DEVELOPING A EUROPEAN DIGITAL LIBRARY FOR ECONOMICS: THE DECOMATE II PROJECT

Thomas Place

Access to the content of electronic journals can be provided in numerous ways, depending on decisions about storage and interface systems. Decomate II has developed a model for economics literature based on local storage and a uniform interface.

Thomas Place is Deputy Librarian at Tilburg University, Library Systems and Development, PO Box 90153, 5000 LE Tilburg, The Netherlands
E-mail: T.W.Place@kub.nl

Introduction

The number of scientific publishers who make their journals available in electronic form is rapidly growing. Journal users in the university community are keen to obtain access to these electronic journals. There are various ways in which access to the electronic journals can be provided to users. Within the European project entitled Decomate II a very specific model for providing access has been developed. This paper starts with an overview of the ways in which access to full text electronic journals can be provided. Next, the model developed in the Decomate II project is described in some detail.

Background information about the Decomate II project

The Decomate II project[1] is funded by the Telematics Applications Programme of the European Commission. The project officially started in February 1998 and will finish in July 2000. The project partners are the libraries of Tilburg University (The Netherlands), Universitat Autónoma de Barcelona (Spain), London School of Economics and Political Science (United Kingdom), European University Institute (Italy) and SilverPlatter Information Ltd (United Kingdom).

The project aims at creating a European Digital Library for economics with mutual access to the heterogeneous distributed and pooled digital resources of the consortium members in the field of economics. Emphasis is on the disclosure of the full text of copyrighted and non-copyrighted material. It will result in a visible example of interconnected library services, integrating various functions to provide a full-scale virtual library service to end-users.

The Decomate II project is based on the results of the successful Decomate I project. This project has resulted in an end-user service that provides users with access to copyright materials. The

Decomate II project builds on this by aiming to provide end-users with access to heterogeneous materials (i.e. copyright and non-copyright materials of different types and formats) in a distributed network (i.e. users can access material in any of the participating libraries) through a single, uniform user interface.

Publishers, libraries and electronic journals

More and more scientific publishers make their journals available in electronic form. Within the Decomate II project a list of 118 key journal titles in the field of economics has been constructed. This list contains no titles from either Elsevier Science Publishers or Kluwer Academic Publishers as the partners in the Decomate II project already have electronic access to the journals of these two publishers. It is estimated that roughly 50% of the journal titles included in the key list of 118 titles are available electronically. If the Elsevier and Kluwer journals are also taken into account, it can be concluded that a large majority of the main journals in economics are available electronically.

Different publishers have different policies when it comes to providing access to electronic journals. Some publishers are willing to provide libraries with both bibliographic information and full text for local storage. This is done either by regularly sending CD-ROMs to the library or by making the material available via an FTP archive from which a library can download the files. Elsevier Science is one of the publishers who deliver electronic full text in this way; Elsevier developed this method in their well-known Tulip project[2]. With this model, it is up to the library to build a service that can make the information available to its end-users.

There is also the model in which the publisher supplies only bibliographic information to the library. The bibliographic information has to be included in a local database maintained by the library. Access to the full text for the end-user is provided via the publisher's Internet site in this model.

Some publishers have their own sites at which end-users can access both bibliographic information and full text; examples are Science Direct of Elsevier Science, Wiley InterScience, Ideal of Academic Press and Springer Link.

A last possible way to provide access to electronic journals is through the Web site of an intermediary; examples are SwetsNet of Swets & Zeitlinger and the First Search Electronic Collection Online of OCLC. In this model, neither bibliographic data nor full text articles are stored locally by the library.

Of course, all access to the material of the publishers comes at a price. For libraries, the first model is the most costly. In this model they have to invest in hardware and software; also, they need specialised staff to maintain the service. The model may be costly, but it does offer a number of obvious advantages to the end-user. We will discuss the advantages in some detail below. Because of the advantages involved, the Decomate II project implemented this model. The project will pay special attention to the costs involved in setting up and running the Decomate system. A report on this topic will become available in due course.

Publishers have their policies when it comes to providing access to electronic journals, but so should libraries. After all, substantial costs are often involved in providing end-user access to electronic journals. Libraries have at least the following options available for providing end-user access to electronic journals. First, a library can decide to include on its own Web site links to the sites of the publishers and intermediaries. This can be done by making a list of the journal titles to which the library subscribes and to which it has electronic access. Clicking a journal title brings the end-user to the site with the full text of the journal. A drawback of such a policy is that it is impossible for a user to search the entire electronic journal collection for relevant articles; searches can only be performed individually for each publisher's site. This disadvantage can be slightly overcome when searches are carried out via the site of an intermediary, as such sites integrate the journals of several, but not all, publishers.

