CHAPTER III

RESEARCH METHODOLOGY

CHAPTER III

RESEARCH METHODOLOGY

3.0 Meaning and Significance of Research

The research is a scientific and systematic approach or investigation for researching the searched, for gaining new insights and knowledge. Research is an inquiry made for the finding out the answer to the questions with regards to the research problem being searched for the answer. The progress of human society is based on the generation of new knowledge. Thus, the research enhances existing knowledge stock. The selection of most suitable research methodology for a studying a problem in a systematic manner is very important and most indispensable step. The term research is very specific in nature which involves various levels and must be used with care. There are various types of research in which it could be classified. Some of the important categories, are (i) Analytical and Descriptive, (ii) Fundamental and Applied Research, (iii) Qualitative and Quantitative. In the light of these categories, we may identify the two approaches to the research viz., qualitative and quantitative, which can be grouped further based on the level of data analysis.

3.1 Qualitative and Quantitative Research

Qualitative research is widely used in the Social sciences. The qualitative research is concerned about, the cognitive and behavioural aspects of the living organism. In library and information science, we used to study the user's behavioural aspects of information gathering, their reading habits, use of learning material etc. The methods for data collection are observations, interviews, survey, and questionnaire etc. The core purpose of qualitative research is to find out the motive of human behaviour. Quantitative research aims to quantify the measurable effect of the causal relation of variables. Statistical inferences help to analyse the data on multiple statistical parameters. However, research methods like observations and questionnaires can yield quantitative and qualitative data as well. The results can be statistically tested. Data can be validated mathematically. In order to accomplish the research topic **"Global Knowledge Diffusion in Vocational Education and Training: A Scientometric Study"** major issues related to the

comprehensive accomplishment of the above study have been focused and analysed. To accomplish the study quantitative research method is decided to use.

3.2 Application of Documentary Research Method

The most of the social science researchers are tend to adopt the traditional survey method for collection of data (more rely on the survey, interviews, questionnaires, observations, etc. but these methods are not only the ones and the most reliable, cost effective and convenient). Therefore, many times social scientists and researchers overlooked the importance of documentary sources for data collection (**Mogalakwe, M., 2009**).

When the tacit knowledge is transformed into the documented form it becomes explicit. Nowadays, knowledge is produced in various forms like books, thesis, journal articles, editorials, reviews, letters, commentaries, technical reports, etc. Here, we must understand what the document is "a written textual form of knowledge in physical form". Guba and Lincoln, in 1981 defined it as "any written material other than a record that was not prepared specifically in response to some requests from the investigator". Further, they have distinguished between documents and records. They defined a record as "any written statement prepared by an individual or an agency for the purpose of attesting to an event or providing an accounting" (Ahmed, J. U., 2010).

According to the **Bailey (1994)**, the *"documentary research method refers to the analysis of documents that contain information about the phenomenon we wish to study".*

Payne and Payne, (2004), explained the *"documentary research method is used to investigate and categorise physical sources, most commonly written documents, whether in the private or public domain"* (Mogalakwe, Monageng., 2009).

3.2.1 Documentary Sources

There are basically two broad categories of sources of documents that could be used for documentary research method; these are namely Primary and Secondary Sources of information.

- **Primary Sources**: includes the first hnd of information sources i.e.historical dociuments, legal papers, outcome of experiments, statistical data, manuscripts, audio visuals, speeches, and artwork are primary sources.
- **Secondary Sources**: Secondary source materials can be newspaper articles, magazines, books, journal articles disscussing others original research, etc.

3.2.2 Classification of Documents

The documents could be classified into various categories like

- **Public documents** popularly include government publications i.e. Acts passed by the Parliament and Assembly, various reports, statistics commission findings, Annual Reports, etc. The implementation of public documents affects the specific geographic region, i.e. country or state
- **Private documents** are organisational based on business documents, board meeting minutes and decision, advertisements, invoices, trade manuals, personal records, Will etc. Private documents are effective within the restricted use only.
- **Open access (free documents) Documents** are those documents which are available in public domain for access, here access means that document is also available to the people those who are other than the authors and there is no restriction on access.
- **Restricted access (fee based) documents**: Priced documents are restricted for use or require authorization for use. E-Journals, Databases, Online Books are the examples of restricted documents.
- **Secret and Confidential Documents:** Secret or confidential documents are those documents which are under closed access. Mass access to these documents may result in damage to the creator. Strategic documents, Trade Secrets, Confidential Reports, Lab findings, etc.

