РОССИЙСКИЙ НАУЧНЫЙ ЭЛЕКТРОННЫЙ ЖУРНАЛ
Электронные библиотеки
Издается с 1998 года
ISSN 1562-5419

Russian Digital Libraries Journal - 1999 - Vol 2 - Issue 3

You've Got Personalised Library Mail! : A Personalised Current Awareness Service for Library and Information Staff

Robin Yeates
LITC, South Bank University

Abstract

There are three main ways of obtaining information: searching, browsing and alerting. The first two are being widely developed by libraries using the Web, but the last has been somewhat neglected. The NewsAgent for Libraries project was originally funded under the eLib Programme by JISC (Joint Information Systems Committee of the UK higher education funding councils) as a two-year collaborative project started in April 1996.

Several small publishers of library and information science journals worked with network specialists, market evaluators and commercial software developers to design an open, distributed architecture for disseminating information via email and personalised Web pages. Dublin Core metadata was used, enhanced by NewsAgent specific keywords, to map stored user subject profiles against information feeds. Metadata was harvested using software robots to build an Oracle database where both user profiles and document attributes were stored.

Users can join the service via a Web page, to receive information updates by email or as a personalised Web page. Users can select predefined Topics in which they are interested, or create new named ones (stored queries). They can also modify existing Topics. Topics are presented in groups, called Channels.

A major part of the project was an extensive study of the potential end users of the service, before and after a prototype service was created. The project was considered a success, although further development of both software and marketing strategy were needed before a full scale launch could be planned. This is now expected in autumn 1999. In addition to this service, the software is being applied to other services by different organisations, targetted at groups such as small businesses, medical information and environmental information. It is expected that a commercial software package will be available from Fretwell-Downing Informatics as a result of the project.

1. Introduction

Librarians are inquisitive, acquisitive, and informative. One of the more speculative projects funded under the Electronic Journals strand of the Electronic Libraries Programme (eLib) started in April 1996, to consider them as a suitable community to research and develop the concept of an open-standards-based, electronic, personalised, current awareness service.

NewsAgent has been developed by a consortium of librarians and software developers in the UK and will be launched as a free service on the Internet in the autumn of 1999. Having completed initial development work in mid 1998, the project has moved into its exploitation phase, with each partner institution free to explore its application further. LITC at South Bank University is leading the creation of the NewsAgent for Libraries service, as it is a respected publisher of traditional printed current awareness journals for library and information staff, and also has experience of library automation and digital publishing.

2. From Search and browse to alerting

Aiming to help users access information more easily, it is probably true to say that almost all digital library projects involve an Information Discovery component. Most projects began by focussing on Search functions, to develop the capability for distributed clients to address queries to target information systems provided by a wide range of different suppliers, using protocols such as Z39.50 (ISO23950). Such projects enable users to send ad-hoc queries to servers provided they are aware that information exists. With the extremely rapid growth of the Web and end-user searching, the advantages of offering a Browse capability have also become clear, particularly from the popularity of directory services such as Yahoo! Browsing gives

users a sense of order and structure and allows them immediately to follow interesting-sounding hyperlinks. Some professional librarians have created browsing tools based on traditional classification schemes such as Dewey to cluster electronic information together in familiar ways.

At the same time as Search and Browse services have been developed, email has become popular, along with distribution lists and discussion fora. One of the major uses for email is to alert individuals and groups to new or newly-discovered information. Major publishers and webmasters alike have realised that a popular service that they can provide to users is a facility to allow readers to register their email addresses to receive information alerts, on a regular or irregular basis. Most large learned e-journal publishers now offer such a service, delivering the contents pages of new issues, or titles of newly published articles to the mailboxes of individuals. Some e-journal aggregation services also offer services across a number of publishers - for example ingenta does this for the UK academic community. There is no equivalent service for the wide range of distributed Web sites which offer access to information, although some experiments have been done using usenet newsgroups technology, which has never really taken off for academic or library purposes.

