CHAPTER - 1

Introduction
Chapter - 1

INTRODUCTION

Introduction

Bibliometric is well established discipline for quantitative study of the various aspects of literature of a given subject in almost all the subject disciplines. Bibliometric research developed a body of theoretical knowledge and group a techniques and application based on distributions of data elements.

Librarians have been observing ever-growing number of biographic units like books, periodicals, research materials corresponding increase in the library collection, library readers, library materials and phenomenon like increasing specially in indexing. Changes in search strategy and so on. All these have been observed and recorded for better management and service it is quite natural that when sizable volume of such observation is available attempts would be made to make some generalization and formulate some empirical laws. This have been the course in all sciences and library science cannot be exception. Citation Analysis is one form of Bibliometric study it is an established research tool used by librarians, information scientist to indicate the relationship that exist between cited and citing document. In fact it is the outcome of realization of the rate of growth and fragmentation of scientific literature the rising cost of journals.

Bibliometrics is of recent origin and relatively a new one, which has emerged as a research front in its own right in Information science. It has now being vigorously pursued and with the result, it has been found that one-fourth of all the articles published in Library and Information Science periodicals are on Bibliometrics and its related topics. It has also been found that many of the Social Science and Science periodicals are also carrying a large number of articles on Bibliometrics. Pritchard and Witting\(^1\) compiled a bibliography on the subject comprising 600 entries covering the period 1874 to 1959 which incidentally rose to 2032 entries in 1980 as per compilation of Hjreppe\(^2\). In 1982, Hjreppe\(^3\) again published a supplement to his 1980 bibliography covering 518 items of information. This increase in the number of publications may have been due to the use of Bibliometric techniques in various disciplines and to the greater interest in the theory of literature structure. In view of the enormous literature available in the field-both in theory and applications, an
attempt has been made in this article to provide an overview of Bibliometrics, highlighting the original and important contributions to the field.

1.1 Bibliometrics

Many attempts have been made to define the term bibliometrics and its analogous terms since the use of the term 'statistical bibliography' in 1923 by Hulme. According to him, "the purpose of statistical bibliography is to 'shed light on the process of written communication and of the nature and course of development of a discipline (in so far as this is displayed through written communication), by means of counting and analysis its various facets of written communication."

"Raising" in 1962 defined it as "The assembling and interpretation of statistics relating to books and periodicals to demonstrate historical movements, to determine national and universal research, use of books and journal and to ascertain in many local situation the general use of books and journals." It is regarded as one of the classical definitions of Bibliometrics. As it was pointed out earlier, the term Bibliometrics was first coined by Pritchard in 1969 in preference to existing terminology 'statistical bibliography' as he felt there is fair likelihood to misinterpret it as bibliography of statistics. He defined 'Bibliometrics' as "The application of mathematical methods to books and other media of communication". According to Fairthorne, it is the "Quantitative treatment of properties of recorded discourse and behavior appertaining to it". The British Standard Glossary of documentation of Terms explained Bibliometrics as the study of "the use of documents and patterns of publication in which mathematical and statistical methods have been applied", which is basically similar to Pritchard's original definition. Hawkins in his on-line Bibliometric study interpreted Bibliometrics as "the quantitative analysis of the bibliographic features of a body of literature." Nicholas and Ritchie in their book entitled 'Literature on Bibliometrics' opined that bibliometrics provided information about the structure of knowledge and how it is communicated?. They further added that bibliometric studies fall mainly into two broad groups those describing characteristics or features of a literature (descriptive studies) and those examining the relationship formed between the components of a literature (behavioral studies). More recently Potter, defined bibliometrics as "the study and measurement of the publication patterns of all forms of written communication and their authorship".
Schrader\textsuperscript{12} said it even more simply bibliometrics is "the scientific study of recorded discourse". Broadus presented a historical overview of various definitions of bibliometrics and proposed an alternative definition. According to him, bibliometrics is the quantitative study of physical published units or of bibliographic units or of surrogates of either. More explicitly Sengupta\textsuperscript{13} defines it as the "Organization, classification and quantitative evaluations along with their authorship by mathematical and statistical calculus".

