

An Overview on Prevailing Bibliographic and Citation Analysis Tools

Priya¹ Kirti Bhatia² Rohini Sharma³ Shalini Bhadola⁴

¹Student ^{2,3,4}Assistant Professor

^{1,2,4}Sat Kabir Institute of Technology and Management, Haryana, India

³Government College for Women, Rohtak, India

Abstract— Now a day, a lot of research work is performed in every field of real world. It can be in the field of science, technology, literature, commerce and medical etc. Lots of researchers are doing their research and publishing it in different journals, magazines, electronic and print media. Some research becomes very popular and followed by many peoples around the world. As everyone wants to excel in her filed and wants to increase the popularity of their work, they have to spread the awareness of their work. A research work can be viewed and followed by its citation number. A research work with higher citation number is considered more reliable than a research work with less citation number. In this paper, we have investigated the importance of measurement of Bibliography and citation in the area of research. We have also analyzed the different approaches available for measurement, their pros and cons.

Keywords: Bibliometric, Citations Analysis Tool, Impact Factor, H-Index

I. INTRODUCTION

To determine the impact and significance of the research citation and Bibliometric are two popular ways. The background of science and research is quickly developing day by day. Everything is now available on the Internet. Currently, the emphasis is progressively changing from if a researcher has published a paper to where she has published it as well as the impact of research work on the scientific society.

Quantitative analysis is the key tool of measuring the impact like counting, measuring, comparing quantities, analyzing measurements. Technical investigation, storing and transmitting results of the research results via publications, have become massive and composite. Particular information and proficiency are no more adequate means for perception of tendencies or for producing conclusions. To emphasize noteworthy or favorable fields of research, and to supervise improved ventures in technology are improving day by day.

Those in universities, government offices and labs, and boardrooms must decide what research should be supported and what should not, or which research projects and researchers should receive more support than others. Until relatively recently, peer review was the main route by which science policymakers and research funders made policy decisions about science. A library faced with collection decisions, a foundation making funding choices, or a government office weighing national research needs must rely on expert analysis of scientific research performance.

Bibliometric (or Scientometrics) proved as a key instrument of quantitative analysis of science. Bibliometric can analyze and gather statistics about publications like journal articles and their associated citation totals. Now a day, computable assessment of publication and citation of data is used in nearly whole universe with a substantial science

enterprise. It is used for research performance assessment, particularly in education departments and research labs, government funded project works, research administrators, information authorities and in libraries.

Citation data can be used for a variety of reasons like in a competitive atmosphere, it becomes necessary for an organization to express its performance which supports its mission. Definitely, the measurements may be a dominant statement instrument to maintain the standards. Now a day, education institutes have to keep track of their research performance. Statistics about research performance from several areas help the university to make a tactical decision and it start to support an area which is lacking behind in terms of research performance. This statistics helps the university to understand its position in context of research nationally and internationally. Like, how much research is being conducted by faculty and students every year? How many research papers are being published in national or International Journals? What is the impact factor [1] of these Journals? Now a day, Bibliometric systems along with a huge number of analysts are steadily set up in numerous countries and the group of analyst release Bibliometric report which are known as science indicators readings at fixed periods. Examples of such groups are the National Science Foundation (United States) and the European Commission.

Analysis of cited work is important in competitive environment. An organization's research and development employees perform recurring assessments of research operation and make a periodic report about it. These organizations level valuations should be complemented with departmental level research valuations for functioning of approval, term, teaching staff and assessment. Figure 1 shows the factors used to measure citation of a research article or research work.

