Department of Library and Information Science

Bharathidasan University
Tiruchirappalli-620024

Name of the Programme: M.Lib.I.Sc

Course - 4.5: Elective – V INFORMETRICS AND SCIENTOMETRICS

Course Code:P21MLS20A

Unit-I: Informetrics – Evaluation – Concept, Bliometrics, Scientometrics, Webometrics, Altmetrics

Presented

By

Course Teacher: DR.C.RANGANATHAN

Professor

e-mail: cranganathan72@gmail.com

Introduction

- The twentieth century may be described as the century of the development of metric science. such as librametrics, bibliometrics, scientometrics, informetrics webometrics, cybermetrics and altmetrics
- Among these different metrics is the most interesting subject area in the field of library and information science, which can be applied to any discipline irrespective of their period of evolution.
- It involves quantitative studies of scientific activities and useful metrics/technique which helps to solve the problems, challenges posed by so called information explosion.

Cont..,

- There is an increasing interest in science communication and promotion of new inventions in this deluge digital information age.
- Expansion of World Wide Web and newer technologies has improved the ways in which science is communicated and its evaluation techniques.
- The use of conventional metric tools gauges the impact of scholarly publications using citation and download counts.
- They are widely used to evaluate articles, authors and disciplines on publishers' platforms. These tools analyses the citation data in natural course of time and make them available over a period of time.

Why quantitative studies?

- Qualitative methods often depend on assertions. ‘authoritative’ statements, anecdotal evidence
- Science searches for regularities
- Success of statistical methods in social sciences
- Need for justification & basis for decisions
- Something can be counted - irresistible

Reasons for quantitative studies of literature

- Analysis of structure and dynamics
  - search for regularities - predictions possible
- Understanding of patterns
  - “order out of documentary chaos”
  - verification of models, assumptions
- Rationale for policies & design.

Type of Quantitative Studies

Quantitative Science and it is divided into two Basic Categories .

- Descriptive Bibliometrics (Productivity Count)
- Evaluative Bibliometrics (Literature usage Count)

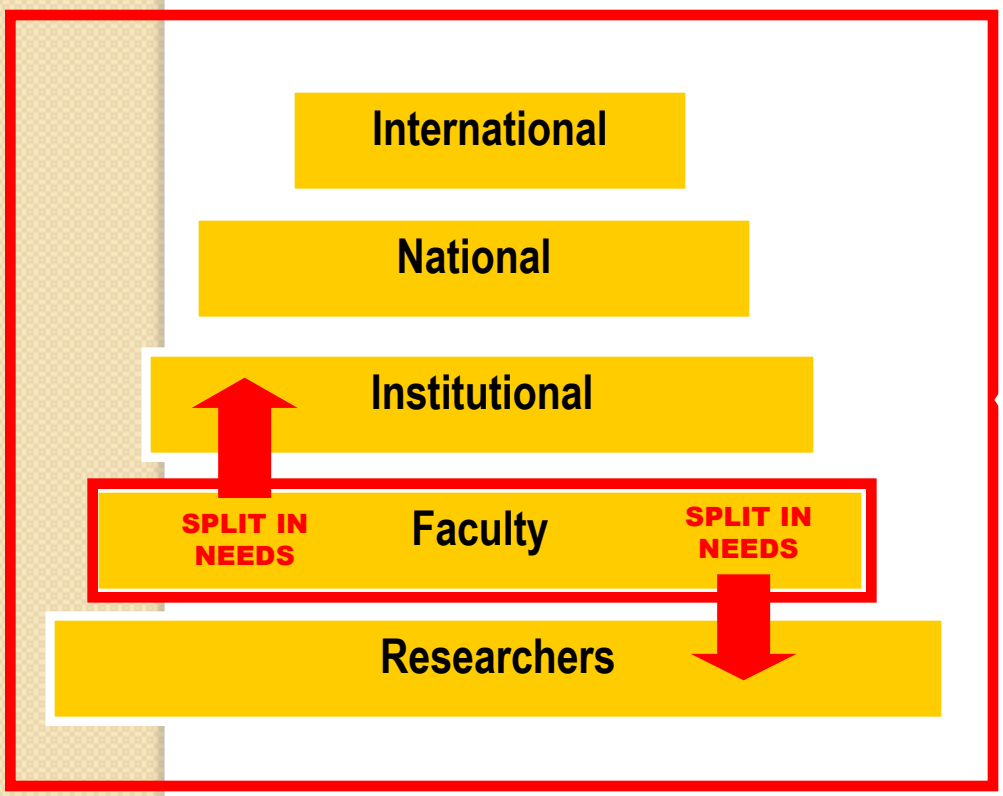
- **Descriptive Bibliometrics**:

  Geographical, Time Period and discipline

  Evaluative Bibliometrics:

  Reference Count and Citation Count

Why do we evaluate scientific output



- Grant Allocations
- Policy Decisions
- Benchmarking
- Promotion
- Collection management
- Funding allocations
- Research

Metric Studies

- The metric studies are used to measures
  - scholarly communication;
  - identify research trends and growth of knowledge;
  - identify users of different subjects;
  - estimate comprehensiveness of secondary periodicals;
  - forecast past, present and future publishing trends;
  - adopt an accurate weeding and stacking policy;
  - develop norms and standardization;
  - identify authorship and its trends in documents on various subjects;
  - predict productivity of publishers, individual authors, organizations, country or that of an entire discipline.