Citation

It is important to make a distinction between the terms reference and citation they are quite often being used as equal even though they denote two different concepts.

Figure 1 reflects the difference between two concepts.

Citation

A received a citation from B

\begin{center}
\begin{tikzpicture}
\node[draw] (A) at (0,0) {A};
\node[draw] (B) at (2,0) {B};
\draw[->] (A) -- node[above] {Reference} (B);
\draw[->] (B) -- node[below] {Time} (A);
\end{tikzpicture}
\end{center}

B makes a reference to A

It is shown that concept reference is used as retrospective term, while the concept citation is a forward looking term.

In other words, a reference is recognition you give another document, while a citation is recognition you receive from another document.

The reference list at the end of a document contains references to other documents to other documents who receive a citation.

Similar variations of the concepts are cited and citing. The figure shows that document A is a cited document, while B is a citing document.
Bibliometrics an Overview

Bibliometrics studies have become very popular these days. The word Bibliometrics is deviled from Latin and Greek words i.e. biblio and metrics so etymologically it stands for the application of mathematics to study of bibliography.

A more elaborate concepts of Bibliometrics has recently been expounded by Eggue\textsuperscript{14} Who defines it as the "development and application of mathematical models and techniques to all aspects of communication."

The origin of the term "Bibliometrics"

The terms bibliometrics and scientometrics have been introduced almost simultaneously by Pritchard and by Nalimov and Mulchenko\textsuperscript{15} in 1969. While Pritchard explained the term bibliometrics as "the application of mathematical and statistical methods to books and other media of communication". Nalimov and Mulchenko\textsuperscript{16} defined scientometrics as "the application of those quantitative methods which are dealing with the analysis of science viewed as an information process". According to these interpretations, scientometrics is restricted to the measurement of science communication, whereas bibliometrics is designed to deal with more general information processes.

The anyhow fuzzy borderline between the two specialties almost vanished during the last three decades, and nowadays both terms are used almost as synonyms. Instead, the field informetrics took the place of the originally broader specialty bibliometrics. The term informetrics was adopted by VINITI and stands for a more general sub-field of information science dealing with mathematical-statistical analysis of communication processes in science. In contrast to the original definition of bibliometrics, informetrics also deals with electronic media and thus includes topics such as the statistical analysis of the (scientific) text and hypertext systems, library circulation, information measures in electronic libraries, models for Information Production Processes and quantitative aspects of information retrieval as well. In his review entitled "Biblio-, sciento-, informetrics Brookes\textsuperscript{17} in 1990 gave an interesting overview about origin and contexts of these metrics of science, literature and information in general. The description given by Glanzel and Schoepflin\textsuperscript{18} in 1994 defines the scope of bibliometric research areas, which is much wider than the usual ones, and thus integrate all presently existing orientations such as applications to
science policy, library science, and information retrieval. According to their approach, bibliometrics and informetrics include "all quantitative aspects and models of science communication, storage, dissemination and retrieval of scientific information".

The definition by Gloria Carrizo-Sainero18 (2000) considers bibliometrics "as the ensemble of methodological knowledge that will serve the application of quantitative techniques in order to evaluate the processes of production, communication and use of scientific information. Its goal is to contribute to the analysis and evaluation of science and research." This gives a clear orientation in direction toward research evaluation that has become the most important application of bibliometric research and technology.

From the above-mentioned general description of the main task of the research field bibliometrics (Scientometrics), the following statement becomes quite obvious. Bibliometrics can be used to develop and provide tools to be applied to research evaluation but is not designed to evaluate research results. Moreover, bibliometrics does not aim at replacing qualitative methods by quantitative approaches and bibliometrics is not designed to override or even to substitute peer reviews or evaluation by experts but qualitative and quantitative methods in science studies should complement each other.