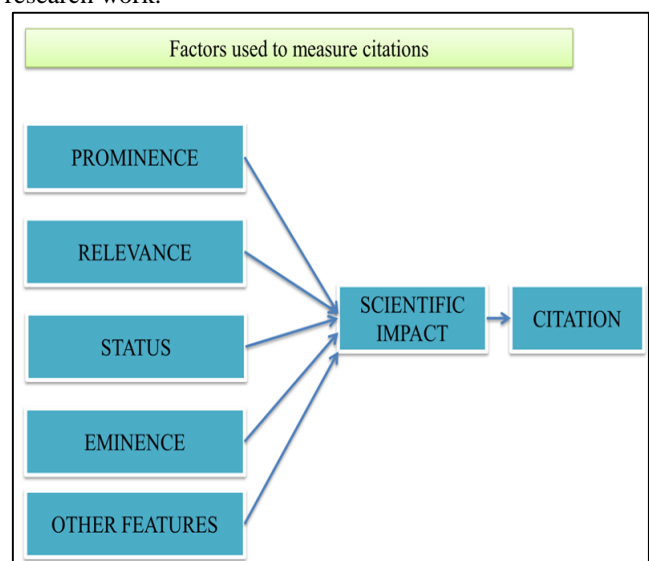


Fig. 1: Factors to Measure Citation

II. RELATED WORK

Authors in [2] have provided the likelihood of utilizing Bibliometric procedures for assessing investigation in the field of humanities. They have evaluated some recent work for the development of Bibliometric approaches for reviewing the humanities and it exhibits the structural, methodological variances and carried separate research in this research field. They have find out the potentials of employing Bibliometric approaches and thy have identified the requirement on contemporaries, integrative and the rustic character of research in numerous humanistic regulations. They have also highlighted certain talented methods and the likelihood of emerging a Bibliometric for the humanities was also inspected. Lastly, the knowledgeable features of precise regulations must be measured while constructing the quality signs, and the significance of containing researchers from the humanities in the development was also considered greatly.

III. BIBLIOMETRIC

Bibliometric is statistical study of written periodicals, like books or articles [2]. The techniques of Bibliometric are regularly utilized in the area of library and information science, containing scientometrics. It mainly provides computable examination of academic literature [3]. Citation analysis is a frequently manipulated Bibliometric technique that is based on creating the citation graph which is a network or graph demonstration of the citations among articles. Numerous research areas make use of Bibliometric approaches to investigate the influence of their area, the influence of a group of researchers, the effect of a specific article, or to categorize predominantly effective articles of a precise area of research[4][5][6]. Bibliometric has an extensive rangé of applications, like expressive linguistics etc. Initially, the word Bibliometric was introduced by Paul Otlet in 1934 and described as the magnitude of all features linked to the periodical and interpretation of volumes and papers. Bibliometric established through the post 1945 period make a greater impact in research, development of publication and organized electronic databases. It came into the research arcade very long ago, through Eugene Garfield's Science Citation Index (started in 1964) and in following years, it had been implemented by the US National Science Organization. In the Europe, the citation analysis emerged into research-management material, predominantly in the Netherlands and Hungary. Since mid-1990s, in US, the Bibliometric started to play the role of policy-building as a manner to standard organizations and academics. It's profitable prospective steered Elsevier to make the Scopus bibliographic database, as well as the basic significance of citation systems for exploration and detection prepared Google Scholar a preferred among various researchers. Due to the chronological wisdom of investigation, the book's management of procedures and mechanism became very popular.

Now a day, the Bibliometric supports research strategy and administration and also it is a well-known fragment of programmed assessment for some national organizations like UK Research Council. But even after so many years, we are not able to completely apprehend the use of citation data. Although some researcher's uses

Bibliometric, but they do not have any apparent measures about its implementation and interpretation. If an instrument that effects research greatly is used inaccurately, then it may crate trouble in research assessment area.

There are some more problems associated with it, like administrative paymasters and research executives need sensible statistics to influence the current trend rather than indicating what has happened in past. However citations statistics relate to effort subsidized years before the Bibliometric has come into the existence. Statistics from Social media can also be used an alternative of Bibliometric. But citation analysis constructed by years of work, and social media has very small life span. Therefore, if considerate observers do not provide improved responses, prompt specialists can haste to gratify policy shareholders. Through carefulness resolutions about the upcoming guidelines, subsidizing and recruitment of science should be cognizant by the methods and progresses conferred in Bibliometric.