Cont..,

- Research output is evaluated in a number of ways by using various metric tools.
- One of the ways to analyses quality of research is to measure the number of times the work has been cited in literature.
- All these conventional tools need reasonable time to compute and publish the data.
- In this measure of 'quality' of research many invisible factors viz.,
  - prestige of the journal, publishing history,
  - frequency, editorial team, peer reviewing process, accessibility of the published article,
  - impact, number of downloads,
  - hits and many more are involved..

Cont..,

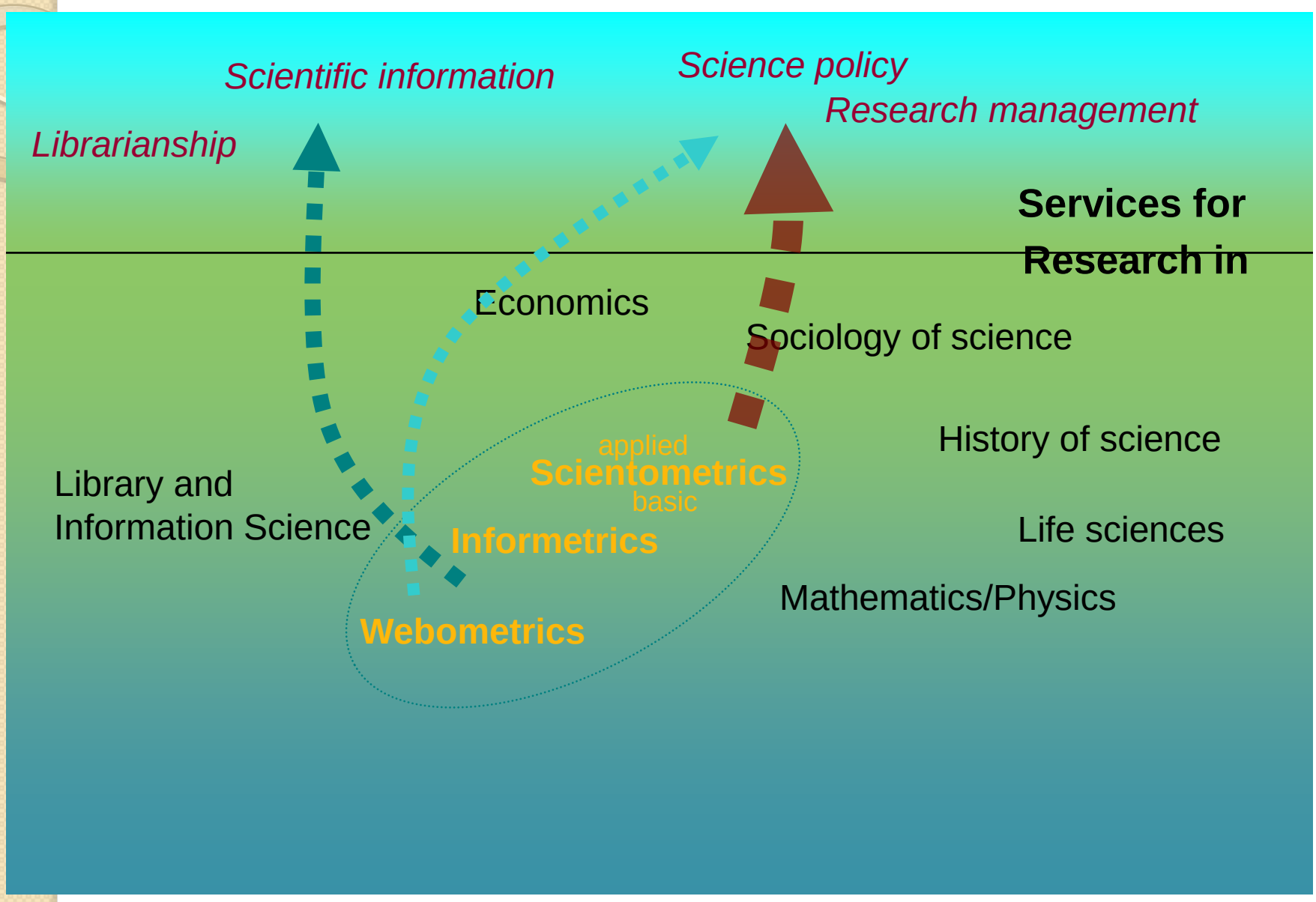
- All metrics are categorized in three classes on the basis of respective time of inception,
  - Classical metrics,
  - Neo-classical metrics and
  - Modern metrics.
- The four metrics, i.e. librametrics, bibliometrics, scientometrics and informetrics are categorized under classical metrics.
- Cybermetrics is regarded as neo-classical metrics
- Webometrics and Altmetrics are evolved in 21st century are recognized as modern metrics.

