WEBSITES OF INDIAN INSTITUTES OF TECHNOLOGY: A WEBOMETRIC STUDY

Dissertation

Submitted to the Department of Library and Information Science, University of Delhi in partial fulfillment of the requirements award of the Degree of

MASTER OF PHILOSOPHY
IN
LIBRARY AND INFORMATION SCIENCE

Submitted by
Mr. Shashi Prakash
Research Scholar

Supervisor
Dr. M. Madhusudhan
Assistant Professor

Department of Library and Information Science
University of Delhi, Delhi, India

December 2012

DECLARATION

I hereby declare that the dissertation entitled "**Websites of Indian Institutes of Technology : A Webometric Study**" is based on the original research work carried out by me for the award of **Master of Philosophy in Library and Information Science**, University of Delhi, Delhi.

Place: Delhi **(Mr. Shashi Prakash)**

Date: - 12 - 2012. Roll No: M.Phil. – 08

CERTIFICATE

This is to certify that the dissertation entitled "**Websites of Indian Institutes of Technology: A Webometric Study**" submitted by **Mr. Shashi Prakash** (Roll No.-08) in partial fulfillment of the requirements of the **Master of Philosophy in Library and Information Science** examination of the University of Delhi, Delhi is his own work carried out under my guidance and is worthy of examination.

Place: Delhi **(Dr. M. MADHUSUDHAN)**
Date: - 12 - 2012.

ACKNOWLEDGEMENTS

I express my deep sense of gratitude to my supervisor **Dr. M. Madhusudhan**, Assistant Professor, Department of Library and Information Science, University of Delhi, Delhi for his invaluable guidance, constant support, patience and encouragement without which this Dissertation would not have seen the light of the day. His down to earth approach inspired and helped me to finish this work comfortably. It was an honor to work under his supervision.

I would like to express my sincere gratitude and appreciation to Head of the Department **Dr. (Mrs.) P.K. Walia,** without her initiative, encouragement and skilled suggestions both theoretical and practical; it would have been difficult to make this study a success.

I also thank all my teachers **Prof. (Mrs.) S. P. Singh, Dr. Shailendra Kumar, Dr. R.K. Bhatt, Dr. K.P. Singh, Dr. (Mrs.) Meera, and Mr. Manish Kumar** for their timely suggestions.

I extend my thanks to all my friends and classmates from whom I have learned more during the study. My thanks also goes to the library staff of study libraries and library website development team for helping me in searching the Information.

I would make mistake if I forget to acknowledge the cooperation and good wishes of my parents and other family members. My precious thanks to the almighty who is watching my every movement right from the start till the end from heaven.

Lastly, I thank all who helped me directly or indirectly in completion of this dissertation.

Shashi Prakash

PREFACE

The Dissertation entitled "**Websites of Indian Institutes of Technology: A Webometric Study"** for the award of degree of Master of Philosophy in Library and Information Science under the supervision of **Dr. M. Madhusudhan**. The primary objective of the study is to examine critically the effectiveness and efficiency of the use of web impact factor and to find out the link patterns i.e. Top Level Domain (TLD), Second Level Domain (SLD) and Webpage Second Level Domain (WSLD) and Web Indicators for Science, Technology and Innovation Research (WISER) ranking applications among the Indian Institutes of Technology under study.

The present Dissertation contains **Five** (5) chapters:

Chapter 1: Introduction - discusses the background study, statement of the problem under study, objectives, hypotheses, scope, methodology used for the study, study period and limitations of the present study.

Chapter 2: Review of Literature – provides a literature review of webometric and websites of Indian Institutes of India on the basis of studies conducted in India as well as in abroad is organized on the basis of importance and relevance of the study. The review of related literature is presented on different aspects of the webometric evaluation of IIT Websites falls into three main areas: (i) Webometrics, (ii) Sociology of academic web spaces, and (iii) Webometrics in academic web spaces.

Chapter 3: Profiles of Indian Institutes of Technology – presents brief stock of all the existing institutions of national importance Indian Institutes of Technology with their landmarks and development in world class educational platform that is dynamically sustained through quality teaching and internationally acclaimed research with excellent infrastructure and the best available minds.

Chapter 4: Data Analysis and Interpretation - analyzes and interprets the data collected for ranking of the sixteen Indian Institutes of Technology in India study with the help of Web Impact Factor (WIF) and Web Indicators for Science, Technology and Innovation Research (WISER).

Chapter 5: Findings, Suggestions and Conclusion - concludes the dissertation with findings, suggestions and conclusion for the effective webometric study and practices and explores the scope for further research. This chapter immediately followed by a Bibliography and appendix.

CONTENTS

LIST OF TABLES USED

LIST OF FIGURES USED

LIST OF ABBREVIATIONS USED

ADM	**-**	Alternative Document Model
BENCO	-	Banaras Engineering College
BHU	**-**	Banaras Hindu University
EAT	-	Earth Air Tunnel
HCEs	**-**	Hit Count Estimates
ICT	-	Information and Communication Technology
IIMs	**-**	Indian Institutes of Management
IIT	-	Indian Institute of Technology
IITs	-	Indian Institutes of Technology
ISI	-	Institute of Scientific Information
IT	-	Information Technology
IT-BHU	**-**	Institute of Technology, Banaras Hindu University
JIF	-	Journal Impact Factor
KIAP	**-**	Kanpur Indo-American Programme
LIS	**-**	Library and Information science
MHRD	**-**	Ministry of Human Resource Development
MLA	**-**	Modern Language Association
MoU	-	Memorandum of Understanding
NER	-	North-East Region
PWD	**-**	Public Works Departments
R&D	-	Research and development
RAE	-	Research Assessment Exercise
SASMIRA	**-**	Synthetic and Art Silk Mills Research Association

SLD	-	Second Level Domain
SME	**-**	School of Military Engineering
THES	-	The Times Higher Education Supplement
TIE	-	Theatre-In-Education
TLD	-	Top Level Domain
TMS	-	Thermal Mass Storage
UNCTAD	-	United Nations Conference on Trade and Development
UNESCO	-	United Nations Educational, Scientific and Cultural Organization
URL	-	Uniform Resource Locator
VRV	-	Variable Refrigerant Volume System
WIF	-	Web Impact Factor
WISER	-	Web Indicators for Science, Technology and Innovation Research
WSLD	-	Webpage Second Level Domain
WTO	-	World Trade Organization
WWW	-	World Wide Web

Chapter-1

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

In the face of growing demand for education and fundamental changes in higher education in most countries, the increasing value gets the creative coordination of institutions of higher education on the basis of a constant exchange of experience between them. “The easiest and the effective way to operational exchange of information is publication on the website. Enhancing a web policy expands the dialogue between the universities, contributes to the formation of new communications in the scientific community, and helps the realization of innovative development. This would shed some light on the use of key communication medium and could lead to more effective academic use of the Web. The World Wide Web (WWW) has now become one of the main sources of information on academic and research activities, and therefore is an excellent platform to test new methods of evaluating webometric activities” (Babu, Jeyshankar and Rao, 2010).

Web resources are apple of information professional’s eye due to its value added services to meet their current and diversified information needs. In the WWW, the web pages are the entities of information, with hyperlinks from them acting as citations. Quantitative analysis on the WWW is being carried out in the same way, as is tradition in citation databases. As information on web increases towards entropy, it’s needed to apply some theory/ metrics (measurement) to develop new methods, modeling techniques and metaphors to examine this emerging complex network. Through webometric study one can observe that how users actually react and use specific web document. The Web is in out of control in growth, which means opportunities exist where good system architecture and diligent analysis can be applied for everyone's benefit. On the basis of the study and conception the definition of webometric is, "The webometric study is based on quantitative measurement – indirectly includes the qualitative aspect also of structure, use of information resources and technologies on WWW drawing on bibliometric and informetric approach"(Goswami, 2007).

The other metric sciences – infometrics, scientometrics, cybermetrics and webometrics are also statistical methods and having their origin on the basis of bibliometrics. But all of them are having their different subject area: (i) *Informetrics* - the study of quantitative aspect of information in any form; (ii) *Bibliometrics* - the study of quantitative aspect of production, dissemination and use of recorded information; (iii) *Cybermetrics* - the study of quantitative aspect of Internet as a whole; (iv) *Scientometrics* - the study of quantitative aspect of science as a discipline or economic activity; and (v) *Webometrics* - the study of quantitative aspect of web/web site. Figure 1.1 depicts the relationship between five metrics.

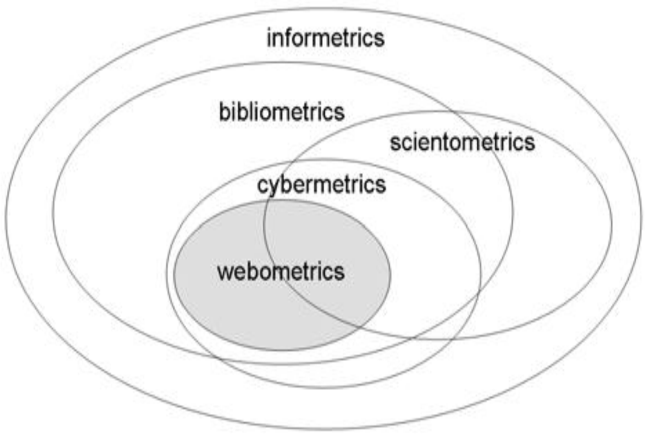

Fig. 1.1: Relationship of different Metrics
(Source: Goswami , Sharma and Shukla,2008: 650)

Figure 1.1 depicts the circle of Informetrics covers all other metrics circles, because according to stub (given above), it is a quantitative aspect of any type of information. This part, which overlaps the circle of bibliometrics, of scientometrics, shows the politico-economical aspects of scientometrics. The economic aspect of science shows the impact of scientific research over the society.

The circle of webometrics in Figure 1.1 overlaps the circle of bibliometrics, but within the boundaries of cybermetrics. Webometrics circle can't overlap the circle of cybermetrics because web is a part of cyberspace. But in the diagram the circle of webometrics ellipse lying outside the bibliometrics, because some aspect of webometrics (link structure, technologies and so on), dose not included in bibliometrics or it is beyond the boundaries of bibliometrics.

1.2 STATEMENT OF THE PROBLEM

The WWW has now become one of the main sources of information on academic and research activities, and therefore it is an excellent platform to test new methods of evaluating webometric activities. However the world scientific community has not yet accepted the Web as a full supplement or a complement to traditional scientific publishing. The increase in the use of the web for research has led to the evolution of web bibliometric, popularly referred to as webometric. Webometric analysis show nature, structure, content features of websites as well as links structure to understand virtual highways and their interrelations (Bjorneborn and Ingwersen, 2001).

In the wake of Internet/Web developments, some bibliometricians drew analogies between Web-based and research documents and came up with the idea that the scientific content of the Web could be analysed in the same way as the science journal system. In the web environment, visibility of the websites and introducing their owners to their users are so important that indicated websites reliability. Websites, through links made to each other, allow users to move from one site to another in the minimum time and access to needed information. Relationship between websites is reinforced via links made to each other which results in more visibility of websites having more in-links. Also, this leads to make more and more links in order to meet user's information needs and access to useful information as well. As a result of increasing information resources available on the web, determining which information resource is more reliable and which one is less important and retrieving these resources are difficult. In spite of these problems, webometric methods can help users to improve them.

The present research aims to study the impact and visibility rate by webometric method of the websites of IITs both quantitatively and qualitatively demonstrates their impact rate and visibility with the help of Web Impact Factor (WIF) of in-link, Web Indicators for Science, Technology and Innovation Research (WISER) and compare their WISER, WIF (in-link) and World rank which can be measured through determining number of web pages, total number of link pages, self-link pages, in-link pages or external link pages and WISER Index Value.

1.3 OBJECTIVES OF THE STUDY

The primary objective of the study is to examine critically the effectiveness and efficiency of the use of web impact factor and to find out the link patterns among the websites of IITs under study. The other inter-linked objectives are, as follows:

(i) to identify and analyze links of websites of Indian Institutes of Technology;

(ii) to investigate relevance of Web Impact Factor (WIF) with reference to Top Level Domain (TLD), Second Level Domain (SLD) and Webpage Second Level Domain (WSLD).

(iii) to calculate the simple Web Impact Factor (WIF), self-link WIF and in-link or external WIF;

(iv) to compare various ranking approaches among websites of IITs;

(v) to compute the correlation between ranking of WISER and WIF(inlink), and

(vi) to rank the IIT websites under study as per WIF, WISER index value, and world rank.

In order to realize the above objectives, the study area is explored with the following Hypotheses, Scope (of the study), Methodology, Reference period, and Limitations of the study.

1.4 HYPOTHESES OF THE STUDY

With the above assumptions the following hypotheses are formulated for verification under the study:

(i) The domain structures of websites of IITs in India are homogeneous;

(ii) India as a whole is having strong value of WIF as per measurement of web presence, and

(iii) Reliability of ranking between WIF (inlinks) and world ranking for IITs are same.

1.5 SCOPE OF THE STUDY

The present study makes a webometric analysis of IITs websites in India. The study examined the websites of 16 IITs in the country and aimed at to establish a kind of academic ranking of these websites by measuring their web impact factor and WISER indicative value with the help of Google, AltaVista and Yahoo search engines. The ranking of websites will help the reader to compare and identify IITs websites in India according to their WIF.Table 1.5 presents the study websites of IITs in India.

Table -1.5: Websites of Indian Institutes of Technology in India

Domain name of hosts of IITs	URL of the Library
Indian Institute of Technology (BHU) Varanasi	http://www.itbhu.ac.in
Indian Institute of Technology Bhubaneshwar	http://www.iitbbs.ac.in
Indian Institute of Technology Bombay	http://www.iitb.ac.in
Indian Institute of Technology Delhi	http://www.iitd.ac.in
Indian Institute of Technology Gandhinagar	http://www.iitgn.ac.in
Indian Institute of Technology Guwahati	http://www.iitg.ac.in
Indian Institute of Technology Hyderabad	http://www.iith.ac.in
Indian Institute of Technology Indore	http://www.iiti.ac.in
Indian Institute of Technology Kanpur	http://www.iitk.ac.in
Indian Institute of Technology Kharagpur	http://www.iitkgp.ac.in
Indian Institute of Technology Madras	http://www.iitm.ac.in
Indian Institute of Technology Mandi	http://www.iitmandi.ac.in
Indian Institute of Technology Patna	http://www.iitp.ac.in
Indian Institute of Technology Rajasthan	http://www.iitj.ac.in
Indian Institute of Technology Roorkee	http://www.iitr.ac.in
Indian Institute of Technology Ropar	http://www.iitrpr.ac.in

In order to determine the scope of Webometrics, it is necessary to know the relationship between Informetrics, Scientometrics, Bibliometrics and Webometrics. It may be noted that Webometrics is associated with Bibliometrics and overlaps Scientometrics to an extent (Bjornborn and Ingwersen, 2004). Informetrics covers both Scientometrics and Bibliometrics (Brookes, 1990), as shown in the above figure 1.1.

1.6 REFERENCE PERIOD

Reference period of the webometric study from October to November 2012, during this period all sixteen websites of IITs in India have been evaluated and observed with the help of Google, AltaVista and Yahoo search engines.

1.7 METHODOLOGY

The methodology of the study is a **Composite one** i.e. it is a combination of more than one research methods, because, the study deals with three important aspects of the problem. The methods of data collection used for the study are observation, evaluation checklist and webometric analysis. However, these three methods adopted serve as compliments to one another, because of the characteristics of the entities (web facilities, web utilization and web presence). The following sections will explain the methodology in detail:

1.7.1 Designing Evaluation Checklist

A checklist with four segments was designed and used to collect the data from study IIT websites on:

(i) Authority and Accuracy;

(ii) Purpose and Content;

(iii) Currency, and

(iv) Design, Organisation and ease of use.

The checklist enabled the researcher to directly seek the information of study websites on the factors relating to their utilization of the web.

1.7.2 Observation

Observation was used as a supplement method to evaluate the websites of IITs and their domain names adopted by the IITs in terms of websites' comprehensiveness, regularity of up-dating them, and their domain names.

1.7.3 Webometric Analysis

Webometrics was triggered by the realisation that the web is an enormous document repository with many of these documents being academic-related (http://webometrics.wlv.ac.uk/). Moreover, the web has its own citation indexes in the form of commercial search engines, and so it is ready for researchers to exploit. In fact, several major search engines can also deliver their results automatically to investigators' computer programs, allowing large-scale investigations. Webometrics includes link analysis, web citation analysis, search engine evaluation and purely descriptive studies of the web. Webometrics, a modern, fast-growing offshoot of bibliometrics and webometric analysis of IIT websites were carried out using Yahoo, AltaVista, and Google search engines. The IIT name was used as search term using the advanced search interface in order to have tighter control on term coordination. Irrespective of the fact that general search engines have not been found comprehensive for this kind of study, it is assumed that "Google should be able to present a result that would enable the researcher present a glimpse of how the University is represented in the web" (Bar-Ilan, 2005 and Baeza-Yates, Castillo and Lopez, 2006).

1.7.4 Calculation of Web Impact Factors

WIF is the web versions of impact factor. There are three types of WIFs: WIF (simple), WIF (self-link) and WIF (in-link). The various types of WIF calculations of IIT websites have been presented in Chapter - 4, Section 4.5.

1.7.5 WISER Ranking

The first web indicator, WIF was proposed by Almind and Ingwersen (1997) based on link analysis that combines the number of external or inlinks and the number of pages of the websites, *a ratio of 1:1 between visibility and size*. This ratio was modified later on by adding two new indicators to the size components, i.e., *number of documents in rich file formats (formats that are used for scholar communications) and number of publications being collected by Google Scholar database*. This new measure is called Web Indicators for Science, Technology and Innovation Research (WISER) and ranking of IITs are ordered according to higher to lower value of WISER that represents WISER ranking. WISER ranking of IITs based on WISER indicator presented in Cahpter-4, Table 4.8.1T1.

1.7.6 Selection of Search Engines

Google, AltaVista and Yahoo have been chosen to collect the required webometric data of the study IIT websites. The selection of these three search engines were based on previous studies and relevant to find the desired values searched against all the domain names and URLs to check whether these domain / sites or not.

1.7.7 Query Syntax

General search engines like Yahoo! and Google are indispensable to cope with the flood of information on the Internet. But the quality of search results is not always good enough due to the vast size of the search space that these engines cover. The Webometric analysis is based on the data collected from the Web using various search engines. In each search engine there are some specific search keywords assigned to retrieve the required information from the Web. To integrate these search sites we need to conceal the difference of the query syntax from the user. But the query syntax of these search sites vary. These specific search keywords along with search syntax are mentioned in Chapter 4, Section 4.2.

1.7.8 Data Collection

AltaVista, Yahoo! and Google had been chosen to collect the data for this study. Data collection was done during 25 October, 2012 to 25 November, 2012. All the domain names ware verified check whether Yahoo!, Google and AltaVista support the domain name or not. For each of these domains a search was carried out to determine the total number of links, total webpages, self-links and inlinks using the commands that presented in Chapter 4, Section 4.3.

1.7.9 Style of Bibliographic References

The references have been given according to Modern Language Association of America (MLA) Handbook for writers of research papers, 7th edition with the help of an easybib web tool <www.easybib.com> for creation of references and all bibliographical entries including review of literature is presented according to alphabetical order by author name in the Bibliography.

1.8 LIMITATIONS OF THE STUDY

The study was focused to analyse the webpages and their links that are accessible from study the Websites of IITs. The advanced search facility of Yahoo!, AltaVista and Google was used to collect the primary data. Data collection was conducted to retrieve the links as well as the web presence of IITs.

The main limitation of the study was collection of primary data with the help of three search engines. The other limitations of data collection on the Web are: Limited number of search keywords (e.g. domain, link) available at present to support webometrics studies. Besides, all the available search engines do not support the entire webometrics search keywords mentioned above. Most of the commercial search engines do not support webometric research keywords and also do not have advanced search features. Some query syntax in two different search engines yield two different data. Therefore, it is very difficult to decide or choose one set of data over the other for further calculation in webometrics research.

