



Document Landscape and Generative AI: A Synthetic Term DHISCREMENT Postulated for Nomenclature of Scholarly Library Resources

Dr Soumen Mallik^a

^aAssociate Professor and Head, Department of Library and Information Science, Vidyasagar University, Midnapore-721 101, West Bengal, India. Email: soumenmallik@mail.vidyasagar.ac.in ; soumenmallik@gmail.com

Received: 28 November 2024; Accepted: 27 March 2025

Libraries are considered as storehouse of knowledge. Library acquires resources to satisfy the thirst for knowledge of the user community. A generic term 'library documents' is designated to the knowledge resources of libraries. Historically, 'documents' were used in legal perspective. Presently the notion 'document' is general in scope. The document horizon is found to be ever expanding. Adoption of the notion 'document' to designate library resources encountered semantic challenges. Defining documents and designating library resources as documents has ever been a dilemma to the librarians. Greatest of the library scientists like Paul Otlet and S R Ranganathan were of separate opinion regarding documents. Ranganathan defined documents on the basis of involvement of human intellect but included instrument generated records in the document category. The contemporary artificial intelligence (AI) influenced information scenario and application of generative AI in information generation extending the document landscape further. A lexicological warrant enforced the libramatic mind for postulation of suitable synthetic word 'DHISCREMENT' to designate library document of traditional sense for the sake of restoring semantic stability amongst documents and 'library documents'.

Keywords: DHISCREMENT, Document Definition, Library Document, Knowledge Generation, Scholarly Publication

Introduction

Libraries are social institutions that provide a platform to satisfy abstract thinking of the users through tangible information resources. Predominantly libraries serve the user community through acquired resources which include manuscripts, treatises, books, primary periodicals, secondary journals, popular magazines, reference tools, newspapers, atlas, globe, maps, audio, video, multimedia information and so on. Amongst these library resources all the textual information sources are commonly designated as 'library documents'. The word 'document' is used as a generic term to identify the variety of library resources irrespective of their nature of information contents. The Oxford Learners' Dictionaries defines document as – "an official paper, book or electronic file that gives information about something, or that can be used as evidence or proof of something". Libraries serve the purpose of supplementing cognitive process to generate knowledge. Conventionally it is the human being who visit libraries as library user to satisfy their intellectual thirst by acquiring existing knowledge of the creators contributed in the library resources for its subsequent assimilation and generation of new knowledge. Thus, in

traditional sense every piece of new knowledge is the result of intellectual processing or exercise of human brain i.e., the creator or author's intellect. Therefore, a generic term 'document' to designate library resources where the universe of knowledge domains is embedded seems to be inappropriate. The semantic interpretation of library resources as library documents is contradictory because documents are not defined to represent knowledge pinpointedly whereas, library resources actually represent knowledge which is generated through intellectual processing of human mind. In the age of artificial intelligence (or AI) the concept of 'documents' became semantically instable since generative AI enabling computers and machines to simulate human intelligence in generating images and texts that look like a human created them. A lexicological warrant enforced the libramatic mind for postulation of suitable synthetic word 'DHISCREMENT's to designate library documents in traditional sense for the sake of restoring semantic stability amongst documents and 'library documents'.

Objectives

As a consequence of widespread application of artificial intelligence, the present article is a diligent

effort to achieve the unitary objective of postulating a suitable nomenclature to the library resources which are traditionally called as 'library documents'.

Scope

The present article explores the views of prominent researchers on 'library documents'.

This article will compare the notion 'documents' with those of traditional 'library documents' based on the human intellectual activities in generation of knowledge.

The Document Landscape

A search into The Oxford English Dictionary reveals that the word 'document' is in use as verb since 1648 while it is recorded to be noun from the Middle English period (1150—1500). The root of the word 'document' could be traced in Latin word 'documentum' which means lesson or proof. Buckland (2018) discuss about documents in a wide range of meanings centered on teaching, instructing, showing, telling, and demonstrating. The qualifying suffix '-mentum' denotes a non-abstract object.