3.2.3 Selection of Documentary Sources

The Selection of documentary sources for research as a research method should be based on the following parameters:

- **Authenticity:** It refers genuiness and reliability of the document. Analysis of authentic evidence is the fundamental requirement. Thus, the researchers must ensure the geniuses and integrity of consulting document are not forged like the interviewer must ensure the identity of the interviewee or the observer is at the right place to manage the observation. The documents should not be taken for granted and close scrutiny of a document needs to be done in following circumstances:
  a) Documents not having anny meaning or errornous
  b) Inconsistent (in style, content, authentication of authors, etc.)
  c) Varoius variant of the document
  d) Documents originated from the dubious, suspicious or undependable source
  e) Document handled by the person or persons with vested interest

- **Credibility:** It refers to whether the evidence is free from error and distortion. It should concern the extent to which a researcher need to record an accurate account from the selected source of the document is free from distortion, deliberately altered or misleading. It is equally important to establish the credibility of documents in documentary research like other research methods. For example, the researcher may rely on newspapers as documentary sources information. While using the reporter/editors comments or opinions on events and occurrences needs to be avoided.
- **Representativeness:** The representativeness applies to some of the documents. Which means whether the evidence is typical of its kind, or if not, is the extent of its untypicality known. Statistical documents i.e. household economic survey is prepared based on the representative sampling method containing the totality of the relevant representative population.

- **Meaning:** It refers whether the evidence is clear and comprehensible. The core purpose of interpretation and examination of the documentary source is to arrive at a conclusion that document contains Significant and meaningful information. For example, information on income per capita is given in factual quantitative information that has the quantitative value, which is not self-evident and the researcher has to derive the meaningful value in the context of countries income per capita. Pictorial descriptions etc. are useless until they have meaningful information attached to them (Mogalakwe, M., 2009).

3.3 Quantitative Research Method for Library and Information Science

For the advancement of any subject domain largely depends on the research made in the domain. The same has happened in the LIS domain and it is advancing day by day. The new methods of research are being inducted into the LIS. Before the application of ICT in the Library and Information, most of the studied were confined to apply qualitative methods of research such as survey research methods, historical research methods, observation studies, case studies etc. The application of ICT and statistical methods made quantitative research more convenient. Largely speaking the science philosophers have contributed more to the development of quantitative research methods. During the 1960-70 the major breakthroughs happened in the development of quantitative research methods which resulted in the development of bibliometrics and scientometrics, its empirical laws.

3.4 Sociology and Philosophy of Science

The sociology and philosophy of scientific knowledge explain the sociological, philosophical, historical, economical aspect of scientific knowledge over the society, institutions and individuals. The scientific knowledge domain of sociology, philosophy and history of Science has various approached namely; sociology of knowledge, normative approach, identification of characteristics of scientific knowledge and the fourth dealt with the scientist using scientific literature. Social interaction plays an important role in the scientific growth. This social aspect of the growth of knowledge could be seen as reflected through the publications and the author’s collaboration in a scientific field.

3.5 Scientometric Research Method

Kumar and Shivrama (2017) explained that “the word metrics derived either from the Latin or Greek word, *"metricus"*, or *"metrikos"* respectively means measurement. When the term metrics is suffixed to a subject domain, it indicates the application of mathematical and statistical measurement techniques to that subject domain. For example, when the term "metrics" is suffixed to the subjects i.e. biology, sociology, psychology, economics, it gives birth to the subjects like biometrics, socio-metrics, psychometrics, and econometrics, where various standard mathematics and statistics techniques extensively used for measurement. Therefore, the terms like librametrics, bibliometrics, informetrics, and scientometrics have been derived by combining the metrics with disciplines such as the library, bibliography, information, and science”.

If it is not possible to measure than it will be very difficult to understand that how much progress has been made by the society. Thus, the concept of measurement of progress remains important. One of the major problems of measurement is not having standard measurement tools/ devices especially in the case of measuring intangible entities like knowledge. Thus, scientist have developed various indexes for measuring tangible/intangible entities like happiness, economic development, IQ level, GDP Indicators, Educational Indicators, etc. The measurement usually represented by a real number by assigning a numerical value to it. The human’s brain understands numeric values very easily as it is easy to communicate. Hence, development of quantified procedures of measurement is essential in any knowledge domain. In the recent past, there had been many information measurement domains covering specific knowledge domain. In the library and information science, domain name of the subject is based on the facets they are associated, which we popularly called as bibliometrics, librametrics, informetrics, scientometrics, cybermetrics, etc. However, we can treat all of them as components of scientoinformetrics.

3.6 Scientometrics

When the term metrics is suffixed to a subject domain, it indicates the application of mathematical and statistical measurement techniques to that subject domain. Therefore, the terms like librametrics, bibliometrics, informetrics, and scientometrics have been derived by combining the metrics with disciplines such as the library, bibliography, information, and science. Where various standard mathematical and

statistical techniques extensively used for measurement. The Scientometrics is commonly known as the sociology of science by the scientists. It is a multidisciplinary subject developed over a period of time. It applies the standard calculation methods for data analysis in a subject domain. It gives birth to a new measurement domain. It is regarded as finest method of science policy and research analysis, evaluation and forecasting It used for studying structural, dynamics, identification of indicators, mapping, scientific growth and knowledge diffusion, etc. in a subject domain. when we combine the the term "metrics" to the subjects i.e. biology, sociology, psychology, economics, new subjects like biometrics, socio-metrics, psychometrics, and econometrics, if formed. However, the modern scientometrics is evolved from the pioneer works of Derek J. de Solla Price (Big Science and Little Science) and Eugene Garfield (Citation Indexing). The Russian scientist V.V. Nalimov coined the Russian term “naukometriya” in 1966 and subsequently used by V.V. Nalimov and Z. M. Mulchenko as the title of their book on quantitative methods of research on the evolution of science. The book was machine-translated into English by the United States Air Force Foreign Technology Division (now the National Air and Space Intelligence Center) in 1971. Thus, the term “Scientometrics” came into existence (**Iris Kisjes et.al., 2011 and Air Force System Command, 1971)**.