The NewsAgent team have realised that there is a need for communities to operate their own alerting services as part of the information landscape. These will ideally have the advantage of operating across a wider range of information providers, and will operate without revealing personal profile information to the publishers themselves. The publishers will, on the other hand, be able to reach specific target groups without dealing with publishing intermediaries who may be competitors, and they will receive more detailed aggregated management information so that they can develop their publishing more effectively. The user information available will be from a wider range of individuals than just the customers of a single aggregator, and they may feel more freedom to express their true information needs in their profiles if the system is not operated by a profit-making commercial entity. The customisation and personalisation of information services is a trend which is certain to continue. Library and information services will need to ensure that they can deliver to individuals remotely across the Internet, as well as communicate with groups and anonymous users.

3. Description of service

3.1 Scenario - Irina, Librarian

Irina has access to a standard Web browser and requires regular information about new knowledge management systems. She visits the NewsAgent for Libraries Web site, registers herself with a username and password and gives the email address where she would like to receive alerts to new information.

Next, Irina chooses which Topics she is interested in, from a predefined list, by checking boxes. She clicks on the topic names to see a list of recent relevant titles, to check that the system will deliver appropriate information to her email box. She adjusts the Topic selections as required.

Finally, Irina selects to receive email alert messages daily or weekly, and decides whether to have text or HTML format mail. She chooses the Daily text-only version.

Next day Irina receives a personal email message from the NewsAgent service, showing the new titles added to the service since she joined. Each title is presented classified under its Topic name. Topics may also be grouped into more general Channels. She can click on the URLs given to access the full-text documents mentioned from the original site, or in the case of some material, from the NewsAgent content store. Even though she is receiving text-only emails, she has email client software capable of recognizing and activating embedded URLs, as do most people.

3.2 Scenario - Robin, Researcher

Robin has been using the daily email alerting service for a week or so. He has desktop access to the Web from his office PC, and is researching the development of digital libraries for children. The email alerts he is receiving are numerous but do not all seem to be relevant to the subject.

Robin visits the NewsAgent Web site by clicking on the URL link provided at the bottom of each email alert message.

Robin clicks on the Configure button to show a list of the currently selected Topics. Although there is one for Digital Libraries, no Topic seems to be specific to children. He decides to create a new Personal Topic. Using the Digital Libraries stored Topic as a template, he edits it by amending the Topic details in his Web browser. When he has typed the subject terms required, he gives the Topic a name 'Digital Libraries for Children' and clicks on the button to save the new Topic in the central NewsAgent system. Next he revisits his Topics list and de-selects the original 'Digital Libraries' Topic Next day, Robin receives his first email containing hits about Digital Libraries for Children.

3.3 Scenario - Andrew, Web newsletter publisher

Andrew works for a small unit which collects intelligence about library automation systems. He issues a bimonthly printed newsletter and wishes a) to deliver news more promptly to subscribers; b) reach more readers; c) encourage regular use of the unit's Web site; d) build a sound foundation for integration of

electronic publishing operations with external information services for the same target audience.

Andrew creates an MS Access 97 database system to store news items, and allow easy indexing with assigned keyterms and links to supplier contact information. Simple Visual Basic code allows news to be published to the unit's Web site several times a day, along with embedded Dublin Core metadata supplemented by controlled terms which should make it easier for readers to filter news. A one or two step publishing operation also includes generation of a separate Resource Description Framework (RDF) file which will enable users to retrieve news items as part of the customised Web page My Netscape service.

No further action is required from Andrew. The NewsAgent for Libraries harvester robot visits the unit's Web site daily to collect the metadata from new news items and store it centrally in a secure robust RDBMS (Oracle under Solaris). NewsAgent then matches the metadata against user profiles created by the users themselves. Users finally receive alerting information mixed with that from other publishers.

Andrew periodically receives a management report giving information about use of the alerting system and level of demand for information by topic.

4. Collaborative effort

The consortium that has developed NewsAgent was fortunate to include LIS academics, librarians, publishing expertise and commercial library software developers.

