World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:6, No:12, 2012

Ontology-based Query System for UNITEN Postgraduate Students

Zaihisma C. Cob, Alicia Y.C. Tang, and Sharifah J. Syed Aziz

Open Science Index, Computer and Information Engineering Vol:6, No:12, 2012 publications.waset.org/10947/pdf

***Abstract*—**This paper proposes a new model to support user queries on postgraduate research information at Universiti Tenaga Nasional. The ontology to be developed will contribute towards shareable and reusable domain knowledge that makes knowledge assets intelligently accessible to both people and software. This work adapts a methodology for ontology development based on the framework proposed by Uschold and King. The concepts and relations in this domain are represented in a class diagram using the Protégé software. The ontology will be used to support a menu-driven query system for assisting students in searching for information related to postgraduate research at the university.

***Keywords*—**Ontology, Protégé, postgraduate program, query system.

I. Introduction

IN 2008, the Malaysian Ministry of Higher Education launched the MyBrain15 scheme that aims to increase the number of postgraduate degree holders – especially at the doctoral level – to promote higher learning to fill the talent gap. The scheme aims to produce 48,000 PhD holders by 2020 and 60,000 PhD holders by 2023, as laid out in the 12th Malaysia Plan [1].

Consequently, the number of postgraduate students has increased about six-fold, from 4,868 in 1990 to 30,477 in 2000. In 2010, local postgraduate enrolments rose by 23% to 99,852. This figure consists of 70.72% Master's, 23.27% PhD, 3.17% postgraduate diploma and 2.82% professional qualification enrolments [2].

Before embarking on a postgraduate program, a candidate will have to research on available programs and the potential supervisors. The task of searching for postgraduate information often returns voluminous amount of data that requires a lot of effort to filter. When a set of keywords is provided by the user, the search engine returns all documents that are associated with these keywords. This is due to the conventional keyword-based searching of relational databases that do not contain metadata to describe the meaning of the data.

Semantic technology can be deployed to provide intelligent support for postgraduate student to search for relevant information. The semantic application will be able to perform various tasks such as enabling students to check entry requirements, and permitting students to select their supervisors.

One of the major components of the Semantic Web is ontology, which is recognized as a fundamental difficulty of any Semantic Web Services [3]. For each domain of the human knowledge, ontology must be constructed, partly by hand and partly with the aid of automation tools. The use of ontology in the e-science community determines the ultimate success of the Semantic Web [4]. Several tools such as ontology editing tools, ontology merging tools, and ontology extraction tools have therefore been proposed to accelerate ontology development. Consequently, the need for an automatic ontology extraction tool has increased in the last two decades and many tools have been developed for this purpose [5].

For ontology based systems, the page is linked to ontology page that defines information about specific department that offer postgraduate programs. For instance, considering a postgraduate student who wants to find a potential supervisor for his title in the area of Artificial Intelligence (AI), the computers can find that "Alicia" is a member of Computer and Agent Technology Center at College of IT, UNITEN, headed several AI research projects, has a particular e-mail address, and so on. All that information is readily processed by a computer since the data are expressed in a meaningful description which describes the concepts and relationships between data. These could be used to answer queries that currently would require human effort to filter and select the relevant answers.

This research will use some of the technologies crucial for semantic applications such as semantic processing, ontology, and SWRL for rule extraction and reasoning. The role of ontology is to provide a shared and common understanding of a domain that can be communicated between people and applications such that more accurate and effective services can be offered [6]. The logic-based description of ontology is carried out with Semantic Web ontology languages such as XMLS, RDF, RDFS, and OWL. This work uses OWL for representing and organizing domain knowledge and the query language SPARQL for supporting user query. SPARQL has the query protocol for accessing RDF designed by the W3C RDF Data Access Working Group. As a query language, SPARQL is "data-oriented" in that it only queries the information held in models; there is no inference in the query language itself. The proposed ontology-based system can assist supervisors and students in problems related to project title selection, expertise browsing, and general querying.

Z.C. Cob is with the College of IT, Universiti Tenaga Nasional (UNITEN), Malaysia (phone: +603-8921 2020; fax: +603-8921 2161; e-mail: zaihisma@uniten.edu.my).

Y.C.A Tang is with the College of IT, Universiti Tenaga Nasional (UNITEN), Malaysia (e-mail: aliciat@uniten.edu.my).

S. J S. Aziz is with the College of IT, Universiti Tenaga Nasional (UNITEN), Malaysia (e-mail: sjunainah@uniten.edu.my).

Open Science Index, Computer and Information Engineering Vol:6, No:12, 2012 publications.waset.org/10947/pdf

In this research, we intend to explore the semantic technology and develop a strong foundation in using ontology for developing or enhancing Web-based applications. The aim of this study is to strengthen the adoption of semantic technologies in university activities.

