The Decade of Metrics? Examining the Evolution of Metrics Within and Outside LIS

by Vincent Larivière

Metrics & ASIS&T

EDITOR'S SUMMARY

Since its start in the mid-19th century, the measurement of document attributes and interrelationships has evolved into the formal field of bibliometrics, becoming solidly established in information science and reaching into other domains. Bibliometric methods have been used to shed light on organizational structure of library and information science (LIS), the reach and influence of LIS research and on interdisciplinarity and the emergence of new fields of study. Data from citation indexes from 1900 through 2011 show the rise of metrics, with the decade starting in 2010 likely to be the most productive and influential. The most cited researchers in information science in 2010 and 2011 focus on metrics, and the use of the term *bibliometri** has surged since 2000. The study of metrics in medicine is increasing rapidly, though less so in the social sciences, humanities and natural sciences. Among LIS subtopics, metrics appears to have the strongest influence outside the field.

KEYWORDS

bibliometrics
information science history
research methods
cross disciplinary fertilization
interdisciplinarity
emerging disciplines
trends
trends

Vincent Larivière is an assistant professor at the École de Bibliothéconomie et des Sciences de l'Information, Université de Montréal, Montréal, QC, Canada, and a research associate with the Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal, Montréal, QC, Canada. He can be reached at vincent.lariviere<at>umontreal.ca.

Bibliometric methods are at the heart of library and information science (LIS). It is one of the few – if not the only – method that arose from LIS scholars and that uses one of their main objects – documents and their characteristics – as its unit of analysis. First created by librarians in the mid-19th century to manage collections and used by statisticians such as Lotka in the 1920s, bibliometric methods were democratized in the mid-20th century with the founding by Eugene Garfield of the Institute for Scientific Information (ISI) and the creation of its various citation indexes [1]. Bibliometrics can be defined as the quantitative analysis of the characteristics of documents (articles, conference proceedings and so forth) published by researchers. Although it can theoretically be applied to the measurement of any type of literature – novels, newspapers and scientific journals – it is generally used for the measurement of science and technology and thus applied to scientific documents [2]. As a consequence, terms such as *scientometrics* or *informetrics* are often used as synonyms. One of the basic premises of bibliometrics is that new knowledge is incorporated in the scientific literature and that we can understand this process by measuring the characteristics of this literature and measuring certain attributes of knowledge production such as its main producers (authors, institutions, countries), research topics (words, journals) and diffusion and integration patterns (citations).

In a previous paper [3], Larivière, Sugimoto and Cronin used bibliometric methods to study the evolution since 1900 of the organizational structure of LIS, its means for diffusing research, patterns of interdisciplinarity and changing research topics. They showed that, despite a growth in the number of papers published, LIS's market share of all social science and humanities research decreased. They also analyzed interdisciplinary patterns of LIS and provided evidence that LIS scholars now cite and receive citations from

other fields more than from LIS itself. Along the lines of Cronin and Meho [4], they also show that the "intellectual balance of trade" of LIS with other disciplines has been shifting from negative to positive since the 1990s, when LIS began to receive a growing number of citations from journals in computer science and management. This paper goes one step beyond and studies the place of metric-related research inside the LIS literature, as well as the exportation of this research outside LIS. It provides data on the evolution of the top authors cited in LIS papers and then analyzes the use of five different metric-related terms – bibliometri*, scientometri*, info[r]metri*, web[o]metri* and altmetri* – inside and outside the LIS literature.

The next section provides a short primer on bibliometric methods and their limitations and details the specific methods used for the compilation of the data presented in this paper. It is followed by the presentation and discussion of the results and ends with a few concluding paragraphs.

Bibliometric Data and Methods

Bibliometric data are typically compiled using citation indexes such as Thomson Reuters' (formerly ISI) *Web of Science* (*WoS*) or Elsevier's *Scopus*. *Google Scholar* is also increasingly used for compiling bibliometric data at the level of individual researchers, although its use for macro-level data is much more problematic. *WoS* indexes the articles published in about 11,500 journals; *Scopus* indexes articles in approximately 17,500 journals. To be indexed in these citation databases, journals have to fulfill several criteria [5], among which citations received is only one of many – although it has historically been the main criteria in the case of the *WoS*. Despite differences in terms of coverage, the results obtained in terms of numbers of papers published and citations received are very highly correlated at the level of countries [6]. Although these data sources identify several types of documents, only articles, research notes and review articles are generally used in bibliometric studies, because they represent the main channels of scholarly dissemination [7]. An additional strength of these two databases is that they index the addresses of all authors, which allows analysis of the regionalization of scientific production – what countries, institutions or cities are the most active in a specific area – and the analysis of collaboration patterns.

