iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** V **Month of publication:** May 2020

DOI: http://doi.org/10.22214/ijraset.2020.5268

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

History of Indexing: An Overview

Pratik Ghosh
Research Scholar (Management), Christ Deemed to be University

Abstract: ***Indexing is the method of keeping the indices or a list of information being published. The modern indexing has a prolonged history and has evolved in many senses from its inception. This article gives an overview of the history of indexing, important people who contributed towards indexing, usage and importance. The article is divided into sections based on the chronology of indexing and its development in stages and aims to provide valuable insights to the reader.***
Keywords: History of indexing, indexing, contributors, research, development of indexing

I. INTRODUCTION

Indexing the term as described in the dictionary is a list, a page, a name or a collection of information listed in alphabetical order (as of bibliographical information or citations to a body of literature). It can be also considered that it aids to count the number of times people have shared/viewed the work etc. Indexing plays a huge importance in research as it facilitates the ease of search for the works of authors, researches etc. as per relevancy and aids citation i.e. giving proper references of the previous literature in the field. Modern-day indexing uses computational and complex algorithms to search as per the keywords, which is developed and has been advanced upon the works of many contributors since ages which are further discussed in this review paper.

II. HISTORY OF INDEXING

The history of indexing is vast and has been documented only in the recent times with the advent of technology. It could be broadly classified as per the invention and availability of the printing press and computers, i.e. the pre-printing era and the print era and the digital era.

A. Pre-printing Era

The pre-printing era was dominated by manuscripts. They were interpreted and written by humans while taking notes from any sermon or wordings. No two manuscripts were same and the process of pagination started through them, in which few of the scripts were documented so immensely that they are considered one of the finest works of literature and are referred till date.

Writing of manuscripts also gave rise to the process of **pagination** which is giving page numbers for documents which helped in enlisting the table of contents. Manuscripts mostly had religious affiliations and all the major religions of the world preserved them as part of divinity and references. Before indexes were coined, concordances were recorded from the seventh century onwards, which are as defined in the dictionary as an index arranged alphabetically of the principal words in a book or the works of an author along with their immediate contexts. The list of words mainly used in any religious text, philosophy or government. For example, the list of words which had reputedly used in the Bible was recorded and these concordances were referred to quickly reach to the verse or text.

B. Printing Era

Modern-day printing started in 1440, which opened new avenues and the need to index the various forms of publications. In 1460 came the first printed back indexes and in the following year, i.e 1470 saw the first print concordance of the Bible.

In the 18th century, the first dictionary of English Language was coined by Dr Samuel Jackson. The dictionary was a result of the growing demand of the language learners as there was a rise of literature books, text, maps, pamphlets, newspapers etc due to the availability of printing press and bookbinding. This was indeed a result of the rise of literacy and the demand for set grammar, definition and spelling of the words which was not standardized then. Before this, though there were few dictionaries complied in the Italian, Latin, French and English the lexicographers were not able to connect to the usage of the word uniformly. Dr Jackson took seven years and almost single headedly worked in the compilation of the words knows in literation and coined A Dictionary of the English Language. Though later the etymologies and spellings of Johnson were much criticized and modified in the later stage which gave rise to much simpler forms like Webster and Oxford dictionaries that were only after a century later.

In the 19th century, William Frederick Poole and American bibliographer and Librarian published a 154-page index for the periodical literature in 1848 which later added in three editions in 1853 and 1882. A Belgian author and bibliographer named Paul Otlet were one of the visionaries who was also referred to as the father of information science and documentation, has an immense

