The journey from librametry to altmetrics: a look back

Bidyarthi Dutta[1]

Asst. Professor
Dept. of Library & Information Science
Vidyasagar University
Midnapore 721 102, WB, India

Abstract

The twentieth century may be described as the century of development of metric sciences, i.e. librametrics, scientometrics, bibliometrics, informetrics, econometrics, technometrics, biometrics, sociometrics, psychometrics, educametrics and so on…Possibly the inception of cybermetrics was the concluding milestone of the metric sciences' voyage in the last century. The internet and open access revolution touched the crest at 1990s that laid down a milestone in 2001. The dawn of new millennium radiated the new spark of light across the globe, which is Wikipedia. The knowledge dissemination process demolished another barricade to ensure people's easy access. It was a new concept that drastically transformed the world of scholarly communication. Various new dimensions were added to the processes of information collection, storage, processing, dissemination and evaluation. The performance and impact measurements of these processes gradually systematized them towards more objectiveness. In this way newer metric sciences are developing incessantly in different subject domains. Recent developments include Wikimetrics, article-level metrics, altmetrics etc. In this paper all metrics generally associated directly or indirectly with the scope and context of library and information science have been discussed along with some brief historical background. Starting from Ranganathan's librametry, it addresses upto most recent metric studies. All metrics are categorized in three classes on the basis of respective time of inception, i.e. classical metrics, neo-classical metrics and modern metrics. The four metrics, i.e. librametrics, bibliometrics, scientometrics and informetrics are categorized under classical metrics. Cybermetrics is regarded as neo-classical metrics while remaining others evolved in 21st century are recognized as modern metrics.

Keywords

Librametry; Bibliometrics; Informetrics; Scientometrics; Webometrics; Cybermetrics; Wikimetrics; Author metrics; Article metrics; Altmetrics; Social metrics; Open source metrics; Google scholar metrics; h index; g index; i-10 index

1 e-mail: bidyarthi.bhaswati@gmail.com
Mobile no. 98308 84794

Introduction

The term metric was originated from the Latin word *metricus*[1] and French word *métrique*[2], which means a measure for something or any mean of deriving quantitative measurement or approximation. This word was first used in 1864[2]. This word envelops number of subject domains, e.g. general relativity under physics, networking, mathematics, software analysis etc. One of the most well-known uses of this word is found as unit of measurement in various subjects and LIS also. Any system of measurement invariably involves one crux subject from methodological viewpoint, which is nothing but statistics. Prof. P.C. Mahalanobis in the early fifties described statistics as a key technology, because it is required for all socio-economic development activities and since statistical techniques are used in all development and forecasting studies. Statistics is applied to almost all major subject areas under broad disciplines such as engineering science, medical science, agricultural science, social science, behavioural science, cognitive science etc. Statistics and related techniques when applied in depth to a subject domain may give rise to a new subject, for instance, application of statistics to social science or sociology gives rise to a new subject, i.e. sociometry. Similar other examples are, econometrics, educametry, psychometry etc. In the field of library and information science, the applications of quantitative techniques are very popular today. From library classificationist's viewpoint the subject *statistics* or *quantitative technique* may thus be looked upon as an isolate idea also apart from its existence as a major science subject. It is also included in the Auxiliary Tables of the major classification schemes (DDC, UDC or CC) just like other subjects such as history, philosophy, management etc. The early works in this subject domain goes back to mid 1920s and the studies involving applications of quantitative techniques to library and bibliographical work are commonly known as *statistical bibliography*. Wittig[3] discussed about the history of statistical bibliography indicating that the origin of the term could be traced to its first use by Cole and Eales[4] in 1917 and by Hulme[5] in 1922. The term "statistical bibliography" signals application of quantitative methods to library and bibliographic works. After Hulme, Gosnell[6] also used the term in his Ph.D dissertation entitled "The rate of obsolescence in college library book collection" (P.16), which was submitted to New York University in 1943. He used this term to put emphasis on the quantitative aspect rather than qualitative features. Raisig[7] also claimed that there was potential utility of statistical bibliography as a method of analysing information needs in medical sciences. The keyword "statistical bibliography" is formed from two words, i.e. statistics and bibliography. The word *statistics* indicates[8] "facts or data of numerical type that can be assembled, classified and tabulated so as to present information about a subject". The word *bibliography*[8] is derived from two Latin words, *biblion* and *graphos*, the former means books while the later means to write. The word *bibliography* indicates a list of authors' writings or the literature dealing with a certain subject or author. Hulme[5] defined bibliography as the science of organization of knowledge. This two words *statistics* and *bibliography* are combined to form the keyword statistical bibliography that broadly connotes application of quantitative methods in library science.

Classical metrics

Ranganathan[9] suggested as early as 1948 at the Aslib conference in Leamington Spa "that it is necessary for librarians to develop "librametry" on the lines of biometry, econometry and psychometry, since many of the matters connected with library work and services involve large number". The word 'librametry' or 'librametrics' is the couple of two separate words, library and metrics, which indicates application of mathematical models and statistical techniques to evaluate library systems and services. It is well-known fact that any library report comprising either library collection or library services can be picturesquely objective only when presented through statistical techniques. Since statistics speaks in an exact voice, there is no scope of any fuzziness. As Rao[9] commented, "In spite of his early attempt to define the scope of librametry, the subject hardly developed until the early 1970s" and Pritchard[10] used the term 'bibliometrics' in 1969 to describe all studies which seek to quantify the process of written communication. He[10] defined the term as follows, "The definition and purpose of bibliometrics is to shed light on the process of written communication and of the nature and course of a discipline (in so far as this is displayed through written communication) by means of counting and analyzing the various facets of written communication". The coining of the term bibliometrics is frequently credited to Pritchard, but Wilson[11] indicated that this term has a French precedent. Fonseca[12] pointed out the use of the French equivalent of the term, i.e. 'bibliometrie' by *Paul Otlet*[13] in 1934 in his *Traitée de Documentation. Le livre sur le Livre. Theorie et Pratique*, which is an obscure work and Section 124, pp.13-22, of this text is entitled 'Le Livre et la Mesure. Bibliometrie.' Sengupta[14] claims that Campbell[15] produced the first bibliometric study, using statistical methods for studying subject scattering in publications. The British standard glossary of documentation terms[16] described bibliometrics "as the study of the use of documents and pattern of publication in which the mathematical and statistical methods could be applied.