Historically use of documents started in legal procedures. Williams (1948) elaborated the forms and formats of materials which were accepted as 'documents' in legal procedure. According to the description of Williams (1948) 'documents' were essentially textual contents written by a person as proof of something. However, the earliest of legal procedures refused to accept pictures and paintings as documents. This is important to remember that the concept of document was confined to the statement represented as textual material only in original without any other scope of expression of the same. Interestingly, textual material which is the statement of a specific person was considered as document, but, any kind of fabrication out of that original text is not accepted as document. It seems that documents were only those evidences which could support to establish a truth and not falsehood in legal procedure.

Transformation in the concept of documents from legal perspective to generalization facilitated its adoption in librarianship through addition of an extra element - 'thought'. The semantic change of the word 'document' could be traced as 'a statement in textual form accepted in legal procedure' to definition given by Ranganathan (1956) as 'expression of thought in a physical form' (expressed embodied thought) with a universal scope of usage'.

Ranganathan's (1973) definition reflects the evolutionary lineage in extending the scope of documents beyond paper based textual materials to include photographs, films, statistical data, graphic and written records representing ideas or objects which actually consolidates the postulations of Otlet, Suzanne Briet's manifesto of documentation 1951 and many other writers. Librarians and documentalists started categorization of documents based on the nature of their information contents. Till then, knowledge was generating as a natural phenomenon of intellectual creativity to enrich mankind and libraries. As knowledge was categorized into primary, secondary or tertiary the nomenclature of document followed the analogy of knowledge categories. Albeit of being general in scope the documents of each category served the purpose of an evidence of nature of their information content i.e., primary, secondary or tertiary. A harmony in nomenclature of document types resembling the type of knowledge content could be possible in an era of unitary root of knowledge generation which involved the exercise of human intellect.

Ranganathan (1956) defined the document landscape with his brilliant conceptualization of the thought content which has been embedded in the document. His conscious exclusion of 'author' from the definition of documents could be justified by the introduction of a new category of documents for which Ranganathan (1956) coined the term 'meta-document'. It seems even Ranganathan was also concerned only about the proof of the knowledge embedded in the meta-document, otherwise it would not have been incorporated within the category of documents. Amongst the document categories of Ranganathan, meta-document is the only exception where 'thought' is not expressed and embodied rather it is the manifestation of phenomena embodied, whereas in conventional and neo-conventional documents nature of 'thought' is the basis of categorization of documents.

Document Definition is Librarian's Disguised Dilemma

The strenuous efforts of Michael Buckland (2018) in consolidating postulations of scientists over the years to develop semantics of 'documents' depicts different phases of evolution. Excerpts from his writings emphasizes the dilemma of librarians in defining scope of 'library documents'. For example,

in 1934, while defining ‘documents’, in addition to printed or written textual materials in papers Paul Otlet accepted photographs, films and statistical data. Later on, Otlet included “natural objects, artifacts, objects bearing traces of human activity (such as archaeological finds), explanatory models, educational games, and works of art” into the purview of documents. Inclusion of variety of information formats into the library resources is a compulsion of the librarians of their time but designating them as ‘documents’ is a dilemma as Otlet included all form of information communication in his definition of documents. Buckland the well-known ambassador of information science expressed the same dilemma while identifying manuscripts to subsumed in the expanding scope of ‘documents’. Nevertheless, his discussion mentions Grenersen et.al. to include animals in a zoo, even a landscape in the document definition which certainly beyond the purview of librarianship. Therefore, accepting the concept of document with its expended horizon is a dilemma for librarianship.

The dilemma in designating library resources as documents deeply rooted into the fundamental philosophy of the profession of librarianship. Professor B K Sen (2008) ushered the librimatic minds through his critical analysis of Five laws of library science postulated by Ranganathan for the first time in 1928 discussing about ‘books’ and later on revision these laws with discussion about ‘documents’ while Ranganathan himself treating ‘books’ and ‘documents’ separately.

