Key issues

Documents in Information Science: a free space for our profession on the internet

1. What is DoIS?

DoIS (Documents in Information Science) is a service for finding and downloading research documents in the area of library and information science. It is part of a broader initiative called RCLIS (Research in Computing, Library and Information Science).

RCLIS is an international co-operative effort to develop a digital library for information science. RCLIS is managed by an international team of volunteers made up of professionals from different areas (such as librarians, computer scientists and economists). They come from a number of countries like Spain, Italy, the United States of America, Bosnia and India. The initiative has two main objectives. Firstly, it aims to compile and to place in the public domain metadata about research documents. The data is freely available for public and private, commercial and non-commercial purposes. It will also serve as a testbed for digital library research. Secondly, to facilitate access to documents freely available on the internet, in order to increase their visibility.

RCLIS is an initiative to export to the library discipline the model defined by RePEc (Research Papers in Economics), a highly successful digital library specialising in economics that provides metadata about more than 200,000 research documents. Both digital libraries have been built on a distributed architecture. They are based on the principle of cost sharing between as many participants as possible, so that each one contributes only a tiny fraction of the work needed to carry out the objectives. Participants in the digital library may be classified in two categories: data providers and service providers. While data providers (or archives) hold metadata about documents, service providers take such metadata in order to provide some added value and to make the result useful for the end user. The interchange

JOSÉ MANUEL BARRUECO CRUZ
University of Valencia

IMMA SUBIRATS COLL
Government of Catalonia

of metadata from archives to services is done using a basic set of rules that are specified in a document called *Guildford Protocol* (Krichel, 1997). Metadata is encoded using a bibliographic format called ReDIF (Research Documents Information Format) (Krichel, 1997). In this way a graphical description of the RCLIS architecture is shown in Figure 1.
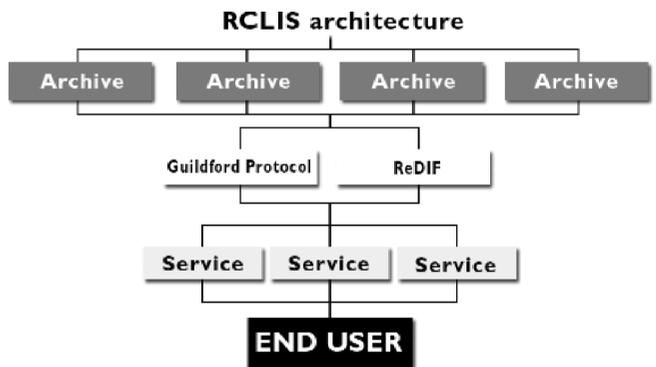

Figure 1. RCLIS architecture

In this architecture DoIS plays the role of service provider or content aggregator for all RCLIS archives. In April 2003, DoIS held metadata on more than 10,530 articles published in journals and 3,144 papers presented at conferences: in total more than 13,600 items with 7,238 of them freely
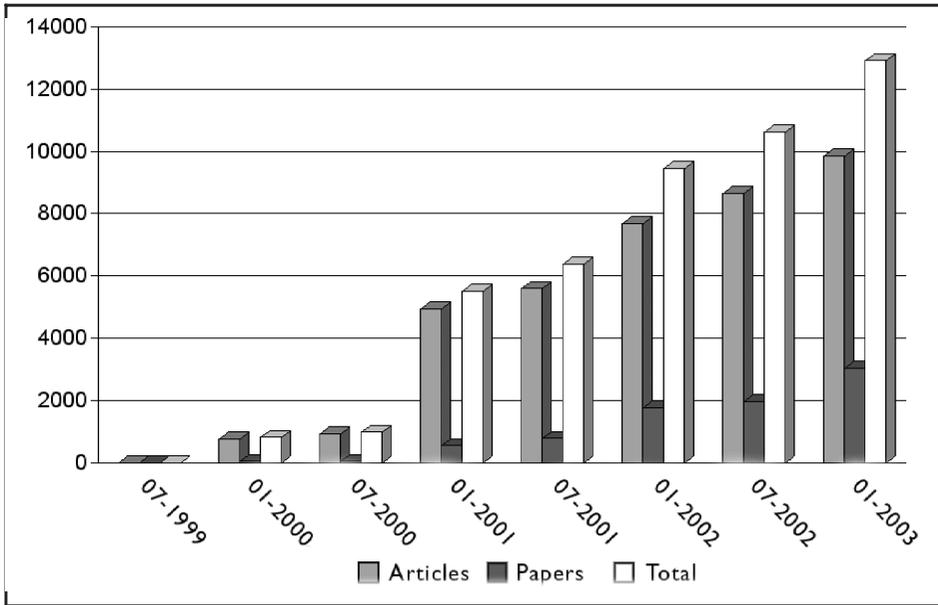
Figure 2.
Evolution of the
number of documents
in RCLIS

available in electronic format on the internet. Since DoIS was born in 1999 the number of documents has been increasing steadily, as can be seen in Figure 2. Figure 3 shows the number of times the service web site was accessed over the same period of time.

