Ontology-Based Information Retrieval for Animal Kingdom

Hnin Ei Ei Cho
Computer University (Monywa) Myanmar
hninei009@gmail.com

Abstract

Retrieving the necessary information about the domain provide the user to acquire the explicit knowledge for that domain. This information about terms and their relations is traditionally provided through dictionaries. But these dictionary entries evolved for the convenience of human readers, not for machine. So, this paper attends to become the central resources for applications in animal kingdom. Various research communities commonly assume that ontologies are the appropriate modeling structure for representing knowledge. This implementation provides a more effective retrieval knowledge for animal kingdom and their conceptual information by implementing ontology. It is based on Web Ontology Language (OWL) which is an ontology markup language and also has its corresponding knowledge representation methods.

1. Introduction

Today's world is a world of information. Many tasks in the area of information technology require the representation of world knowledge. The history of artificial intelligence shows that knowledge is critical for intelligent systems. In many cases, better knowledge can be more important for solving a task than better algorithms. Information and knowledge are essential for human. The method how to easily retrieve information and knowledge are needed. The capability of representation with knowledge base on text is important. To have truly intelligent systems, knowledge needs to be captured, processed, reused, and communicated. Ontologies support all these tasks. The ontology model serves as a model for human information processing. The most prominent domain for the convenience of biology is a huge ontological structure. Ontological hierarchies might also serve as structuring tools.

This thesis represents the knowledge by using ontology concept. The main focus of this system is study how to crate domain ontology. Whereas, ontology can offer a way to address meaning of terms (concepts and relations) required for the domain. This system accepts the query as input and the user can retrieve the relevant information. So, our retrieval system with the structure of ontology will produce not only the resources but also the conceptual relations for animal.

Ontology allows users to organize information into taxonomies of concepts, each with their attributes, and the describe relationship between concepts. When data is marked up using ontologies, software agents can better understand the semantics and therefore more intelligently locate and integrate data for a wide variety of tasks.

2. Related Work

Three principal reasons to use Information Retrieval in ontology are to represent knowledge, to store constraints and to find information by making reasoning with the concept of the ontology to better facilitate for Information Retrieval, finding the name that is nearly same with the existing information when the user enters the searched name. So, the implementation of Animal Ontology becomes an essential knowledge for ontological researches. Moreover, using this ontology can be used in zoo for observing the animals although the existing tasks are time consuming to define its concepts and their relations.

3. Theory background

3.1 Ontology

An ontology is an explicit specification of the conceptualization. The term is borrowed from philosophy, where an Ontology is a systematic account of Existence. For AI systems, what "exists" is that which can be represented. When the knowledge of a domain is represented in a declarative formalism, the set of objects that can be represented is called the universe of discourse. This set of objects, and the describable relationships among them, are reflected in the representational vocabulary with which a knowledge-based program represents knowledge [1].

In Computer Science, Ontology is a data model that represents a set of concepts within a domain and the relationships between these concepts. It is

used to reason about the objects within that domain. Ontologies are used in Artificial Intelligence, Semantic Web, Software Engineering, Biomedical Informatics and information architecture about the world or some part of it [5].

3.2. Knowledge base

A knowledge base is a special kind of database for knowledge management, providing the means for the computerized collection, organization, and retrieval of knowledge [6].

Machine-readable knowledge bases store knowledge in a computer-readable form. They contain a set of data, often in the form of rules that describe the knowledge in a logically consistent manner. An ontology can define the structure of stored data - what types of entities are recorded and what their relationships are. Some machine-readable knowledge bases are used with artificial intelligence. Such knowledge bases are also used by the semantic web.

Human-readable knowledge bases are designed to allow people to retrieve and use the knowledge they contain. They are commonly used to complement a help desk or for sharing information among employees within an organization. They might store troubleshooting information, articles, white papers, user manuals, or answers to frequently asked questions. Typically, a search engine is used to locate information in the system, or users may browse through a classification scheme.

3.3 Developing an ontology

Developing an ontology includes the following**:**

- Defining classes in the ontology,
- Arranging the classes in a taxonomic (subclass-super class) hierarchy,
- Defining properties and describing allowed values for these properties,
- Filling in the values for instances,
- Creating a knowledge base by defining individual of these classes, filling in specific property value information and additional property restrictions.