Potter[17] defined it as "the study and measurement of the publication patterns of all forms of written communication and their authors". Sengupta[18] described it as “organization, classification and quantitative evaluation of publication patterns of all macro and micro communication along with their authorships by mathematical and statistical calculus”. Hertzel[19] described it as “the science of recorded discourse, which uses specific methodologies, mathematical and scientific, in its research, is a controlled study of communication. It is the body of a literature, a bibliography quantitatively or numerically or statistically analysed – a statistical bibliography; a bibliography in which measurements are used to document and explain the regularity of communication phenomenon.” Bibliometric studies are generally categorized in two classes, descriptive bibliometrics and evaluative bibliometrics. Descriptive studies concern geographic distribution or temporal distribution of productivity count. Evaluative bibliometrics includes references or citations that are known as literature usage count. According to Nicholas and Ritchie[20], “bibliometrics provides information about the structure of knowledge and how it is communicated”. They clearly distinguished between two types of bibliometrics, i.e. “those describing the characteristics or features of the literature (descriptive) and those examining relationships formed between components of a literature (behavioural). Evaluative studies use the references to literature used by research workers in a field. The scope of bibliometrics includes studying the relationship within a literature or describing a literature, focusing on consistent patterns involving authors, monographs, journals or subject.” Fairthorne[21] in 1969 defined bibliometrics as “the quantitative treatment of the properties of recorded discourse and behaviour pertaining to it”.

According to Rao[9], the concept of bibliometrics in East Europe was known as scientometrics in early seventies. Hood and Wilson[22] mentioned that in 1969, Vassily V. Nalimov & Z. M. Mulchenko[23] coined the Russian equivalent of the term ‘scientometrics’ (‘naukometriya’)[24]. The name of this term is self-explanatory and mainly used for the study of all aspects of the literature of science and technology. This term had achieved extensive acknowledgment by the foundation in 1978 of the journal *Scientometrics* by Tibor Braun in Hungary. According to its subtitle, *Scientometrics* includes all quantitative aspects of the science of science, communication in science, and science policy[25]. The objectives and scope of *scientometrics* domain were also crystallized in some early papers by Nalimov which helped to nurture the then nascent discipline[26,27]. The scope of scientometrics thus envelops several subjects like quantitative aspects of science, science policy, science administration etc. Actually the subject area scientometrics deals with the quantitative studies of output of all disciplines of science. The inner sense of the term *bibliometrics* signals quantitative studies or statistical analysis of bibliographies and its various uses. The scope of scientometrics is thus wider than the scope of bibliometrics. Rajan and Sen[28] outlined the scope of scientometrics, "Etymologically scientometrics means the study relating to the measurement of science. Science can be measured from a number of points of view like the production of graduates, post-graduates or Ph.Ds of science; the establishment of research institutions, the institutions of study and teaching of science; the deployment of scientific manpower, brain drain; expenditure of R & D; founding of the media of scientific communication, e.g. primary and secondary scientific periodicals; scientific literature and scientific information system, services and products. The metric studies of all these aspects fall within the ambit of scientometrics. The area of scientometrics which deal with scientific information is also covered by informetrics. It is to be noted that a very large share of the literature of informetrics pertain to scientometrics". According to Tague[29], "Scientometrics is the study of quantitative aspects of science as a discipline or economic activity. It is part of the sociology of science and has application to science policymaking. It involves quantitative studies of scientific activities, including among others, publication, and so overlaps bibliometrics to some extent". It is clear from the basic definition that the scope of scientometrics is limited to studies of science, whereas the informetrics studies are spread over all fields of knowledge. According to Lancaster[30], the subject scientometrics has grown from simple data analysis to well defined subject involving applied statistics, modelling, simulation, cluster analysis, study of citation network etc.

The word *information* is inbuilt in the term informetrics and information is just like air, that is here, there and everywhere or ubiquitous. In the context of any subject the concept of information may be conceived as the fundamental building block, because if there is no information, no learning or thinking or knowledge sharing/cultivation process can start. The inbuilt units in the previously discussed three words, i.e. librametrics, bibliometrics and scientometrics are library, bibliography and science respectively. These three words are not as fundamental as information. The first word indicates an institution where information is available; the second one indicates an object comprising list of systematic entries of some entities bearing information and the last one indicates a broad discipline of study and research. An instant tally of these three words with information instantly signals the originality or fundamentalism of the word informetrics compared to either librametrics or bibliometrics or scientometrics. Now, the term informetrics was coined by Otto Nacke[29] of