Another way to provide end-users with access to electronic journals is by including links in the library catalogue. These links are added to the journal description. In most cases, a link brings the user to the list of issues of the journal that are electronically available on a site of a publisher or intermediary. Access is thus provided at the level of the journal. This solution is of course only

possible when the OPAC of a library has a Web interface.

Libraries can also decide to provide access to electronic full text at the article level. The link to the full text is then provided in the bibliographic description of an article. There are two possibilities here. One is that a library uses the linking technology of, eg SilverPlatter with their *SilverLinker* technology or the Institute for Scientific Information with *ISI Links*. In both cases, the bibliographical and citation databases have to be implemented with the Web solutions offered by these companies (respectively, *WebSpirs* and *Web of Science*). The secondary information offered by SilverPlatter and ISI is linked to the primary information stored on the Web sites of the publishers and intermediaries.

The second possibility is that a library has its own article database that corresponds with the journal collection of the library. Such an article database can be seen as an extension of the catalogue: it provides an inventory of the journal collection owned by the library. This is the option implemented by the libraries working in the Decomate project.

The Decomate model: starting point

The main aim of the Decomate II project is to develop a European Digital Library for economics. The system must allow end-users to simultaneously search multiple databases - these databases may contain heterogeneous material. All searching is done via one uniform user interface. Giving access through one single, uniform user interface to a range of databases and resources is the main goal of the Decomate project. The project will result in a working demonstrator for the field of economics.

In order to realise the aim of the project a number of decisions have been made. These decisions together constitute the Decomate model. Within the Decomate project, a conscious decision has been made to provide access to electronic access at the level of the article. In the Decomate model, the bibliographic information including abstracts coming from various sources is stored locally in one database. This information is accessed by the end-user through the Decomate Web interface. For this, each library runs its own Decomate server. Storage of the electronic full text can be local or remote.

The reason why so much emphasis is placed on the local storage of bibliographic data is the following. The bibliographic data in question provide a description of the library's own journal collection. A local reference database of journal articles articulates the collection development policy of the library. The collection is offered to the end-users as the local digital library, as an integrated whole. Just as the book collection is covered by the OPAC, the journal collection of a library is covered at the level of the articles in the local digital library. High quality requirements apply to the article descriptions. Examples of such requirements are the presence of keywords or abstracts in bibliographic records, links to the journal holdings of the library (which requires integration with the catalogue) and the integration with other local and remote information. Other requirements arise from the decision to offer users an electronic current awareness service and other tailor-made services. Local control over the bibliographic data allows for more sophisticated and discipline-oriented access to the collection, e.g. via a journal browser or integration with a thesaurus. With local storage and local indexing of the information, a high level of performance (in terms of response times) is obtained. It also provides better facilities for collecting data about the usage of the electronic collection. User studies are an important part of the Decomate project.

Another important aspect of the Decomate model is the use of a single, uniform user interface for the different resources made available to the end-user. Apart from database specific features, there is no difference between searching a database with references to electronic full text, searching the OPAC or searching a bibliographical database like Econlit or ABI/Inform. Simultaneous searching of multiple databases is possible. When the difference between the databases is insignificant to the user the databases can be treated by the system as one virtual database. In addition, access to databases at remote locations will be possible in the Decomate project. The magic word here is Z39.50.

The Decomate model: technical details

The Decomate system will offer end-users two scenarios for searching for information. The first scenario, in a sense the 'simpler' or traditional

scenario, is particularly suitable to more advanced users. In this scenario the user is presented with a search screen in which a single search term or a complex Boolean query can be entered. The search results are presented to the user in the form of a list. The user can browse this list. By clicking on a title in this list, the full bibliographic description is obtained. If the article is available electronically, a link to the full text is included. If the article is not available electronically, the user has the option of ordering the document. Different document ordering systems will be supported.

In itself, the search facility within the Decomate system is rather standard. However, the project has defined a work package for the development of advanced access to federated digital libraries. A special graphical tool is developed; the *concept browser* that allows navigation through the concept space of a discipline. Such a conceptual network consists of several loosely coupled vocabularies and thesauri that are implemented as Z39.50 databases. For the digital library for economics, several vocabularies and thesauri are available: the commonly used JEL classification codes of the *Journal of Economic Literature* and thesauri maintained by the project partners. The links between related concepts are graphically displayed and the links can be followed from concept to concept. Relevant concepts can be marked and irrelevant concepts can be crossed out. While navigating through the conceptual space, the information need of the user becomes known to the system. After the user has finished navigating the conceptual space, the system generates an optimised query that is passed to relevant databases. In principle, the advanced access system is multi-lingual. The results from the different databases are merged and ranked in order of relevance. The documents can be depicted in the conceptual space.