The coverage of journals includes the most important publications in the discipline, such as *Journal of Documentation* (1991-) and *Journal of the American Society for Information Science* (1986-). It includes significant electronic journals like *D-Lib Magazine, Ariadne, IFLA Journal* or *Bulletin des Bibliothèques de France. Serials*, the journal of the United Kingdom Serials Group, has recently been added to the dataset. At the moment we have only

metadata for the last two years but more back issues will be included soon.

Even though the coverage of conference proceedings is centred mainly in the Spanish world, it also includes international meetings such as ISKO and IFLA Council and General Conference.

2. How to access the information?

The user has two different ways to access the DoIS contents: browsing and searching.

Firstly one can browse the documents by lists of publications. After a first choice between journals or conference proceedings, a list of publications
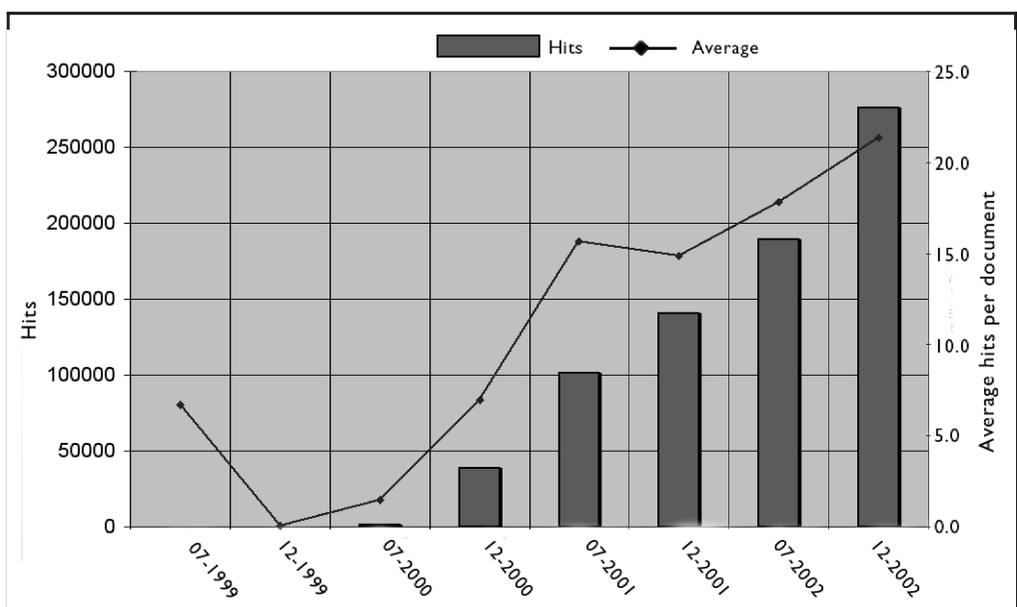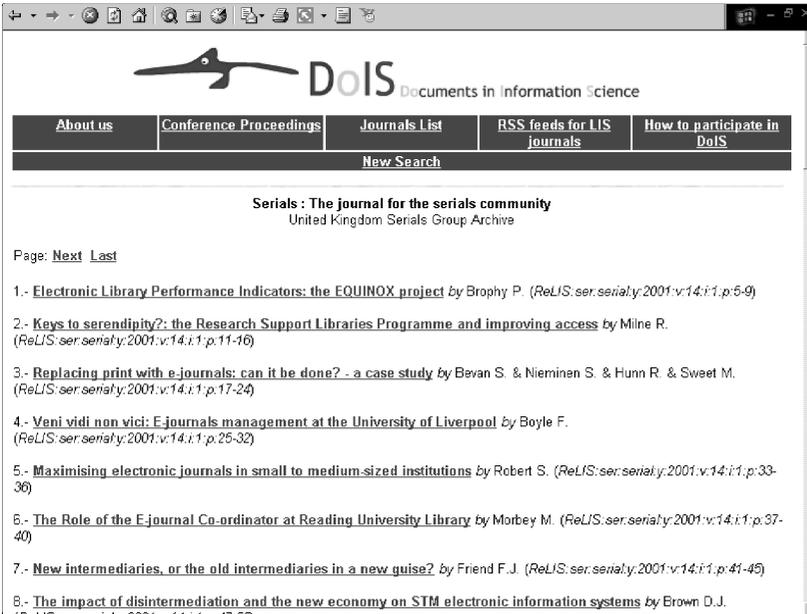