Development of Librametrics to Altmetrics

The quantitative studies in Library and Information Science at different time period known by various names such as

- Statistical bibliography (1920s)
- Librametry (1940s)
- Bibliometrics (late 1960s)
- Scientometrics (1960s)
- Informetrics (1980s)
- Webometrics (1990s)
- Cybermetrics (2000s)
- Altmetrics (2011s)

Links of bibliometrics with related research fields and application services



Statistical Bibliography

Hulme (1923) has introduced the term “Statistical Bibliography”.

It is defined as “the application of mathematical and statistical methods to books and other means of printed communication”.

This term continued to be used for statistical measures till the end of 1960s.

Librametrics

- Dr. S. R. Ranganathan coined the term librametry and presented his concept in 1948 at the ASLIB conference held at Lemington Spa.
- He used the term to include statistical approaches to the study of library and its services. However, the practice of using quantitative method to measure information sources.
- Ranganatham coined the term “Librametry” to denote the measurement of various library activities and services using mathematical and statistical techniques.
- “ Quantitative analysis of various facets of library activities and library documents by the application of mathematical and statistical calculus to seek solution to library problems.”

Scope in…

- Reader statistics
- Statistics used documents
- Statistics of incoming documents
- Calculation of shelf space of the library
- Annual cost of purchase material
- Circulation statistics
- Inter library loan statistics
- Statistics of footnotes and references

Application in…

- To determine optimal strength of library staff in difficult section and their judicious distribution
- To develop library staff to dispose of references queries during different library hours
- To systematic circulation of library documents
- To develop and organising local ,state ,regional and national library
- To determine the optimal size of a service library
- To design library building along with its fittings and furniture
- To distinguish service and dormitory libraries
- To analyse reader's queries
- To streamline acquisition system by analysing library documents .i.e books, monographs journals, etc.,
- To determine the most helpful sequence for classification of macro and micro documents
- To test the accuracy of cataloguing entries by sampling techniques
- To ensure accurate selection of library documents .

Bibliometrics

What is?

- "… all studies which seek to quantify processes of written communication."-Pritchard
- "… the quantitative treatment of the properties of recorded discourse and behavior pertaining to it."-Fairthorne
- Recorded communication - 'literature'->quantitative methods

Alan Pritchard 1969

- Coined the term "bibliometrics"

"the application of mathematics and statistical methods to books and other media of communication".

"the application of mathematical and statistical methods to books and other printed media, process of written communication and of the nature and course of discipline".

Concepts

Basic (primitive) concepts:

1. Subject
2. Recorded communication -> document, information object
3. Subject literature

- Bibliometrics related to:
  - ✓ science of science
  - ✓ sociology of science - numerical methods

Literature studies

- Qualitative
  often in humanities, librarianship
- Quantitative
  bibliometrics
- Mixed

Scope and Purpose

- The scope of bibliometrics includes:
  - *all quantitative aspects and models of science communication,*
  - *storage,*
  - *dissemination and retrieval of scientific information.*
- Process of written communication counting and analysing like various facts of written communication structure of knowledge and how it is communicated.

Type

- Bibliometrics is Classified as
  - Descriptive
  - Behavioural
  - Citation Studies

Bibliometrics studies are mainly employed to investigate the following areas.

- Scattering of articles
- Author productivity
- Citation analysis
- Measures of journal productivity

Application in ...

- **Scientific output evaluation**
  - Impact
  - Citations
- **History of science**
- **Publication strategies**
- **Sociology of science**
- **Science policy; resource allocation**
- **Collection management**

  Library selection, weeding, policies
- **Information organization**
- **Information management & Utilization**

Cont…,

- To identify research trend at growth of knowledge
- To estimate comprehensiveness of secondary periodicals
- To identify users of different subjects
- To identify authorship and Its trends in documents on various subjects
- To measure the usefulness of the adhoc and retrospective SDI services
- To forecast past, present and future publishing trends
- To develop experimental models correlating or by passing its existing one
- To adopt an accurate weeding stacking policy

Cont..,

- To formulate an accurate need – based acquisition policy with in the limited budgetary provision
- To identify core periodicals in different discipline
- To initiate effective multilevel network system
- To study obsolescence and dispersion of scientific literature (clustering and coupling of scientific papers)
- To predict productivity of publisher individual author organisation , country or that of an entire organisation
- To design automatic language processing for auto indexing ,auto abstracting and auto classification
- To develop harms for standardization

What is Scientrometrics?

- It is the science of measuring the **“quality** **“**of science.
- It is often done using bibliometrics which is a measurement of the impact of scientific publication.
- It includes all quantitative aspects of the science of science, communication in science, and science policy.
- The term ***Scientometrics*** originated as a Russian term for the application of quantitative methods to the history of science.
- This term was introduction and came into prominence with founding of the journal named ***‘Scientrometrics’***
- It deals with analysis, evaluation and graphic representation of science and technology information. It tells “Who is doing what and where?”

