

Evolving Ecological Niches: Technological Change and the Transformation of the Libraries Role in Publishing

Andrew Treloar

Senior Lecturer in Information Management, School of Computing and Mathematics, Deakin University, 662 Blackburn Road, Clayton, 3168, Australia.
Email: andrew.treloar@deakin.edu.au. WWW: <http://www.deakin.edu.au/~aet/>

Abstract

Print has been the most significant scholarly communication technology for the last three hundred years (at least). Kaufer and Carley's Ecology of Communicative Transactions analyses print communication in ecological terms. This paper applies this perspective to the changes now occurring in scholarly communication.

The theory of punctuated equilibrium proposes that evolution of new species occurs both in bursts and in response to changes in environments. Rapid changes in the scholarly communication environment have occurred over the last fifty years, and most particularly since the rise of the Internet. Viewing the Internet as a new ecological niche, this paper looks at five university libraries that are redefining their roles in the scholarly communication ecology. They are acting as facilitators for electronic scholarly publishing rather than just as access points for content created by others.

The five projects (Highwire Press, Internet Library of Early Journals, Project EDUCATE, Project Muse and the Scholarly Communications Project) all demonstrate different organizational models, funding sources and types of content. In interviewing project team personnel, two clusters of issues emerged related to libraries and their changing roles. With respect to the nature of publishing, most respondents emphasized access and the need for an ongoing commitment to content. With respect to the libraries role, all saw their activities as consistent with their responsibilities to their user communities; they saw no barrier to libraries moving into electronic publishing.

The past of our planet can give us pointers to how changing systems might respond now. Past episodes have been characterized by rapid diversification followed by a locking in of a few choices. We are currently seeing a rapid development of new species of scholarly publishing artifacts, with some being selected against and disappearing. Will the next decade see a return to stasis?

Introduction

It has become a truism to say that technology is changing everything. For libraries and readers over the last three hundred years or so, the predominant information delivery technology has been print, typically instantiated as monographs and serials. For most libraries this is still the most important technology and will be into the foreseeable future for many applications. The technology of print has of course had a significant influence on communication in all fields.

The communication ecology

David Kaufer and Kathleen Carley in their book *Communication at a Distance* (Kaufer & Carley 1993) discuss the advent of print and its influence on the world of communication we inhabit. Their work places the focus on the entire interaction cycle of communication as the basic unit of analysis. This shifts the emphasis from particular elements in the communicative transaction to a single communication ecology.

Briefly summarized, this communication ecology consists of agents, communicative transactions, content and context. *Agents* may be individuals, communications artifacts or software and may operate autonomously or in a well-defined way. *Communicative transactions* are the process whereby agents exchange messages and may occur using any media. Messages in turn contain *content*, and are situated within some overall *context*. The term ecology evokes images of richness, interdependence, and evolving complexity. It also allows the analysis of print to move beyond the traditional open systems model.

The idea of the ecological structural framework is to build on this simple communication ecology model by regarding all these components of communication as mutually defining, co-adaptive, and co-evolving. In the context of research into new forms of communication, such a model means that attempting to consider any one component in isolation is fraught with difficulties. Changes in one area will almost certainly prompt unpredictable changes in another. Later work by Kathleen Carley (Carley 1995) has moved beyond the print focus of the earlier work and started to suggest some ways in which this analysis might deal with some of the new communication technologies.

This idea of communication as being an evolving ecology is one that potentially provides a rich range of analytical possibilities. Kaufer and Carley (1993) provide a range of examples that explore possible aspects of this ecological perspective through modeling simulations. In this paper, I have chosen to explore communication by seeing what insights the study of real ecologies can provide in considering the changes now occurring in a specific area of communication; scholarly communication. In particular, I want to consider the process whereby evolution occurs and new species come into existence.

Punctuated equilibrium and speculation

Evolution in real ecologies occurs on a time scale that is too long for an individual to observe. If one wishes to look at how organisms in an ecology adapt and change, it is necessary to look at the past. Perhaps the most exciting new concept in the fields of paleontology and evolution in the last thirty years has been the idea of punctuated equilibration. This idea first proposed by Niles Eldredge and Stephen Gould (Eldredge & Gould 1972) suggests that rather than occurring gradually, evolution proceeds in fits and starts. They propose that the development

of new species occurs when parts of the breeding population become cut off from the rest of the population in different environments. They adapt to the new environmental challenges by evolving into new species that no longer need to change (Benton 1993, p. 33). If the surrounding environment then also changes, a few of the newly developed species are already pre-adapted and will quickly out-compete their ancestral relatives. In the fossil record, this then shows up as a sudden change from one organism to another.

