Electronic SDI : State of the Art at IIT, Kharagpur

V K Jagajeevan and **A K Katna**
Central Library, Inidan Institute of Technology, Kharagpur - 721 302

Abstract

The evolution of computer based matching of context profiles of information with interest profiles of users has helped in the transformation of the traditional SDI service to an electronic form in the central library of IIT, Kharagpur. Electronic SDI was introduced in 1997. Based on the Current Contents On Diskette (CCOD) database on three of the major thrust areas i.e. i) physics, chemistry, earth science ii) engineering, computing and technology, and iii) social, behavioural and earth sciences, the authors have described the methodology adopted in the Central Library. The problems faced and the future prospects are also mentioned. Besides opening new vistas for paperless correspondence, the electronic SDI also provides opportunity for the development of the expert database of faculty interest profiles.

Introduction

Traditional methods of notifying researchers regarding the relevant literature in a SDI service have dramatically changed due to the evolution of computer based matching of content profiles of information with interest profiles of users, using Boolean and other techniques.

Electronic SDI Service at IIT, Kharagpur

One of the most important services provided by the Central Library of IIT, Kharagpur is the timely access to a wide range of journals. Such access is an essential ingredient for the success of the IITs in state-of-the-art teaching as well as R & D. The retrieval of information from the primary and abstracting journals in print format is becoming increasingly inefficacious due to time lags in printing, postal delays and inadequate methods of retrieval in view of enormous quantum of literature generated.

Electronic SDI was introduced in May 1997. It essentially consists of scanning keywords (research profile of individual users) in the current literature, filtering of information, and transmitting the retrieved information to the users on their desks. The construction of profiles according to the standardised terms is an intellectual job and can form a

foundation for routine SDI service. A continual tuning of the profiles is desired as per the feedback. A constant monitoring mechanism of the service is evolved out of useful feedback and constructive comments. An optimum level of recall and precision should be ensured by constant interaction with the users to know their current research interests. It must be able to serve the twin purposes of alerting the users about the latest research being published in their research field and also provide sufficient inspiration to acquire the relevant article (s) either by photocopy or by resource sharing arrangements.

Current Contents are available in electronic format on weekly basis and contain information about literature published/made available at ISI Office during that week. The floppies popularly knows as CCOD are airmailed and reach clients in India within two weeks. This publication is also available at its Web site may be downloaded for a really current SDI. The Central Library subscribe to three out of the seven parts of Current Contents on Diskette (CCOD) database published by Institute for Scientific Information, U.S.A. Considering the major thrust areas taken-up by our Institute these satisfy majority of the teaching and research needs. They are

a) Physics, Chemisty & Earth Sc. (PCES with abstracts),

b) Engineering, Computing and Technology (ECT), and

c) Social, Behavioural and Earth Sciences (SBS).

Each of these includes information on articles published each week out of the over 1000 core journals.

Methodology adopted

The methodology consisted of constructing 'keyword' profiles relating to the current areas of research being undertaken in various departments by our faculty. These profiles are scanned for relevant literature on monthly basis from CCOD issues received in a particular month. Each user is communicated the output received for his known profile with a request to furnish the comments on its utility and relevance. Based on the remarks received, the precision for SDI is enhanced.

The current faculty strength of the institute is about 450 serving in its 25 departments. To initiate the service, sample profiles are constructed from the subject interests of faculty listed in the recent Annual Report of the Institute. The subject terms are standardised using the 'keywords' adopted in the CCOD. They are then combined with boolean operators AND, OR, NOT and THRU (repeated OR), to manifest the term relationships and research requirements. An optimum number of output

records is also achieved by using them. The subject profiles are stored for future use by assigning filenames same as the faculty names (including a preceding symbol for the department). A request for collecting research interests of the faculty is also issued through e-mail, alongwith the first sample output. The research area profile is sophisticated using keywords based on bibliographic details like co-authors, their addresses, source-journals etc., as per feedback. Thus tuning of profiles is a continuous process helping to maintain the dynamism of the service. The response received during the last six issues of the service has been highly encouraging. Every month, new updates of suitable databases are matched with faculty profiles and outputs mailed.

Before the electronic SDI assumed its present shape, individual researchers would attend the Central Library and use CCOD to scan for relevant literature based upon their desired "keywords" and download the information on a floppy. The CCOD's window version offered searching for last 16 weeks at one session. The CCOD transcends from the same publications in print form with a user friendly sophisticated searchware. We can search the database in two modes - Quick search for simple queries and Full search to answer complicated information needs. The searchware supports the construction of user interest profiles to repeatedly search different updates or versions. Another feature of special significance is record selection (filtering) during output stage using Special Interest Checklist. The first feature essentially supports the batch processing operation of SDI while the second one further ensures screening the output of irrelevant literature.

Problems & Prospects

As the CCOD are published weekly and cover a wide range of journals published world-wide, the information generated for any individual user would contain literature from sources that may not be acquired by our library; information pointing towards the availability of the desired article locally would be of great help in decision making for the procurement of full text/document. An attempt is being made to link-up our database of subscribed journals with the SDI output so as to enhance the document delivery process.

Many Indian journals are not covered by the CCOD, and hence information contained therein is not available through this facility. We are planning to create a database covering disciplines relating to Science & Technology literature produced in India (with retrieval facility) and to add as an interface with the existing facility.

Some operation constraints may develop like shortage of computing resources due to the increase in data transmission which causes choking of campus LAN.

The efficacy of Boolean logic widely followed in most of the databases and retrieval searchware is somewhat not exacting enough to manifest the increasing precision requirement of researcher in inter-disciplinary and multi-disciplinary areas. Balancing the conflicting IR parameters of recall and precision is a never ending phenomenon.

We also face problem of undesired retrievals due to the difference between 'what is needed' and 'what is demanded' by the user, which ultimately leads to imprecise formulation of the search strategy. An expert system needs to be developed to match natural terms with those of standardized terminology. An inquiry may also be required to understand the exact needs of the researcher.

The interpretations of copyright act in providing such services are quite diffused and may pose a serious problem if not taken up in right perspectives. Luckily most of the publishers of electronic databases allow the use of the information for the obvious benefit of promotion in the use of electronic media and in quick access. The libraries are, however optimising the use for which they subscribe such services and hence remain within the prescribed "fair-use". The library however do not commercially exploit by spreading the usage. Libraries do have vast resources of information and by adopting new technology, they can provide value added information services.

A "push technology" application to put SDI output automatically at faculty's computer each time the database is updated, using Sybase and WWW is also under testing.

Conclusion

The electronic SDI has opened up a new vista for paperless correspondence. Several off-shoots of this exercise would emerge, like the expert database of faculty interest profiles.

Acknowledgments