Semantic Web Oriented framework for Knowledge Management in Agriculture Domain

Nidhi Malik, Aditi Sharan
JNU, New Delhi
nidhimalik14@gmail.com
aditisaharan@gmail.com

ABSTRACT: *Agriculture plays a vital role in the economic growth of a country. With the evolution of internet, the agriculture sector is also becoming IT focused. In recent years, Semantic web has added altogether a new dimension to the World Wide Web. It has changed the way the internet was used by people. The Semantic web has provided solutions for intelligent services by providing machine understanding of the contents. This paper attempts to showcase the utility of semantic web in the agriculture domain by elaborating the use of ontologies for knowledge management. The Semantic web is a promising platform for knowledge management. Ontologies provide a shared and common understanding of a domain that can be communicated across people and applications.*

Keywords: Semantic Web, Knowledge, Ontology, Knowledge Management, SPARQL

Received: 12 May 2016, Revised 19 June 2016, Accepted 25 June 2016

1. Introduction

The Semantic web has changed the scenario for World Wide Web. The Internet had made tremendous changes for the society. The Semantic web has added altogether a new dimension to www. It has changed the way the internet was used by people. With the amount and increased number of sources of information, there came problems which arise while handling knowledge which is heterogeneous, scattered and unstructured. So, a new architecture named as Semantic web was built. The notion of semantic web enables to look for information in the context of meaning rather than traditionally looking for information based on the keywords. In other words, semantic web enhances the visibility of information on the web.Semantic Web was introduced with a vision to semantically interpret the information present on the web so that its consumption can be increased.

Ontologies are the central concept of Semantic Web (Ontologies Handbook, 2007). In recent years, Ontology has emerged as a research area and has applications in almost all fields. Technically, an Ontology is defined as a set of concepts and relations between them in a particular domain (Gruber, 1995). In today's scenario, ontologies are being developed in all domains such as biomedical, tourism, health, etc. In agriculture sector also, various efforts are going on in the direction of increasing productivity, improving farmer's conditions by providing them quick solutions through SMS, emails, portals, etc. The volume of data is increasing rapidly, so it needs to be efficiently managed and represented in a way so that it can be effectively utilized.

The paper is organized in five sections. The first section gives the general introduction about semantic web. The second section, we have discussed related work particularly in agriculture domain and the research scope of this work. In the third section, we have discussed the challenges of ontology development in agriculture domain and the research methodology followed. Section 4 gives the framework for design and development. The criteria followed for each step in the ontology development process is clearly stated. The next section concludes the proposed work.

2. Related Work And Research Scope

During the literature review, it has been found that there are already few ontologies developed in the agriculture domain. The Thai Rice Ontology (Thunkijjanukij, 2009) is a prototype ontology for plant production using Thai rice as a case study. This ontology covers all the stages of rice production in Thailand starting from cultivation to harvesting. Since there are no well-structured and organized repositories for plant production, researchers face many problems while trying to find relevant information for their research purposes. Hence, Thai Rice ontology has been designed with an aim to facilitate the process of knowledge acquisition and information retrieval for research purposes. The User Centered Ontology for Sri Lankan Farmers (Walisadeera, 2014) is an ontology whose main aim is to provide agricultural information and relevant knowledge that is complete and structured and specific to user context. This ontology has been developed considering the farmers' needs and also taking into account the questions that vary from farmer to farmer such as farm environment, types of farmers, etc. Some systems are available (Sriswasdi, 2008) which provides tailor built fertilizers to farmers. Such systems are made to target specific problems; they are not generic in nature. (Walisadeea, 2013) have developed an ontology for social life networks. (Walisadeera, 2014) have studied the life cycle of crops and identified different stages of the crop lifecycle to produce an ontology. AGROVOC is a multilingual, structured, controlled vocabulary/thesaurus designed to cover concepts and terminology in agriculture, forestry, fisheries, food and related domains developed by the Food and Agriculture Organization (FAO) of the United Nations and the Commission of the European Communities. Agropedia(is an online knowledge repository for information related to agriculture in India backed by the Government of India and sponsored by the World Bank through the *National Agricultural Innovation Project* of the Indian Council of Agricultural Research (ICAR). It is more useful for those users whose needs are mainly based on crops as the repository is maintained in such a way that it is focused on crops and the related information crop wise. Its main aim is to keep alert on different crops from the scientists and make the farmers aware of it by keeping them updated through text messages. It is a crop wise based knowledge repository(Das, 2012).