The other limitation of the study is to establish suitable reasons for web situation and categorization of situations of newly established seven IIT websites i.e. IIT Bhubaneshwar, IIT Gandhinagar, IIT Indore, IIT Mandi, IIT Patna, IIT Rajasthan, and IIT Ropar and that are having less number of webpages therefore; they may not qualify for comparative webometric studies, especially for ranking purposes, due to underdeveloped websites.

Chapter-2

REVIEW OF LITERATURE

This chapter gives an overview of review of related articles dealing with various aspects of webometrics and websites have been conducted in India as well as in abroad is organized on the basis of importance and relevance of the study. However, an attempt is also made to highlight some of the recent developments related to webometrics analysis, particularly with reference to websites. The review of related literature reveals the existence of a substantial body of literature about evaluation of websites.

A search has been conducted with combination of various key terms such as "webometrics", "websites", "IIT" and " evaluation of university websites" in Library and Information Science Abstracts (LISA), Emerald database, Ebsco database, Science Direct and other databases, to complete review of literature for the present study, including search on the e-journals websites and search engines. In addition to above searches, bibliographies, journal article are also reviewed for more sources, as well as websites of IITs and consulted eminent experts in the field of webometrics.

The review of related literature is presented on different aspects of the evaluation of websites with the help of webometrics falls into three main areas: (i) Webometrics, and (ii) Sociology of academic web spaces, (iii) Webometrics in academic web spaces. The reviewed articles under each heading are presented in the following sections:

2.1 WEBOMETRICS

Library and information science and related fields in the sociology of science and science and technology studies have developed a range of theories and methodologies, now including webometrics – concerning quantitative aspects of how different types of information are generated, organized, distributed and utilized by different users in different contexts. Historically, this development arose during the first half of the twentieth century from statistical studies of bibliographies and scientific journals (Hertzel, 1987). These early studies revealed bibliometric power laws like Lotka's law on productivity distribution among scientists (Lotka, 1926); Bradford's law on the scattering of literature on a particular topic over different journals (Bradford, 1934); and Zipf's law of word frequencies in texts (Zipf, 1949). Similar power-law distributions have been identified on the Web, e.g., the distribution of TLDs (top level domains) on a given topic (Rousseau, 1997) or inlinks per web site (Albert, Jeong & Barabási, 1999; Adamic & Huberman, 2000; 2001).

Decisive for the development of bibliometrics and scientometrics was the arrival of citation indexes of scientific literature introduced by Garfield (1955) that enabled analyses of citation networks in science (e.g., Price, 1965). Access to online citation databases catalyzed a wide range of citation studies, especially mapping scientific domains, including growth, diffusion, specialization, collaboration, impact and obsolescence of literature and concepts (cf. e.g., White & McCain, 1989; Borgman & Furner, 2002).

The breakthrough of online citation analysis parallels the later avalanche of webometric studies enabled by access to large-scale web data. In particular, the apparent yet ambiguous resemblance between citation networks and the hyper-textual inter-document structures of the Web triggered much interest from the mid-1990s (e.g., Bossy, 1995; Moulthrop & Kaplan, 1995; McKiernan, 1996; Kuster, 1996; Larson, 1996; Downie, 1996; Rousseau, 1997; Almind & Ingwersen, 1997; Pitkow & Pirolli, 1997; Spertus, 1997; Ingwersen, 1998). Further, the central bibliometric measures of co-citation (Small, 1973) and bibliographic coupling (Kessler, 1963) have been applied to studies of web clustering, web growth and web searching (e.g., Larson, 1996; Weiss et al., 1996; Pitkow & Pirolli, 1997; Efe et al., 2000; Ding et al., 2002; Menczer, 2002).

Since its advent, the Web has been widely used in both formal and informal scholarly communication and collaboration (e.g., Cronin et al., 1998; Harter & Ford, 2000; Hurd, 2000; Zhang, 2001; Thelwall & Wilkinson, 2003; and Wilkinson et al., 2003). As noted earlier, webometrics thus offers potentials for tracking aspects of scientific endeavor traditionally more hidden from bibliometric or scientometric studies, such as the use of research results in teaching and by the general public (Björneborn & Ingwersen, 2001; Cronin, 2001; Thelwall & Wilkinson, 2003; Thelwall, Vaughan & Björneborn, 2005), but also the actual use of scientific web pages.

A range of new terms for the emerging research field were rapidly proposed from the mid-1990s, for instance, netometrics (Bossy, 1995), webometry (Abraham, 1996), internetometrics (Almind & Ingwersen, 1996), webometrics (Almind & Ingwersen, 1997), cybermetrics (journal started 1997 by Isidro Aguillo), web bibliometry (Chakrabarti et al., 2002). Webometrics and cybermetrics are currently the two most widely adopted terms, often used as synonyms.

Basically, Webometrics is derived from the field of information science. Therefore, content analysis, hyperlink analysis, automatic web pages evaluation, classification of links; various motivational factors for linking the websites are covered in webometric studies. Besides, the above main areas, webometrics deals also with web data collection and its methodology in terms of quality assessment of search engines, web space studies, informetric analysis etc.

The scope of webometrics is closely associated with the dynamic changing nature of the Web which must be taken into account when carrying out informetric analysis of the Web. Printed documents, the main data source of traditional bibliometrics are relatively more permanent compared to Web documents which are constantly changing in several ways. The contents of webpages change, documents are often removed, URLs change, websites disappear, and some documents are temporarily inaccessible. Search engines are the main sources of webometrics data collection. Therefore, knowledge about the availability of commercial search engines, their performance, coverage of the Web, advanced search query formulation etc. form the core of webometric studies.

A special issue in the 50th volume of Scientometrics was dedicated to Internet studies containing, amongst others, a paper on different perspectives of webometrics by Björneborn & Ingwersen (2004) attempting to point to selected areas of webometric research that demonstrate interesting progress and space for development, for instance with regard to graph theoretic approaches to web studies including small-world phenomena, as well as to some more problematic areas, for instance, the so-called Web Impact Factor by Ingwersen (1998) facing methodological difficulties both with regard to reliability due to the dependence on secondary data from commercial search engines with opaque data coverage and performance, and with regard to validity due to problems with defining comparable units of analysis. These matters including the Web Impact Factor are further elaborated in Section 2.3.

A special issue of JASIST on webometrics is set to appear in 2004, containing a wide variety of different approaches to the quantitative study of the Web, including a basic conceptual framework of webometrics proposed by Björneborn & Ingwersen (2004). Bar-Ilan & Peritz (2002) give an excellent review of "Informetric theories and methods for exploring the Internet" with focus on general informetric techniques to be applied in both web studies and non-web Internet research. Furthermore, Bar-Ilan (forthcoming) reviews search engine research in a forthcoming ARIST chapter. Data collection techniques in general on the Web are covered by Bar-Ilan (2001) and Thelwall (2002d).

Henzinger (2001) reviewed link structures analysis from a computer science perspective, showing how links could be used in search engine ranking algorithms. Barabási (2002) and Huberman (2001) have written popular science books explaining current research into mathematical modelling of the topology and growth of the Web including graph theoretic approaches to small-world phenomena on the Web.

Another webometric review article by Park & Thelwall (2003), compared information science approaches to studying the Web to those from social network analysis. It was found that information science tended to emphasize data validation and the study of methodological issues, whereas social network analysis suggested how its existing theory could transfer to the Web.

Aguillo (2002) points out that the webometrics is still in its infancy as a scientific domain -“with its own different theories to be built, tasks to be done, units to be defined, methods to be developed and problems to be solved.”

Besides the few selected examples above of more broadly covering webometric reviews, etc., it is beyond the scope of the project report to provide a more detailed review of the rich diversity of the increasing amount of webometric research.

2.2 SOCIOLOGY OF ACADEMIC WEB SPACES

As noted earlier in Section 2.1, a webometric ‘tradition’ has evolved – in less than 10 years of webometric research – for investigating academic web spaces. This webometric ‘tradition’ may be traced to similar bibliometric and scientometric focus on scholarly publication activities. Being a research field that has grown out of bibliometrics as described in Section 2.1, it is thus not surprising that webometrics shows similar inclinations. The fact that the Web was initially developed for scholarly use (Berners-Lee & Cailliau, 1990) and today has become an exceedingly important platform for both formal and informal scholarly communication and collaboration as mentioned earlier, naturally has contributed to this webometric research interest in academic web spaces.

The Web has thus had a significant – some say revolutionary by Goodrum et al. (2001) – impact on the entire scholarly communication process. According to many researchers in the sociology of science by Cronin & McKim (1996); the Web is reshaping the ways in which scholars communicate with each other. As Cronin & McKim (1996) put it with regard to the ways in which the Web may support and alter the conduct of scholarship:

> *"The Web is much more than a virtual analogue of existing archival and library institutions. It is a dynamic, interactive and evolving environment that supports new kinds of foraging and communication, in which scholars are anything but passive participants."* (p. 163)

In an earlier work, the early LIS recognizers of hypertext potentials, Davenport & Cronin (2000) emphasize how hypertext may affect the conduct and creativity of science by "the freedom of movement inside and across texts" enabled by hypertext that thus allows readers to see referenced sources instantaneously. Early pre-Web hypertext researchers as also build on the 'Memex' vision for creating what they call "webs of information" as hypertextual research literatures enabling scholars "to both create connections and follow those made by others" and thus "link scholars together".

Today, the Web has enabled such world-wide inter-linkage of scholars. On the Web, new kinds of scholarly and proto-scholarly publishing are emerging, implying that work-in-progress, early drafts, preprints and refereed articles are now almost immediately sharable. The Web thus provides fast and efficient means of disseminating and accessing scientific information, with scientists, institutions, and archives making formal research as well as work-in-progress publicly available on their web sites by Goodrum et al. (2001). In other words, the Web offers scholars "instantaneous and interlinked access" by Miles-Board et al. (2001) to large research literatures and other scholarly resources available on the Web. In this context, there is a clear trend, especially for younger researchers, to bypass subscription barriers and rely almost exclusively on what they can find free on the Web which often includes working versions posted on the home pages of the authors. This finding is supported by many author, who found a clear correlation between the number of times an article is cited and the probability that the article is online. As also stated by Zhang (1999):

> *"[...] scholars are using e-sources [Internet-based electronic resources] as a channel to communicate with colleagues, known or unknown; to elicit research ideas from exchanges on mailing lists or newsgroups; to download preprints or reprints; and to seek research-related information. [...] communicating through the network allows researchers to reach broader audiences in an efficient way; hence, it extends the traditional "invisible colleges" model for scholarly communication in the networked environment. Scholars are also relying on e-sources as unique, useful, and current sources of information for research. They often consult e-sources when they need to find some factual, background, or contact information for their research. E-sources also provide efficient ways for scholars to track the progress of related research to stay current." (p. 644)*

One may thus say, that the Web – in spite of the ongoing massive 'colonization' by commercial and other non-academic web players – also still reflects the original idea behind the World-Wide Web developed by Berners-Lee (1989/1990): as a global platform for interlinking academic research by facilitating researchers' information sharing through easy access to online publishing and browsing. In that respect, the Web still reflects and facilitates scholarly activities. In other words and in continuation of the quotation from Zhang (1999) above, the Web provides a richer and more easily accessible as well as more diversified, muddled and cluttered – picture of scholars' scholarly and non-scholarly activities than print media do: (i) curriculum vitae; personal research interests and profiles; ongoing, finished, or planned research projects; links to research partners ('invisible college' cf. Crane, 1972), publication lists, links to work-in-progress, preprints, conference presentations, course syllabi, tutorials, resource guides, bookmark lists, etc. (ii) personal hobby interests, family relations, friends, etc.

This blending of scholars' scholarly and non-scholarly activities made visible on the Web is also stressed by Thelwall (2002) who states that the Web "often provides a public unrefereed creative space that is used for informal research, teaching and recreational information, for example in personal home pages" (p. 563).

Academic websites thus are populated by pages designed for a mixture of purposes and targeted at different audiences (Middleton et al., 1999; Thelwall, 2001). According to Middleton et al. (1999), university websites function as a tool for communication, providing access, and promotion targeted at a variety of users, both internal and external. The latter audience includes prospective students, prospective staff, other academics, alumni, news media, donors/benefactors, and legislators.

Furthermore, web links reflect a diversity of interests, preferences, navigation means and actions of web actors. Thus, motives for making links are more diverse than motives for making references in scientific articles by Wilkinson et al. (2003) and Thelwall (2003). In other words, link structures represent human annotations reflecting cognitive and social structures more extensive than those represented in scientific citation networks because there exist no convention for link creations as for citations in the scientific world.

As stated by Wilkinson et al. (2003), the lack of understanding why web links are created is a major obstacle in webometrics and one “that must be directly addressed in spite of its evident complexity”. Further, they state that the study “has really only scratched the surface of the topic of academic linking motivations”. Using a random sample of 414 inter-university links from the UK academic web space, i.e. the ac.uk domain, Wilkinson et al. (2003), investigated web authors’ motivations for creating links between university web sites. The study showed that over 90% of the links were created for broadly scholarly reasons, including teaching activities.

In this context, it is also important to note that the sociology of academic web spaces differs between scientific domains. Different scientific domains have developed and use distinctly different communicative forums, both in paper and electronic arenas (including the Web), cf. Kling & McKim (2000). The sharing of preprints and other un-refereed papers is thus frequent in some fields, for example, in physics or computer science, but not universal for all fields. Webometrics may cast light on such domain differences in use of the Web.

As made clear in the present section and also noted in Section 1.1, the Web offers obvious new possibilities of tracking and ‘mining’ aspects of scientific endeavour traditionally more hidden for bibliometric or scientometric studies, for instance, the use of research results in teaching and by the general public by Björneborn & Ingwersen (2004); Thelwall & Wilkinson (2003). The realization of such possibilities for new approaches on studying the sociology of science strongly spurred the emergence of the new research field of webometrics. The next section gives a brief overview of webometric research performed in academic web spaces.

2.3 WEBOMETRICS IN ACADEMIC WEB SPACES

Webometrics grew out of a realization that quantitative methods originally designed for bibliometric analysis of citation patterns of scientific journal articles could be applied to the Web by using commercial search engines to provide the raw data. Especially, AltaVista's (www.altavista.com) search interface that allowed complex Boolean search strings including properties of links and URLs triggered this approach.

As further outlined in the review on webometrics by Thelwall, Vaughan & Björneborn (2005), a considerable number of research articles have been published concerning scholarly communication on the Web, mostly originating in the hope that web links could be used to provide similar kinds of information to that extracted from journal citations by Larson (1996), Almind & Ingwersen (1997), Rousseau (1997), Ingwersen (1998), Davenport & Cronin (2000), Cronin (2001), Borgman & Furner (2002a) and Thelwall (2002b). The major difference between the two is that journal citations occur in refereed documents and therefore their production is subject to quality control and they are part of the mainstream of academic endeavor, whereas hyperlinks are none of these things. This makes web links also in the multipurpose sociology of academic web spaces as outlined in the previous section – a more complex phenomenon than journal citations. Thus, several authors like Meyer (1999) and Björneborn & Ingwersen (2001) warn against taking the analogy between citation analyses and link analyses too far.

Bossy (1995) suggested how netometrics, as she called it, could supplement bibliometrics and scientometrics in observing “science in action” on the Internet, enabling “new ways of measuring the impact of scientific contribution that take into account the cooperative aspect of science”. However, she made no empirical investigation of academic web spaces in the paper.

Larson (1996) was one of the first information scientists to perform an investigation of link structures in academic web spaces. In his paper ‘Bibliometrics of the World Wide Web: an exploratory analysis of the intellectual structure of Cyberspace’. He used AltaVista in a co-citation analysis of a set of earth science related websites and could produce clustering of web sites that had topical similarities.

Shortly after, Almind & Ingwersen (1997), in a paper introducing the term webometrics, applied a variety of bibliometric-like methods to the Nordic portion of the Web in order to observe the kinds of page connections and define the typology of web pages found at national Nordic level. The methodology involved stratified sampling of web pages and download for local analysis purposes. The contribution also attempted a comparison between the estimated share of scientific web pages and the distribution found in the citation indexes between the Nordic countries. Clearly, the visibility on the Web was quite different from that displayed in the citation databases. Norway, for instance, was much more visible on a Web scale than in the printed scholarly world at the time of analysis.

'sitations' – using a term coined by McKiernan (1996) for site inlinks – Rousseau (1997) analyzed the patterns of distribution of web sites, site inlinks and site selflinks ('self-citations'). Rousseau's (1997) study operated with 343 web sites retrieved in AltaVista with the search string, informetrics OR bibliometrics OR scientometrics. The study showed that the distribution of TLDs (top level domains, such as .edu, .uk, .dk) for the investigated sites followed the ubiquitous power-law-like Lotka distribution. Similarly, Rousseau demonstrated that the distribution of inlinks to the 343 sites also followed a Lotka distribution.

Ingwersen (1998) introduced the concept of the Web Impact Factor (WIF) for national domains and individual web sites with parallels to the Journal Impact Factors published by the Institute of Scientific Information (ISI) for scientific journals receiving citations from scientific journals indexed in the ISI citation databases (e.g., Hjortgaard Christensen, Ingwersen & Wormell, 1997). In this context, it should be noted that prior to Ingwersen, Rodriguezi Gairin (1997) had introduced the concept of information impact on the Internet in a Spanish documentation journal.

The so-called external WIF for a given web site (or TLD, top level domain) was defined by Ingwersen (1998) as the number of external pages (i.e. pages in other sites or TLDs) with links to the given site (or TLD) divided by the number of web pages at the site (or TLD). However, the fluctuating performance of AltaVista at the time of the study yielded problematic variations in the calculated WIF measures.

Subsequently, Thelwall has developed the WIF measure in several papers in order to find possible correlations to traditional research productivity indicators by Thelwall (2000; 2001a; 2001d; 2001e; 2002a; 2003a), Smith & Thelwall (2001; 2002) and Thelwall & Tang (2003).

As stated by Thelwall, Vaughan & Björneborn (2005), the goal underlying almost all of the research reported above was to validate links as a new information source. One of the key tasks is to compare the link data with other related data in order to establish the degree of correlation and overlap between the two. With links between university web sites, for instance, a positive correlation between link counts and a measure of research would provide some evidence that link creation was not completely random and could be useful for studying scholarly activities.

Thelwall (2001a) showed that the counts of inlinks to a set of 25 UK universities correlated significantly with their average research productivity using the five-yearly UK government Research Assessment Exercise (RAE) cf. HERO (2001). The WIF delivering the highest correlation with the RAE research rankings was the ratio of web pages with links pointing at research-based pages to the number of full-time academic members of staff. This finding provided the first concrete evidence of a real association between research and links, although no cause and effect relationship was claimed. A comparable relationship was later found for Australia by Smith & Thelwall (2002) and Taiwan by Thelwall & Tang, (2003) using different national measures of research productivity.

Thelwall (2001a; 2001e) introduced an important methodological improvement for webometric investigations of academic web spaces; the employment of specially designed web crawlers for collecting primary web data directly from the investigated academic web sites, instead of having to rely on secondary data collected in the big commercial search engines with opaque coverage, update frequency, indexing rules, computing performance, and ranking algorithms, etc. cf. Lawrence & Giles (200), Snyder & Rosenbaum (1999), Björneborn & Ingwersen (2001) and Bar-Ilan (2002).

The interest in the Web Impact Factor thus catalyzed an avalanche of webometric research, especially into links in academic web spaces. In parallel with studies of interlinking between universities, there have been studies of departments within a scientific domain. Thomas & Willett (2000) studied UK library and information science departments, finding no significant correlation between inlink counts and research ratings. An earlier small study of 13 Scottish computer science departments by Chen et al. (1998) revealed a number of correlation relationships between structural connectivity measures and the organizational profile based on research assessment exercise ratings, teaching quality assessments, student-staff ratios and funding levels. Furthermore, linkage patterns from the 13 Scottish academic sites to commercial sites in UK and America highlighted the impact of culture and the appropriateness of information technologies on the acceptance of the Web. The study by Chen et al. (ibid.) has later been criticized by Thelwall, Vaughan & Björneborn (2005) for not taking departmental size into sufficient account.

