

## Sustainable Agriculture through ICT innovation

**Development of Simple Web Systems to Support Characterization of Information Resources: A Step towards Knowledge Representation and Sharing**Yolanda M. Fernández-Ordóñez<sup>1</sup>, and Jesús Soria-Ruiz<sup>2</sup>

<sup>1</sup>Applied Computing Program. Postgraduate College in Agricultural Science (COLPOS)  
Campus Montecillo, Carr. México-Texcoco km 36.5, Montecillo, Texcoco 56230,  
Mexico. [yfernand@colpos.mx](mailto:yfernand@colpos.mx)

<sup>2</sup>Geomatics Lab. National Institute of Research for Forestry, Agricultural and Livestock  
(INIFAP). Vialidad A. López Mateos Km. 4.5 Carr. Toluca-Zitácuaro.  
Zinacantepec 52107, Mexico. [soria.jesus@inifap.gob.mx](mailto:soria.jesus@inifap.gob.mx)

**ABSTRACT**

Knowledge representation and sharing are part of the long term vision of the Semantic Web. Knowledge domains related to agriculture are present in these issues and many problems remain to be addressed concerning knowledge storage and processing. Information resources representation and retrieval through meta knowledge is addressed in this paper. We present our work towards information resources characterization and associated software retrieval related to forests and forestry. One aim is the enhancement of the resource descriptors by allowing users to help define new search terms. This approach will contribute to the knowledge engineers' understanding of how users desire to search for information based on concepts relevant to them.

**Keywords:** Information resources, knowledge representation, knowledge sharing, forest, forestry, Mexico.

**1. INTRODUCTION**

Knowledge representation and sharing are part of the long term vision of the Semantic Web (SW) which originally “described the evolution of a Web that consisted largely of documents for humans to read to one that included data and information for computers to manipulate” (Shadbolt *et al.*, 2006). As in many other domains, knowledge related to agriculture and forestry is present in the SW concerns, particularly from the point of view of sustainable agriculture. Knowledge is mostly generated and systematized in the scientific research communities but is requested by many end users, from decision makers to academicians and students in different fields and even by common people. Many problems remain to be addressed concerning knowledge representation and management to promote effective communication and consultation between the producers of research and the end-users. Glossaries, standard domain vocabularies, ontologies, support models and languages have been proposed to make semantic contents of web resources available for machine processing (Uschold, 2003) and have been applied to agricultural domains (FAO, 2013) (Maliappis, 2007). Comprehensive

## Sustainable Agriculture through ICT innovation

platforms and systems are also been proposed for other domains with similar justifications of effective communication requirements (Pang *et al.*, 2003). Among the main issues is the representation of information resources so that they can be retrieved, not directly through the knowledge they contain but through meta knowledge. Our line of work is oriented towards developing and testing information resources characterizations. In this paper we present our approach to describe information resources for forests and forestry (IRFFs), and we how we propose to enhance the descriptors by allowing users to define new search terms through associated software. This will contribute to the understanding of our knowledge engineers of how real users desire to search for information based on concepts relevant to them and in the longer term this understanding will support building of advanced constructs such as ontologies.

## 2. MATERIAL AND METHODS

### 2.1. Information resources for forests and forestry

Information resources for forests and forestry (IRFFs) come from a variety of producers such as government and academic institutions, commercial businesses, and professional consultants. We considered producers which are of interest to domain researchers, academicians and students in Mexico, such as the National Forestry Commission, CONAFOR ([www.conafor.gob.mx](http://www.conafor.gob.mx)); the Ministry of Agriculture SAGARPA ([www.sagarpa.gob.mx](http://www.sagarpa.gob.mx)); the National Institute of Statistics, Informatics and Geography institute, INEGI ([www.inegi.gob.mx](http://www.inegi.gob.mx)); the Ministry of Environment and Natural Resources, SEMARNAT ([www.semarnat.gob.mx](http://www.semarnat.gob.mx)); and the Commission for the Use and Knowledge of Biodiversity, CONABIO ([www.conabio.gob.mx](http://www.conabio.gob.mx)). These organizations provide official maps and statistics, training materials, technical documents, maps, photographs, satellite images, multimedia files, and other kinds of digital documents through web sites. As IRFFs we have considered information systems, databases and digital multimedia libraries and other sources such as the Forestry Compendium (CABI, 2013). One categorization criterion for the IRFFs is the storage media and format, and another is the contents (theme or main subject matter). Forest researchers suggested that main queries to IRFFs existing in the sites we would use refer to two broad themes: timber or non-timber tree species. These criteria were used for structuring our IRFFs repository and for the implementation of the management system.

