> Semantic Web Use Cases and Case Studies

Case Study: The Semantic Web for the Agricultural Domain, Semantic Navigation of Food, Nutrition and Agriculture Journal

Gauri Salokhe, Margherita Sini, and Johannes Keizer, Food and Agriculture Organization of the United Nations, Italy

May 2007

General Description

Introduction

The Food and Agriculture Organization of the United Nations (FAO) leads international efforts to defeat hunger with information. FAO's activities comprise four main areas: (a) putting information within reach, (b) sharing policy expertise, (c) providing a meeting place for nations, and (d) bringing knowledge to the field.

With the advent of the Web in the 1990s, there have been enormous advances in “putting information within reach” and “bringing knowledge to the field”, however this has been at the cost of increasing complexity. The information resources available from the international agriculture arena differ in degree of coverage. For example, some resources focus on generic agriculture and food safety; some focus on sub-disciplines such as nutrition, or animal and plant health; and others focus on the type of data represented for example images or news announcements. As information is so distributed, and heterogeneous in nature, there is not a single search engine that can effectively retrieve a comprehensive set of the resources. As the current Web architecture does not allow integration of similar data from different sources, we are exploring the use of the Semantic Web.

Current Applications

Current applications in the agricultural domain that use the Semantic Web include:

- **The AGRIS Application Profile (or AGRIS Application Ontology)** consists of concepts describing document-like resource objects or DLIOs such as title, creator, publisher. It is a root ontology and controlled vocabularies provide values for the attributes of the resource, e.g., language, keyword, etc. This ontology is used to resolve the issue of semantic heterogeneity that is associated with distributed data repositories. The assignment of metadata to the repositories enables services such as: (a) search by organization, resource type, subject category, (b) issue complex queries, such as checking to see if a given author wrote any articles written within a given time period, (c) present the user with information related to the query, such as titles associated with a particular author, and co-authors. In the future, the same technique may be applied for sharing information on other types of resources, namely, events, news, projects, organizations and experts.
- **Use of Application Ontology and Thesaurus** – The AGRIS Application Ontology is used with AGROVOC, the multilingual agricultural thesaurus, to create an information retrieval application that performs query expansion (Figure 1). In contrast to the previous scenarios, where semantic relations exist only between metadata descriptors, the relationship between thesaurus terms allow for development of intelligent applications.

