DESIDOC Bulletin of Information Technology, Vol. 24, No.1, January 2004, pp. 3-8

A Bibliometric Approach towards Mapping the Dynamics of Science and Technology

B M Gupta & Sujit Bhattacharya

Abstract

This paper attempts to highlight the role of bibliometrics in studying the dynamics of science and technology. Tools and techniques available in bibliometrics to address and understand the complexities of scientific fields are explored. The paper concludes that for wider acceptance among academicians and policy makers, bibliometric approach should ingrain itself within sociology and philosophy of science in studying the different facets of science and technology.

1. BACKGROUND

Science and technology studies are in a state of continuous flux. Research areas/sectors once strategic become stagnant or obsolete, while new and promising fields appear. In these new fields, research communities develop, new journals are created to address research problems, new areas compete with established fields for funding support, and other related developments occur. Policy makers and academicians are often confronted with these ever increasing and changing demands of science and technology. Studying the dynamics of science and technology is thus important as it provides the necessary clues to understand how the fields are changing and the inception and subsequent growth of new fields. Well-constructed studies can provide the necessary clarity to academicians and policy makers to get a more informed judgment of a research field. This helps in decision-making or in proper formulation of research themes and programmes.

There has been a long tradition of studying scientific disciplines in philosophy of science based on discourses and debates (*see for instance* Popper[1]; Lakatos[2]; Toulmin[3]; and Hesse[4]). Emergence of a discipline, origins of problems therein, and how problems are solved have been some of the major issues investigated by scholars in philosophy of science. Research in sociology of science, by embedding a particular social structure in context, focused on growth of disciplines (e.g. Mullins[5]) scientific controversies (e.g. Collins[6] and Mackenzie[7]); work in scientific laboratories (e.g. Latour and Woolgar[8]). Historians of science (e.g. Kuhn[9]; Shapin[10]; and Thackray[11]) have tried to understand the changes in a field by analyzing the developments over long time-periods. Writers on science policy and political dimensions of science (Weingart[12] and members of the Starnberg School; Rip[13]; Bohme, Van Den Daele, Hohlfeld, Krohn and Schafer[14]) have used these debates to highlight the current concerns.

In each period, various problems and objects were at the center of science investigation and new models of development were discussed. By the second half of 20th century, methodological conceptions and models were more actively discussed in the philosophy of science, where even greater importance was attached to such concepts, as scientific specialty, science paradigm, research programme, and social group.

Parallel to this, some important research works appeared on the horizon focusing on the problems in the formulation and development of new scientific directions and producing maps of scientific disciplines, as well as studying scientific knowledge and scientific activity in their sociological aspects, leaning upon various forms of scientific publications or studying group of scientists and types of communication in science. At the same time, one also started witnessing convergence and interaction between the philosophical and sociological approaches in science (Marshakova[15]).

A new approach in studying scientific disciplines emerged when scholars started substantiating their debates with quantitative investigations of research publications, and using other codified attributes within these research publications like authors, affiliations, citations in papers, etc. Quantitative studies based on research publications became commonly known as 'bibliometric' studies. Under, bibliometric approach, a series of new methods and techniques were developed and used by scholars such as D. de Solla Price,[16] Eugene Garfield,[17] Belvith Griffith,[18] Diana Crane.[19] New concepts and terms like exponential growth of science, invisible college, gatekeepers, etc., emerged and became popular over a wider scientific domain. Using the bibliometric approach, knowledge representation at the level of scientific specialty, intellectual structures, informal and formal networks in both natural and social science were investigated.

2. BIBLIOMETRIC APPROACH IN STUDY OF SCIENCE AND TECHNOLOGY

The bibliometric approach to science and technology is primarily based in quantitative characteristics; attributes of research publications such as article titles (including keywords and phrases), authors (including their addresses, co-authors and reputation), books, journals (titles, subjects, and origin of country), etc. The objects of study are often grouped under the various attributes as mentioned above. The methodology of science views scientific knowledge and its development through the presentation of objects described in a scientific publication. Scientific publication is by no means an invariable piece of information that can be published, stored, retrieved and delivered on request. A scientific publication is a kind of written material containing information with respect of scientific activities—either in its physical form, or its electronic equivalent in a computerized database (Tijssen[20]). Social scientists consider scientific publication as a web of science, produced under social conditions.