These databases, however, have several limitations in terms of coverage, and the proportion of published literature that they index varies considerably across the spectrum of disciplines. As a consequence, bibliometric indicators are generally considered to be very reliable for the natural sciences, engineering and health sciences, but much less so for the social sciences and humanities. These differences in coverage reflect the diversity of ways in which scholars in the social sciences and humanities (SSH) disseminate new knowledge, compared to scholars in the natural or medical sciences. Several researchers have emphasized the fundamental differences between the communication practices of researchers in those two domains [8]. This divergence is reflected by greater use of monographs and conference proceedings in SSH as well as a lower use of journal articles [9, 10]. Unfortunately, no database covers these other forms of publication as systematically and comprehensively as *WoS* or *Scopus* do for journal articles. Another source of limited coverage is the fact that research subjects in SSH are often more local [8, 11], and, consequently, researchers publish more often in their native language and in journals with more limited distribution [12]. Given that these "local" journals are often not indexed in the *WoS* or *Scopus*, the coverage of SSH research from non-English-speaking countries is much weaker than for English-speaking countries.

Data presented in this paper are from *WoS*, which includes the *Century of Science* and the *Century of Social Science* for the period 1900-1944, as well as the *Science Citation Index Expanded* (SCIE), the *Social Sciences Citation Index* (SSCI) and the *Arts and Humanities Citation Index* (AHCI) for the period 1945-2011. This database was preferred over Elsevier's *Scopus* because of its better historical coverage of LIS. LIS is defined here as all papers published in journals to which the field "information science & library science" was assigned in the classification created by the patent board for the Science and Engineering Indicators Series of the National Science Foundation (NSF). Other field groupings used (medical sciences, natural sciences and other SSH) are also drawn from this classification scheme. For the years 1900-2011, the LIS dataset comprises 160 journals and about 320,000 documents, of which slightly less than a third are research articles. However, the analysis of the place of metrics outside LIS uses the full *WoS* dataset, which comprises approximately 37 million papers and 820 million references.

Results and Discussion

Table 1 presents by decade the top 10 authors that are the most cited in the LIS literature with those who are mostly known for their contribution to LIS-metrics indicated in bold face. It clearly shows the increase of the importance of metrics within LIS: although no metric-related author made it to the top 10 in the 1950s – an obvious reflection of the fact that ISI had yet to release its first version of the *Science Citation Index* – Eugene Garfield makes it to the top 10 in the 1960s and 1970s, where he is joined by Derek de Solla Price in the 1980s and then by Blaise Cronin and Christine Borgman in the 1990s. The proportion of metric-related researchers among the top 10 most-cited authors increases to five in the 2000s, with Wolfgang Glänzel and Leo Egghe joining the list, but with Derek de Solla Price leaving it. Peter Ingwersen could also be added to the list, as part of his work dealt with metrics.

Although the 2010s are far from complete, these results suggest that this decade might be a decade of metrics. Indeed, for the years 2010 and 2011, nine of the top 10 most cited researchers are mostly known for their contributions to LIS-related metrics. On the whole, this table clearly shows the increasing importance over the last six decades or so of metrics in the LIS citation landscape. The importance of this area is also reflected by the creation of new journals such as *Scientometrics* in 1978 and the *Journal of Informetrics* in 2007.

Figure 1 presents the annual number of papers having a metric-related keyword in its title (panel A) or abstract (panel B). One can easily see a steep increase in the use of bibliometri* since 2003. More specifically, while about 40 papers had bibliometri* in the titles in 2000, this number increased by more than three times to 130 in 2011. The increase of the use of bibliometri* is even steeper in abstracts, which increased from 50 to almost 250 papers over the same period. The two panels also show an increase of the use of scientometri* for the same period, although it is not as steep as that of bibliometri*. Interestingly, the use of bibliometri* and scientometri* in titles (panel A) was quite similar in the mid-1990s, which is likely a reflection of their use in other areas of social sciences.

