23

Semantic Web Technologies for the Adaptive Web

Peter Dolog\textsuperscript{1} and Wolfgang Nejdl\textsuperscript{2}

\textsuperscript{1} Department of Computer Science, Aalborg University, Fredrik Bajers Vej 7E, DK-9220 Aalborg, Denmark, dolog@cs.aau.dk
http://www.cs.aau.dk/~dolog
\textsuperscript{2} L3S Research Center, University of Hannover Appelstrasse 9A, 30167 Hannover, Germany nejdl@l3s.de

Abstract. Ontologies and reasoning are the key terms brought into focus by the semantic web community. Formal representation of ontologies in a common data model on the web can be taken as a foundation for adaptive web technologies as well. This chapter describes how ontologies shared on the semantic web provide conceptualization for the links which are a main vehicle to access information on the web. The subject domain ontologies serve as constraints for generating only those links which are relevant for the domain a user is currently interested in. Furthermore, user model ontologies provide additional means for deciding which links to show, annotate, hide, generate, and reorder. The semantic web technologies provide means to formalize the domain ontologies and metadata created from them. The formalization enables reasoning for personalization decisions. This chapter describes which components are crucial to be formalized by the semantic web ontologies for adaptive web. We use examples from an eLearning domain to illustrate the principles which are broadly applicable to any information domain on the web.

23.1 Introduction

Information access on the web is realized through the hypertext paradigm. Hypertext interlinks related pieces of information (pages) and allows the user to browse through the information space. The links are provided either explicitly, encoded by authors of the pages, or they are generated automatically, for example based on the results of a query.

Personalized information access in this context is concerned with user-centered bias of the hyperlinks to better support the current user context. Generating links automatically, taking user profiles into account, is a very attractive option but creates challenges as well. According to [5], adaptive web systems extend the adaptive navigation and presentation techniques from closed corpus adaptive hypermedia to the open corpus information resources available on the web and thus supporting personalized access on the web. In this chapter we discuss solutions based on semantic web techniques to
realize personalized link generation. Key aspects of this solution are ontologies and reasoning techniques. Ontologies represent shared and agreed upon conceptual models in a domain, which describe the main concepts of the domain and their relationships. Ontologies can thus serve as reference models for generating links in this domain, and represent hypertext, content and user information. Reasoning techniques can then work on metadata based on these ontologies, and generate links based on content, user context and user background.

As discussed in Chapter 8 [4] of this book, hypertext is a collection of text fragments interconnected by active links, used to access the information fragments addressed by them. Research in the hypertext community has concentrated on how to improve navigation in hypertext systems. The hypertext community has been concerned with several ways of browsing [18, 17]. Information retrieval concepts have been studied together with hypertext concepts [1, 35].

We can distinguish between two link concepts in hypertext: links maintained within the text (embedded links) and links maintained externally to the text as first class entities. Hypertext which utilizes the first view is often denoted as a closed hypertext, the latter one is denoted as an open hypertext [29]. Hypertext is used also in connection with hypermedia, i.e. text is augmented with other media types like pictures, video or audio.

The advantage of the embedded links is that they are bound directly to the information which utilizes the links to access related information. The advantage of the second kind of links is that we can maintain and exchange links which link information in different contexts and possibly for different users, thus providing a more flexible solution ready for personalized access. This separation of text/media items from link structures is now widely accepted in hypermedia systems [18, 17].

Information retrieval systems (especially the content-based ones) rely on index structures with terms from the documents they index. The index structures are used for making retrieval more efficient (see Chapter 10 [31] of this book for more details on content-based recommender systems). Advanced information retrieval systems maintain additional relationships between the index entries. Such structures can be seen as document models which are based on conceptual modeling approaches, semantic net approaches, Bayesian network approaches and so on (see Chapter 5 [6] of this book on document modeling). Open hypermedia research deals with links which are external to the content items. Such links can be seen as indexes of the content helping to browse and navigate the content items they index and map in an efficient way. Therefore, such conceptual structures are related to the document models and information retrieval approaches.

A notion of conceptual open hypermedia has been developed [8, 26, 33, 16]. Conceptual open hypermedia deals with knowledge representation of access structures to content items for particular context from a browsing point of view. Current semantic web technologies are very close to this notion of hypermedia, i.e. they can be used to model and represent such link structures and related objects for reasoning, querying, and processing purposes.

Though the domain ontologies are useful to generate links suited for a particular domain context, with huge corpuses it might result in too many links. Knowledge about