IV. VARIOUS MEASURES OF BIBLIOMETRIC

A. Impact Factor (IF):

The IF or journal impact factor (JIF) of an educational journal is a measurement of annual average number of citations to latest research work issued in that journal. Often, it is utilized as a representation of the comparative significance of a journal within its discipline. A journal with higher IF than other journals is considered to be more significant than those with lower ones. It was invented by Eugene Garfield, the originator of the Institute for Scientific Information. The computation of IF started in 1975 for academic journals registered in the Journal Citation Reports. In any specified year, the impact factor of a journal is the number of citations, obtained in that particular year, by the research work issued in that journal for the duration of two previous years, divided by the total number of research work printed in that journal for the same duration and it is shown by the equation number (1):

$$IF_y = \frac{\text{Citations}_{y-1} + \text{Citations}_{y-2}}{\text{Publications}_{y-1} + \text{Publications}_{y-2}} \quad (1)$$

Fresh journals, which have released their first issue, will be able to get an impact factor past two years of indexing; henceforth, the citations for the first two years will remain zero and known as zero values. To receive an impact factor, a Journal will have to wait for at least three years. Sporadically, Journal Citation Reports allocates an impact factor to fresh journals by not as much of two years of indexing, according to statistics of partial citation [7][8]. The estimate constantly usages two whole and identified years of article tally, however for fresh articles count number is zero. Sometimes Almanacs and other non-regular journals print no article in a specific year and it influence the count number. The IF links to a precise time era; it is probable to compute it for an anticipated epoch. Moreover, the Journal Citation Reports comprises of a 5 year IF, that is computed by dividing the number of citations to the journal in a specified period by the number of research work printed in that journal in the preceding 5 years. Within a specific field, the impact factor is used to compare two or more than two journals. Thousands of science and social science journals are indexed in web of science [9]. Often, the impact factors of Journal are employed

to estimate the worth of specific articles and specific researchers. It is difficult to publish a paper in a journal with high impact factor and there are some good quality journals which exist even before establishment of impact factor. An IF is a journal level metric rather than an article level metric. An IF can be manipulated by journal, an individual or by a specific subject, i.e. impact factor of some subject specific article is more than other subject specific articles. Figure 2 shows the computation of Impact factor.

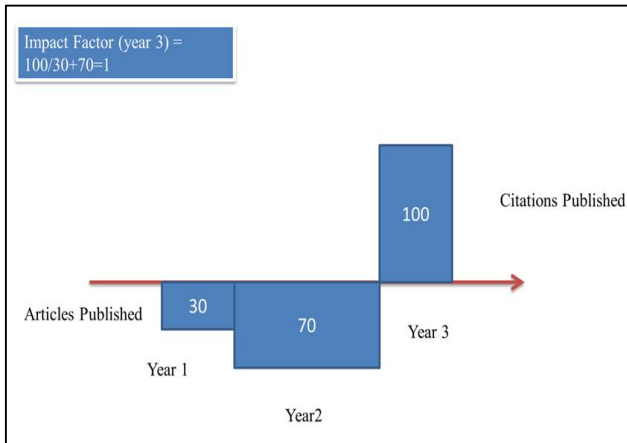


Fig. 2: Computation of Impact Factor

B. H-index:

The Hirsch index, or H-index, is a dissemination centered marker which relates to the number of articles at or above a specified citation level equivalent to the value of the citation threshold. It was proposed by Jorge E. Hirsch in 2005 [10]. Figure 3 demonstrates the computation of H-index. It indicates the number of papers (N) in a dataset which has N or more citations. E.g., if H-Index is 10, then at least 10 papers in a dataset have 10 or more than 10 citation. It alone reveals both the number of papers and the number of citations. It depends upon time period and field of research. Let f be a function in decreasing order from greatest to the least one, the H- index number can be computed as follows:

$$h - \text{index}(f) = \max \min(f(i), i) \quad (2)$$

The H- index can be used as an alternative to the impact factor to assess the performance of a researcher. It can be computed either through citation databases or through automatic tools. Some subscription based databases like WOS and Scopus provides automatic calculator for h-index computation. Due to different coverage, different databases may produce different h-index. E.g., WOS mainly covers high IF based Journals but it does not cover good conferences. On the other hand, Scopus has good analysis of conferences but it does not have good coverage of journals before 1990. However, Google Scholar covers both conference proceedings as well as Journal articles.