II. Related Works

Researchers in artificial intelligence first developed ontology in the early 1990s to facilitate knowledge sharing and reuse. Currently the topic becomes the talk of researchers in other fields such as Database, Knowledge Representation, Bioinformatics, Semantic Web, and so on. Even though ontology spurred interests in many researchers across the world, the meaning of this concept still generates a lot of controversy in discussions.

Ontology is the term referring to the shared understanding of some domains of interest, which is often conceived as a set of classes, relations, functions, axioms and instances [4]. They are meta-information or information about information. In the context of the Semantic Web, the relationships between the various terms within the information are encoded using an ontology language. Meanwhile, in knowledge representation, ontology is a description of the concepts and relationships in an application domain. However, no matter how they define ontology, it should be understandable not only by humans but also by the machines.

For an ontology-based application, data is envisioned to be annotated using ontology. Ontology conveys background information which enriches the description of the data and explicitly describes the context of the information. The same ontology can be used to annotate multiple data sources, including Web pages, collections of XML documents, relational databases, and etc. Since several data sources share the same specifications about the data, it enables a certain degree of inter-operation between these data sources.

Many successful stories related to how ontology is used to solve problems in different domains are discussed in the literature. For example, the most recent work in [7] implemented ontology-based search approach for accessing software architecture knowledge. An ontology-based semantic search mechanism is used in the framework to retrieve not only architecture properties, but also the rationale behind decisions taken during the construction of the software. The ontologies defined are also used in the visualization of the search results, providing an interactive environment for users to explore, discover, and analyze the information.

Tang Lijun and Chen Xu [8] presented ontology for college teaching management domain using Protégé. They proposed a semantic retrieval system model based on college teaching management ontology by applying Jena. With the present system, the users are allowed to search information with natural language and the information retrieval based on semantic level is achieved, which helps improve the retrieval efficiency.

A similar work by Memon and Khoja [9] discussed the development of ontology framework for UAE University program administration. The paper described how semantic web technologies based on ontology framework enables the integration of various sub-system functionalities within the university to support various program management functions. Protégé is used to simulate semantic web management activities such as ontology management, semantic annotation and semantic query of annotation triples.

In another domain, EgeoIT, a consulting company based in Spain, developed RDi-Advise by using semantic technologies to support R&D process decisions [10]. The ontology for RDi-Advise was built on the domain of Information Technology research and has been defined using OWL. The project highlighted several lessons learned during the project implementation and adoption.

Another project by [11] developed a tool implemented on the Android platform to help tourists in finding relevant travel information with minimum effort. The tool's main component, STAAR (Semantic Tourist information Access and Recommending), is able to support requests from heterogeneous environments. The ontology formalized in RDFS/OWL to describe travel-related information and to support integrating data from linked data sources that allow tourists to express their needs in semantic queries with different levels of complexity, and then get access to the relevant information.

III. Ontology Building Methodology

The main research activities of this work are shown in Fig. 1. This paper only presents the results of the 5th and 6th phases of the activities (refer to the highlighted boxes).

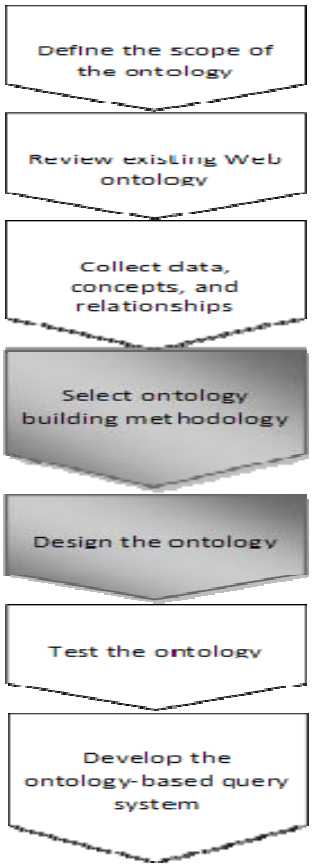


Fig. 1 Research activities

Open Science Index, Computer and Information Engineering Vol:6, No:12, 2012 publications.waset.org/10947/pdf

A framework adapted from the Uschold and King [12] is used to construct the ontology models for this work. Other prominent ontology development methodologies are Gruninger and Fox [13] and *Methontology* [14]. A comparative study for these methodologies can be found in [15, 16, 17, 18].