Publishers in the field of library automation formed the core of the project team, including:

- LITC at South Bank University, which publishes VINE, and also Library Technology (jointly with the Library Association);
- UKOLN, the UK Office for Library and Information Networking at the University of Bath, which publishes Ariadne;
- University of Wales at Aberystwyth, Department of Information and Library Studies (UWA/DILS), from where staff edit the Aslib journal Program and the Bowker-Saur Journal of Librarianship and Information Science.

Each of these partners in the consortium also took on additional roles in the project, including project management and dissemination (LITC), metadata handling and pilot service operation (UKOLN) and user needs analysis, editorial issues and vocabulary control (UWA/DILS).

These partners were joined by CERLIM, the Centre for Research in Library and Information Management, originally at the University of Central Lancashire and moved in 1998 to Manchester Metropolitan University. CERLIM acted mainly as the internal evaluator of the project, setting and monitoring adherence to criteria in the light of the user needs analysis and its extensive evaluation experience.

Last, but by no means least, Fretwell-Downing Informatics Ltd (FDI) provided the main software development effort, although in the end UKOLN also played a significant part in creating, implementing and evaluating some aspects of the technology, such as the metadata harvesting system.

5. An open solution

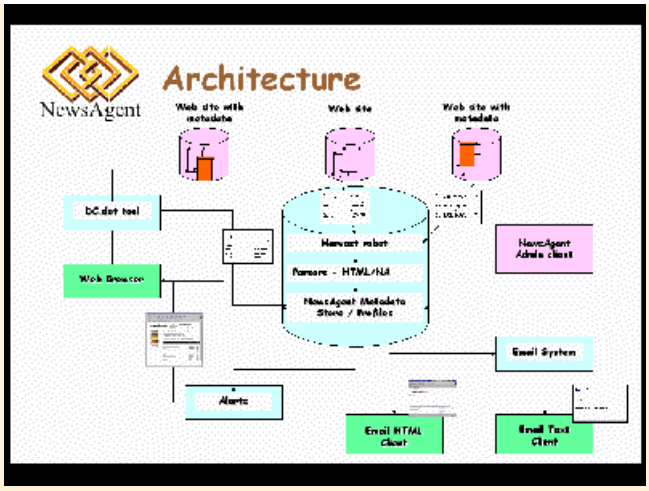
Though this co-operation, we have developed an open solution. It is capable of integrating with integrated library management systems, in particular OLIB7 from FDI. NewsAgent also integrates with metadata tools such as DC.dot. DC.dot was developed by UKOLN and enhanced for NewsAgent to provide the capability for any Web user to send metadata directly to the system for any HTML page.

We spent nine months studying potential user needs and the wider IPR and commercial environment. We discovered that the idea of selective electronic alerting was apparently more popular than printed delivery, and that email should form the main focus of the service, enhanced by a Web-based profile management capability.

We were keen to ensure that, if successful, the software could be marketed by FDI for other purposes beyond the NewsAgent for Libraries service. In this way we might have long-term support for the system and widespread acceptance of our metadata publishing and harvesting model, on which any such automated alerting system depends. Other alerting and information management systems will also benefit from improved metadata provision by authors and publishers.

The overall architecture of the prototype system is as follows:

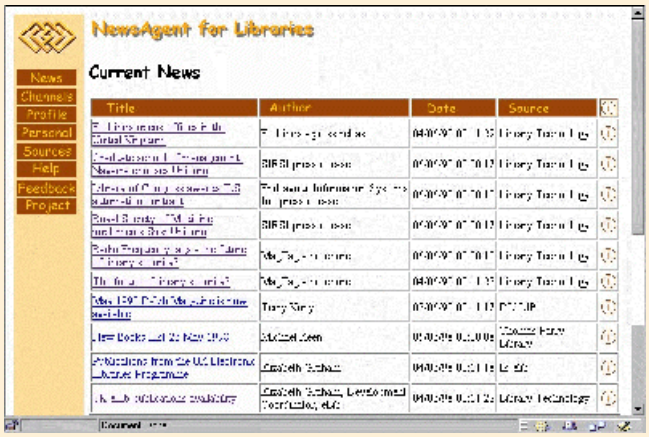
Figure 1: The overall NewsAgent architecture.



