

Scottish Academic Periodicals Implementing an Effective Networked Service (SAPIENS) project

Jake Wallis

Author

Jake Wallis is the researcher on the SAPIENS project. He has worked as a researcher at the Centre for Digital Library Research within the University of Strathclyde since 2000 when he completed his MSc at the University's Graduate School of Informatics. Centre for Digital Library Research, Department of Computer and Information Sciences, University of Strathclyde, Room 12.12, Livingstone Tower, 26 Richmond Street, Glasgow G1 1XQ, United Kingdom
E-mail: jake.wallis@strath.ac.uk

Abstract

This article describes the aims and continuing progress of the Scottish Academic Periodicals Implementing an Effective Networked Service (SAPIENS) project which has been running at the University of Strathclyde's Centre for Digital Library Research since September 2001. Initially funded for two years, the project has been extended until October 2004.

The rationale behind SAPIENS is the concern that small Scottish publishers, operating on limited budgets, are in danger of finding themselves marginalised in the modern information environment. The project's primary objectives are to explore the viability of, and launch, an electronic publishing service to assist small-scale Scottish publishers of academic and cultural periodicals to publish online. It has achieved these aims by implementing a demonstration service which is gradually moving into an operational mode, delivering current journals.

Aims

The Scottish Academic Periodicals: Implementing an Effective Networked Service (SAPIENS) project began in October 2001, funded for two years by the Scottish Higher Education Funding Council (SHEFC). The project had three main aims:

- to explore the case for a centralised electronic publishing service which could assist small scale Scottish publishers of academic and cultural periodicals to enter the online environment
- to design and build an online demonstration service which would deliver journals from a selection of publishers
- to develop and launch a self-sustaining, operational service.

Project rationale

The background to the project was a growing movement within academia to make use of information and communications technology (ICT) to remove commercial

publishers from the publishing chain in the process of research dissemination. The thinking behind SAPIENS was influenced by various factors.

- Limitations on the funds available to Higher Education institutions have prevented academic libraries from being able to afford the increasing volume of research publications
- A trend towards acquisitions and mergers within the academic publishing industry resulted in concerns over monopolistic practices, to the extent that the sector was investigated by the UK Office of Fair Trading (Office of Fair Trading, 2002). The marketing of bundled packages of journals from large publishers, for example to UK further and higher education institutions through the National e-Journals Initiative¹, has had the result of leaving independently published journals more susceptible to cancellation by academic libraries attempting to reduce costs. Institutions have little choice but to participate in the purchasing of packages of journals from companies such as Elsevier, Springer and Wiley in order to attain a critical mass of provision for their end users.
- Developments in the use of ICT have resulted in significant changes in modes of teaching and learning. In today's educational environment much learning activity is undertaken by using a computer to access educational resources over a network. Access to information is less mediated and more direct than ever before. The other side of the coin is that access to vast information resources such as the Internet, combined with other electronic forms of communication, can produce the phenomenon now commonly referred to as information overload. This state might be described as a sense of confusion resulting from the retrieval of information in such quantities that to extract its meaning seems impossible. For the information seeker, faced with a bewildering amount of electronic information, there is a delicate balance to be found between the ease of access to a resource and its quality. The ease of searching using a tool such as Google² to produce a large number of hits, whatever the quality or level of relevance of the retrieved resources, may seem more appealing than the more complex search procedures required to track down an article in a printed journal, when the library's single copy of that journal may be out on loan anyway. The preference for easily accessible full text articles available on the Web has been described as a "Full Text On the Net bias" in information-seeking activities (Wentz, 2002).
- Academically valuable material from small-scale Scottish publishers may find itself on the fringes of this competitive information environment if it is not available and retrievable in electronic form. With educational institutions forced to allocate significant amounts of their library budgets to packages from the large companies, smaller publishers are more susceptible to

¹ NESLI2 (<http://www.nesli2.ac.uk/>) is a national electronic journal licensing initiative for UK higher and further education. It offers a licence model, a defined set of publishers to seek agreements with and a pre-defined negotiation process.

² <http://www.google.com/>

cancellation. This may have an impact on the survival of publishers at this level and on the availability of the special interest material that they publish.

- The final factor relates to the Scottish Parliament's "Digital Scotland" initiative, which "aims to ensure that Scotland obtains and retains maximum economic and social advantage from information and communication technologies" (Scottish Executive, 2000). The SAPIENS project endeavours to further this aim by promoting Scotland's knowledge economy in the global information environment. Furthermore it was intended that the project would add a distinctively Scottish presence on the Web, promoting Scottish culture and heritage.

