INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Development of Ontology for Sports Domain

Sudha Ramkumar[1], Dr. B. Poorna[2]

[1]*Research Scholar, Bharathiar University, Coimbatore, India*
[2]*Principal, Shri Shankarlal Sundarbai shasun college, Chennai, India*

Abstract: Ontology has been used as a knowledge representation method and it has been widely used to improve the information retrieval. The ontology is used to efficiently capture the semantics of the information which significantly improves the performance of the information retrieval process. The importance of ontology has been proved in the areas of semantic data mining, knowledge management and information retrieval. Protégé is the most commonly used tool for developing ontology to describe the concepts and the relationship between those concepts to create class hierarchy. In this paper, a methodology for the development of ontology for sports domain is proposed and it consists of 6 phases. An overview of these development phases has been described. Finally, using Jena the sports ontology Owl model is converted into Jena ontmodel which is further used for document clustering process.
Keywords: Ontology, Information retrieval, protégé, knowledge, Jena

I.   INTRODUCTION

According to World Wide Web consortium domain ontology is a formal description of data and descriptive knowledge of a domain of interest. Ontology is defined as a formal description of concepts in a particular domain of interest, properties of each concept describes various features and attributes of the concepts are otherwise called as slots or roles or properties and restriction on those slots are called as facets [2]. Ontology along with a set of individual instances of classes together forms a knowledge base. A class can have subclasses that represent concepts that are more specific than the super class. For example, the sports class is divided into five subclasses such as cricket, athletics, football, rugby and tennis. In practical terms, developing ontology includes,

A.   Defining classes in the ontology
B.   Arranging the classes in a hierarchy
C.   Defining slots and describing allowed values for these slots
D.   Filling in the values for slots for instances.

Definition of each class, subclass concept and list of references to each class and subclass are attached to the node of class hierarchy. The onto Graf tab is used to visualize the ontology concepts like graph. The DL query tab of protégé is a powerful and easy to use feature for searching a classified ontology. This query language is based on the OWL syntax that is fundamentally based on collecting all the information about a class, property or instances into a single frame.

The Main Components of Ontology are Class, Properties, Relations, Individuals, Axioms and Rules. In this paper, Sports Domain Ontology has been developed using protégé 4.2 with the related classes as concepts and sub concepts. Totally there are 840 classes and subclasses have been defined in the class hierarchy and all these concepts are related to the sports areas such as athletics, cricket, football, rugby and tennis with data properties, object properties and annotation properties. The object property defines the relationship between the concepts and the data properties used to define the relationship between the concept individual and the data literal. The annotation properties such as comment, see also, is defined by etc are used to annotate the concepts of the domain ontology of interest. The rest of the paper is organized as follows: section 2 describes some of the previous work related to this paper, section 3 describes the methodology used for the development of sports ontology and section 4 describes the use of ontology for further step and section 5 concludes the paper.

II.   BACKGROUND

The semantic web aims to extend the current web standards and technology so that the semantics of the web is machine understandable [12]. The use of domain ontology in the semantic information retrieval system enables users to represent a meaningful knowledge of a particular domain of interest which makes the machine more intelligent to understand the semantics of the information. A methodology proposed [1] is for developing ontology according to software engineering principles, to achieve high quality ontology with minimum cost. The methodology [3] defined five issues that much be addressed to use a proper methodology for developing domain ontology. The proposed methodology of [4] and [5] demonstrated the development of a university domain ontology using protégé editor in their work. The various aspects like Super class-Subclass, creating a class

hierarchy, instances for classes, query retrieval and visualization of ontology using graph tools have been demonstrated. The approach [7] presented a new method for semantic information retrieval for cricket domain by developing cricket domain ontology and it is stored in RDBMS. Querying, inference and searching techniques have been developed.

The research work [8] developed a semantic search engine for the university domain. It uses ontology as a knowledge base for the information retrieval. This paper applied query expansion and ranking technique to improve the accuracy of the information retrieval process. A new methodology [9] is presented which consists of 4 phases such as requirement analysis, development, implementation, evaluation and maintenance. In this paper, they compared existing methodologies and based on the comparison they proposed a new methodology with clear explanation of each and every activity involved within each phase. This helps the users in developing high quality ontology. An approach developed [10] human resource management ontology which provides a knowledge base for the applied field. This human resource management ontology can be evaluated against many criteria: coverage area addressed in ontology, complexity and granularity of that coverage area, the consistency and completeness of the ontology and the representation language in which the ontology was modelled. Ontology development methodology [11] presented for traditional Chinese medicine(TCM) system. This work aims to build a unified ontology framework to effectively organize and integrate the terminological knowledge of TCM. The main effort of this work is to provide ontology approach to standardize TCM medical terminology.