Technology and ecological change

What might be the equivalent of such an environmental change within a scholarly communication ecology? I would like to suggest that changes in the scholarly communication space caused by new technologies are a good candidate. The most important of these technologies for scholars since the middle of this century have been the communication and computing technologies. These have provided the potential for transformations of both professional practices and intermediary processes. The stakeholders who have been affected are publishers, scholars (as both consumers and producers of content), scholarly societies (as representatives of scholars and as publishing intermediaries in their own right) and librarians.

Most recently, the rise of the Internet and its increased possibilities for electronic publishing have provided a new ecological niche that is being colonized by all sorts of communications products, scholarly journals included. Is it possible that the move to the new electronic environment might involve the development of new species of scholarly communication artifacts? The research presented in this paper sought in part to examine these changes and to identify new ecological niches that might be becoming available.

Research design

This paper examines these issues through the lens of five projects where a university library has taken on a new role within the ecology. Since the advent of printing, libraries have been traditionally associated with providing access to content created by other people. This access provision has involved a range of interrelated roles: acquisition, cataloguing and archiving primarily print materials. Few libraries have seen this role extending to publishing the material themselves or becoming actively involved in such publishing. The libraries chosen for this project are acting as significant facilitators for electronic scholarly publishing, rather than just as access points for material provided by a publisher. To examine a range of possible responses to this opportunity I designed my research to select library-facilitated projects that differed on a range of measures: organizational structure, geographical location, and type of published product.

The research methodology was to first investigate potential candidate projects via a structured literature and Web search. Once the candidates had been identified, I contacted the most suitable personnel via electronic mail, outlining my research and requesting their cooperation. I also asked if I should be talking to anyone else involved in the project. Once I had received agreement, I undertook intensive research into the project based on publicly available material. Each site was visited and senior informants (directors or assistant directors) from each project took part in a structured interview.

Project overviews

The geographical range of the projects selected ended up encompassing both coasts of the United States, the United Kingdom and Europe. No Australian projects (of which there are a number) were included in this phase; the funding arrangements for the initial phase of the research required that projects studied be located outside Australia. Given the extremely fluid nature of the electronic publishing field, specific comments about the projects are only valid as of the conclusion of this phase of the research (end of 1997). For those wishing more details on the specific projects themselves, in a previous paper (Treloar 1998a) I have discussed the projects more fully and concentrated on how the projects came to exist, how they are organized, what they provide, their level of commercial sustainability, the lessons learned by the participants about might be applicable to similar projects, and the likely future prospects for each project. An overview article summarizing this information will appear in the April, 1998 *Communications of the ACM* special issue on digital libraries (Treloar 1998b).

Highwire Press¹

Highwire Press is an initiative of Stanford University Libraries/Academic Information Resources. It was selected because it is implementing leading edge Web-based electronic journal technologies and because it is commercially successful (i.e. not just a pilot).

Highwire currently operates as a separate cost centre within the library, with the Publisher of Highwire, Michael Keller, also being the University Librarian, and Director of Academic Information Resources. At the time of writing the Highwire Press team listed on their homepage consists of 26 people (including support staff). The more senior Highwire staff also fulfil other positions within Stanford University. The project is currently commercially sustainable, in line with Stanford's policy of extensive charge-back for services. It is not however seen as a way for the university or library to make a profit. Rather, it is viewed as a cost-recovery exercise with both tangible and intangible benefits for the university. Because of its organizational location within the library and physical location in Silicon Valley, the infrastructure was already in place to ease the startup process.

Implicit in the Highwire Press mission statement is their intention to provide a model for re-engineering scholarly communication. To this end, they are working in partnership with scholarly societies to bring existing print journals online. The first of these, *The Journal of Biological Chemistry*² has now been joined by 24 largely biomedical publications which are available in both print and electronic form. An additional five or six **dozen** more titles will be available online within the coming year (consult their homepage for details).

These journals are at the leading edge of Web-based journal publishing and are progressively adding a number of additional value features that are only possible in an online environment. These include:

- direct searching by author, title keywords or text words, both within journals and across journals

¹ Available online at <<http://highwire.stanford.edu/>>

² Available online at <<http://www.jbc.org/>>

- display in PDF (best for printing) or HTML (best for navigation)
- automatic creation of hyperlinks to MEDLINE citations provided by the National Library of Medicines PubMed service
- links from Genbank accession numbers to full Genbank records
- bi-directional links between citing articles and cited references (where available) (Newman 1997)
- ‘toll-free’ links between the references from one journal article to the full text of the cited article” (Reich 1997).