The foregoing dilemma in librarianship is confined specifically to the adoption of the notion ‘document’ which is general in its scope. The cognitive aspect of ‘Library documents’ differ from generic ‘documents’ as the former represents the resources of libraries, hence, qualified to be designated as source of knowledge that contribute to generation of knowledge. The semantic inequality perhaps philosophically encountered by librarians over the decades. Even Ranganathan (1956) found to be imposing size and weight as parameters while considering or rejecting a material in the document categories.

Analysis of contributions of researchers over the years reflect a general impression that involvement of human intellect in generation of a material is the basis for inclusion into library resources. A conspicuous reflection of such conscious or subconscious general

consideration attains its height while Ranganathan (1956) defining instrument generated records as meta-documents and clearly mentions that the knowledge is generated through by-passing human intellect. According to Buckland (2018) cognitive aspects of the notion ‘documents’ was also considered as the basis in the educational program title “documentation studies” introduced in 1996 by the University of Tromsø.

Documents in the age of Artificial Intelligence

The emergence of artificial intelligence (AI) technologies has received attention in writing of research reports and scholarly articles. Increasing application of machine learning and natural language processing in AI-based tools to assist researchers in their writing process emphasizing quality enhancement along with the ease in the process of research writing. Salvagno et.al.(2023) discussed in details about the use of AI Chatbot in scientific writing for example ChatGPT, developed by OpenAI, that uses the Generative Pre-trained Transformer (GPT) language model to understand and respond to natural language inputs. AI chatbot especially ChatGPT is undoubtedly a useful tool in scientific writing that assist in organizing materials, generates preliminary draft for proofreading. The article given example of ChatGPT to generate text on different topics and cited that an entire letter to the editor was written by ChatGPT, answering the author's questions. Fang (2023) describes a step-by-step methodology for AI based story writing.

The volume of published literature about AI application in document creation is enormous. The documents solely created by Generative AI technology are required to be identified for library services. The dilemma of librarianship is still lies on the consideration of variety of knowledge sources as ‘documents’. The AI generated documents will not differ in form or formats of traditional documents. A cognitive approach in nomenclature of resources necessitated to decide about their inclusion into the libraries. Accordingly, this article postulates a nomenclature generic to such resources as ‘DHISCREMENT’s.

DHISCREMENT

Libraries are fundamentally deals with knowledge. Librarianship ensures the authority and authenticity of their stored or accessed knowledge resources.

Conventionally knowledge is accepted as the product of human intellect. Librarianship is the custodian of scholarly output of human intellect.

The intellectual processing of human brain initiates with some research question or noble cause. Therefore, the inquiry to resolve the intellectual thirst is the stimulus that triggers the human brain for its processing. The mechanism resembles endocrine system of human body where a research question stimulates the human mind to think and intellectual processing resembles secretion of hormones in human endocrine system.

DHISCREMENT – the term has been synthesized for designating the scholarly library resources which are ultimately contributing to knowledge generation. The synthetic term ‘Dhiscrement’ finds its root in the traditional Indian culture of knowledge generation as described in Veda Samhita written in Sanskrit language. Etymology of the synthetic term ‘Dhiscrement’ has been described in following steps:

DHI - In Sanskrit language the word ‘Dhi’ means thought or intellect. The Hindu Encyclopedia³ explains the etymology of the word ‘Dhi’ as an organ that ‘holds’ the intellect or a ‘place’ where intellect resides. To make it unambiguous a comparative analysis with another Sanskrit word ‘Mati’ has been defined as more generalized word which means thought as a mental process. Dhi indicate the activity of the intellect on a particular thought.

The Sanskrit word ‘Dhi’ has been considered as base word to synthesize a term which will replace the word ‘document’ in library perspective. Since Ranganathan conceived the concept of ‘thought’ in his definition of documents the postulated synthetic term in the present study also restoring the ‘thought’ into intellect.