3.4 Web Ontology Language

OWL is a language for making ontological statements and may include the description of classes, along with their related properties and instances. OWL can be used to explicitly represent the meaning of terms in the vocabularies and the relationships between these terms. This representation of terms and their interrelationships is called ontology. It facilitates greater machine interpretability of web content than that supported by XML, RDF and RDFS by providing additional vocabulary [3].

3.4.1. Requirements for OWL. There are eight requirements for OWL [4] :

- Ontologies are objects on the web.
- They have their own meta-data, versioning, etc.
- Ontologies are extendable.
- Ontologies contain class, property, data type , range and domain, individuals.
- Equality (for class , individuals).
- Classes as instances.
- Cardinality constraints.
- XML syntax.

3.4.2. Design goals for OWL. In OWL, there are several facts to design. They are following as:

- Sharable
- Changing overtime
- Interoperability between ontologies
- Inconsistency detection
- Balancing expressivity and complexity
- Cast of use
- Compatible with existing standards
- Internationalization

4. Analysis of animal kingdom

There are two kingdoms in our living world: Animal Kingdom and Plant Kingdom. Retrieving the information for animal kingdom based on ontology can be analyzed as follow:

The analysis of animal kingdom, its related relations and how they relate each other are performed in this phase. In ontology-based information retrieving, animal names can be same class and relation. So, we can create a link between two entries by applying these relations. Here some examples of relations between animals.

4.1 Anatomy

Animals can be grouped by their anatomy. They are: Vertebrates and Invertebrates. Every animals of the world belong to one of them. Vertebrates are animals with backbones and invertebrates are without backbones.

4.2 Eating-Habit

Animals can be further subdivided according to their eating-habit: cannibal, carnivore, omnivore, detritivore, scavenger and herbivore.

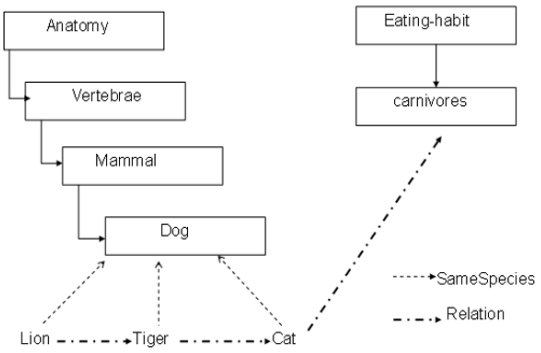

Figure .2 Representation of an Animal

5. Construction of the ontology-based information retrieval

Information Retrieval is a concerned with the process involved in the representation, storage, searching and finding of information which is relevant to a requirement for information desire a human user. There are two methods of finding interesting information on the web which are querying and browsing. Querying is keyword based search Browsing or navigation through a hierarchy likes the set of links toward topic. Browsing is suitable when user cannot express his goal explicitly by a Boolean combination of the keywords [2].

OWL is used to build the storage of names. Moreover, the information about the animals are in system database as a list. Implementation of the Animal Ontology

The goal of this is to apply the developed ontology- based information retrieval system as a central resource for building user-readable electronic dictionary. It constraints the two main phases: (1) LOOK Up and (2) DATA ENTRY.

"LOOK UP" portion is any user enable to access. This portion is very useful as an electronic dictionary to search for the information of animal. The objective of this phase is to exploit the knowledge resources from the ontology. Retrieving appropriate Information and relevant relations for the given animal name can be done.

The other portion, Data Entry can only be used by administrator who can enter data and save knowledge by entering the correct password. All the necessary information such as classification of animal, concept hierarchy, related animals and its example usage. The next portion, Create SameSpecies relation can be done when two animals have same anatomy and eating-habit. Ontology-based information retrieval which is used to map Microsoft office visual studio c# 2008 and OWL for developing the interfaces. Detail architecture of the system implementation is described in the following algorithms.

ALGORITHM: DATA ENTRY
(1) - Enter the required information for Na
(2) - if the name has same species field
-Retrieve the name of Ns
-Saving the data and properties of Na
where, "samespecies" relation of Na is Ns
else
- Saving the data and properties of Na in the ontology as a new field.

ALGORITHM: LOOKUP
(1) - Input the animal name, Na, you would like to search.
(2)- Find Na in the ontology.
(3)- if (found)
-Retrieve specific name, according to its Phylum, Class, Family, Order, Binomial Name, Genus, Species, Anatomy, Group and Eating Habit
-Display detail information of the animal concerned with the structure of the body.
- if (Na has the "SameSpecies" relation)
- Retrieve and show Ns of Na
else return "Not Found" information to user.