Our first prototype faced difficulties mainly relating to the design of the Web interface for end users. It proved difficult to get across simply enough the method of defining profiles. However, a complete working system was tested in early 1998. Since then, we have enhanced the software, produced a version running under Windows NT, and prepared a business plan.

The original user interface at UKOLN was designed more or less as a standalone service.

Figure 2: The first prototype Web page display of alerts.



Further details of our Phase 1 work are given in the Final Report[1].

LITC is currently developing a plan to sign up content providers for the launch of a service in autumn 1999, and is awaiting the second release of the system software expected in July and already almost complete. The full system will be designed to integrate with other parts of a Web site, providing a value-added 'Update' service for those who wish to register their email address for alerts. A number of services are being developed separately by NewsAgent partners, some of which are based on a regional service for small business in Yorkshire. They have redesigned the interface along the lines expected for the LITC service. The following series of screenshots shows the main features of the service:

Figure 3: Joining or entering to update a profile.

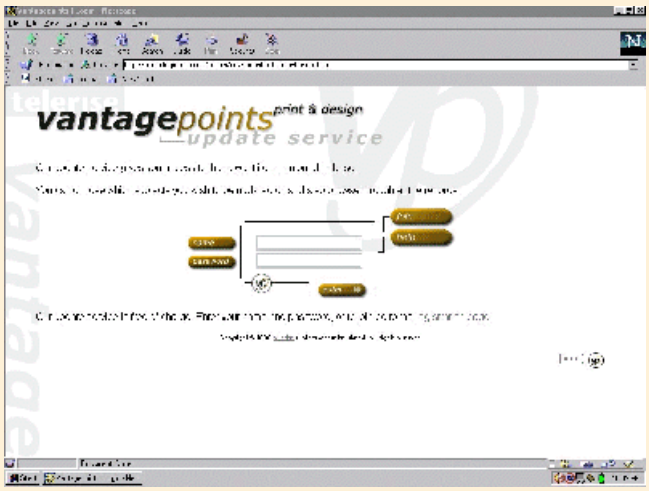


Figure 4: Choosing Topics to store in your profile.

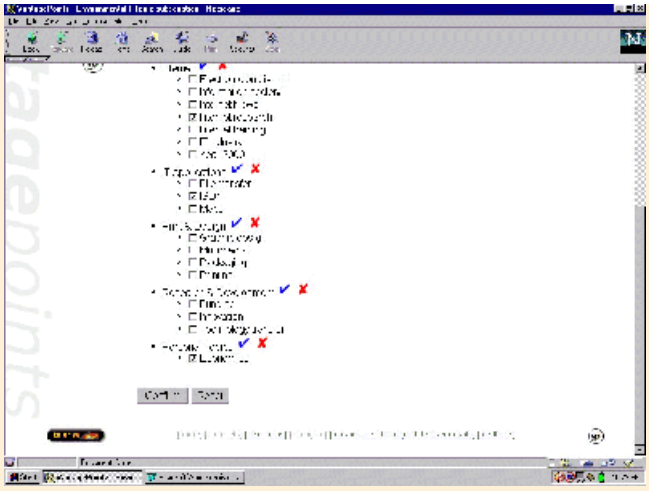


Figure 5: Checking the list of Topics currently subscribed to.

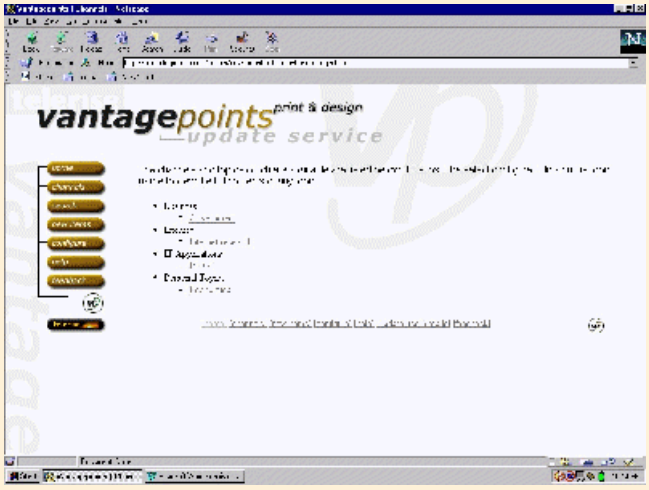


Figure 6: Testing current Profile Topic ideas by searching the database...