Uses of Bibliometrics

Bibliometrics is used in pure research to map knowledge structures to determine such things as:

- level and nature of collaboration between scientists and disciplines (i.e. between industry and the academy).
- to study technology transfer.
- to map knowledge drift and migration (i.e. vertical, horizontal, and lateral integration and coupling).
- To determine disciplinary encroachment, inter-disciplinarily, and other geographic boundary flows, and for cognitive (i.e. human, machine, and neural network) development.

Bibliometrics is also a primary tool for analysis of scientist / researcher behaviors (citation, decision processes, user needs etc.) It is used for design and evaluation of information Retrieval (IR) systems (i.e. search engines, online databases).
Bibliometrics is Used In Applied Research For

- Thesaurus construction
- Taxonomy and ontology development
- Meta-data development
- Design of metric-based management, product, technology and communication integration systems to relate cause and effect, such as patent citation metrics as a measure of transfer of technology from science to industry.
- policy studies.

Bibliometrics is a decision-support methodology used by LIS and KM professionals in and across virtually every discipline, including computer science, telecommunication, political science, agriculture, biology, neuroscience, education, mathematics, business, industry, religion, and art.

Bibliometrics is used in action research in the areas of

- Collection development
- indexing and abstracting
- design and execution of knowledge and community flow designs.
- evaluation studies for research funding and training programs.

1.1.1 Informetrics

According to Brookes\(^20\), the fourth term 'Informetrics' was first proposed Otto Nacke of West Germany in 1979. An FID Committee with very broadly defined objectives in the provision of research and technical data was subsequently given this name. However, the term was not widely adopted until 1987, when B.C. Brookes at the First Conference of Bibliometrics and Theoretical Aspects of Information Retrieval held ad Diepenbeek Belgium, suggested that the term 'Informetrics' be included in the named of the Second International Conference on the subject. Egghe and Rousseau\(^21\) in the proceedings of the conference state that "the term 'Informetrics' was favored rather that the terms (or beside) the terms 'Bibliometrics' and 'Scientometrics'. Hence, in promoting a new name, it was decided to use the name 'Informetrics' together with the name 'Bibliometrics in the title of the following conferences and in the title of the published conference proceedings also". The second meeting was thus named as 'International Conference on Bibliometrics,
Informetrics and Scientometrics' and the term 'Informetrics' continues to be used in this series of biennial conference.

Brookes\textsuperscript{22} opines that "this new term is being used to cover both sciento-and biblio-metrics impartially. It has produced no distinctively new ideas of its own but as it implicitly covers both documentary and electronic forms of communication, it may have a feature" As a general field of study, it includes the earlier field 'Bibliometrics' and 'Scientometrics'.

Tague-Sutcliffe\textsuperscript{23} defines the term, as "the study of the quantitative aspects of information in any form, not just records or bibliographies, and in any social group, not just scientists. It can incorporate, utilize and extend the many studies of the measurement of information that and lie outside boundaries of both Bibliometrics and Scientometrics and she continues to say that, "Although in practice the scope of 'Informetrics is very broad, two phenomena that have not, in the past, been as part of 'Bibliometrics' and 'Scientometrics', but fit within the scope of 'Informetrics' are:

- Definition and measurement of information, and
- types and characteristics of retrieval performance measures".

The former has been studied in 'Electrical Engineering' and 'Mathematics' science Shannon and Weaver\textsuperscript{24} (1949), and versions of the original measure have, indeed, been developed in many fields. Retrieval performance measures have been studied by information retrieval theorists in 'Information and Computer Sciences' for both its theoretical and practical aspects. Thus, it can be said that the scope of 'Informetrics' is both practical and theoretical. However, the primary emphasis has been on the development of mathematical models, and a secondary emphasis on the derivation of measures for the diverse phenomena studies. The value of a model lies in its ability to summarize, in terms of a few parameters, the characteristics of many data sets the overall shape, concentration, scatter, and the way the data sets change over time. Such models provide a basis for practical decision making.