ALGORITHM: Create SameSpecies Relation
(1) - Enter Na, G and E as input.
(2) - Find name that have same A and E properties in the ontology.
if (found)
- Saving the "SameSpecies" relation.
else Display "Not Found" message.

Where Na = Name of animal you want to search
E = Name of Eating-Habit
Ns = Name of animal (SameSpecies)
G = Group of animal in anatomy

6. System design

The system searches the name of animal that the user entered. If the name is involved in animal ontology, the system extracts the relevant information of that animal. If the information is not found, the system will display "Not Found" information. In Admin site, the administrator must login to update the information. When the password is true, information in animal ontology can be updated. Finally, the updated information is displayed. If the information is found, the system will display the searching results.

At the start of the system, the administrator must register the information and then add to the database. After that, the user can search information from such resources.

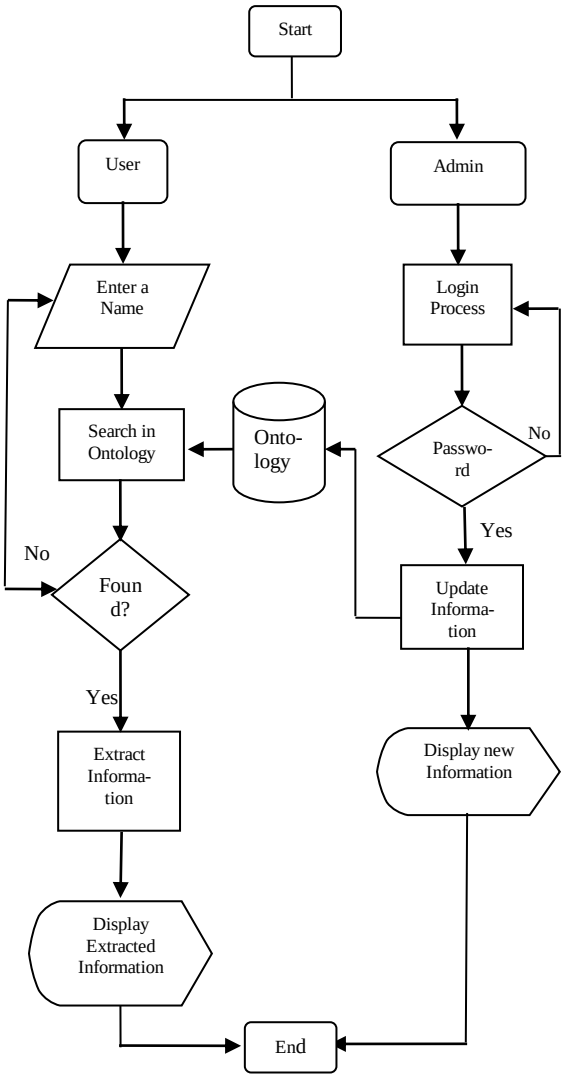

Figure 2. System flow diagram of the system

5.1 Searching result

The system will retrieve the Phylum, Class, Genus, Order, Family, Species and Binomial Name Of the searched animal. Moreover, the picture of that animal will be showed. Then, the detailed information of the searched animal will be described. The system can further retrieve the anatomical structure of the searched animal. Then the system displays the animals name that are same in "Class" of the typed animal. The system can also describe the animal names that are same in both "Eating-Habit" and "Anatomy".

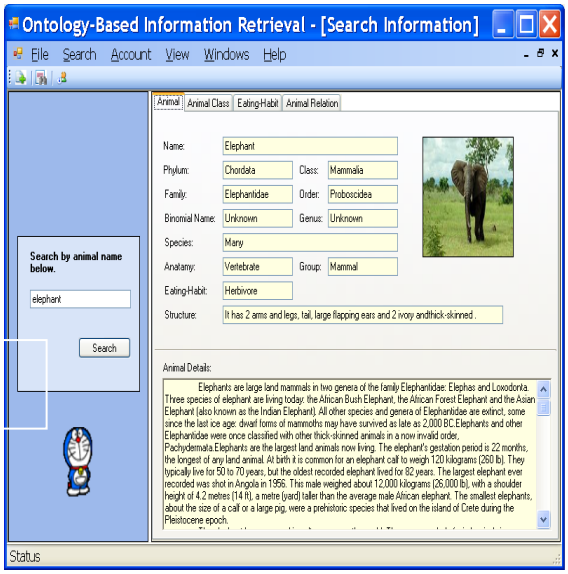

Fig 3. Information search result

6.2 Updating information

To update the desired information, the administrator must enter the true password. If the password is false, the information cannot be update. In this session, the information can be added, deleted and prepared to the desired result. The administrator can update the existing information by using move next item, move first item, move last item and move previous item to reach the desired animal name. To add the new animal in existing ontology, use add button. If the administrator wants to remove the animal, delete button can be used.