West Germany in 1979, while the other three terms were coined in the years 1948 (librametrics) and 1969 (bibliometrics and scientometrics). Although *informetrics* is recent than other three terms, but it is logical to accept it contextually and conceptually more fundamental and preceder to libra-, biblio- and sciento- metrics. Informetrics may be conceived as a most generalized term that subsumes scientometrics and more specifically bibliometrics. Rajan and Sen[28] from INSDOC (Now NISCAIR), New Delhi, framed the objectives of informetrics in 1985 to be the provision of reliable data for research and development; for policy-making and planning; and for the measurement of institutions, projects, programmes and activities. Brookes[31] in 1989 pointed out that this definition of informetrics is the widest and deepest of the three metrics terms, i.e. biblio-, sciento- and info- metrics. Also, Rao[9] pointed out that Informetrics is basically used to connote the use and development of variety of measures to study and analyze several properties of information in general and documents in particular. A brief definition was given by Egghe & Rousseau[32] in the subtitle of their book: "Informetrics: Quantitative Methods in Library, Documentation and Information Science." Informetrics covers the empirical studies of literature and documents, as well as theoretical studies of different mathematical laws and properties along with distributions that have been discovered. Tague-Sutcliffe[29] defined this term as, "Informetrics is the study of the quantitative aspects of information in any form, not just records or bibliographies, and in any social group, not just scientists. Thus it looks at the quantitative aspects of informal or spoken communication, as well as recorded, and of information needs and uses of the disadvantaged, not just the intellectual elite. It can incorporate, utilise, and extend the many studies of the measurement of information that lie outside the boundaries of both bibliometrics and scientometrics. … Two phenomena that have not, in the past, been seen as a part of bibliometrics or scientometrics, but fit comfortably within the scope of informetrics are: definition and measurement of information, and types and characteristics of retrieval performance measures." Ingwersen & Christensen[33] provided following definition: "The term informetrics designates a recent extension of the traditional bibliometric analyses also to cover non-scholarly communities in which information is produced, communicated, and used." Wilson[25] (2001) concludes the latest ARIST review with the following definition: "… informetrics is the quantitative study of collections of moderate-sized units of potentially informative text, directed to the scientific understanding of informing processes at the social level." Rajan and Sen[28] put forth some new viewpoints and pointed out the inadequacies of Nacke's definition of informetrics. Nacke defined informetrics as, "the application of mathematical methods to the investigation of information science objects, with the aim of describing and analysing their properties and laws in order to optimize these objects in decision making[34]." Rajan and Sen[28] opined, "This definition is inadequate as the definition considers only the investigation of 'information science objects' and not 'information science activities' and 'information itself'. According to the definition, informetrics aims to describe and analyse properties and laws of information science objects. It needs to be pointed out here that informetrics not only describes and analyses the properties and laws of information science objects, but also establishes laws employing mathematical and statistical methods relating to the growth, propagation, use and decay of information. Morales[35] defined informetrics as 'the metric discipline concerned with the study of mathematical and statistical methods and models and their application to the quantitative analysis of their structure and properties of scientific information and the patterns and laws of scientific communication processes including identification of the laws proper'." Bonitz[36] found out the reasons behind the advent of this term. He compared 'informetrics' with 'bibliometrics' and 'scientometrics' and noticed that the introduction of this new term was necessary to differentiate the focussed areas of informetrics, i.e. science communication, from the science of science and library science.

In 1984, the All-Union Institute for Scientific and Technical Information (VINITI) established a Fédération Internationale de la Documentation (FID) Committee on Informetrics under O. Nacke's chairmanship, where 'informetrics' was considered as a generic term for both bibliometrics and scientometrics. This usage was adopted in the VINITI monograph by *Gorkova* in 1988 with the Russian title *Informetriya* [Informetrics]. At the First International Conference on *Bibliometrics and Theoretical Aspects of Information Retrieval* in 1988, *Brookes* suggested that an 'informetrics' which subsumes both bibliometrics and scientometrics, for the documentary and electronic information as well, may have a future[22]. *Informetrics 87/88* was adopted as the short title for the published conference proceedings. The editors noted that "in promoting a new name, it is a classical technique to use the new name together with the old one"[37]. At the second conference in 1990, Brookes[38] endorsed 'informetrics' as a generic term for both scientometrics and bibliometrics, with scientometrics taken as leaning to policy studies and bibliometrics conceded more to library studies. The term 'informetrics' gained more importance in the *Third International Conference on Informetrics*[39]. In the fourth conference entitled *International Conference on Bibliometrics, Informetrics, and Scientometrics* all the three terms were equally emphasized. The proceedings of the fourth conference were published in four separate volumes, three of which were whole issues of regular journals in English[40,41,42]. This conference is historically very important, since at this conference, the International Society for Scientometrics and Informetrics (ISSI) was founded, and subsequent conferences[43,44,45] have been held biennially under the society's auspices. The term 'informetrics' was thus widely recognised by the early 1990s[25].

One of the stalwarts in the context of classical metrics was an American scientist, Eugene Garfield, who was born in 1925 in New York City[46]. He received a PhD in Structural Linguistics from the University of Pennsylvania in 1961 and established the Institute for Scientific Information (ISI) in Philadelphia at Pennsylvania, USA. ISI now forms a major part of the science division of Thomson Reuters Company. Garfield is responsible for many innovative bibliographic products, including *Current Contents*, the *Science Citation Index* (SCI), and other citation databases, the *Journal Citation Reports*, and *Index Chemicus*. He is the founding editor and publisher of *The Scientist*, a news magazine for life scientists. In 2007, he launched HistCite, a bibliometric analysis and visualization software package. Following ideas inspired by Vannevar Bush's famous 1945 article "As We May Think", Garfield undertook the development of a comprehensive citation index showing the propagation of scientific thinking; he started the Institute for Scientific Information in 1955. The *Science Citation Index* made it possible to calculate impact factor, immediacy index, cited half life, citing half life and many other indicators to measure the importance of scientific journals. It led to the unexpected discovery that a few journals like *Nature* and *Science* were core for all of hard science. The same pattern however, does not occur in case of humanities and social sciences.

Cybermetrics: the neo-classical metric

The world of scholarly communication was facing some radical changes since the beginning of 1990s. Paper-based printed communication was gradually sidelined to leave track for then newcomer electronic-media based online (offline also) communication. The last decade of the 20th century may be reckoned as the opening gateway to internet and communication revolution, which added new dimensions to information and knowledge society. Cybernetics is a frequently used term for internet and online communication, which was coined by Norbert Wiener[46] in 1948 in the context of "the scientific study of control and communication in the animal and the machine." The concept of internet was far ahead at that time. Actually the word cyber is a commonly used general prefix term and its first use as prefix was observed in the word cybernetics. Sen[47] remarked that "Weiner derived the word from Greek that means steersman wherein the idea of control is embedded. Thereafter, William Gibson used the word in a story published in Omni magazine. He once again used the word in his novel *Neuromancer* published in 1984. In the novel, the word carried the sense of electronic space. Over the years the meaning of the word has evolved and now embraces information technology, internet and virtual reality". It is well-known that applications of mathematical and statistical methods are heart and soul of the subjects like bibliometrics, scientometrics, informetrics etc. The same is true for cybermetrics also. This particular field is closely related to bibliometrics, informetrics and scientometrics. In cybermetrics the websites play the same role as the documents in bibliometrics, informetrics or scientometrics. The advent of internet and World Wide Web in 1990s aroused plenty of concepts like cyberspace, cyber security, cyber laws, cyber crime etc. Printed documents were gradually becoming extinct to leave space for electronic counterparts. The bibliometric or informetric studies are basically traditional printed document-based, whereas cybermetrics studies are web-based. Sen[47] described the scope of cybermetrics as follows, "Every now and then new websites are being created in cyberspace. Some of them are dynamic undergoing changes quite often, some stable changing very little or not at all, and a few are vanishing sometimes without any prior notice. Many are enjoying long life and some short. As different periodicals on the same subject differ in quality, in the same way different web sites on the same topic differ in quality. Mechanisms have evolved and are also evolving to rank the websites, to calculate Web Impact Factors and to study cited sites. These entire phenomena bring websites within the jurisdiction of cybermetrics".