In another domain study, significant associations between inlink counts and newspaper rankings (US News) were found for US LIS schools, giving the first statistical evidence that departmental level studies could give information about scholarly communication by Thelwall, Vaughan & Björneborn (2005). Subsequently, significant research and inlink count correlations have been found for UK computer science departments, in US psychology and US chemistry departments by Tang & Thelwall, (2003). The latter study found that interlinking between US history departments was too low for patterns to be extracted and that there were significant disciplinary differences in patterns of interlinking between all three domains. This finding supports the earlier mentioned Kling & McKim (2000) who stress the large differences between different scientific fields in the way electronic media, including the Web, are implemented and utilized.

Geographic factors for interlinking in academic web spaces have also been investigated. For example, the degree of interlinking between pairs of UK universities decreases with geographic distance as found by Thelwall (2002d). In particular, neighbouring institutions were much more likely to interlink than average. This shows that despite the existence of collaboratories and other tools for virtual collaboration on the Internet, and its undoubted use for global computer mediated communication, “the Web is not divorced from the physical reality” by Thelwall, Vaughan & Björneborn (2005).

The above examples are mostly from studies employed within national university systems. However, the aforementioned study by Chen et al. (1998) also included cross-national and cross-sectorial link connectivity studies. Another example of a cross-national link structure analysis – while within academia – is the co-inlink (called “cositation”) analysis of 791 university sites from 15 European countries by

Polanco et al. (2001). Smith & Thelwall (2002) compared linking patterns between UK, Australian and New Zealand universities, and found that New Zealand was relatively isolated on the Web, in line with a previous bibliometric study for journals by Glänzel (2001). A larger follow-up study mapped the interlinking between universities in the Asia-Pacific region by Thelwall & Smith (2002) showing that Australia and Japan were central web players in the region, with smaller countries attracting attention disproportionate to their size.

Adamic & Huberman (2000) have used a similar method for aggregating web documents. They studied a crawl of 260,000 web sites each representing a separate domain name. Two sites were considered connected if any of the pages at one site linked to any page in the other. Adamic & Huberman (2000) found that the distribution of inlinks between the sites followed a power law.

The majority of studies in academic web spaces have analyzed inter-university links as outlined above, either within a discipline, within a national university system, or in international university comparisons. However, some studies have focused instead on links between universities and other sectors of society, such as commerce, industry and government, for example, the aforementioned study by Chen et al. (1998). Another example of a cross sectoral study is Leydesdorff & Curran (2000) who mapped university industry government relations on the Internet.

Another important evolving webometric research area in academic web spaces is concerned with scientific journals available on the Web, also known as e-journals cf. Harter & Ford, 2000; Kling & Callahan (2003). Since much of webometrics has been motivated by citation analysis, a natural step has been to see if the kinds of techniques that are applied to journals and authors could also be applied to e-journals by Thelwall, Vaughan & Björneborn (2005). For example, Smith (1999) and Harter & Ford (2000) have investigated e-journals finding no significant correlation between link measures and ISI impact factors for the journals. However, in a large study that incorporated different degrees of web site content, Vaughan & Thelwall (2003) compared inlink counts and ISI's Journal Impact Factors of 88 law and 38 LIS journals indexed by ISI. The findings confirmed that in both law and LIS, counts of links to journal web sites correlated with Journal Impact Factors. Not surprisingly, journals with more extensive online content tended to attract more links than older journal web sites.

Direct web citations between scientific papers were investigated by Vaughan & Shaw (2005). Using Google to collect web citation data to 46 LIS journals, the authors found predominantly significant correlations with traditional citations, suggesting that online and offline citation impact are similar phenomena. A classification of 854 web citations to papers in the LIS journals indicated that many web citations "represented intellectual impact, coming from other papers posted on the Web (30%) or from class readings lists (12%)".

Other webometric studies of academic web spaces have chosen scholars rather than journal web sites or journal papers as the basic unit of analysis, for example, collecting data based upon invocations: the mentioning of a scholar's name in any context in a web page. Cronin et al. (1998) found academics to be invoked on the Web in a wide variety of contexts, including informal documents such as conference information pages and course reading lists. This was used to support a claim that web invocations could help to highlight “the diverse ways in which academic influence is exercised and acknowledged”.

In this context, a study by found that it was the quantity of research produced by scholars that was the main reason for attracting inlinks: universities with better researchers attract more inlinks because the researchers produce more web content, rather than because the content produced is of a higher ‘link attractiveness’ by Thelwall, Vaughan & Björneborn (2005). This is in contrast to the case for formal scholarly publications, where better scholars tend to produce papers that attract more citations by Borgman & Furner (2002).

Thelwall and Sud (2011) estimating the online impact of an organisation is to count links to its web site. Link counts have been available from commercial search engines for over a decade but this was set to end by early 2012 and so a replacement is needed. This article compares link counts to two alternative methods: URL citations and organisation title mentions. New variations of these methods are also introduced. The three methods are compared against each other using Yahoo. Two of the three methods (URL citations and organisation title mentions) are also compared against each other using Bing. Evidence from a case study of 131 UK universities and 49 US Library and Information Science (LIS) departments suggests that Bing's Hit Count Estimates (HCEs) for popular title searches are not useful for webometric research but that Yahoo's HCEs for all three types of search and Bing’s URL citation HCEs seem to be consistent. For exact URL counts the results of all three methods in Yahoo and both methods in Bing are also consistent. Four types of accuracy factors are also introduced and defined: search engine coverage, search engine retrieval variation, search engine retrieval anomalies, and query polysemy.

A survey was conducted in order to test the coverage of search engines to calculated WIF in Indian researchers. Some of the related Indian articles which dealt with the subject are: Jeyshankar and Ramesh Babu (2009), examines and explores through a webometric study the webistes of 45 universities in Tamil Nadu comprising of 27 state and 18 private universities. Identifies the domain systems of the websites; analyses the number of web pages and link pages; and calculates the simle WIF, self link WIF and external WIF of the university webiste in Tamil Nadu and ranks the websites as per the WIF. Reflects that some universities in Tamil Nadu have higher number of web pages but correspondingly their link pages are very small in number and websites fall behind in their simle, self link and external link WIF.

Jalal, Biswas and Mukhopadhyay (2010a) attempt has been made to rank central universities in India using appropriate webometric indicators. Results revealed that based on WISER ranking University of Delhi gets top rank; Sikkim University occupied the last rank among central universities in India. The study also examined the link relationship through inlinks and outlinks among central universities using personal web crawler i.e. SocSciBot 3.0. The reason for hyper-linking is traced among the universities under study. Finally, co-link matrix was formed to detect the link pattern among the central universities in India.

Another article above same authors, investigates the effectiveness and relevance of web impact factor for Indian universities' websites. Reviews WIF as to how this link-based metrics is developed and is applied. SocSciBot 3.0 is used to generate link data in order to develop/form micro-link topology under study. Result shows that all the NITs are closely related in the topology framework/their activities whereas nodes are not linked significantly for the case of state universities and central universities (Jalal, Biswas and Mukhopadhyay, 2010b).

Another article by Jalal, Biswas and Mukhopadhyay (2011) describes the trend from bibliometrics to webometrics. Also discusses some important application areas of Webometrics research, the methodology adopted for data collection, techniques and tools of Web analysis and the problems encountered in Webometric research. Methods of computing Web Impact Factors and highlights the research possibility in Webometric study are considered. A webometric case study of 13 Indian Institutes of Technology (IITs) and Indian Institutes of Management (IIMs) is presented. The analysis shows that the IIT, Bombay occupies the first place among the 13 IITs and the IIMs based on both absolute and logarithm values.

Ratha, Joshi, and Naidu (2012) found significant differences according to some important point of view such as the user supporting services, number of hyperlinks on home pages and whole websites, number of images, location of images, In-active links and web pages, etc. The paper finally looks the design and structure of the library websites of IITs.

Shukla and Poluru (2012) analyze web presence of Indian State Universities (173) on the World Wide Web (WWW) and also to find ways to get high web links that further help to improve presence on Web. The data was collected from Yahoo Site Explorer and Google Scholar. WISER (Web Indicators for Science, TEchnology and Innovation Research) ranking method was applied to know about the visibility and connectivity of universities on the Web. A WebCrawler Socscibot 4 was used to examine the pattern of link visibility and to create ADM (Alternative Document Model) Count Summary. Pajek, software for large network analysis was used as a graphical tool for creating inlinks topology. The study shows that some state universities have more visibility compared with their counter parts. It was also noted that the impact of maintaining institutional repositories, promotion of open access, academic researches, collaboration with other universities, online communities etc., are helpful in increasing the more visibility of the a university on the Web.

Thanuskodi (2012) presents the web based information re-sources. A huge amount of data in every subject stream is available on different websites. Internet is a network of networks carrying information on almost any subject under the sun. Everybody today would like to be on internet because of wealth of information that lies there to be exchanged. And, with its global connections and millions of users, the internet is world's biggest electronic library and publish gathering place, which contain a vast amount of information. Information professionals and users face a number of challenges in networked information resources and service environment. But problem is only one that how to retrieve desired information. Many attempts were made and solution comes as 'webometrics'. In this paper an attempt has been made to describe this concept.

Vijayakumar, Kannappanavar, and Santosh (2012) focuses on the identification of web presence and their links among SAARC countries. The research explores that

India possesses maximum of 14, 10, 00, 000 webpages; 58, 20, 000 external links; 1, 18, 00,000 internal links; and 9, 83, 00,000 over all links. In case of web impact factor of external and internal links, Sri Lanka claims highest and for over all links once again India claims highest. One can note that, except India no SAARC country possesses all sub-domains, but these possess only a few sub-domains like .edu, .gov, .net, and .org. When analysis of the links was done, it was found that Pakistan has a maximum of 3,610 links to India as compared to other SAARC countries. India once again claims top position among the SAARC countries for Wiser ranking.

All the above forgoing articles and studies indicate that a considerable work has been carried out on the subject matter of Webometric analysis and university websites, it is clearly inferred that despite the bonanzic literature available in the sphere of webometrics pertaining to the webometric analysis of Websites of IIT institutes. This literature will be pertinent for carrying out a study on evaluation of university websites. In other words, this literature can be applied for studying the **Webometric Study of Websites of Indian Institutes of Technology**.

Chapter-3

PROFILES OF INDIAN INSTITUTES OF TECHNOLOGY

The Indian Institutes of Technology (IITs) are a group of autonomous public engineering institutes of higher education. The IITs are governed by the Institutes of Technology Act, 1961 which has declared them as "institutions of national importance", and lays down their powers, duties, framework for governance etc.

The sixteen institutes located at Bhubaneswar, Chennai, Delhi, Gandhinagar, Guwahati, Hyderabad, Indore, Jodhpur, Kanpur, Kharagpur, Mandi, Mumbai, Patna, Ropar, Roorkee and Varanasi. Each IIT is an autonomous institution, linked to the others through a common IIT Council, which oversees their administration. Over the years, IITs have created world class educational platform that is dynamically sustained through quality teaching and internationally acclaimed research with excellent infrastructure and the best available minds. The faculty and alumni of IITs continue to make huge impact in all sectors of society, both in India and abroad. IITs are Institutes of National Importance established through special acts of Indian Parliament. The success of the IITs led to the creation of the Indian Institutes of Information Technology (IIIT) in the late 1990s and in the 2000s.

This chapter provides brief description of sixteen Indian Institutes of Technology (IITs) with websites in the following sections:

3.1 INDIAN INSTITUTE OF TECHNOLOGY (BHU) VARANASI

Indian Institute of Technology (BHU) Varanasi (IIT-Varanasi or IIT-V), formerly known as Institute of Technology, Banaras Hindu University (IT-BHU), is an Indian public engineering institute located in Varanasi, Uttar Pradesh, India. Founded in 1919 as part of the Banaras Hindu University, it was designated an Indian Institute of Technology in 2012. IIT -Varanasi has 13 departments and three inter-disciplinary schools.

IIT-Varanasi has a residential and co-educational campus within the larger BHU campus which is spread over nearly 1,300 acres (5.3 km^2) at the southern end of Varanasi on the banks of the River Ganges. In 1971, three faculties of BHU viz., Banaras Engineering College (BENCO), College of Mining & Metallurgy (MINMET) and College of Technology (TECHNO) were merged to form the Institute of Technology, Banaras Hindu University (IT-BHU). IT-BHU was designated as an IIT by The Institutes of Technology (Amendment) Act, 2011 which was passed by the Lok Sabha on 24 March 2011 and by the Rajya Sabha on 30 April 2012. The President signed the Bill on 20 June 2012 and was notified in the Gazette of India on 21 June 2012. Figure 3.1 depicts the snapshot of the website of Indian Institute of Technology (BHU), Varanasi.

Figure 3.1: Snapshot of Indian Institute of Technology (BHU) Varanasi
(Source: http://www.itbhu.ac.in, accessed on: 25-10-2012)

3.2 INDIAN INSTITUTE OF TECHNOLOGY BHUBANESHWAR

Indian Institute of Technology Bhubaneswar is an "engineering and technology" higher education institute, located at Bhubaneswar, Orissa, India. It is one of the eight new Indian Institutes of Technology (IITs) established by the Ministry of Human Resource Development, Government of India under The Institutes of Technology (Amendment) Act, 2011 which declares these eight IITs as well as the conversion of Institute of Technology, Banaras Hindu University to IIT. The Act was passed in the Lok Sabha on 24 March 2011 and by the Rajya Sabha on 30 April 2012. It started functioning from the campus of IIT Kharagpur on 23 July 2008 and shifted its operation to the city of Bhubaneswar on 22 July 2009. Figure 3.2 presents the snapshot of the website of Indian Institute of Technology, Bhubaneswar.

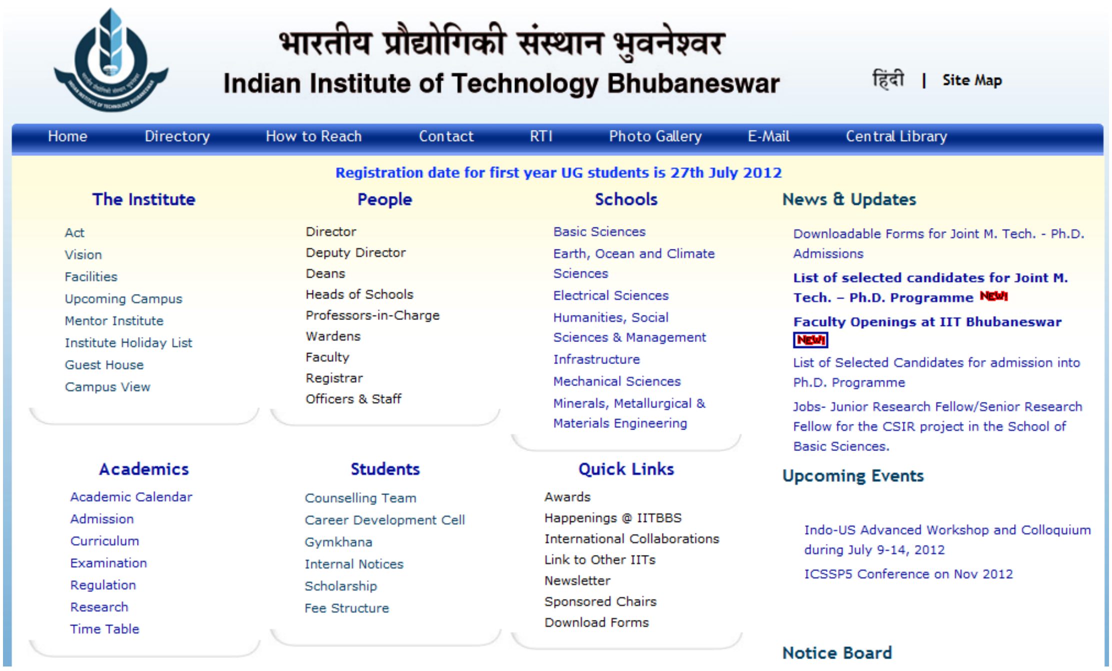

Figure 3.2: Snapshot of Indian Institute of Technology Bhubaneswar
(Source: http://www.iitbbs.ac.in, accessed on: 26-10-2012)

3.3 INDIAN INSTITUTE OF TECHNOLOGY BOMBAY

IIT Bombay was the second IIT to be established in 1958 with assistance from UNESCO and with funds contributed by the Soviet Union. UNESCO agreed to provide equipment and technical experts mainly from the Soviet Union, while the Government of India accepted the responsibility for all other expenses including the cost of the building project and recurring expenses.

The students were selected from over 3,400 applicants for admission to the first year undergraduate engineering programmes of Aerospace, Chemical, Civil, Computer, Electrical, Engineering Physics, Energy, Mechanical, Metallurgical Engineering and M.Sc. Chemistry. One of the main objectives of establishing the Institute was to develop facilities for studies in a variety of specialized engineering and technological sciences. The need for establishing adequate facilities for postgraduate studies and research was kept uppermost in mind in the founding years.

While the Institute was functioning provisionally at Worli, an effort was made to expedite the progress of the building project at its permanent location and Jawaharlal Nehru laid the foundation stone of the Institute at Powai on March 10, 1959. Figure 3.3 presents the snapshot of the website of Indian Institute of Technology, Bombay.

Figure 3.3: Snapshot of Indian Institute of Technology Bombay
(Source: http://www.iitb.ac.in, accessed on: 26-10-2012)

3.4 INDIAN INSTITUTE OF TECHNOLOGY DELHI

The Government of India negotiated with the British Government for collaboration in setting up an Institute of Technology at Delhi. The British Government agreed in principle to such collaboration, but was inclined initially to start in a modest way. It was therefore agreed that a College of Engineering & Technology should be established at Delhi with their assistance. A trust called the Delhi Engineering College Trust was established with the help of the UK Government and the Federation of British Industries in London. Later H. R. H. Prince Philips, Duke of Edinburgh, during his visit to India, laid the foundation stone of the college at Hauz Khas on January 28, 1959.

The College of Engineering & Technology was registered as a Society on 14th June, 1960, under the Societies Registration Act No. XXI of 1860 (Registration No.S1663 of 1960-61). The first admissions were made in 1961. The students were asked to report at the College on 16th August, 1961, and the College was formally inaugurated on 17th of August, 1961, by Prof. Humayun Kabir, Minister of Scientific Research & Cultural Affairs. The College was affiliated to the University of Delhi.

The College of Engineering & Technology established in 1961 was declared an institution of National Importance under the "Institute of Technology (Amendment) Act 1963" and was renamed "Indian Institute of Technology Delhi". It was then accorded the status of a University with powers to decide its own academic policy, to conduct its own examinations, and to award its own degrees.

IIT Delhi is an autonomous statutory organisation functioning in terms of the Institutes of Technology Act, 1961 amended vide the Institutes of Technology (Amendment) Act, 1963, and the Statutes framed thereunder. Figure 3.4 presents the snapshot of the website of Indian Institute of Technology, Delhi.