Uschold and King prescribed four main stages in the development of the semantic models, namely (1) Identify ontology scope, (2) Building the ontology, (3) Ontology evaluation, and (4) Ontology documentation. We extend the framework to include two additional activities, namely *designing informal competency questions* (helps in designing possible queries) in the ontology capture stage, and *ontology refinement and structuring* after checking the integrity of the ontology. The proposed framework for ontology modeling and construction is depicted in Fig. 2.

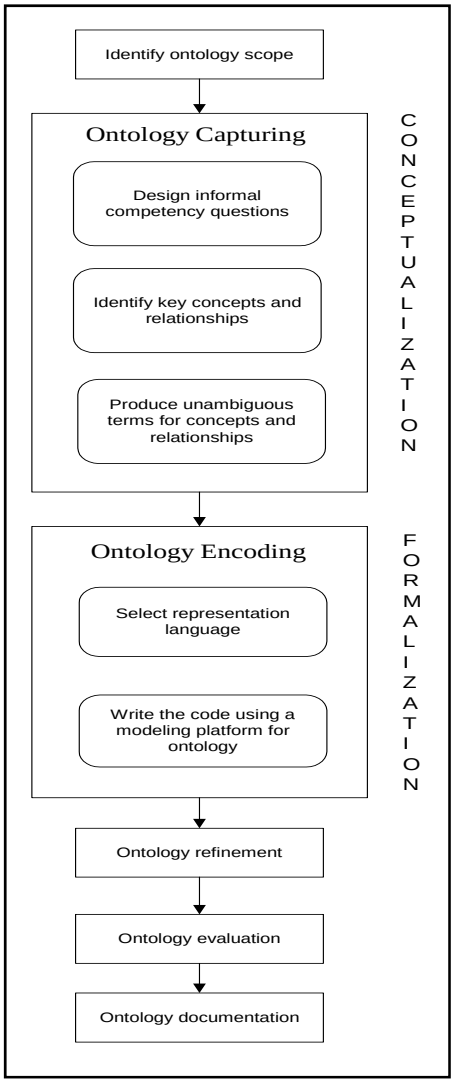


Fig. 2 The ontology development framework

IV. Ontology Component Design

In the simplest sense, ontology contains a body of concepts that is formally represented (machine-readable). Ontologies can support many computing activities and services that require mapping, matching, and easy composition of information that are required to achieve an enterprise's goals. The conceptualization phase began by defining the following key elements for the domain ontology:

- Concepts - a set of entities within a domain
- Relationships – interactions between concepts of concepts' properties
- Instances – concrete examples of concepts in the domain
- Axioms – explicit rules to constrain the use of concepts

In the ontology capturing stage, work commenced by conducting qualitative data collection based on interviews. Observation was employed to identify the main concepts of the problem domain. Informal competency questions were designed to collect further requirements. Since the ontology will be used in a query system, such questions can help in the design of the conceptual model more accurately.

TABLE I
Some Domain-Related Questions Supported by the Ontology

Q	Competency questions
1	What is the list of staff members at the graduate college?
2	What are the expertise areas of the staff members?
3	What is the maximum number of supervision allowed?
4	Are the lecturers from other colleges allowed in the supervision of COIT postgraduate projects?
5	What study structure or modes are available for MIT?
6	What is the duration of research for the PhD program?
7	What is the frequency for progress report submission?
8	What are the reporting techniques available?
9	Which research area has the most project students?
10	What are the projects currently being supervised or assigned?

Fig. 3 shows the main concepts and sample relations developed for the said domain, while a sample of the ontology is shown in Fig. 4 and Fig. 5. The concepts and relations are modeled using the Protégé ontology development tool [19]. Protégé is an open source software. It is used as the platform for modeling the domain knowledge of this work. The tool is selected due to its ease of use (graphical user interface) and its abstraction capabilities (enables ontology developers to concentrate on conceptual modeling).

Open Science Index, Computer and Information Engineering Vol:6, No:12, 2012 publications.waset.org/10947/pdf

Concept		Constraints
Superclass	**Subclass**	
Lecturer	Main Supervisor	**Must** be a staff member in the department that offers the postgraduate program
	Co-supervisor	Nil
Examiner	External Examiner	**Must** not be a UNITEN lecturer
	Internal Examiner	Nil
Postgraduate Student	Masters Student	**Must** possess a first degree in a relevant field
	PhD Student	**Must** possess a Masters degree in a relevant field
Postgraduate Program	Phd Program	Nil
	Masters Program	Nil
Assessments	Written Exam	Nil
	Viva Voce	Nil
	Thesis	Nil
Field	Nil	Nil

(a) Main "Concepts" used in the domain

Relation	Description
supervise	A main supervisor supervises a postgraduate student(s).
undertake	A postgraduate student undertakes a postgraduate program.
appoint	A postgraduate student appoints an internal examiner.
sit	A postgraduate student sits for a written exam(s).
submit	A postgraduate student submits a thesis.
teach	A lecturer teaches a postgraduate course(s).
profess	A lecturer professes in a field (s).
research	A postgraduate student researches in a field.

