



## Web based software for the study of USDA soil taxonomy and classification of newly found soil

MANORANJAN DAS<sup>1</sup>, P K MALHOTRA<sup>2</sup>, SUDEEP MARWAHA<sup>3</sup> and R N PANDEY<sup>4</sup>

Indian Agricultural Statistics Research Institute, New Delhi 110 012

Received: 31 January 2012; Revised accepted: 4 July 2014

### ABSTRACT

United States Department of Agriculture (USDA) Soil Taxonomy is based on soil properties that can be objectively observed and measured in the natural conditions as they exist today. There are many soil classification systems but USDA Soil Taxonomy is most accepted worldwide. Ontologies are the new form of knowledge representation that acts in synergy with agents and Semantic Web Architecture. Soil ontology developed for USDA soil taxonomy has been used to develop a query interface that will help in detailed study of soil taxonomy, classification of new soil as well as exchange knowledge between software agents and systems. This is a web based application having N-tier architecture. Application development environment is NetBeans 6.9 editor and Protégé. Web development technology is Java Server Pages (JSP). Programming languages JAVA and SPARQL are used for querying. Client interface is developed with Hyper Text Markup Language (HTML), Cascading Style Sheet (CSS) and JavaScript. Third tier of software consist of database which is in MS-SQL server 2005. Other two layers are Web Ontology Language (OWL) Ontology layer and Semantic Web Framework layer. OWL layer contains soil taxonomy information in the form of Ontology. Semantic Web Framework layer is implemented using JENA. In the search panel user can search anything related to USDA Soil Taxonomy, which comprises of twelve orders. However, this software contains information about seven soil orders reported in India. Domain experts can see and edit the knowledge base (i.e. Soil Ontology) or can suggest anything related to the creation of Soil Taxonomy Ontology through WebProtégé.

**Key words :** Ontology, OWL, USDA Soil Taxonomy, WebProtégé

Soil Taxonomy is based on soil morphology that can be objectively observed and measured in the field. There are many soil classification systems but the United States Department of Agriculture (USDA) Soil Taxonomy (Soil Survey Staff 2010) is most accepted worldwide. Ontologies are the new form of knowledge representation that acts in synergy with agents and Semantic Web Architecture (Berners-Lee *et al.* 2001). Ontologies define domain concepts and the relationships between them, and thus provide a domain language that is meaningful to both humans and machines. Building ontologies in different domains of USDA Soil Taxonomy which is presently available in text form will help to convert unstructured knowledge into structured one that can be shared across applications. Soil Ontology has been developed for USDA soil taxonomy from Order to Sub group level (Das 2010). A query interface, Soil Taxonomy Ontology, developed using soil ontology for USDA Soil Taxonomy will help in detailed study of soil taxonomy, classification of new soil as well as exchange knowledge between software agents and systems. In this

paper development of such a web based application has been presented for seven soil orders extensively reported in India (Sehgal 1996).

### MATERIALS AND METHODS

While designing the software, the aspect of exchange of knowledge between software agents and systems has been taken into consideration. The tools and techniques have been chosen accordingly. The basic design of software is as shown in Fig 1.

N-tier architecture is used to develop the software. Application development environment is NetBeans 6.9 editor and Protégé, the Web Ontology Language (OWL) ontology editor for creating Soil Ontology. Web development technology is Java Server Pages (JSP). Programming languages JAVA and SPARQL (Clark 2008) are used for querying. Client interface is developed with Hyper Text Markup Language (HTML), Cascading Style Sheet (CSS) and JavaScript. Third tier of software consist of database

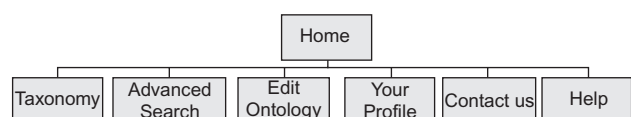


Fig 1 Design of Software

<sup>1</sup>Senior Research Fellow (email: cac.mano@gmail.com); IARI, New Delhi; <sup>2</sup>Principal Scientist (email: pkm@iasri.res.in); <sup>3</sup>Senior Scientist (email: sudeep@iasri.res.in), <sup>4</sup>Principal Scientist (email: rnpandeyssaciari@rediffmail.com), IARI, New Delhi

which is in MS-SQL server 2005. This makes the software login secured. Other two layers are OWL Ontology layer (Smith *et al.* 2004) and Semantic Web Framework layer (Berners-Lee *et al.* 2001). OWL layer contains soil taxonomy information in the form of Ontology, which is human as well as machine readable and Semantic Web Framework layer is implemented by using JENA, a java framework for building Semantic Web applications.