III. METHODOLOGY

Ontology is defined as a formal description of concepts in a particular domain of interest. Classes are the main focus of the ontology. Classes are otherwise called as concepts. Properties of each concept describing various features and attributes of the concepts are called as slots or roles and restrictions on slots are called as facets. Ontology along with a set of individual instances of classes together forms a knowledge base. A class can have subclasses that represent concepts that are more specific than the super class.

The onto Graf tab is used to visualize the ontology concepts like graph. The DL query tab of protégé is a powerful and easy to use feature for searching a classified ontology. This query language is based on the OWL syntax that is fundamentally based on collecting all the information about a class, property or instances into a single frame. There is no correct method or methodology for developing ontology. This work uses the Natalya F. Noy methodology for developing ontology and according to him "Ontology development is an iterative process" [6]. This methodology consists of the following steps,

A. Determine the domain and scope of the ontology.
B. Enumerate important terms in the ontology
C. Define the classes and subclasses in the class hierarchy.
D. Define the properties of classes.
E. Define the restrictions on the slots or properties.
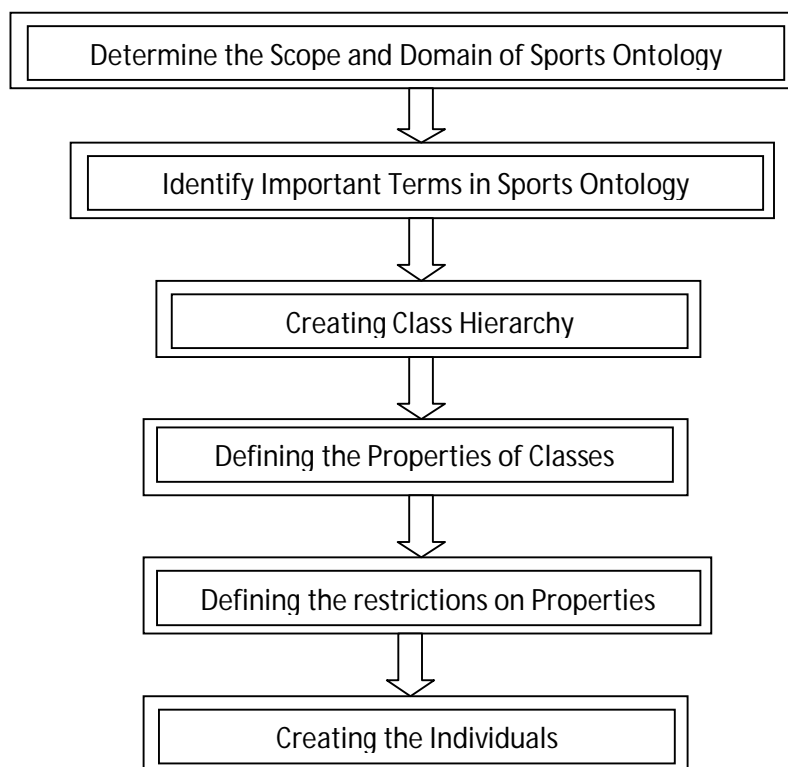F. Create individuals.

Fig 1: Methodology for developing Sports Ontology

In the first step, the domain and scope of the sports ontology has been identified with the support of domain experts. The details of the five sports has been collected from the domain experts and the important terms of the five games has been identified as classes and subclasses based on the relationship between those terms. The properties of the classes and subclasses have been defined carefully. The restrictions on the properties are defined and finally the individuals for the classes are created.

A. Determine the scope of ontology

Many researchers suggested that ontology development starts with identifying the scope of ontology. By defining the scope will make describing classes, subclasses and relationship between them easier. Sports ontology developed in this work is used to assist the clustering process. Hence, the concepts related to the sports domain are very important and the synonyms of those concepts need to be included for the clustering process.