Methodology

An operational electronic publishing service would be built up in several stages. The starting point was the literature review, which provided the project team with the most recent research upon which it could build. The next stage was a survey of the attitudes of small-scale Scottish publishers to online publishing. Using the results of the survey, a subset of the publishers who had expressed interest in the project was selected. This group provided journal articles for the project team to experiment with. A pilot demonstration service was set up to deliver this sample content online³. Currently the project is moving into an operational phase. This will allow the SAPIENS service to deliver current content and implement its cost recovery models.

The literature review established that the work of four projects funded by the Electronic Libraries Programme in the UK offered some background research on the promotion of an electronic publishing culture and the delivery of electronic academic journals:

- the Electronic Publishing Resource Service project at the University of Surrey⁴
- the Electronic Law Journals project at the University of Warwick⁵
- Internet Archaeology at the University of York⁶
- Birmingham University Integrated Library Development and Electronic Resource project⁷

³ The SAPIENS pilot online publishing service is available at <http://sapiensdemo.cdlr.strath.ac.uk/>

⁴ The EPRESS project (<http://www.epress.ac.uk/>) developed software tools for electronic journal management and production. The tools produced by the project provide the production environment for the online social science journals *Sociological Research Online*, *Journal of Artificial Societies and Social Simulation* and *Journal of Social Work*.

⁵ The ELJ project (<http://elj.warwick.ac.uk/>) was established to promote the development of an electronic culture in legal academic publishing. The project publishes two online law journals; *Journal of Information, Law and Technology* and *Law, Social Justice & Global Development*.

⁶ *Internet Archaeology* (<http://intarch.ac.uk/>) is an online archaeology journal that publishes articles in digital formats in order to make use of the electronic medium's potential for enhancing and extending the content (for example by incorporating clickable maps, databases, graphics and multimedia).

⁷ Amongst other electronic library resources, the BUILDER project (<http://builder.bham.ac.uk/main.asp>) developed two electronic journals; *Forensic Linguistics* and *Midland History*.

The use of electronic publishing to take control of the process of research dissemination is visible at an international level in the form of various projects and initiatives. The pan-European FIGARO project⁸ has similar aims to the SAPIENS project but is international in scale. In the United States the HighWire journal hosting service⁹, based at the University of Stanford, also provides a large-scale, successful model to aspire to.

International initiatives such as SPARC Europe¹⁰ and the Budapest Open Access Initiative¹¹ offer guidance and support whilst acting as advocates for the movement towards open access publishing.

An important first step was to make contact with the community for which the project hoped to develop a service. We compiled a database of Scottish publishers of cultural, literary and humanities periodicals from various information sources such as the Scottish Publishers Association's *Directory of Publishing in Scotland* (Scottish Publishers Association, 2001), and the Scottish Library and Information Council's *Scottish Library and Information Resources* (Scottish Library and Information Council, 2001). These publishers, due to the niche markets that they serve, tend to be small in scale, operating with limited resources.

A survey of the attitudes of these publishers to online publishing was undertaken. 86 questionnaires were sent out and a response rate of 43% was achieved. As part of the survey a number of publishers expressed interest in further involvement in the project (Table 1).

Definitely interested in involvement with the SAPIENS project	12
Interested in finding out more about the project	17
Not interested	8

Table 1: Publisher interest in developing electronic publishing in conjunction with the SAPIENS project

The survey was also useful in that it provided qualitative data on the attitudes to electronic publishing amongst small-scale Scottish publishers. Many had concerns about the impact that an electronic version of their publication might have on sales of the printed edition, which were a significant income stream. In the case of associations and societies where the publication was a benefit of membership there were similar fears over the potential loss of membership subscription income that might result from online publication. It was clear from the survey results that uncertainty over the most appropriate business model for online publication was a

⁸ Figaro (<http://www.figaro-europe.net/index.html>) is an international project aiming to develop a technical and organisational infrastructure to support scholarly publishing by the European academic community.

⁹ HighWire Press (<http://highwire.stanford.edu/>) is an electronic publishing service set up to enhance scientific scholarly communication.

¹⁰ A group of European research libraries and institutions formed SPARC (The Scholarly Publishing and Academic Resources Coalition) Europe (<http://www.sparceurope.org/>) in order to encourage new models of scholarly communication and to advocate change in the academic publishing market.

¹¹ The Budapest Open Access Initiative (<http://www.soros.org/openaccess/>) is an international statement of support for, and commitment to, open access publishing.

significant barrier to publishers at this level. Limited resources in terms of finances and skills were also factors in hindering movement into the electronic environment.

In addition the survey provided information on the production processes these publishers used to publish their periodicals (Table 2). This meant that the project had information about whether the content was stored electronically as part of the production process, and, if so, in what format. This was important in thinking about how the SAPIENS project could tap in to the existing production cycle in order to develop an e-publishing process for that journal.