The IRFFs were separated by category according to the ten types shown in Table 1 and metadata were defined to describe their contents. Subtypes were considered which inherit the metadata definition but can allow additional properties to be assigned. For reasons of space only some subtypes are shown. The type of an information resource is an attribute. Common values are used to assign a resource to a category. Subtypes may relate to themes or the main subject matter of the information object, or they can have other identifying properties. For example, the subtype chapter of a book will inherit the definition of the publishing enterprise for the type book but can have different authors

---

### C0196

Y.M. Fernández-Ordóñez. J. Soria-Ruiz. "Development of Simple Web Systems to Support Characterization of Information Resources: A Step towards Knowledge Representation and Sharing" EFITA-WCCA-CIGR Conference "Sustainable Agriculture through ICT Innovation", Turin, Italy, 24-27 June 2013.

## Sustainable Agriculture through ICT innovation

from another chapter in the same book. The information resources were then grouped according to sub-themes. Each sub-theme has its associated information resources and its own attributes (@), and elements are established via a conceptual schema. Comments to some entries in Table 1 appear in square brackets. This provides for a better organization of the information in order to identify and retrieve it.

Table 1. Information object categories.

INFORMATION OBJECT CATEGORIES	TYPES / *SUBTYPES
1. Document	Same as types/subtypes of e-document
2. Document-Like Object (DLIO) [Refers to any discrete information resource object formally described by Dublin Core metadata.]	Text
	Image
	Movie
3. Electronic document (e-document)	Book / *chapter; *page
	Document / *index, *glossary, *thesaurus
	Image / *video; *photograph; *map [Multimedia <sup>1</sup> : video, image and map.]
	Software
	Library
	Magazine
	Manual
	Mass media [Mass media: newspaper, interview, magazine, news.]
	Newsletter
	Newspaper article
	Scientific article
	*published
	*unpublished
	Technical report
Thesis	
4. Event	Conference
	Congress
	Webinar
5. Information Provider	Organization / *government; *university; *ngo; *private enterprise
6. Learning resource	Same as types/subtypes of e-document [The LR category includes any e-material used by teachers as support for student learning.]

<sup>1</sup> In a previous version, the underlined appeared as categories with shown types.

### C0196

Y.M. Fernández-Ordóñez. J. Soria-Ruiz. "Development of Simple Web Systems to Support Characterization of Information Resources: A Step towards Knowledge Representation and Sharing" EFITA-WCCA-CIGR Conference "Sustainable Agriculture through ICT Innovation", Turin, Italy, 24-27 June 2013.

## Sustainable Agriculture through ICT innovation

(“continued”) Table 1. Information object categories.

7. Ontology	Knowledge object [An object constituted by a body of concepts, formally described and published]
8. Person	Academician
	Entrepreneur
	Expert
	Journalist
	Researcher
	Student
	Teacher
9. Software	Other
	Algorithm
	Database
	Corporate memory
	Knowledge base
	Program
	System
10. Web site	Tool
	Blog
	Discussion forum
	Gallery
	Portal
	Wiki

### 2.2. Repository of IRFFs

A repository of IRFFs is built following perspectives from corporate memories, it constitutes the semantic forest memory or MSF (from the equivalent acronym in Spanish of “Memoria Semántica Forestal”). Different strategies could be followed to append semantics to the IRFFs, such as corporate memories, an empirical strategy used in certain enterprises to manage their knowledge assets (D. Te’eni and H. Weinberger , 2000). The integration of corporate knowledge pertaining to business organizations and the organization of meaning of information resources for semantic querying over the web, pursue similar aims in terms of the conceptual organization of the IRs and the tools needed to manage and query large repositories (Dieng-Kuntz, R. 2002) . If one considers information resources restricted to the particular domain of forest and forestry –or indeed to any other domain—the corporate memory perspective is useful.