©IJRASET: All Rights are Reserved

contribution towards indexing. He created a Universal Decimal Classification and founded the Institut International de Bibliographie in 1895. He also created index cards which thus created a repository known as *Repertoire Bibliographique Universel*
Index society was formed in 1877 whose main purpose was to create a general index of universal literature. There was an equal increase of indexing and broadly there was either citation or legal. For legal citation Frank Shepard, a citatory in the United States Legal research indexed as per the case, title etc. The term Shepard's Citations was coined in 1873 was trademarked and also the verb *Shepardisizing* was coined for the use of those citations.
In the 20th century with more and more literature getting published around the world and also due to the advent of technology, a need for more mechanised indexing was required. By the '40s, an Italian Jesuit priest called Fr. Roberto Busa, who could be undoubtedly called the pioneer in computational linguistics complied the indexes for Saint Thomas Aquinas works and named it as *Index Thomisticus* in 1946. In 1949 he collaborated with Thomas Watson the founder of IBM and persuaded him to sponsor for the same which became the turning point in the history of indexing such a huge corpus of works. The Society of Indexers, a professional society based in the UK was also founded in 1957 which had members worldwide and worked together for indexing, quality of indexes and related areas.

C. Citation Indexing

Citation indexing is a simple concept, it recognises the value of information which is determined by the users, and the best way is to measure the quality of the work than by measuring the impact on the community at large. The widest conceivable population inside the academic community (for example any individual who uses or refers to the source material) decides the impact or effect of the thought and its originator on our collection of information. Due to its effortlessness, one will in general overlook that citation referencing is a genuinely on-going type of data the board and recovery.
Development of citation index was caused by these three back in the 1950s. Following the World War II, there was a remarkable conjunction of government dollars into research work the exploration community normally started to openly archive its discoveries through the acknowledged channel of the published scientific journal literature. The subsequent mushrooming of the literature created a need for a method of indexing and retrieval that would be more cost-effective and efficient than the then-current model of human indexing of materials for subject-specific indices. While the inconspicuous choices made by subject experts were important in offering profundity to a subject index, manual indexing was both an additional tedious procedure and labour extensive. Its costs expanded with respect to the development of material to be filed. So the requirement for a superior method of managing data was the main factor.
The second factor was the increasing dissatisfaction with the volume of subject indexing to meet the requirements of the researcher. As this point in time, a subject index could have exorbitant slack occasions in adding materials to the files of the time; months could go before scientists in a single field would learn of published works in some other field that had significance to their study. Besides, there were constraints to the subject ordering as far as recovery. Appropriate terminology to one explicit control would not have importance to specialists in another, maybe covering, discipline. Simultaneously, researchers perceived that they must know about, if not acquainted with, work in various distinctive subjects teaches to be sure that they had appropriately grounded the exploration through a proper review of the literature.
Alongside this need, the third and last factor in the improvement of citation indexing was the expectation that automation may hold the appropriate responses. Computerization during the 1950s was far expelled from the work area condition of today, yet there was excitement over potential advantages derived from the use of machines to the generation and accumulation of data. The U.S. government trusted that automation could alleviate or even wipe out the troubles of manual indexing. Various projects were propelled by the United States with the expectation of researching these conceivable outcomes.
One of the contributors for citation indexing is Dr Eugene Garfield, founder and now Chairman Emeritus of ISI (now Clarivate Analytics), who was deeply involved in the research relating to machine-generated indexes in the mid-1950s and early 1960s. Perhaps the earliest purpose of contribution was an undertaking supported by the Armed Forces Medical Library (predessor to the present day National Library of Medicine). The Welch Medical Library Indexing venture, as it was called, was to research the job of mechanization in the association and recovery of medical literature. The expectation was that the issues related with abstract human choice in the choice of descriptors and ordering terms could be wiped out. By evacuating the human component, one may subsequently speed up with which data was joined into the records. It may likewise build the cost-adequacy of the records. Garfield got a handle on at an opportune time that survey articles in the diary writing were vigorously dependent on the bibliographic references that eluded the perused to the first distributed hotspot for the eminent thought or idea. By catching those references, Garfield accepted, the specialist could promptly get a perspective on the methodology taken by another researcher to help a thought

©IJRASET: All Rights are Reserved

or strategy dependent on the sources that the distributed essayist had counselled and referred to as appropriate in the list of sources. As recovery terms, references could work just as catchphrases and descriptors that were astutely doled out by an expert indexer.