Webometrics, the synonymous term to cybermetrics, is also frequently used in the context of measuring links, structure and usage patterns of World Wide Web. Björneborn and Ingwersen[48] defined webometrics as "the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web drawing on bibliometric and informetric approaches." The term *webometrics* was first coined by Almind and Ingwersen[49] in 1997. Another definition was given by Thelwall[50], i.e. "the study of web-based content with primarily quantitative methods for social science research goals using techniques that are not specific to one field of study". This definition emphasizes the development of applied methods for use in the wider social contexts. This definition extended the scope of webometrics to various other areas of social sciences beyond information and library science. One important measure is the "Web Impact Factor" (WIF) introduced by Ingwersen[51] in 1998. It may be defined as the number of web pages in a web site receiving links from other web sites, divided by the number of web pages published in the site that are accessible to the crawler. Various other indicators to measure web impact have also been developed today.

Citation analysis

According to Garfield and Rubin, citation analysis may be defined as the thorough investigation of the frequency, patterns and graphs of citations in articles and books.[52,53] It uses citations in scholarly works to establish links to other research works. Citation analysis is one of the most widely used methods of bibliometrics. For example, bibliographic coupling and co-citation are association measures based on citation analysis. Automated citation indexing[54] has added new dimension to citation analysis research, allowing millions of citations to be analyzed for large-scale patterns and knowledge discovery. The first example of automated citation indexing was CiteSeer, followed by Google Scholar. Today citation analysis tools are easily available to compute various impact measures for scholars based on data from citation indices. These have various applications, from the identification of expert referees to review papers and grant proposals, to providing transparent data in support of academic merit review, tenure, and promotion decisions.

The first recorded citation analysis study so far observed based on counting and analyzing citations was due to Gross and Gross[55] reported in 1927. They counted and analyzed the citations appended to articles in a chemistry journal and by ranking the journal titles according to the number of citations received, they produced a list of journals they considered 'indispensable in chemical education'. This is the first recorded study of citation analysis. It is to be noted that the studies by Cole and Eales were based on entries in bibliographies, but not on citations. Citation analysis gradually became an inseparable area of bibliometrics with applications in many fields of knowledge.

The phenomenon of scattering

The bibliographic components like authors or sources (journals, books, monographs etc.) are always scattered for any subject domain over large number of items. For instance, the articles are scattered over journals, the authors are scattered over journals or articles etc. This phenomenon is known as bibliographic scattering. One of the major objectives of classical metric studies is to measure such scattering phenomenon. The sets of data generally handled in classical metric studies are found to conform rather closely to a number of laws and mathematical distributions. The three most popular such laws are Bradford's Law, Lotka's Law and Zipf's Law. These laws and their various mathematical forms have been reviewed by several authors including Fairthorne[56], Brookes[57], Leimkuhler[58], Bookstein[59] and Price[60]. Some authors expressed these laws in readily usable forms in case of practical situations, while some others investigated the similarities of these laws to standard statistical distributions. Some scientists viewed these three laws as special cases of a basic unique distribution. In particular, Price[60] proposed a cumulative advantage distribution which models statistically the situation in which success breeds success. Price opined that this distribution is shown to be “an appropriate underlying probabilistic theory for the Bradford Law, the Lotka Law, the Pareto Law, the Zipf law and for all the empirical results of citation frequency analysis”.

Modern metrics

Wikimetrics: another milestone

The term Wikimetrics is an amalgamation of wiki and metrics. Till now, the Wikipedia is perhaps the one of the most exciting leader project of 21st Century. Jimmy Wales and Larry Sanger launched Wikipedia on January 15, 2001 and the name given to it was a portmanteau of ***wiki*** (the name of a type of collaborative website, from the Hawaiian word for "quick") and *encyclo**pedia***. Just after thirteen years of beginning, i.e. in February 2014, *The New York Times* reported that Wikipedia is ranked fifth globally among all websites stating, "With 18 billion page views and nearly 500 million unique visitors a month..., Wikipedia trails just Yahoo, Facebook, Microsoft and Google, the largest with 1.2 billion unique visitors[61]." Wikimetrics facilitates data analysis of Wiki pages, establishing standardized metrics across the movement and improved workflow between data stakeholders. Wikimetrics is a webtool (formerly known as UserMetrics) designed to simplify the measurement of on-site user activity based on a set of standardized metrics. Using this tool, different metrics can be selected and applied to an arbitrarily defined combination of users to measure their overall productivity, retention, quantity and quality of wiki work. The platform is language- and project-agnostic (it can retrieve data from any Wikimedia project), extensible (adding new metrics, modifying metric parameters) and designed to make data collection for various types of cohort analysis and program evaluation more user-friendly[62].

Open source metrics

A software metric is a measure of some property of a piece of software or its specifications. Since quantitative measurements are essential in all sciences, there is a continuous effort by computer science practitioners and theoreticians to bring similar approaches to software development. The goal is obtaining objective, reproducible and quantifiable measurements, which may have numerous valuable applications in schedule and budget planning, cost estimation, quality assurance testing, software debugging, software performance optimization, and optimal personnel task assignments[63]. An important software metric function was undertaken by the FLOSSMetrics project, where FLOSSMetrics stands for Free/Libre Open Source Software Metrics. The main objective of FLOSSMETRICS is to construct, publish and analyse a large scale database with information and metrics about libre software development coming from several thousands of software projects, using existing methodologies, and tools already developed. The word *libre* is a loanword in English language from French, which means 'state of being free' or 'as in having freedom' or liberty. FLOSSMetrics is providing access to dumps of this database (along with charts, tables and other quantitative information about FLOSS development projects) in the Melquiades website. FLOSSMetrics is also providing the FLOSS Guide for SMEs (small and medium enterprises), which explains the benefits of FLOSS form SMEs, and how to take advantage of it[64].