Finally, the updated information can be saved by using save button as shown in Figure 4.

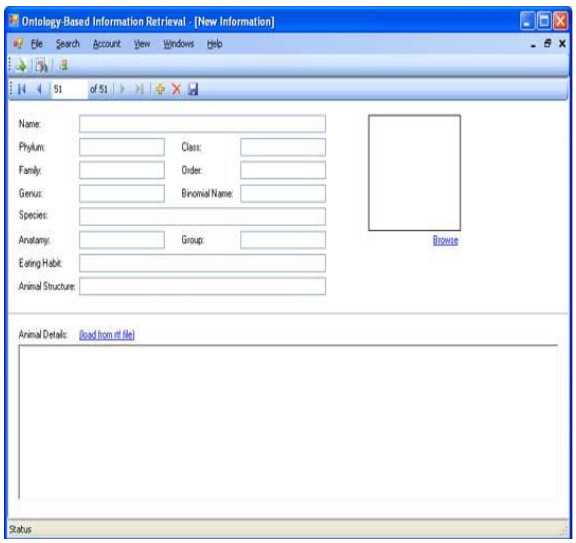

Figure 4. Update the new information by administrator

7. Conclusion and further works

Development of Information Retrieval for animal kingdom will facilitate many tasks such as syntax and semantic, Information Extraction and expressive power. OWL information can easily be changed between different types of computers using different types of application languages [4].In this paper, we describe our efforts in building ontology. Since this implementation provides not only information but also relations, it may become the central resources of animal research although developing animal dictionaries are very tedious and time consuming task. Moreover, we will also provide the user to exploit these resources. So, it can serve as User- readable electronic retrieval, as well as machine-readable.

Since the concepts in the ontology can optionally be classified and organized, the system can possibly be extensible to hold additional concepts or properties or both. And the system can be modified to specific domain, particularly for technical domain, medical domain, agricultural domain, etc. New data entries can also be entered appropriately.

Because of an Animal Electronic Dictionary among World Wide Web, extracting the data from Animal Ontology is often insufficient to support retrieving. So, this can be further extendable by using Web. In further work, the next generation of the web will be more organized, informative, searchable, accessible and useful, World Wide Web has a much wider potential: one can imagine an enormous number of different domain-specific and general applications. The work and research of Information Retrieval for animal kingdom can significantly help in search and retrieval, data selection and therefore will inevitably impact on Expert System as well as e-dictionary.

6. REFERENCES

[1] A. Cnomex-Perz, M.F.Lopez, O.Corcho: "**Ontological Enbgineering**" Facultad de Informatica, Universidad Politecnica de Madrid, Campus de Montegancedo sn., Bodilla del Monte, 28660 Madrid, Spain.

[2] K. M ,Myo,**"Information Providing for Requirements Management in COTS Software"** University of Computer Studies, Yangon, Myanmar, Proceeding of the third International Conference on Computer Applications 2005,pp 321 to 331.

[3] K.S Candan, H.liu, & R. Survana **"Resource Description Framework: Metadata and its Applications"** Department of Computer science & Engineering Arizona State University Tempe, AZ 85287-5405 {candan, hliu, reshma.survarna} &asu.edu

[4] N. Memon, O. I. Eldai and M.A. Uquali, **"Semantic Web Languages: A Comparison"**, Department of Software nad Media Technology, Aalborg UniversityNiels Bohrs Vej8,6700,Esbjerg,Denmark,Institute of Information Technology, Mehran University if Engineering and Technology Jamshoro Sindh, Pakiston, Proceeding if the Third International Conference on Computer Applications 2005,pp 385 to 395.

[5] http://en.wikipedia.org/wiki/knowledge base/

[6] http://en.wikipedia.org/wiki/ontology/