This historical account confirms that the contemporary approaches in 'Scientometrics / Informetrics' are marked by extensive use of quantitative and qualitative techniques. In recent years, such quantitative methods are being extensively used for studying the structure of literature (of discipline). These
techniques have been helpful in enabling the mapping of disciplines as also a study of any transition in the structure and composition of a discipline. In fact, in the last one decade a number tools techniques and indices for this purpose have been developed and applied to various areas.

1.1.2 Webometrics

In recent years, however the number of electronic activities has increased drastically. In most cases this also means that data are gathered automatically and hence one is inclined to think that it has become easier to collect data. This is not true. It is true that data are gathered in a much faster way but at the same time their accuracy has dropped. One reason can be the fact that these data are delivered by the computer via and third person who might have another idea of what the exact definition of certain attribute is.

In short, Webometrics is Informetrics on World Wide Web. According to Lennart Bjorneborn and Peter Ingwersen\textsuperscript{26} Webometrics consists of four main research Areas –

1) Web page content analysis
2) Web link structure analysis
3) Web usage analysis
4) Web technology analysis

There are definite connections to other Informetrics activities. Web page content analysis, web link structure study has its roots in citation analysis, web usage analysis is part of a more general user and usage research, and web technology analysis refers to information system evaluation.

Webometrics meets its subject on the World Wide Web. But this is only one of the Internet's services. If we include all those services such as e-mail, discussion, groups and chats it is possible to speak about "Cybermetrics". We can define special branches of Webometrics. There are close relations between general descriptive and normothetic informetrics and special applications like scientometrics and webometrics for example, in general informetrics. Co-citation analysis is a way to map the intellectual structure of scientific field. In Webometrics a co-link analysis also leads to the production of a map. But this map does not necessarily re-present
intellectual or cognitive structures. So the application of informetrics methods in special fields of empirical Information Science is not always the same, but sometimes only a procedure by analogy. In early days of Webometrics link between web pages and citations were seen as two sides of the same coin. Web pages are entities of information on the web, with hyperlinks from them acting as citations. Today we have to recognize specific differences between links and citations.

There are different conceptions of Informetrics, bibliometric and Scientometrics. the diagram in Fig. 2 shows the field of Informetrics embracing the overlapping fields of bibliometric and Scientometric following widely adopted definitions by, Brookes (1990), Egghe and Rousseau (1990) and Tague-Sutcliffe (1992). According to Tague-Sutcliffe (1992), 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". Further more, bibliometrics is defined as "the study of the quantitative aspects of the production, dissemination and use of recorded information" and scientometrics as "the study of the quantitative aspects of science as a discipline or economic activity". In the figure, politico-economical aspects of scientometrics are covered by the part of the scientometric ellipse lying outside the bibliometric one.

The figure further shows the field of webometrics entirely encompassed by bibliometrics, because web documents, whether text or multimedia, are recorded information stored on web servers. This recording may be temporary only, just as not all paper documents are properly archived. Webometrics is partially covered by scientometrics, as many scholarly activities today are web-based whilst other such activities are even beyond bibliometrics, i.e. non-recorded, like of cybermetrics as defined above.

In the diagram, the field of Cybermetrics exceeds the boundaries of bibliometrics, because some activities in cyberspace normally are not recorded, but communicated synchronously, like in chat rooms. Cybermetric studies of such activities still fit in the generic field of informetrics as the study of the quantitative aspects of information "in any form" and "in any social group" as stated above by Tague-Sutcliffe (1992).
Fig. 2 Relationships between the LIS fields of info-/biblio-/scien-/cyber-/webo-metrics. Sizes of the overlapping ellipse are made for sake of clarity only.

Naturally, the inclusion of webometrics expands the field of bibliometrics as webometrics inevitably will contribute with further methodological developments of web-specific approaches. As ideas rooted in bibliometrics, scientometrics and informetrics contributed to the emergence of webometrics, ideas in webometrics might now contribute to the development of these embracing fields.

1.2 Scope of the Study

Bibliometric study may be studied from various aspects such as author citation, subject citation, journal citation etc. Various analytical studies have been carried out in recent past, this has developed very important branch of study called "citation analysis”.

