STATE OF ART AND MOTIVATION FOR THE PRESENT WORK

2.1 GENERAL

A good deal of literature is available pertaining to the knowledge management and e-learning environment that lead to think on applicability of knowledge management practices in e-learning environment. This study has derived to focus on expert’s knowledge externalization in e-learning environment resulted to improve the communication mechanism available in e-learning system. The idea has triggered to find advanced expert finding techniques and automatic query routing mechanism to reach appropriate expert for knowledge transformation. In spite of the remarkable amount of publications in this area, it was felt that there is want of research improvement and enhancement of algorithms that tune the environment with unpredictable changes.

Many research papers have been reported that the externalization of knowledge take place by the transfer of knowledge with appropriate subject experts. The most of the reports deal with manual query routing and expert finding techniques. These are based on self-classification and document-based relevance approach, which doesn’t use the Internet as a potential factor to find the updated information about the expert’s expertization for expert ranking and automatic transfer of user query to the best subject expert. The collaborations in online group-based learning provide better opportunities to develop skills and knowledge compared with individual courses. The learning collaborations and communications with subject experts are highly practiced in corporate sectors by maintaining a knowledge portal for the stakeholders. An approach based on e-learning and its mechanism (Lloyd, 2003) is used to set up an organizational memory of the scientific, technical and administrative assets of the university and those who are interested in the construction of a warehouse of resources deliberately intended for training and for research. The major component of the learning
process is the result of the learner’s collaborations and communications with subject experts.

The corporate sectors utilize the knowledge management techniques for sharing, capturing and storing of knowledge for effective usability. Koskinen (2003) discussed the introduction of a new model in which the business management can evaluate the type of role that tacit knowledge plays in their organizations. The model structure is discussed with four different systems, memory, communication, motivational, and situational systems. This includes numerous factors that affect tacit knowledge utilization in organizations. To combine the knowledge management concepts on e-learning environment, it is required that the e-learning system must contain a communication process integrated with knowledge clusters (Dolog, Peter, et al 2004).

As per Nonaka & Takeuchi knowledge management concept, the transformation of knowledge categories such as: Socialization, Externalization, Combination and Internalization.

Socialization is a process which focuses on tacit to tacit knowledge linking and more over socialization is also known as converting into new knowledge through shared experiences. New knowledge is created by using the process of interactions, observing, discussing, analyzing, spending time together or living in same environment. The possibility of tacit to tacit conversion is more in direct practices for example, son learns the technique of wood craft from his father by working with him (rather than from reading from books or manuals).

Externalization is a process focuses on tacit to explicit knowledge linking. Thus the standardized storage is required to store the lessons learned from situations and mechanism has to be enabled for an on demand service to the searcher (Yi, J, 2006). Externalization of tacit knowledge requires a threat-free and emotionally secure environment based on trust and
mutual respect (Babur, et al. 2006). Tacit knowledge is regarded as an essential input of the innovation process, since it regularly instigates the idea creation process in new product development (Sigala, M. 2007). Quality circles are formed in Information Technology sectors where workers put their learning and experience, they have to improve or solve the process related problems. This topic has attracted considerable interest by providing a variety of benefits to learners, educational institutions, and organizations by removing the barriers of time and space in the development of knowledge and skill; providing just-in-time learning, convenient access, and flexible learning processes; enabling real-time content updates while avoiding information overload; reducing travel, off-site training costs and time away from the job; and facilitating the interconnectivity of people for knowledge transfer (Ozdemir and Abrevaya, 2007).

Combination is a process of transforming the explicit knowledge to explicit knowledge. Creative use of database to get business report, sorting, adding, categorizing are some examples of combination process.

Internalization is a process of transforming the explicit knowledge into tacit knowledge. The internalization of newly created knowledge is the conversion of explicit knowledge into the organization's tacit knowledge. This requires the individual to identify the knowledge relevant for one's self within the organizational knowledge. That again requires finding one's self in a larger entity. Learning by doing, training and exercises allow the individual to access the knowledge realm of the group and the entire organization.

Recent years have witnessed the enormous growth of e-learning system in the education and corporate sectors. The advancement in e-learning system has begun to reshape the learning approaches from traditional learning methodology to smart learning solutions (Craswell et al. 2007). The e-learning system plays a significant role on knowledge transformations, which includes interchanging information, opinions, experiences and perceptions. The success rate of an e-learning system can
be determined by the factors such as expert's expertization, quality of content, experts and learner connectivity, adaptability and ease of use. A major portion of these factors can be achieved by any e-learning system on fine-tuning the expert finding approaches (Li, G, 2007).