FLOSSMetrics main targets may be summarized as:

- To identify and evaluate sources of data and develop a comprehensive database structure, built upon the results of CALIBRE
- To integrate already available tools to extract and process such data into a complete platform
- To build and maintain an updated empirical database applying extraction tools to thousands of open source projects
- To develop visualisation methods and analytical studies, especially relating to benchmarking, identification of best practices, measuring and predicting success and failure of projects, productivity measurement, simulation and cost/effort estimation
- To disseminate the results, including data, methods and software
- To provide for exploitation of the results by producing an exploitation plan, validated with the project participants from industry especially from an SME perspective

Journal metrics

The foremost journal metric *Impact Factor* was introduced by Eugene Garfield in 1975 for those journals indexed in Journal Citation Reports. It is highly dependent on the academic discipline, possibly on the speed with which papers get cited in a field. The percentage of total citations occurring in the first two years after publication widely varies among disciplines from 1–3% in the mathematical and physical sciences to 5–8% in the biological sciences[65]. Thus it is logically feasible that impact factors cannot be used to compare journals across disciplines. The impact factor is based on the arithmetic mean number of citations per paper, but it is commonly observed that citation counts follow a Bradford distribution or power law distribution. The arithmetic mean thus is a statistically inappropriate measure. For example, about 90% of *Nature*'s 2004 impact factor was based on only a quarter of its publications, and thus the importance of any one publication will be different from the overall number[66]. Also, the strength of the relationship between impact factors of journals and the citation rates of the papers therein has been steadily decreasing since articles began to be available digitally[67]. Today, there are various other metrics to measure quality or standard of journals as listed below:

1. **Immediacy Index**: It is defined as the ratio of number of articles cited in one year to the number of articles published that year.
2. **Cited Half Life**: it is defined as the median age of articles cited.
3. **Aggregate Impact Factor**: It is defined as the impact factor for an entire subject.
4. **Eigenfactor Score**[68,69]: The Eigenfactor score is a rating of the total importance of a scientific journal. It was developed by Jevin West and Carl Bergstrom at the University of Washington in 2008. Journals are rated according to the number of incoming citations, with citations from highly ranked journals weighted to make a larger contribution to the eigenfactor than those from poorly ranked journals. As a measure of importance, the Eigenfactor score scales with the total impact of a journal. The journals generating higher impact to the field have larger Eigenfactor scores.It may be defined as the number of times articles published in a journal over 5 years are cited in a year. Here Citations to the same journal are removed. Bergstrom and West[68,69] interpreted this indicator as a modified 5-year Impact Factor also. For a JCR year, the Eigenfactor algorithm effectively ranks journals according to citations and the length of time that researchers are logged on to a journal's website. It is, in effect, a journal website citation research.
5. **Article Influence Score**: The Article Influence Score calculates measures the relative importance of the journal on a per-article basis. It is the journal's Eigenfactor Score divided by the fraction of articles published by the journal. That fraction is normalized so that the sum total of articles from all journals is 1. The mean Article Influence Score is 1.00. A score greater than 1.00 indicates that each article in the journal has above-average influence. A score less than 1.00 indicates that each article in the journal has below-average influence. For a JCR year, Article Influence of an ISI journal is defined as "Eigenfactor score divided by the fraction of all ISI articles published by the ISI journal." In 2006, the top journal by Article Influence score was *Annual Reviews of Immunology*, with an article influence of 27.454. This means that

the average article in that journal has twenty seven times the influence of the mean journal in the JCR[70].

6. **h5-index**: It is defined as the largest number h such that at least h articles in that publication were cited at least h times. A journal with h5-index of 8 has, over 5 years, at least 8 articles that were cited 8 times.
7. **SJR indicator– SCImago Journal Rank**: It is defined as number of citations received by number of articles, weighted depending on the prestige and subject area of the journal. It measures scientific influence of scholarly journals that accounts for both the number of citations received by a journal and the importance or prestige of the journals where such citations come from.
8. **SNIP- Source-Normalized Impact per Paper**: SNIP measures a source's contextual citation impact by weighting citations based on the total number of citations in a subject field. It helps to make a direct comparison of sources in different subject fields. SNIP takes into account characteristics of the source's subject field, which is the set of documents citing that source. SNIP especially considers:
    - the frequency at which authors cite other papers in their reference lists
    - the speed at which citation impact matures
    - the extent to which the database used in the assessment covers the field's literature

SNIP is the ratio of a source's average citation count per paper and the citation potential of its subject field. The citation potential of a source's subject field is the average number of references per document citing that source. It represents the likelihood of being cited for documents in a particular field. A source in a field with a high citation potential tends to have a high impact per paper. Citation potential is important because it accounts for the fact that typical citation counts vary widely between research disciplines. For example, they tend to be higher in life sciences than in mathematics or social sciences. If papers in one subject field contain an average of 40 cited references while those in another contain an average of 10, then the former field has a citation potential that is 4 times higher than that of the latter. Citation potential also varies between subject fields within a discipline. For instance, basic journals tend to show higher citation potentials than applied or clinical journals, and journals covering emerging topics tend to have higher citation potentials than periodicals in well established areas[71].