(b) Sample "Relations"

Fig. 3 Ontology components of "Concepts" and "Relations"

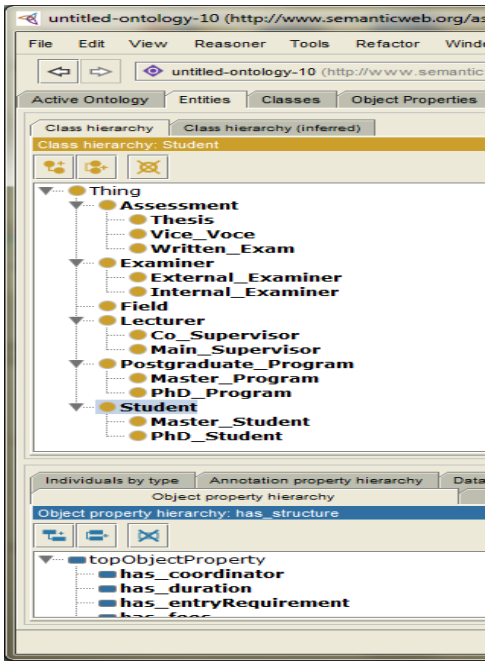


Fig. 4 The main "concepts" and the "properties" for postgraduate program definitions modeled using Protégé

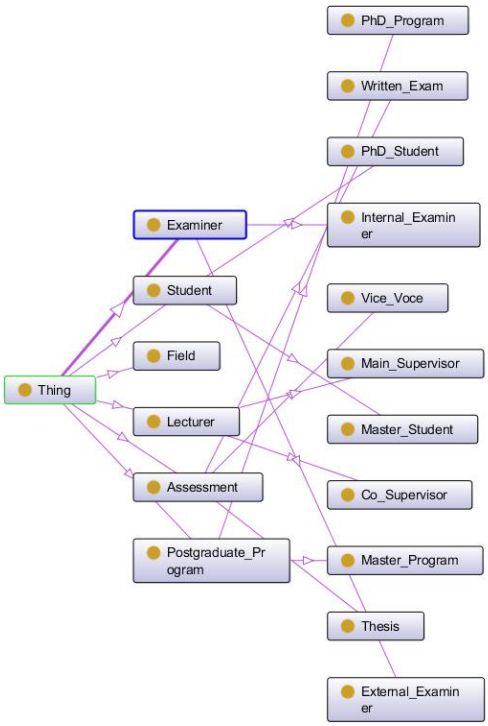


Fig. 5 The hierarchical view of the main and sub classes of the concepts

V. Conclusion

This paper presents a framework for modeling a specific ontology to support a query system for information searching by postgraduate students. The methodology used for building the ontology is discussed and the conceptual model with sample concepts and relationships are presented. The importance of designing domain related competency questions is also explained. It is hoped that this study will help promote the adoption of semantic technologies in academic administrative processes at Universiti Tenaga Nasional. This study would also be of interest to novice Semantic Web developers who might use it as a starting point for further investigations.

References

[1] Malaysian Ministry of Higher Education, MyBrain15. Putrajaya, 2008.
[2] S.K. Yeong, "Postgraduate Trends and Outlook 2011," *In Postgradasia.com*, Retrieved September 28, 2012, from http://www.postgradasia.com/adviceresources2/postgraduate-trends-and-outlook-2011.
[3] P. Lord, S. Bechhofer, M. D. Wilkinson, G. Schiltz, D. Gessler, D. Hull, C. Goble, and L. Stein, "Applying Semantic Web Services to Bioinformatics: Experiences Gained, Lessons Learnt," Springer-Verlag Berlin Heidelberg, 2004.
[4] N. Shadbolt, W. Hall, and T. Berners-Lee, "The Semantic Web Revisited", *IEEE*, 2006.
[5] J. Park and W. Cho, and S. Rho, "Evaluating ontology extraction tools using a comprehensive evaluation framework," *Elsevier*, vol. 69, No. 10, pp. 1043-1061, 2010.
[6] J. Davies, D. Fensel, and F.V. Harmelen (ed.), "Towards the Semantic Web: Ontology-driven Knowledge Management," *John Wiley & Sons*, 2003.
[7] A.M. Figueiredo, J.C. dos Reis, and M. A. Rodrigues, "Improving Access to Software Architecture Knowledge An Ontology-based Search