This work has been driven by the fact that there is a scope of developing ontologies for representing and extracting information about the fertilizers. Following are motivations of thiswork:

• AGROVOC is the most exhaustive and well-established thesaurus available today in the agriculture domain. But it covers agriculture as whole. It covers fisheries, animal husbandry, etc. and information related to fertilizer is not properly provided in this thesaurus. Also, the work that has been done in agriculture domain is mostly crop-specific. For example, Agropedia, Thai Rice, etc.

• It is important to organise the scattered information on fertilizers so that general purpose questions on fertilizers can be answered.

• The fertilizer ontology can always be upgraded by adding more information to it and also it can be made to merge with other generic ontologies such as soil ontology and crop ontology so that questions related to applications of fertilizer based on soil type and crop can be answered.

3. Challenges And Research Methodology

The building of ontology from scratch is considered as a time consuming and challenging task. Here are the challenges that are usually faced while developing domain ontology in agriculture domain:

3.1 Unavailability of authentic resources

The first step in building any domain ontology from scratch is to have an adequate amount of relevant data which is authentic. In a specific domain, finding the relevant, authentic data is a big task. Firstly, it requires one to contact the domain experts and need to collect the data upon their suggestions. Secondly, there are a lot of ambiguities when it comes to defining the hierarchy of the terms. For example, in the fertilizer sub-domain, defining a proper hierarchy of the types of fertilizers is a confusing one. Thus, the process of extracting and exploiting the resources becomes harder, which leads to problems such as time consuming.

3.2 Question of which methodology to follow

A number of methodologies have been proposed for ontology development over the years by different researchers. However, due to the relatively immature nature of the field of Ontology Engineering, there is no specific ontology development methodology defined for a particular domain or for a particular task or application for that matter. It is the onus on the ontology developer to find out the methodology that best suits the domain as well as the use of the ontology to be developed.

3.3 Classification of classes and properties

The task of defining classes of ontology in the agricultural domain is a challenging one. Firstly, there are no proper hierarchy available that we can follow. We have to consider information from all the available sources and see which one or combination of which data best suits our purpose. This task is a very tedious and time consuming task. Again, once the classes are defined, defining its properties is also a big task. We have to read each and every line of the available resources and find out which information can be represented in the form of ontology.

3.4 Incorporation of statistical data

The domain of agriculture, especially the subdomain of fertilizer has an enormous data in the form of statistical records. There are no specific methods of incorporating such data in ontology. Since the amount of statistical data is too huge, it takes up a lot of time and is a never ending process. The statistical data can always be updated into the ontology.

4. Design Process

During literature survey, we did not come across an ontology developed for fertilizers. Since Agrovoc is the most well established and publicly available resources for agriculture, we started our work by capturing some information from Agrovoc regarding the terms and taxonomical knowledge. In addition to Agrovoc, we also considered other knowledge repositories (digital and non-digital) such as Agropedia (Agropedia, 2000), textbooks. Information in Agrovoc is not structured as per the ontological notations. It was time consuming to gather information which is suitable to be included in the ontology. Therefore, we decided on the properties of the classes and restrictions to be applied in consultation with domain experts. Competency questions were identified after much brainstorming since these derive the overall development process.

There are a number of methodologies available as discussed in. Every methodology has certain pros and cons. We have used the design methodology described in (Gruinger, 1995). The reason behind using this methodology is the high degree of formalism that it provides. The use of FOL enables the user to incorporate semantics which will enhance the structure and functionality of the ontology. The figure below depicts the conceptual framework of the whole process.