Cont..,

- by T . Braunin 1977, originally published in Hungary and currently from Amsterdam.
- Scientrometrics used to mean communication process in science including social- culture aspects and appears to be almost synonymous to science of science.
- Scientrometrics can be treated as an analogue concept to bibliometrics.
- The term Scientrometrics is a field which applies quantitative methods to the study of science as an information process.
- Thus scientrometrics is a part of the sociology of science and has application to science policy making.

Definitions

- **Nalmov and Mulchenko (1969) of USSR defined** as the quantitative methods which deals with analysis of science viewed as an information process.
- According to **Pouris (1989),** 'Scientrometrics is a application of quantitative techniques(system analysis, mathematical and statistical techniques etc.) to scientific communication(science output, science policy, science administration etc.) with the objectives of;

I. Developing science indicators;
II. Measuring the impact of science on society; and
III. Comparing the output as well as the impact of science at national and international levels.

Cont…,

- **Tgue- Sutcliffe pointed** "Scientometrics "as the study of quantitative aspects of science as a discipline or economic activity.
- **Dobrov and Karennnol defined** it as " the measurement of informatics processes".
- **While Mikhilov defined it** " as that scientific discipline that studies the structure and properties of scientific information and the laws of processes of scientific discipline devoted to all quantitative aspects of science and scientific research".
- **Scientometrics (bibliometrics) -** The measurement of scientific output activity through statistics on academic publications

Cont..,

- scientists etc. – from different branches of physics and the natural sciences I,e science managers and policy makers, and administrators of government and non-governmental organizations, among many others.
- Thus Scientometrics involves studies in :

I. Sociology of science
II. History of science
III. Growth of science and scientific institutions
IV. Behaviour of science and scientists.
V. Science policy and decision- making

Scientometrics tools

- The quantitative as well as qualitative analysis of any Scientometrics study, such as
- citation mapping, visualization,
- bibliographic coupling,
- co- authorship network,
- co-words mapping etc.
- This tools are very much useful for Scientrometricians for mapping their parameters in any accept of their study.

Informetrics

Introduction

Areas of applications of statistics to various disciplines have given new names such as Econometrics, Sociometris and Informetrics etc. Informetrics was formed to develop statistical and mathematical method in order to study and analyses the characteristics of all kinds of information.

Example: Text, digital image, videos, spoken documents and music etc.

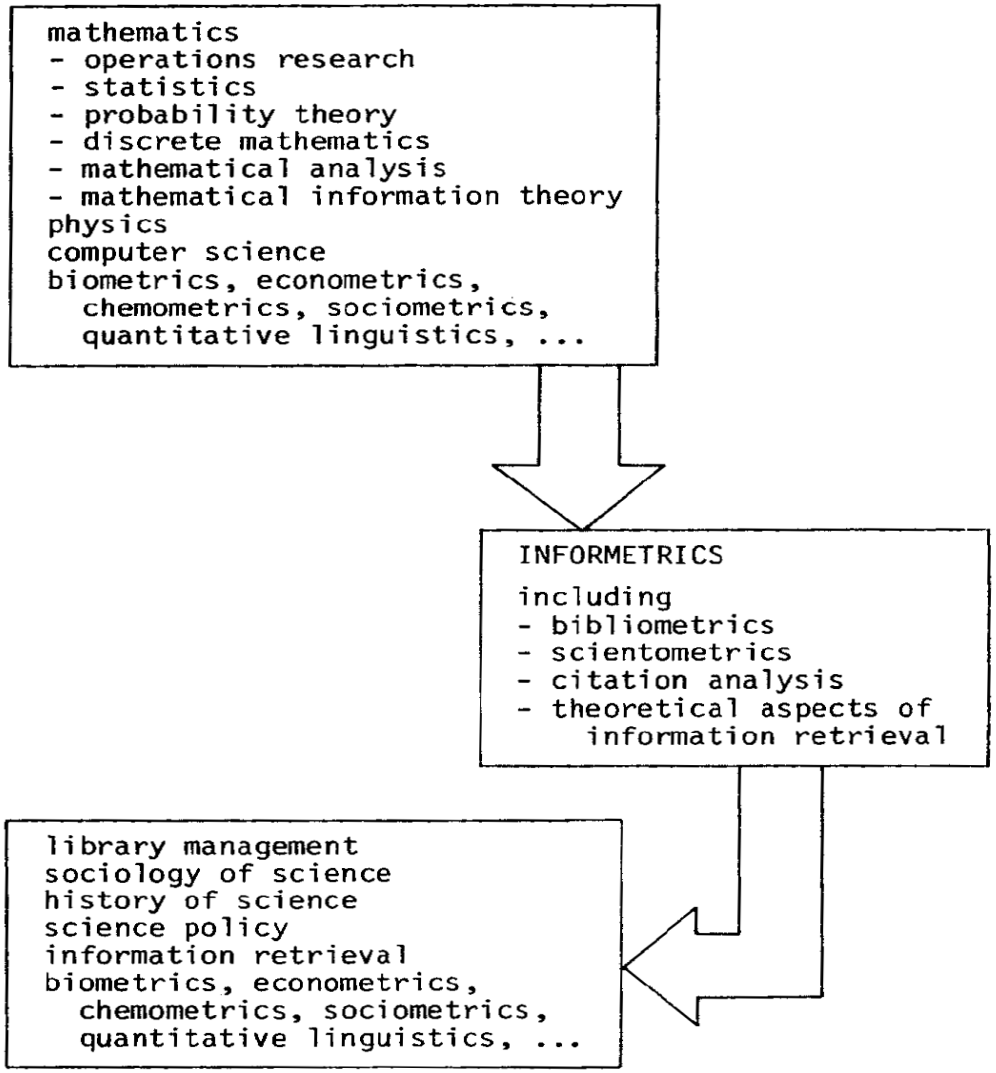
Informetrics:

The terms 'Informetrics' was introduced by Blackert, Siegal and Otta Nacke in 1979 but gained popularity by launch of the International Informetrics Conferences in 1987.

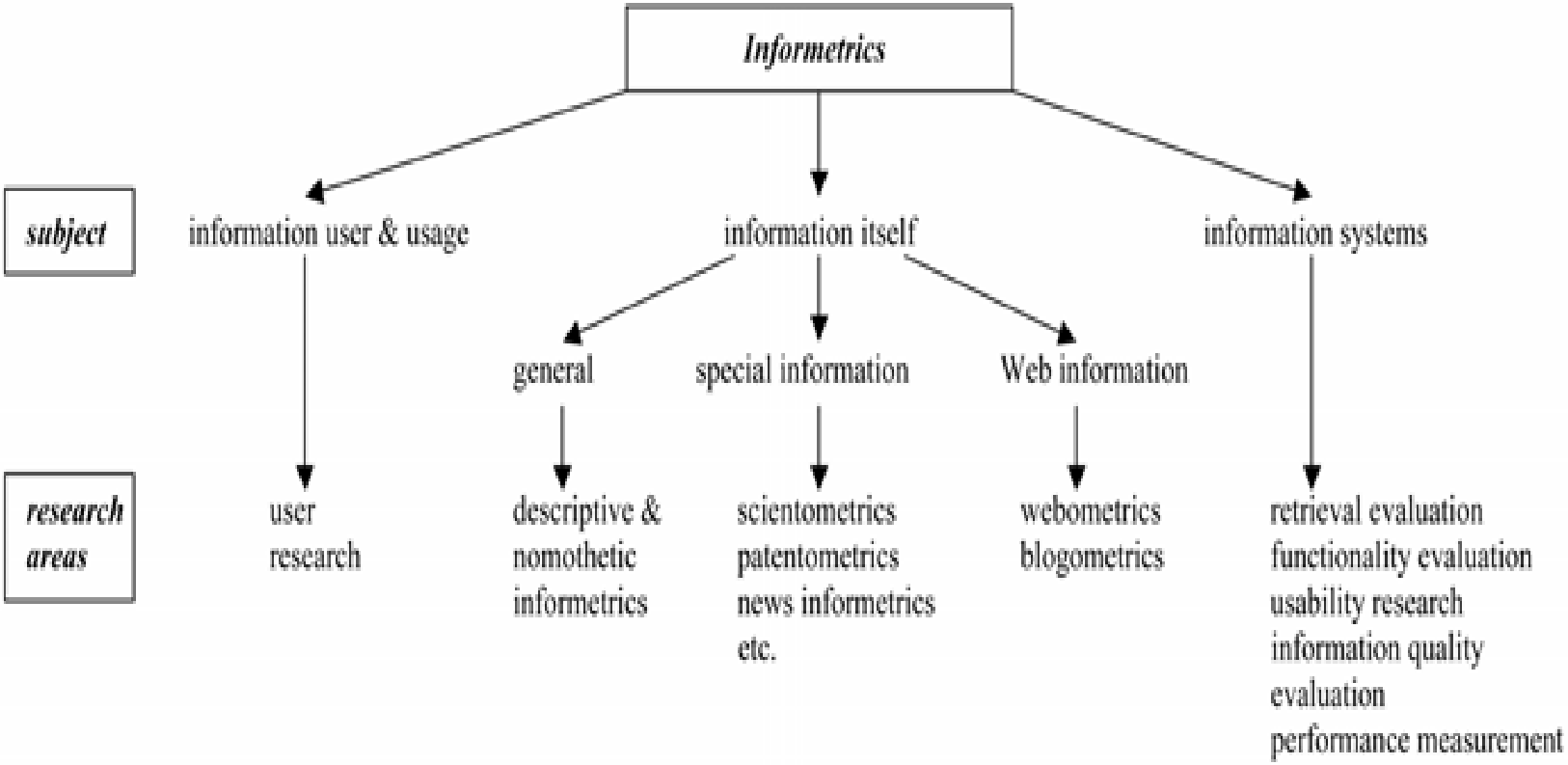
Definition:

According to Tague – Sutclifle, "Informetrics is the study of the quantitative aspects of information in any form, not just records or bibliographies and any social group, not just scientists".

