Chapter 2.16
Organizational Semantic Webs

Jean-Yves Fortier
University of Picardie Jules Verne, France

Gilles Kassel
University of Picardie Jules Verne, France

INTRODUCTION

The main subject tackled in this article is the use of knowledge technologies to develop corporate memories or (stated more generally) “organizational memories” (OMs) (Dieng, Corby, Giboin, & Ribièrê, 1999).

At the end of the 1990s, AI technologies, in general, and knowledge technologies, in particular, were recognized as pertinent and promising tools (in addition to information technologies) for the design of OMs (Buckingham Shum, 1997; O’Leary, 1998; Milton, Shadbolt, Cottam, & Hammersley, 1999). These very diverse technologies (concepts, methods, and tools) have been conceived to assist knowledge acquisition, modeling, and discovery, as well as the development of knowledge-based systems (Studer, Benjamins, & Fensel, 1998). In this article, we focus on knowledge modeling and formalization techniques, since our prime interest is the preservation of knowledge within OMs and its impact on the exploitation of this knowledge.

In practice, the use of these technologies generates two complementary proposals: (1) the formalization of a part of knowledge to be preserved, which means considering hybrid memories in terms of specification modes (formal, semi-formal, and informal); and (2) the introduction of a formal ontology of the domain in question, in order to facilitate the expression, comprehension, and access to capitalized knowledge. Formalization thus relates to both (1) knowledge (as propositional knowledge) and (2) meaning (as conceptual knowledge).

Regarding the balance between formal and informal specification, a broad spectrum of OM architectures have been proposed, ranging from informal annotation of formal knowledge bases (Euzénat, 1996) to the formal annotation of informal documents (Buckingham Shum, Motta, & Domingue, 2000). It should be noted that these extremes (i.e., the development of a text-documented knowledge base and the publication of scientific articles on the Web, respectively) correspond to atypical OM applications.
The knowledge technologies used in 2004 to develop OMs are generally those of the Semantic Web, where languages like OWL (Antoniou & van Harmelen, 2004) allow us to exchange knowledge bases on the Web. One particular asset of OWL is its ability to offering several dialects with different expressive powers—the choice of the dialect depending on the specific application in question.

A review of the state of the art (cf. section 2) shows that current OM architectures rely on “lightweight” knowledge models, corresponding to formal annotations of textual resources. These approaches focus on document “enrichment” (Motta, Buckingham Shum, & Domingue, 2000), since the knowledge models and ontologies are used to facilitate access to textual resources and the dissemination of the latter to interested users.

In contrast to these initiatives (or rather by extending them), we recommend giving more importance to formalization, by going back to Buckingham Shum’s original proposal (1997) of formalizing a part of the knowledge to be capitalized. Such an approach requires us to improve the knowledge technologies used, in order to make it possible to apprehend and reason on the contents of the resources independently of the specification modes (cf. sections 3, 4, and 5).

**BACKGROUND**

Our current work concerns the conception and development of organizational Semantic Webs (OSWs), that is, OMs whose implementation exploits Semantic Web technologies. The evolution of the Web into a Semantic Web is currently the subject of numerous research programs (Berners-Lee, Hendler, & Lasilla, 2001). The principal aim is to enable software agents to exploit the contents of textual resources present on the Web so that users can ultimately be relieved of certain information searching and combination tasks (Fensel, Wahlster, Lieberman, & Hendler, 2003). The developed technologies apply as much to the Web as a whole as to OSWs in particular.

Current OSW architectures rely on the coupling of a collection of textual resources with formal resources, the latter also being qualified as “semantic” resources. Of these, one can distinguish annotations of textual resources or “metadata” (which express knowledge about textual resources) (Handschiuh & Staab, 2003) on one hand, and ontologies (which stipulate the meaning of the terms used to express the textual resources and the metadata) (Davies, Fensel, & van Harmelen, 2003; Abecker & van Elst, 2004) on the other hand. Again, one finds a distinction between knowledge and meaning. In terms of the contribution of these semantic resources, various approaches are being explored. They may thus be used for:

- navigating within a network of annotations, in order to help discover documents and apprehend their contents (Buckingham Shum et al., 2000)
- furnishing the user with the documents likely to interest him or her, by taking into account his or her centers of interest expressed in terms of ontological concepts (Davies, Duke, & Sure, 2003; Middleton, De Roure, & Shadbolt, 2004; Uschold et al., 2003)
- ranking answers to queries by taking into account the annotations’ contents (Stojanovic, Studer, & Stojanovic, 2003) and/or the memory’s uses such as previous consultations

The study of these architectures shows that they force formal resources into a precise role: constituting an index for textual resources. This type of coupling can be qualified as “weak,” to the extent that the only aim of these formal resources is to facilitate the exploitation (access, dissemination) of the textual resources — the capitalized knowledge being only present in the latter. When a