Chapter VII
A Semi-Automatic Semantic Annotation and Authoring Tool for a Library Help Desk Service

Antti Vehviläinen
Helsinki University of Technology (TKK), Finland

Eero Hyvönen
Helsinki University of Technology (TKK) and University of Helsinki, Finland

Olli Alm
University of Helsinki and Helsinki University of Technology (TKK), Finland

INTRODUCTION

This chapter discusses how knowledge technologies can be utilized in creating help desk services on the Semantic Web. To ease the content indexer's work, we propose semi-automatic semantic annotation of natural language text for annotating question-answer (QA) pairs, and case-based reasoning techniques for finding similar questions. To provide answers matching the content indexer's and end-user's information needs, methods for combining case-based reasoning with semantic search, linking, and authoring are proposed. We integrate different data sources by using large ontologies. Techniques to utilize these sources in authoring answers are suggested. A prototype implementation of a real life ontology-based help desk application, based on an existing national library help desk service in Finland, is presented as a proof of concept.

Help Desk Services

Companies and public organizations widely use help desk services to solve problems for their customers. The classic example of a help desk

service is a call center, where support persons answer questions by phone or by e-mail. As help desk services are being transferred to the Web, it is more and more common that the customers have also the possibility to solve their problems by themselves by using the knowledge and content accumulated at the service, without contacting a support person directly (Foo, Hui, Leong, & Liu, 2000). A simple approach, for example, is to publish Frequently Asked Questions (FAQ) lists on the Web. The option to use a simple and fast question-answer (QA) self-service is appreciated not only by the customers, but by the authors of the answers, too. Their time is saved if the QA service can automatically provide an answer to the customer. Furthermore, the author can use the accumulated QA knowledge of the service by herself, which helps in authoring the answers and improves the quality of the answers.

Research Problem

In the following, we consider application of Semantic Web technologies to QA help desk services, where the database of the service is composed of previously answered questions, that is, QA pairs. In such a service, the end-user has a question in mind, and the service has to answer to it. This involves two major tasks:

1. **Finding relevant previous answers.** A search method is needed to find the already answered relevant QA pairs from the repository.
2. **Authoring a new answer.** An existing QA pair may satisfy the customer's information need, but usually some kind of adaptation of the old answer case is needed. Usually answers are created and modified manually by a human editor.

The research problem is to investigate how to support semi-automatic answer authoring in a QA help desk service. Our methodology is to use Semantic Web technologies in content annotation, in utilizing the QA repository, and in integrating information available online on the Web with the authoring process and the answers.

The term *indexing* will usually refer to the old, existing way of doing indexing where *index terms* are just strings without an ontological reference. The term *annotation* will refer to the new way of using *annotation concepts* that have an ontological reference as content descriptions.

The Existing Service

The research is based on a real life case study: we use the data set of the operational *Ask a librarian* service[1] offered nationally in Finland by the editors of the Libraries.fi[2] portal. In this service the clients can send questions to a virtual librarian via e-mail, and a librarian of the service provides an answer within three working days. Some of the questions that the clients send are simple and the librarian can answer them straight away. These include questions about the opening times of a library, how to make an interlibrary loan, and so forth. However, most of the questions require that the librarian uses more time to investigate the subject of the question. For example, following questions have been asked:

I'm wondering where I could find information about studies of the library and information science?
I'm giving a presentation about Nokia. Where could I find helpful information?

Answers to these questions span typically a few paragraphs of text and contain some links to useful Web sites. The librarians report that on average they use working time from half an hour to an hour to compose such an answer.

Each QA pair in the service has been indexed using the YSA thesaurus[3] of some 23,000 common