Phase 1	Phase 2
1. Scope Identification • Formulate the competency questions	1. Ontology Formalization Concept definitions Class Declarations Property Axioms
2. Identify the knowledge sources • Research journals, textbooks, fertilizers statistics prepared by Fertlizers association of India, mass media etc.	2. Actual Implementation OWL/RDF Protege SPARQL
3. Indentify the Ontology concepts	3. Ontology Evaluation Internal Evaluation Domain Experts

Figure 1. Framework for design and development process

Phase 1 : Ontology Specification, Knowledge Acquisition and Ontology Conceptualization

The first step in the ontology construction is to define domain and scope. In order to define the domain and scope of the ontology, two kinds of questions are sketched:

• Basic questions – defines a list of questions which define the purpose of the ontology and help to limit the scope of the domain such as what is the domain that the ontology will cover, for what we are going to use the ontology, for what type of questions the information in the ontology should provide answers, who will use and maintain the ontology.

• Competency questions – these are the questions that the knowledge base based on the ontology should be able to answer. These questions let us know if the ontology has enough information about the domain or if there is a need a particular level of detail or representation of a particular area. For our research, the list of competency questions was framed by consulting with the domain experts. Some of the competency questions along with the corresponding generalized questions are given below:

Generalized Question	Competency Question
What is the concerned nutrient content in the corresponding fertilizers?	What is the calcium content in various calcium fertilizers?
Which type of phosphatic fertilizers are suitable for a specified type of soil?	Which phosphatic fertilizers are suitable for acidic soil?
Which nitrogenous fertilizers use a specified type of application method?	Which nitrogenous fertilizers use Topdressing type of fertilizer application method?
List the phosphatic fertilizers that contain phosphorus in the form of a specified type of phosphate form.	List the phosphatic fertilizers that contain phosphorus in the form of dicalcium phosphate.

Table 1. List of competency questions with the corresponding generalized questions

Looking at the competency questions above, the ontology will contain information on the various types and subtypes of fertilizers available, the application time of different fertilizers and the statistical information of the fertilizers.

Knowledge Acquisition

This step is about extracting the domain specific knowledge from the knowledge resources as much as possible. In order to fulfil this need, we have gone through numerous online sources, books and contacted domain experts. We have extracted the knowledge required for the fertilizer ontology development from the following reliable sources:

• Domain experts from the NCAP, ISRI, New Delhi;

• Research journals and papers, Fertilizer use by crop in India (2005);

• Textbooks (A handbook of Fertilizer, Soil and Manure; Agricultural Handbook);

• Fertilizer Statistics 2011-12, prepared by The Fertilizer Association of India;

• Online data sources from Authoritative Organizations (the Ministry of Agriculture - India)

• Mass media (newspaper, television, radio).

Ontology Conceptualization

This step involves building of the conceptual model by following the specification listed in the previous step. Ontology is a data model that represents the concepts related to the domain and the relationships among these concepts. And so, as pointed out earlier also, concepts and relations are the main components of an ontology. Relationships can be of two types: hierarchical relationships and associative relationship. Hierarchical relationships are those between the concepts of the same hierarchy, i.e., between superclass and subclass. Associative relationships are the relationships between concepts which are in different hierarchies.

For conceptualization, much of the work is done following the different steps proposed by Noy and McGuinness (2001).

Define the classes and class hierarchy

We have followed the combination of top down and bottom up approaches for development of ontology. We start with a top-level concept 'Fertilizer', and a more specific concept 'AmideFertilizer'. And then we can relate them to a middle-level concept, 'Nitrogenousb Fertilizer'.

Noy and McGuinness had discussed several guidelines to keep in mind while developing a class hierarchy. We consider these guidelines and use them to check against the class hierarchy that we have created for our ontology.They had set various rules of thumb that help us in deciding the class hierarchy.

Subclasses of a class usually
(1) have additional properties that the superclass does not have, or
(2) restrictions different from those of the superclass, or
(3) participate in different relationships than the superclasses.