A scientific paper or text not only reveals the world-building strategy of its authors, but also the nature and force of the building blocks derived from the domain of science from which it draws and to which it contributes. Thus, it provides access to the dynamics of science to the shared worlds that constitutes a means of mutual (or evolving) control (Callon *et. al.*[21]). Each publication is uniquely represented by the bibliographical information it contains. In addition, each distinct type of publication has its own set of common characteristics and attributes. Two basic units of bibliographic information are generally used: (a) items, which refer to information in or about publications. The principle items of analysis are author names, addresses, citations, and keywords, as well selected words from the title, abstract, or the full text. Abstracting and indexing services often add further bibliographical information such as type of publication, language, one or more controlled terms (keywords given by data producers), subject classification terms or codes and (b) entities, which refer to (aggregates of) publications. These entities represent a set of publications, starting from the micro-level (a single author) via meso-levels (varying in size from university departments, scientific journals to research specialties) to macro-level entities (scientific sub-fields or countries). The information required for bibliometric analysis is derived not only from primary scientific journals, but also from indexing and abstracting journals and databases.

There are two common methodological approaches to the quantification of information flow. The first approach builds directly on the work of Price and his predecessors and uses publication as a whole or its attributes such as authors' names, addresses, citations and keywords, etc.; and the second approach consists in identification of links between objects, their co-occurrences and networks. Callon, Courtial, Turner, Bauin (see for instance Callon, *et. al*.[22]), were some of the eminent scholars who helped to define this approach.

In the first approach, depending upon the choice of object for a bibliometric study, different pictures of dynamics can be obtained as regards to the state of science and technology. One-dimensional or scalar techniques are generally used. The scalar techniques are based on direct counts (occurrences) of specific bibliographic elements such as publications and patents. For example, one can model the growth of a scientific field and the direction it is going to follow, by measuring annual publication data. Similarly, if the bibliometric study has its object of study, e.g. keywords, then alternations in keyword frequencies can reflect some substantial regrouping of research directions in science, and (much more rarely) linguistic changes. Similarly, the study of the distribution of productivity of authors and articles in journals in different time periods can be considered in terms of indicators, which can throw light on the level of development in the scientific specialties. The use of citation data and their age distribution over different time periods can also reveal the level of obsolescence and the structural changes taking place in the field from time to time. Using this approach, it is now possible to build a bibliometric monitoring system capable of diagnosing important characteristics of research performance (output, impact, international collaboration and so on) and trends in these performance characteristics over time. Such bibliometric monitors are useful for institutional research management as well for creating (inter)national benchmarks.

Under the second approach, the analytical bibliometric procedures are not directed at obtaining characteristics, but to identify relations among constituting elements in a research field. The two-dimensional or relational indicators are constructed from co-occurrences of specific items, such as the number of times keywords, citations and authors are mentioned together in publications in a particular field of science. This can be extrapolated in technology studies by investigating patents in a similar manner.

Relational data are constructed from co-authored publications, their addresses, co-occurrences of keywords from the text, and the reference lists. Each aspect reflects a different aspect of science. For example, address can be used for assessment of international cooperation, while citation relation may indicate intellectual links between sets of publications. Bibliometric maps can be built for exploring the underlying structure of similarities and interrelationships between items and/or entities.

2.1 Frequently Used Bibliometric Maps

So far, the following four types of bibliometric maps have been frequently used, constructed from research journals/articles:

(a) *Journal-to-journal citation maps* to uncover journals having strong linkages with each other, journals that are central implying most of the other journals refer to them, etc. (Leydesdorff[23]). By juxtaposing the fields/subfields of each journal in this map, it is possible to reveal the linkages among the fields or sub-fields. These maps thus can reveal the macro-level structure of a scientific field.

(b) *Co-citation maps* depicting the extent to which two articles are cited together in other articles (Small[24]). Clustering procedures make it possible to obtain co-citation clusters which reflect the structure and content of research-front specialties.