Authors Macdonald, C. & Ounis, I., (2008) and Serdyukov, P.,(2009), have surveyed the most important concepts and representative previous works in the expert finding. The two of the most popular and well-performing types of methods are the profile-centric and the document-centric approaches. It is emphasized that profile-centric approaches build an expert profile as a pseudo document by aggregating text segments relevant to the expert. These profiles are later indexed and used to support the search for experts on a topic. Document-centric approaches are typically based on traditional document retrieval techniques, using the documents directly. The people all over the world are in need of gaining knowledge which is achieved by means of searching information through web services. Even though search outcome obtained from web services are relevant, the user is trying to obtain the best information providers (Li et al 2009).

The self-classification and document-based relevance are widely used method to find the experts. Self-disclosure requires expert candidates to explicitly declare their expertise in their posted profiles. Expert recommendation systems that employ this approach include yellow pages. com, guru.com, 88 owls.com, and other opt-in directory listings of experts. The manual process is time-consuming and expertise profiles are also unlikely to remain current as each user’s expertise is continuously expanding. The difficulty with this approach which requires manual entry to keep the expert profile updated. Studies have revealed that even if efforts are frequently put into updating the expert profile, they do not provide efficient and exact outcome. The reasons for this insufficiency are that the people who are entrusted the assignment of maintaining the expert profile do not have enough knowledge of the experts. Also, manual information entry of every individual in an e-learning system was extremely laborious and
expensive. Thus, the manual entry of expert profiles into an e-learning system did not prove to be a feasible strategy. E-learning focuses on the use of computer and network technologies to create and deliver a rich learning environment that includes a broad array of instruction, information resources and solutions with the goal of enhancing individual and organizational performance (Sammour et al. 2008). E-learning suggests that efficient e-learning system should make accessible not only content but also authors who are the foundation of the provided content (Bouguessa et al. 2008).

As computer technologies are highly emerged in today’s world, the leading organizations have digitized their maximum data transferred in their organizations. These lead to digital data maintenance issue as well as issue of retaining the knowledgeable data, which is considered as organization intellectual asset. Organizations are taking higher steps to build knowledge repositories to retain their knowledgeable data on centralized shared storages. The importance of maintaining knowledge repositories is to preserve the intellectual capitals which have made the institutions to create knowledge repositories for holding institutional intellectual assets. Rajalakshmi and Wahidabanu (2011) stated that learning mechanism in higher education has to be customized in accordance to the new requirements and changes of people, culture and economy. This requires new models of management for its logical capital. Knowledge Management can facilitate universities to ensure knowledge collection by providing a way to reach the appropriate expert at the right time to take the worthy decisions using Knowledge Management Systems.

Rajalakshmi and Wahidabanu (2011) proposed a model known as Info-Ca-Sh, a contributed knowledge portal of dynamic web content activities. The design flow serves as an exchange of knowledge among users by providing the users with a range of open-source tools. The tacit and explicit knowledge of the users are captured and externalized as a knowledge repository. The learning mechanism in higher education must be customized in accordance with the new requirements and current shifts in users, the
culture and the economy. This effort requires new models for management of logical capital. Knowledge management can facilitate universities to find the appropriate people at the right time to make the best decisions using knowledge management systems.

A number of studies have investigated the use of intelligent tutoring techniques, which have personalized learning interfaces and adaptive learning. These efforts generally emphasize technology development but exhibit little concern for effective instruction or pedagogy to enhance learning performance (Wang et al 2011). Venkataramanan et al (2011) presented their the idea on the authorities and staff of the Education who should pay special attention to better and more effective use of ICT to enjoy more of the positive and useful applications of these tools in educating and solving their problems.

The major component of the learning process is the result of the learner’s collaborations and communications with subject experts. In an expert search task, the learners require to discover people who have related expertise to a topic of concern. An expert search method predicts and ranks the expertise of a set of people with respect to the users’ query. Ahmad Kardan and Fatemeh Hendijanifard (2011) discussed the topic of finding subject experts for problem solving as an important issue in an e-learning environment. In an e-learning environment, there is no direct approach to finding the superior individuals. The current methods of analyzing the discussions or considering the learner require a large amount of data and contain limitations. In this work, concept maps were used to define the experts in an e-learning environment.