A fragment of the categorization of IRFFs is shown in Figure 1 where the subject matter or theme appears in the construction of the hierarchy. Names in the figure have been translated from the original Spanish names. The conceptualization of IRFFs based on

#### C0196

Y.M. Fernández-Ordóñez. J. Soria-Ruiz. “Development of Simple Web Systems to Support Characterization of Information Resources: A Step towards Knowledge Representation and Sharing” EFITA-WCCA-CIGR Conference “Sustainable Agriculture through ICT Innovation”, Turin, Italy, 24-27 June 2013.

## Sustainable Agriculture through ICT innovation

objects with properties could be used later on for obtaining elements for an ontology. In our work, as many IRFFs as were considered representative of important sets were gathered. These IRFFs were described, organized, and incorporated into a first version of the semantic memory, MSF-1 according to the categorization described in the previous section. Once the characterization by themes and sub-themes was validated by experts (species: timber, non-timber; agroforestry practices: alley cropping, zonal planting, tree-crop interface), an XML schema definition (XSD) was produced.

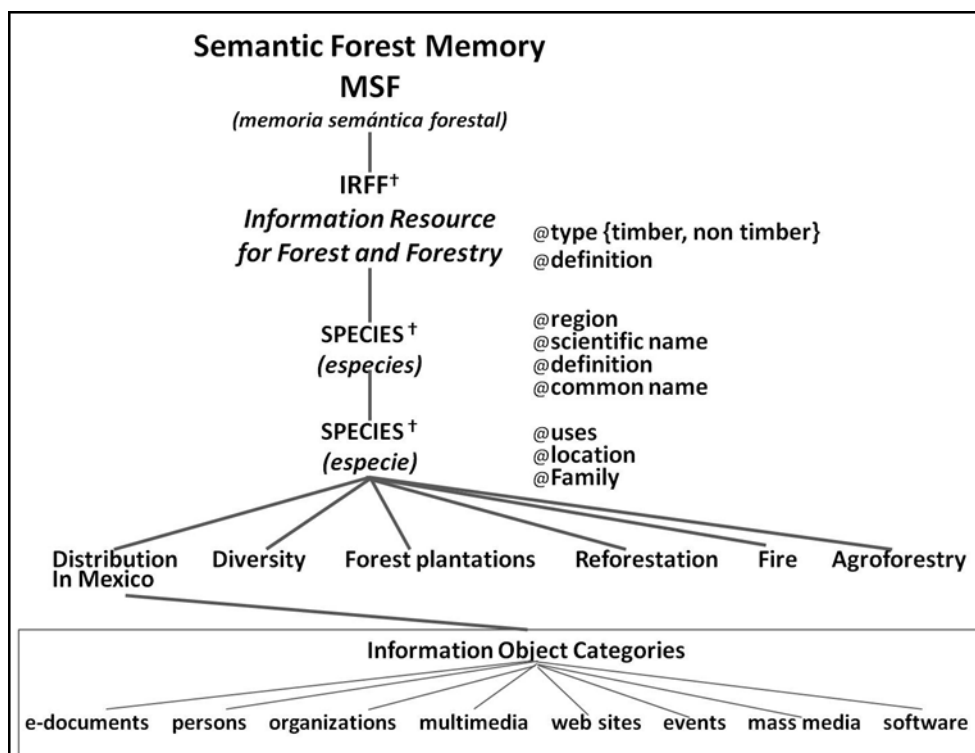


Figure 1. Conceptual schema for the IRFFs repository.

Figure 2 shows a fragment of the XSD schema with the original names in Spanish. The XSD is used to describe the structure and content restrictions of the XML documents in a precise manner and in such a way that the documents are considered valid according to the established schema (Walmsley, 2012). The conceptualization of IRFFs based on objects with properties could be used later on for obtaining elements for an ontology.

### 3. SYSTEM ARCHITECTURE

As explained in the previous section, the implementation approach considers two main parts: (1) XML documents corresponding to the categories of the IRFFs classification; (2) Validated XML documents through the XSD schema which besides structure describe contents restrictions. Figure 3 summarizes the architecture of the pilot system named SISFOR to visualize and update the MS. It accepts three types of

#### C0196

Y.M. Fernández-Ordóñez, J. Soria-Ruiz. "Development of Simple Web Systems to Support Characterization of Information Resources: A Step towards Knowledge Representation and Sharing" EFITA-WCCA-CIGR Conference "Sustainable Agriculture through ICT Innovation", Turin, Italy, 24-27 June 2013.