(c) *Co-word maps* are based on identifying pair of keywords, which occur together in large number of documents (Callon and Rip[25]). The framework assumes that when two words appear together in titles or abstracts or in full text for a same document it indicates possible link between two distinct centers of interest designated by these words. Links like this when repeated in large number of documents can delineate a research sub-topic or intellectual sub-domain that may not have entered the mainstream of research delineation of that field.

(d) *Co-classification maps* are based on the analysis of the co-occurrences of subject classification terms. The classification terms mostly represent cognitive elements, which relate to scientific topics, specialties, or fields (Van Raan and Peters[26]).

Using sufficiently large amount of publications data, this enormous network of linkages combines to 'abstract' structures. With mathematical techniques based on matrix algebra these structures can be displayed on two-dimensional space – maps of science. The underlying objectives of these spatial mapping representations are to understand the regularities, i.e., patterns underlying the data. Among the common dimensionality-reduction techniques used in such mapping are: factor analysis, cluster analysis, multidimensional scaling (MDS), and correspondence analysis. The first three methods, i.e., factor, cluster and MDS share a common feature – they operate on a square, symmetric matrix. The rows and columns of the matrix refer to the same entities, and contents of the matrix consist of the same measure of association or proximity between pairs of these entities. Correspondence analysis allows one to analyse the association among the rows and columns separately and also project them jointly in a low dimensional subspace. Another technique utilized to understand the multivariate relationship among the variables in the data is through the social network approach. In this approach, the various concepts of the network like centrality, density, nodes, etc. are used to reveal the linkages and other attributes of the data.

Both the scalar and relational indicators can be constructed on micro (research specialties or groups), meso (large scientific fields or organisations, companies), or macro (national, international) level. One can also do mapping of research fields, which allow us to show the representation of structural relations within a specific field, with all its sub fields and specialties. Further, these bibliometric maps can be constructed for successive years, thus representing the temporal developments (dynamics) in the field, together with the (changing) role of research groups or institutes concerned (Van Raan[27]).

Bibliometric analysis, in addition, can act as a useful tool for research management. For example, taking the field of polymer chemistry, bibliometric indicators can help us find answers to questions like: (i) what is the level of scientific activity (in terms of research output) and trends—overall as well as in the different countries or regions in the field of polymer chemistry? (ii) what impact this activity has overall, within the academic and business-sector research? the most recognized groups, and what are their specialties in this field? (iii) influence of research activities on R&D developments in specific sub-sectors within this field, say for example in new materials? (iv) collaboration intensity of different countries in this field? (v) map this field showing its most important research areas and the linkages with neighboring fields? etc.

However, bibliometric methods and studies would be more acceptable if they are able to alien and bridge the gap with disciplines mainly sociology of science, history of science and philosophy of science. Empirical findings in bibliometrics should be synthesized so that researchers in other disciplines can be motivated to use their empirical findings. Laydesdorff (1989) prescription aptly sums up the task in hand:

"Only by further specifications of what empirical studies contribute to specific theoretical questions can we systematically further the relations between qualitative theory and scientometric (refer note below) methods in S&T studies. (Leydesdorff[28])."

Note**:** Bibliometrics is a sub-domain of scientometrics. Scientometric research is devoted to quantitative studies in science and technology. The quantitative indicators mainly utilized in these studies include input indicators like R&D funds, S&T manpower, etc. ; the output indicators are journal articles, patents, etc. The output indicators are mainly addressed through bibliometrics.