Options to navigate the system are available to users only after they logon to system which are available on home page (Fig 1). This system offers two types of user levels, one is *Normal User* and other is *Domain Experts* (those having detailed soil taxonomic knowledge) during sign up to the system. Domain Experts are those users who can edit, view or add anything new to the knowledge base with prior permission of administrator. After log in all options except “*Edit Ontology*” are available to users those who signed up as a normal user. “*Edit Ontology*” option is available to those users who signed up as a *Domain Expert*.

RESULTS AND DISCUSSION

After sign up, the first page appears (Fig 2). By using *Taxonomy* navigation key one can study USDA Soil Taxonomy in detail up to Subgroup level of the orders (*Alfisol*s, *Aridisol*s, *Entisol*s, *Inceptisol*s, *Mollisol*s, *Ultisol*s and *Vertisol*s) available in India (Fig 3). It will show the definitions, keywords in the definition and photograph.

By using *Advance Search* navigation key one can easily classify a newly found soil up to Subgroup level. This module is menu driven and the user can choose the option up to what level he/she wants classification. User have to answer some of the questions (*listed below*) for classifying the soil. After selecting any probable result it will give the taxonomy and detailed information of that soil, so that user can easily compare the property of the soil which he/she found with the property of probable soil generated by system. After comparison user can ensure whether the newly found soil exactly matches with the system generated soil or not, if not the user should move to other probable results for proper classification. A set of questions asked to the user for classification of soil are:

- Presence of
- 1. Which diagnostic surface horizons (epipedons)?
- 2. Which diagnostic subsurface horizons (endopedons)?
- 3. Which temperature regimes?
- 4. Which moisture regimes?
- 5. What other properties?

A searching technique has been implemented to search soil Orders, Suborders, Great groups and Subgroups by mentioning their properties (Fig 4).

The soil ontology (knowledge base) can be edited by domain experts if there is any wrong entry by the system developer or any new information is available regarding a