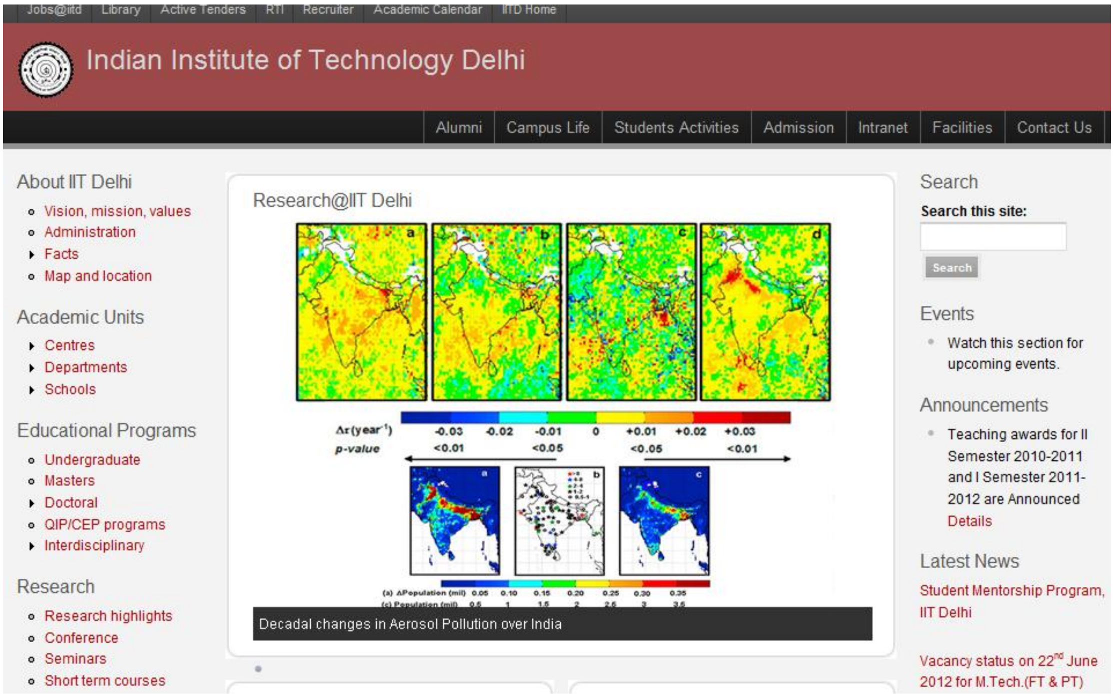

Figure 3.4: Snapshot of Indian Institute of Technology Delhi

(Source: http://www.iitd.ac.in, accessed on: 26-10-2012)

3.5 INDIAN INSTITUTE OF TECHNOLOGY GANDHINAGAR

The Indian Institute of Technology Gandhinagar (IITGN) was founded in the year 2008. It is currently located in Chandkheda, Ahmedabad in Gujarat. Ahmedabad is one of the oldest living cities in India, known both for a rich cultural past as well as state-of-the-art infrastructure, thriving industries, and many modern amenities.

IITGN has excellent facilities in terms of classrooms, laboratory spaces, faculty offices, etc. IIT Gandhinagar currently offers a 4-year B.Tech. programme in Chemical, Electrical and Mechanical Engineering and an M.Tech. programme in Chemical, Civil, Electrical, Mechanical Engineering and in Metallurgy Materials Science. In addition, the Institute offers the students the option to do a Minor in Computer Science or Management along with their B.Tech. degrees. It also offers PhD. progammes in several disciplines in Engineering (Chemical, Civil, Computer

Science, Electrical, Mechanical, and Metallurgy & Materials Science), Sciences (Chemistry, Physics and Mathematics) and Humanities & Social Sciences (Cognitive Science, English, Philosophy, Sociology and others). IITGN goes many steps beyond the call and definition of a technology Institute, it aims to create excellence in science, technology as well as the Humanities and Social Sciences and thereby create rounded and nuanced minds. Figure 3.5 presents the snapshot of the website of Indian Institute of Technology Gandhinagar.

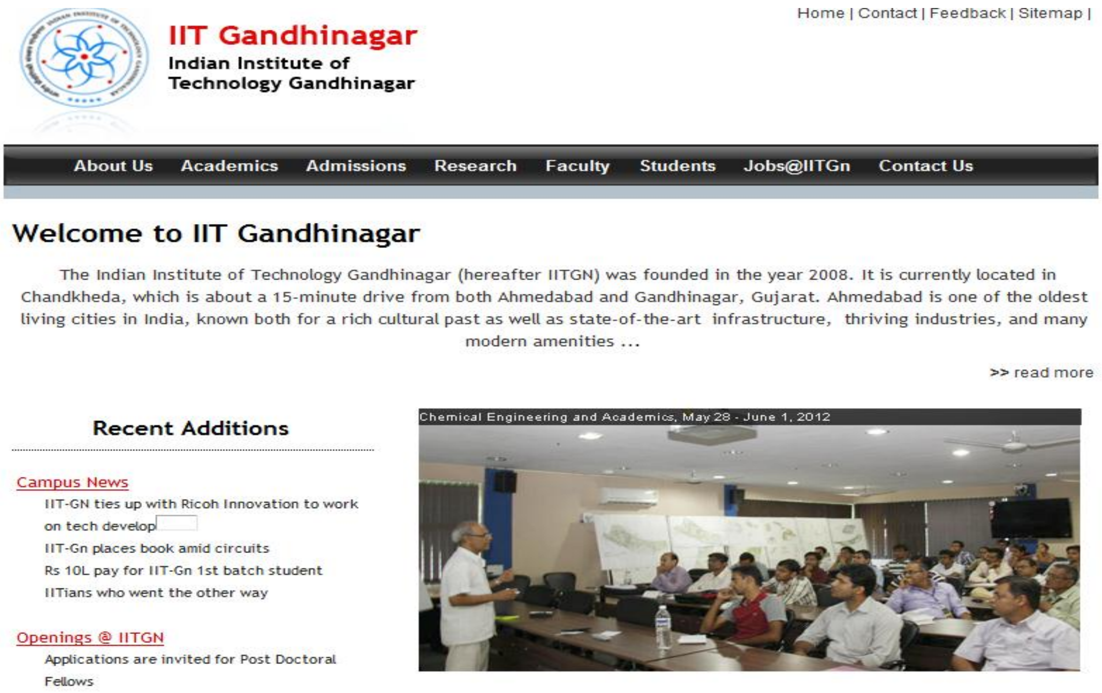

Figure 3.5: Snapshot of Indian Institute of Technology Gandhinagar
(Source: http://www.iitgn.ac.in, accessed on: 26-10-2012)

3.6 INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

Indian Institute of Technology Guwahati, the sixth member of the IIT fraternity, was established in 1994. The academic programme of IIT Guwahati commenced in 1995.

At present the Institute has eleven departments and three inter-disciplinary academic centres covering all the major engineering, science and humanities disciplines, offering BTech, BDes, MA, MDes, MTech, MSc and PhD programmes. Within a short period of time, IIT Guwahati has been able to build up world class infrastructure for carrying out advanced research and has been equipped with state-of-the-art scientific and engineering instruments.

Indian Institute of Technology Guwahati's campus is on a sprawling 285 hectares plot of land on the north bank of the river Brahmaputra on one side, and with hills and vast open spaces on others, the campus provides an ideal setting for learning. Figure 3.6 depicts the the snapshot of the website of IIT Guwahati.

Figure 3.6: Snapshot of Indian Institute of Technology Guwahati
(Source: http://www.iitg.ac.in or http://www.iitg.ernet.in, accessed on: 27-10-2012)

3.7 INDIAN INSTITUTE OF TECHNOLOGY HYDERABAD

IIT Hyderabad (IITH) is part of history in the making. In a very short time, just about a year, IIT Hyderabad has made significant strides. IIT Hyderabad admitted the first batch of 111 B.Tech., students, and started functioning on August 20, 2008. Three departments, CSE, EE & ME were initiated. The first year had its highs and lows but was an exciting period that will be etched in the minds of the pioneer batch. In January 2009, IIT Hyderabad admitted 11 PhD students. On February, 27, 2009, the foundation stone of IIT Hyderabad was laid by Smt. Sonia Gandhi. For the 2009-10 academic year, IITH will be admitting 120 B.Tech. students, 35 M.Tech. students and 10-15 Ph.D. students.

From the very first year IITH has embarked on research and development; it received two funded projects: One on "Development of High Energy Density Li-on Batteries for Mobile Applications" jointly with International Advanced Research Centre for Powder Metallurgy and New Materials (ARCI). The second is a joint project with Kieo University and University of Tokyo on "Information Network for Natural Disaster Mitigation and Recovery". IITH, through its Master Plan, has embarked on creating a signature campus which will provide the ambience for fostering inventions and innovations. Figure 3.7 presents the snapshot of the website of Indian Institute of Technology Hyderabad.

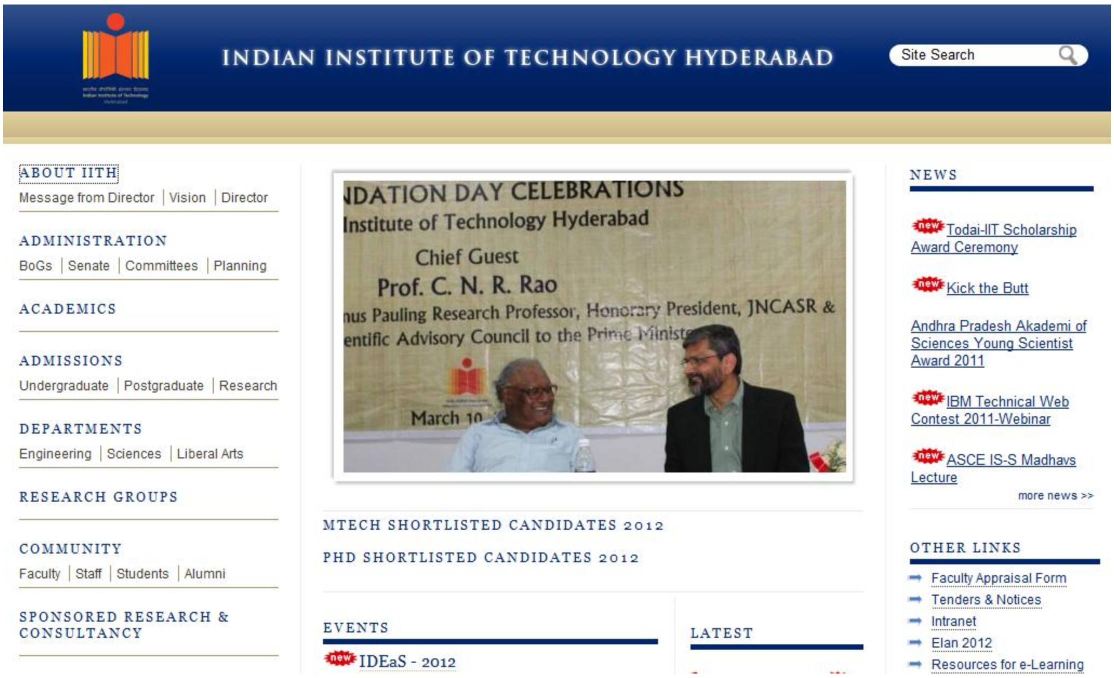

Figure 3.7: Snapshot of Indian Institute of Technology Hyderabad
(Source: http://www.iith.ac.in, accessed on: 27-10-2012)

3.8 INDIAN INSTITUTE OF TECHNOLOGY INDORE

Indian Institute of Technology Indore (IITI), located in Madhya Pradesh, is an institute of national importance established by the Government of India in 2009. It is one of the eight new Indian Institutes of Technology (IITs) established by the Ministry of Human Resource Development, Government of India under The Institutes of Technology (Amendment) Act, 2011 which declares these eight IITs as well as the conversion of Institute of Technology, Banaras Hindu University to IIT. The Act was passed in the Lok Sabha on 24 March 2011 and by the Rajya Sabha on 30 April 2012.

The institution started functioning from 2009-10 in a temporary campus at Institute of Engineering and Technology of Indore University under mentorship of IIT Bombay. Figure 3.8 presents the snapshot of the website of Indian Institute of Technology Indore.

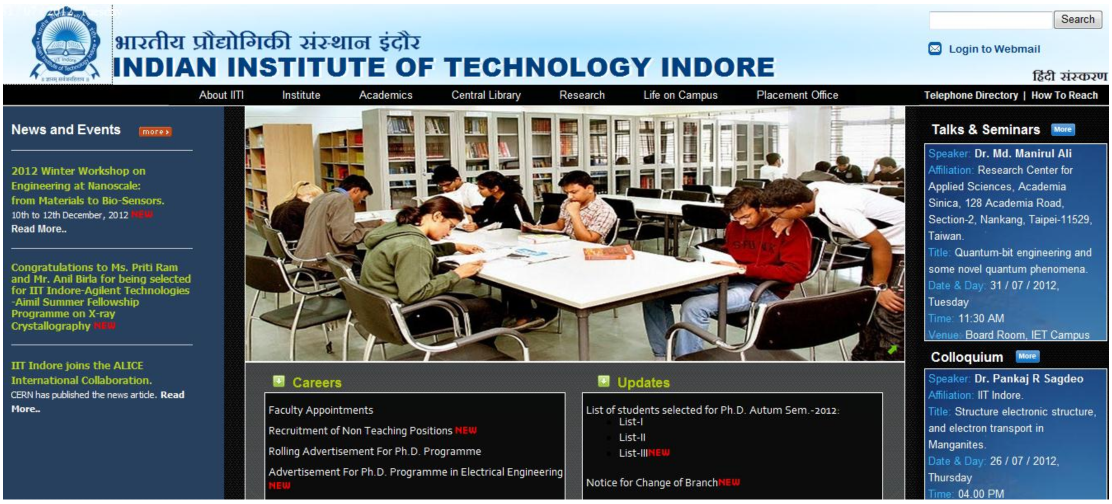

Figure 3.8: Snapshot of Indian Institute of Technology Indore
(Source: http://www.iiti.ac.in, accessed on: 27-10-2012)

3.9 INDIAN INSTITUTE OF TECHNOLOGY KANPUR

IIT Kanpur was established by an Act of Parliament in 1959. The institute was started in December 1959 in a room in the canteen building of the Harcourt Butler Technological Institute at Agricultural Gardens in Kanpur. In 1963, the institute moved to its present location, on the Grand Trunk Road near the village of Kalyanpur in Kanpur district.

During the first ten years of its existence, IIT Kanpur benefited from the Kanpur Indo-American Programme (KIAP), where a consortium of nine US universities (namely M.I.T, University of California, Berkeley, California Institute of Technology, Princeton University, Carnegie Institute of Technology, University of Michigan, Ohio State University, Case Institute of Technology and Purdue University) helped set up the research laboratories and academic programmes. The first Director of the Institute was P. K. Kelkar (after whom the Central Library was renamed in 2002).

Under the guidance of economist John Kenneth Galbraith, IIT Kanpur was the first institute in India to offer Computer Science education. The earliest computer courses were started at IIT Kanpur in August 1963 on an IBM 1620 system. The initiative for computer education came from the Electrical Engineering department, then under the chairmanship of Prof. H.K. Kesavan, who was concurrently the chairman of Electrical Engineering and head of the Computer Centre. In 1971, the institute initiated an independent academic program in Computer Science and Engineering, leading to M.Tech. and Ph.D. degrees.

In 1972 the KIAP program ended, in part because of tensions due to the U.S. support of Pakistan. Government funding was also reduced as a reaction to the sentiment that the IIT's were contributing to the brain drain. Since then IIT Kanpur has increased collaboration with industry and has achieved a reasonable level of stability. Figure 3.9 presents the snapshot of the website of Indian Institute of Technology Kanpur.

Figure 3.9: Snapshot of Indian Institute of Technology Kanpur
(Source: http://www.iitk.ac.in, accessed on: 27-10-2012)

3.10 INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

The Indian Institute of Technology Kharagpur (IITK) is an autonomous public engineering, technology and management oriented institute of higher education established by the government of India in 1951. The first IIT to be recognized as an Institute of National Importance by the government of India.

The institute was established to train scientists and engineers after India attained independence in 1947. It shares its organisational structure and undergraduate admission process with sister IITs. The students and alumni of IIT Kharagpur are informally referred to as KGPians. Among all IITs, IIT Kharagpur has the largest campus (2,100 acres), the most departments, and the highest student enrollment. IIT Kharagpur is known for its festivals: Spring Fest (Social and Cultural Festival), Kshitij (Techno-Management Festival) and Shaurya (sports festival). Figure 3.10 presents the snapshot of the website of Indian Institute of Technology Kharagpur.

Figure 3.10: Snapshot of Indian Institute of Technology Kharagpur
(Source: http://www.iitkgp.ac.in, accessed on: 27-10-2012)

3.11 INDIAN INSTITUTE OF TECHNOLOGY MADRAS

The Indian Institute of Technology Madras (IITM) is an autonomous institution of higher education and research, located in Chennai (formerly Madras), Tamil Nadu. It is recognized as an Institute of National Importance by the Government of India. Founded in 1959 with technical and financial assistance from the government of the former West Germany, it is third among the Indian Institutes of Technology established by the Government of India through an Act of Parliament, to provide education and research facilities in engineering and technology.

IIT Madras is a residential institute that occupies 620 acres campus that was formerly part of the adjoining Guindy National Park. The institute has nearly 460 faculty, 6,000 students and 1,250 administrative and supporting staff. Growing ever since it obtained its charter from the Indian Parliament in 1961, Much of the campus is a protected forest, carved out of the Guindy National Park, home to chital (spotted deer), black buck, and other wildlife. A natural lake, deepened in 1988 and 2003, drains most of its rainwater. Figure 3.11 presents the snapshot of the website of Indian Institute of Technology Madras.

Figure 3.11: Snapshot of Indian Institute of Technology Madras
(Source: http://www.iitm.ac.in, accessed on: 29-10-2012)

3.12 INDIAN INSTITUTE OF TECHNOLOGY MANDI

Indian Institute of Technology Mandi (IIT Mandi) is an engineering and technology higher education institute located in Mandi. It is one of the eight new IITs established by the Ministry of Human Resource Development (MHRD), Government of India under The Institutes of Technology (Amendment) Act, 2011 which declares these eight IITs as well as the conversion of Institute of Technology, Banaras Hindu University to IIT. The Act was passed in the Lok Sabha on 24 March 2011 and by the Rajya Sabha on 30 April 2012.

IIT Mandi is mentored by IIT Roorkee, which hosted the first batch of students. Presently, IIT Mandi functions from transit campus at Mandi town in Himachal Pradesh. Permanent campus (about 12 km from historic city of Mandi) is under construction along Uhl River (a tributary of River Beas) at Kamand, Mandi. The institute admits students into various B.Tech., M.S, and Ph.D. programs (as of 2010). Figure 3.12 presents the snapshot of the website of Indian Institute of Technology Mandi.

Figure 3.12: Snapshot of Indian Institute of Technology Mandi
(Source: http://www.iitmandi.ac.in, accessed on: 29-10-2012)

3.13 INDIAN INSTITUTE OF TECHNOLOGY PATNA

The Indian Institute of Technology, Patna (IIT Patna) is an autonomous institute for education and research in science, engineering and technology located in Patna, India. It is recognized as an Institute of National Importance by the Government of India. It is one of the eight new IITs established by the MHRD, Government of India under The Institutes of Technology (Amendment) Act, 2011 which declares these eight IITs as well as the conversion of Institute of Technology, Banaras Hindu University to IIT. The Act was passed in the Lok Sabha on 24 March 2011 and by the Rajya Sabha on 30 April 2012. Figure 3.13 presents the snapshot of the website of Indian Institute of Technology Patna.

Figure 3.13: Snapshot of Indian Institute of Technology Patna
(Source: http://www.iitp.ac.in, accessed on: 29-10-2012)

3.14 INDIAN INSTITUTE OF TECHNOLOGY RAJASTHAN

The Indian Institute of Technology Rajasthan represents a distinctive culture of engineering and science education along with other areas of scholarship and generates research to best fulfill the current and emerging needs of society, enhancing its prosperity through environmentally sustainable means. Founded in 2008 in collaboration with France, IIT Rajasthan functions through centres of excellence with innovative curriculum design, interdisciplinary research and thought leadership, delivering to society leaders who are entrepreneurial and well-rounded.