REFERENCES

1. Popper, Karl. The logic of scientific discovery. Hutchinson, London. 1959.
2. Lakatos, Imre. Falsification and methodology of scientific research program. *In* Criticism and Growth of Knowledge, *edited by Lakatos, Imre & Musgrave, A*. Cambridge University Press, 1970, 91-195.
3. Toulmin, S.E. Human understanding. Part 1. Oxford University Press, 1972.
4. Hesse, Mary B. The structure of scientific inference. MacMillan, London, 1974.
5. Mullins, N.C. A model for the development of a scientific specialty: The phage group and the origins of molecular biology. *Minerva, 1972,* **10**, 51-82.
6. Collins, H.M. The seven sexes: A study in the sociology of a phenomenon or the replication of experiments in physics. *Sociology,* 1975, **9**, 205-224.
7. MacKenzie, Donald. Statistical theory and social interests: A case study. *Social Studies of Science,* 1978, **8**, 35-83.
8. Latour, Bruno and Woolgar, Steven. Laboratory life: The social construction of scientific facts. Sage, London, 1979.
9. Kuhn, Thomas S. The structure of scientific revolution. Chicago University Press, Chicago, 1970.
10. Shapin, Steven. The politics of observation: Cerebral anatomy and social interests in the Edinburgh phrenology disputes. *In* On the Margins of Science: The Social Construction of Rejected Knowledge, *edited by* Wallis, Roy G. University of Keele. 1979, 139-178 (Sociological Review Monograph)
11. Thackray, Arnold. John Dalton: Critical assessments of his life and science. Cambridge, Harvard University Press, Mass. 1972.
12. Weingart, Peter. The scientific power elite: A chimera. The deinstitutionalization and politicization of science. *In* Scientific Establishments and Hierarchies, *edited by* Elias, Norbert, Martins, Hermino, & Whitley, Richard M. Reidel, Dordrecht and Boston, 1982, 71-87 (Sociology of Science Yearbook, Vol. 6).
13. Rip, Arie. A cognitive approach to science policy. *Research Policy,* 1981, **10**, 294-311.
14. Bohme, Gernot; Daele, Wolfgang van den; Hohlfeld, Rainer; & Schafer, Wolf. Finalization in Science: The Social orientation of scientific progress. Reidel, Dordrecht, Boston and Lancaster, 1983.
15. Marshakova-Shaikevich, I.V. Bibliometrics as a research technique in epistemology and philosophy of science. *International Forum on Information and Documentation,* October 1993, **18**, 3-9.
16. Price, D. de S. Little science, big science. Columbia University Press, New York, 1973.
17. Garfield, E. Citation indexing its theory and applications in science, technology and humanities. Wiley, New York, 1979.
18. Graffith, B.C., Miller, A.J. Networks of informal communication among scientifically productive scientists. *In* Communication among Scientists and Engineers, *edited by* C. Nelson & D. Pollock. 125-140, Laxington M.A., D.C. Heath, 1970.
19. Crane, D. Invisible colleges. Chicago University Press, Chicago, 1972.
20. Tijssen, Robert J.W. Introduction to scientometric studies of science. *In* Cartography of Science, Tijssen, Robert J.W. DSWO Press, Leiden University, 1992.
21. Callon, Michel; Law, John & Rip, Arie. Quantitative scientometrics. *In* Mapping the Dynamics of Science and Technology, e*dited by* Callon, Michel,

Law, John & Rip, Arie. The MacMillan Press, London, 1986, p.103.

22. Callon, M.; Courtial, J.P.; Turner, W.A. & Bauin, S. From translations to problematic networks: An introduction to co-word analysis. *Social Science Information,* 1983, **22**, 191-235.
23. Leydesdorff, L. Various methods of mapping of science. *Scientometrics* 1987, **11**, 295-324.
24. Small, H. Co-citation in the scientific literature: A new measure of the relationship between publications. *Journal of the American Society for Information Science,* 1973, **24**, 265-69.
25. Callon, M.; Courtail, J.P.; Laville, F. Co-word analysis as a tool for describing the network of interactions between basic and technology research: The case of polymer chemistry. *Scientometrics*, **22**(1), 153-203.
26. Van Raan, A.F.J. & Peters, H.P.F. Dynamics of a scientific field analyzed by co-subfield structures. *Scientometrics* 1989, **15**, 607-20.
27. Van Raan, A.F.J. Advanced bibliometric methods to assess research performance and scientific development: Basic principles and recent practical applications. R*esearch Evaluation,* 1993, **3**, 151-66.
28. Leydesdorff, L. The relations between qualitative theory and scientometrics method in science and technology studies. Introduction to the topical issues. *Scientometrics,* 1989, **15**, 333-47.

Contributors: **Sh. B. M. Gupta,** Scientist F, National Institute of Science, Technology and Development Studies (NISTADS), K.S.Krishnan Marg, New Delhi 110 012.
e-mail: bmgupta@nistads.res.in

Sh. Sujit Bhattacharya, Scientist E1, NISTADS, K.S.Krishnan Marg, New Delhi 110 012. e-mail: sujit@nistads.res.in