B. Enumerating important terms in ontology

The terms plays major role in the text document clustering. The terms related to the sports such as cricket, athletics, football, rugby and tennis are collected from the domain experts. Along with the terms, the relationship between the terms is identified and according to the relationship, it will be placed in the class hierarchy.

C. Creating Class hierarchy

Once the classes and subclasses are identified along with their relationship, the classes are organized into a class hierarchy which is otherwise called as taxonomy. Class hierarchy consists of super class-subclass relationships. Subclasses specialize their super classes. For example, consider the classes kicking and corner kick, Corner kick is a subclass of kicking which implies corner kick is a type of kicking which in turn is a subclass of foul in football class. Initially, the class hierarchy consists of one class called thing. All classes are subclasses of thing. The class hierarchy consists of equivalent classes, disjoint classes or named classes. Two classes can be disjoint with each other if the member or individual of one class cannot be a member/individual of other class. Two classes can be equivalent with other when the member or individual of one class can be member/individual of other class.

D. Properties in class

The properties of ontology are categorized into the following three types,

1) Data properties,
2) Object properties,
3) Annotation properties.

The object property defines the relationship between the concepts of Sports Ontology and the data properties used to define the relationship between the concept individual and the data literal. The annotation properties such as comment, see Also, is defined by etc are used to annotate the concepts of the domain ontology of interest. Definition of each class, subclass concept and list of references to each class and subclass are attached to the node of class hierarchy. For example, the Sports class is divided into five subclasses such as Cricket, Athletics, Football, Rugby and Tennis.

E. Restrictions on Properties

After creating properties, they are used to describe and define the sports ontology classes. Properties are used to create restrictions which restrict the individuals that belong to a class. Restrictions in OWL consist of 3 categories such as Quantifiers, Cardinality, and has Value. Quantifier is a type of restrictions which consists of 3 parts as Quantifier, Property and Filler. Either existential quantifier (some) or universal quantifier (only) may be used according to the purpose. The figure 2 shows the example for existential quantifier some used in Hurdles Class.
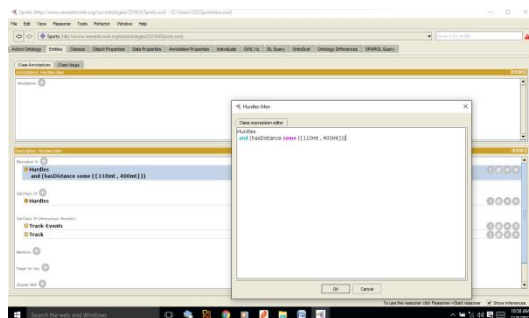

Fig.2: Existential Quantifier some in Hurdles

To describe the class of individuals those have at least, at most or exactly a specified number of relationships with other individuals or data type values. Cardinality restrictions are used to specify the number which allowed in a class that defines that class. Either maximum cardinality or minimum cardinality can be used in a class. For example, In Athletics class, categories class has 5 subclasses such as Junior, Masters, Open-Class, Under-23 and Youth. The has Age object property is used to differentiate these 5 subclasses. For Junior class, the age must be $<=20$ which is described in protégé as Junior has Age max 20 is shown in the figure 3. In Masters Class, the age must be 35 and is described as Masters has Age min 35.
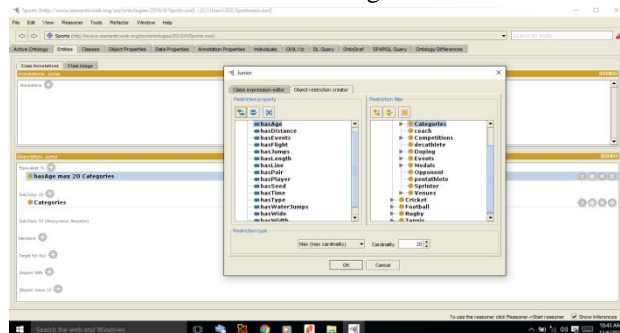

Fig.3: Max Cardinality Restrictions

F. Individuals

Individuals are instances of ontology classes. It is otherwise called as instances. In order to define individuals in protégé, select the required class and then select individuals, individuals window will be displayed then create its instances. For example, to create

individual for winners' class, select winners then create its instances as Steffi and Agassi. An individual can belong to many classes.