The Indian Institute of Technology Rajasthan focuses on imparting education that prepares students toward entrepreneurship and public service. Administering its programs through Centres of Excellence, IIT Rajasthan transcends the compartmentalization of education brought about by departments, and aims at excellence in teaching and research that would benefit both the academia and the industry. The institute emphasizes multi-disciplinarily in its undergraduate curriculum and interdisciplinary specializations at the postgraduate level. Additionally, IIT Rajasthan's close association with leading Indian and foreign industries and research institutes effectively helps researchers gain hands-on experience and solve technical challenges. Having recognized the importance of international collaborations in academia, IIT Rajasthan has undertaken a joint venture with a French consortium comprising of twenty highly-ranked and diversified academic institutions of France. The aim of this collaboration is to upgrade Indo-French cooperation on capabilities of knowledge management, higher technical education, training, applied research, technology, and knowledge dissemination. Today, as IIT Rajasthan continues to expand, it is driven by a zeal to impart education wherein innovation and entrepreneurship are the chief goals. Functioning presently from its temporary campus at MBM Engineering College in Jodhpur Rajasthan, the institute in the forthcoming years will shift to a sprawling residential campus near Nagaur in Jodhpur. It is envisioned that the new campus of IIT Rajasthan would stand as a symbol of academic excellence while creating a multi-cultural ethos with centres such as an Eco Village, an Arts and Culture Centre and an International Inter-Cultural Activity Centre, all of which would lead to a holistic development of its community. Figure 3.14 presents the snapshot of the website of Indian Institute of Technology Rajasthan.

Figure 3.14: Snapshot of Indian Institute of Technology Rajasthan
(Source: http://www.iitj.ac.in, accessed on: 29-10-2012)

3.15 INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

The institution has its origins in a class started in 1845 to train local youth in engineering to assist in public works then beginning. In 1847 it was officially established. It was renamed as the Thomason College of Civil Engineering in 1854 in honour of its founder, Sir James Thomason, lieutenant governor 1843–53. The first Indian to pass out from the Roorkee college was Rai Bahadur Kanhaiya Lal in 1852.

Initially, the college had engineers' class only for Europeans, upper subordinate class for Europeans and Indians and lower subordinate class for Indians only. Such was the reputation of the college, that the recruitment of the Engineering students was directly controlled by the Public Works Departments (PWD). Every student was guaranteed a post in the PWD/irrigation departments. Roorkee pass outs played a role in all the areas of engineering primarily civil, including maintenance of the Ganges canal, construction of dam and irrigation projects like Bhakra Nangal, the Rajasthan canal, the Aswan dam on the Nile in Egypt, and construction of Chandigarh.

Between 1934 to 1943, officers of the Indian Army Corps of Engineers received training at the Thomason College of Engineering and, even after the establishment of the School of Military Engineering (SME) at Roorkee in 1943, they continued to receive technical training at Thomason. In 1948 when SME was moved to Dapodi, Pune. It was given the status of University by Act No. IX of 1948 of the United Provinces (Uttar Pradesh) and was titled University of Roorkee. Jawahar Lal Nehru, the first Prime Minister of India, presented the Charter in November 1949, elevating the erstwhile college to the first engineering university of independent India. Soil scientist Jagdish Narain was the first student to be admitted into the university under this act.

On 21 September 2001, an ordinance issued by the Government of India declared it as the nation's seventh Indian Institute of Technology, renaming it to the current name, Indian Institute of Technology Roorkee. The ordinance was converted into an act by the Parliament to make IIT Roorkee an "Institution of National Importance". Figure 3.15 presents the snapshot of the website of Indian Institute of Technology Roorkee.

Figure 3.15: Snapshot of Indian Institute of Technology Roorkee
(Source: http://www.iitr.ac.in, accessed on: 30-10-2012)

3.16 INDIAN INSTITUTE OF TECHNOLOGY ROPAR

Indian Institute of Technology, Ropar is one of the eight new IITs set up by the MHRD, Government of India, to expand the reach and enhance the quality of technical education in the country. This institute is committed to providing state-of-the-art technical education in a variety of fields and also for facilitating transmission of knowledge in keeping with latest developments in pedagogy. These two areas of focus will enable students to gain exposure to recent trends in their chosen domains of study and gain practical experience through a wide variety of activities the institute facilitates in its own campus and arranges for in collaboration with industry and other institutes. At present, the institute offers B.Tech programme in the following disciplines: Computer Science and Engineering, Electrical Engineering, and Mechanical Engineering. This programme is spread over a period of eight semesters and the institute admits forty students in each branch, selected through IIT Joint Entrance Examination conducted every year. In addition, the institute now offers doctoral programme in several disciplines. Figure 3.16 presents the snapshot of the website of Indian Institute of Technology Ropar.

Figure 3.16: Snapshot of Indian Institute of Technology Ropar
(Source: http://www.iitrpr.ac.in, accessed on: 30-10-2012)

Chapter-4

DATA ANALYSIS AND INTERPRETATION

Websites of IITs are supposed as important communication tools and making interaction and collaboration among persons and organizations doing academic and research affairs, creating professional unity, defending rights of individuals and organizations, contributing to the development of infrastructure in the countries as well as facilitating access to information is undeniable. Therefore, in-links and use rate of these websites, their visibility, group collaborations, their impact factor and identifying the most important websites are problems to be studied in the present research.

Evaluation of web sites is a formidable but necessary task considering the wide range of choices available. The WIF & WISER, as explained in the above, is a useful tool for evaluation of web sites, but it must be used discreetly. Considerations include the amount of webpages or other types of material published in a web site, contents, and variations between disciplines. The web sites' status in regard to coverage in the search engines' databases as well as the occurrence of a domain name change is also very important. The WIFs are always approximate and not absolute. The WIF of a site is not stable, because every day some webmasters are deleting the old inlinks to several web sites and others are linking to new ones. The WIF would still be far from being a quality indicator: link impact is primarily a measure of scientific utility rather than of scientific quality. For evaluation of scientific quality, there seems to be no alternative to qualified experts reading the web site resources. All WIF studies should be normalized to take into account variables such as field, or discipline, country, language, and link practices.

The WIF is not a perfect tool to measure the quality of web sites but there is nothing better and it has the advantage of already being in existence and is, therefore, a technique for quantitative evaluation of web sites. Despite the recognition that the WIF is an imperfect measure and eight years of criticism, there is no obvious alternative. Thus, those forced to use this tool for direct web site comparison should be encouraged to remain open-minded and cautious, with an awareness of the inherent limitations of its use. While the WIF is arguably useful for quantitative intra-country comparison, application beyond this (i.e., to inter-country assessment) has little value. In the future, there may be more sophisticated ways of assessing the quality of web sites. Therefore, it seems that IITs have made remarkable progress in developing their websites the study also reflected the comparison of different methodological approach and visibility of IITs.

This chapter deals with the analysis and interpretation of data, which have been collected from Indian Institutes of Technology websites. The collected data are organized and tabulated by using webometric methods.

Evaluation of sixteen (16) websites of IIT was carried out during October 25, 2012 to November 25, 2012 with the help of Google, AltaVista and Yahoo search engines.

4.1 TOP LEVEL DOMAIN OF WEBSITES OF IIT

As per the objective of this study, Top Level Domains **(**TLDs) of IITs, Second Level Domains (SLDs) related to education and research domain under the TLD for India (i.e .in) and the Universal Resource Locators (URLs) of 16 IITs have been collected by conducting Internet searching using Google, AltaVista and Yahoo search engines.

The selected search engine is then searched against all the domain names and URLs to check whether the AltaVista databases include these domain / sites or not. Each URL has been checked by visiting at least twice daily during the period of evaluation period. Table 4.1T1 through 4.1T3 presents the TLDs, SLDs under .in and URLs of the study websites of IITs in India.

Table - 4.1T1: Top Level Domain of Websites of IITs (Group-I)

TLD of IITs	
India	**.in**

Table - 4.1T2: Second Level Domain of Websites of IITs (Group-II)

SLD related to education and research activities under TLD for India	
Academic activities	**ac.in**

Table - 4.1T3: Domain name of hosts of IITs (Group- III)

Domain name of hosts of IITs	URL
Indian Institute of Technology (BHU) Varanasi	iitbhu.ac.in
Indian Institute of Technology Bhubaneshwar	iitbbs.ac.in
Indian Institute of Technology Bombay	iitb.ac.in
Indian Institute of Technology Delhi	iitd.ac.in
Indian Institute of Technology Gandhinagar	iitgn.ac.in
Indian Institute of Technology Guwahati	iitg.ac.in
Indian Institute of Technology Hyderabad	iith.ac.in
Indian Institute of Technology Indore	iiti.ac.in
Indian Institute of Technology Kanpur	iitk.ac.in
Indian Institute of Technology Kharagpur	iitkgp.ac.in
Indian Institute of Technology Madras	iitm.ac.in
Indian Institute of Technology Mandi	iitmandi.ac.in
Indian Institute of Technology Patna	iitp.ac.in
Indian Institute of Technology Rajasthan	iitj.ac.in
Indian Institute of Technology Roorkee	iitr.ac.in
Indian Institute of Technology Ropar	iitrpr.ac.in

4.2 USE OF APPROPRIATE QUERY SYNTAX

The Webometric analysis is based on the data collected from the Web using various search engines. In each search engine there are some specific search keywords assigned by the search engines to retrieve the required information from the Web.

These specific search keywords along with search syntax have been presented in Table 4.2.

Table 4.2 : Webometric query syntax with results

Search Command	Results	Supported By
domain:~abc	Total number of WebPages	AltaVista, Yahoo!
site:~abc	Total number of WebPages	Google
domain:~abc NOT linkdomain:~abc	Total number of external links or inlinks	AltaVista, Yahoo!
site:~abc NOT linkdomain:~abc	Total number of external links or inlinks	Google
domain:~abc AND linkdomain:~abc	Total number of Self-links	AltaVista, Yahoo!
site:~abc AND linkdomain:~abc	Total number of Self-links	Google
linkdomain:~abc	Total number of links	AltaVista, Yahoo!, Google
site:~abc filetype:pdf	Report total number of pdf files	Google
domain:~abc fileType:pdf	Report total number of pdf files	Yahoo!

Note: ~ *denote space to use various commands in the search engines.*

4.3 DATA COLLECTION

All the domain names were verified to check whether Google and AltaVista support the domain name or not. For each of these domains a search was carried out to determine the total number of links, total Webpages, self-links and in-links using the following commands:

- The total number of web pages of IITs in India to TLD (for example) →
  - *domain*: domain name (for Yahoo)
  - *link*: domain name (for AltaVista)
  - *site*: domain name (for Google)
- The number of total links of IITs at TLD (for example) →
  - *link*: domain name (for Yahoo)
  - *domain*: domain name (for AltaVista)
- The number of in-link can be calculated using the command (for example) →
  - *link*: domain name NOT domain: domain name (for AltaVista)
  - *domain*: domain name NOT link: domain name (for Yahoo)
- The number of self-links can be measured using the command (for example)→
  - *link*: domain name AND domain: domain name (for AltaVista)
  - *domain*: domain name AND link: domain name (for Yahoo)

A series of online snapshot searches over one month (25/10/2012 to 25/11/2012) have been performed on the selected Google, AltaVista and Yahoo search engines by keeping the various search conditions constant. The three search statements that have been used to collect various data for each TLD, SLD and Webpage Second Level Domain (WSLD) or URL may be illustrated with one example from each group. Table 4.3 T1 presents the SLD ac.in under TLD .in (Group-I) and Table 4.3 T2 presents the WSLD iitd.ac.in under SLD ac.in (Group-II).

Table - 4.3 T1: SLD ac.in under TLD .in (Group- I)

domain: ac.in	It will report number of web pages under *ac.in* domain (SLD under .in) included in the AltaVista/Yahoo databases that provide number of webpages.
linkdomain: ac.in	It will report number of web pages in AltaVista/Yahoo databases that provided total number of links to *ac.in* domain (SLD under .in)
(domain: ac.in AND linkdomain: ac.in)	It will report number of web pages under *ac.in* domain (SLD under .in) included in the AltaVista/Yahoo databases that provided hyperlinks i.e. self-link pages
(domain: ac.in NOT linkdomain: ac.in)	It will report number of web pages not under *ac.in* domain but provided hyperlinks to *ac.in* domain (SLD under .in) included in the AltaVista/Yahoo! i.e. external-link pages or inlink.

Table - 4.3 T2: WSLD iitd.ac.in under SLD ac.in (Group-II)

domain: iitd.ac.in	It will report number of web pages under *iitd.ac.in* domain (SLD under .in) included in the AltaVista/Yahoo databases that provide number of webpages.
linkdomain: iitd.ac.in	It will report number of web pages in AltaVista/Yahoo databases that provided total number of links to *iitd.ac.in* domain (SLD under .in).
(domain: iitd.ac.in AND linkdomain: iitd.ac.in)	It will report number of web pages under *iitd.ac.in* domain (SLD under .in) included in the AltaVista/Yahoo databases that provided hyperlinks i.e. self-link pages.
(domain: iitd.ac.in NOT linkdomain: iitd.ac.in)	It will report number of web pages not under *iitd.ac.in* domain but provided hyperlinks to *iitd.ac.in* domain (SLD under .in) included in the AltaVista/Yahoo! i.e. external-link pages.

4.4 RECORDING AND CHECKING

AltaVista reported number of web pages retrieved against each search and these data are recorded in a tabular sheet. The search engine allows visiting a maximum of 200 web pages in 20 sets. Each set contains 10 records. One record or web page selected randomly from each set has been visited at the time of searching. This is done for checking appropriateness of search results.

4.5 CALCULATION OF WEB IMPACT FACTORS

WIF is the web versions of impact factor. There are three types of WIFs: WIF (simple), WIF (self-link) and WIF (in-link). The WIF introduced by Ingwersen (1998) is the ratio of the number of backlinks to a site, divided by the number of webpages at the site, as follows:

A = Total number of webpages to a particular site

B = Number of external backlinks to a given site

C = Number of self-links to a given site

D = Total number of links to a site

Therefore, **WIF (simple) = D/A**; **WIF (in-link) = B/A** and **WIF (self-link) = C/A**

4.6 MEASURING WEB PRESENCE

Web presence can be measured according to several Web-based indicators, some of which include the number of webpages, number of inlinks or external links, number of self-links and the number of total links. The data relating to the web presence of IITs have been retrieved using the webometric query syntax (Table 4.2) as supported by the commercial search engines. WIFs were calculated and reported in order to compare the institutions web influence. Tables 4.6.T1 & T2 presents the various types of WIF calculations related to study websites of IITs in India.

Table 4.6T1: Calculation of WIFs for India (i.e. .in) on 20th November, 2012

	VALUES				RESULTS		
Search Engines	Webpages (A)	inlinks (B)	self-links (C)	total links (D)	WIF (simple) (D/A)	WIF (inlink) (B/A)	WIF (self-link) (C/A)
AltaVista	34100000	33700000	7200	200000	0.005865	0.988269	0.000211
Google	14100000	1760000	1870000	786	0.000055	0.124822	0.132624
Yahoo	34100000	33700000	7200	200000	0.005865	0.988269	0.000211

Note: AltaVista and Yahoo search engine giving same result just because of the AltaVista website is up, and it is stated on the website that they are using the Yahoo search engine.

Table 4.6T1 reflects that India as a whole is having strong value of WIF e.g. 0.988269 through AltaVista search engine. This unexpected result is due to the fact that lower value of webpages as compared to external-links or inlink.

Table 4.6T2: Calculation of WIFs for Indian Academic Web Space (i.e. .ac.in)

	DATA				RESULT		
Search Engines	webpages (A)	inlinks (B)	self-links (C)	total links (D)	WIF (simple) (D/A)	WIF (inlink) (B/A)	WIF (self-link) (C/A)
AltaVista	2450000	141000	1890000	141000	0.057551	0.057551	0.771428
Google	1370000000	1160000000	414000000	1020000	0.000744	0.846715	0.302189

Note: Since AltaVista and Yahoo search engine giving same result that is data has been taken only AltaVista not Yahoo! Search engine in Table 4.5T2.

It has been found from above table 4.6T2 that Google search engine reported very higher result than AltaVista. This unexpected result is due to the fact that higher number of webpages, huge value of inlinks and around 15 times higher value than AltaVista search engines generated.

4.7 CALCULATION OF WIFS FOR IITS BASED ON WIF (INLINK)

The Web Impact Factor (WIF) for each TLD, SLD and WSLD has been calculated at three levels – WIF (simple) by considering all the link pages, WIF (self-link) by considering only self-link pages and WIF (inlink) or WIF (external) by considering only external-link pages. The ranking is done on the basis of WIF (inlink) as it is the true reflection of the degree of impact of the domain spaces on the WWW.

WIF for each TLD, SLD and WSLD (selected for this study) have been calculated on the basis of formula given above in various groups. WIF for each selected web space is calculated in three different ways. These are WIF (simple) - *a ratio of number of total link pages and number of web pages*; WIF (self-link) - *a ratio of number of self link pages and number of web pages* and WIF (inlink) - *a ratio of number of self link pages and number of web pages*. A matrix may represent the calculation of WIF for different web spaces at different levels.

Table 4.7: Calculation of WIFs for IITs based on WIF (Inlink)

	DATA				RESULT	
WSLD	webpages (A)	inlinks (B)	Self-links (C)	Total links (D)	WIF(simple) (D/A)	WIF(inlink) (B/A)
iitbhu.ac.in	131	02	131	00	00	0.015267
iitbbs.ac.in	139	06	139	00	00	0.043165
iitb.ac.in	3630	21500	4400	28	0.007713	5.922865
iitd.ac.in	6390	15600	6600	17	0.002660	2.441314
iitgn.ac.in	153	11	155	00	00	0.071895
iitg.ac.in	12000	03	14500	00	00	0.000250
iith.ac.in	19500	51	20100	00	00	0.002615
iiti.ac.in	370	06	372	02	0.005405	0.016216
iitk.ac.in	3590	10	08	11	0.003064	0.002785
iitkgp.ac.in	10200	05	04	02	0.000196	0.000490
iitm.ac.in	6650	34300	8120	41	0.006165	5.157894
iitmandi.ac.in	109	05	109	02	0.018348	0.045871
iitp.ac.in	141	02	143	04	0.028368	0.014184
iitj.ac.in	43000	13	43400	01	0.000023	0.000302
iitr.ac.in	2390	07	2420	05	0.002092	0.002928
iitrpr.ac.in	121	02	122	02	0.016528	0.016528

***Source*:** AltaVista November 20, 2012 – November 22, 2012.

Table 4.7 reveals that IIT Bombay is having highest WIF (inlink) i.e. 5.922865 due to its least value of webpages with respect to value of inlinks available but IIT Guwahati is having least value of WIF (inlink) i.e. 0.000250, because of very largest value of webpages with respect to available inlinks.

4.8 RANKING OF IITs IN INDIA

There are various approaches for ranking Institutions. The study of ranking helps the reader to compare and identify the websites of IITs in India by their WISER, WIF (in-link) and World Rank. Four main areas of research/content in the field of webometric studies are:

- *Webpage content analysis* - includes automatic categorization of web pages and texts using different search engines and tools for web analysis;

- *Web link structure analysis* - includes the categorization of hyperlinks and in-links, self-links and external links to a particular website, patterns of linking, etc.;

- *Web usage analysis* - which includes the exploitation of log files for users' searching and browsing behavior, and

- *Web technology analysis* - includes the performance of search engines with respect to information retrieval and supporting webometric analysis.

Some of the methods have been explained in detail.