The task of expert finding is intended to identifying people with relevant expertise skill or experience for the given topic. The topic of finding subject experts for problem solving is an important issue in an e-learning environment. In this environment, there is no direct approach to finding the superior individuals. Kardan, K. and Hendijanifard, K.,(2011) discussed the
current methods of analyzing the discussions forum and considered the learners' expertise which require a large amount of data and contain limitations. In this work, concept maps are used to define the experts in an e-learning environment.

In review with Mirsaeed Ghazi, et al (2011), it is been identified that there is no solid approach has been carried out to connect the expert finding system with Internet resources, were the system can verify expert expertise skill set entered manually by the experts and update the expert profile automatically on expansion of the expert skill sets. However, the term e-learning is used in this study to encompass computer-based learning, computer-based training, technology-enhanced learning, technology-mediated learning, web-based education, or virtual learning environment (Waheed et al 2012).

A knowledge portal acts as an access tool for other information sources to provide internal and external information beyond the organization's own resources that can be made available to its staff. The portal also serves as a communication tool to enable individuals, teams and communities of practice to share and discuss ideas and knowledge (Venkata Subramanian, 2013). With the rapid development of globalization and technology, the importance of knowledge has also gradually increased. The enterprises can effectively create, accumulate, utilize and manage knowledge and convert knowledge into a competitive weapon (Chen et al 2014).

Senthil kumaran et al (2014) discussed that the success of any e-learning system depends on the quality and quantity of assistance provided to its students in the learning process. Hence, it is essential to analyze a student's academic skills to personalize the education that is provided both vertically and horizontally.
To create a collaborative knowledge network, it is necessary to build a cluster of knowledge providers in the e-learning system. Fuzzy C-Means (FCM) algorithm is highly suggested to create clusters. In faster fuzzy clustering (also referred to as soft clustering), data elements can belong to more than one cluster, and a set of membership levels is associated with each element. These clusters indicate the strength of the association between that data element and a particular cluster. Fuzzy clustering is a process of assigning these membership levels and using them to assign data elements to one or more clusters. In many situations, fuzzy clustering is more natural than hard clustering. Objects on the boundaries among several classes are not forced to fully belong to one of the classes but rather are assigned membership degrees between 0 and 1 that indicate their partial membership. The fuzzy c-means algorithm is used in computational geometry, data compression, vector quantization, pattern recognition and pattern classification. FCM is an unsupervised clustering algorithm which has been applied to a wide range of problems involving feature analysis, clustering and classifier design. The FCM clustering, which constitutes the oldest component of software computing is quite suitable for handling issues related to understanding patterns, incomplete/noisy data, mixed media information, and human interactions. It can provide approximate solutions (Yogeshwari & Balamurugan, 2014).

Expert's finding is comparatively a new area of research with ample scope. In recent times, a large number of researchers have put their significant efforts in this area (Zhu et al 2014 and Zhao et al 2013). The automatic search for knowledgeable people in the scope of specific user communities with basis on documents describing people’s activities is an information retrieval problem that has been getting increasing attention. Usually referred to as expert finding, the task involves taking a short user query as input, denoting a topic of expertise, and returning a list of people sorted by their level of expertise in what concerns the query topic. Several effective approaches for finding experts have been proposed, exploring different retrieval models and different sources of evidence for estimating
expertise. However, Moreira M et al (2015) discussed that the current state-of-the-art is still lacking in principled approaches for combining the multiple sources of evidence that can be used to estimate expertise. Documents which are written or reviewed by an expert candidate are also a useful expertise indicator (Liu et al 2015). There are a number of expert finding systems that use text mining techniques to automatically capture the authors’ expertise.

The collaborations in online group-based learning provide better opportunities to develop skills and knowledge compared with individual courses (Liu et al 2015). The hybrid approach was carried to effectively find experts for the category of the target question in question-answering websites. Different from the conventional approaches that only consider user profile or user authority, but their study considered user subject relevance, user reputation and authority of a category in finding experts. However, to improve existing e-learning applications, smart learning environments must provide personal services to assist a learner in using, managing and interacting with the learning system (Judrups 2015).

Many schools have availed course management softwares (e.g., Blackboard, WebCT, and Moodle) to complement traditional classroom-based instruction. Many empirical studies have been conducted to demonstrate how IT supports learning improves student learning outcomes and enhances student information literacy (Nitnaware et al 2016). Despite the variation in research findings, there is a consensus that substantial gains in student talent are achievable if the use of Information Technology in schools is planned, structured and integrated effectively.