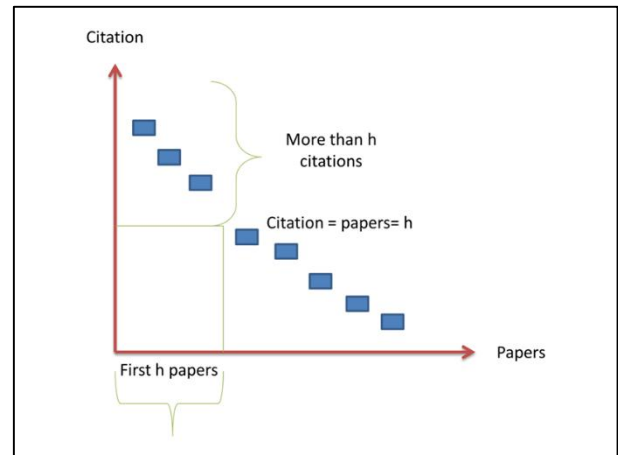


Fig. 3: H-Index Computation

C. SCImago Journal Rank:

It is an evaluation of scientific impact of academic journals and it describes both the number of citations obtained by a journal and the significance or status of the journals from which these citations are received.

The thoughtless use of research database and tools may pose big problems in the field of research. The h-index, which measures the contribution of a researcher's citations, is distinctively defective as it is not approachable to that area, however, its easiness can cause an excess of copied directories. The IF provides information on journals, but not of separate periodicals, however research executives use them to evaluate a person's individual performance. Technical persons issue conference proceedings, but present merely journal article for valuation. Statisticians publish occasionally and cite very profound into the historic research and environmentalists did exactly opposite of it. Economists and atom physicists have entirely dissimilar standards of writing a project. Only very less users of Bibliometric realize the inferences of these degrees, however most of them contemplate that they may grasp the numbers, or can encourage others to recite them in their manner.

V. TYPES OF DATA USED IN BIBLIOMETRIC

The standard of citation analysis is based on the type of data used for analysis. E.g. Thomson Reuters' citation index [11], available through Web of Science [12], is used as the universal benchmark for Bibliometric. Since half decade, the Thomson Reuters is assembling data from various fields like science, social science, technology and medical science. The Science Citation Index was developed in 1961 by Eugene Garfield, through this index, a researcher can be aware of the importance of his published work. This database is also significant in analyzing the growth and structure of science. Web of science (WOS) database covers journals from science, social science, Art and humanities and chemical reactions. The WOS covers approximately 9300 highly recognized journals from every filed. The WOS also comprises of all authors' name and their affiliations. A comparatively fewer basic of research writings is regularly cited, whereas the rest of the literature is either not cited or very irregularly cited. Additionally, some significant research that receives a huge number of citations in one area is also commonly cited in more areas. To cover such types of

citations which are related to more than one field, Thomson Reuters' multidisciplinary citation index was produced. In pursuance of maintaining the statistical coherence, Thomson Reuters' citation index comprises of entire work circulated in the reputed journals. It involves the basic means of distributing research outcomes, like editorials and analyses, articles, communication letters and books. A database may have a Bibliometric analysis which only includes articles and reviews, or it may include the various trivial article categories. Some popular databases are as follows:

A. Scopus:

It is a contribution based database and popular as a substitute of Web of Knowledge. It suggests an analogous article, author, and journal-level criterion, however it make use of marginally dissimilar procedures to compute these criterion. Important metrics are times cited and h-index, and innovative metrics such as SJR and SNIP from SCImago. Recently Scopus has started a third party application Altmetric for Scopus which implements the sidebar of Scopus pages to trace references of papers through social media websites, technology blogs, broadcasting channels, and recommendation executives.

B. Web of Science:

It is a Thomson Reuters contribution database which helps the users in presenting the Bibliometric through h-index which was introduced in 1982. It comprises Web of Science, for article and author inquiries, and Journal Citation Reports, for discovery of recognized and influenced journals. The key metrics are impact factor, times cited, h-index, Eigen factor, and area-based journal statuses. Although most of these metrics are not able to completely signifying scholarly standards in some subjects or areas, still these are deliberated as the effective benchmark in conventional Bibliometric.

C. Google Scholar:

It provides an easy way for authors through which they can keep track of citations of their research work. They can assess the number of citations of their articles, graph citations over a specific duration, and can calculate numerous citation measurements. They may create a public profile, henceforth if a scholar search an author by his name, the name appears in Google Scholar results. Moreover, it is simple to create and maintain it, No matter, how much articles belongs to an author or even if an article is written by more than one author, articles and the author' name appear in Google scholar search. The article or groups of linked articles can be added automatically or manually. As soon as there is a citation to any of the author's work, the citation metrics are calculated and revised spontaneously by the Google Scholar. An author may select to have his record of research work revised spontaneously or evaluate the updates by himself or to revise his work at any instance of time.