4.8.1 Ranking of IITs through WISER

IIT institute activity is multi-dimensional and this is reflected in its web presence. So the best way to build the ranking is combining a group of indicators that measures these different aspects. Almind & Ingwersen (1997) proposed the first Web indicator, Web Impact Factor (WIF) based on link analysis that combines the number of in-links or external links and the number of pages of the website i.e. Webpages, a ratio of 1:1 between visibility and size. This ratio is used for the ranking but adding two new indicators to the size component: Number of documents, measured from the number of rich files in a web domain, and number of publications being collected by Google scholar database. As it has been already commented, the four indicators were obtained from the quantitative results provided by the main search engines as follows:

- **Size (S)** - number of pages recovered from four engines: Google, Yahoo, Live search and Exalead. For each engine, results are log-normalised to 1 for the highest value. Then for each domain, maximum and minimum results are excluded and every institution is assigned a rank according to the combined sum, as follows:

- **Visibility (V) -** the total number of unique external links received (in-links) by a site can be only confidently obtained from Yahoo search, Live search and Exalead. For each engine, results are log-normalised to 1 for the highest value and then combined to generate the rank.
- **Rich Files (R)** - after evaluation of their relevance to academic and publication activities and considering the volume of the different file formats, the following were selected: Adobe Acrobat (.*pdf*), Adobe PostScript (.*ps*), Microsoft Word (.*doc*) and Microsoft PowerPoint (.*ppt*). These data were extracted using Google and merging the results for each file type after log-normalising in the same way as described before.

- **Scholar (Sc)** - Google scholar provides the number of papers and citations for each academic domain. These results from the Scholar database represent papers, reports and other academic items.

The four ranks were combined according to a formula where each one has a different weight:

Webometrics Rank (position) = 4*RankV + 2*RankS + 1*RankR + 1*RankSc ,

Where, **V**= **Visibility; S**= **Size; R**= **Rich Files and Sc**= **Google Scholar.**

Another formula mentioned below is a modification of the above prescribed in November 2012 by the Webometrics Research Group (www.webometric.info) which has been accessed on 20 November, 2012. Figure 4.8.1T1 presents the WISER Ranking.

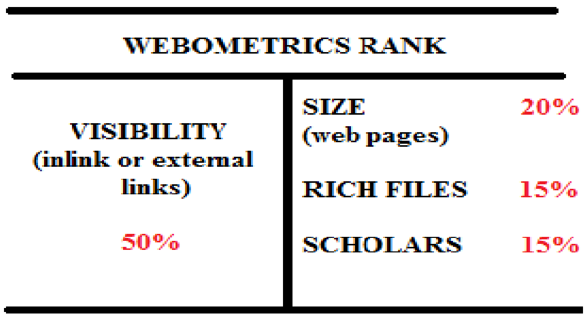

Figure 4.8.1: WISER Ranking (http://www.webometrics.info/en/Methodology)

The above Figure 4.8.1 reveals that the **Webometrics Rank (position) =50% *RankV + 20%*RankS + 15%*RankR + 15%*RankSc.**

Where, **V**= **Visibility; S**= **Size; R**= **Rich Files and Sc**= **Google Scholar.**

Aguillo, et al. (2008) has given the formula for WISER ranking as:

WISER ranking = log (Visibility 50%) + log (Size 20%) + log (Rich files 15%) + log (Scholars 15%).

The volume of contents is measured by the number of pages freely accessible and their visibility by the number of incoming links. The number of rich files is used as an indicator because rich files are preferred formats for scholarly communications. Total number of documents indexed in Google scholar is also considered as an important indicator for scientific publications on the Web. Each web domain is ranked by the linear aggregation of these indicators for their ranking.

The following ranking of institutions has been done based on the formula:

WISER ranking = log (Visibility 50%) + log (Size 20%) + log (Rich files 15%) + log (Scholars 15%) as prescribed by the World Webometrics Group (www.webometrics.info). WISER is Web Indicators for Science, Technology and Innovation Research and it is popular for ranking of academic institutions. Table 4.8.1T2 presents the ranking of IITs based on WISER indicator.

Table 4.8.1T1: Ranking of IITs based on WISER indicator

Name of IIT	webpages (A) **[S]**	inlinks (B) **[V]**	total links (D)	Rich Files [R]					Google Scholar (F) [SC]	WISER Index Value
				DOC	PDF	PS	PPT	TOTAL		
IIT(BHU) Varanasi	131	02	00	02	02	00	02	06	02	0.849665
IIT Bhubaneshwar	139	06	00	01	00	00	00	01	63	2.072689
IIT Bombay	3630	21500	28	54	131	00	53	238	10600	11.646410
IIT Delhi	6390	15600	17	14	128	00	11	153	2880	10.994891
IIT Gandhinagar	153	11	00	02	05	00	00	07	44	3.066817
IIT Guwahati	12000	03	00	02	02	00	00	04	187	4.782386
IIT Hyderabad	19500	51	0	8	23	0	2	33	204	7.177931
IIT Indore	370	06	02	02	06	00	00	08	47	3.273723
IIT Kanpur	3590	10	11	18	28	00	15	61	8130	7.602697
IIT Kharagpur	10200	05	02	05	16	00	03	24	738	6.308020
IIT Madras	6650	34300	41	77	196	00	37	310	6200	11.994051
IIT Mandi	109	05	02	02	02	00	02	06	24	2.246941
IIT Patna	141	02	04	01	02	00	00	03	191	3.338737
IIT Rajasthan	43000	13	01	04	08	00	03	15	42	5.898934
IIT Roorkee	2390	07	05	03	06	00	03	12	36	4.211162
IIT Ropar	121	02	9790	02	02	00	00	04	63	2.137398

4.8.2 Ranking of IITs through WIF (inlink)

Ranking of IITs can be made based on WIF (inlink) indicator. The result is explained (Table- 4.6) where it has been reflected that IIT Bombay become the top position with the WIF(inlink) value 5.922865 and IIT Guwahati is the last place with the WIF (inlink) value 0.000250.

4.8.3 Comparison of Ranking of IITs in India

The comparison of ranking of existing studies being done using WISER, NAAC and WIF (inlink). In NAAC, there is various grading system for ranking the universities based through quality assessment. The latest method is CGPA (Cumulative Grade Point Average) method with 5 point scale, assigned grade A, B, C and D (very good, satisfactory and unsatisfactory respectively). For this study, the comparisons of ranking of Indian Institutes of Technology (IITs) have been done using WISER, WIF (inlink) and World Rank and Table 4.8.3T1 through 4.8.3T3 presents the same.

Table 4.8.3T1: Comparison of Ranking of Indian Institute of Technologies in November, 2012

Name of IIT	Domain	WISER	WIF (inlink)	World Rank
IIT(BHU) Varanasi	iitbhu.ac.in	16	09	8033(8)
IIT Bhubaneshwar	iitbbs.ac.in	15	06	11357(10)
IIT Bombay	iitb.ac.in	02	01	492(1)
IIT Delhi	iitd.ac.in	03	03	890(4)
IIT Gandhinagar	iitgn.ac.in	12	04	12100(12)
IIT Guwahati	iitg.ac.in	08	16	2485(6)
IIT Hyderabad	iith.ac.in	05	13	8042(9)
IIT Indore	iiti.ac.in	11	08	12936(15)
IIT Kanpur	iitk.ac.in	04	12	614(3)
IIT Kharagpur	iitkgp.ac.in	06	14	2045(5)
IIT Madras	iitm.ac.in	01	02	513(2)
IIT Mandi	iitmandi.ac.in	13	05	12678(13)
IIT Patna	iitp.ac.in	10	10	12032(11)
IIT Rajasthan	iitj.ac.in	07	15	14420(16)
IIT Roorkee	iitr.ac.in	09	11	2720(7)
IIT Ropar	iitrpr.ac.in	14	07	12837(14)

Note: WISER= Web Indicators for Science, Technology and Innovation Research and world ranking data are retrieved on November 22, 2012 from http://www.webometrics.info/en/Asia_Pacifico/South%20Asia

The world ranking (Table 4.8.3 T1) implies that all the sixteen IITs are having world rank as mentioned in the parenthesis of each IIT. Table 4.8.3 T2 presents the correlation between ranking of WISER and WIF (inlink).

Table 4.8.3T2: Correlation between ranking of WISER and WIF(inlink)

Name of IIT	WISER (X)	WIF (Inlink) (Y)	Square (X)	Square (Y)	XY	x=(X-Xbar)	y=(Y-Ybar)	xy	Square (x)	Square (y)
IIT(BHU) Varanasi	16	09	256	81	144	+7.5	-0.5	-3.75	56.25	0.25
IIT Bhubaneshwar	15	06	225	36	90	+6.5	-2.5	-16.25	42.25	6.25
IIT Bombay	02	01	04	01	02	-6.5	-7.5	+48.75	42.25	56.25
IIT Delhi	03	03	09	09	09	-5.5	-5.5	+30.25	30.25	30.25
IIT Gandhinagar	12	04	144	16	48	+4.5	-4.5	-20.25	20.25	20.25
IIT Guwahati	08	16	64	256	128	-0.5	+7.5	-3.75	0.25	56.25
IIT Hyderabad	05	13	25	169	65	-3.5	+4.5	-15.75	12.25	20.25
IIT Indore	11	08	121	64	88	+2.5	-0.5	-1.25	6.25	0.25
IIT Kanpur	04	12	16	144	48	-4.5	+4.5	-20.25	20.25	20.25
IIT Kharagpur	06	14	36	196	84	-2.5	+5.5	-13.75	6.25	30.25
IIT Madras	01	02	01	04	02	-7.5	-6.5	+48.75	56.25	42.25
IIT Mandi	13	05	169	25	65	+4.5	-3.5	-15.25	20.25	12.25
IIT Patna	10	10	100	100	100	+1.5	+1.5	+2.25	2.25	2.25
IIT Rajasthan	07	15	49	225	105	-1.5	+6.5	-9.75	2.25	42.25
IIT Roorkee	09	11	81	121	99	-0.5	+2.5	-1.25	0.25	6.25
IIT Ropar	14	07	196	49	98	+5.5	-1.5	-8.25	30.25	2.25
Total	**136**	**136**	**1496**	**1496**	**1175**	**0**	**0**	**-0.50**	**347.75**	**348.00**

Hence, Mean for the variable (X & Y) can be calculated as:

$$\mathbf{Xbar = 1/N \sum_{i=1}^{N} x_i = 1/N(x_1+x_2+^{\cdots\cdots\cdots}+ x_N).}$$

In this case mean (X & Y) are same i.e. **Xbar = Ybar = 8.5**. Standard deviation will be calculated with the help of following formula:

$$\boldsymbol{\sigma}_{\mathbf{x}} \mathbf{= Sqrt} \left[\mathbf{1/N \sum_{i=1}^{N}(X_i - Xbar)^2}\right]$$

Where N=16.

In such a situation, standard deviations of X (i.e. σ_x) & Y (i.e. σ_y) are 4.6620140 and 4.6636895 respectively.

The correlation coefficient relates the strength and direction of linear relationship between two variables. The coefficient of determination represents the percent of the data that is the closest to the line of best fit. Correlation will always between -1.0 and +1.0. If the correlation is positive, we have a positive relationship. If it is negative, the relationship is negative. The coefficient of determination (i.e. r^2) is such that $0< r^2 < 1$, and denotes the strength of the linear association between x and y. The formula can be given as follows:

$$\mathbf{Correlation(r) =} \frac{\mathbf{N\Sigma XY - (\Sigma X)(\Sigma Y)}}{\mathbf{Sqrt\ ([N\Sigma X^2 - (\Sigma X)^2][N\Sigma Y^2 - (\Sigma Y)^2])}} \;;$$

Or

$$\mathbf{r^2 = [COV(X, Y)/\ \sigma_x * \sigma_y] = [(1/N \sum XY - mean(X) * mean(Y))/ \sigma_x * \sigma_y];}$$

Where, N=16; ΣX =136; ΣY = 136; ΣXY =1175; ΣX^2 =1496; ΣY^2 =1496 (For upper one Equation i.e. For **r)**

Or

mean(X) = mean(Y) = 8.5 ; σ_x = 4.6620140 and σ_y = 4.6636895 (For lower one Equation i.e. For $\mathbf{r^2}$)

Therefore, the calculated value of r would be = +0.0558824.

Where, N is the number of pairs of data and r denotes correlation coefficient. σ_x is the standard deviation of X and σ_y standard deviation of Y.

The correlation between WISER ranking and WIF (inlink) is having correlation i.e. +0.0558824 which implied that there is much association or closeness between two ranking methods. In other words, there is a very less difference between two ranking methods. Table 4.8.3T3 presents the reliability of ranking methods in comparison with world ranking for IITs.

Table 4.8.3T3: Reliability of ranking methods in comparison with world ranking for IITs

Name of IIT	Domain	WISER	WIF(inlink)	Inlinks	World Rank
IIT Bombay	iitb.ac.in	02	01	21500(2)	492(1)
IIT Madras	iitm.ac.in	01	02	34300(1)	513(2)
IIT Kanpur	iitk.ac.in	04	12	10(7)	614(3)
IIT Delhi	iitd.ac.in	03	03	15600(3)	890(4)
IIT Kharagpur	iitkgp.ac.in	06	14	5(12)	2045(5)
IIT Guwahati	iitg.ac.in	08	16	3(13)	2485(6)
IIT Roorkee	iitr.ac.in	09	11	7(8)	2720(7)
IIT(BHU)Varanasi	itbhu.ac.in	16	09	2(15)	8033(8)
IIT Hyderabad	iith.ac.in	05	13	51(4)	8042(9)
IIT Bhubaneshwar	iitbbs.ac.in	15	06	6(9)	11357(10)
IIT Patna	iitp.ac.in	10	10	2(16)	12032(11)
IIT Gandhinagar	iitgn.ac.in	12	04	11(6)	12100(12)
IIT Mandi	iitmandi.ac.in	13	05	5(11)	12678(13)
IIT Ropar	iitrpr.ac.in	14	07	2(14)	12837(14)
IIT Indore	iiti.ac.in	11	08	6(10)	12936(15)
IIT Rajasthan	iitj.ac.in	07	15	13(5)	14420(16)

Note: Data source are the same with Table 4.7.3T1. The numbers outside the parenthesis were global ranks, those inside were the country ranks as per (http://www.webometrics.info).

Table 4.8.3T3 clearly indicates that the conditions under which it is possible to evaluate in a reliable and valid way the research strengths of IITs will highly automated procedures, within a reasonable time perspective. So the most crucial question came in mind is one: *how much effort is a reliable evaluation of an entire IIT, and, as a consequence, will such an evaluation be possible for all IITs in India, in a short period of time? Are rankings a reliable means of benchmarking universities against a global standard?*

This finding casts severe doubts on the reliability of this expert-based formula ranking. But, as per the reliability is concerned, WIF(inlink) rank gave much closer value to the world rank, because, IIT Bombay and IIT Madras secured 1st position and 2nd position respectively in WIF(inlink) as well as world rank, whereas, WIF(inlink) position of IIT Rajasthan shows one position better rank than the world ranking i.e. 15th position. It concluded that the WIF (inlink) value is more reliable than other value like WISER.

4.9 MOTIVATION FOR HYPERLINKS

Kim (2000) investigated motivations for creating links in electronic publications in order to find out the relationship between citations and scholarly e-journals. He identified three factors- scholarly, social and technological reasons. Harrison (2002) identified some principles of link creation and proposed a classification of links. Park (2002) conducted a survey of 64 Korean webmasters of commercial websites to assess their motivations for linking to other websites. He found that webmasters were more likely to hyperlink to websites possessing practical content, information or services. Chu (2005) analyzed sample of links and generated list of reasons of hyperlinks.

He found that only 27% of the links were made out of research or teaching motivations. Kousha and Horri (2004) made a survey in Iranian university and found that 63% hyperlinks were made for navigational purpose. Links between UK universities in the field of Mathematics, Physics and Sociology were analyzed. Wilkinson et al. (2003) surveyed 414 links between UK university websites and classified them. They found that less than 1% of hyperlinks targeted formal scholarly articles in journals or conferences; 90% of targeted materials were some way or rather related to scholarly activity. Bar-Ilan (2004) made an academic link studies and included categories for the type of sources and target pages of inter-university links in Israel. He found that 20% links related to research category while Wilkinson et al. found 27% links related to research. Thelwall (2003) surveyed a sample of 100 random inter-site links to UK university homepage and found four types of motivations: ownership, social, general and navigation reasons. Thelwall (2001) made an attempt to distinguish links between research related and non-research oriented.

Chapter-5

FINDINGS, SUGGESTIONS AND CONCLUSION

Webometrics is the quantitative analysis of web phenomena, drawing upon informetric methods, and typically addressing problems related to bibliometrics. Webometric evolved because of the impact of the web to scholarship. It has become a challenge to Indian Institutes of Technology, especially the newly evolving IITs, as they must face the challenges of adopting the web and related technologies for academic activities in a meaningful and systematic manner. Since webometric seeks to categorize IITs to those that have meaningfully adopted the web for research, teaching and learning and those that have not, research into it has become imperative.

This chapter discusses the research findings which are related to the research objectives and hypotheses. Thereafter, conclusions of the study are drawn and some recommendations for improvement of WIF, WISER index value and World ranking and future research are suggested.

5.1 FINDINGS OF THE STUDY

The study yielded useful findings and achieved its objectives. The following inferences are drawn on the analysis of data and presentations made in previous chapters. The main objective of the study is to examine critically the effectiveness and efficiency of the use of web impact factor and to find out the link patterns among the websites of IITs under study. The following sections present the finding on objectives of the study.

Objective 1: To identify and analyze links of websites of Indian Institutes of Technology.

The present study identified sixteen Indian Institute of Technology (IITs) websites in India and all websites of IITs are fully functional. Hence all sixteen IITs websites are taken into consideration for the present study and presented in Chapter 1, Table 1.5.

Objective 2: To investigate relevance of web impact factor (WIF) with reference to Top Level Domain (TLD), Second Level Domain (SLD) and Webpage Second Level Domain (WSLD).

The Web Impact Factor (WIF) for each Top Level Domains (TLDs) of IITs, Second Level Domains (SLDs) related to education and research domain under the TLD for India (i.e .in) and the Universal Resource Locators (URLs) of 16 IITs have been collected by conducting Internet searching using Google, AltaVista and Yahoo search engines.

All the sixteen IITs have their own websites and all websites working under Domain Name System (DNS) ".ac.in". Only IIT Guwahati home page is also working under Domain Name System ".ernet.in" but, same home page is also accessible through DNS ".ac.in". Altavista search engine database retrieve different value of webpages, total links, Self-links and inlinks or external links of various WSLD of different IITs.

Table 4.1T1 through 4.1T3 presents the TLDs, SLDs under .in and URLs of the study websites of IITs in India in Chapter-4.

Objective 3: To calculate the simple Web Impact Factor (WIF), Self-link and in-link or external WIF.

WIF (self-link) by considering only self-link pages and WIF (inlink) or WIF (external) by considering only external-link pages. All search engines database retrieve different value of webpages, total links , Self-links and inlinks or external-links, but, this study selected three popular search engines, such as: AltaVista, Yahoo and Google for finding various WIF (i.e. WIF(simple), WIF(inlink) and WIF(self-link)) and WISER Index value ranking .

AltaVista search engine database taken to finding webpages, inlinks , Self-links, total links and Rich File(R) and in addition to that Google scholar search engine is also used to finding different IITs article for obtaining the value of WISER index value that presented in Chapter 4, Table 4.7.

Objective 4: To compare various ranking approaches among websites of IITs.

For this study, the comparisons of ranking of Indian Institutes of Technology (IITs) have been done using WISER, WIF (inlink) and World Rank and presented in Table 4.8.3T1 through 4.8.3T3.

Objective 5: To compute the correlation between ranking of WISER and WIF(inlink).

For this study, the comparisons of ranking of Indian Institutes of Technology (IITs) have been done using WISER, WIF (inlink) and World rank.

The correlation between WISER ranking and WIF (inlink) is having correlation i.e. +0.0558824 which implied that there is much association or closeness between two ranking methods. In other words, there is a very less difference between two ranking methods and presented in Chapter 4, Table 4.8.3T2.

Objective 6: To rank the IIT websites under study as per WIF, WISER index value, and world rank.

Ranking of sixteen IITs under the study made based on WIF (inlink), WISER index value and World Rank. The result is explained (Table 4.8.3T3) where it is been reflected that IIT Bombay become the top position with the WIF-inlink value (01) and IIT Rajasthan is the last place with the value of WIF-inlink (15) and world rank (16) as per URL: http://repositories. webometrics.info/en/Asia/India.