The scientometrics can be defined as “*measurement of science which deals with the study of development of science as an information process*” According to A.F.J. van Raan “*scientometric research is devoted to quantitative studies of science and technology”*.

Main subjects of scientometrics are individual scientific documents, authors, scientific institutions, academic journals, and regional aspects of science. Scientometrics expended the boundaries of information science. We see a rapid addition of scientometrics studies. These aspects may include productivity, subject analysis of the documents, reception of citations and formal communication. Scientometrics is focused on scientific information only **(Wolfgang, G. S. and Sonja, W. D., 2006)**. Based on the bibliometricians / scientometricians views scientometrics can be categorised into following categories:

Descriptive Scientometric and Behavioural Scientometrics: Descriptive Studies describe the characteristics features of documents or literature. This includes the study

of bodies producing documents individual authors and research institutions, Geographical origins. Documents that carry information, Quantity of information, Time and frequency of studies, Form of transmission such as books, journals etc. Behavioural Studies examines the relationship between various elements such as among the Authors, Subject, Citation, Form and content, related documents (**Velmurugan, C. and Radhakrishnan, N., 2014**).

- Borgman (1989) classified three types of analysis 1) analysis of *producers* (authors or institutions), 2) *artifacts* (books, journal articles, websites), and 3) *concepts* (topics or subject areas).

- Stevens (1953) proposed two types 1) *descriptive* (productivity by author, organization, geographical region, time period, or subject) and 2) *evaluative* (usage data, citation behaviour).

- Nicholas and Ritchie (1978) used two related categories 1) *literature characteristics* (authorship, publication year, etc.) and 2) *literature relationships* (behavioural studies such as citing or co-citing patterns) **(Teresa S. Welsh, Ed., 2015)**

3.7 Research Design

A research design is an anatomical framework or blueprint for conducting the research by identifying the most suitable and ideal methods and procedures. This study is based on exploratory scientometric research design attempting to explore the growth of knowledge, and collaboration among knowledge producers. Identification of possible indicators to establish a causal relationship between the various interrelated variables. The scientometrics research design is analytical that adopt the detailed analysis of secondary data using a range of Scientometric tools, techniques, formula and laws along with standard statistical techniques. The quantitative approach of documentary research methods i.e. scientometric research design examines various aspects of published knowledge in the form of literature (documents) in a subject domain. There are various scientometric variables (indicators) as discussed above are applied to study the research problem **"Global Knowledge Diffusion in Vocational**

Education and Training: A Scientometric Study". The following levels of scientometric analysis are performed to explore the communication phenomenon such as the growth of literature, determining research productivity, exploration of citation pattern, co-authorship and collaboration pattern, prolific authors, organisations and countries, the form of literature, validation of classical laws of Bradford's Law and Lotka Law, etc., are tested. The correlation analysis was selected to understand the relationship between GDP and publication productivity of various countries in the field of vocational education and training.

Based on the systematic review of the prior studies as a part of the review of the literature, none of the studies was conducted in the VET domain. Therefore, the study is proposed to explore the knowledge diffusion in VET domain based on the literature indexed in WoS database.

3.7.1 Data

Selection of data for research is the vital foundation of the research work. The data mining has been made by exploring standard databases Web of Science Core Collection Database, which is further discussed below.

3.7.2 Web of Science Core Collection Database

Web of Science™ Core Collection database a licensed fee based subscribed by the Tata Institute of Social Sciences, Mumbai, was used for retrieving the data on VET for the period of twenty-five years for study from 1992-2016. The researcher has applied the search strings "Vocational education* OR Vocational training*" (see figure 3.1) in the topic with the combination of Boolean operator OR to obtain the maximum number of records for analysis. The filed search string has used for the data extraction from the Web of Science Core Collection comprising Science Citation Index (SCI)-Expanded, Social Science Citation Index (SSCI), Arts and Humanities Citation Index (A&HCI), and Emerging Sources Citation Index (ESCI) for the period of 25 years, based on the above strings. As of May 2017, a total of 6719 publications (see figure 3.2) were published during the period 1992-2016 and 64065 citations received by these publications. The data analysis was performed according to the objectives of the study.