## Sustainable Agriculture through ICT innovation

users: Administrators (all privileges), Domain Experts (forestry experts with update privileges), and General User (anyone searching for IRFFs).

```
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:element name="memoria_forestal" type="T_memoria_forestal"/>
  <xsd:complexType name="T_memoria_forestal">
    <xsd:sequence>
      <xsd:element name="recurso" type="T_recurso" maxOccurs="2"/>
    </xsd:sequence>
  </xsd:complexType>
  <xsd:complexType name="T_recurso">
    <xsd:sequence>
      <xsd:element name="especies" type="T_especies"/>
    </xsd:sequence>
    <xsd:attribute name="tipo_recurso" type="Tipo_recurso"/>
    <xsd:attribute name="definicion" type="xsd:string"/>
  </xsd:complexType>
  <xsd:complexType name="T_especies">
    <xsd:sequence>
      <xsd:element name="especie" type="T_especie" maxOccurs="unbounded"/>
    </xsd:sequence>
  </xsd:complexType>
  <xsd:simpleType name="Tipo_recurso">
    <xsd:restriction base="xsd:string">
      <xsd:enumeration value="maderable"/>
      <xsd:enumeration value="no_maderable"/>
    </xsd:restriction>
  </xsd:simpleType>
  <xsd:complexType name="T_especie">
    <xsd:sequence>
      <xsd:element ref="subtema" maxOccurs="unbounded"/>
    </xsd:sequence>
    <xsd:attribute name="region" type="T_region"/>
    <xsd:attribute name="nombre_cientifico" type="xsd:string"/>
    <xsd:attribute name="nombre_comun" type="xsd:string"/>
    <xsd:attribute name="usos" type="T_usos"/>
    <xsd:attribute name="ubicacion" type="xsd:string"/>
    <xsd:attribute name="familia" type="xsd:string"/>
  </xsd:complexType>
</xsd:schema>
```

Figure 2. XSD schema

The query interface allows for insertion of items into the MSF. Queries are processed in a web browser. The technique to retrieve information from XML documents is based on the application interface of the Document Object Model (<http://www.w3.org/DOM/>), a platform and language-neutral interface, which allows programs and scripts to dynamically access and update the content, structure, and style of documents.

---

**C0196**

Y.M. Fernández-Ordóñez, J. Soria-Ruiz. "Development of Simple Web Systems to Support Characterization of Information Resources: A Step towards Knowledge Representation and Sharing" EFITA-WCCA-CIGR Conference "Sustainable Agriculture through ICT Innovation", Turin, Italy, 24-27 June 2013.

## Sustainable Agriculture through ICT innovation

The XSLT processor allows for the transformation of XML documents into XHTML documents. XSL style sheets are created, which are representations of the XML document close to the final format. The style sheets contain the appropriate text in the right order; the appearance look and formatting details are incorporated using the cascading style sheets language (CSS) for font types, sizes, and colors. When a query is presented by the user, the Apache server with PHP accepts the petition and the item is searched for in the MSF. The server is in charge of transforming the query results from XML into XHTML format via the XSL style sheets. The validation of data stored in the MSF is performed via an XML schema (XSD).