Author metrics

Author metrics essentially measures research impact of scholarly publications by respective authors. All authors have limitations. The metric should emphasize the exact parameter to be measured. As different subject areas or disciplines have different types of publishing patterns, therefore research impact of authors belonging to all disciplines cannot be measured on equal standard. According to Kaur[72], author impact analysis is increasingly playing crucial role in grant evaluation, hiring and tenure decisions. The central objective of the metrics are however to assesses scholarly visibility and social visibility of the authors. The major author metrics parameters are enumerated below[73]:

1. **Citations**: Number of times cited in the literature
2. **Usage**: Number of times viewed on a website (publishers); Number of times downloaded; How often the supplemental data has been accessed
3. **Captures**: How often it has been bookmarked/shared (CiteULike/ Mendeley)
4. **Mentions**: Number of times blogged about; How many news stories; Mentions in Wikipedia etc.; Comments on publishers website & elsewhere
5. **Social Media**: Facebook shares/likes; LinkedIn shares; Tweets
6. **h-index**: A scientist has index h if h of his/her N_p papers have at least h citations each, and the other (N_p-h) papers
   have no more than h citations each. It was introduced in 2005 by J E Hirsch.
7. **g-index**: Given a set of articles ranked in decreasing order of the number of citations that they received, the g-index is the (unique) largest number such that the top g articles received (together) at least g^2 citations. It was introduced in 2006 by Leo Egghe.
8. **i-10 index**: The i10-index indicates the number of academic publications an author has written that have at least ten citations from others. It was introduced in July 2011 by Google as part of their work on Google Scholar.
9. **h-core**: The h-core of a publication is a set of top cited h articles from the publication. These are the articles that the h-index is based on.
10. **h-median**: The h-median of a publication is the median of the citation counts in its h-core.

11. **h5-index, h5-core** and **h5-median**: these indicators of a publication are, respectively, the h-index, h-core and h-median of only those of its articles that were published in the last five complete calendar years.

Web of Knowledge encounters following author-evaluation parameters:

1. Results found
2. Sum of the times cited
3. Sum of times cited without self-citations
4. Citing articles
5. Citing articles without self citations
6. Average citations per item
7. h-index

Article-level metrics

Article-level metrics are metrics for measuring the usage and impact of individual research articles. Traditionally the usage and impact of research were evaluated on the basis of bibliometrics or informetrics, which was normally focused on journals, such as the impact factor or immediacy index. Some researcher-level metrics such as the h-index, g-index or i-10 index were developed in the last decade. Article-level metrics, unlike journal or author metrics did not focus on journals or authors but on the individual article. This is related to, but distinct from, altmetrics. The open access publisher PLOS provides article level metrics for all of its journals including downloads, citations and altmetrics.

Altmetrics

Altmetrics are new metrics proposed as an alternative to the widely used journal impact factor and personal citation indices like the h-index, g-index or i-10 index. The term altmetrics was coined by Jason Priem[74,75] in 2010, as a generalization of article level metrics, and rooted in the twitter #altmetrics hashtag. Although altmetrics are often thought of as metrics about articles, they can be applied to people, journals, books, data sets, presentations, videos, source code repositories, web pages, etc. Altmetrics does not cover just citation counts, but also other aspects of the impact of a work, such as how many data and knowledge bases refer to it, article views, download, or mentions in social media and news media. Altmetrics are a very broad group of metrics, capturing various parts of impact a paper or work can have. A classification of altmetrics was proposed by ImpactStory in September 2012, and a very similar classification is used by the Public Library of Science (PLOS)[76,77] as listed below:

- **Viewed** - HTML views and PDF downloads
- **Discussed** - journal comments, science blogs, Wikipedia, Twitter, Facebook and other social media
- **Saved** - Mendeley, CiteULike and other social bookmarks
- **Cited** - citations in the scholarly literature, tracked by Web of Science, Scopus, CrossRef and others
- **Recommended** - for example used by F1000Prime

Inference

In this paper a brief historical account of all metrics in the context of library and information science has been outlined. The discussion was started from librametrics, the foremost metrics in library science and continued upto the contemporary one, i.e. altmetrics that is a topic of active research today. All metrics are categorized in three classes on the basis of respective time of inception, i.e. classical metrics, neo-classical metrics and modern metrics. The four metrics, i.e. librametrics, bibliometrics, scientometrics and informetrics are categorized under classical metrics. Cybermetrics is regarded as neo-classical metrics while remaining others evolved in 21^{st} century are recognized as modern metrics. Continuous conceptual progress and necessary amendments in the ideas like altmetrics, article-level metrics, sociometrics etc. are going on to achieve new dimensions in the areas of performance measurement and impact analysis of research and scholarly communication.

References

1. *Metric*. (2014). Retrieved June 03, 2014, from Wikipedia: http://en.wiktionary.org/wiki/metric.
2. Metrics. (2014). In *Merriam-Webster Online: Dictionary and Thesaurus*. Retrieved June 3, 2014 from http://www.merriam-webster.com/dictionary/metric.
3. Wittig, G. R. (1978). Statistical Bibliography - a historical footnote. *Journal of Documentation*, *34(3)*, 240-241.
4. Cole, F. J. & Eales, N. B. (1917). The history of comparative anatomy Part 1: A statistical analysis of the literature. *Science Progress*, *11*, 578-596.
5. Hulme, E. W. (1923). *Statistical Bibliography in relation to the growth of modern civilization*, London: Grafton.
6. Gosnell, C. F. (1943). *The rate of obsolescence in college library book collection* (Ph.D dissertation). New York University, New York City, USA, 16.
7. Raisig, L. M. (1962). Statistical bibliography in health sciences. *Bulletin Medical Libraries Association, 50, 450*.
8. Friend, J. H. & Guralnik, D. B. (Ed.). (1964). *Webster's New World Dictionary of Americal Language*. New York: The World Pub Co. 144.
9. Rao, I. K. R. (1998). Informetrics: scope, definition, methodology and conceptual questions, *Workshop on Informetrics and Scientometrics*, 16-19 March, Bangalore, organized by Documentation Research and Training Centre, Indian Statistical Institute.

