A Framework for the Construction of Ontology for ICT Experts Implementation

Authors

Akmal Aris, Juhana Salim, Shahrul Azman Mohd Noah and Kamsuriah Ahmad

Universiti Kebangsaan Malaysia, Bangi, Malaysia
Abstract

Systems that help to find answers to suitable experts have received the attention of many researchers. Pioneer researchers on the development of search system for experts emphasized the importance of a search system that required returning a list prioritizing the names of individuals. Among the issues raised related to locating experts are, the critical problem of maintaining up-to-date information in the expert database and the inadequacy in expert finding systems in returning search results that are expected to account for not only the list of the names of the experts, but any information related to the experts and those involved with them. In the last decade, researchers have examined the search for experts from various research perspectives such as expert tracking system and construction of
expert profiles involving ontology. This paper aims to describe a new approach in designing a framework for the construction of ontology that will be used in the directory of ICT experts. The researchers propose to incorporate thesaurus in their construction of ontology on ICT experts by providing a profile of the experts including their social profile and whatever concerns that may be associated with the expert. The researchers constructed the ontology on ICT experts by extracting information from sources such as their resumes and personal web sites to obtain the index glossary of words that characterized experts to be used in the directory of experts system. In this study, the standards contained in the Performance Evaluation System of the National University of Malaysia are used to extract information on the academicians in the Faculty of Information Science and Technology. The researchers employed the index
glossary, metadata of experts and integrate appropriate taxonomy, thesaurus and classification schemes; such as Association for Computing Machinery (ACM) taxonomy, web classification schemes such as Standard Industrial Classification (SIC), North American Industry Classification System (NAICS) to enrich the ontology of ICT expertise. The proposed framework aims at helping users find information on the expert they require and at the same time obtain other information related to the experts from various perspectives encompassing research, consultancy, links with research partners, other interest related to a particular field and other resources. With a semantically driven directory of ICT experts, matters related to ICT can be referred to the right experts. The framework will be validated by developing the directory of ICT experts prototype and by involving domain experts to evaluate the content of the ontology
constructed. In addition, the researchers will evaluate the search results of users who will use the prototype to search for ICT experts.

**Keywords:** Expertise Tracking, Expertise Ontology, Profiles of Experts, Access to Experts.
Introduction

Experts are individuals who play an important role in the success of an organization and are considered critical in creating a value to the organization. The expert skills and knowledge very often can be channeled through consultation, mentoring systems and corporate memory. According to D'Amore (2008), experts are generally located in the formal or informal workspace and their work is based on the domain and culture of the organization. Yiman-Seid and Kobsa (2003) have identified several goals in allocating the experts based on defining the problem, evaluation and analysis, filtering information and project assignments. However, there are problems in the placement of experts in an area where there are difficulties in finding them based on the work context and work locations that are not parallel (Yiman-
Seid and Kobsa 2003). Studies on the search for experts have focused on the task of seeking individuals who have the skills and knowledge that are particularly suited to answer the question "Who is an expert in the field of X?" The task to find answers to suitable experts have received the attention of many researchers and most of them concentrated on experts profiling with search applications using query that describe the areas of expertise being sought. Pioneer researchers on the development of search system for experts emphasized the search system that required to return a list that prioritized the names of individuals (Balog and de Rijke 2007, Hawking 2004, Karimzadehgan, White and Richardson 2009, Stankovic et al. 2010). Among the issues raised in relation to locating experts are, the critical problem of maintaining up-to-date information in the expert database and the inadequacy in expert finding systems in returning search
results that are expected to account for not only the list of the names of the experts, but any information related to the experts and those involved with them. For example, in tracking an expert, the details that one would be interested in, include: the area of expertise, who are working with the expert, other information needed to contact the expert and the individual / organization or others that have the same expertise and relationships with the expert. Therefore, to develop a system that allows the tracking of experts, there is an imperative need to identify the complete metadata in profiling an expert. These include describing the collaborative environment of the expert.

