has raised many questions regarding:
 - The validity of findings based exclusively on data from Web of Science
 - The value and necessity of using multiple databases (or data sources) for assessing the scholarly impact of research
 - The appropriateness of using Scopus and/or Google Scholar as an alternative source of citation and bibliometric data to Web of Science

Why asking Web of Science, Scopus, or Google Scholar?

	Web of Science	Scopus	Google Scholar
Breadth of coverage	10,000 peer-reviewed journals, including over 600 open-access titles and over 110,000 conf. proceedings 38m ref. enhanced records back to 1900 (several millions are ed. material and reviews to different types of material)	18,000 peer-reviewed journals, including over 1,200 open-access titles and approximately 4 m conf. papers 38m records back to 1823, of which 19m are ref. enhanced back to 1996 (great majority are j. art. and conf. papers)	No information provided. Theoretically, all electronic resources available on the Web
Subjects covered	All	All	All
Languages	A&HCI (47) ; 71% English SSCI (42) ; 92% English SCI (48) ; 92% English	55 languages; 82% English	Unlimited

Studies comparing WoS, Scopus, and GS

- Bibliographic coverage
- Impact on citation counting and ranking
- Impact on *h*-index scores and rankings
- Scholarly impact of authors and research domains

Bibliographic coverage

Study	Method / Topic	Findings
Norris & Oppenheim (2007)	33,000 j. art. submitted for the UK 2001 RAE in the social sciences; 2,800 journals indexed in <i>IBSS</i>	Scopus provides the best coverage of social science literature. Scopus could be used as an alternative to WoS as a tool to evaluate research impact in the social sciences
Gavel & Iselid (2008)	Compared coverage of all active journal titles in Scopus and WoS	84% of the WoS titles are indexed in Scopus whereas only 54% of the Scopus titles are indexed in WoS. Study was not based on manual searching of the databases to determine breadth and depth of coverage of each title
Gorraiz & Schloegel (2008)	Compared the coverage of top pharmacology and pharmacy journals in Scopus and WoS	All of the 100 highest impact WoS-covered journals in the field are indexed in Scopus; Scopus covers additional high impact journals not indexed by WoS
López-Illescas, de Moya-Anegón, & Moed (2008)	Compared the coverage of oncology journals in Scopus and WoSc	All 126 WoS-covered oncology journals are indexed in Scopus, with additional 106 journals covered in Scopus (which have lower impact factors than those commonly covered by both dbs.

Citation counting and ranking

Study	Method / Topic	Findings
Bar-Ilan, Levene, & Lin (2007)	Compared ranking of pub. of 22 highly-cited Israeli researchers as measured by citation counts in GS, Scopus, and WoS	High similarity between Scopus and WoS and lower similarities between GS and Scopus and WoS
Meho & Yang (2007)	Compared citation counts and rankings of 25 library and information science faculty in Scopus and WoS	The addition of Scopus citations to those of WoS significantly altered the relative rankings of those faculty appearing in the middle
Meho & Rogers (2008)	Examined differences between Scopus and Web of Science in the citation counting and ranking of 22 top human-computer interaction researchers	Scopus provides sig. more coverage of HCI literature than WoS, due to broader coverage of conf. proc.; Scopus can be used as a sole data source for citation-based research and evaluation in HCI
Levine-Clark & Gil (2009)	Examined citation coverage in WoS, Scopus, and GS for 15 business and economics journals	All three dbs should be used to get a more complete picture of the scholarly impact of an article in these journals
Torres-Salinas, et al. (in press)	Compared citation count of 2,299 papers pub. between 1999 and 2005 by all 864 health sciences researchers of the University of Navarra (Spain)	Papers received 15% more citations in Scopus than in WoS, with diff. being greater in clinical depts. Both dbs generate similar rankings of researchers

h-index

Study	Method / Topic	Findings
Meho & Rogers (2008)	Examined diff. between Scopus and WoS in the <i>h</i> -index of 22 top HCI researchers	Scopus generates significantly higher <i>h</i> -index scores than WoS for HCI researchers, primarily bcs of better coverage of ACM and IEEE literature.
Bar-Ilan (2008)	Compared the <i>h</i> -index scores of 40 highly cited Israeli researchers based on citation counts found in WoS, Scopus, and GS	Except for three researchers, the differences in the <i>h</i> -index between Scopus and WoS was insignificant. In many cases, the results obtained through GS were considerably different than those based on Scopus and WoS.

h-index, cont'd (Jorge E. Hirsch)

	<i>h</i> -index	Total records	Total citations	Av. no. of citations per paper
WoS	52	192	6,417	33
ADS	47	245	8,661	42
PROLA	37	143	5,462	38
Scopus	35	194	3,287	17
GS	33	387	5,040	13

ADS=Astrophysics Data System; PROLA=Physical Review Online Archive.
Data for GS was not cleaned for duplicate entries of the same paper. Total citation counts for GS and PROLA are sum of citation count of all papers.