10. Pritchad, A. (1969). Statistical bibliography and bibliometrics. *Journal of Documentation, 25(4)*, 348-349.
11. Wilson, C. S. (1995). *The formation of subject literature collections for bibliometric analysis: the case of the topic of Bradford's Law of Scattering*(Ph.D. dissertation). The University of New South Wales, Sydney, Australia. Retrieved from http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN1999.0056.
12. Fonseca, E. N. D. A. (1973), In Portuguese: *Bibliografia Estatistica e Bibliometria*: *Uma Reivindicacao de Prioridades*. [*Statistical bibliography and bibliometrics: a re-indication of priorities*], *Ciencia da Informacao, 2(1)*, 5–7.
13. Otlet, P. (1934). *Traite de Documentation. Le Livre sur le Livre. Theorie et Pratique.* [*Treatise on documentation. The book on the book: Theory and practice*], Brussels: Van Keerberghen.
14. Sengupta, I. N. (1992). Bibliometrics, informetrics, scientometrics and librametrics: an overview. *Libri*, *42*, 75–98.
15. Campbell, F. B. F. (1896). *The Theory of the National and International Bibliography: with Special Reference to the Introduction of System in the Record of Modern Literature,* London: Library Bureau.
16. British Standard Institution. (1916). *British Standard glossary of documentation terms*. Prepared under the direction of Documentation Standards Committee, 7.
17. Potter, W. G. (1981). Introduction to bibliometrics, *library trends*, *30(2)*, 5-7.
18. Sengupta, I. N. (1985). Bibliometrics, a bird's eye view, *IASLIC Bulletin*, *30*, 167-174.
19. Hertzel, D. L. (1987). Bibliometrics, history of the development of ideas in, In: *Encyclopaedia of Library and Information Science, 42(7)*, 144-219.
20. Nicholas, D & Ritchie, M. (1978). *Literature and bibliometrics*, London: Clive Bingley, 9-11.
21. Fairthorne, R. A. (1969). Empirical hyperbolic distributions (Bradford-Zipf-Mandelbort) for bibliometric description and predication, *Journal of Documentation, 25(4),* 319-343.
22. Hood, W. W & Wilson, C. S. (2001). The literature of bibliometrics, scientometrics and informetrics, *Scientometrics*, *52(2)*, 291-314.
23. Nalimov, V. V. & Mulchenko Z. M. (1969a). *Eshche raz k voprosu o kontseptsii eksponentsia1'nogo rosta*. [*A word to add on the exponential growth concept*.] *Nauchno-Tekhnicheskaya Informatsiya. Seriya 2(8)* , 12–14. [English translation in: *Automatic Documentation and Mathematical Linguistics*. 3 (1969) 37–40.]
24. Nalimov, V. V. & Mulchenko, Z. M. (1969b). *Naukometriya. Izuchenie Razvitiya Nauki kak Informatsionnogo Protsessa.* [*Scientometrics Study of the Development of Science as an Information Process*], Nauka, Moscow, (English translation: 1971. Washington, D.C.: Foreign Technology Division. U.S. Air Force Systems Command, Wright-Patterson AFB, Ohio. (NTIS Report No.AD735-634).
25. Wilson, C. S. (2001), Informetrics. In: M. E. Williams, (Ed.), *Annual Review of Information Science and Technology, Vol.34*, (pp. 3-143). Medford, NJ: Information Today Inc. for the American Society for Information Science.
26. Nalimov, V. V. (1970). Influence of mathematic statistics and cybernetics on the methodology of scientific investigations, *Zavodskaya Laboratoriya, 36(10)*, 1218–1226. [English translation in *Industrial Laboratory, 36(10)*, 1549–1558.]
27. Nalimov, V. V., Kordon, I. V. & Korneeva, A. Y. A. (1971). *Geograficheskoe Raspredelenie Nauchnoi Informatsii*. [*Geographic Distribution of Scientific Information*.] Informatsionnye Materialy. Moscow: an SSSR Nauchnyi Sovet po Kompleksnoi Probleme Kibernetiki. [Informational Papers. Moscow: Soviet Academy of Science, Scientific Council on Cybernetics.] 2, 3–37. [English translation in: V. V. Nalimov, *Faces of Science*. Philadelphia, Institute for Scientific Information, 1981, 237–260 (chapter 11).]
28. Rajan, T. N & Sen, B. K. (1986). An essay on informetrics: a study in growth and development. *Annals of Library Science and Documentation*, *33(1-2)*, 1-12.
29. Tague-Sutcliffe, J. M. (1992). An introduction to informetrics, *Information Processing and Management, 28(1)*, 1-3.
30. Lancaster, F. W. (1991). Bibliometric method in assessing productivity and impact of research. *Sarada Ranganathan Endowment for Library Science*, Bangalore, 52.
31. Brookes, B. C. (1990), Biblio-, sciento-, infor-metrics??? What are we talking about? In: L. Egghe & R. Rousseau (Eds), *Informetrics 89/90. Selection of Papers Submitted for the Second International Conference on Bibliometrics, Scientometrics and Informetrics,* (pp. 31-43). Amsterdam: Elsevier.
32. Egghe, L. & Rousseau, R. (1990). *Introduction to informetrics: quantitative methods in library, documentation and information science,*
Amsterdam: Elsevier Science Pub.