Initial approach used to search for experts involved the development of a database that stored information about the skills and knowledge of individuals in an organization (Maron et
al. 1996; Davenport and Prusak 1998), and merely focused on how to integrate the database containing the same information found in a data warehouse where it can be mined to obtain information on the expert. Most of the earlier researches were carried out by communities in the area of knowledge management and their research output includes yellow pages and systems for searching experts (Becerra-Fernandez 2000).

In relation to the research on the profiling of experts to support the development of the directory of experts, another approach used by previous researchers is the use of ontology (Stankovic et al. 2010). Potential use of ontology in data related to the experts has been demonstrated in studies by Meza et al. (2007) where the researchers have combined Resource Description Framework (RDF) vocabulary to search for experts. Meza et al. (2007)
designed a framework for the reuse and expansion of vocabulary / thesaurus in the semantic web. The framework is designed to support efforts to develop the ontology considered important in allowing the expert tracking system to find information about an expert not only in terms of qualifications and brief resume, but whatever concerns that may be associated with the expert. As long as such information exist on the Internet, the system can find it. To date, literature on expert tracking system indicates that very little research has been done in the building of ontology to support expert tracking.

This study will develop an ontology-based approach applied in previous studies that used thesaurus and ontology as a method to extend the search for information in various domains. The researchers propose to incorporate thesaurus in their
construction of ontology on ICT experts by providing a profile of the experts including social profile where other things and other parties related to the expert could be found.

The body of this paper is organized as follows: Section 2 discusses the literature review; Section 3 defines the research objective followed by Section 4 which describes the research method. In Section 5, the researchers offer their conclusion regarding the potential outcome of their framework and the feasibility of their approach in tracking experts and linking other aspects related to the expert.
Studies in search of experts focused on the task of seeking individuals with appropriate skills and knowledge. Since the last decade, researchers have examined the search for experts from various research perspectives such as expert search system (Hawking 2004; Fu et al. 2006; Balog and de Rijke 2007; Chen et al. 2007; Haruechaiyasak 2009; Karimzadehgan, White and Richardson 2009; Tu et al. 2010), construction of expert profiles (Whittaker et al. 1997; Ackerman and Halverson 1998; Crasswell et al. 2001; Krulwich and Burkey 1995; Mockus and Herbsleb 2002; Yiman-Seid and Kobsa 2003; Trajkova and Gauch 2004; Liu et al. 2005; Fu et al. 2006; Yang et al. 2008; Serdyukov 2009; Reichling and Wulf 2009; Stankovic et al. 2010), identification of expert (Hinds 1999; McDonald 2001; Pipek et al. 2003; Zhang et
al. 2007) and ontology. This literature review has critically evaluated expert tracking system, expert profile and ontology on expert.

**Expert Tracking System**

Development in research on expert tracking system shows that it is a fast-growing field and the issues discussed above have been given much attention especially on the approaches used to address the related issues and problems. Previous researchers focused on the task of creating a complete profile of experts and integrate ontology development to be applied with semantic web technology. Research effort in profiling experts is a new endeavour.
Studies on the profiling of experts in the expertise search systems have attracted the attention of many researchers. Previous researchers have examined the profile of experts from various angles such as research in the construction of expert profiles (Whittaker et al. 1997; Ackerman and Halverson 1998; Crasswell et al. 2001; Mockus and Herbsleb 2002; Yimam-Seid et al. 2003; Trajkova dan Gauch 2004; Liu et al. 2005; Fu et al. 2006; Serdyukov 2009; and Reichling and Wulf 2009) and the identification of expert and use of research information (Wu et al. 2010; Latif et al. 2010). Expert search approach begins with the search by using profile information to obtain results matching the experts (Crasswell et al. 2001; Liu et al. 2005). The search for expertise based on profile is the first step in automating the search for expertise in the organization and to avoid manual maintenance of personal profile information such as resume and
personal web pages (Serdyukov 2009). According to Trajkova and Gauch (2004), profiles can be constructed based on questions raised or based on the observation of the user activity. A user profile is usually delivered using keywords or vector concepts.