The heart of the Decomate system is the *Broker* that is accessed by the Web browser of the end-user. The Broker mediates between the user and information services that are outside the Decomate system. The Broker is assisted by so
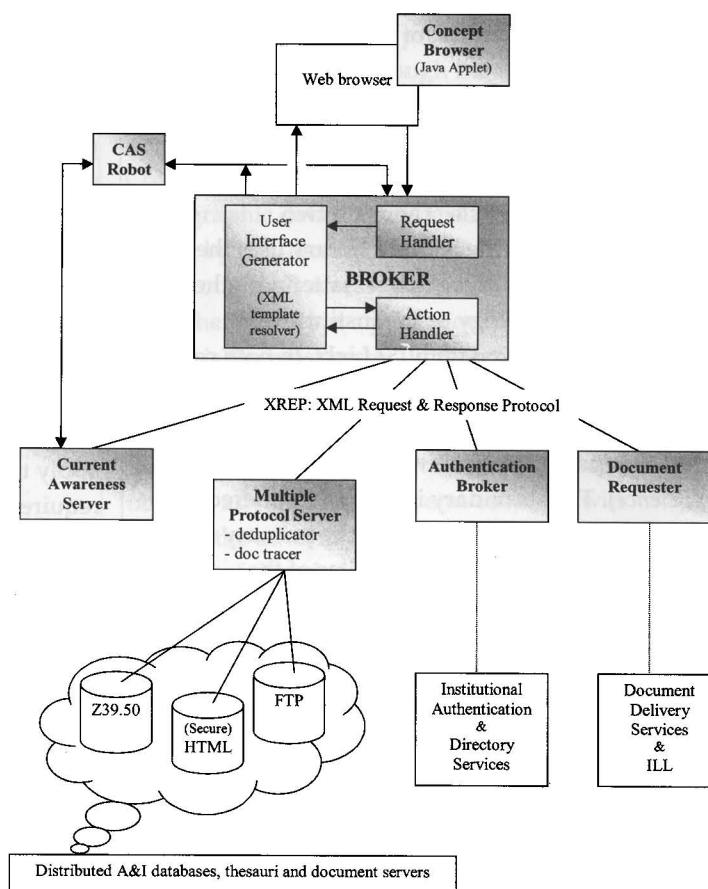

Figure 1. The Decomate model

called Broker services that connect the Decomate system to external servers. In the project, the following Broker services will be implemented: Current Awareness Server, the Multiple Protocol Server, the Authentication broker and the Document Requester. The Broker services can run as independent servers. Communication between the Broker and its services makes use of a communication protocol in which XML messages are exchanged; this is called XML Request and Response Protocol (XREP).

Because the Decomate system is supposed to be highly customisable, a library that implements Decomate can choose its own layout of the screens. It is also possible for a library to add new functionality. In order to achieve this, the Broker has a *User Interface Generator* that is driven by XML templates that are easy to maintain by a library. A site maintainer is developed for designing and maintaining a Decomate site. From the XML templates, the User Interface Generator can produce HTML pages. In addition, other formats can be supported.

Separate from the Broker is the *Multiple Protocol Server* (MPS). The MPS connects to the different database systems. The preferred protocol used by the MPS is Z39.50 for searching and (secure) HTTP for the fetching of full text documents. The MPS allows for simultaneous searching of Z39.50 databases. It can keep the connections with the Z39.50 databases open, even though the Broker will disconnect from the MPS each time it receives a response from the MPS. When the Broker reconnects to the MPS, the connections with the Z39.50 databases are ready to be used for sending new requests. All the information that is retrieved by the MPS is returned in XML to the Broker. The MPS has translation facilities for converting USMARC (and other Marc formats) and records in Generic Record Syntax into XML documents that can be processed by the Broker.

Other protocols can be added to the MPS. To facilitate this there is an API (Application Programming Interface) for interfacing the MPS with software modules that implement these other protocols. An example is the addition of the Pica protocol to the MPS by Tilburg University. In this way, it is possible to achieve complete integration of the Pica catalogue in the Decomate system as it is running at Tilburg University. The catalogue can be searched via the Z39.50 protocol. For information about which books a user has borrowed, which books s/he has reserved and what fines are outstanding, the Pica protocol is used. The same protocol can also be used to connect the Decomate system of TU with the national ILL system of the Netherlands. Integration with the local systems for document delivery is another objective of the Decomate project. This will be implemented by a subsystem called the *Document Requester*.

