ISSN: 2456-9348
Vol (03) _Issue (07) Impact Factor: 4.520

iJETRM
International Journal of Engineering Technology Research & Management

ONTOLOGY-BASED RETRIEVAL SYSTEM

Thi Thi Soe[*1]
Zarni Sann[2]
[*1]Faculty of Computer Science
[2]Faculty of Computer Systems and Technologies
[*1, 2]University of Computer Studies (Mandalay), Myanmar
[*1]hlaing.nn@gmail.com
[2]zarnisann@gmail.com

ABSTRACT
Today, Information Technology (IT) is rapidly improving. Large-scale available data and information are demanded by information society. To organize large information repositories and access these repositories efficiently, metadata can be used. Metadata is closely related to Ontology. Ontology is a formal explicit description of concepts in a domain of discourse, properties of each concept describing various features and attributes of the concept, and restrictions on slots. This work intends to realize how to create an ontology for image annotation and retrieval system. In this paper, we simulate a search engine based on the ontology concept for retrieving the desired image and related information. As a case study, the system uses the student information from University of Computer Studies (Mandalay). The retrieval system is implemented by using Java, and Java API for developing Resource Description Framework (RDF) files. These RDF files can give the image and related information for the user's requirement.

KEYWORDS:
Ontology, Information and Image retrieval, University student

INTRODUCTION

Information is crucial in the process of planning and making decisions. Many researchers concentrating on retrieval system to automate the process of selecting, filtering and searching the desired information. In this way, much of the information will be reduced and the result will consist of higher information density more compact and valuable information. Information Retrieval (IR) deals with the representation, storage, organization and access to information items. The representation and organization of the information items should provide the user with easy access to the information in which user is interested. The purpose of the IR system is to capture the desired items and to filter out unwanted items. To build the large information repositories and access these repositories efficiently, metadata can be used. Metadata is closely related to Ontology. The idea of ontologies is that they are conceptually motivated, i.e., can be used to express the intended meaning of things, and not just words as textual strings. In [1] described how ontologies can be used to create better image annotation and retrieval systems. In a nutshell, ontologies are used to overcome the problems that evolve from traditional text-based information retrieval when it is applied to images. [2] discussed an ontology-based image retrieval approach that aims to standardize image description and the understanding of semantic content. Ontology-based image retrieval has the potential to fully describe the semantic content of an image, allowing the similarity between images and retrieval query to be computed accurately. This work is to design and develop an architecture that can retrieve desired image and related information from the student ontology. The system uses the information from first year to fourth year students in University of Computer Studies (Mandalay). The organization of remaining sections is as follows: Section II describes the basic aspect of ontology. In section III, we devote the design process to build the ontology-based retrieval system. Preliminary testing is carried out on the system, which is explored in section IV. Finally, we conclude the paper in section V.

iJETRM
International Journal of Engineering Technology Research & Management

METHODOLOGY

Ontology is a specification of an abstract, simplified view of the world [3]. Ontology defines a set of representational terms called concepts. Ontology can be constructed in two ways: domain-dependent or generic. Generic ontologies are definitions of concepts in general; such as WordNet, which defines the meaning and interrelationships of English words. A domain-dependent ontology generally provides concepts in a specific domain, which focuses on the knowledge in the limited area, while generic ontologies provide concepts more comprehensively.

Ontology Markup Language

The first ontology markup language was SHOE. SHOE is language that combines frames and rules. It was built as an extension of HTML. It used tags different from those of the HTML specification, thus allowing the insertion of ontology in HTML documents. Later its syntax was adapted to XML. The rest of ontology markup language presented here is based on XML. RDF is a semantic-network based language to describe Web Resources. The RDF Schema language was also built by the W3C as an extension to RDF with frame-based primitives [4].

Resource Description Framework

The Resource Description Framework (RDF) is a language for representing meta-data. The RDF data model defines the structure of the RDF language. The data model consists of three data types:

Resources: All data objects described by a RDF statement are called resources.

Properties: A specific aspect, characteristic or relation of a resource is described by a property.

Statements: A statement combines a resource with its describing property and the value of the property. RDF statements are the structural building blocks of the language.

A RDF statement is typically expressed as "resource-property-value" – triple, commonly written as P (R, V): A resource R has a property P with value V. These triple can also be seen as object-attribute-value triple. Statements can also be expressed as graphs with nodes for resources and values where directed edges represent the properties. Figure 2 shows the graph of the resource R with an edge for the property P directed to the property value V [5].

Jena

Jena is a programming toolkit, using the Java programming language. It is written for the programmer who is unfamiliar with RDF and who learns best prototyping, or, for other reasons, to move quickly to implementation. The advantage to ontology is that it is an explicit, first-class description. Jena offers a simple abstraction of the RDF graph as its central internal interface. This is used uniformly for graph implementations, including in-memory, database-backed, and inferred graphs. The main contribution of Jena is a rich API for manipulation RDF graphs.

The main building blocks of Jena are its RDF API with RDF/XML parser and persistence subsystem, its ontology subsystem, its reasoning subsystem, and the RDQL query language. The RDF API provides methods for manipulating the set of RDF triples. It also provides a way for extending the behavior of resources. An important part of RDF API is ARP, the Jena's RDF/XML Parser. Another building block of Jena is its persistence subsystem that provides a way to store models using database. The persistence subsystem supports also a capability for RDQL queries that dynamically generates SQL queries in order to perform as much of the RDQL query within an SQL database engine as possible [6].

SYSTEM ARCHITECTURE

The system architecture is shown in fig. 1. In this system, there are two main modules. The first module is the creation of the triple forms to build students information ontology. The second module is the querying from the ontology that is created in first module. These modules are explained as follows.