Word	13
Rich Text Format	2
Quark Xpress	3
Adobe Portable Document Format	4
Adobe Pagemaker	4
HTML	2
SGML	1
No electronic format detailed	8

Table 2: Software formats used in production

Working with a group of small-scale publishers who had expressed interest in the project allowed the SAPIENS team to experiment with sample content that the publishers provided. The publishers supplied their content in different software formats, depending on what they used in their existing publishing processes. This proved to be problematic in that each periodical required a specific process for converting its content into a digital format suitable for delivery over the Web.

It was decided that when creating online articles the SAPIENS project would follow the guidance of the World Wide Web Consortium¹² in using international standards to ensure that the content would be accessible to the widest possible audience. The intention was to minimise any usability, accessibility or technological hindrances that might act as barriers to the journals' online content.

A significant period of experimentation was required in order to develop the implementation of appropriate content standards within the SAPIENS project. Currently articles are prepared using Hypertext Mark-up Language (HTML) which validates according to the World Wide Web Consortium's standards for HTML 4.0. Presentation (aspects such as font and page layout) is defined by cascading style sheets. In addition, where desired by the publishers, articles have been delivered in Adobe's Portable Document Format (PDF), as well as HTML. Whilst it is a proprietary format, PDF is commonly used in the transfer of electronic documents (such as e-journals) over the Internet. The reader software, required to view documents stored as PDF, is freely available from the Adobe web site¹³. Some publishers felt that there was some value in having an electronic version of their

¹² The World Wide Web Consortium (<http://www.w3c.org/>) was founded in 1994 to define standards which would promote the Web's evolution and ensure that it is universally accessible.

¹³ <http://www.adobe.com/products/acrobat/readstep.html>

publication which could be downloaded from the Web and would mirror the printed version in terms of its look and feel.

The issue of conversion from one electronic file format to another proved problematic. Publishers supplied the SAPIENS project team with their journal articles in Microsoft Word, Quark Xpress and Adobe PDF. None of these software packages could export its content into HTML of the standard required by the project. An electronic publishing process therefore had to be tailored to each periodical.

Currently the conversion from the original file format into standard HTML is carried out manually. A template or style sheet is applied to the file in order to give a standardised, branded presentation. One advantage of electronic publishing is the capacity to enhance the searchability of content. Therefore, in addition to the conversion into web-ready digital formats, metadata is applied to the articles in order to make them searchable at both local and global levels. The project team has developed an online search facility which draws upon this metadata. In line with the project's approach to content creation, appropriate international metadata standards are used. This approach allows online journals to be searched internally and also offers scope for enabling the service to interoperate with other approaches to search and retrieval, such as that proposed by the Open Archives Initiative¹⁴. Dublin Core metadata¹⁵ is included in the HTML version of individual articles. Within the metadata, Library of Congress Subject Headings¹⁶ are used to provide the descriptive terms. This process is labour intensive and therefore adds to the cost of online publishing within the SAPIENS service.

A pilot service was set up to experiment with the delivery of the converted content. This allowed the project to investigate the information architecture that the service would require and the technology that could be used to deliver it, as well as some of the issues surrounding access control. This was very important, as those publishers who had agreed to provide sample articles for the project to experiment with, did not wish to allow unrestricted online access to their content.

Achievements

What has the SAPIENS project team achieved?

- It has established contact with the small-scale periodical publishing community in Scotland, and collected data on their publishing processes and attitudes to electronic publishing.
- A pilot electronic publishing service has been developed and is currently online delivering sample content from a range of publications.
- It has investigated the costs of electronic publishing and the means of their recovery.

¹⁴ The Open Archives Initiative (<http://www.openarchives.org/>) promotes interoperability standards with the aim of increasing scholarly communication.

¹⁵ Dublin Core is an online metadata standard.

¹⁶ Library of Congress Subject Headings are part of the Library of Congress system of classification.

- At the time of writing the service is moving in to an operational phase. It will deliver current journal issues, charge publishers for the electronic publishing service that the SAPIENS project provides, whilst assisting these publishers to implement a business model that will recover the costs of online publishing.