Hypotheses of the study are tested with the help of the data and related literature and views of the eminent personnel in the field. The following section presents the verification of the hypotheses of the study and the final findings.

Hypothesis 1: The domain structures of websites of IITs institutes in India are homogeneous.

Through the discussion in Chapter 4, Section 4.1 with Table 4.1T1 through 4.1T3 identified the various home pages of IITs.

The above findings clearly indicate that all the study websites work under homogeneous Domain Name System that is **".ac.in"** .

Therefore, based on the above analysis, the study thus ***proves*** first hypothesis.

Hypothesis 2: India as a whole is having strong value of WIF as per measurement of web presence.

Through the discussion in Chapter - 4, Sec. 4.6 and Table 4.6T1 reflects that India as a whole is having strong value of WIF e.g. 0.988269 through AltaVista search engine. This unexpected result is due to the fact that lower value of webpages as compared to external-links or inlink.

Based on the above findings, the study thus ***proves*** the second hypothesis.

Hypothesis 3: Reliability of ranking between WIF (inlinks) and world ranking for IITs are same.

In the present study, higher web impact factor of inlink taken into the consideration for top to bottom ranking of various IITs institute websites that are presented in Chapter 4, Table 4.8.3T1. WIF (inlink) rank, WISER rank and World Rank are compare with each other.

Table 4.8.3T3 reveals that WIF(inlink) rank gave much closer value to the world rank. Hence, the IIT Bombay and IIT Madras secured 1st position and 2nd position respectively in WIF(inlink) as well as world rank, whereas, WIF(inlink) position of IIT Rajasthan shows one position better rank than the world ranking i.e. 15th position.

The reliable and valid way the research strengths of IITs highly automated procedures, within a reasonable time perspective. The findings of the above table clearly indicated that the **WIF (inlink) value much closer value to the world rank.**

Therefore, based on the above findings, the study ***partially proves*** third hypothesis. i.e. the reliability of ranking between WIF (inlinks) and world ranking for IITs are same .

5.2 CONCLUSION

Indian Institute of Technology (IITs) are having a good web presence in general. IIT Rajasthan is having highest number of webpages among various IITs followed by IIT

Hyderabad and IIT Guwahati but IIT Madras got 5th position based on number of webpages. IIT Rajasthan witnessed highest number of self-link counts 43400 webpages and also this institute secured 5th position as per inlinks or external-links count is concerned i.e. 13 webpages but the institute has not been performed well in the table just because of huge number of webpages and also self-links are more in number than webpages i.e. webpages are 43000 and self-links are 43400. IIT Madras is having 4th position as per selflink counts is concerned i.e. 8120 webpage in the table and this institute secured 1rd position as per inlink count is concerned i.e. 34300 webpages but the institute secured 2nd top position due to highest value of WIF (inlink) i.e. 5.157894 but this institute secured top position as per WISER index value is concerned i.e. 11.994051 value in the table. Also, IIT Madras occupied 2nd position with 513 rank in World Rank list. Having World Rank, IIT Bombay occupied top 492 rank among various reputed institutes websites in the world but as per IITs is concerned it occupies top position whereas IIT Rajasthan occupied lowest position i.e. 14420. There is very low correlation i.e. +0.0558824 between WISER Rank and WIF (inlink) for the case of Indian Institute of Technology (IITs). Therefore, volume of webpage is an important indicator for influencing WIF as well as WISER Index Value of any institutions.

5.3 SUGGESTIONS

Web Impact Factor and link analysis of Indian Institute of Technologies is an unexplored area of webometic research. The present study, hopefully, provides a fair idea and information about the website of all the 16 IITs of the country. There is a scope for further improvement in webometric research in this area.

Based on the findings of the study, in this section an attempt has been made to suggest a fewer commendations, which will help to improve the WIF (inlink), WISER index value and World ranking of IIT Websites in India. The main suggestion for improvement of IIT Websites in India with regard to WIF (inlink), as follows:

(i) The WIF (inlinks) of the study reveals that only three IIT Websites having good web presence and they secure highest WIF (inlink) values, such as: IIT Bombay, Delhi and Madras. The other thirteen IIT Websites needs to improve the Web presence by way of least value of their webpages with respect to value of inlinks.

5.4 ISSUES FOR FURTHER RESEARCH

One of the key issues for any future research would be to explore the formalized bridge between hyperlinking and hyperlinked Web sites' authors, serving as social symbols or signs of communication hyper linkage among themselves.

The literature suggests how hyperlinks networks may in some circumstances reflect off-line connections among social actors, and be unique to online interactions in other cases. Further, hyperlink networks among Web sites and social relations in the offline world may be seen as co-constructing each other to some extent, so that offline relationships can influence how online relationships are developed and established. In terms of the development of methods, hyperlink analysis has been able to apply social networks analysis techniques to collections of Web sites and draw conclusions based upon an assumption of actor relationships.

The shortcomings noted in the previous section are not inherent properties of the web or of the web searching tools. We are using a tool for web-link analysis that is not specifically meant for the task. Search engines are designed for contents retrieval and not analysis of web page links. These problems are technical and could be resolved, if the search engine programmers had incentives to work on them. Anyway, the future application of Webometrics may include the electronic databases, which will contain not only published record but also large volume of unpublished data. As the data gathering mechanism is quite easy to follow by using commercial search engines, Webometrics has all the potential to be evolved as a tool for performance evaluation of any web site instantly.

Similarly, research should be conducted along more longitudinal lines; I hope that the present study has succeeded in a starting point for such making an investigation. The study identified certain issues for further research as follows:

(i) Impact of data sources on citation counts and rankings of Indian Institutes of Technology faculty in India: Web of Science versus Scopus and Google Scholar.

(ii) How much information do search engines disclose on the links to a web page? A longitudinal case study of the 'cybermetrics' homepage of IIT websites.

(iii) A Comparison of sources of links for academic web impact factor calculations of IIT websites and IIM Websites.

(iv) Longitudinal trends in IIT web links.

BIBLIOGRAPHY

Abraham, R.H. Webometry: Measuring the complexity of the world wide web. Web. 14 Oct. 2012. <http://www.ralph-abraham.org/vita/redwood/vienna.html>

Adamic, L.A., & Huberman, B.A. "Power-law distribution of the world wide web." *Science* 287(2000):2115a. Print.

Adamic, L.A., & Huberman, B.A. "The web's hidden order." *Communications of the ACM* 44.9(2001): 55–59. Print.

Adamic, L. A., & Huberman, B.A. " Zipf's law and the Internet. " *Glottometrics* 3 (2002): 143-150. Print.

Aguillo, F.I., Ortega, L.J. & Fernandez, M. "Webometric ranking of world universities: Introduction, methodology, and future developments." *Higher Education in Europe* 33.2&33.3 (2008): 233-244. Print.

Aguillo, I. *Cybermetrics'97 (Jerusalem, Israel)* (2002). Web. 10 Oct. 2012. <http://www.cindoc.csic.es/cybermetrics/cybermetrics97.html>

Ajiferuke, I., and D. Wolfram. "Modelling the characteristics of webpage outlinks." *Scientometrics* 59.1 (2004): 43-62. Print.

Albert, R., Jeong, H., & Barabási, A.L. Diameter of the world-wide web. *Nature* 401(1999): 130–131. Print.

Almind, T.C., & Ingwersen, P. “Informetric analysis on the world wide web: Methodological approaches to ‘webometrics’.” *Journal of Documentation* 53.4 (1997): 404-426. Print.

Arroyo, N. "What is the invisible web? A crawler perspective." *Proceedings of the AoIRASIST 2004 Workshop on Web Science Research Methods* (2004). AoIR-ASIST. Web. 10 Oct. 2012. <http://cybermetrics.wlv.ac.uk/AoIRASIST/arroyo.html>.

Babu, B. Ramesh, Jeyshankar, R., & Rao, P.Nageshware. “Websites of central universities in india: A webometrics analysis.” *DESIDOC Journal of Library & Information Technology.* 30.4 (2010): 33-43.Print.

Baeza-Yates, R., and C. Castillo. "Crawling the infinite web: Five levels are enough." *Algorithms and Models for the Web-Graph: Third International Workshop* (2004): 156-67. Print.

Baeza-Yate, R., Castllo, C. and Lopez, V. “Characteristics of the web of spain.” *Cybermetrics* 9(2006): Paper3. Web. 13 Oct. 2012 <http://www.cindoc.csic.es/Cybermetrics/articles/v9ilp3.html>

Bar-Ilan, J. "A microscopic link analysis of academic institutions within a country – the case of israel." *Scientometric* 59.3 (2004): 391-403. Print.

Bar-Ilan, J. “Data collection methods on the web for informetric purposes.” *Scientometrics* 50.1 (2001): 7-32. Print.

Bar-Ilan, J. "Expectations versus reality: Search engine features needed for web research at mid 2005. " *Cybermetrics* 9.1 (2005). Cybermetrics. Issues Contents: Vol. 9 (2005): Paper 2. Expectations versus reality Search eng. Web. 12 Oct. 2012. < http://www.cindoc.csic.es/cybermetrics/articles/v9i1p2.html >

Bar-Ilan, J. "Self-linking and self-linked rates of academic institutions on the web.” *Scientometrics 59*.1 (2004): 29-41.Print.

Bar-Ilan, J. "The use of Web search engines in information science research." In B. Cronin (Ed.), *Annual Review of Information Science and Technology* 38 (2004): 231-288. Print.

Bar-Ilan, J. "What do we know about links and linking? A framework for studying links in academic environments." *Information Processing and Management 41*.4 (2005): 973-986. Print.

Bar-Ilan, J., & Echerman, A. "The anthrax scare and the web: A content analysis of web pages linking to resources on anthrax." *Scientometrics* 63.3 (2005): 443-462. Print.

Bar-Ilan, J., & Peritz, B. "Evolution, continuity, and disappearance of documents on a specific topic on the web: A longitudinal study of "informetrics." *Journal of the American Society for Information Science and Technology* 55.11 (2002): 980-90. Print.

Bar-Ilan, J., & Peritz, B. "School of library, archive, and information studies" *Library Trends* 50.3(2002): 371-392. Print.

Bar-Ilan, J., & Peritz, B. The lifespan of "informetrics" on the web: An eight year study (1998-2006). In D. Torres-Salinas, & H. F. Moed (Eds.), *Proceedings of the 11th International Conference of the International Society for Scientometrics and Informetrics 1* (2007): 52-62. Print.

Barabasi, A. L. Linked: The new science of networks. Cambridge, *MA:* Perseus, 2002. Print.

Barjak, F. The role of the internet in informal scholarly communication. *Journal of the American Society for Information Science and Technology*, *57*.10(2006), 1350 -1367. Print.

BBC News. *US duo in first spam*. BBC. 2005. Web. 11 Oct. 2012. <http://news.bbc.co.uk/1/hi/technology/3981099.stm>

Beaulieu, A., & Simakova, E. Textured connectivity: An ethnographic approach to understanding the timescape of hyperlinks. *Cybermetrics 10*.1(2006). Web. 12 Oct. 2012. < http://www.cindoc.csic.es/cybermetrics/articles/v10i1p5.html>

Berners-Lee, T., & Cailliau, M. Weaving the Web: The past, present and future of the world wide web by its inventor. London: Orion Business Books, 1990. Print.

Berners-Lee, T., & Cailliau, R. World wide web: Proposal for a hypertext project. Web. 11 Oct. 2012. < http://www.w3.org/Proposal.html>

Bjorneborn, L. Small-world link structures across an academic web space: A library and information science approach. Ph.D. Thesis. Copenhagen : Royal School of Library and Information Science, 2004. Print.

Bjorneborn, L. & Ingwersen, P. Handbook of quantitative science and technology research. Dordrecht: Kluwer Academic Publishers, 2001. Print.

Bjorneborn, L., & Ingwersen, P. Perspectives of webometrics. *Scientometrics*, 50.1 (2001): 65-82. Print.

Bjorneborne, L & Ingwersen, P. “Toward a basic framework for webometrics.” *J. Ameri. Soc. Inform. Sci. Technol.* 55.4 (2004): 1216–27. Print.

Borgman, C. L., & Furner, J. “Scholarly communication and bibliometrics.” *Annual Review of Information Science & Technology* 36 (2002): 372-376. Print.

Borrull, A. L., & Oppenheim, C. Legal aspects of the web. *Annual Review of Information Science & Technology 38*.1 (2004): 483-548. Print.

Bossy, M. J. *The last of the litter: “Netometrics*. 1995. Web. 13 Oct. 2012. <http://biblio-fr.info.unicaen.fr/bnum/jelec/Solaris/d02/2bossy.html>

Bradford, S.C. “Sources of information on specific subjects.” *British Journal of Engineering* 137(1934): 85–86. Print.

Brown, C. The Mathew effect of the annual reviews series and the flow of scientific communication through the world wide web. *Scientometrics 60*.1(2004): 25-36. Print.

Butcher, J., & Jeffrey, P. The use of bibliometric indicators to explore industry-academia collaboration trends over time in the field of membrane use for water treatment. *Technovation 25*(2005): 1273-1280. Print.

Calluzzo, V. J., & Cante, C. J. Ethics in information technology and software use. *Journal of Business Ethics 51*.3(2004): 301-312. Print.

Carey, P. *Data protection: A practical guide to UK and EU law*. Oxford: Oxford University Press, 2004. Print.

Cassiman, B., Glenisson, P., & Van Looy, B. Measuring the industry-science links through inventor-author relations: A profiling methodology. *Scientometrics 70*.2 (2007): 379-391. Print.

Chakrabarti, S., Joshi, M.M., Punera, K., & Pennock, D.M. "The structure of broad topics on the web." *Proceedings of the WWW2002 Conference* Web.13 Oct. 2012. <http://www2002.org/CDROM/refereed/338/>

Chau, M., Shiu, B., Chan, I., & Chen, H. Redips: Backlink search and analysis on the web for business intelligence analysis. *Journal of the American Society for Information Science and Technology 58*.3(2007): 351-365. Print.

Chen, C., Newman, J., Newman, R., & Rada, R. How did university departments interweave the web: A study of connectivity and underlying factors. *Interacting with Computers* 10.4(1998): 353-373. Print.

Chu, H. “Taxonomy of inlinked web entities: what does it imply for webometric research?” *Library and Information Science Research* 27.1(2005): 8-27. Print.

Cothey, V., Aguillo, I., & Arroyo, N. Operationalising websites: lexically, semantically or topologically? *Cybermetrics, 10*.1(2006). Web. 14 Oct. 2012. <http://www.cindoc.csic.es/cybermetrics/articles/v10i1p3.html>

Crane, D. Invisible colleges. Diffusion of knowledge in scientific communication. Chicago: The university of Chicago Press, 1972. Print.

Cronin, B. “Bibliometrics and beyond: Some thoughts on web-based citation analysis.” *Journal of Information Science*, 27.1(2001):1–7. Print.

Cronin, B., & McKim, G. “Science and scholarship on the world wide web: A North American perspective.” *Journal of Documentation* 52.2(1996): 163-171. Print.

Cronin, B., Snyder, H. W., Rosenbaum, H., Martinson, A., & Callahan, E. “Invoked on the web.” *Journal of the American Society for Information Science* 49.14 (1998): 1319 -1328. Print.

Davenport, E., & Cronin, B. The citation network as a prototype for representing trust in virtual environments. In B. Cronin & H. B. Atkins (Eds.), *The web of knowledge: A Festschrift in honor of Eugene Garfield* (2000): 517-534. Medford, NJ: Information Today Inc., 2000. Print.

Davenport, E., & Cronin, B. The Web of Knowledge: A Festchrift in Honor of Eugene Garfield. Medford, NJ: Information Technology Inc., 1990. Print.

Ding, C., Zha, H., He, X., Husbands, P., & Simon, H. “Link analysis: Hubs and authorities on the world wide web.” *LBNL Tech Report 47847*. Web. 15 Oct. 2012. <http://crd.lbl.gov/~cding/papers/hits5.pdf>

Downie,J.S. “Informetrics and the world wide web: A case study and discussion.” *Proceedings of the 24th Annual Conference of the Canadian Association for Information Science* (1996): 130–141. Print.

DTI. *Innovation* (2007a). Web. 13 Oct. 2012. <http://www.dti.gov.uk/innovation/index.html>

Edgerton, D. The linear model did not exist. In K. Grandin, N. Worms, & S. Widmalm (Eds.), *The Science-Industry Nexus: History, Policy, Implications* Sagamore Beach: Science History Publications, 2004. Print.

Efe, K., Raghavan, V., Chu, C.H., Broadwater, A.L., Bolelli, L., & Ertekin, S. "The shape of the web and its implications for searching the web." *Proceedings of the International Conference on the Advances in Infrastructure for Electronic Business, Science, and Education on the Internet, L'Aquila, Italy, July 31–Aug 6, 2000.* Web. 15 Oct. 2012. <http://citeseer.nj.nec.com/317732.html>

Etzkowitz, H., & Zhou, C. "Triple helix twins: Innovation and sustainability." *Science and Public Policy, 33*.1 (2006): 77-83. Print.

Fagerberg, J. "Innovation: A guide to the literature." In F. Fagerberg, D. C. Mowery, & R. R. Nelson (Eds.), *The Oxford Handbook of Innovation.* Oxford: OUP, 2005. Print.

Fleck, J. "The structure of technological evolutions: Linear models, configurations, and systems of development." In K. Grandin, N. Wormbs, & S. Widmalm (Eds.), *The Science-Industry Nexus: History, Policy, Implications* . Canton, MA: Science History Publications, 2004. Print.

Garfield, E. Citation indexes for science: A new dimension in documentation through association of ideas. *Science* 122(1955): 108–111. Print.

Gau, Y., & Vaughan, L. Web hyperlink profiles of news sites: A comparison of newspapers of USA, Canada, and China. *Aslib Proceedings: New Information Perspectives* 57.5(2005): 398-411. Print.

Glanzel, W. “National characteristics in international scientific co-authorship” *Scientometrics*, 51.1(2001):69-115. <http://www.steunpuntoos.be/WG_Papers/Scientometrics_51_1_69.pdf>

Glanzel, W., & Schubert, A. “Analysing scientific networks through co-authorship.” In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* (pp.257-276). Dordrecht: Kluwer Academic Publishers, 2004. Print.

Godin, B. “The Linear Model of Innovation: The historical construction of an analytical framework.” *Project on the History and Sociology of S&T Statistics. Working paper No.30.* 2005. Web. 14 Oct. 2012. < http://www.csiic.ca/PDF/Godin_30.pdf>

Goodrum, A. “Terrorism or civil disobedience: Toward a hacktivist ethic.” *Computers and Society* 30.2 (2001): 14-19. Print.

Goodrum, A. , McCain, K. W., Lawrence, S., & Giles, C. L. “Scholarly publishing in the internet age: A citation analysis of computer science literature.” *Information Processing & Management* 37.5(2001): 661-676. Print.

Google Librarian Central. *Back from WebSearch University*. 2007. Web 14 Oct. 2012. <http://librariancentral.blogspot.com/2007/05/back-from-websearch-university.html>

Gopal, R. D., Sanders, G. L., Bhattacharjee, S., Agrawal, M., & Wagner, S. C. "A behavioural model of digital music piracy." *Journal of Organisational Computing and Electronic Commerce 14*.2(2004): 89-105. Print.

Goswami, P. *Webometric study of universities websites of Uttar Pradesh* MPHIL Dissertation. Ujjain: Vikram University, 2007. Print.

Goswami, P. , Sharma, U. and Shukla, A. K. "The webometrics" *International CALIBER-2008* Ahmedabad: Inflibnet Center , 2008. Print.

Grupp, H., & Mogee, M. E. "Indicators for national science and technology policy: Their development, use and possible misuse." In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* Dordrecht: Kluwer Academic Publishers., 2004. Print.

Harries, G., Wilkinson, D., Price, L., Fairclough, R., & Thelwall, M. "Hyperlinks as a data source for science mapping." *Journal of Information Science 30*.5(2004):236-447. Print.