Eugene Garfield and Associates, in the early 1960s, developed two pilot projects that will be helpful in testing the viability and efficiency of citation indexing. The first project was for two private pharmaceutical companies and it involved database creation that would index the citations of 5,000 chemical patents held by them The referenced references in this example were to earlier licenses, the documentation sources that the administration patent analysts were utilizing to help a choice to concede or deny a patent. The associations that the patent citation index made were then broke down with two tantamount groupings and ordering frameworks that were at present being utilized by the members. In light of this examination and investigation, the venture supports discovered that citation indexing allowed the recovery of significant writing across discretionary characterizations such that subject-situated ordering proved unable.

In 1962, the second pilot project involved, the Institute for Scientific Information (now Clarivate Analytics) which was recently incorporated by Garfiled, with the United States National Institutes of Health in building an index to the published literature on genetics. This project was unquestionably more perplexing than the licenses record. Three databases were worked to cover the literature for more than 1 year, 5 years and 14 years with a changing number of source distributions listed in each. While this venture was to test the plausibility and utility of a restricted, discipline-arranged reference record, at fulfilment, it was reasoned that the database with the most extensively based arrangement of source distributions shaped the most exhaustive and helpful manual for the distributed writing in the field of genetics. The database for the single-year term had drawn not simply on journals that were essentially committed to the field of genetics inquire about however had drawn too from an enormous pool of journals that distributed hereditary qualities papers on a progressively fringe or periodic premise. Also, while the mechanized framework required a specific degree of exertion in normalizing the sections from a wide assortment of distributed materials, the project exhibited the cost-adequacy of citation indexing instead of the cost of traditional subject indexing processes.

During the completion of the project, the government sponsors decided not to sponsor the improvement of a national citation database, Eugene Garfield was urged to push forward with the private publication of his multidisciplinary citation index as the first edition of the Science Citation Index (SCI). Accessible for buy since 1963, the SCI at that point and now speaks to the most far-reaching citation index to the scientific journal literature. The Web-based variant of that index in today's date covers 5,600 journals across more than 150 scientific disciplines.

Garfield’s achievement laid the way in establishing the utility and objectivity of a citation index in pulling up related papers in the published literature that at first glance might not have seemed pertinent to the researcher's inquiry. Till date it is viewed as one of the most reliable resources in following the advancement of thought over the large number of disciplines that are a piece of our collection of scientific knowledge and many projects take an inspiration from his work to develop the algorithms etc.