Leo Egghe and Ronald Rousseau have identified a place for Informetrics among other fields



Subject and Research Areas of Informetrics:



1. Information user and Usage: ‘User research’ is a prime area of research in this subject area which mainly deals the topics on

- Human and their information behavior.
- Information seeking behavior on web (especially usage of search engine).
- Usage of library services.

2. Information System: ‘Retrieval evaluation’, ‘Usability research’ and ‘Information quality’ are major research areas in Information system.

- Retrieval evaluation is mainly involved in studying recall and precision of retrieval system.

- Usability research deals with tests of accessibility, accuracy, relevance, completeness, objectivity and timeliness of users.

3. Information Itself: Bibliometrics, Scientometrics and Webometrics are main research area in Information Itself.

Informetrics Distribution or Laws:

Informetrics studies in library and information science are based on the mathematical expression of the three laws such as Lotka's law, Bradford's law and Zipf's law.

- **Lotka's law (1926):**

  Applications:

  1. Collaboration and communication pattern of an Individual, an Institution and a country
  2. Trends in scientific collaboration

- **Bradford's law (1934):**

  Applications:

  1. To study the dispersion of literature mostly in the fields of science, engineering and medicine

- **Zipf's Law (1949):**

  Applications:

  1. Number of authors contributing in a discipline or other field
  2. Distribution of articles in a set of journals
  3. Ranking of word frequency in a set of documents

Ravichandra Rao I.K. stated that the research areas of following topics are called as librametry, or as scientometrics or as bibliometrics or as infometrics:

- Quantitative aspects of library and information science, especially use and user studies.
- Quantitative studies related to book usage, acquisition, age distribution of documents etc.
- Circulation studies.
- Citation studies / analyses (impact factors and other measures)
- Journal productivity (by coverage, by use, by citation, etc.)
- Author productivity.
- Obsolescence and growth studies.
- Quantitative analysis of science (- science indicators, country-wise, language-wise, subject-wise etc.).
- Identifying relations among various disciplines
- Structure of subjects / disciplines
- Evaluation of scientific research (by institutions, by individuals, by countries)

Webometrics

Almind and Ingwerson coined the name Webometrics.

According to Bjorneborn & Ingwerson "Webometrics is defined as the study of quantitative aspects of construction and use of information resources, structures and technologies on the web drawing on bibliometrics and informetrics approaches".

Webometrics is based on two indicators:

- Volume of published materials of institutions/individual on the web, and
- The visibility and impact of the web pages measured by the citation (site citations or links they receive).

There are four areas of webometrics research as follows.

- Web page content analysis.
- Web link structure analysis (e.g. Hyperlink, Self link and External link)
- Web usage analysis (Including log files of users' search & browsing behavior).
- Web terminology analysis (Including search engine performance).

Web Page Content Analysis or Citation Analysis

- The Webpage Content Analysis can be used to analyze content of the websites.
- It provides hits on the systematic organization of web based information sources.
- It enables the users to reduce their time in the choice of right sources.
- It compares the efficiency of search engines in retrieving the required information sources.
- It will be useful for students, researchers, scientists who seek information through www.
- Simplistic counts and content analysis of web pages are like traditional publication analysis.

Web Link Structure Analysis

- Link analysis has been used successfully for deciding which
- The web pages add to the collection of documents (i.e., which pages to *crawl*), and
- How to order the documents matching a user query (i.e., how to *rank* pages).
- It has also been used
  - to categories web pages,
  - to find pages that are related to given pages,
  - to find duplicated websites, and
  - various other problems related to web information retrieval.
- The Web Link Structure Analysis provides ***hyperlinks or selflinks*** between documents and records of user behavior. To be precise, *hypertexts* (i.e., collections of documents connected by hyperlinks).
- It provides counts and analysis of outgoing links from web pages, here named called ***outlinks or external links***.
- It provides links to web pages or links coming from the other websites called ***inlinks or incoming links.***

Web Usage Analysis

The following components cover the web usage analysis

- Log files for users searching.
- Browsing behavior.
- Log analysis for security applications.
- Web usage pre-processing.
- Novel techniques for discovery and analysis of Web usage patterns.
- Integrating semantics and domain knowledge in Web usage mining and analysis.
- Reliability and consistency of Webometrics.
- Integration of click stream data with back-end data and related metrics.
- Intelligent summarization/explanation of changes in Web usage metrics

Web technology analysis: Search engine performance

The search engines performance determines following aspects:

- Measuring the search engines performances.
- Total number of hits retrieved.
- Number of relevant hits retrieved.
- The content of the page like what is the page all about and so on.
- Ranking of search engines.
  - **LexiURL Searcher** is a programme to conduct automatic analysis of the impact of collections of documents or websites, or to create network diagrams of collections of websites. It has the ability to automatically submit queries to search engines and process the results. The four main features of LexiURL Searcher are:
- A Web Impact Report of the number of times each of a set of words, phrases or documents have been mentioned online.
- A Link Impact Report of the number of web pages and websites that link to one or more websites or web pages.
- A Network Diagram of the links between a collection of websites.
- A Web Environment Network of an individual website.