Figure 3.
DoIS usage

Figure 4.
Browsing of
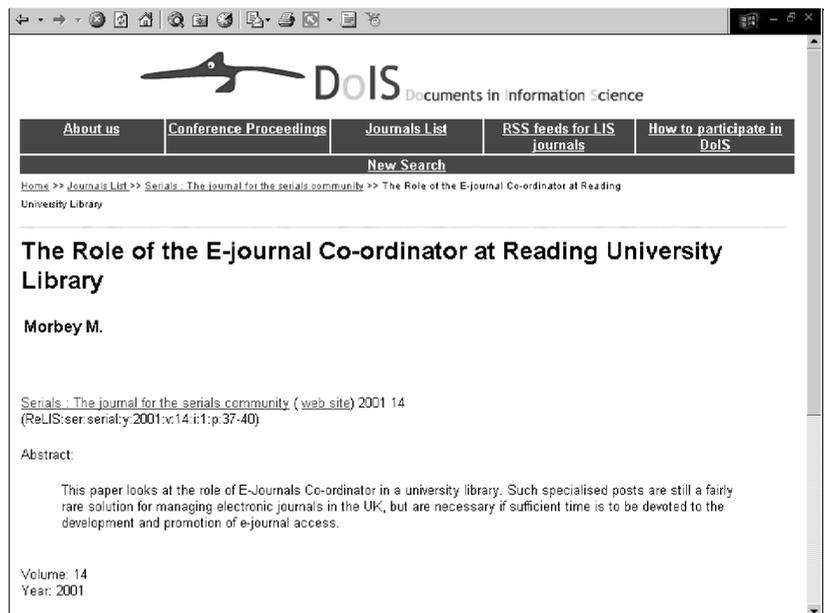*'Serials' articles*

Figure 5.
Article description

sorted alphabetically is presented. For each publication there is an icon stating whether or not the publication has electronic contents. The number of items for each entry is also presented. Once the user has chosen a title, he is presented with abbreviated information for each article or paper published. The user may browse the paper titles and by selecting one of them, can move forward to the full bibliographic information including links to the paper full text when available. Figures 4 and 5 show this process.

Secondly, one can use a search engine to seek for documents containing particular words. DoIS uses the Unix based Swish++ engine to index the web site. It allows boolean searches on the full text of the pages, which is sufficient for the majority of queries, but if necesary it allows the user to make fielded searches by title, author and abstract. This is possible because Swish++ is able to understand the META tags of the html pages. Three different tags have been created in each page to allow this type of search.

Recently DoIS has started to provide a new service which is called 'RSS feeds for LIS Journals'. Rich Site Summary (RSS) is a lightweight XML format designed for sharing headlines and other web content. Basically it is a service that distributes, using a standard format, the table of contents of several electronic journals from the moment they are made available on the internet.

In this way, it may be used by third party content aggregators to provide new awareness services for their users. At the moment only feeds for *Ariadne, D-Lib* and *Information Research* are provided but more will be added in the future.

3. How to participate in DoIS?

DoIS is different from similar initiatives in that it is grounded on voluntary work by people from a variety of different backgrounds, and it has a non-commercial orientation (there is neither funding nor interest in gaining revenue or making a profit). It is based in the spirit of the open source initiative in computing where people worldwide work together to build software that is in the public domain. In our case we work to create a digital library freely available for the LIS community.

In this way we need collaboration from two areas. Firstly, if you publish a journal or organise a conference and want to increase its visibility in the research community, please get in touch with us. Secondly, if you share our goals and want to become a member of an international team, please also get in touch with us.

4. More information:

http://dois.mimas.ac.uk

http://rclis.org

http://repec.org

Thomas, K. (ed.), *Guildford Protocol*, 1997. Available at:

ftp://netec.mcc.ac.uk/pub/NetEc/RePEc/all/root/docu/guilp.html

Thomas, K. (ed.), *Research Documents Information Format,* 1997. Available at:

ftp://netec.mcc.ac.uk/pub/NetEc/RePEc/all/root/docu/redif_1.html

José Manuel Barrueco Cruz
University of Valencia
E-mail: barrueco@uv.es

Imma Subirats Coll
Government of Catalonia
E-mail: immasubirats@myrealbox.com