VI. FUNCTIONS OF BIBLIOMETRIC

Citations are the references scholars attach to their articles to clearly mention previous research related to their work. Following citations and considering their developments in background is a key factor for evaluation of the impression and effect of research work. Citation not only influences an

individual research work, it also affects a research organization. An article with high number of citation is considered a very significant and useful research work. Paper counts determine productivity and are the primary measure of Bibliometric and offer the raw material for complete analysis of citation. The paper counts aids ranking institutes to equate the efficiency and size of research outcome. While comparing publication counts among institutes, the number of investigators at every institute, attribute of research paper, publication year and contents must be considered.

VII. REASONS OF CITING A RESEARCH WORK

- Giving respect to inventors.
- Providing acknowledgment for linked work.
- Classifying procedure and tools.
- Offering contextual interpretation.
- Revising one's own research.
- Modifying the research of others.
- Assessing preceding research.
- Demonstrating privileges.
- Notifying researchers to upcoming research.
- Offering clues to inadequately distributed, defectively indexed, or un-cited research.
- Detecting unique publications in which a notion or idea was conferred.
- Classifying the inventive publications explaining an appellation idea or context.
- Discussing the work or concepts of others.

VIII. CONCLUSION

Bibliometric evaluates academic influence in every form like in citations to journal papers, data sets, conference proceedings and letters. Its symbols and tools were used for better understanding of and relevance of research. The outcome of the Bibliometric is considered as a landmark for the subject and a recommendation for other scholars. The results can be used by reviewers, groups and research grant societies who use Bibliometric to take decisions. Bibliometric assisted to lay research strategy and organization on a suggestion founded balance that might confront the expectations and preconceptions of well-known scientists. We have mentioned why it is necessary to measure the citation of a research. We have also provided the popular databases which are used to measure the impact of a journal, conference proceedings and letters. There are various tools (IF, H-index, SCI Imago rank) to determine the impact of a research in a particular field. Bibliometric is popular in all fields, like science, humanities and social science etc.

REFERENCES

- [1] https://en.wikipedia.org/wiki/Impact_factor.
- [2] Björn Hammarfelt, Beyond Coverage: Toward a Bibliometrics for the Humanities, Research Assessment in the Humanities pp 115-131, Springer, 2016
- [3] "Bibliometrics". OECD Glossary of Statistical Terms. OECD. 28 March 2013. Retrieved 5 January 2015.
- [4] De Bellis, Nicola (2009). Bibliometrics and citation analysis: from the Science citation index to cybermetrics. Scarecrow Press. p. 417. ISBN 978-0-8108-6713-0.

- [5] Yeung AWK, Heinrich M, Atanasov AG. Ethnopharmacology-A Bibliometric Analysis of a Field of Research Meandering Between Medicine and Food Science? *Front Pharmacol.* 2018 Mar 15;9:215. doi: 10.3389/fphar.2018.00215.
- [6] Alon, Ilan; Anderson, John; Munim, Ziaul Haque; Ho, Alice (2018). "A review of the internationalization of Chinese enterprises". *Asia Pacific Journal of Management.* 35 (3): 573–605. doi:10.1007/s10490-018-9597-5.
- [7] Maditati, Dhanavanth Reddy; Munim, Ziaul Haque; Schramm, Hans-Joachim; Kummer, Sebastian (2018). "A review of green supply chain management: From bibliometric analysis to a conceptual framework and future research directions". *Resources, Conservation and Recycling.* 139: 150–162. doi:10.1016/j.resconrec.2018.08.00.
- [8] "RSC Advances receives its first partial impact factor". *RSC Advances Blog.* 24 June 2013. Retrieved 16 July 2018.
- [9] "Our first (partial) impact factor and our continuing (full) story". *news.cell.com.* 30 July 2014. Archived from the original on 7 March 2016. Retrieved 21 May 2015.
- [10] https://en.wikipedia.org/wiki/Web_of_Science.
- [11] <https://en.wikipedia.org/wiki/H-index>
- [12] <http://mjl.clarivate.com/>
- [13] <https://clarivate.com/products/web-of-science/>