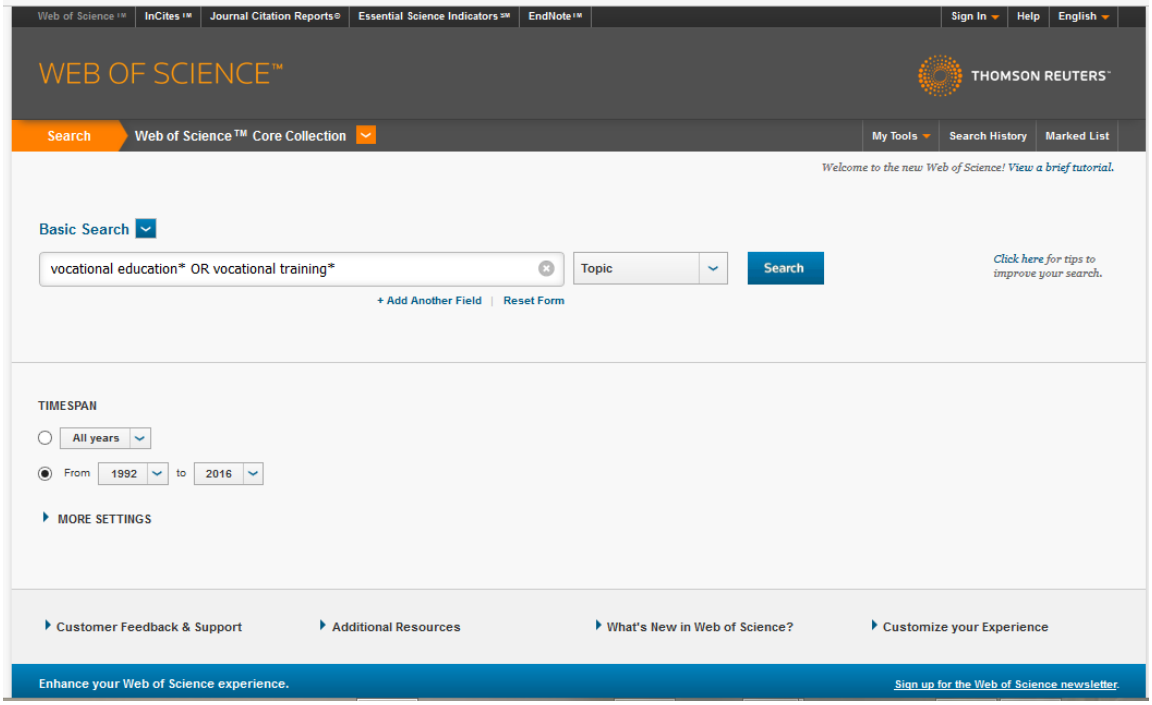

Figure 3.1: Screenshot of Web of Science Database

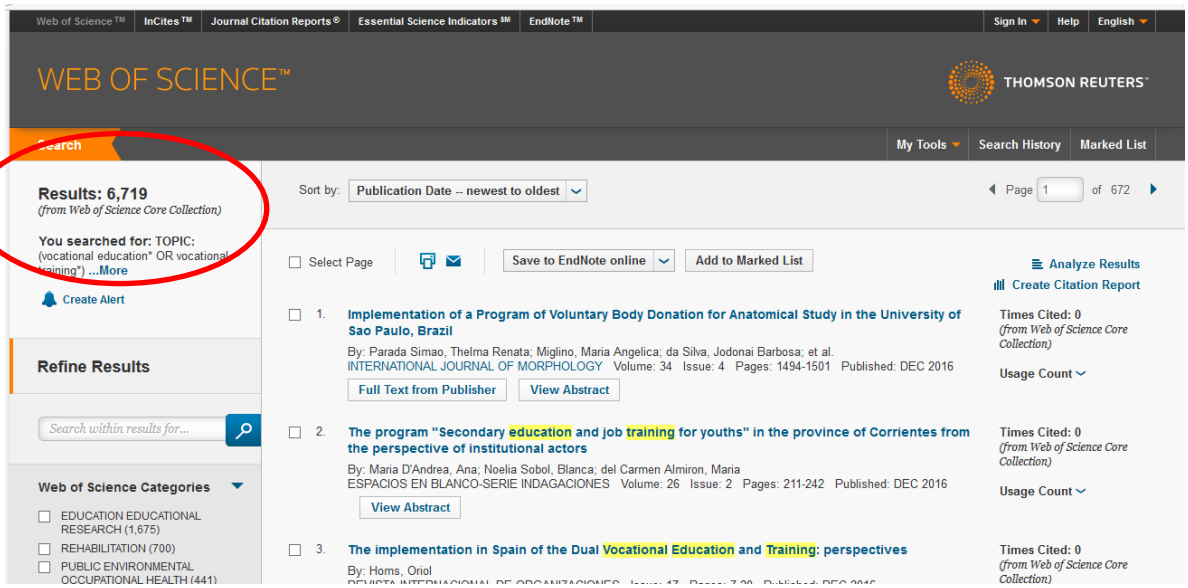

Figure 3.2: Screenshot of Search Result of Web of Science Database

3.7.3 UIS, UNESCO Economic Data

The Unesco Institute of Statistics, UNESCO, has provided the data of demographic and socio-economic conditions of the world. Thus, the socio-economic indicators data

of GDP per capita (current US$) for the year 2015 is compared with the publication output the World for analysis (UIS, UNESCO, 2017).