Keeping these rules in mind, we define our class hierarchy as follows:

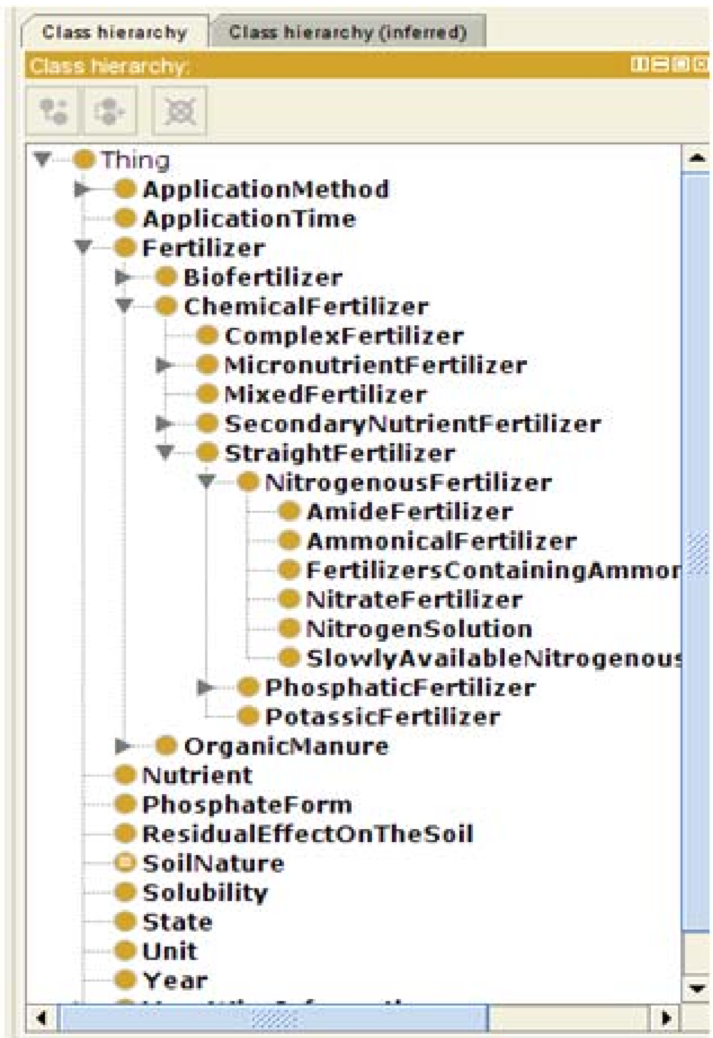

Figure 2. Fertilizer class hierarchy

Identify relationships

• There is only one relation for the hierarchical relationships, namely "hasSubclass". This relation is defined between all of the hierarchical concepts. Some of the hierarchical relationships are given in the table below:

Subject concept	Relation	Object concept
Fertilizer	hasSubclass	Chemical Fertilizer
Chemical Fertilizer	hasSubclass	Straight Fertilizer
Straight Fertilizer	hasSubclass	Nitrogenous Fertilizer
Nitrogenous Fertilizer	hasSubclass	Amide Fertilizer
Amide Fertilizer	hasSubclass	Urea

Table 2. Concept and Hierarchical relationship

• Associative relationships are assigned between concepts belonging to different hierarchies. These relationships are defined by identifying verbs related between concepts and assigning relation name that would form a meaningful statement with the name of the concepts. Some of the examples of associative relationships are given in the table below:

Subject concept	Relation	Object concept
Fertilizer	hasApplicationMethod	Application Method
Phosphatic fertilizer	hasSuitableFertilizer	Soil Nature
Phosphatic Fertilizer	containsPInTheFormOf	Phosphate Form
Urea	hasChemicalFormula	$CO(NH_2)_2$
Urea	hasPercentN	46

Table 3. Concept and Associative relationships

Define the properties of classes-slots

In order to answer the competency questions defined in step 1, we must describe properties of the classes. There are two types of properties:

• ***Data Properties:*** These are the properties that relate individuals to a user-defined value.
• ***Object Properties:*** These are the properties that relate the individuals of a class to individuals of another class.