Scholarly impact of authors and research domains

Study	Method / Topic	Findings
Meho & Rogers (2008)	Examined differences between Scopus and Web of Science in mapping the scholarly impact of top HCI researchers	Scopus generates significantly different maps of scholarly impact of individual scholars from those generated by WoS
Meho & Sugimoto (in press)	Citations to the work of 80 randomly selected full-time information studies faculty members from North America	<p>When analysis is based on smaller citing entities (e.g., journals, conference proceedings, and institutions), the two databases produce considerably different results</p> <p>When analysis is based on larger citing entities (e.g., research domains and countries), the two databases produce very similar pictures of scholarly impact</p>

Some strengths and weaknesses

	Strengths	Weaknesses
Google Books	<ul style="list-style-type: none">- Available for free- Covers 5,000,000 books- Easy to search- Linked to WorldCat	<ul style="list-style-type: none">- Inaccurate search results (ISBN)- Very time-consuming for cit. search- Displays only a max. of 500 records- Lacks data downloading features- Lacks advanced cit. searching options- Lacks citation analysis features
Google Scholar	<ul style="list-style-type: none">- Available for free- Covers all types of documents- Covers hundreds of m of items- Easy to search	<ul style="list-style-type: none">- No quality control- Very time-consuming for cit. search- Covers non-scholarly items- Does not cover all journals- Coverage uneven across disciplines- Doesn't perform well for older pubs.- Displays a max. of 1,000 records- Allows downloading one item at a time- Lacks advanced searching options- Lacks citation analysis features

Some strengths and weaknesses, cont'd

	Strengths	Weaknesses
Scopus	<ul style="list-style-type: none">- Advanced citation searching features- Advanced citation analysis features- Covers 18,000 Js. (incl. 3,500 A&H)- Covers many conference proceedings- Easy to search- Advanced downloading options	<ul style="list-style-type: none">- Does not cover books, chapters, diss., etc.- Citation searching from 1996 on only- Covers many low-impact journals
Web of Science	<ul style="list-style-type: none">- Advanced citation searching features- Advanced citation analysis features- Covers mainly high impact Js.- Citation searching from 1900 on- Advanced downloading options- Long history	<ul style="list-style-type: none">- Does not cover books, chapters, diss., etc.- General Search limited to ISI Js.- Limited coverage of non-English- CR searches can be time-consuming- CR searches can be quite complicated- Allows downloading 500 records at a time

Sample of other databases that can be used for citation searching

BIOLOGICAL SCIENCES: BioOne

BUSINESS: Business Source Premier

CHEMISTRY: Chemical Abstracts/SciFinder Scholar

COMMUNICATION: Communication & Mass Media Complete

COMPUTER SCIENCE: ACM Digital Library, CiteSeer, IEEE Xplore

ECONOMICS: EconLit

ENGINEERING AND TECHNOLOGY: IEEE Xplore

ENVIRONMENTAL STUDIES: GreenFILE

GENERAL/MULTIDISCIPLINARY: [Amazon.com](#), EBSCO's Academic Search, Books@Ovid, Cambridge University Press Journals, Emerald, [Google Books](#), [Google Scholar](#), Journals@Ovid, Oxford Journals, Sage Full-Text Collections, ScienceDirect, SpringerLink

GEOLOGICAL SCIENCES: GeoBase

HISTORY: America: History & Life, Historical Abstracts

LAW: HeinOnline

MATHEMATICS: MathSciNet

MEDICINE: [PubMed/PubMed Central](#)

MUSIC: Oxford Music Online

NURSING: CINAHL Plus with Full Text

PATENTS: [Citation Bridge \(US Patents\)](#), [US Patent and Trademark Office](#)

PHYSICS: [Institute of Physics](#), [ADS \(Astrophysics Data System\)](#), [PROLA \(Physical Review Online Archive\)](#)

POLITICAL SCIENCE: Worldwide Political Science Abstracts

PSYCHOLOGY: PsycARTICLES, PsycBOOKS, PsycCRITIQUES, PsycINFO

SOCIOLOGY: Sociological Abstracts

Concluding remarks

- Users of citation-based metrics for analyzing intellectual trade among research domains and examining the evolution of journals, growth of scientific fields, and the impact of research must first determine the appropriate citation database(s) to use in such exercises, and explain why they have made their choice, before carrying out their research and disseminating their findings and conclusions
- Without first verifying whether the use of Scopus, Web of Science, or both is necessary in assessing and/or mapping research impact, it will be hard to claim the validity and accuracy of findings of these studies
- The need to use one, two, or more citation databases in research assessment varies from one research domain to the other
- Despite the emergence of several large-scale, free citation data sources, Scopus and/or WoS remain irreplaceable tools for citation and bibliometric analyses of research