Digital Libraries: Did You Ask the Users?

Tiziana Catarci
Dipartimento di Informatica e Sistemistica
Università di Roma "La Sapienza"
Via Salaria 113, 00198 Roma, Italy
catarci@dis.uniroma1.it

Digital Libraries basically store materials in electronic format to be then accessed by different kinds of users, very often connected via the Internet. At this purpose, it is fundamental to develop both the infrastructure and the user interfaces. Such interfaces must support a range of functions including query formulation, presentation of retrieved information, feedback and browsing. Unfortunately, user interfaces encountered in the digital libraries, archives, and electronic journal editions that have been considered, almost look the same. They allow the user to access catalogs by submitting queries through HTML fill-in forms, and in most cases offer the choice between "simple" and "advanced" query modes. The bibliographic search seems to be usually considered a quite easy task, which does not deserve very sophisticated user interfaces.

During the last years, digital library systems have not made many efforts to solve the user-interaction problems. Only recently, new projects (e.g. University of Stanford\(^1\) and University of Michigan\(^2\) DL Projects) are developing a more complex model of information-seeking tasks. Display of information, visualization of, and navigation through large information collections, as well as linkages to information manipulation/analysis tools and structured thesauri can be identified as key areas for research.

The necessity for a more comprehensive understanding of user needs, objectives, and behavior in employing digital library systems is stressed repeatedly in new projects as the basis for designing effective systems. However, not many such projects seem to follow a rigorous rigorous "user-centered" design approach, with a strong involvement of a key library user: the librarians.

A strong cooperation with librarians from all over Europe is what we have experimented in the LAURIN project\(^3\) which involves seventeen participants from several countries, including a large group of libraries. LAURIN has two major goals: 1. To set up a network of digitized newspaper clipping archives that can be accessed through the Internet in a centralized fashion, for searching and retrieving clippings. 2. To provide a generic model to be used by individual libraries for scanning, digitizing, storing, and indexing newspaper clippings, and making them accessible via the LAURIN network.

Concerning objective 1, the LAURIN interface offers, besides traditional keyword based search methods, also the possibility of browsing the clipping collection by argument. Such an activity is supported by the use of an integrated multilingual Thesaurus, which plays a central role in the LAURIN system. The user sees a unified search space and therefore she can ignore the existence of different information sources, i.e., libraries (however, she could also select a library on demand). Requests can be formulated in any of the languages supported by the system.

Concerning objective 2, the integrated Thesaurus system will support librarians in indexing and handling the clippings. The LAURIN project had the opportunity of exploiting the presence of several "extremely expert" users, i.e., librarians, plus many casual users, i.e., people simply accessing the participating libraries. We learned the most from librarians. For instance, they have the very precise idea of what a digital library is, and do not accept something different from computer scientists. From a librarian point of view, a digital library is very different from an XML repository! Also, librarians are very familiar with classifications, thesauri and taxonomies. They used to have their own classifications for many years and do expect something "better" from information technology, but very often this does not seem to be the case.

Among other things, librarians involved in LAURIN have stressed the importance of having an indexing system and especially a thesaurus reflecting both their requirements and the needs of people who want to access the clipping archives to retrieve information of interest. Up to now the lack of structured thesauri, supporting a semantic classification of clippings, has prevented final users from accessing the clipping libraries themselves. What usually happens (at least in the many European libraries we have analyzed in LAURIN) is that the user asks a vague query to a "human interface" (i.e., the librarian) and she first tries to refine the query and then searches for the clipping potentially matching the user's interests in the archive. In order to realize an Internet-based service which aims at replicating at least the efficiency (even if it could never get all other qualities of a human-human interaction) of the librarian-mediated retrieval, the LAURIN project concentrated on two crucial components of any interactive digital library system, namely the multilingual thesaurus and the visual (indexing and retrieval) interfaces.

\(^1\)http://www-diglib.stanford.edu/diglib/
\(^2\)http://http2.sils.umich.edu/UNDL/
\(^3\)Libraries and Archives Collecting Newspaper Clippings Unified for their Integration into Networks (project). LAURIN is an EU-funded Project Telematics Program. Libraries Project LB-5629/A.
http://laurin.ubk.ac.at/