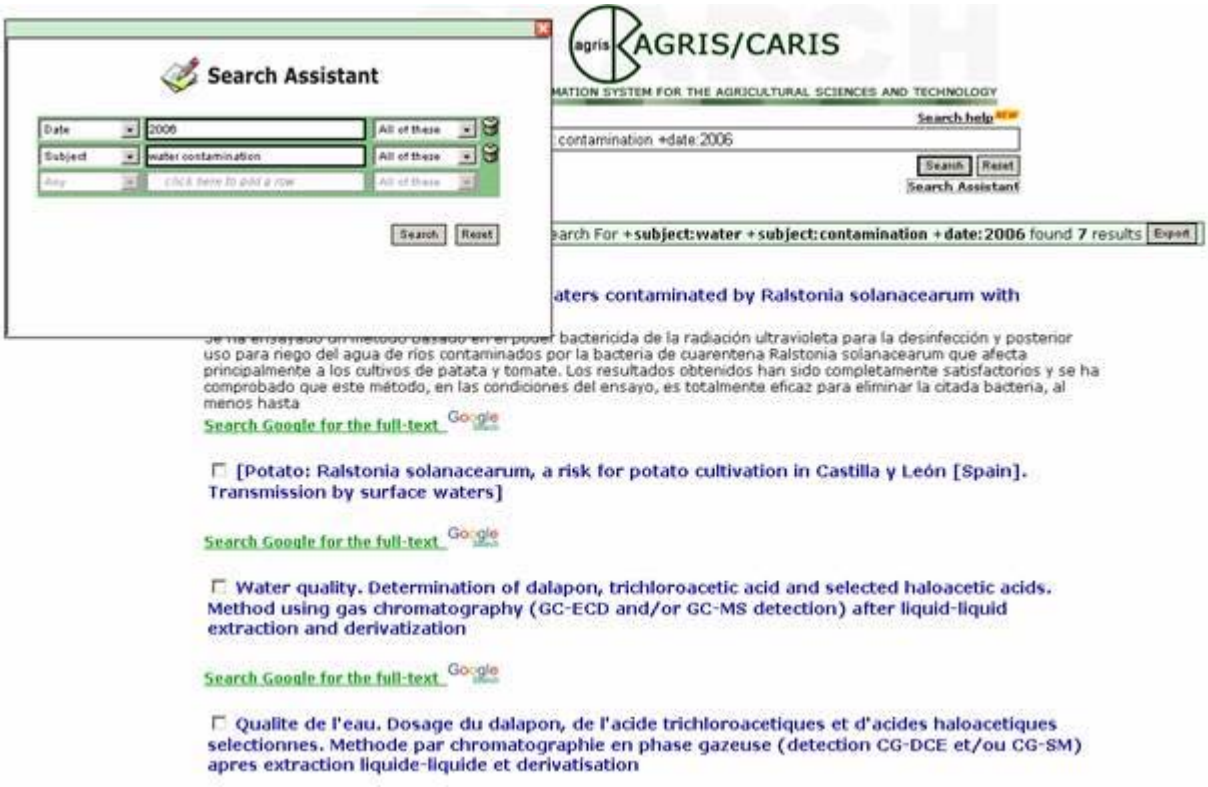

Figure 1: Searching XML files with Keywords from AGROVOC Thesaurus

Use of Ontologies to improve search results

The articles in the Food, Nutrition and Agriculture (FNA) Journal cover topics such as community nutrition, food quality and safety, nutrition assessment, nutrient requirements, food security and rural development. The full-text articles may be in English, French or Spanish. The FAO provides information in this journal about nutrition and food safety to audiences around the world for over four decades.

Metadata about each article in the FNA Journal was available in FAO's Corporate Document Repository and the FAO's bibliographical database. The two metadata sets contained slightly overlapping metadata information about each resource. Work has been undertaken to combine the metadata and to convert it to a single RDFS format. The bibliographic multilingual data was managed and stored using the Karlsruhe ontology (KAON) suite of tools.

An ontology was developed in RDF using metadata about articles from FAO's cataloguing and indexing systems. The AGRIS Application Ontology provided the backbone on which agricultural resources, namely the individual articles from the FNA Journal, could be organized and described (Figure 2). Following the creation of this ontology, instances were created taking the actual metadata merged from the FAO's Corporate Document Repository and the FAO Bibliographical Database.

A search application (see Figure 3) was created on top of the ontology and the instance data. When a user enters a phrase in the query box, the system searches all the objects in the ontology which may be lexicalized with the query string. The objects can be articles, publications, authors, keywords, subject categories, languages or geographical areas. When one or more matches are found, the user is provided with the list of objects.

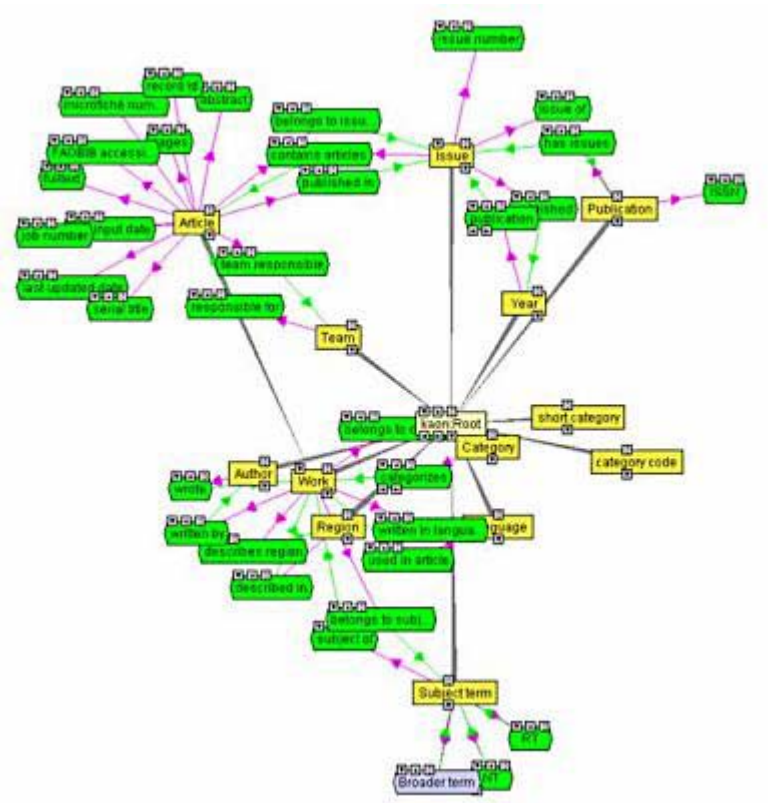

Figure 2: Application Ontology and its relationships expressed in RDF

The portal is predominantly browse-based although users can also search the metadata using a free-text search. A user is guided through the navigation of data by following the links that connect the different metadata elements, such as articles within a specific issue, authors, languages, or keyword. The user is able to get answers with a single click to questions such as “give me all articles published in this issue”, “give me all authors that wrote about this article”, “list all articles about ‘allergens’”.

