

# **Applications of Traditional Knowledge Organization Systems in Digital Libraries: A Study on Current Status**

## **Abstract**

The present study “Applications of Traditional Knowledge Organization Systems in Digital Libraries: A Study on Current Status” was primarily designed to understand the potential applications of traditional KOSs in digital libraries (DLs). The first objective of the research was to identify and to describe the potential applications of traditional KOSs in DLs. The second objective was to investigate the effectiveness of traditional KOSs in DLs. And, the third objective was to assess the current status of the applications of traditional KOSs in existing DLs of the world. Moreover, it was also envisioned to know the possible reasons of not using traditional KOSs in existing DLs. Furthermore, it was also intended to investigate the perception of administrators of DLs towards potentials of KOSs in DLs.

Looking at the nature of the objectives set out for the study, a qualitative research approach has been adopted. Within the framework of qualitative approach chosen for the study, the research employed multiple method approach and used two different research methods - systematic literature review and survey-based exploratory study. This study has been carried out in two different, but inter-related phases – systematic literature review and survey-based exploration. In Phase I of the study, more than 200 case studies, reported in research publications, has been analyzed to identify and to characterize the potential applications of different traditional KOSs in DLs. Moreover, more than 50 earlier evaluation studies have been analyzed to illustrate the effectiveness of various traditional KOSs in DLs. In Phase II of the study, the administrators of existing DLs from around the world have been surveyed through online questionnaire. In total, 333 responses from 72 countries from around the world have been received and analyzed.

This study has identified, and described, many potential applications of traditional KOSs in DLs which may be broadly grouped into two – (i) the common primary applications and (ii) other applications which are comparatively less common. The primary applications are indexing, browsing and searching DLs. Other identified applications are alerting service / user Notification (Selective Dissemination of Information), automatic annotations/indexing of contents/documents, automatic classification of contents/documents, developing Concept Space for enhanced browsing, developing Ontologies for enhanced browsing/navigation, interoperability among DLs, linked data creation, mapping multiple DLs for browsing and searching, mapping multiple KOSs for browsing, multilingual information retrieval, publication venue authority control and visualization for browsing. Moreover, probably the most important finding of the present study is identification of two additional functionalities - Automatic Query Expansion, and Search Term Recommendation (STR) System which can exclusively be achieved by the application of traditional KOSs.

The investigation on the effectiveness of traditional KOSs in DLs has shown that they can enhance the effectiveness of DLs mainly in terms - retrieval effectiveness of the system and usability of the system. Thesaurus, taxonomy, and classification scheme have shown more acceptability in earlier usability studies. In general, all the studies have reported positive reactions of participants in terms of identification and use of these KOSs. Moreover, thesaurus has significant potential to enhance retrieval effectiveness of DLs if used with other techniques to rank the results like vector space model or Jaccard similarity measure.

All the five traditional KOSs viz. thesaurus, taxonomy, classification schemes, subject heading lists and authority files probed in the study were found pertinent in DLs. The results of this study show that there has been more research on implementing thesaurus and taxonomy in DLs and in other digital information systems as compared to that of classification schemes, subject heading lists and authority files. However, contrary to expectations, it was found that subject heading lists, classification schemes, and authority files are more used in existing DLs than thesaurus and taxonomy.

In general, applications of traditional KOSs has been found in lesser number of existing DLs. Majority of existing DLs use subject heading lists and authority files primarily used in its conventional application i.e. indexing of documents. The real potential of thesaurus and taxonomy in digital environment i.e. browsing and searching, as identified in Phase I of the study, is not being exploited by majority of existing digital repositories.

The identified reasons for not using traditional KOSs, in order of their rank in terms of their frequency, are lack of resources (funds, manpower, etc.), it was not thought of using traditional KOSs, non-availability of this feature in the software platform used for the repository, it was not known whether this feature is available in the software used for the repository, other knowledge organization/representation system/s is being used, it was not known that traditional KOSs can be of use in digital repository and few other reasons have also been identified. The results of perception analysis shows that more than 50% repository administrators agreed with the potentials of KOSs in indexing, browsing and searching digital repositories. The results show highest agreement of repository administrators with the potentials of thesaurus in indexing, browsing and searching digital repositories. This is followed by taxonomy, authority files, subject heading lists and classification schemes.

This research extends our knowledge on applications of traditional KOSs in DLs by identifying and by providing description of potential applications. Despite its exploratory nature, this study offers some insight into the effectiveness of traditional KOSs in DLs. As such, it has significant implications for administrators of DLs, developers and researchers of DLs, and educators of DLs. Administrators of DLs including digital repositories, subject gateways, online information

systems, etc. may make more informed decisions towards implementation of traditional KOSs, developers and researchers of DLs or any digital information retrieval systems may proceed from more informed standpoints.

The results of perception analysis demonstrate that more than 50% repository administrators agrees with the potentials of KOSs in indexing, browsing and searching digital repositories. And, lack of resources (funds, manpower, etc.) has been found to be the most prominent reason for not using traditional KOSs in existing DLs. As such, the findings have significant implications for authorities to decide upon adequate fund allocation towards exploiting the potentials of traditional KOSs in DLs.

Potential applications and effectiveness of thesaurus in DLs have been established in the present research. Moreover, readily available thesauri can be used while indexing the digital documents. As such, it may be suggested to the practitioners of academic DLs to select best suitable thesaurus for their purpose, a general purpose thesaurus like UNESCO thesaurus or a domain thesaurus like AGROVOC, and use it while indexing digital documents. This may also be practiced even if the repository software platform used for DLs doesn't have advanced features of using KOSs in browsing and searching.

The thesis is presented in 7 chapters. Apart from the introduction and conclusion, *Chapter 2* provides conceptual framework of the study and review of related literature. *Chapter 3* describes the research methodology comprising the two phases of the research. *Chapter 4* reports on the results of systematic literature review carried out for the study. *Chapter 5* reports on the results of survey of academic open access digital repositories. It portrays the current status of application of traditional KOSs, reasons for not using it, and perceptions of administrators of existing DLs towards potentials of traditional KOSs in digital repository, and *Chapter 6* presents the analysis and discussion.

A number of important limitations of the study need to be considered. First, the scope of the study was limited to the applications of only traditional KOSs in DLs. Applications of other forms of KOSs i.e. Knowledge Representation Systems (KRSs) could not be studied. Second, systematic review only included earlier case studies which have been published in English language and which have been indexed in LISA and Scopus only. Third, only academic open access repositories which are registered with OpenDOAR have been studied to assess the current status of applications of traditional KOSs in DLs.

It would be interesting to carry out a research study to investigate the potentials of ontologies – the most popular knowledge representation system – in DLs. Also, a more comprehensive study with larger sample of open access repositories would produce more generalized results.