The knowledge management system refers to a system for customizing knowledge in organizations in terms of capture, storage information and broadcasting of knowledge (Pushpalatha A.M and A.Nirmal Kumar, 2016). In the last century to the present, the matter of management of intangible assets as part of the vital resources is considered. Among these
assets, knowledge is a most important among other intangible assets as stated by Monisha et al (2016).

Existing expert finding techniques often rely on the following indicators to find one’s expertise area and the level of expertization: Self-classification and document-based relevance. Self-Classification basically includes the expert’s expertization or skill sets are manually entered by the expert and document-based relevance is evaluated based on relevant document available in the system. Most automated expert finding techniques rely on document-based relevance to measure the expertise level of an expert to assign a user query. This technique assumes that the relevance of one’s authored documents to the query is positively related to their expertise level on the query.

These perceptions clearly depict that the expert’s expertise levels were only measured with the data available in the system. Therefore, the existing expert finding approaches cannot be considered as an accurate evaluation on expert’s expertization. Since the existing expert finding techniques doesn’t use the Internet as a potential factor to find the updated information about the expert’s expertization.

It is required to have the expert’s expertise keywords along with the expert identity data such as expert name, co-authors and email id to find the expertization level of an expert using Internet. The expert given expertise keywords along with expert identity data have to be cross verified with the data available in the Internet to ensure the contribution made by the expert on the given expertise area. The verification has to be made in various resources in the Internet such as websites, blogs and journals etc. But most of the websites and journals require authenticity to access the resources available in their database. These constraints create a barrier to the researchers on utilizing the Internet as an important data source to evaluate an expert’s expertise level.
To overcome this problem, an approach has been proposed to ensure the expertise level of an expert by utilizing the search engines as an agent to extract the expert's data available in the Internet and weight according to their contribution made towards the given expertise area. The search engines are used as a data retrieving agents, since they can access Meta data or data of any websites without any authenticity of the accessing website. The expert's keywords have to be searched along with the expert name to extract the data of the expert related to the given keyword. This action will provide a bulk of results linked to the expert contribution towards the expertise keywords.

The results produced by the search engines would contain many junk data. So the entire result of the search engine cannot be considered as a valuable data input to measure expert's expertise level. Hence it is required to remove the junk data and extract the records related to the expert and the given keyword. In addition, these results have to be processed with constraint parameters to find the exact record of an expert from the extracted records. The proposed data extraction technique gives a valuable input data with reasonable accuracy to measure the expertise level of an expert towards the given expertise keyword. This method can be considered as the best solution to extract the updated expertization level of an expert towards the expertise area.

The need of knowledge seekers to directly connect with the domain expert for collection of tacit knowledge and transforming it into externalized knowledge has motivated to carry out this research work.

2.2 OBJECTIVES

The primary objective of this research is to create a dynamic query handling framework to externalize the experts’ knowledge in an e-learning environment by streamlining the communication process between knowledge
seeker and knowledge provider for capturing and sharing of knowledge with the right person at the right time. The problems considered are

- Construction of Basic Dynamic Query Handling Model (DQH Model) that can automatically transfer a user query to the best subject expert available in the system.
- Implementation of Internet Data Relevance and Expertise Mapping model (IDREM) which can rank the experts from relevant expert data available in the Internet than using traditional method of self-classification and document based relevance expert ranking.
- Development of Improved Internet Data Relevance and Mapping model (IIDRM) considering the experience and qualification of the expert as a leading factor on expert ranking along with Internet verified expertization level of the expert to suggest the best subject expert for query transformation.

2.3 SCOPE OF THE THESIS

The thesis entitled “Externalization of Experts’ Knowledge in E-learning Environment “delineates the applications of knowledge sharing in e-learning paradigms such as expert finding and dynamic query routing mechanism and the development of new approaches and algorithms to explore knowledge externalization in e-learning environment.

This dissertation is divided into six chapters. Chapter 1 introduces the thesis. Well established and the very significant communication tools of e-learning system have been reviewed in chapter 2. A dynamic query handling model (DQH) is proposed in chapter 3, the DQH provides the ease of use to the learning community in e-learning environment. It has the significance of taking a query from a user to a suitable subject expert automatically. It offers a path to reach a potential expert for solution of a problem query and
opportunity to collect valuable knowledge from the expert and store it in the knowledge repository.