This area of cost recovery is acknowledged as being difficult. Both SPARC and the Budapest Open Access Initiative have acknowledged this by producing guides to the business planning of non-profit making electronic publishing ventures. The electronic publishing process is certainly not cost free. It requires skilled staff and a technical infrastructure which must be maintained and intermittently upgraded. Value-added services, such as the metadata creation and searching facilities offered by the project, add to the costs of the process. A subscription model was initially envisaged as being the most effective means of cost recovery. However, as the open access movement grows in momentum the approach to cost recovery taken by the SAPIENS project team has evolved. This is an area that must be explored further by electronic publishers. The difficulty is in moving from the most prevalent existing model (subscription charges for access) to a model that is better suited to today's academic publishing environment. Alternative means of recouping costs, such as the open access model, with charges for dissemination rather than access (as operated by the online biomedical research journal provider BioMed Central¹⁷), are becoming more widely used. In practice this publication charge might be covered by the author's institution or funding body. What might emerge could be a hybrid business model, with costs recovered partly by subscriptions and partly by submission charges for articles which the authors decide to provide on an open access basis. Statistics on article downloads could offer useful guidance as to the circulation of closed and open access papers. SPARC Europe advocates this approach as a means of moving a journal from a closed to open access model whilst continuing to generate sufficient revenue to fund the publishing process (Prosser, 2003).

The SAPIENS project has identified the functional components of the small-scale publishing environment. These components include:

- document formatting
- document storage
- sample content
- metadata for the retrieval of individual articles by author, subject and title
- browsing of journal titles, issues and articles
- free-text retrieval of article content
- user authentication
- subscription management.

The identification of individual components allows the SAPIENS project to offer an electronic publishing service that has a modular framework. Publishers can select the individual e-publishing modules that they require. This provides the flexibility

¹⁷ BioMed Central (<http://www.biomedcentral.com/>) is an online publishing service offering open access to biomedical research. As a service it has taken the decision that open access publishing in the online environment offers a more rapid and efficient method of disseminating scientific research.

necessitated by the varying constraints and resource limitations of small-scale publishing. Publishers can take up whichever elements of the electronic publishing process are unavailable to them at the local level, or whichever fit with their business plans and marketing strategies. For example, publishers may wish to store their content locally yet enlist the SAPIENS project to provide the article level metadata which would enable retrieval in a global information environment. For publishers who wish to increase sales of their printed periodicals the SAPIENS team could provide sample content online. This flexibility allows publishers to pace their take up of digital content delivery in line with their existing practices, procedures and financial constraints.

Conclusions

It has taken time to understand fully the intricacies of the electronic publishing process, in particular, to address the need to:

- put the appropriate technical infrastructure in place
- overcome the difficulties in converting content from proprietary to open standard formats
- investigate the costs of the process and the means of recovering them.

The project has also encountered more intangible barriers to the uptake of electronic publishing. In addition to the uncertainties over the most appropriate business model, the online market and the level of IT skills required to succeed within it, there is, amongst some, a cultural perception of publishing limited to a print paradigm, which can obfuscate the benefits that electronic publishing can provide (such as improved access, search facilities, heightened presence, preservation of back issues). An online issue of a publication can complement rather than threaten the printed version. SAPIENS does not wish to do away with print but rather to offer another means of delivery which may open up new markets for existing publications.

The approach espoused by the project is to define content (such as that of a journal issue) as being separate from its means of distribution, with printing and electronic publishing as simply two different delivery mechanisms. Publishers are being asked to change their publishing paradigm; this is both a cultural and process change that will take time. This view is supported by the Budapest Open Access Initiative:

“electronic-only publication...may require that you overcome considerable systemic inertia to change the perceptions of various stakeholders” (Budapest Open Access Initiative, 2003)

As the information environment changes, there is more pressure for small-scale publishers to engage with electronic publishing if their publications are not to be lost in the margins. Clearly identifiable benefits of electronic publishing include adding value to content through enhanced retrievability and heightened presence in the online environment, improved access and the potential of new markets. SAPIENS, having developed a flexible e-publishing model, must demonstrate these benefits to the publishing community if it is to have its support.

References

Budapest Open Access Initiative. (2003). *Guide to business planning for launching a new open access journal*. New York: Open Society Institute. Available at: <http://www.soros.org/openaccess/oajguides/index.shtml> (checked 14 August 2003).

Office of Fair Trading. (2002). *Can the scientific journals market work any better?* London: Office of Fair Trading. Available at: <http://www.offt.gov.uk/news/press+releases/2002/pn+55-02+can+the+scientific+journals+market+work+better.htm> (checked 14 August 2003).

Prosser, D.C. (2003). From here to there: a proposed mechanism for transforming journals from closed to open access. *Learned Publishing* 16 (3), 163-166.

Scottish Executive. (2000). *Digital Scotland*. Glasgow: Scottish Executive. Available at: <http://www.scotland.gov.uk/digitalscotland/default.asp> (checked 14 August 2003).

Scottish Library and Information Council. (2001). *Scottish library and information resources*. Hamilton: Scottish Library and Information Council.

Scottish Publishers Association. (2001). *Directory of publishing in Scotland 2001*. Edinburgh: Scottish Publishers Association.

Wentz, R. (2002). Visibility of research: FUTON bias. *The Lancet* 360, 1256.

WORD COUNT: 3269