Citation analysis is the analysis of citation or bibliographical references that are appended with the research communication. It can be bring out useful information like the authorship pattern in different disciplines, the relative use of different kinds of document such as books, periodicals, reports, theses, patents, standards etc. Reference is the acknowledgement that is document receives from another. A relationship is implied between the cited documents and citing documents, all or part. Thus citations are references given at the end of scholarly papers, to the articles of documents previously published. The very fact that work is cited in paper indicates that there is a relationship between two papers, in other words. If an author
is citing reference in his paper, there is a strong possibility that his own paper is related to it. So, there is necessary to study literature use pattern in Doctoral Research of the North Maharashtra University, Jalgaon, so this topic has been taken as an area for the study. In short the scope of the study is confine to the "Literature used pattern in Doctoral Research at North Maharashtra University, Jalgaon : A bibliometric Study during 1995-2000".

Limitations of the Study
1. The Doctoral Thesis which have been awarded / submitted from 1995-2000.
2. The Doctoral Thesis awarded 2000 to 2002 which have been registered during 1995-2000.
3. The present study is undertaken as a sample case study of 93 theses of Social sciences, Commerce and Science faculties.

1.3 Objectives and Hypotheses of the Study

Many university and college libraries have begun looking at student projects, theses and dissertations as a source of evidence of library research skills. Citation analysis has been used for decades as an analytical tool for examining reference citations in Bibliometric studies.

User satisfaction is the highest objective of the collection development program of any library. For the proper and systematic planning and development of information resources and services the citation studies are most important. There are no University accepted methodologies for measuring cost for the purpose and assessing user effectiveness, bibliometrics is one of the effective methods for measuring cost effectiveness. Its effectiveness has been equally disputed and defended. Citations are particularly appealing because they can be treated as objects that can be described and counted. Moreover, because dissertations and theses references list can be obtained without the participation of the authors, the data to collection relatively easy and has the benefit of note being uncontaminated by participant responses and opinions. However inspite of the fact that citations can be used as a partial indicator of quality of students. So it was felt necessary to study literature use pattern of Ph. D. scholars of Various disciplines of North Maharashtra University, Jalgaon.
1.3.1 Objectives

i) To study of Research Trend in Doctoral Dissertations at North Maharashtra University, Jalgaon.

ii) To find out the most favored media for research scholars conducting in their research.

iii) To find out core literature cited by research scholars of North Maharashtra University, Jalgaon.

iv) To know the institutional journal collection citation from the references based on citation analysis.

1.3.2 Hypotheses

The following hypotheses were formed:

i) Journal citations are (In science Researcher) more than Book citation.

ii) Book citation are (In social science and commerce Researcher) more than Journal citation.

iii) Foreign authors are preferred than Indian authors in Science faculty.

iv) Half-life literature of journal is 5-8 years in science faculty and 10-15 years in social science and commerce faculty.

v) In house (Host Institution) Journal citations are less in number.

1.4 Purpose of the Study

In North Maharashtra University, Jalgaon has been established in last few years. There is no sufficient library staff as per guideline of UGC. Due to lack of grant library has not acquire sufficient books, journals, reference books etc. Therefore it has become very necessary to study of the thesis submitted to North Maharashtra University Library, Jalgaon for right literature, right information to be acquire by library for researcher.

The purpose of present study is to provide the information about quantification of literature used by research scholars. Present study will help to understand the
literature used patterns of the Researcher and to know the following Information through the citation analysis.

1) Ranking of Journals
2) Half-life literature.
3) Geographical Distribution of Literature cited.
4) Authorship Pattern.
5) Most Favored Media of a Literature for Citation.
6) In house Literature Cited.
7) Chronological Distribution.

1.5 Research Methodology

The Present study in concerned with the sources referred to in a Ph. D. Theses submitted to North Maharashtra University, Jalgaon during 1995-2000 and analysis of bibliographical unit which include the footnotes and bibliographical references which appended in the thesis. In this study following methods is adopted to collect the data.