3.8 Statistical Methods and Analysis

To study the growth of knowledge, knowledge productivity, the degree of collaboration, citation pattern, various Scientometric indicators and laws mentioned below have been implemented to explore knowledge diffusion at the global level. The researcher has applied percentage analysis and average score analysis and applied the following statistical tools:

3.8.1 Compound Annual Growth Rate

Investopedia (2017) define that "the compound annual growth rate (CAGR) is the mean annual growth rate of an entity during a specific period". It can be calculated using following formula:

$$CARG = \left(\frac{Ending\ Value}{Beginning\ Value}\right)^{\left(\frac{1}{no.of\ years}\right)} - 1$$

3.8.2 Relative Growth Rate

Aswathy and Gopikuttan (2015) defined "Relative Growth Rate (RGR) is the rate of growth relative to the size of the population. In simple words, RGR is the growth in the number of articles diffused per unit of time". It can be calculated from the below formula:

$$\mathbf{R\ (P)} = \frac{\text{Log}_e 2P - \text{Log}_e 1P}{2^T - 1^T}$$

Whereas

R (P) = RGR of articles over the specific period of time.

Log $_e$ 1P = log of initial number of articles.

Log $_e$ 2P = log of final number of articles.

$2^T - 1^T$ = Unit difference between the initial time and final time.

3.8.3 Doubling Time

According to the Aswathy and Gopikuttan (2015) "The Doubling time is the period required to double the number of articles in a subject domain during a specific period of time. The doubling time is directly related to the Relative Growth Rate (RGR). When the relative growth rate (not the absolute growth rate) is constant, the quantity undergoes exponential growth and has a constant doubling time or period, which can be calculated directly from the growth rate".

The formula of calculation of Doubling Time is as follows:

$$\text{Doubling time (Dt)} = \frac{\text{Log e } 2}{R} = \frac{0.693}{R} \qquad R = \text{RGR}$$

3.8.4 Exponential Growth Model

The Exponential model is associated with the name of Thomas Robert Malthus (1766-1834). As per the exponential growth model, the population in a domain increases over time is at an exponential rate, continuing upward a result of the number of individuals available to reproduce without regard to resource limits. Thomas Robert Malthus realized that any entity may potentially increase in numbers according to a geometric series. There are 3 possible model outcomes (see figure 3.3):

a) Population exponentially declines ($r < 0$)
b) Population exponentially increases ($r > 0$)
c) Population does not change ($r = 0$)

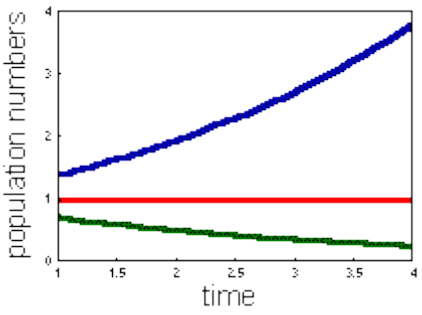

Figure 3.3: Exponential Growth Model
Source: Alexei Sharov (1997)

Parameter r in the exponential model represents the difference between the birth (reproduction) rate and the death rate:

$$\frac{dN}{dt} = (b - m)N = rN$$

where

b is the birth rate
m is the death rate.

The rate of population growth (*r*) is equal to birth rate (*b*) minus death rate *(m)* (Alexei Sharov, 1997).

3.8.5 Citations Per Paper (CPP)

According to Harzing, Anne-Wil.(2010-2013) "this can be a very useful metric to assess the average impact of publications made by a author or in a journal. The Citations Per Paper (CPP) is the average number of citations per paper received by the papers in a subject. It can be calculated by dividing the total number of citations by the total number of papers".

.

$$\text{CPP} = \frac{\text{total number of citations}}{\text{total number of papers}}$$

3.9 Authorship Pattern

Stephen, G. and Balamurugan, T. (2015) have defined "that authorship the important bibliometric measurement reflecting contemporary communication patterns, productivity and collaboration among the researchers".

3.10 Degree of Collaboration

The collaboration in academic writings is very common activity. According to Subramaniyam, K. (1983) "degree of collaboration is the ratio of the number of collaborative research papers to the total number of research papers in the discipline in a specific time period". The following formula was suggested by Subramanyam for the calculation of the degree of collaboration:

$$C = \frac{NM}{NM + NS}$$

Where

C = Degree of collaboration
NM = Number of multi authored papers
NS = Number of single authored papers

3.11 Publication Efficiency Index

The Publication Efficiency Index of the leading countries of the world can be calculated according to the formula suggested by the Guan and Ma (2007) in their study as "a measure of research quality is used for calculation of PEI. The PEI reflects the research impact of publications in a country is compatible with the global research efforts in a specific research domain. The value of PEI > 100 for a country indicates that the impact of publications is more than the research efforts devoted to it for that particular country and vice versa" (Guan, Jiancheng and Ma, Nan 2007).

$$PEI = \frac{TNC_i / TNC_t}{TNP_i / TNP_t}$$

Where:
TNC_i denotes the total number of citations of country **i**
TNC_t denotes the total number of citations of all countries
TNP_i denotes the total number of papers of country **i**
TNP_t denotes the total number of papers of all countries

3.12 Relative Citation Index

To assess the impact of research output the Relative Citation Index (RCI) can be examined by using scientometric indicators. The Institute of Scientific Information developed RCI to calculate the science and engineering indicators. According to Kumari, (2009) "RCI is a measure of both the influence and visibility of a nation's

research in global perspective. It is defined as the ratio of a country's share of world citations to the country's share of world publications (Citation% / Publication%)".