33. Ingwersen, P. F. & Christensen, H. (1997). Data set isolation for bibliometric online analyses of research publications: fundamental methodological issues. *Journal of the American Society for Information Science, 48,* 205–217.
34. Nacke, O. (1979). Informetric: Ein nuer Name fuer eine Disziplin. *Nachrichten fuer Dokumentation*, *30(6)*, 219-226.
35. Morales, M. (1985). Informetrics and its importance. *International Forum of Information and Documentation, 10(2)*, 15-21.
36. Bonitz, M. (1982). Scientometrie, bibliometrie, informetrie. *Zentralblatt für Bibliothekswesen, 96 (2)*, 19-24.
37. Egghe, L. & Rousseau, R. (1988), (Eds) *Informetrics87/88: Select Proceedings of the First International Conference on Bibliometrics and Theoretical Aspects of Information Retrieval*; 1987 August 25-28; Diepenbeek, Belgium. Amsterdam: Elsevier.
38. Egghe, L. & Rousseau, R. (1990), (Eds) *Informetrics89/90: Selection of Papers Submitted for the Second International Conference on Bibliometrics, Scientometrics and Informetrics*; 1989 July 5-7; London, Ontario, Canada. Amsterdam: Elsevier.
39. Rao, I. K. R. (1992). (Ed.) *Informetrics - 91. Selected Papers from the Third International Conference on Informetrics*; 1991 August 9-12; Bangalore, India. Bangalore: Sarada Ranganathan Endowment for Library Science.
40. Glänzel, W. & Kretschmer, H. (1992). (Eds) Selected Papers Presented at the Fourth International Conference on Bibliometrics, Informetrics and Scientometrics; 1993 September 11-15; Berlin, Germany, *Research Evaluation*, *2 (3)*, 121-188.
41. Glänzel, W. & Kretschmer, H. (1994). (Eds) Selected Papers Presented at the Fourth International Conference on Bibliometrics, Informetrics and Scientometrics; 1993 September 11-15; Berlin, Germany. *Scientometrics*, *30(1)*.
42. Glänzel, W. & Kretschmer, H. (1994). (Eds) Selected Papers Presented at the Fourth International Conference on Bibliometrics, Informetrics and Scientometrics; 1993 September 11-15; Berlin, Germany, *Science and Science of Science*, *3(5)*.
43. Koenig, M. E. D. & Bookstein, A. (1995). (Eds) *Fifth Biennial Conference of the International Society for Scientometrics and Informetrics*; 1995 June 7-10; River Forest, IL. Medford, NJ: Learned Information Inc.
44. Peritz, B. C. & Egghe, L. (1997). (Eds) *Sixth Conference of the International Society for Scientometrics and Informetrics*; 1997 June 16-19; Jerusalem, Israel. Jerusalem : The School of Library, Archive and Information Studies of the Hebrew University of Jerusalem.
45. Macías-chapula, C. A. (1999). (Ed.) *Seventh Conference of the International Society for Scientometrics and Informetrics*; 1999 July 5-8; Colima, México. Colima: Universidad de Colima.
46. *Cybernetics*. (2014). Retrieved June 21, 2014, from Wikipedia: http://en.wikipedia.org/wiki/Cybernetics.
47. Sen, B. K. (2004). Cybermetrics-meaning, definition, scope and constituents. *Annals of Library and Information Studies, 51(3)*, 116-120.
48. Björneborn, L. & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology, 55(14)*, 1216–1227.
49. Almind, T. C. & Ingwersen, P. (1997). Informetric analyses on the World Wide Web: Methodological approaches to 'webometrics'. *Journal of Documentation, 53(4),* 404–426.
50. Thelwall, M. (2009). *Introduction to webometrics: quantitative web research for the social sciences*. California: Morgan & Claypool.
51. Ingwersen, P. (1998). The calculation of web impact factors. *Journal of Documentation, 54(2)*, 236–243.
52. Garfield, E. (1983). *Citation indexing - its theory and application in science, technology and humanities.* Philadelphia: ISI Press.
53. Rubin, R. (2010). *Foundations of library and information science* (3rd ed.). New York: Neal-Schuman Pub.
54. Giles, C. L., Bollacker, K. D. & Lawrence, S. (1998). CiteSeer: an automatic citation indexing system. *Digital libraries 98: the Third ACM Conference on Digital Libraries, June 23–26, 1998, Pittsburgh, PA.* (New York: Association for Computing Machinery), 89–98.
55. Gross, P. L. K. & Gross E. M. (1927). College libraries and chemical education. *Science 66*, 1229-1234.
56. Fairthorne, R. A. (1969). Empirical hyperbolic distributions (Bradford-Zipf-Mandelbrot) for bibliometric description and prediction. *Journal of Documentation, 25*, 319-343.
57. Brookes, B. C. (1969). Bradford's law and the bibliography of science. *Nature, 224*, *5223*, 953-956.
58. Leimkuhler, F. F. (1967). The Bradford distribution. *Journal of Documentation, 23(3)*, 197-207.

59. Bookstein, A. (1976). The bibliometric distributions. *Library Quarterly, 46(4),* 416-423.
60. Price, D. J. S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science, 27(5)*, 292-306.
61. Wikipedia. (2014). Retrieved June 22, 2014 from http://en.wikipedia.org/wiki/Wikipedia.
62. Wikimetrics. (2014). Retrieved June 22, 2014 from http://www.mediawiki.org/wiki/Analytics/Wikimetrics.
63. Software metric. (2014). Retrieved June 22, 2014 from http://en.wikipedia.org/wiki/Software_metric.
64. Flossmetrics. (2014). Retrieved June 23, 2014 from http://www.flossmetrics.org.
65. Nierop, E. V. (2009). Why do statistics journals have low impact factors?. *Statistica Neerlandica 63(1)*, 52–62.
66. Not-so-deep impact. (2005). *Nature 435(7045)*, 1003–1004. doi:10.1038/4351003b. PMID 15973362.
67. Lozano, G. A., Larivière, V. & Gingras, Y. (2012). The weakening relationship between the impact factor and papers' citations in the digital age. *Journal of the American Society for Information Science and Technology, 63(11)*, 2140.
68. Bergstrom, C. T., West, J. D. & Wiseman, M. A. (2008). The eigenfactor™ metrics. *Journal of Neuroscience, 28(45)*, 11433–11434.
69. Bergstrom, C. T. (2007). Eigenfactor: measuring the value and prestige of scholarly journals. *College & Research Libraries News, 68(5)*.
70. Retrieved June 24, 2014 from http://ox.libguides.com/content.php?pid=207971&sid=1733765.
71. Retrieved June 24, 2014 from http://help.scopus.com/Content/h_jrnlsnip.htm.
72. Kaur, J., Radicchi, F. & Menczer, F. (2013). Universality of scholarly impact metrics. *Journal of Informetrics.*
73. Retrieved June 24, 2014 from http://wiki.lib.sun.ac.za/images/3/3d/2013-OA-BBL-presentation_METRICS.pdf.
74. J. Priem (@jasonpriem), I like the term #articlelevelmetrics, but it fails to imply *diversity* of measures. Lately, I'm liking #altmetrics., 4:28 AM - 29 Sep 2010, Tweet.
75. Priem, J., Taraborelli, D., Groth, P. & Neylon, C. (2010). Altmetrics: a manifesto, Retrieved June 25, 2014 from http://altmetrics.org/manifesto.
76. A new framework for altmetrics. ImpactStory Blog. 2012-09-14. Retrieved 2013-08-13. Retrieved June 25, 2014.
77. Lin, J. & Fenner, M. (2013). Altmetrics in evolution: defining and redefining the ontology of article-level metrics. *Information Standards Quarterly, 25(2),* 20.