In particular, researchers have focused on topical and social profiling of experts. Yang et al. (2008) highlighted some of the questions to be answered in building the profile of expert which includes: "What is the area specialization of the expert?", "Who is the expert in a particular area?", "Who is familiar with this type of expertise?", “With who is the expert working?”, “What is the detailed information to enable the experts to be contacted?” and “Who are the individuals
organizations or parties that have the same expertise and relationships related to the experts?\”

To answer the above questions, previous studies have developed systems that can assess the expertise in individuals and determine the expert best suited with whatever issues that need to be addressed and problems to be solved. Mc Donald (2001) points out that there are problems faced in terms of matching the appropriate documents and expert on the problems faced. He proposed several changes to be made in order to match the human expert with issues and problems to be solved. Mc Donald proposed that humans can assess the expertise in individuals and systems can determine the appropriate expert based on the performance demonstrated in solving a particular problem and the level of performance measurement of the expert to determine
the appropriate experts with topics searched. Zhang et al. (2007) highlighted two steps in the search for experts in the social network; that is, by using personal information to determine the scores and the propagation-based approach to identify existing relationships with the experts. Li et al. (2006) look at personal contact information from the perspective of social networks that can be connected through four types of relationship: knows, collaborates, collaborated and consulted by.

Balog et al. (2006) proposed two strategies in the search for experts. First, directly model an expert’s knowledge based on the documents that they are associated with and second, locate documents on the topic and then find the expert associated with the topic. According to Krulwich and Burkey (1995), the involvement of individuals with forums discussing particular
topic on the Internet can help in building expert profile. Haruechaitayak et al. (2009) stressed on the fact that the main purpose of searching for an expert is to identify persons who have specialized knowledge. As expert profiling refers to the study that focuses on the identification of areas of expertise with a specialist, it is important to note that expert profiling needs to be explored in the research involving experts searching.

Previous researchers have developed a number of tools that can help in the search for experts. Fu et al. (2006) have designed the description document model known as ‘Person Description Document' for a more effective search for experts. Information such as the features and context of the relevant experts are extracted to build an expert profile known as self description documents for experts. However, several issues were discovered
relating to the information resources that are of various types such as websites, electronic mail and databases. In addition, documents containing information of different experts that are found in different formats such as in HTML, PDF and WORD, complicate the extraction process and the consolidation of useful information. Fu et al. (2006) found two major problems encountered in searching experts in the organizations. First, the information on experts are distributed across organizations from different sources and in different formats. Second, most of the information on experts is not fully documented and only part of the document contains information on the experts involved. To overcome these problems, Fu et al. (2006) used Person Description Document (PDD) to compile expert information to a central unit that can be used to profile experts. The advantage of using the self description document is to enable the extraction of
expert information to be more flexible and easier to read. In other words, self description document is the index information about the expertise that allows efficient access to the documents related to the experts.

Reichling and Wulf (2009) proposed the application of expert search incorporating two mechanisms in the construction of the expert profile that are: the generation of keywords semi automatically and the use of Yellow Pages approach. Through the construction of keyword profiles, a list of large-scale keywords acquired from the doc., pdf., html and txt file selected from the file system users can be generated. In building the right profile and to protect the privacy issue, experts are allowed to choose their own files or folders containing their documents during the generation of keywords. To obtain a more accurate keywords profile, there
are words such as stop words that need to be filtered and disposed. The second mechanism proposed by Reichling and Wulf (2009) in the construction of expert profiles is to use the Yellow Pages approach where experts will enter the contact information and other information, such as educational background, job description, qualifications, competency of other language and other personal information. According to Reichling and Wulf (2009), using the Yellow Pages approach, experts will develop a profile of their expertise.