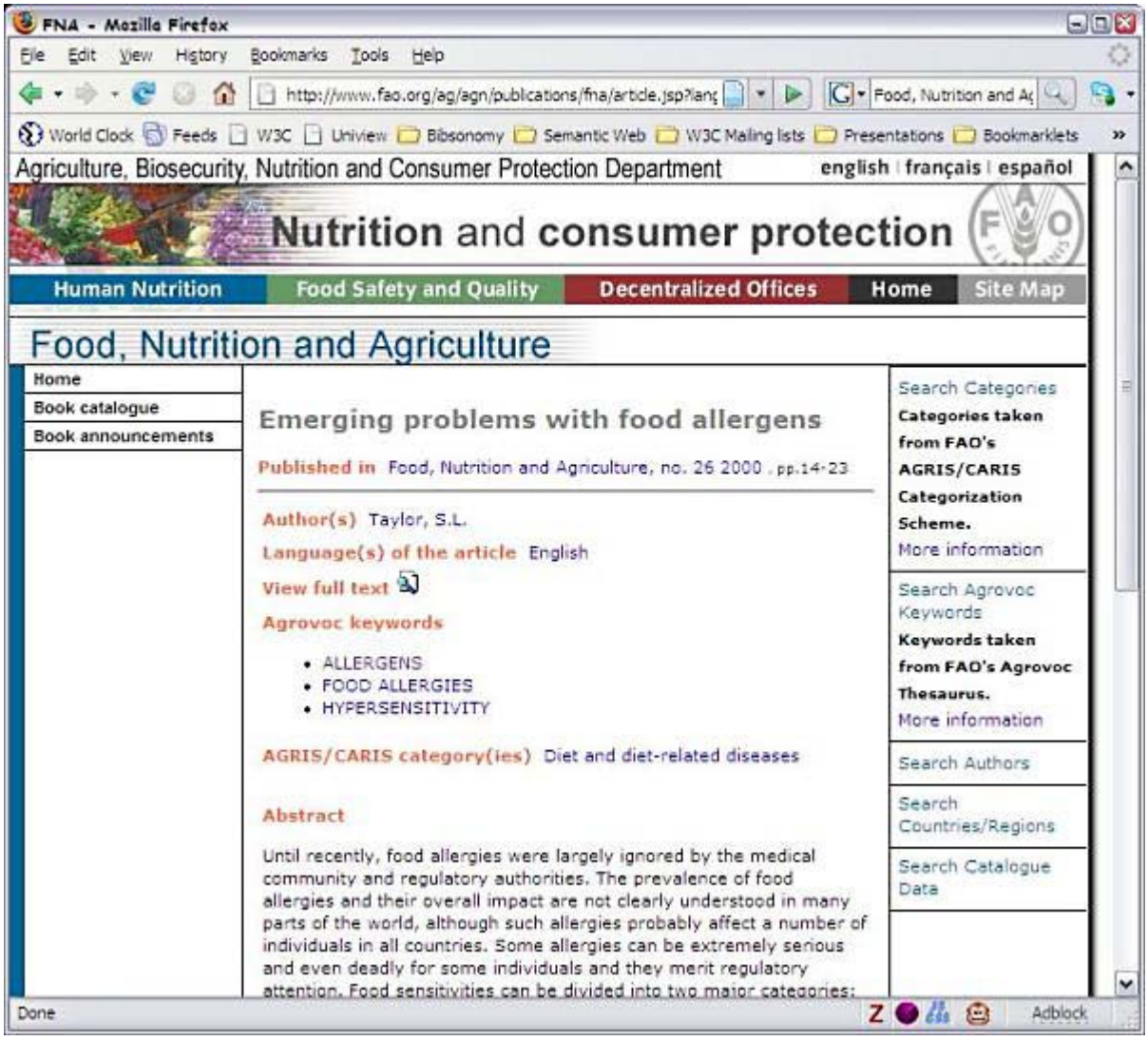

Figure 3: The browser interface to the FNA journal

Key Benefits of Using Semantic Web Technology

- One major advantage of an ontology-based system over a traditional system is in its ability to perform concept-based searches. The FNA portal allows searches for specific concepts; for example, a search for “child” or “children” will give same result even though these are two different lexicalizations of the same concept, that is one is the plural of the other.
- The portal also allows multilingual searches: “children” or “enfants” will retrieve the same result.