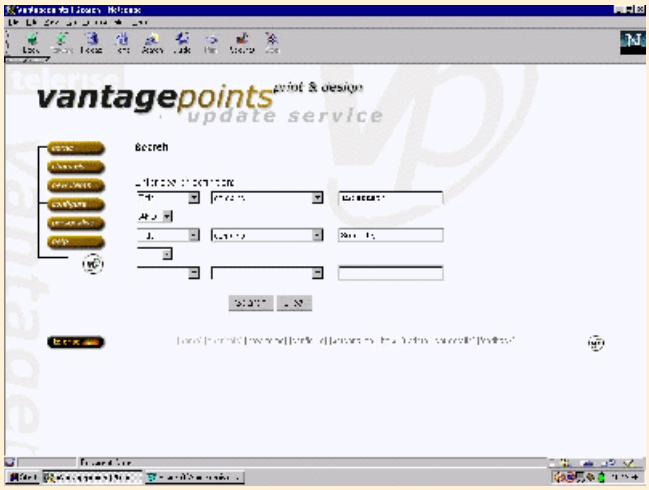


Figure 7: ...and viewing the results interactively via the Web.

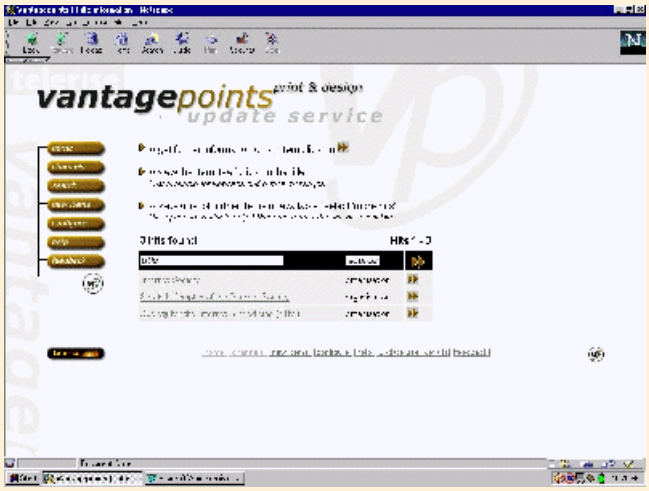


Figure 8: Users can examine the stored metadata about items in the database to check quality/origin.

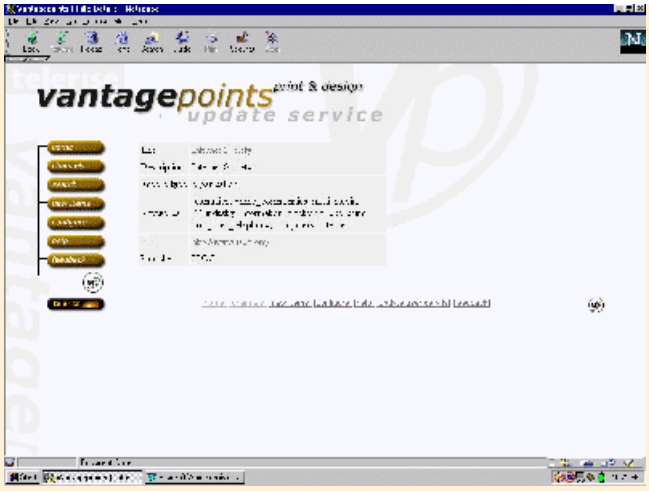


Figure 9: Users select how often to receive alerts and in HTML or plain text email format.