RCI > 1 indicates that country's citation rate is higher than world's citation rate:
RCI = 1 indicates that country's citation rate is equal to world citation rate; and
RCI < 1 indicates that country's citation rate is less than world's citation rate.

3.13 Journal Performance Indicators

There are various journal performance indicators available in today's world. Most of these indicators are broadly based on the number of citations and number of publications. Some of the popular journal performance indicators are discussed below.

3.13.1 Journal Citation Report

The Journal Citation Reports (JCR) is the product of Clarivate Analytics (formerly Thomson Reuters). The JCR gives the one of the widest coverage of more than11,000 indexed journals of nearly 250 disciplines, across 81 countries. The coverage includes evaluation of 2.2 million articles, reviews, and other items. The JCR measures research on various indicators.

3.13.1.1 Journal Impact Factor (JIF): The Journal Impact Factor (JIF) is a journal level metrics to assess the journal quality. In simple words, JIF is mean average of number citations received by each published paper of a journal. The JIF was proposed by the Eugene Garfield, the founder of the Institute for Scientific Information. Following is the JIF calculation formula:

$$\text{Journal Impact Factor (2016)} = \frac{\text{Number of citations in 2016 to the articles published in 2014+2015}}{\text{Total citable articles published in 2014+2015}}$$

$$\text{Example: JIF (2016)} = \frac{2900+4800}{850+860}$$

$$\text{JIF (2016)} = 4.50$$

3.13.1.2 Eigenfactor: The Eigenfactor (EFC) Score is also a journal level metrics developed by Jevin West and Carl Bergstrom. "Journals are rated according to the number of incoming citations, with citations from highly ranked journals weighted to make a larger contribution to the eigenfactor than those from poorly ranked journals" (**University of Washington. n.d.)**

- **Total Cites (TC)** **:** The number of times a journal is cited as listed in JCR
- **World Rank (WR)** **:** World rank is the global rank of a journal based on scientometric indicators i.e. JIF, h-Index, etc.

3.13.2 SCImago and CWTS Journal Ranking

SCImago is a research group. The SCImago Journal & Country Rank is an openly accessible portal of scientometrics indicators based on the Scopus Database. The analysis could be made at journal level and country level as well. The coverage includes 27 major thematic subject area or by country. The citation data is obtained from more than 21,500 titles of 5,000 international publishers and country performance metrics across 239 countries. This platform takes its name from the SCImago Journal Rank (SJR) indicator, developed by SCImago from the widely known algorithm Google PageRank (SCImago, 2007). CWTS Journal Indicators (www.journalindicators.com) offers a number of scientometric indicators on scientific journals (**CWTS, Leiden University, 2017**).

Following four types of Journal Indicators are provided by the Centre for Science and Technology Studies (CWTS) at Leiden University:

- **Papers:** number of papers of in last three years of a journal
- **Impact per Paper (IPP):** similar to the JIF
- **Source Normalized Impact per Publication (SNIP):** Ratio of a journal' raw impact per paper and the relative database citation potential
- **% self cit**: (percentge of self citatitons of a source)

3.14 Citation Classics

Eugene Garfield explained Citation Classic as a highly cited publication based on the citation data of Science Citation Index (SCI), Social Sciences Citation Index (SSCI), and Arts & Humanities Citation Index (A&HCI). According to him publications having 400 more citations may be considered a classic. But in some disciplines, such as social sciences, arts and humanities having fewer researchers, the publications having 100 and more citations may be rendered as the citation classics in that domain. Another reason could be that citation rates differ from discipline to discipline (Garfield, Eugene n.a.).

3.15 Productivity Index

The productivity index reflects the ability of an individual to produce. The popular productivity indexes are discussed below.

3.15.1 h-index

J.E. Hirsch in 2005 proposed authors productivity index based on the number of papers produced and citation received by these papers popularly known as h-Index. Similarly, it may be applied to the journals, organisations or countries. According to Hirsch, the *h* index is defined as: "A scientist has index h if h of his or her Np papers have at least h citations each and the other (Np – h) papers have ≤h citations each" (Hirsch, 2005). The h-Index could find out when the number of publications ranked in descending order according to the number of time cited (see figure 3.4). Example: If a researcher has 25 papers, each of which has at least 25 citations, their h-index is 25.