From the terms that we have enumerated in step 1, after selecting the classes, most of the terms that are left will be the properties of these classes. For example, for the term *Fertilizer*, its properties would be a fertilizer's *nutrient content, application method, application time,* etc.

Create instances

Creating instances of classes is the last step. We select a class and create an individual of that class and fill the values of its properties. For example, we create an individual instance 'BasicSlag' of the class 'PhosphoricFertilizer'. Few properties of this instance are as below:

canBeMixedWith: RockPhosphate
canBeMixedWith: PotassiumSulphate

hasApplicationTime : WellBeforeSowingTheCrop
hasApplicationMethod : BroadcastingAtPlanting
containsPInTheFormOf : DicalciumPhosphateForm
containsPInTheFormOf : CitricAcidSolublePhosphoricAcidForm
hasPercentP: 3-8

Phase 2 Ontology Formalization, implementation and Evaluation :

In this phase, the concepts are defined through axioms. The axioms in OWL ontology provide explicit logical assertions about three types of things - classes, individuals and properties. Many types of axioms can be expressed in OWL ontology. We have used the OWL functional syntax format for easy understanding of the axioms.

We have made use of First Order Logic to represent the competency questions. Below we have shown an example from table 1:

Competency Question: List the phosphatic fertilizers that contain phosphorus in the form of dicalcium phosphate.

(∃x)(NitrogenousFertilizer(x)) ^ (∃ y)(PhosphateForm(DicalciumPhosphateForm)) ^ containsPInTheFormOf(x, DicalciumPhosphateForm);

Ontology Implementation and Evaluation

There are many ontology implementation languages available. The language that we are using is OWL (Web Ontology Language). OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDF Schema (RDFS) by providing additional vocabulary along with a formal semantics. *Components of OWL*

- Individuals - Individuals are also known as instances. Individuals can be referred to as being instances of class.
- Properties - Properties are roughly equivalent to slots .They are also known as roles in description logics and relations in UML and other object oriented notions and attributes in other formalisms.
- Classes - The word concept is sometimes used in place of class. Classes are a concrete representation of concepts.

The tool used to create the ontology is Protégé Ontology Editor.The domain experts have also validated the components of the ontology.SPARQL (Prud'hommeaux, 2008) has been used to query information from the ontology. It has been done to check whether the ontology is following the ontology construction principles in addition to the reasoning and inferencing functionalities provided by the reasoned available in Protégé (Knublauch, 2004). Below are two examples of SPARQL queries.

Which chemical fertilizers can be mixed physically and used as mixed fertilizers?
```
SELECT DISTINCT ?ChemicalFertilizer ?canBeMixedWith
        WHERE { ?ChemicalFertilizerfer:canBeMixedWith ?canBeMixedWith }
```
Which nitrogenous fertilizers use Topdressing type of fertilizer application method?
```
SELECT ?NitrogenousFertilizer
        WHERE { ?NitrogenousFertilizerfer:hasApplicationMethod ?Topdressing }
```

Conclusion

Knowledge management is a crucial activity for any organization in any domain. Information is of no use if it is not managed and presented in the required form(Kebede, 2010).Ontology surely is a preferred medium for knowledge representation and management. We chose to develop an ontology for fertilizers as currently there is no ontology available for fertilizers. (Poveda, 2010)has been studied to know about the common pitfalls in ontology development. To show the feasibility and usefulness of the ontology, we have also shown how to query information from ontology. We have explained in detail the overall process and different concerns involved in design and development of an ontology. We have evaluated the ontology internally and resolved the problems encountered in consultation with agriculture scientists.

We further wish to integrate our ontology with soil and crop ontology so that we able to extract more specific information. We have got valuable suggestions from domain experts andwill further refine the ontology.

Acknowledgements

We have been in touch with scientists from Indian Agriculture Research Institute (IARI). We thank Dr. Rajni Jain, Principal Scientist, NCAP (IARI) and Ms. Pavithra, Scientist, NCAP (IARI) for giving valuable suggestions on the kind of information that should be included in the ontology.