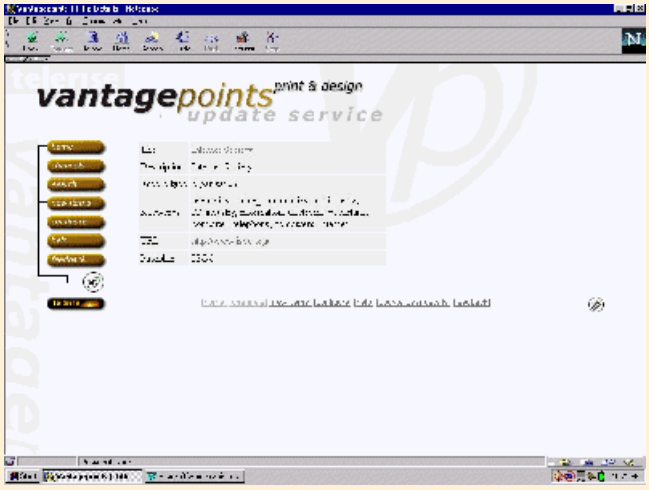


Figure 10: A Netscape Communicator or Internet Explorer 4 user can receive an HTML email alert.

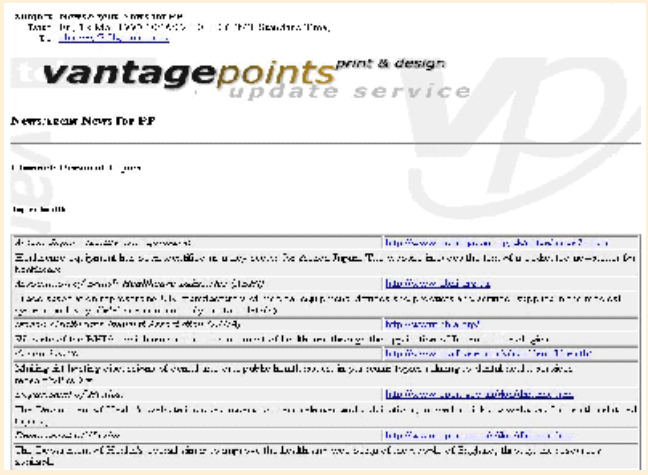


Figure 11: Users can choose plain text email alerts, which may still have live links in many cases, depending on the email client.

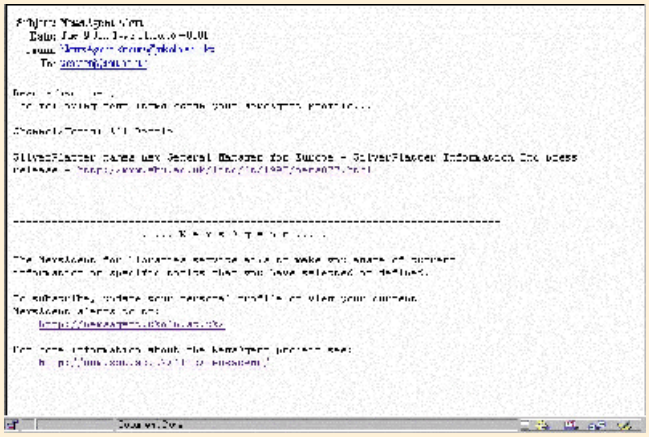


Figure 12: At any time users may visit the Web site to amend their profiles.