The mathematical measurements derived for the query transformation is given in the Equation (1.1).

\[ QT = ((KY \cap ED) \rightarrow XP((ER>OXP) \text{ AND } (BC<TH))) \]  \hspace{1cm} (1.1)

The query (QT) is transferred to the expert (xp), and the keywords (KY) extracted from the query using the NLP parser are used to map or match with the expert domain list (ED), that is, the intersection of the keywords and expert domain list contains a common word. The resultant expert's rating (ER) must be greater than those of other experts (OXP), and the query bucket (BC) must be less than the threshold limits (TH).

The algorithm generated for the query transformation is aimed to provide automatic query routing facilities to the learning community to connect the learner with the subject expert for knowledge transformation. In this algorithm, learner raised query is processed with NLP parser to remove stop-words and retrieve informative words. The extracted words (EW) are compared with expert stored expertise skill sets or expert domain to relate the query with the appropriate expert. Once a match is found between the extracted words and the expert keywords, the user query is mapped with the expert and expert domain and the mapped expert's query bucket threshold limit is checked before transmitting the query to the concerned expert. If the expert's query bucket is full or the threshold has been reached, then the query will be redirected to the next expert. If the query is not matched with any of the expert keywords, then the query will be moved to the dispatcher. The dispatcher or the administrator will divert the query to an expert by manually finding the expert or the query will be forwarded to the appropriate knowledge management team to address the issue and resolve it.
The mechanism of DQH creates a way to the learner, where the user does not need to know the subject or expert group to which the query should be routed, and the experts in the dynamic query handling system are analyzed based on work performance and the solution delivery ratio that they uphold. The experts are ranked and re-ranked based on the performance that they maintain in the system, which retains the best performing expert as first priority on the topic or subject. The dynamic query handling system leads to a method for ranking experts, capturing expert knowledge, ensuring the best solutions for the query and providing a user-friendly environment to learners.

Chapter 4 has emphasized the ranking of experts based on expert data available in Internet rather than self classification or document based relevance methods. Since the existing expert finding techniques do not use the Internet as a potential factor to find the updated information about the expert's expertization for expert ranking, this thesis shows the path to find the expertization level of an expert from the Internet, using the expert’s expertise keywords along with the expert identity data such as expert name, co-authors, email id etc.. Internet data relevance and expertise mapping is achieved by implementing an unbiased expert ranking system to cross verify the experts’ expertise level globally from data available in Internet.

Introduction of this new approach, ensures the expertise level of an expert by utilizing the search engines as data retrieving agents to extract the expert’s data available from the Internet and assigns weight according to their contribution made towards the given expertise area. This action will provide a bulk of results linked to the expert contribution towards the expertise keywords. The results produced by the search engines would contain many junk data. So the entire result of the search engine cannot be considered as a valuable data input to measure expert’s expertise level. Hence it is required to remove the junk data and extract the records related to the expert and the given keyword.
The algorithm generated to find data related to the expert uses the results retrieved from the search engine. The retrieved records are processed with NLP parser. The cleaned data records are then compared with expert supporting parameters like (Expert Name (EN), Expert Current Designation (ECD), Email id of the expert (EID), Expert Previous Designation (EPD), Current Employer (CE), Previous Employer (PE), Co-Authors (COA), Current Organization Website (CWEB), and Previous Organization Website (PWEB). To ensure the number of records belong to the expert. This could be made possible on comparing each and every output record with the expert parameters list. The highest match could be considered as the actual record of the expert from the multiple records. The same process has to be continued to retrieve the actual records of the expert and keyword.

The weightage for each expert keyword is obtained by subtracting the number of records actually belongs to Expert from the Total Number of Records Retrieved using Search Engine. The Target Keyword Relevance and Partial Keyword Relevance are calculated by comparing the expert keyword with the Records actually belongs to Expert. The Expertise Level is then generated by adding the Target Keyword Relevance by multiplying with 0.9 and Partial Keyword Relevance with multiplying 0.1.

In Chapter 5, it is highlighted that consideration of experts’ experience and qualification which are also one of the important factor in expert finding decision. The consideration of experts’ experience and qualification along with Internet Relevance Data will give high accuracy in expert finding. Multiple regression analysis has been used to consider the experience and qualification of the expert as a leading factor on expert ranking along with Internet verified expertization level of the expert. Multiple regression analysis has been used to obtain the best subject expert for query transformation.

In Chapter 6, a review of work reported, major conclusions related and contributions made are dealt with. Recommendations for future research are also stated.