IV. EVALUATION

The Sports domain ontology created in protégé tool as a lightweight ontology which is a structured representation of knowledge where the concepts are arranged in a hierarchy with a relationship between them. It consists of five games such as Athletics, Cricket, Football, Rugby and Tennis and is created with the help of domain expert knowledge. The concepts and sub concepts of this sports ontology were called through Jena – using eclipse luna IDE.Jena is the commonly used Java API for RDF, OWL and XML, providing services for model representation, parsing, database persistence and visualization tools. Protégé-OWL is closely related with Jena. The Protégé OWL model is converted into a Jena OntModel, to get a snapshot of the model at the runtime is shown in the figure 4. The retrieved concepts are used for document clustering which is based on ontology to yield better retrieval results.
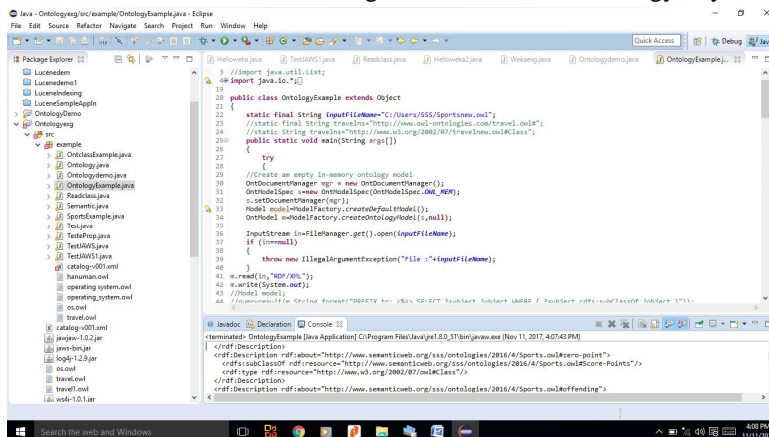

Fig 4: Jena OntModel retrieved in eclipse

V. CONCLUSION AND FUTURE WORK

This work aims to provide detailed information about how to create domain ontology and what are the steps one has to follow to create it. Sports domain ontology is created with the help of widely used ontology editing tool called protégé. The details of the games are collected from the domain expert. This Sports ontology will be useful for sharing the knowledge among people and it can be used in many fields such as Semantic searching and Semantic document clustering. In future, this ontology can be updated by adding more games details.

REFERENCES

[1] Farooq, Amjad, and Abad Shah. "Ontology Development Methodology for Semantic Web Systems." Pakistan Journal of Life Social Sciences 6.1 (2008): 50-58.

[2] Horridge, Matthew, et al. "A Practical Guide To Building OWL Ontologies Using The Protégé-OWL Plugin and CO-ODE Tools Edition 1.0." University of Manchester (2004).

[3] Jones, Dean, Trevor Bench-Capon, and Pepijn Visser. "Methodologies for ontology development." (1998): 62-75.

[4] Malik, Sanjay Kumar, Nupur Prakash, and S. A. M. Rizvi. "Developing an university ontology in education domain using protégé for semantic web."International Journal of Science and Technology 2.9 (2010): 4673-4681.

[5] Malviya, Naveen, Nishchol Mishra, and Santosh Sahu. "Developing university ontology using protégé owl tool: Process and reasoning." International Journal of Scientific & Engineering Research 2.9 (2011): 1-8.

[6] Noy, Natalya F., and Deborah L. McGuinness. "Ontology development 101: A guide to creating your first ontology." (2001).

[7] Patil, S. M., and D. M. Jadhav. "Semantic Search using Ontology and RDBMS for Cricket." International Journal of Computer Applications 46 (2012).

[8] Rajasurya, Swathi, et al. "Semantic information retrieval using ontology in university domain." arXiv preprint arXiv:1207.5745 (2012).

[9] Saad, Aslina, and Shahnita Shaharin. "The Methodology for Ontology Development in Lesson Plan Domain." Methodology 7.4 (2016).

[10] Szekely, Anamaria. "An Approach to Ontology Development in Human Resources Management." Proceedings of the 5th International Conference on Virtual Learning. 2010.

[11] Zhou, Xuezhong, et al. "Ontology development for unified traditional Chinese medical language system." Artificial Intelligence in Medicine 32.1 (2004): 15-27.

[12] Shadbolt, Nigel, Wendy Hall, and Tim Berners-Lee. "The semantic web revisited." Intelligent Systems, IEEE 21.3 (2006): 96-101.

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ☺ (24*7 Support on Whatsapp)