On the other hand, the term info[r]metri*, which emerged in the late 1970s and early 1980s has been much less used by authors throughout the period and actually has decreased since the mid-2000s. This might be due to a more precise use of words by authors, as info[r]metri* can be considered as more generic than scientometri* or bibliometri*. Although web[o]metrics is on the rise and since 2009 is used almost as often as info[r]metrics in the titles and abstracts of papers, no paper was found having the word altmetri* in its title or its abstract. This lack might be due to the novelty of these types of metrics, as well as to the alternative – and perhaps reflexive – approach of its advocates, who seem to prefer diffusing their papers outside of journals.

TABLE 1. Top 10 authors most cited in LIS papers, by decade.

1950s	1960s	1970s	1980s	1990s	2000s	2010s
PERRY-JW	VICKERY-BC	LANCASTER-FW	LANCASTER-FW	SALTON-G	**GARFIELD-E**	**GLANZEL-W**
SHERA-JH	LUHN-HP	MIKHAILOV-AI	**GARFIELD-E**	LANCASTER-FW	SALTON-G	**GARFIELD-E**
SHAW-RR	MIKHAILOV-AI	SALTON-G	SALTON-G	**GARFIELD-E**	SPINK-A	**LEYDESDORFF-L**
TAUBER-MF	SHERA-JH	**GARFIELD-E**	LINE-MB	SARACEVIC-T	**CRONIN-B**	**EGGHE-L**
RICE-SO	KENT-A	LINE-MB	MIKHAILOV-AI	BELKIN-NJ	INGWERSEN-P	**HIRSCH-JE**
WILSON-LR	TAUBE-M	VICKERY-BC	WILLIAMS-ME	**CRONIN-B**	WILSON-TD	**MOED-HF**
TAUBE-M	**GARFIELD-E**	BROOKES-BC	WHITE-HS	**BORGMAN-CL**	KUHLTHAU-CC	**VANRAAN-AFJ**
SHANNON-CE	SALTON-G	CHERNYI-AI	**PRICE-DJD**	BATES-MJ	**GLANZEL-W**	**CRONIN-B**
METCALF-KD	RANGANATHAN-SR	SHREIDER-YA	SARACEVIC-T	**PRICE-DJD**	**BORGMAN-CL**	**BORNMANN-L**
BROWN-CH	CLEVERDON-CW	SHERA-JH	BROOKES-BC	DERVIN-B	**EGGHE-L**	WILSON-TD

FIGURE 1. Number of papers with bibliometri*, scientometri*, info[r]metri* or web[o]metri* in the title (panel A) or abstract (panel B), 1972-2011.

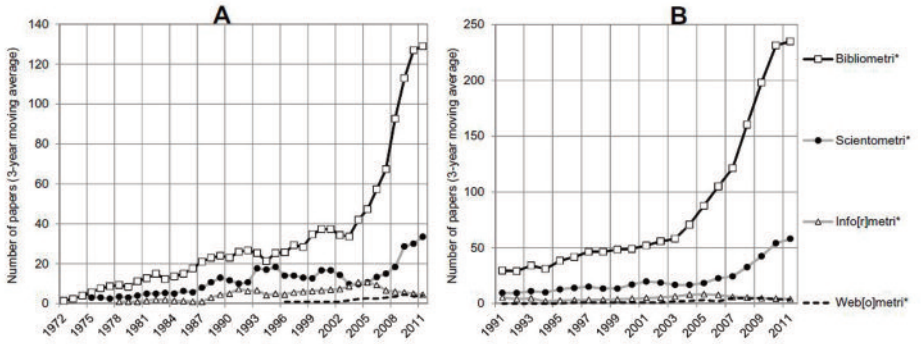


FIGURE 2. Percentage of all papers having bibliometri*, scientometri*, info[r]metri* or web[o]metri* in the title (panel A) or abstract (panel B) by discipline of the journal, 1972-2011.