**Ontology on Expertise**

Another subject of concern in relation to some previous researches in expert tracking system is ontology. Ontological approach in the search for expert is to identify appropriate
expertise to the keywords found with additional information related to the results given by the expert search system. Among the researchers who have studied the use of ontology in the expert tracking system are Liu et al. (2007); Meza et al. (2007); Stankovic et al. (2010) and Punnarut and Sriharee (2010).

Liu et al. (2007) proposed the use of ontology on experts to integrate the various indicators of expertise from diverse data sources and the use of domain ontology to replace the search for experts based on concept rather than on keyword. The purpose of Liu’s study is to increase the chances of finding the relevant experts and assist users to select the appropriate expertise by providing more detail information for each expert. According to Liu et al., domain ontology is constructed not only to save the key
concepts but also the concepts related to a specific domain found in a collection of documents.

Xing et al. (2009) whose research incorporated the use of thesaurus on ontology construction explained that there are four key elements in ontology development; they are: terms, hierarchies, semantic network and ability reasoning. According to Chang and Lu (2001), hierarchy is another important element of ontology. Hierarchy among different object classes refers to inheriting relationship (is-a, kind-of, part-of), while hierarchy among different classes refers to combination relationship (intersection, union, inverse set, complementary set of other classes). By this way, terms (concepts) can be connected together by hierarchy. The fact that thesaurus as vocabulary table also has hierarchy makes it, after a little change and process, possible to
be used in ontology. Xing et al. (2009) justified the reason why thesaurus is useful in building ontology as follows: (i) the standard of terminologies and professional division of thesaurus can satisfy the request of clarity, the completeness and coherence on ontology; (ii) the extendibility of terms so that one can continue adding new terms without changing original terms; (iii) to save time in building ontology because it consumes a lot of time if we fully depend on domain expert. However, the domain expert is still needed to add more attributes and relationships to the ontology because thesaurus is found to be lacking of relationship (Lauser et al. 2006 and Xing et al. 2009). Thus, thesaurus is necessary in building efficient ontology on expertise.

Such approach is to overcome the weaknesses faced in the traditional way of finding experts that commonly requires
database of expertise/skills. Punnarut and Sriharee (2010) used data mining techniques to identify the expertise of researchers with skill classification ontology. Skill classification ontology is a model of expertise that contains the expertise in the field of computer and information science. The purpose of Punnarut and Sriharee using skill classification ontology and research profile of experts is to develop an expertise search system with semantic matching.

In this context, the present study will identify the information related to the search for experts on ICT based on previous research studies resources. Based on the analysis of information search on experts, the researchers found that previous researches have focused on semantic search of experts using the names of the experts. There is also a research effort to search for
experts based on the experts’ publications searchable from the database containing articles and papers of experts. But for a good expertise search system, the system not only need to look for the expert that need to be tracked to solve issues, but should be able to make relationships with the experts from various perspectives such as research, consultancy, links with research partners, the interest in a particular field and other resources related to the experts. In the context of this study, this study will not be doing the identification of the status of the expert or measure the levels of expertise, but will focus on providing results that match experts with the information required by the user. Different perspectives need to be taken into account to enrich the relationship which exists in connection with experts to get results that are relevant to a search topic. In the context of this study, sources of information on the researchers in the research groups
of the Faculty of Information Science and Technology, National University of Malaysia will be used. All the information available to the researcher's resume will be extracted and stored in the word index to be used in the construction of ontology of ICT experts.

**Research Objective**

In order to fill in the research gaps, the main objective of this research is to design a framework for the construction of ontology on ICT experts based on expert profile. To achieve the main objective, the following sub-objectives were identified: i) to explore the approaches in profiling experts as the basis to the construction of ontology on ICT experts; ii) to integrate thesaurus and classification systems in the development of a semantically
driven directory of experts; and iii) to evaluate the effectiveness of the expert directory.

Methodology

To achieve our first research objective, the researchers of the present study conducted literature search from various databases such as ACM, Science Direct, Springer Link and Citeseer to acquire research based articles on expert tracking and the profiling of experts. They then analyzed the approaches that past researchers used in solving the problems related to tracking experts and the profiling of experts. In relation to past researches on the profiling of experts, it is discovered that there were researchers that used ontology. Specifically, it was found that researchers used thesaurus and classification schemes in
constructing ontology. Therefore, to achieve our second research objective, an expertise ontology will be constructed which will be explained in more detail in our framework.