SCRE – ‘Scre’ has taken from the word ‘secretion’. The root of the word secretion finds in French word ‘sécrétion’ which is taken from Latin word ‘secretion’ where ‘secretio’ means separation, and ‘secernere’ means to separate. According to Merriam-Webster Dictionary (2024) the meaning of the word ‘secretion’ is the process of segregating, elaborating, and releasing some material either functionally specialized (such as saliva). The meaning elaborates as a product of secretion formed by an animal or plant especially one performing a specific useful function in the organism.

The intellectual processing of human beings is a specific useful function to generate ‘knowledge’ that resembles ‘secretion’ of human endocrine system.

The prefix of the word secretion - ‘scre’ has been taken to mean the production of cognitive substance and thereby, added to the Sanskrit word ‘Dhi’, because, according to Collins English Dictionary (2024) the suffix ‘-tion’ indicating state, condition, action, process, or result which is irrelevant in the context.

-MENT – According to Merriam-Webster Dictionary (2024) the suffix ‘-ment’ is a noun suffix which means state or condition resulting from a (specified) action.

The synthetic term ‘Dhiscrement’ ended with the suffix ‘-ment’ to mean the result of an action i.e., human intellect (Dhi) production of cognitive (liquid) substance (the prefix ‘scre’ of secretion) to create knowledge (-ment) =Dhiscrement

Conclusion

Libraries are the recipient of the knowledge generation cycle. The knowledge requires formats to be embedded for its subsequent dissemination. Librarianship has mechanism to categorize knowledge contents and nomenclature of formats keeps pace with the knowledge categories like primary periodicals contains primary information. Being social institution libraries are susceptible to be invaded with semantically instable terminologies and the word ‘document’ is one of them. Dealing with the notion ‘documents’ is a century old saga of dilemma in librarianship. When the expanding horizon of documents incorporates even a zoo in its definition whether a library can afford to agree with that? The height of dilemma in librarianship reflected by the fact that Dr S R Ranganathan who defined documents as ‘expressed embodied thought’ was found to be discussing size and weight as parameters of a material to be called document. In the era of Generative AI application in knowledge generation it was felt inevitable for a cognition-based nomenclature of library resources which are independent of invasion by semantically instable terminologies. Accordingly, the synthetic nomenclature ‘Dhiscrements’ postulated in the present article to designate scholarly library resources of knowledge generation cycle.

References

- 1 Buckland M, Document theory, *Knowledge organization*, 45(5) (2018) 425-436.
- 2 Collins Dictionary. Available at: <https://www.collinsdictionary.com/dictionary/english/tion> Accessed on August 1, 2024
- 3 Fang X, Ng D T K, Leung, J K L and Chu S K W, A systematic review of artificial intelligence technologies used

- for story writing, *Education and Information Technologies*, 28(11) (2023) 14361-14397.
- 4 Hindu Encyclopedia. Available at: <https://vedapurana.org/single.php?s=1679&word=Dhi%3E&secure> Accessed on August 5, 2024
 - 5 Merriam-Webster Dictionary. Available at: <https://www.merriam-webster.com/dictionary/secretion> Accessed on August 1, 2024
 - 6 Oxford Learners Dictionaries. Available at: <https://www.oxfordlearnersdictionaries.com/> Accessed on August 1, 2024
 - 7 Ranganathan S R, Documentation in many lands 4. Documentation: Conventional and non-conventional, *Annals of library and information science*, 3(1) (1956) 22-32.
 - 8 Ranganathan, S R, Documentation: genesis and development Vol. 3, (Vikas Publishing House; Delhi) 1973.
 - 9 Salvagno M, Fabio S T, and Alberto G G, Can artificial intelligence help for scientific writing?, *Critical care* 27(1) (2023) 75.
 - 10 Sen B K, Ranganathan's five laws, *Annals of Library and Information Studies*, 55 (2008) 87-90.
 - 11 Williams G L, What is a Document?, *The Modern Law Review*, 11(2) (1948) 150-162.