Approach," *International Journal Multimedia and Image Processing (IJMIP)*, vol. 2, Issues 1/2, March/June 2012.
[8] T. Lijun and C. Xu, "The Study of Semantic Retrieval Based on the Ontology of Teaching Management," *Procedia Engineering*, vol. 15 pp. 1555 – 1559, 2011.
[9] Q. Memon and S. Khoja, "Academic Program Administration via Semantic Web - A Case Study," *in Proceedings of World Academy of Science, Engineering, and Technology*, vol. 37, 2009.
[10] R. Colomo-Palaciosa, Á. García-Crespoa, P. Soto-Acostab, M. Ruano-Mayoralc, and D. Jiménez-López , "A case analysis of semantic technologies for R&D intermediation information management," *International Journal of Information Management*, vol. 30, Issue 5, pp. 465-469, October 2010.
[11] T.D. Cao and N.D. Tuan, "Improving Travel information access with semantic search application on mobile environment", *MoMM2011*, 5-7 December, 2011, Ho Chi Minh City, Vietnam.
[12] M. Uschold and M. King, "Towards a Methodology for Building Ontologies," *In Proceedings of IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing*, Montreal, Canada, pp. 1-13, 1995.
[13] M. Gruninger and M.S. Fox, "Methodology for the design and evaluation of ontologies", *in IJCAI workshop on basic ontological in knowledge sharing*, Montreal, Canada, 1995.
[14] M. Fernandez, A.G. Perez, and N. PaJuristo, "METHONTOLOGY: From ontological art towards ontological engineering," *AAAI technical report*, Madrid, Spain, 1997.
[15] C. Calero, F. Ruiz and M. Piattini, Ontologies for Software Engineering and Software Technology, *Calero.Ruiz.Piattini (Eds.)*, Springer-Verlag Berlin Heidelberg, 2006.
[16] M. Fernandez-Lopez, "Overview of Methodologies for Building Ontologies," *In Proceedings of the IJCAI-99 workshop on Ontologies and Problem-Solving Methods (KRR5)*, Stockholm, Sweden, 2 August, 1999.
[17] H. Beck and H. S. Pinto, "Overview of Approach, Methodologies, Standards, and Tools for Ontologies," *Agricultural Ontology Service (UNFAO)*, 2003.
[18] S. Mustaffa, R.Z. Ishak, and D. Lukose, "Ontology model for herbal medicine knowledge repository," *in proceedings of the 3rd Semantic technology and knowledge engineering conference*, pp. 20-22, pp. 85-94, July, 2011.
[19] Protégé website, http://protege.standard.edu

Zaihisma C. Cob is a Lecturer at the Information Systems Department, College of Information Technology (COIT), UNITEN. She obtained her Bachelor in Information Technology from Multimedia University, Cyberjaya and Master in Management (Information Technology) from Universiti Putra Malaysia. Her research interests include Knowledge Management and Semantic Web. Currently, Zaihisma works on applying data mining to business applications, ontology-based KMS and ontology development to Semantic Web applications. She is a member of IEEE, ACM Computer Society and the Knowledge Management Research Group at COIT.

Alicia Y.C. Tang is a Senior Lecturer at the Systems and Networking Department, College of Information Technology (COIT), UNITEN. She received her Master of Science from University of Manchester, and her Ph.D. in Artificial Intelligence from University of Malaya. Her research interests include qualitative reasoning, data mining, ontology, AI in education, and agent technology. Currently, Dr. Alicia works on applying data mining to business applications, agent technology to Learning Management Systems (LMS), and ontology development to Semantic Web applications. She is a member of the Asia Pacific Society of Computer in Education, IEEE, AIED Society, and the Centre for Agent Technology at COIT. Alicia has served as a reviewer for a number of International Conferences and Journals, and has been member of the program committees for several workshops and conferences.

Sharifah J. Syed Aziz was born in Kuala Lumpur in 1981. She earned a Bachelor in Information Technology (Hons.) from Universiti Tenaga Nasional, Malaysia in 2004, and undertook the M.Sc. in Information Systems at the University of Sheffield, United Kingdom in 2011. Her major field of study is information systems, particularly databases and development. She is currently teaching at Universiti Tenaga Nasional for the Department of Information Systems.

Open Science Index, Computer and Information Engineering Vol:6, No:12, 2012 publications.waset.org/10947/pdf