Cybermetrics

- Cybermetrics is a both an Electronic-only Journal and a Virtual items to the study of the quantitative analysis of scholarly and scientific communications in the Internet.
- It provides open to world-wide researchers to publish and discuss their findings. Internet offers them new and increased capabilities to distribute their results to a greater audience.
- Cybermetrics deals with the quantitative measures of the Internet backbone technology, topology and traffic.
- Cybermetrics is the discipline dedicated to the quantitative description of the contents and processes of the communication that take place in the cyberspace

Cont…,

- The Cyber-scientometric is the sub-field more developed, for practical reasons it is named with the more general term of Cybermetrics or the more specific of Webometrics.
- Some of the topics are used in cybermatrics study such as
  - Informetric distributions and topology
  - Dynamics and evolution
  - Link & “sitation” analysis
  - Indicators: R&D evaluation
  - E-journals and documents repository bibliometrics
  - Applied webometrics: web data mining and positioning

Cybermetrics Vis – a – Vis Webometrics

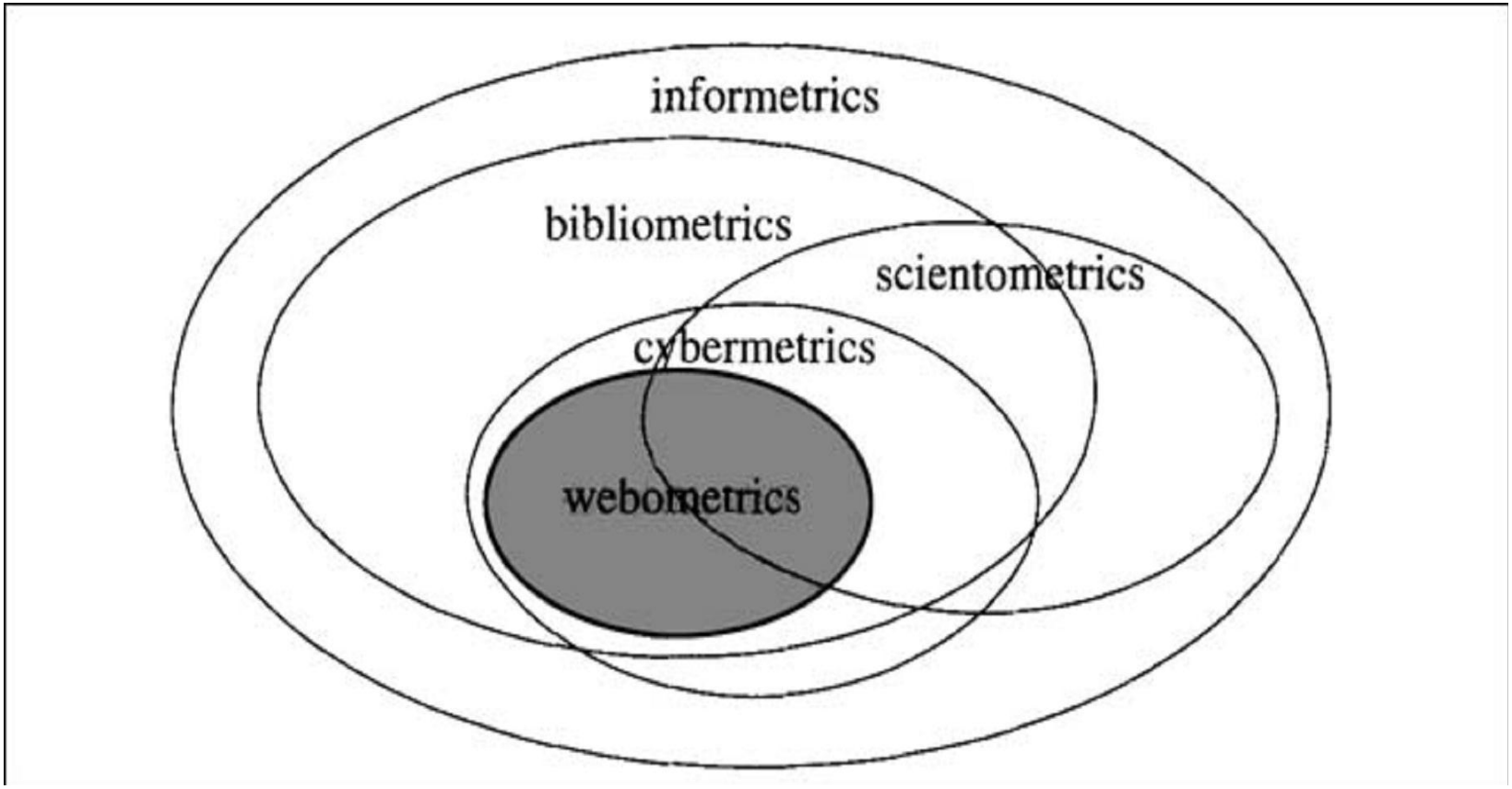
- Cybermetrics has been proposed for analysis of web documents before the term Webometrics.
- Cybermetrics is “the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the whole internet, drawing on bibliometric and informetric approaches”.
- Cybermetrics thus encompasses statistical studies of discussion groups, mailing lists and other computer mediated communication on the internet including the web.
- The breadth of coverage of Cybermetrics is limited and where as webometrics implies large overlaps with proliferating computer science based approaches in analysis of web contents, link structures and web usage and web technologies.