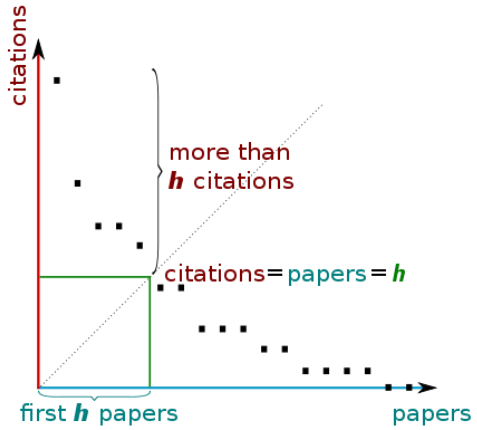

Figure: 3.4: *h*-index (plotting of citations of papers in decreasing order)
Source: Wikipedia, 2017

3.15.2 g-Index

The Egghe, Leo (2006) defined "*g*-index is introduced as an improvement of the *h*-index of Hirsch to measure the global citation performance of a set of articles. If this set is ranked in decreasing order of the number of citations that they received, the *g*-index is the (unique) largest number such that the top *g* articles received (together) at least *g*2 citations".

3.16 Scientometric Laws

The Lotkas law of frequency distribution of publication and Bradford's Law of Scattering of journals are discussed in this section.

3.16.1 Lotka's Law

The Alfred J. Lotka in 1926 describes the frequency of publication by authors in a given field. The law describes that "the number (of authors) making *n* contributions is about $1/n^2$ of those making one; and the proportion of all contributors, that make a single contribution, is about 60 percent". In simple words, the law states that in a specific domain among all the authors, 60 percent authors will have just one publication, and 15 percent will have two publications ($1/2^2$ times 60), 7 percent will have three publications ($1/3^2$ times 60), and so on. As per the observation of this law (law of scientific productivity), only six percent of the authors will produce more than 10 articles in a given field (**Tamilselvan, N., 2013**).

The general formula says:

$$X^n Y = C$$

Or

$$Y = C / X^n$$

Where:

X is the number of publications,
Y the relative frequency of authors with *X* publications, and
n and *C* are constants depending on the specific field.

3.16.2 Bradford's Law

Samuel Clement Bradford formulated the law of scattering of journals. According to **Kanakaraj, (2016)** "when a set of journals arranged in an order of decreasing productivity will result in the journals which yield most productive articles come first while the most unproductive tail last". As per the equation of the law, the journals need to be categorised (into the nucleus of periodicals) in the number of zone each producing a similar number of articles. However, the number of journals in each zone will increase rapidly. Thus the relationship among the zones when the numbers of periodicals in the nucleus and succeeding zones will be $1:n:n^2$".

3.17 Softwares

The Microsoft Excel package was used for the data analysis. Various charts i.e. bar charts, pie charts, bubble charts, etc. have been prepared with MS Excel. The Excel package is also used for determining various averages and mathematical calculations. Further based on Citation Per Paper (CPP) is applied to find out the quality of the research output. This research study explores the Compound Annual Growth Report (CAGR) and Relative Growth rate during the study period. The study aims to analyse the thrust areas of research concentration on VET. It is analytical in nature with the suitable statistical tools applications in strengthening the empirical validity.

3.18 Hypothesis

The Hypothesis is a shrewd guess or mere assumption, which the study wants to accept and reject**.** The hypothesis could be of two types Null Hypothesis and Alternate Hypothesis**,** which can be symbolised as follows:

Null Hypothesis	-	H_0
Alternate Hypothesis	-	H_1

The study intended to statistical test the following null hypothesis:

- Journals are the preferred source of knowledge diffusion than other sources
- Is there any direct or indirect relation between country's GDP and its publication productivity.

3.18.1 Statistical Testing

To test the formulated hypothesis two tests have been performed in this section.

3.18.1.1 *t*-Test

The *t*-Test was done for testing the level of significance of the hypothesis "*Journals are the preferred source of knowledge diffusion than other sources".* Investopedia (2017) defines that "t-test is an analysis of two population means through the use of statistical examination; a t-test with two samples is commonly used with small sample sizes, testing the difference between the samples when the variances of two normal distributions are not known"

Use the following formula to calculate the t-score:

$$t = \frac{(\sum D)/N}{\sqrt{\frac{\sum D^2 - (\frac{(\sum D)^2}{N})}{(N-1)N}}}$$

ΣD: Sum of the differences (Sum of X-Y from Step 2)
ΣD^2: Sum of the squared differences (from Step 4)
$(\Sigma D)^2$: Sum of the differences (from Step 2), squared.

3.18.1.2 Pearson Correlation Coefficient

The Pearson Correlation has been measured and presented to validate the relationship between publication productivity and the GDP of most prolific countries.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

3.19 Limitations

The possible limitation of the study is that it is related to the study of data indexed in Web of Science Core Collection database only. The study is limited to the publications indexed in Web of Science Core Collection database brought out from 1992 to 2016.

References

Ahmed, Jashim Uddin. 2010. “Documentary Research Method: New Dimensions.” *Indus Journal of Management & Social Science (IJMSS)* 4 (1): 1–14.

Alexei Sharov. 1997. “Exponential Model.” Academic. *University of Texas at Austin*. https://www.ma.utexas.edu/users/davis/375/popecol/lec5/exp.html.

Anne-Wil Harzing. 2010. “2.2.1 Basic Metrics Based on Papers and Citations.” *The Publish or Perish Book*. https://harzing.com/popbook/ch2_2_1.htm.