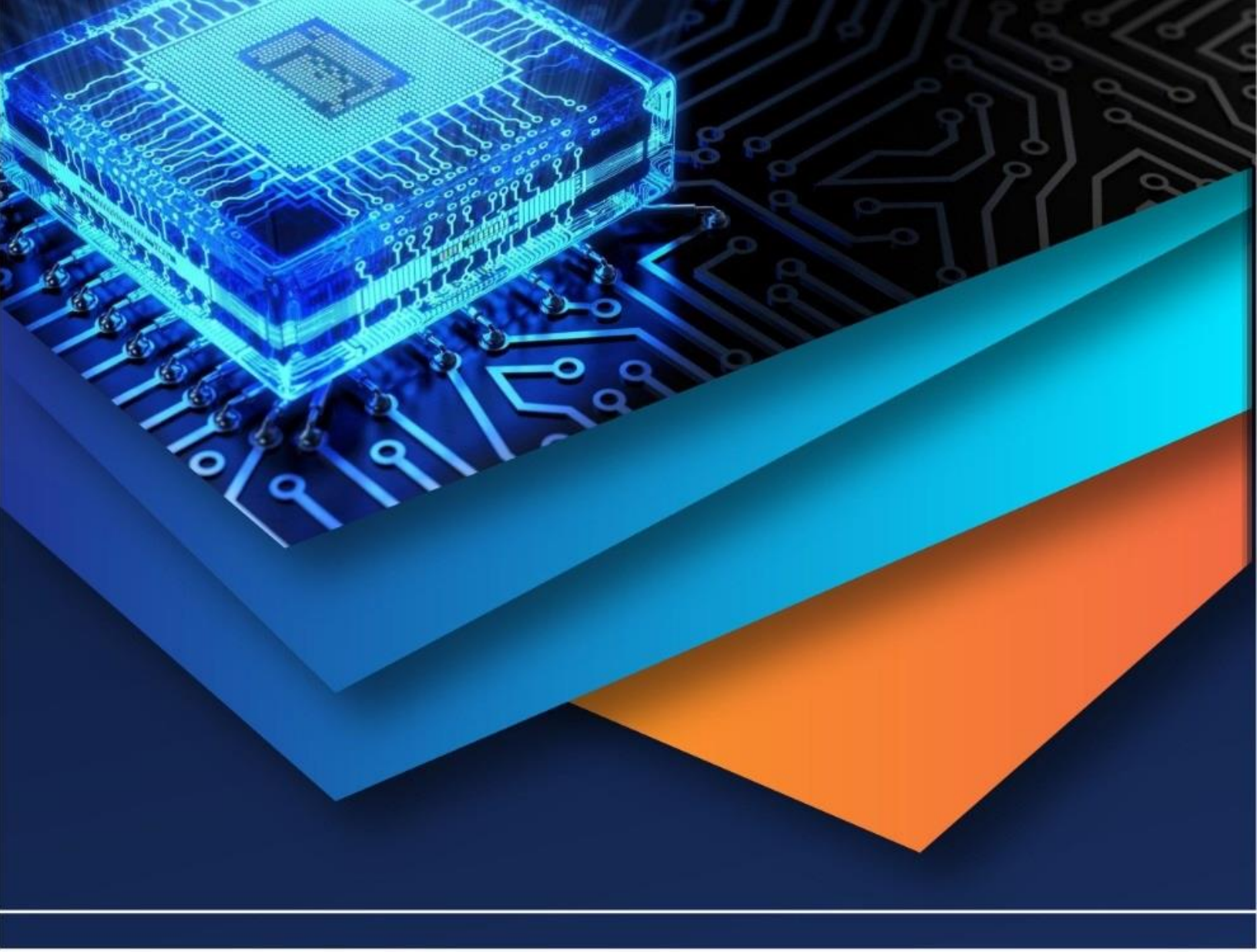
III. CONCLUSION

Indexing as a concept and a process is extremely helpful for all the researchers. It has indeed advanced through stages of developments, as stated in the article, i.e. from the era of manuscripts to the computer-based citation indexing. The process has become convenient and hassles free for all the involved persons in the research process and with the advent of internet and technology the ease of access has been also profound. This article could help future researchers to have an overall view of the history of indexing and open up the scope of further research areas like the study of pre-printing era, the printing era advancements or more recent developments in detail

REFERENCES

[1] History of citation indexing. (2019, August 5). Retrieved May 13, 2020, from https://clarivate.com/webofsciencegroup/essays/history-of-citation-indexing/
[2] A brief history of indexing. (n.d.). Retrieved May 13, 2020, from https://www.anzsi.org/about-indexing/a-brief-history-of-indexing/
[3] Pasler, M., & Pasler, R. (1971). U.S. History: Periodical Abstracting and Indexing. *RQ, 10*(3), 232-236. Retrieved May 14, 2020, from www.jstor.org/stable/25824193
[4] Garfield, E. (1963). Citation indexes in sociological and historical research. *American Documentation, 14*(4), 289–291. https://doi.org/10.1002/asi.5090140405
[5] Index. (n.d.). Retrieved May 14, 2020, from https://www.merriam-webster.com/dictionary/index
[6] Concordance. (n.d.). Retrieved May 14, 2020, from https://www.merriam-webster.com/dictionary/concordance
[7] Pagination. (n.d.). Retrieved May 14, 2020, from https://www.merriam-webster.com/dictionary/pagination

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)