Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 116 (2014) 132 – 138

Procedia
Social and Behavioral Sciences

5th World Conference Educational Sciences -WCES 2013

Emerging of Academic Information Search System with Ontology-Based Approach

Norasykin Mohd Zaid [a 1], Sim Kim Lau [b]

[a]Faculty of Education, Universiti Teknologi Malaysia, Skudai 81310, Malaysia
[b]School of Information Systems and Technology, Faculty of Informatics, University of Wollongong, Wollongong 2500, Australia

Abstract

This paper describes the comparison of ontology development tools for development of academic information search system that assists inexperienced research students at a local university in Malaysia to search for academic resources in the local language context (Bahasa Malaysia). The cohort of inexperienced research students faces two main problems when using current system comprises of keyword search. Firstly the language barrier-limiting students' capabilities to conduct keyword search in foreign language (such as English). Secondly limited research experience in querying often results in obtaining irrelevant search results. The proposed semantic search system aims to apply ontology-based search to overcome the above two problems. The paper presents the first phase of system development; ontology design and ontology development tool.

Selection and/or peer-review under responsibility of Academic World Education and Research Center.
Keywords:

1. Introduction

Information and knowledge are increasingly becoming shareable and searchable resources, particularly in the current digitized world. Since 1996, the World Wide Web (WWW) has become a primary source for information offering online resources that are available 24/7. Traditionally library is an important source of information, particularly as academic resources and has become important source of reference for academic researchers. Library classification system has migrated from Dewey Decimal Classification System (DDC) to a new digitized format such as Online Public Access Catalog (OPAC) system that can be accessed through the web. The OPAC system is based on known-item search (Antelman et al., 2006).

Furthermore, keyword search and Boolean operators can also be used to facilitate the search process. Undoubtedly digital library provides an improved source of information access that include digital document creation and storage, documents classification and data indexing, documents searching and retrieving, distribution, administration and access control (Garza-Salazar et al., 2003). However human interpretation is still required when records matching the search criteria (such as keywords) are returned to determine its relevance and usefulness. For example, in searching for a programming textbook, which we do not know the exact title, we tend to type the word programming in the search box. When search results are returned, we scroll down the list of titles to look for the one that we search for. This is commonly encountered by students who are inexperienced in literature search. The motivation of this paper is to propose the development of an ontology-based information retrieval system to assist

[1]Corresponding Author: Norasykin Mohd Zaid. Tel.: +60-10-406-1060
E-mail address: norasykin@utm.my

Selection and/or peer-review under responsibility of Academic World Education and Research Center.
doi:10.1016/j.sbspro.2014.01.180

inexperienced research students at a local university in Malaysia to search for academic resources in the local language context (Bahasa Malaysia). The rest of the paper is organized as follows. Section 2 discusses ontology design and development of ontology. Section 3 discusses ontology development tools and follows by conclusion in Section 4.

2. Ontology

Ontologies have been known as a database of terms that justified a domain to be used and shared in a global area (Borst, 1997). Ontology becomes a model of real word to represent a domain of knowledge. This new technology has been used in the Semantic Web although the original word of ontology is being borrowed from the philosophy discipline, which defines the concepts of things. Thomas (1993) explains the real definition of ontology is a systematic account of existence, however in computer science, ontology is a representation of precise specification to form a concept. Thus, ontology is described as formal specification of terms in the define domain and identifying any relations existing in between the terms. Ontology enables people or machines to retrieve the desired information with an understanding of the meaning of terms used in the domain and share common vocabularies used in the same domain (Wang et al., 2008a). Therefore, the use of ontology is about using, reusing and sharing domain knowledge of terms concept. Many ontology classes have been developed recently and are kept in a database to be used or referred to by others as knowledge/resource sources. Ontologies are not only used in the field of Semantic Web but also in many others fields such as artificial intelligence, software engineering, biomedical informatics, library science, and information architecture.

There are two types of ontologies according to two dimensions of perception: the amount and type of structure of the conceptualisation and the subject of the conceptualisation. The first dimension, according to Heijst et al. (1995), includes: (i) terminological ontologies, (ii) information ontologies, and (iii) knowledge modelling ontologies; whereas the second dimension includes: (i) domain ontologies, (ii) generic ontologies, (iii) representation ontologies, and (iv) application ontologies. The first dimension with terminological ontologies is referred to as ontology that defines the terms to represent knowledge in the domain of discourse, such as medical or biological domains. Information ontologies are defined as records structure of a database, which is a flat structure, unlike the knowledge modeling ontologies, which have a richer structure of database, such as involving distinction and decision-making processes. To refer to the second dimension of ontologies, domain ontologies refer to specific particular area while generic ontologies refer to domain ontologies across many areas. Representation ontologies are supposed to be naturally present in general contrast to application ontologies, which are specifically designed to the particular application such as the Marine Metadata Interoperability Project (MMI) (https://marinemetadata.org/).

2.1. Ontology design

Holsapple and Joshi (2002) present five approaches to ontological design: (1) inspiration, (2) induction, (3) deduction, (4) synthesis, and (5) collaboration. Inspirational approach starts the design idea by collecting individual personal views and creativity to construct the domain context. Inductive approach is based on the observation and analysing of current or specific domains to apply to particular domains. Deductive approach adopts some general principles to construct a new domain while the synthetic approach applies some potential characterisation from the existing ontologies. With the collaborative approach, the approach relies on human participation, which involves individual reflection and viewpoints to get along with the collaborative process.

2.2. Ontology development

How these ontologies can be developed depends on how or what method is being used. Uschold and Gruninger (1996) conclude that there are five steps in the process of ontologies development: (i) identify purpose and scope, (ii) building the ontology, (iii) evaluation, (iv) documentation, and (v) guidelines for each phase. In the second step of building ontology, it includes: (a) ontology capture, (b) ontology coding, and (c) integrating existing ontologies (Uschold and Gruninger, 1996). The first step in building the ontology is by considering when there is a clear idea