Aswathy, S., and A. Gopikuttan. 2015. “Bibliometric Observation of Publication Output of University Teachers: A Study with Special Reference to Physics.” *Journal of Scientometric Research* 4 (1): 14–19.

Centre for Science and Technology Studies. 2017. “CWTS Journal Indicators.” *CWTS Journal Indicators*. http://www.journalindicators.com.

Chandran Velmurugan, and Radhakrishnan Natarajan. 2017. “Quantitative Techniques for Library and Information Science Students and Research Scholars: A Conceptual Approach.” In *A Conceptual Approach, Perception Changes of Librarianship in Digital Era*, edited by Dr. Kaushal Kishor Chaudhary and Dr. Sudhir Kumar Jena, 1sted., 109–28. New Delhi: S.K. Book Agency. Accessed July 26. https://www.researchgate.net/publication/289522253_Quantitative_Techniques_for_Library_and_Information_Science_Students_and_Research_Scholars_A_Conceptual_Approach.

Egghe, Leo. 2006. “Theory and Practise of the G-Index.” *Scientometrics* 69 (1): 131–52.

Eugene Garfield. 2017. “What Is a Citation Classic?” *Citation Classics*. Accessed July 31. http://garfield.library.upenn.edu/classics/.

Guan, Jiancheng, and Nan Ma. 2007. “A Bibliometric Study of China’s Semiconductor Literature Compared with Other Major Asian Countries.” *Scientometrics* 70 (1): 107–24.

Hirsch, J. E. 2005. “An Index to Quantify an Individual’s Scientific Research Output.” *Proceedings of the National Academy of Sciences of the United States of America* 102 (46): 16569–72.

Investopedia. 2017. “T-Test.” *Investopedia*. http://www.investopedia.com/terms/t/t-test.asp#ixzz4miALALJE.

Investopedia. 2017. “Compound Annual Growth Rate (CAGR).” *Investopedia*. http://www.investopedia.com/terms/c/cagr.asp#ixzz4mpenVvld.

Iris Kisjes, Gert Jan Geraeds, Andrew Plume, David Tempest, Judith Kamalski, Sarah Couffignal-Szymczak, and Michelle Pirotta. 2009. “Scientometrics: The Birth of the Science of Science.” *Research Trends* (11):1-12.

Kanakaraj, S. 2016. “Aquaculture Research Outputs from Scopus Database: A Bibliometric Analysis” *Indian Journal of Information Sources and Services* 6 (1): 20-29.

Kumar, A., & Shivarama, J. (2017). Knowledge Diffusion to Knowledge Dissemination: A Theoretical Study. *International Research Journal of Multidisciplinary Studies, 3*(7). Retrieved from http://www.irjms.in/sites/irjms/index.php/files/article/view/484

Kumari, G. Lalitha. 2009. “Synthetic Organic Chemistry Research: Analysis by Scientometric Indicators.” *Scientometrics* 80 (3): 559–70.

Mogalakwe, Monageng. 2009. “The Documentary Research Method – Using Documentary Sources in Social Research.” *Eastern Africa Social Science Research Review* 25 (1): 43-58.

Scimago Lab. 2017. “SJR SCImago Journal & Country Rank.” SJR SCImago Journal & Country Rank. http://www.scimagojr.com.

Stephen, G., and T. Balamurugan. 2015. “Open Access Literature Productivity of Library and Information Science: A DOAJ Perspective.” In *Handbook of Research on Inventive Digital Tools for Collection Management and Development in Modern Libraries*, edited by S. Thanuskodi, 153–69. https://www.igi-global.com/chapter/open-access-literature-productivity-of-library-and-information-science/133963.

Subramanyam, K. 1983. “Bibliometric Studies of Research Collaboration: A Review.” *Information Scientist* 6 (1): 33–38.

Tamilselvan, N., Shivakumar, N., and Sevukan, R. 2013. “Lotka’s Law a Study with Reference to the Literature by faculties of National institute of Technology, India.” *International Journal of Library and Information Science* 2(1): 19-34.

Teresa S. Welsh, ed. n.d. “Qualitative and Quantitative Methods in Libraries (QQML) Special Issue Bibliometrics and Scientometrics.” *Qualitative and Quantitative Methods in Libraries* 3: 1–3.

Unesco Institute of Statistics, UNESCO. 2017. “Demographic and Socio-Economic : Socio-Economic Indicators.” *UIS.Stat*. January 21. http://data.uis.unesco.org/Index.aspx.

University of Washington. n.d. “EigenfactorTM Score and Article Influence TM Score: Detailed Methods.” www.eigenfactor.org. http://www.eigenfactor.org/methods.pdf.

V. V. Nalimove. 1971. “Measurement of Science.” Translated by Z.M. Mulc’henko. Air Force System Command. http://www.garfield.library.upenn.edu/nalimov/nalimovmeasurementofscience/book.pdf.

Wikipedia. 2017. “H-Index.” *Wikipedia*. https://en.wikipedia.org/w/index.php?title=H-index&oldid=798100329.

Wolfgang G. Stock, and Sonja Weber. 2006. “Facets of Informetrics.” *Information* 57 (8): 385–89.