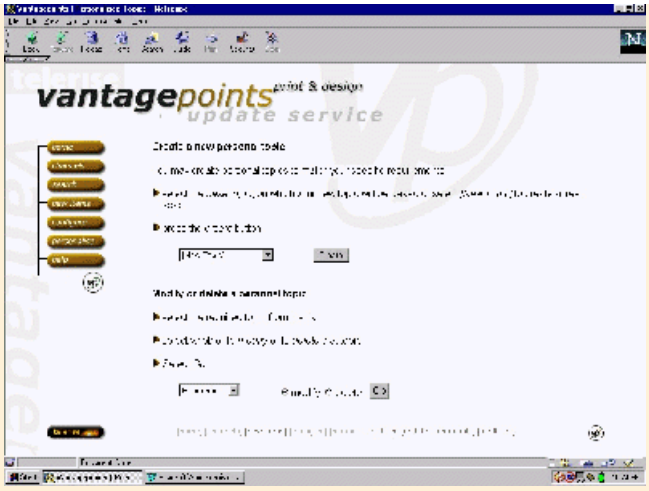
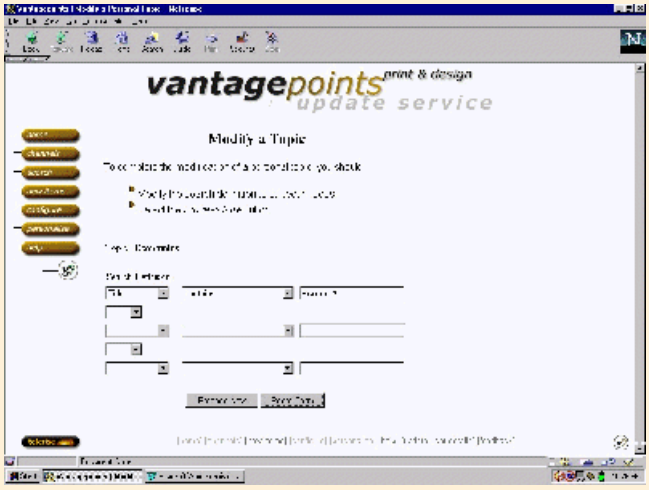


Figure 13: Users can even create or modify their own Topics instead of relying on the predefined Topics.



6. Issues and opportunities

Within the time-scale of the initial two-year project, we were only able to muster a limited amount of content during the project, because we were uncertain of the nature and format of metadata we wished information providers to supply to us. Neither did we know the commercial conditions under which such a service might continue. We have now made the decision to concentrate on freely available Web resources, because we wish to offer a low-cost, convenient and open service to further test reaction before we introduce premium content.

The business model we have chosen to adopt will be based on the service being free at the point of use, funded initially by a mixture of transitional exploitation funding from the JISC and sponsorship by a library organisation (or several organisations) with an existing large user base. We also intend to explore advertising as a means of generating ongoing revenue. The system itself would require relatively minor enhancements to introduce a form of end-user charging. At present, after examining many economic models, we feel that the cost of administering full user billing outweighs the likely income, so we do not intend to follow such a path. We are keen to attract those wishing to deliver information to our audience of library and information professionals, and we shall soon be issuing information for such information providers, who will not have to pay to use the service at present. Those wishing to take the whole system up in other sectors are encouraged to contact us to discuss other exploitation. We expect that FDI will be marketing a commercial version of the system, but no date for its availability has yet been set.

7. Conclusions

One current Internet trend is that towards 'Portals' or 'Virtual Communities', which go beyond the simple provision of material by a publisher to include sharing of electronic mail, discussion list comments, votes and preferences. Although single publisher and single aggregator services will be important, users will always seek to move aggregation functions towards themselves on the network, and librarians will be there helping them to do so, just as they have always held collections from multiple sources. A key difference from the already 'traditional' subject gateways and search engines such as Infoseek or directories such as Yahoo is the need for an individual to join. We hope that systems such as NewsAgent will overcome most individuals' reluctance to join, at the same time as providing an additional channel for communications, with content not restricted to one publisher but controlled by the user him- or herself. Such systems challenge us all to develop more user-centred collaborative business models to support effective and efficient communication.

8. References

1. *NewsAgent for Libraries: Phase 1 Final Report*. A report covering April 1996 to March 1998, JISC Electronic Libraries Programme Project 2/18. Available at http://agent.sbu.ac.uk/newsagent/archive/nafinalapr983.html [18th May 1999]

9. Further reading

- Tedd, Lucy A. and Robin Yeates. A personalised current awareness service for library and information services staff: an overview of the NewsAgent for Libraries project. *Program*, vol.32, no.4, October 1998, pp.373-390.
- Yeates, Robin. Have you heard the library news? Personalised Net alerts for librarians. Aslib Proceedings, vol.51, no.5, May 1999, pp.137-143.
- NewsAgent for Libraries Web site. - URL: http://agent.sbu.ac.uk/newsagent/ [18th May 1999]
