

A Virtual Community Library: SICS Digital Library Infrastructure Project

Andreas Rasmusson, Tomas Olsson and Preben Hansen

Swedish Institute of Computer Science (SICS), Box 1263, SE-164 29 Kista, Sweden
{ara, tol, preben}@sics.se

1 Introduction

In this project¹, we aim to create an agent-based digital library architecture for a Virtual Community Library (VCL) where each user has a personal library and, at the same time, is part of a larger community. The community is dynamically composed of the users' personal libraries and, through intermediators, other digital libraries.

We want to stress the fact that the users participate in a large dynamic decentralised community where they continually interact with each other. Being a part of a community means that each user can benefit from the work put into the other libraries. For example, by obtaining documents through search queries or recommendations using social filtering, but also by getting help to organise the personal library.

In the VCL, we try to combine the best aspects of the WWW, the library and the personal library. For example, ease to publish documents, personal information space, decentralised control of the document collection and ability to search for documents.

We have currently implemented two prototypes of the system, one for the personal library and one for visualising the information spread between the users.

2 Agent Architecture

The foundation for the VCL is an agent architecture where the users are represented by self-interested agents. This is an open-ended knowledge system, which supports creation, inferring, manipulation and sharing of knowledge about information objects ("metadata"), and supports (enables automatisations of) interaction between agents pertaining to the information and knowledge management related business processes.

Interaction (compatibility) with other systems in a number of formats and protocols will be investigated, such as Z39.50, MARC, DIENST, Dublin Core, BibTeX, etc. [1], [2]. Many existing bibliographic formats are ambiguous or limited in the

¹ See <http://www.sics.se/isl/diglib>

way they represent knowledge. This introduces difficulties when the information is to be used in machine-machine conversations (in contrast to human-machine). In our architecture, emphasis is put on handling translations between partially incomplete, incoherent or incrementally developed ontologies and particular collections of knowledge.

The work is to be based on software-agent research and platforms developed in the SICS Intelligent Systems Laboratory.

3 Personal Library

A user is represented by a personal library agent. By emphasising that the information put into it is relevant to the individual users we address the problem that users will rarely bother to register information they themselves have no interest in. Classification of information by end-users usually introduces the problem of having non-librarians doing classification according to the best of their own knowledge. This makes it harder to develop usage-conventions in classifications. Hence, we still need librarians that classify larger collections to get homogeneous usage-conventions.

Within the information architecture, no library has any special status as keeping complete, correct or authoritative information. Each digital library is autonomous and free to negotiate which other libraries it depends on to find new information. However, large cohesive collections (institutions) should still be "influential" in defining authoritative descriptions of objects since they are more often consulted for information than a single-user personal library.

4 Information Dissemination and Retrieval

By seeing our digital library not just as a static entity but as a dynamic virtual community we strive to get properties of real communities. In real communities, information is spread in a number of ways, such as by looking in somebody's personal library, asking or telling other persons about something or searching in a database. The control of the information flow is decentralised to the individuals of the community and they can retrieve information both actively (by asking or looking) and passively (by being told). We implement active retrieval as explicitly formed search queries and we implement passive retrieval as decentralised social filtering based on trust [3], which is built on the interactions between personal recommender agents performing content-based filtering. Hence, we try to combine the advantages of both filtering techniques.

References

1. Library of Congress: Metadata, Dublin Core and USMARC: A Review of Current Efforts. MARBI Discussion Paper no. 99, Library of Congress, January 21 (1997)
2. Heery, R.: Review of Metadata Formats. In: Program, Vol. 30, No. 4 (1996) 345-373
3. Olsson, T.: Decentralised Social Filtering based on Trust. In: The Working Notes of the AAAI-98 Recommender Systems Workshop, Madison, Wisconsin (1998)