Future work

These steps are a starting point for further exploitation of other semantic relationships available in a bibliographic metadata record. Making use of existing semantic relationships between, for example, author and keyword, that are not normally exploited in bibliographic databases allowed for more meaningful and hence user-friendly browse experiences. The benefits of converting from RDFS to OWL are currently being explored.

A best case scenario would be a question and answer service which brings together different resource types (people, experts, projects etc.) and exploits the relationships between them to provide answers to questions such as:

- Who can I contact for more information on “water pollution in Nigeria”?

- Have there been any projects in collaboration with the FAO in Eastern Thailand?

This would require the heterogeneous data repositories are made available in RDF along with the controlled vocabularies in RDF/OWL. The access to information also should be languages independent and the tools on top should be able to exploit the multilingual semantics encoded in the RDF/OWL subject ontologies.

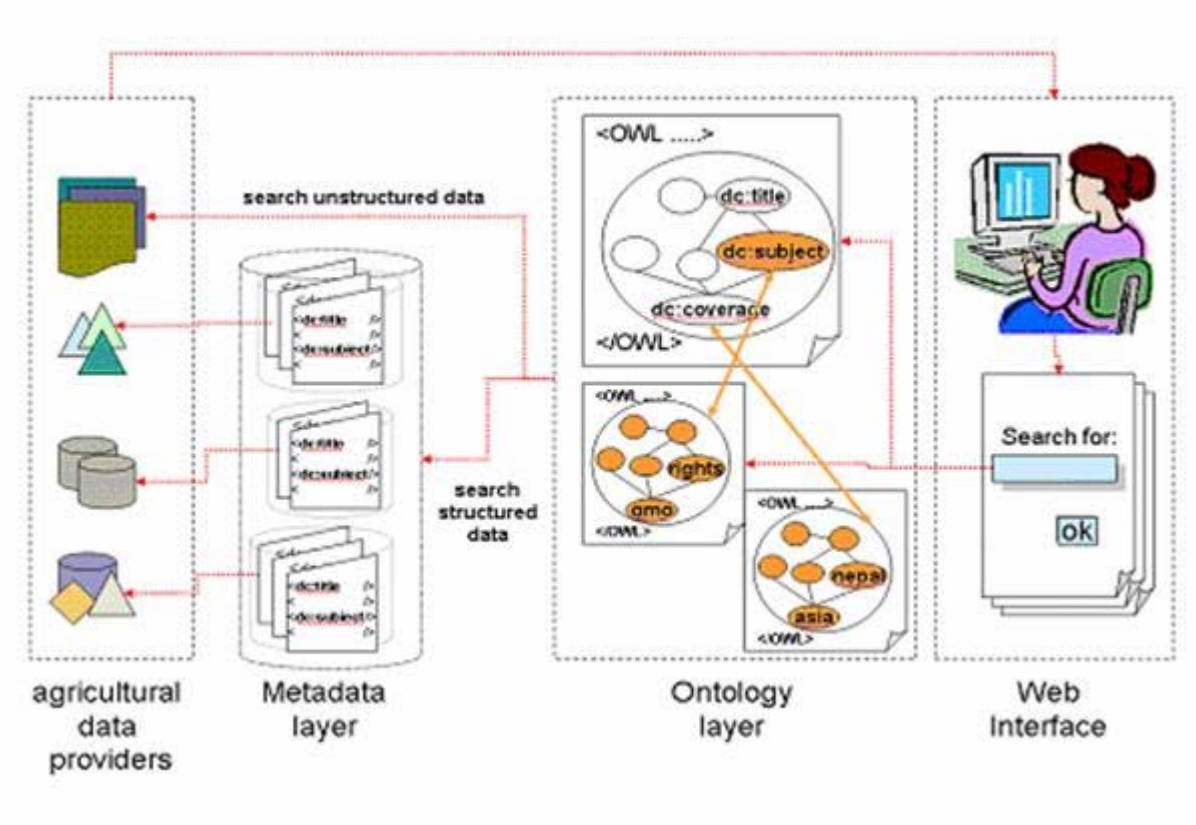

Figure 4: Seamless integration of the ontology layer, metadata layer and data repositories

© Copyright 2007, Food and Agriculture Organization of the United Nations (FAO)