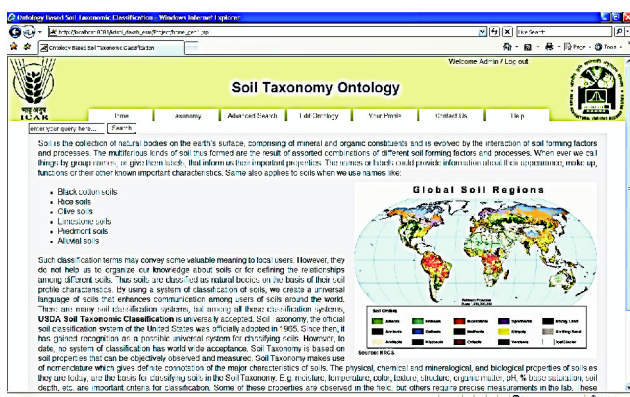


Fig 1 First page after log in

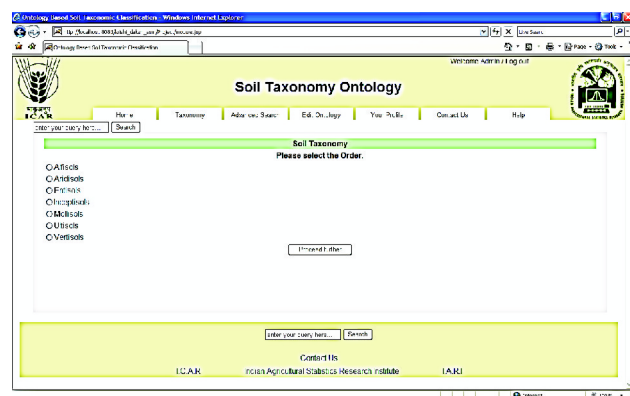


Fig 2 Selection of soil orders for detailed information

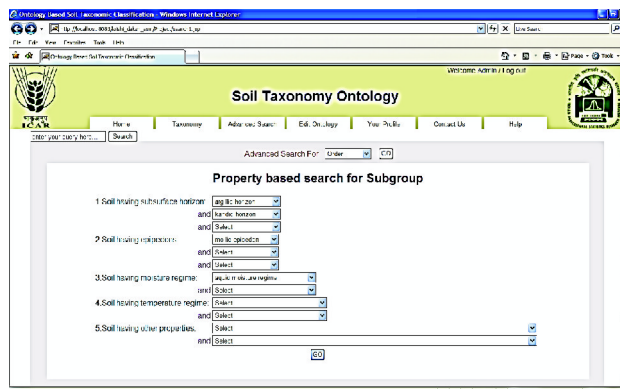


Fig 3 Property based search for subgroup level

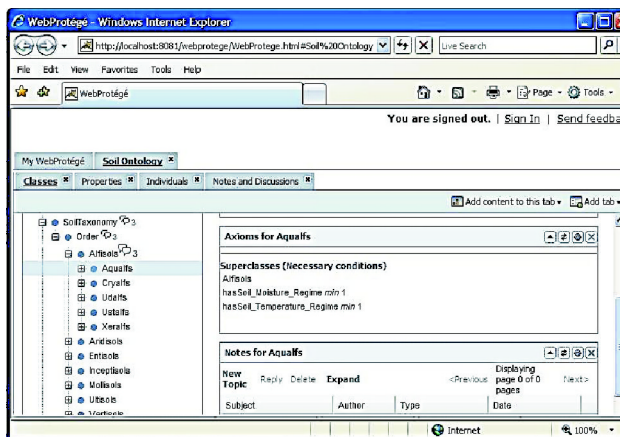


Fig 4 WebProtégé window showing classes of Soil Ontology in Class Tab

particular topic later. The domain expert has the right to edit the knowledge base of this software and he can make corrections in entries as well as add new entries. This can be done by using *Edit Ontology* navigation key (Fig 4). On clicking this link, WebProtégé (*web version of protégé for editing ontologies online*) will open on a blank page and domain expert need to log in again for security reason to edit the ontology. No need to log in again for simply viewing the Ontology. In user interface of WebProtégé, users can customize the content and layout of the User Interface (UI). The WebProtégé user interface is made up of tabs (e.g. Classes Tab, Property Tab, etc.). Each tab contains several portlets (e.g. the class tree, property values, restriction view, notes, etc.). The user is able to resize and rearrange the portlets in a tab by drag-n-drop. *My WebProtégé* tab contains the Soil Ontology which is available for viewing and/or editing. Ontology will be opened simply by clicking on ontology name in the name column. After opening of ontology, one can view the ontology class hierarchy on the Classes tab. By selecting a class in the *class tree*, one can view the class properties, axioms, notes, etc. in the portlets to the right of the class tree.

As the software allows domain experts to edit, view and add new information in the knowledge base, the administrator must have right to block any fake “domain expert” from making wrong changes in knowledge base. In this software this right is given to administrator only. The software provides online help for editing the knowledge base (*i.e. Soil Ontology*) and for using software.

#### REFERENCES

- Berners-Lee T, Hendler J and Lassila O. 2001. The Semantic Web. *Scientific American* **284** (5): 34–43.
- Clark K. 2008. *SPARQL Protocol for RDF, W3C Recommendation*, <http://www.w3.org/TR/rdf-sparql-protocol/>.
- Das M. 2010. ‘Building and querying soil ontology for agriculture.’ M Sc thesis, IARI, New Delhi.
- Smith M K, Welty C, and McGuinness D L. 2004. *OWL Web Ontology Language Guide (W3C Recommendation)*. <http://www.w3.org/TR/owl-guide/>.
- Sehgal J. 1996. *Pedology Concepts and Applications*. Kalyani Publishers, Ludhiana.
- Soil Survey Staff. 2010. *Keys to Soil Taxonomy*, 11<sup>th</sup> Edition. United States Department of Agriculture, -Natural Resources Conservation Service, Washington, DC.