Based on the critical evaluation of past research and literature related to the areas mentioned earlier, a framework was designed for the construction of ontology on ICT expertise based on expert profile (See figure 2). As shown in the framework, the area that will be studied includes:

i. Experts system and Profiling of experts.

ii. Semantic Web Technology.

iii. Ontology on ICT expertise.
iv. Thesaurus, taxonomy and classification scheme.

v. Directory of experts.

The framework begins with the identification of major problems faced by users in finding experts and expertise based on user demand. Semantic web technologies will be studied in order to apply in the construction of ontology on ICT experts to provide more relevant search results for expert tracking. The information needed will be extracted from various sources such as resume and personal website to be used in the construction of ontology on ICT experts.

To achieve our second objective, based on previous studies by Fu et al. (2006) who used Person Description Document (PDD),
the ontology on ICT experts will be constructed by extracting information from sources such as resume and personal web sites to obtain the index glossary of words that characterized experts to be used in the directory of experts system. In this study, the researchers of this paper will use the standards contained in the Performance Evaluation System of the National University of Malaysia to extract information on the academicians which include the following information: personal, education, teaching, supervision and reviewing, publications, research, consultancy, conferences, invention / innovation / product, or contributions to the university administration, community services / activities, student services, information services, professional qualifications, membership of professional bodies, awards and insignia and training / short courses / workshops [See Figure 1].
Figure 1. Information Available on University Performance Assessment System
By having this ontology of ICT experts, the expert directory will not only be used to help accelerate the search but also show the relationship between experts with other areas such as publication, research, conferences and other activities and other people associated to the experts. The researchers will use the index glossary, metadata of experts and integrate appropriate taxonomy, thesaurus and classification schemes, such as Association for Computing Machinery (ACM) taxonomy, web classification such as Standard Industrial Classification (SIC), North American Industry Classification System (NAICS) to enrich the ontology of ICT expertise. The ontology of ICT expertise is the main engine for semantic web retrieval which is a technology for semantic web.
Figure 2. Ontology Development Framework for the Directory of Experts
Once the ontology on ICT experts had been constructed, a prototype directory of expert is then developed as a source of reference to facilitate the tracking of experts and their areas of specialization. It will be developed using System Development Life Cycle (SDLC) in order to validate our framework. The ontology constructed will be integrated into the directory of experts to enable the directory to provide a semantically driven search capability. The semantically driven search capability will be tested by matching the search result with the requirements of users that used the directory of experts.

There are two validation processes involved in order to validate our framework. Firstly, the ontology of ICT experts will be validated by the experts in the ICT domain to evaluate the concept, attribute and relationship which existed in the ICT
experts ontology that have been constructed. The second validation process will involve content evaluation. The users will be instructed to use the ICT expert directory prototype to search an expert based on their query. The result from the query term which is entered by the user in the directory will be tested to know whether the results are relevant and meet the requirements of users. A qualitative study will be conducted to obtain feedback from users about the effective use of ICT experts ontology in directory of ICT experts prototype.

**Conclusion**

An important outcome from this research is a framework for building ontology of ICT experts for an ICT expert directory which applies a new approach in creating ontology. This research
involves the application of a new approach to create ontology by using well established thesaurus such as Library Congress Subject Heading and ACM Taxonomy including web classification schemes such as NAIC and SIC and various references sources. This proposed framework aims at helping users find information on the expert they required and at the same time obtain other information related to the experts from various perspectives encompassing research, consultancy, links with research partners, other interest related to a particular field and other resources. With a semantically driven directory of ICT experts, matters related to ICT can be referred to the right experts.
References


International Joint Conference on Artificial Intelligence (IJCAI'07), San Francisco, CA, USA, 2657-2662.