Altmetrics

- Altmetrics is a concept and term coined by Jason Priem, University of North Carolina-Chapel Hill, North Carolina, USA.
- It is a short form for 'Alternative to Metrics' or in other words it is an alternate to conventional metric tools.
- Altmetrics is the creation and study of new metrics based on social web for analyzing and informing scholarship.
- Altmetric tools capture the article level scholarly data that is shared in social media and measures the impact of content in real time basis and the data is presented with visual effects.
- Altmetrics aren't citations, nor are they webometrics. They are relatively unstructured and closed.

Relation between Different Metrics:

Thelwall, Vaughan and Bjorneboren in 2005 attempted to explained relationship between different metrics through the following diagram.



- The field of informetrics embracing the overlapping filed of bibliometrics and scienctometrics.

- Webometrics is seen as entirely encompassed by bibliometrics because web documents in their various forms such as text, multimedia are all recorded information stored on the web server.

- Webometrics is partially covered by scientometrics because many scholarly activities today are web-based.

- Cybermetrics completely covers webometrics but exceeds the boundaries of bibliometrics because some activities in the cyberspace are not normally recorded.

Conclusion

All metrics as a subject has seen tremendous growth in recent times. The availability of information through Internet has made it much more interesting topic and eased the data collection process. Today large data can be analyzed easily with assistance web based technologies. The policy markers, research institution and other major stake holders involved in research and innovations have been very much depended on any metrics for granting funds for research. As developing countries thrusting more on research and innovation, there is need to use any metrics for measuring the research output of an organization/institutions or individuals. Many of the universities know measure the research output of its scientists through scientometric tools such has h-index, g-index and citations for research publications.

References:

1. Almind, T.C. and Ingwersen, P. (1997). Informetric analysison the World Wide Web:methodological approachesto “webometrics”. Journal of Documentation, 53 (4), 404- 426.

2. Bar-Ilan, Judit. and Peritz, Bluma C. (2009.) A method for measuring the evolution of a topic on the web: The case of “informetrics”. *Journal of the American Society for Information Science and Technology, 60*(9), 1730–1740.

3. Bar-Ilan,Judit.(2008). Informetrics at the beginning of the 21stcentury: A review. *Journal of Informetrics, 2*(1), 1-52.

4. Björneborn , Lennart and Ingwersen, Peter. (2004). Toward a basic framework for webometrics.*Journal of the American Society for Information Science and Technology, 55*(14) 1216-1227.

5. Burrell,Quentin L. (2005). Measuring similarity of concentration between different informetric distributions: Two new approaches. *Journal of the American Society for Information Science and Technology, 56*(7), 704–714.

6. Egghe, L. (2005). Expansion of the field of informetrics: Origins and consequences. Information *Processing and Management, 41*(6), 1311-1316.

7. Egghe, Leo and Rousseau, Ronald. (1990). *Introduction to informetrics: Quantitative methods in library, documentation and information science*. Amsterdam: Elsevier Science Publishers.

Cont…

8. Erar, Aydn. (2002). Bibliometrics or informetrics: Displaying regularity in scientific patterns by using statistical distributions. *Hacettepe Journal of Mathematics and Statistics, 31,* 113- 125.

9. Jacobs, Daisy. (2010). Demystification of bibliometrics, scientometrics, Informetrics and webometrics. Paper presented at *11th DIS Annual Conference 2010, 2nd – 3rd September, Richardsbay, University of Zululand, South Africa*.

10. Nisonger, Thomas E. (2001). Report on the 8th international conference on scientometrics and informetrics in Sydney, Australia*. Library Collections, Acquisitions, and Technical Services, 25*(4), 353-494.

11. Rao, I.K. Ravichandra. (1998). Informetrics: Scope, definition, methodology and conceptual questions. Paper presented at *Workshop on Informetrics and Scientometrics 16-19 March 1998, Bangalore.*

12. Sangam, S.L et. al. (Eds.) (2010). *Webometrics, informetrics and scientometrics : Measuring scientific and technological progress of India*. Dharwad: Department of Library and Information Science, Karnataka University.

13. Stock,Wolfgang G. and Weber,Sonja. (2006). Facets of informetrics. *Information, 57*(8), 385-389.

14. Wilson, Concepcion s and Hood, William W. (2001). The literature of bibliometrics, scientometrics,and informetrics. *Scientometrics, 52*(2), 291-314.

15. Wormell, Irene. (1998).Informetrics: an emerging subdiscipline in information science. *Asian Libraries, 7*(10), 257-268.

THANK YOU