By personal scanning of theses submitted to North Maharashtra University, Jalgaon during 1995-2000 and citations appended to each dissertation under the study have been collected using predefined worksheet, in this process, every citation made at one time has been counted as a one citation. If same reference occurs again using the ibid. considered as a separate reference and countered for the study. All the citation were arranged with decreasing rank of bibliographic forms such as books, journals, conference proceedings, reference sources etc. It helps to understand the literature use pattern by the researcher of North Maharashtra University, Jalgaon. Thus the data collected has been analysis based on the objectives of the study and collected data has been analyzed with the help of computer software MS-ACCESS and presented in tabular form whenever necessary graphs and charts, tables is presented. The analysis has been carried out based on the total citations appended to the thesis. This facilitates comparison of the results, and also enables better
insight into the literature used pattern in different disciplines. The methodology employed for the analysis of citations is described in the following subsections.

a) Source of Data

The sources from references appended to 93 Ph. D. theses registered / accepted by North Maharashtra University, Jalgaon during 1995-2000. In this study the citation appended to each thesis under the study have been collected using pre-defined worksheet. Thus total of 12892 citations are arrived from the 93 theses.

b) Quantitative Analysis

The present study consider the every Bibliographic unit made at one time has been counted as one citation from the theses and this study considers the quantity of references attached to the theses which are compiled, tabulated and analyzed.

c) Core list of Journal

The list of journals most frequently cited by research scholar and journals are arrange in the order of their rank to find out core literature of journal from citations for each subject area. The applicability of Bradford's law of scattering (Graphical Formulation) to citation data was tested by plotting the cumulative number of citations in the Y - axis and cumulative number of Journal in X - axis, where the Journal are cumulative in decreasing order of productivity.

d) Productivity of Journals

The productivity of Journals was measured by grouping all the Journal citation into 4 groups, so that each group contained approximately an equal number of citations. The number of journals contributing to the related citations in each group was computed, the average rate of productivity of Journals in each group by the number of Journals in that group.

e) Collaboration Pattern

The days of individual research are by gone and it will be no longer be possible for the researcher to refrain themselves from collaborating with others. Present study makes a complete count of all authors whose contributions are cited in thesis.
f) Obsolescence of Literature

Present study considers the age wise distribution of citations cited by the researcher, citations are grouped and analyzed on the basis of date of publication of citations.

g) Availability of Journals

The present study also attempts to assess the impact of availability of required literature for the research productivity of science, social science and commerce researchers. The Journal literature cited by the researcher in North Maharashtra University Library. It focuses on both printed and electronic journals available accessible in the library.

1.6 Chapterisation

1. Introduction

In this chapter of Introduction, definition of bibliometric uses of bibliometric applied research, informetrics, scientometrics, objective of the study, scope of the study, hypotheses, research methodology was discussed.

2. Review of Literature

Second chapter reviewed from 1920-2005 relevant literature that supported study from various Indian and foreign periodicals, books, conferences, proceedings, seminars etc.

3. Growth and Development of North Maharashtra University

In this chapter growth and development of North Maharashtra University, Jalgaon and necessary information about bibliometric methodologies such as citation analysis, use of citation analysis, bibliographic coupling, law of bibliometric are given.
4. Literature use pattern in Doctoral Research at North Maharashtra University, Jalgaon: A Bibliometric Study during 1995-2000 An Analysis

In this chapter collected data analyzed with the help of bibliometric methodologies, ranking of journals, obsolescence of journals, and books, Half-life period of books and journals, bradford’s law of scattering, productivity of journal citation classics was discussed.

5. Conclusion

In this chapter, Doctoral theses which are the product of research activity have been examined through citation analysis with view to finding out their effectiveness on the collection development of university library. For this purpose a case study of the Doctoral dissertations of North Maharashtra University Jalgaon from 1995-2000 has undertaken and results have been recorded and validation of hypotheses, future area of research, bibliographic references given in this chapter.

Annexure

List of Journals available in the North Maharashtra University Library

List of Ph. D. Theses
REFERENCES