References

[1] Agropedia, http://agropedia.iitk.ac.in

[2] AGROVOC Multilingual Thesaurus (2000) http://aims.fao.org/agrovoc .

[3] Brewster, C., O'Hara, K. (2007). Knowledge representation with ontologies: Present challenges—Future possibilities. *International Journal of Human-Computer Studies*, 65 (7) 563-568.

[4] Das, M., Malhotra, P. K., Marwaha, S., Pandey, R. N. (2012). Building and querying soil ontology. *J. of ISAS*, 66 (3) 459-464.

[5] Fisseha, F. (2002).Towards better Semantic Standards for Information Management AGROVOC and the Agricultural Ontology Service (AOS). UN FAO, Rome, Italy.

[6] Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-computer Studies,* 43 (5) 907-928.

[7] Grüninger, M., Fox, M. S. (1995). Methodology for the Design and Evaluation of Ontologies.

[8] Gupta P.K. (2010). Handbook of Soil, fertilizers and manure, Agrobiospvt. Limited, India.

[9] Handbook of Agriculture. (2011). Published by Indian Council of Agriculture Research. New Delhi.

[10] Igoche, Bernard (2014). Using ontologies for knowledge management, *International Journal of Scientific and Research Publications,* Volume 4 (1) January 2014.

[11] Kebede, G. (2010). Knowledge management: An Information Science Perspective, *International Journal of Information Management*, 30, p. 416-424.

[12] Kishore (2007) Ontologies : A handbook of Principles, Concepts and Applications in Information Systems, Springer.

[13] Knublauch, H., Fergerson, R. W., Noy, N. F., Musen, M. A. (2004). The Protégé OWL plugin: An open development environment for semantic web applications. *In: The Semantic Web–ISWC 2004* (p. 229-243). Springer Berlin Heidelberg.

[14] Noy, N. F., McGuinness, D. L., (2001). Ontology development 101: A guide to creating your first ontology.

[15] Poveda, M., Suárez-Figueroa, M. C., Gómez-Pérez, A., (2010). Common pitfalls in ontology development. In *Current Topics in Artificial Intelligence* (p. 91-100). Springer Berlin Heidelberg.

[16] Prud'hommeaux, E., Seaborne, A., 2008. SPARQL query language for RDF. W3C Recommendation.Retrieved from http://www.w3.org/TR/rdf-sparql-query.

[17] Semantic Web, http://www.w3.org/standards/semanticweb

[18] Shadbolt, N.R., Lee Berners T., Hall, W. (2006). The Semantic Web Revisited, *IEEE Intelligent Systems,* 21 (3) 96-101.
Sriswasdi, W., Luengsrisagoon, S., Lorsuwansiri, N., Wuttilerdcharoenwong, S., Khunthong, V., Suksaengsri, T.Pusittigul, A. (2008, August). A smart mobilized fertilizing expert system: 1-2-3 personalized fertilizer. *In: World Conference on Agricultural Information and IT, Tokyo, Japan* (p. 397-404).

[19] Uschold, M., Gruninger, M. (1996). Ontologies: Principles, methods and applications. *The Knowledge Engineering Review*, 11 (02) 93-136.

[20] Thunkijjanukij, A., Kawtrakul, A., Panichsakpatana, S., Veesommai, U. (2009). Rice production knowledge management: criteria for ontology development. *Thai Journal of Agricultural Science*, 42 (2) 115-124.

[21] Walisadeera, A. I., Ginige, A., Wikramanayake, G. N. (2014). User centered ontology for Sri Lankan farmers. *Ecological*

Informatics.

[22] Walisadeera, A. I., Wikramanayake, G. N., Ginige, A. (2013). Designing a Farmer Centered Ontology for Social Life Network. *In: DATA* (p. 238-247).

[23] Walisadeera, A. I., Ginige, A., Wikramanayake, G. N. (2014). Conceptualizing Crop Life Cycle Events to Create a User Centered Ontology for Farmers. *In: Computational Science and Its Applications–ICCSA 2014* (p. 791-806). Springer International Publishing.