Chapter 44
Effectiveness of Visualization for Information Retrieval through Ontologies with Entity Evolution:
The Impact of Ontology Modeling

Akrivi Katifori
National and Kapodistrian University of Athens, Greece

George Lepouras
University of Peloponnese, Greece

Costas Vassilakis
University of Peloponnese, Greece

Elena Torou
National and Kapodistrian University of Athens, Greece

ABSTRACT

Incorporating digital tools in the business and scientific research workflows is at the moment an on-going process, challenging and demanding as every domain has its own needs in terms of data models and information retrieval methods. The information in some domains involves entity evolution, a characteristic that introduces additional tasks, such as finding all evolution stages of an entity, and poses additional requirements for the information retrieval process. In this paper, we present a user study aiming to investigate how the different aspects of ontology modelling affect the performance and effectiveness of users regarding information retrieval tasks that are carried out using visualization methods. The results of the user study are analyzed and guidelines for ontology design are offered.

1. INTRODUCTION

In recent times, the continuing progress in data capturing, data storage and network technologies has enabled the digitization and dissemination of huge amounts of documents. The need for more effective information retrieval has led to the creation of the semantic web and personalized information management notions. These research areas take advantage of the semantic context of documents to facilitate their management. In many of the proposed solutions in this field, it is common to

DOI: 10.4018/978-1-4666-9562-7.ch044

include the use of an ontology. Consequently, the need for effective ontology visualization for design, management and browsing has arisen.

Ontologies have been identified as prominent information representation methods to facilitate information retrieval (Nagypal, 2005); (Zhuhadar, Nasraoui, & Wyatt, 2009); (Sy, et al., 2012). In this context, the modeling of the ontology may play an important role in the effectiveness of supporting the users. This is especially true in cases where the domains modeled by the ontology is to be employed contain time-evolving information, such as ontologies for Historical Archives (Torou, Katifori, Vassilakis, Lepouras, & Halatsis, 2010) and ontologies modeling the user domain (personal ontologies) (Catarci, Dix, Katifori, Lepouras, & Poggi, 2007). To this end, the overall model and ontological information, as well as the modeling of time-evolving information in particular, needs to be extensively examined. Considering existing theoretical works, time-evolving information can be accommodated either via versioning (Klein & Fensel, 2001) or the four-dimensional perdurantist approach (Welty & Fikes, 2005), according to which each entity is considered to be an event and has a start and an end point, and can be seen as a "space-time worm", with the slices of the worm being temporal parts (time slices) of the entity. The existence of time-evolving information introduces new tasks regarding the information retrieval process: in addition to the IR tasks they perform in "ordinary" repositories (i.e. repositories not involving evolution), need to be able to

A. Locate and mentally link together "worm slices" corresponding to different phases the same entity and
B. Follow links between "worm slices", exploring entity history.

The current work presents the results of an experiment focused on investigating the impact of ontology modeling on information retrieval activities carried out using visualizations, through ontologies with entity evolution. In the experiment, users were requested to perform a set of information retrieval tasks over an ontology employing visualizations, using two different ontology models. This experiment offered insight on how different modeling decisions affect user effectiveness, and our results can serve as guidelines to ontology designers, so as to produce ontologies on top of which information retrieval through visual means can be more effective. In order to ensure that the results are not biased by the visualization method employed, four different visualization methods were used.

The following sections provide useful definitions and related work, followed by a presentation of the ontology and the visualization methods used in the experiment. Afterwards, the evaluation process is described and the results are presented and discussed. Finally, the last section concludes the paper and outlines future work.

2. BACKGROUND AND RELATED WORK

According to (Gruber, 1993), an ontology is an explicit specification of a conceptualization. The term "conceptualization" is defined as an abstract, simplified view of the world that needs to be represented for some purpose. It contains the objects, concepts and other entities that are presumed to exist in some area of interest and the relations that hold between them. The term "ontology" is borrowed from philosophy, where an ontology is a systematic account of existence.

As defined in (Noy & McGuinness, 2001), an ontology is a formal explicit description of a domain, consisting of classes, which are the concepts found in the domain (also called entities). Classes form a specialization/generalization hierarchy, with each class being connected to its parent(s) through *is-a* or *inheritance* links; classes may have one or more parents. Each class has properties (also called slots, roles or attributes)