Jena Statement:

Create Resource

```
Model modelName = ModelFactory.createDefaultModel();
Resource resourceName=modelName.createProperty (location);
```

Adding Property

```
Property propertyName=modelName.createProperty (property);
```

Create Statement

```
Statement statementName = model.createStatement (subject,property.object);
modelName.add(statementName);
```

iJETRM
International Journal of Engineering Technology Research & Management

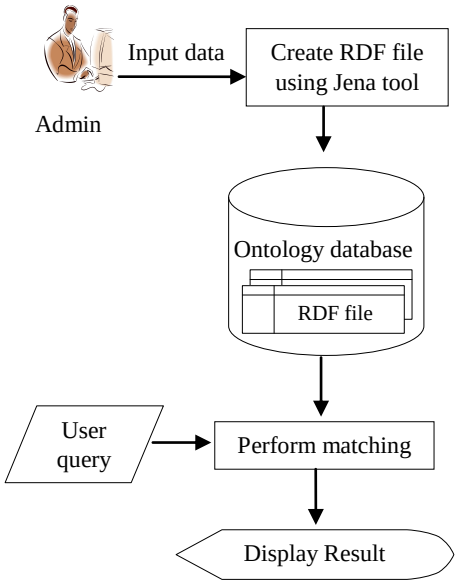

Fig. 1: System Architecture

RDF Creation Module

Jena package is used to create RDF tag. Firstly, triple form of the images is created with their porperties and the values of the properties. Figure 2 is an example triple form of an image. In this figure, "Image1" is the resource or subject, "name", "rollno", "year", "acyear", and "major" are properties or predicates. And "Mg Mg", "23", "Second", "2003-2004", and "CS" are the values of the properties. Each triple form is stored with the RDF format to build student ontology. RDF file for the student "Image1" is mentioned below. To develop RDF file, Jena Package, Java API for RDF will be used.

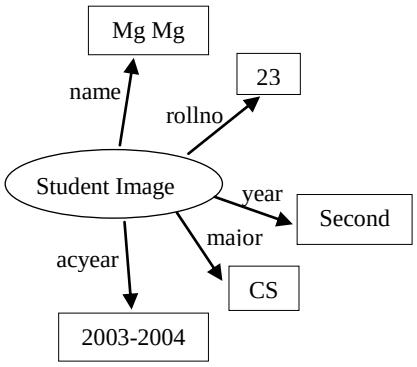

Fig. 2: A RDF Description for an Image

```
<rdf:RDF
xmlns:rdf= "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:Image= "http://www.ucsm.edu.mm/2009/student-rdf/1.0#" >
<rdf:Description rdf:about= "http://www.ucsm.edu.mm/student-affair/students/Image1">
   <Image:acyear>2003-2004</Image:acyear>
   <Image:rollno>23</Image:rollno>
   <Image:name>Mg Mg</Image:name>
   <Image:year>Second</Image:year>
   <Image:major>CS</Image:major>
  </rdf:Description>
```

iJETRM
International Journal of Engineering Technology Research & Management

</rdf:RDF>

RDQL Creation Module

RDQL sentences are written according to the RDQL syntax as describing in section 2. Each RDQL statement is used for one searching button. Example RDQL statement that queries the student of the image, name is "Mg Mg".

```
SELECT ?x, ?name, ?image
WHERE
(?x Image:image ?image), (?x Image:name ?name), (?x Image:name 'Mg Mg')
USING Image FOR <http://www.ucsm.edu.mm/2004/student-rdf/1.0#>
```

PRELIMINARY INVESTIGATION

The retrieval system is implemented in the form of Web Sites. In this system, the user has the ability to search the student information with their name, roll number, year, major and academic year. If the user types desired word in the respective text box, the resultant image within the page appears as in fig. 3 (a). If the user taps the image from the result of searching student page, the detail information of the student appears as shown in fig. 3 (b).

Fig. 3: (a) Result of Student Image, and (b) Related Information

CONCLUSION

Ontology-based information retrieval system is useful for easily finding user information need more effective and relevant, and giving better response to user. The main objectives of this work are to realize how ontologies can be used to create better image annotation and retrieval systems, to study the ontology concept and the development of Resource Description Framework (RDF), and to query and retrieve student information from Ontological database. In our system, user would like to search the student information in the University of Computer Studies (Mandalay) and obtain their images. So Ontology is built by using Resource Description Framework (RDF). We build Ontology but this is not visible by user, then searching process is performed to demonstrate our developed ontology. The system can further extend by including the information of staff and teachers in the University.

REFERENCES

[1] Avril Styrman, *"Ontology-Based Image Annotation and Retrieval"* Department of Computer Science, University of Helsinki, 2008.
[2] Wei Zheng, Yi Ouyang, James Ford, Fillia S. Makedon, *"Ontology-based Image Retrieval"* Department of Computer Science, Dartmouth College, USA.
[3] D. Brickley and R. V. Guha. Resource description framework (RDF) schema specification 1.0 *In W3C Candidate Recommendation* 2000-03-27
[4] Asuncion Gomez-Perez, Mariano Fernandez-Lopez, Oscar Corcho, *"Ontological Engineering"*.
[5] Valerie Bonstrom, Annika Hinze, Heinz Schweppe, *"Storing RDF as a Graph"* Institute of Computer Science, Freie University Berlin, Germany
[6] Kevin Wilkinson, Craig Sayers, Harumi Kuno, Dave Reynolds, *"Efficient RDF Storage and Retrieval in Jena2"*