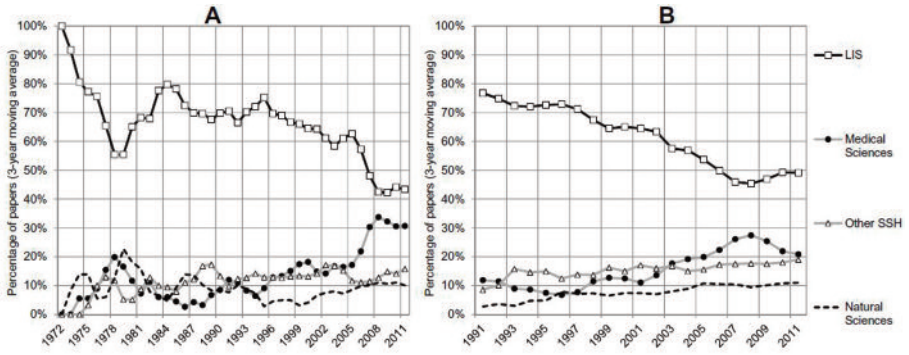


Figure 2 provides evidence of the use of the metric-related keywords within LIS and in other disciplines – categorized as other SSH (all SSH excluding LIS), medical sciences and natural sciences. Taking all papers having at least one of the five keywords as the denominator, it presents the distribution of the use of metrics terms in the titles (panel A) or abstracts (panel B) of papers across these four groups of disciplines. For a given year, the sum of percentages obtained in each of the four groups of disciplines will be 100%. Unsurprisingly, at the beginning of the 1970s, almost all papers related to metrics were published in LIS journals. After some fluctuations at the beginning of the period, caused by the small number of papers involved, this proportion slowly decreases from 80% in the mid-1980s to about 40% in 2008 and has been stable since then. This percentage is slightly higher for abstracts, which suggests that a larger proportion of metrics papers published in LIS do not incorporate one of the metric-related terms in their title.

The area outside LIS where *metrics* is most often used today is medicine. Although medical journals accounted for less than 5% of all metrics papers in the mid-1980s, they represent about a third of all papers having metrics terms in their title in 2008. It is worth noting that this increase might be, at least in part, due to LIS-related scholars publishing in these disciplines, such as Eugene Garfield who has published regularly in medical journals since the 1980s. In any case, this analysis clearly shows an increase in the interest in bibliometric measures from researchers in these disciplines, with many discussions surrounding the impact factor, citation analysis and research evaluation in general. Bibliometric methods are also increasingly discussed in the natural sciences – especially in physics – although to a lesser extent than in medical sciences. In 2011 natural sciences journals accounted for about 10% of all papers on the topic.

Within the disciplines included in other SSH, journals assigned to the field broadly defined as social studies of science (STS) were publishing a large proportion of the papers using bibliometric methods at the beginning of the period. The STS field greatly contributed to the legitimacy of citation analysis in the 1970s and 1980s by providing a framework for studying citations and their functions, as well as performing several analyses on the social stratification of science, cumulative advantage and other structuralist analyses of the scientific community, such as those of the Coles [13], Merton [14] and Zuckerman [15]. During this period, bibliometric methods were considered fundamental to the field. Since the 1990s, however, we have observed a change in the preferred methods of STS scholars, which shifted to ethnomethodology and other qualitative research methods that are more adapted to what is now mostly a case study-based literature. As a consequence, the recent handbook of the discipline – the *Third Handbook of Science and Technology Studies* – does not discuss bibliometrics at all, nor does it contain a single reference to Derek de Solla Price, the editor of the first *Handbook* [16]. This decrease in the use of LIS-metrics by STS scholars was compensated by an increase in other areas of SSH, mainly economics and policy studies related to research and innovation as well as research evaluation. Hence, the proportion of other SSH within all bibliometric literature is relatively stable throughout the period and represents about 10-20% of all bibliometric papers.