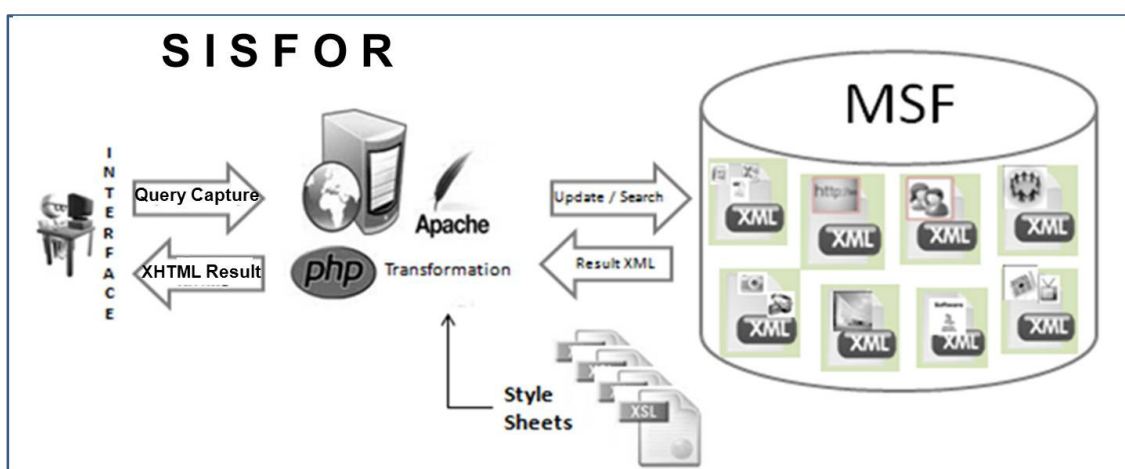


Figure 3. System architecture

#### 4. RESULTS

The results of this research and development project consist of a semantic forest memory MSF built from a corporative memory perspective which contains information resources concerning forest and forestry related themes (IRFFs), ranging from standard e-documents to multimedia resources. These are stored and organized in the MSF which can be accessed via web through a simple interface of the managing system SISFOR. The development has been based on readily available technologies which allow for rapid prototyping.

Both components, the MSF repository and the SISFOR software, are being evaluated by users to produce an improved version which not only covers basic themes such as tree species but which can be expanded to other themes such as those related to environmental concerns. Through user feedback the knowledge engineers' understanding will be improved concerning how to enhance the IRFFs descriptors and what new search terms are required. A more comprehensive system will contribute to research on sustainable development approaches through dissemination and sharing of forest and forestry knowledge.

#### C0196

Y.M. Fernández-Ordóñez. J. Soria-Ruiz. "Development of Simple Web Systems to Support Characterization of Information Resources: A Step towards Knowledge Representation and Sharing" EFITA-WCCA-CIGR Conference "Sustainable Agriculture through ICT Innovation", Turin, Italy, 24-27 June 2013.

## Sustainable Agriculture through ICT innovation

## 5. ACKNOWLEDGEMENTS

The first author acknowledges the funding support of Colegio de Posgraduados through the Research Line LPI9 *Geomatics applied to the study and management of natural resources and agrosystems*.

## 6. REFERENCES

- CABI - Centre for Agriculture and Biosciences International. 2013. *Forestry Compendium*. <http://www.cabi.org/fc/>
- Dieng-Kuntz, R. 2002. Corporate Semantic Webs. *ERCIM News* 51: 19-21.
- FAO. 2012. *Agricultural Information Management Standards - Vocabularies*. <http://aims.fao.org/>
- Maliappis, M. T. 2007. Technological Aspects of Using Agricultural Ontologies. In *Proceedings of the Conference on Metadata and Semantics Research (MTSR 2007), Ionian Academy, Corfu, Greece. CCIS Series. October.*
- Pang, T., R. Sadana, S. Hanney, Z. A. Bhutta, A.A. Hyder, A. A. and J. Simon. 2003. Knowledge for better health: a conceptual framework and foundation for health research systems. *Bulletin of the World Health Organization*, 81(11): 815-820.
- Shadbolt, N., W. Hall and T. Berners-Lee. 2006. The Semantic Web Revisited. *IEEE Intelligent Systems*, 21(3): 96-101.
- Te'eni, D., and H. Weinberger. Systems development of organizational memory: A literature survey. In H. R. Hansen, M. Bichler, and H. Mahrer, editors, *ECIS 2000: A Cyberspace Odyssey. Proceedings of the 8th European Conference on Information Systems*, Vienna, Austria, July 2000. Vienna University of Economics and Business Administration (1): 219-226.
- Uschold, M. 2003. Where are the Semantics on the Semantic Web, *AI Magazine*, 24(3), Fall: 25-36.
- Walmsley, P. 2012. *Definitive XML Schema*, 2nd Ed. Prentice Hall, Saddle River, NJ, USA.

---

**C0196**

Y.M. Fernández-Ordóñez. J. Soria-Ruiz. "Development of Simple Web Systems to Support Characterization of Information Resources: A Step towards Knowledge Representation and Sharing" EFITA-WCCA-CIGR Conference "Sustainable Agriculture through ICT Innovation", Turin, Italy, 24-27 June 2013.