The Broker mediates between the user and the local systems for authentication and directory services (information about the users: names, addresses, e-mail addresses, affiliations) via the *Authentication broker*. Different authentication mechanisms can be used in Decomate. Preferably, a mechanism that allows for single-logon is used: the user logs on the network once and after that all services on the network are available to him or her. The use of X509 certificates, SSL (Secure Socket Layer) or password daemons, and the linking to an LDAP server with directory information are planned.

A last feature of the system that has to be mentioned is the *Current Awareness Service* that periodically informs (alerts) users about new documents that fit their intellectual interests. The Current Aware Service is implemented by a *Robot* that regularly sends search requests to the Broker and by the *CAS Server* for maintaining the interest profiles and accessing the new results. Users can save queries as interest profiles that they can edit later if necessary. After an interest profile has been run, the user is alerted by e-mail. In addition to e-mail, the results can be consulted by entering the system. The last option gives the full functionality of the system with respect to access to the electronic full text and ordering documents.

Current state of affairs: April 1999

The Decomate project started in February 1998. Since then, a lot of work on the development of the actual system itself has been carried out. In March 1999 a so-called interim release of the Decomate II system was made available. The software is installed on the four sites of the project partners. In this release, both cross-site access and the simultaneous searching of multiple databases is available. By using the MPS, it is now possible to search simultaneously in the distributed article databases of the London School of Economics, Tilburg University and the Universitat Autónoma de Barcelona and in a database with references to the full text of economic research papers. After the search, the user can consult the results by database. Currently, the project is experimenting with the detection of duplicates in the result sets of the different databases. In the next release of the software, it will be possible to search in a virtual database that consists of the distributed article databases. Duplicates are filtered from the result set, being a merge of the result sets of the different databases.

Another feature of the March 1999 release is the cross-site access; not only can a user that connects to the local Broker search in a remote database, but they can also retrieve a full-text document from a remote document server. Thus, a user from Barcelona can, using the local Decomate user interface of Barcelona, not only search in the article databases of London and Tilburg, but also fetch full text documents from the document servers of London and Tilburg. In this way, the

Barcelona user has not only access to the full text of the journals that his/her local library subscribes to, but also to the full text of non-subscribed journals that is stored on the document servers of other libraries. The project has permission from Elsevier Science and Kluwer Academic Publishers for this cross-site access experiment.

Content in Decomate

Having an advanced system such as the Decomate system is one achievement. Its speed is impressive now. However, to really attract users to the system, it is not enough for it to be technically advanced. The system will only be successful if it offers enough material, enough content, to its users.

In the Decomate project Tilburg University, the London School of Economics, the Universitat Autónoma de Barcelona and the European University Institute in Fiesole act as a *de facto* consortium. The partners have made license agreements with Elsevier Science and Kluwer Academic Press for the journals to which they currently subscribe. As mentioned above, both publishers allow cross-site access to the various journals within the framework of the project.

The Decomate partners have identified the 100 (in fact 118) most important journals in the field of economics that they subscribe to and that are not published by Elsevier Science and Kluwer Academic Press. In this model, the consortium has focused on a specific discipline and on material it really needs. As a trans-national consortium, it deviates from the more common national and regional consortia. Because of its disciplinary focus, the Decomate consortium is not interested in package deals that some publishers are so keen to offer. All members of the consortium get access to all electronic journals independent of the subscriptions of the individual libraries.

The Decomate partners have asked Swets and Zeitlinger to act as their agent and to provide access to the electronic files of these journals from different publishing houses. According to Swets, more than 50% of the titles on the key list are available electronically. A problem turns out to be the Decomate requirement to store the bibliographic information including abstracts locally. Even if Swets stores this information on its own *SwetsNet* site, then Swets is not allowed without the consent of the publisher to deliver the bibliographic files to the Decomate partners. Currently, Swets is in the middle of the negotiations with the publishers.

Conclusion

This article has tried to give an impression of what the Decomate II project involves. The project illustrates that there are many hurdles on the road to integrated access to digital information. A library has to deal with many players in the digital information domain, each with its own interest. The use of standards is not well developed, so digital products and services are difficult to integrate in one consistent overall service. The Decomate project is trying to contribute to the clarification of how such an integrated service could look. This is achieved by producing a working system that offers an improved access to (digital) information in the field of economics.

References

1 The official Decomate II Web site: http://www.bib.uab.es/decomate2

2 Elsevier Science, TULIP – The University Licensing Program: http://www.elsevier.nl/homepage/about/resproj/tulip.htm