Harrison, C. "Hypertext links: Whither Thou Guest and Why." *First Monday* 7.10(2002). Web. 24 Nov. 2012. <http://firstmonday.org/issues/issue7_10/harrison/index.html>

Harter, S. P., & Ford, C. E. “Web-base analyses of E-journal impact: Approaches, problems, and issues.” *Journal of the American Society for Information Science* 51.13 (2000): 1159-1176. Print.

HEFCE. *Response to consultation on successor to research assessment exercise.* 2006. Web. 15 Oct. 2012. < http://www.hefce.ac.uk/News/HEFCE/2006/rae.htm>

Heimeriks, G., & van den Besselaar, P. Analyzing hyperlinks networks: The mean of hyperlink based indicators of knowledge production. *Cybermetrics 10*.1(2006). Web. 15 Oct. 2012. <http://www.cindoc.csic.es/cybermetrics/articles/v10i1p1.html>

HERO. “Guidance to panel chairs and members: Criteria and working methods” *Research Assessment Exercise 2001* (2001). Web. 10 Oct. 2012. <http://195.194.167.103/PanGuide/ >

Henzinger, M. “Hyperlink analysis for the web.” *IEEE* Internet Computing, 5.1(2001): 45-50. Print.

Hertzel,D.H. History of the development of ideas in bibliometrics. Encyclopedia of Library and Information Science. New York: Marcel Dekker, 1987. Print.

Hinze, S., & Schmoch, U. Opening the black box: Analytical approaches and their impact on the outcome of statistical patent analysis. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* Dordrecht: Kluwer Academic Publishers, 2004. Print.

Hjortgaard C. F., Ingwersen, P. & Wormell, I. "Online determination of the journal impact factor and its international properties." *Scientometrics* 40.3(1997): 529-540. Print.

Huberman, B. A. The laws of the web: Patterns in the ecology of information. Cambridge, MA: MIT Press, 2001. Print.

Huberman, B.A. "The web's hidden order." *Communications of the ACM* 44.9 (2001): 55-59. Print.

Hurd, J.M. "The transformation of scientific communication: A model for 2020." Journal *of the American Society for Information Science* 51.14(2000): 1279–1283. Print.

Ingwersen, P. "The calculation of web impact factors." *Journal of Documentation* 54.2 (1998): 236-43. Print.

Ingwersen, P., & Björneborn, L. "Methodological issues of webometric studies." In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research*. Dordrecht: Kluwer Academic Publishers, 2004. Print.

Jalal S. K., Biswas S.C., and Mukhopadhyay P. S. “Bibliometrics to webometrics.” *Information Studies* 15.1 (2009): 3-20. Print.

Jalal S. K., Biswas S.C., and Mukhopadhyay P. S. “Web impact factor and link analysis of selected Indian universities.” *Annals of Library and Information Studies* 57.2 (2010b): 109-121. Print.

Jalal S. K., Biswas S.C., and Mukhopadhyay P. S. “Webometric Analysis of central universities in India: A study. *IEEExplore*. Web. 20 Oct. 2012 <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5402605>

Jeyshankar, R. & Babu, B. Ramesh. “Websites of universities in Tamil Nadu: a webometric study.” *Annals of Library and Information Studies* 56.2 (2009):69-79. Print.

Jeyshankar, R., Babu, B. Ramesh, & Gopalakrishnan, S. “Basic Frame Work of Webometric Study: A Study.” *KELPRO Bulletin* 13.1 (2009): 41- 48. Print.

Jeyshankar, R., Sujitha, I. Maria, and Valarmathi A. **“**Web pages of ICMR institutes websites: A webometric analysis.” *Global Advanced Research Journal of Library, Information and Archival Studies* 1.1 (2012): 6-18. Print.

Kessler, M.M. “Bibliographic coupling between scientific papers.” *American Documentation* 14.1(1963): 10–25. Print.

Kim, H.J. “Motivations for hyper-linking in scholarly electronic articles: A qualitative study.” *JASIS* 51.10(2000): 887-99. Print.

Kling, R. & Callahan, E. “Electronic journals, the Inteinet, and scholarly communication”. *Annual Review of Information Science and Technology*, 37 (2003). Print.

Kling, R., & McKim, G. “Not just a matter of time: Field differences and the shaping of electronic media in supporting scientific communication.” *Journal of the American Society for Information Science* 51.10 (2000): 1306-1320. Print.

Kousha, K., and Horri, A. “The relationship between scholarly publishing and the counts of academic inlinks to Iranian university web sites: Exploring academic link creation motivations.” *Journal of Information Management and Scientometrics*, 1.2(2004): 13-22. Print.

Krippendorff, K. *Content Analysis: an introduction to its methodology* (2nd ed.). London: Sage, 2004. Print.

Kuster, R.J. “A bibliometric study of the remote hypertext links in public library world wide web sites.” *Proceedings of the ASIS Mid-Year Meeting, San Diego, California* (1996): 338–343. Print.

Larson, R. R. “Bibliometrics of the world wide web: An exploratory analysis of the intellectual structure of cyber space." *Proceedings of the 59th annual meeting, ASIS* Ed. S. Hardin. Baltimore: Learned Information Ltd. (1996). 71-79. Print.

Lawrence, S., & Giles, C. L. “Searching the world wide web.” *Science* 280.5 (2001): 98-100. Print.

Leydesdorff, L. & Curran, H. “The dynamics of innovation: from National Systems and “Mode 2” to a Triple Helix of university-industry government relations.” *Research Policy* 29.2(2000): 109-123. Print.

Leydesdorff, L., & Curran, M. “Mapping university-industry-government relations on the internet: An Exploration of Indicators for a Knowledge-Based Economy.” *Cybermetrics* 4.1(2000): Paper 2 at <http://www.cindoc.csic.es/cybermetrics/articles/v4i1p2.html>.

Lotka, A.J. “The frequency distribution of scientific productivity.” *Journal of the Washington Academy of Sciences*, 16(1926): 317–323. Print.

McKiernan, G. CitedSites(sm): Citation indexing of web resources. Web. 15 Oct. 2012. <http://www.public.iastate.edu/~CYBERSTACKS/Cited.htm>

Menczer, F. “Growing and navigating the small world web by local content.” *Proceedings of the National Academy of Sciences*, 99.22 (2002): 14014–14019. Print.

Meyer, M. “Patents cited in the scientific literature: An exploratory study of ‘reverse’ citation relations.” *Scientometrics* 58.2 (1999): 415-428. Print.

Middleton, I., McConnell, M., & Davidson, G. “Presenting a model for the structure and content of a university world wide web site.” *Journal of Information Science* 25.3 (1999): 219-227. Print.

Miles-Board, T., Kampa, S., , Carr, L., and Hall, W. Linking with Meaning: Ontological Hypertext for Scholars, Technical Report ECSTR-IAM01-005 Southampton: Electronics and Computer Science Department, University of Southampton, 2001. Print.

Moulthrop, S., & Kaplan, N. “Citescapes: Supporting knowledge construction on the web.” *Poster at the WWW4 Conference*. (1995). Web. 15 Oct. 2012. <http://iat.ubalt.edu/moulthrop/essays/citescapes/citescapes.html>

Park, H.W. “Examining the determinants of who is hyperlinked to whom: a survey of webmasters in Korea.” *First Monday* 7.11(2002). Web. 24 Nov. 2012. <http://firstmonday.org/issues/issue7_11/park/>

Park, H. W., & Thelwall, M. “Hyperlink analysis: Between networks and indicators.” *Journal of Computer-Mediated Communication* 8.4(2003). Web. 11 Oct. 2012. <http://www.ascusc.orgtjcmclvol8/issue4/park.html>

Pitkow,J.E.,& Pirolli,P. “Life, death, and lawfulness on the electronic frontier.” *CHI 97 Electronic Publications.* Web. 15 Oct. 2012. <http://www.acm.org/sigchi/chi97/proceedings/paper/jp-www.htm>

Polanco, X., Boudourides, M., Besagni, D., & Roche, I. “Clustering and mapping European university web sites sample for displaying associations and visualizing networks” *NTTS&ETK 2001 Conference* (2001). Print.

Price,D. de Solla. “Networks of scientific papers.” In M. Kochen (Ed.), The growth of knowledge: Readings on organization and retrieval of information (pp. 145–155). New York: Wiley, 1965. Print.

Ratha, B., Joshi, L., and Naidu, G.H.S. “Webometric Study of IIT Libraries Websites.” *DESIDOC Journal of Library & Information Technology* 32.3 (2012): 249-254. Print.

Rodríguez, I. & Gairín, J.M. "Valorando el impacto de la información en internet: AltaVista, el Citation Index de la Red" [Impact assessment of information on the Internet: AltaVista, the Citation Index of the Web]. *Revista Española De Documentación Científica* 20.2(1997): 175-181. Print.

Rousseau, R. "Sitations: an exploratory study." *Cybermetrics* 1.1 (1997). Web. 17 Oct. 2012 <http://www.cindoc.csic.es/cybermetrics/articles/v1i1p1.pdf >

Sangam, S.L. and Munavalli, S.B. "IT literature: A webometrics study." *Library progress (international)* 24 .1(2004): 15-21. Web. 17 Oct. 2012. <http://eprints.rclis.org/archive/00006264/>

Shukla S. H., and Poluru L. "Webometric analysis and indicators of selected indian state universities." *Information Studies* 18.2 (2012): 79-104. Print.

Small, H. "Co-citation in the scientific literature: A new measure of the relationship between two documents." *Journal of the American Society for Information Science,* 24.4(1973): 265–269. Print.

Smith, A. G. "A tale of two web spaces." *Journal of Documentation* 55.5(1999): 577-592. Print.

Smith, A., & Thelwall, M. “Web impact factor for Australasian universities.” *Scientometrics* 54.3(2002): 363-380. Print.

Snyder, H., & Rosenbaum, H. “Can search engines be used as tools for web link analysis? A critical view.” *Journal of Documentation* 55.4 (1999): 375-384. Print.

Spertus, E. “ParaSite: Mining structural information on the Web.” *WWW6 Conference* Web. 15 Oct. 2012. <http://decWeb.ethz.ch/WWW6/Technical/Paper206/Paper206.html>

Thanuskodi s. “A webometric analysis of selected Institutes of national importance websites in india.” *International Journal of Library Science* 1.1 (2012): 13-18. Print.

Thelwall, M. “A comparison of sources of links for academic web impact factor calculations.” *Journal of Documentation* 58.1 (2002a): 66-78. Print.

Thelwall, M. “A free database of university web links: Data collection issues.”*Cybermetrics* 6/7.1(2003a). Web. 20 Oct. 2012. <http://www.cindoc.csic.es/cybermetrics/articles/v6i1p2.html>

Thelwall, M. A publicly accessible database of UK university website links and a discussion of the need for human intervention in web crawling. (2001e). Web. 10 Oct. 2012 <http://www.scit.wlv.ac.uk/~cm1993/papers/a_publicly_accessible_database.pdf>

Thelwall, M. “A web crawler designs for data mining.” *Journal of Information Science* 27.5 (2001a): 319-325. Print.

Thelwall, M. “An initial exploration of the link relationship between UK university websites.” *ASLIB Proceedings* 54.2 (2002b): 118-126. Print.

Thelwall, M. “Commercial web site links.” *Internet Research* 11.2(2001b): 114-124. Print.

Thelwall, M. “Conceptualizing the documentations on the web: An evaluation of different heuristic-based models for counting links between university web sites.” *Journal of the American Society for Information Science and Technology* 53.12 (2002d): 995 -1005. Print.

Thelwall, M. “Exploring the link structure of the Web with network diagrams.” *Journal of Information Science* 27.6(2001c): 393-401. Print.

Thelwall, M. “Extracting macroscopic information from web links.” *Journal of the American Society for Information Science and Technology* 52.13 (2001d): 1157-1168. Print.

Thelwall, M. “Methodologies for crawler based web surveys.” *Internet Research: Electronic Networking and Applications* 12.2 (2002f): 124 -38. Print.

Thelwall, M. “Methods for reporting on the targets of links from national systems of university Web sites.” *Information Processing and Management*. (2003a). Print.

Thelwall, M. “Results from a web impact factor crawler.” *Journal of Documentation* 57.2(2001e): 177-191. Print.

Thelwall, M. “The top 100 linked pages on UK university web sites: High inlink counts are not usually directly associated with quality scholarly content.” *Journal of Information Science, 28*.6(2002a): 485-493. Print.

Thelwall, M. “Web impact factors and search engine coverage” *Journal of Documentation* 56.2 (2000): 185-189. Print.

Thelwall, M. “What is this link doing here? Beginning a fine-grained process of identifying reasons for academic hyperlink creation.” *Information Research* 8.3(2003d). Web. 21 Oct. 2012. <http://informationr.net/ir/8-3/paper151.html >

Thelwall, M & Chu, H. “Search engines for the world wide web: A comparative study and evaluation methodology.” *Proceedings of the Annual Meeting of the American Society for Informaiton Science* 33(2002): 127-135. Print.

Thelwall, M., Harries, G., & Wilkinson, D. “Why do web sites from different academic subjects interlink?” *Journal of Information Science* 29.6 (2003): 453-71. Print.

Thelwall, M. and Sud, P. “A comparison of methods for collecting web citation data for academic organizations.” *Journal of the American Society for Information Science and Technology* 62.8 (2011): 1488-1497. Print.

Thelwall, M., & Tang, R. “Disciplinary and linguistic considerations for academic web linking: An exploratory hyperlink mediated study with Mainland, China and Taiwan.” *Scientometrics* 58.1 (2003):153 -179. Print.

Thelwall, M., & Wilkinson, D. “Graph structure in three national academic Webs: Power laws with anomalies.” *Journal of the American Society for Information Science and Technology* 54.8 (2003a): 706-712. Print.

Thelwall, M., & Wilkinson, D. “Three target document range metrics for university Web sites.” *Journal of the American Society for Information Science and Technology* 54.6(2003): 490–497. Print.

Thelwall, M., Vaughan, L. & Bjorneborn, L. “Webometrics”. In: *Annual Review of Information Science and Technology* 39(2005): 81-135. Print.

Thomas, O., & Willett, P. “Webometric analysis of departments of librarianship and information science.” *Journal of Information Science* 26.6(2000): 421-428. Print.

Vaughan, L., & Shaw, D. “Web citation data for impact assessment: A comparison of four science disciplines.” *Journal of the American Society for Information Science and Technology* 56.10(2005): 1075–1087. Print.

Vaughan, L. & Thelwall, M. "Scholarly use of the web: What are the key inducers of links to journal web sites?" *Journal of the American Society for Information Science and Technology* 54.1 (2003): 29-38. Print.

Vijayakumar M., B.U. Kannappanavar, and Santosh Kumar K.T. “Webometric analysis of web presence and links of SAARC countries.” *DESIDOC Journal of Library & Information Technology* 32.1(2012): 70-76. Print.

Weiss,R., Vélez,B.,Sheldon,M.A.,Namprempre,C.,Szilagyi,P .,Duda, A., et al. “HyPursuit: A hierarchical network search engine that exploits content-link hypertext clustering.” *Proceedings of the Seventh* ACM Conference on Hypertext and Hypermedia (1996): 180–193. Print.

White, H.D., & McCain, K.W. “Bibliometrics.” *Annual Review of Information Science and Technology* 24(1989): 119–186. Print.

Wilkinson, D., Harries, G., Thelwall, M., & Price, L. “Motivations for academic Web site interlinking: Evidence for the web as a novel source of information on informal scholarly communication.” *Journal of Information Science* 29.1(2003): 49–56. Print.

Wilkinson, D., Thelwall, M. and Li, X. “Exploiting hyperlinks to study academic web use.” *Social Science Computer Review* 21.3(2003): 340-351. Print.

Zhang, Y. “Graduate school of library and information science.” *Library Trends* 47.4(1999): 746-770. Print.

Zhang, Y. “Scholarly use of internet-based electronic resources.” *Journal of the American Society for Information Science and Technology* 52.8(2001): 628–654. Print.

Zipf, G.K. *Human behavior and the principle of least effort: An introduction to human ecology*. Cambridge, MA: Addison-Wesley, 1949. Print.

Appendix-I

CHECKLIST

1. DIRECTIONS

- ***What is the URL or Web address of the Web site you are evaluating?***

  http://

- ***What is the title of the Web site?***

2. AUTHORITY AND ACCURACY

Anyone who knows a little HTML coding and has access to a server can create and load a Web site. It is important to find out who the author is and what are the author's qualifications or expertise, in order to determine the credibility and reliability of the information.

- ***Who is the author of the Web site?***

☐ **I couldn't tell.**

☐ **The author is:**

- **What authorship clues did the URL (Web address) provide?** *Check all that apply:*

☐ **company(.com)**

☐ **academic institution (.edu)**

☐ **U.S. Government agency (.gov)**

☐ **U.S. military site (.mil)**

☐ **network of computers (.net)**

- ☐ **non-profit organization (.org)**
- ☐ **country-sponsored site (e.g., .in)**
- ☐ **personal Web page (e.g., http://www.jamieoliver.com/)**
- ☐ **Other? Please describe:**

❖ ***What are the qualifications of the author or group that created the site?***

- ☐ **I couldn't find this information.**
- ☐ **The author's qualifications are:**

3. PURPOSE AND CONTENT

Determine the purpose of the Web site by looking closely at the content of the information. Some sites provide links to information (e.g., About Our Organization or a Mission Statement) detailing the purpose in creating the Web site, while the purpose of others might not be obvious at first. Take the time to thoroughly explore a Web site to determine if the information is mostly subjective (biased or opinionated), objective (factual) or mixed.

❖ ***What is the purpose of the Web page or site?*** Check all that apply:

- ☐ **a personal Web page.**
- ☐ **a company or organization Web site.**
- ☐ **a forum for educational/public service information.**
- ☐ **a forum for scholarly/research information.**
- ☐ **for entertainment.**
- ☐ **an advertisement or electronic commerce.**
- ☐ **a forum for ideas, opinions or points of view.**

☐ **Other? Please explain:** [ ]

- ***In your own words, briefly describe the purpose of the Web site.***
- ***What does the Web site provide?*** *Check one:*

☐ **Balanced, objective or factual information.**

☐ **Biased, subjective or opinionated statements.**

(Are the arguments well supported? ☐ **Yes.** ☐ **No.)**

☐ **Both objective and subjective information.**

☐ **I couldn't tell.**

☐ **Other? Please explain:** [ ]

- ***Does the Web site provide any contact information or means of communicating with the author or Webmaster?***

☐ **No.**

☐ **Yes, the site provides:** [ ]

4. CURRENCY

The currency or regularity of updating information is vital for some types of Web sites, and not so important for others. For example, Web sites that provide historical information, such as the presidential papers of George Washington, do not have to be updated as often as sites that provide news stories or stock market information.

- ***When was the Web site last revised, modified or updated?***

☐ **I couldn't tell.**

☐ **It was updated:** [ ]

❖ ***Is currency important to the type of information available on this Web site?***

☐ **Yes.** *Please explain:*

☐ **No.** *Please explain:*

❖ ***Is the site well-maintained?***

☐ **I couldn't tell.**

☐ **Yes.**

☐ **No.**

❖ ***Are links broken (Error 404 messages)?***

☐ **I couldn't tell.**

☐ **Yes.**

☐ **No.**

5. DESIGN, ORGANIZATION AND EASE OF USE

Design, organization and ease of use are important considerations. Web sites can provide useful sources of information. However, if they are slow to load or difficult to navigate, search and read, then their contribution and usefulness will be lost.

In your opinion, how does the Web site appear overall? *Check all that apply:*

☐ **Well designed and organized.**

☐ **Poorly designed and organized.**

☐ **Easy to read and navigate.**

- ☐ **Difficult to read and navigate.**
- ☐ **HELP screens are available.**
- ☐ **HELP screens are unavailable.**
- ☐ **A search feature/site map is available.**
- ☐ **A search feature/site map is unavailable.**