Development of a methodology for guided construction of ontology by domain experts and information scientists: the case of pests in agriculture / Chaim Mograbi

Abstract

The aim of the research is to define a methodology for building an ontology by domain experts who are nonprofessional ontologists, and non-domain experts users, and to develop a convenient and simple interface designated to convert a thesaurus into an ontology, by cooperation between domain experts, ontologies professionals, and users who have no background in the domain of knowledge. While the firsts may have a unique expertise in their domain of knowledge, they usually have no formal education in ontologies. As their expertise in narrow domains other than ontologies deepens, it is reasonable to assume that they would rather not deepen in exploration of ontologies, ontologies building, logic or programming. The vocabulary used by them may consist of many homonyms, and they might call certain phenomena in various epithets. Ergo, the need to facilitate domain experts to build ontologies for their use. This study provides a friendly and intuitive graphical user interface, which allows even ontology laymen to deal with the task of building an ontology. The assumption of this research is that taxonomies and thesauri which deal with different domains do exist. The aim of the developed interface is to enable domain experts and users to link concepts from different thesauri and taxonomies by rich semantic relationships and thus to build ontologies. The study examined how and to what level of precision and extraction a group of domain non-experts can build a detailed domain ontology based on the infrastructure which had been developed by the experts.

In order to achieve the study goals, a group of ontology experts worked together with a group of domain experts. The ontology experts group tracked the ontology building process steps in order to define the domain infrastructure. The domain experts group chose domain thesauri and defined between which parts of the thesauri relationships can exist. The domain experts also filtrated concepts and advised the ontology experts about the preferred visualization, defined basic relationships and constraints, and assisted the ontology experts in gathering resources.

As a first step in building the ontology, a thesaurus and reliable sources of information relevant to the ontology theme were selected. In the next phase the upper schema was constructed. For this purpose, the domain experts were advised to cherry-pick terms from the above mentioned thesaurus and information sources. The chosen terms used as upper classes of the ontology. Then other terms were selected and served as sub

classes of each upper category. Later the domain experts chose terms from the glossary tree which they built by the choices they had made, and defined suitable cross-thesauri and cross-information sources semantic relationships, taking into account the future specific use of the ontology.

The case study of this research was implemented on a user friendly event oriented ontology for the pests and pesticides branch of agriculture, in collaboration with agricultural experts from the Agricultural Research Organization – Volcani Center. The system met two requirements:

1. Building ontologies out of thesauri;
2. Setting up semantic relationships between pairs of concepts which were selected from hierarchical concepts from chosen thesauri.

The developed graphical system is accessible at http://ramsegev-001-site1.htempurl.com/#/. To serve the system, a convenient and simple interface of converting thesauri into ontologies was developed, by cooperation between agricultural experts, ontologists, and Information Science students who had no agricultural background, but had background in building ontologies. The study examined whether a group of users who were not domain experts could expand an ontology using the developed interface and a basic ontology which was developed and built by experts. Another issue this research tried to examine was whether a group of domain experts who were not ontologists would be able to expand the ontology. Finally, the degree of users' satisfaction with the developed interface was checked.

As part of the discussions with the agricultural experts, and in response to their questions about the ontological solution, an event oriented ontology model was proposed. This model enables to link more than two ontological concepts, and to define properties for this complexed relationship. The complexity of data in the pesticides domain stems from the fact that every combination of crop-pest-pesticide is characterized differently by its properties; hence emerged a need to define a "crop therapeutic event" - an entity or one ontological class which role was to couple these three concepts. Every concept out of crops, pests and pesticides was linked to a concept of compatible type, which coupled the three concepts of crop, pest and pesticide.

As part of this study an experiment was conducted aiming to build a detailed bilingual ontology (English and Hebrew), which encoded information about allowed amounts of crop pesticides according to standards in different countries (Israel, USA, Europe). The ontology is event oriented so that all the terms in it - crop, pest and pesticide - were linked through the concept of "crop therapeutic event" that created a framework for the triple: crop-pest-chemical. The ontology integrates information from various disconnected professional web sites, which contain relevant information. 26 subject participants completed all the stages of building the ontology (out of 31 subjects who

started it), and during the experiment on the graphical system they defined 75 events, and in the course of the Web-Protégé experiment 19 events. The success rate of creating an event on the graphical system was 87%, compared to 73% on Web-Protege. This result indicates that it is easier to understand how to build an event on the graphical system compared to Web-Protégé.

This work endeavours to bridge the extant gap between the agricultural sector users' familiarity with the world of information, and the application of the technology in their professional world, by offering a user-friendly tool for developing ontologies, which can also be used as a foundation for research and development activity. The ontology is capable of sustaining harm dwindling caused by pests, while preserving public health, environment, and other living organisms which are not target for pest control. This ontology amalgamates information from various sources, thus responses to the prevailing need for information regarding pests and their treatment. It can foster the capability to provide cross-enterprises and cross-systems information, as well as adjusting it to a wide range of applications, cultures, and term. The implementation will serve all stakeholders in agricultural information and in particular the Israeli public. The proposed ontology allows concentration and sharing of variety of information from many sources, and includes terms and relationships among them in Hebrew and English.

The main research findings were that non-domain-experts users can use the developed interface to expand an ontology. The average satisfaction rate with the developed system found to be higher than that obtained from the state of the art Web-protégé that was used for comparison purposes.

System No.
002468493