Conclusion

This short paper provides an historical account of the use of bibliometrics and other related metric research inside and outside of LIS. It provides evidence that metrics are increasingly important in LIS literature, as metric-

related authors occupy a growing proportion of top-cited authors. Concurrent with this increase, a larger proportion of metrics-related research is being published outside of LIS, contributing to a leveling of the balance of trade of LIS vis-à-vis other disciplines. Although this paper did not provide any comparison of the export of LIS-metrics to other contributions made by LIS literature, it seems likely that metrics are indeed one of the main exports of LIS. As a social science with deep professional roots, LIS research focused for most of the 20thcentury on classification, cataloguing and other practical aspects of the profession [3] that are less likely to be of interest to other disciplines. Similarly, most theories developed by scholars of the field are generally quite LIS-focused (information retrieval, information needs and so forth), so it seems unlikely that these theoretical contributions are the cause of the shifting balance of trade observed in the literature.

It is worth noting that for the last five years analyzed (2007-2011) the majority of bibliometric papers were published outside of LIS, with medical journals publishing almost as many as LIS journals. Given the wide interest in one of the main applications of bibliometrics – research evaluation and monitoring – it can be expected that its importance will continue to increase in the medical and natural sciences. In the social sciences, however, the tendency is less clear. Although research on science and innovation policy is still using bibliometric methods, its future for use in domains such as STS is ambiguous. As Bourdieu [17] argued, dominant agents of a discipline determine the legitimacy of its research objects and methods. And that observation is no truer anywhere than it is in SSH, where the appropriateness of a topic is governed by whether an author can persuade the community of its importance. Although the current dominant agents of STS – such as journal editors – are clearly in favor of other research methods, if not *against* metrics, it is possible that LIS-based quantitative methods will regain popularity in those disciplines, as editorships are not forever. Scientific revolutions sometimes come full circle. ■

Resources Mentioned in the Article

[1] Wouters, P. (1999). *The citation culture.* [Unpublished doctoral dissertation, University of Amsterdam]

[2] Moed, H.F. (2005). *Citation analysis in research evaluation.* Dordrecht: Springer.

[3] Larivière, V., Sugimoto, C., & Cronin, B. (2012). A bibliometric chronicling of library and information science's first hundred years. *Journal of the American Society for Information Science and Technology, 63*(5), 997-1016.

[4] Cronin, B., & Meho, L.I. (2008). The shifting balance of intellectual trade in information studies. *Journal of the American Society for Information Science and Technology 59*(4), 551–564.

[5] See http://thomsonreuters.com/products_services/science/free/essays/journal_selection_process/ and www.info.sciverse.com/scopus/scopus-in-detail/content-selection

[6] Archambault , É., Campbell , D., Gingras , Y., & Larivière, V. (2009). Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology, 60*(7), 1320-1326.

[7] Moed, H.F. (1996). Differences in the construction of SCI based bibliometric indicators among various producers: A first overview. *Scientometrics, 35*(2), 177–191.

[8] Hicks, D. (2004). The four literatures of social science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.) *Handbook of quantitative science and technology research* (pp. 476-496). Dordrecht: Kluwer Academic.

[9] Larivière, V., Archambault, É., Gingras, Y., & Vignola Gagné, É. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology, 57*(8), 997-1004.

[10] Lisée, C., Larivière, V., & Archambault, É. (2008). Conference proceedings as a source of scientific information: A bibliometric analysis. *Journal of the American Society for Information Science and Technology, 59*(11), 1776–1784.

Resources continued on next page

Resources Mentioned in the Article, cont.

[11] Nederhof, A.J., Zwaan, R.A., Debruin, R.E., & Dekker, P.J. (1989). Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioral sciences: A comparative study. *Scientometrics, 15*(5-6), 423-435.

[12] Gingras, Y. (2002). Les formes spécifiques de l'internationalité du champ scientifique. *Actes de la recherche en sciences sociales*, 141-142, 31-45.

[13] Cole, J.R., & Cole, S. (1973). *Social stratification in science.* Chicago: University of Chicago Press.

[14] Merton, R.K. (1973). *The sociology of science: Theoretical and empirical investigations.* Chicago: University of Chicago Press.

[15] Zuckerman, H. (1977). *Scientific elite: Nobel Laureates in the United States.* New York: The Free Press.

[16] Fuller, S. (2009. March). [Review of the *Handbook of Science and Technology Studies* by Edward J. Hackett], *Isis 100*(1), 207-209.

[17] Bourdieu, P. (2004). *Science of science and reflexivity.* Chicago: University of Chicago Press.