Internet Library of Early Journals (ILEJ)³

The Internet Library of Early Journals (ILEJ) is a joint project between the Universities of Birmingham, Leeds, Manchester and Oxford. It aims to digitize a critical mass (defined as at least 20 consecutive years) of three eighteenth century journals (*Gentleman's Magazine*, *The Annual Register*, and *Philosophical Transactions of the Royal Society*) and three nineteenth century journals (*Notes and Queries*, *The Builder*, and *Blackwood's Edinburgh Magazine*). While not extremely rare, there are only perhaps 20-25 sets of each journal extant. The digitization will therefore need to be done on a non-destructive basis. ILEJ was selected as an example of a digitization project working with non-scientific and older serial material.

The project aims to explore the issues associated with making this sort of material available as well as providing access to it. The variables they are particularly interested in are image creation, indexing techniques and Web access to page images. A number of their working decisions have been made with an eye to reducing the cost of doing this sort of work as far as possible. The intention behind providing access to the material in digital form is to facilitate access by researchers (through desktop access and search mechanisms) and to reduce the need for physical handling of the originals. The project aims to mount 120,000 page images in all.

The project is not currently commercially sustainable, and may in fact never be given the nature of the material being digitized. The project team are actively considering the best way to proceed once the initial funding has been allocated.

The selected journals consist of a wide mixture of intermingled article types laid out on the page in a somewhat unstructured manner. This means that providing page images is the simplest way to present the material. Each image is between 100 and 200K in size. Pages can be browsed by volume/number/page or searched directly. The input to the search database is OCR text from the page scans. This OCR is done using off the shelf software (designed for modern typefaces and layout) and without manual intervention (because of the prohibitive cost of correcting errors). The result is quite ‘dirty’ OCR, which would normally present problems in searching. As a partial solution, the Excalibur Technologies EFS search engine is being used for full text fuzzy matching on the dirty OCR text.

Project Educate⁴

Project EDUCATE (End-user Courses in Information Access through Communication Technology) is a joint initiative of Limerick University in Ireland, the École Nationale des Ponts et Chaussées in France, the University of Barcelona in Spain, Chalmers University of

³ Available online at <<http://www.bodley.ox.ac.uk/ilej/>>

⁴ Available online at <<http://educate.lib.chalmers.se>>

Technology in Sweden, Imperial College of Science Technology and Medicine, and Plymouth University (both in the United Kingdom). The overall aim of the project is to help students, research workers and practitioners to develop their information literacy. EDUCATE was selected because it was publishing online teaching support materials (rather than journals) and because it provided a Nordic/European perspective.

In practice, Chalmers University of Technology Library provides the day to day technical and administrative management. Imperial College are doing some demonstration versions and courseware. Limerick did some of the interface design and Web development. Translation (and adaptation - using and linking to different resources) is being done into French and Spanish at Barcelona and Ponts et Chaussées.

The intention was that EDUCATE should be commercially sustainable immediately. There is a need for a revenue stream for two reasons: maintenance of existing materials and development of new ones. The experience of other electronic publishing projects has been that initial funding is easier to get than maintenance funding. The director of the project wanted to avoid this and has specified that new modules can only be developed with new money. The revenue stream for maintenance comes from license fees. These have been set as low as possible consistent with getting enough funds. The total from these fees is not yet at the break-even point but the trend is looking promising.

The main product of Project Educate has been a series of Web-based self-paced user education courses called Into Info. These provide training in the selection and use of information tools and resources in particular subject areas (chemistry, physics, electrical and electronic engineering, and energy to date). All the Into Info modules are based on a multi-level hierarchical structure with rich internal and external hyperlinking.

Project Muse⁵

Project Muse is an initiative of the Johns Hopkins University (JHU) Press and the Milton S. Eisenhower Library at JHU. It provides worldwide networked access to the full-text of over 40 of the Press's electronic journals. Project Muse was chosen as it is one of the very first Web-based electronic journal projects, and has a predominantly humanities coverage in contrast to Highwire.

Both the Library and the Press recognized that significant seed funding was required to start the project off. They made the decision to seek grants from the National Endowment for the Humanities (NEH) and the Andrew Mellon Foundation, and to use the Web (then very new) as the technology platform. These grant applications were both successful, and development commenced. In effect, the Press provided the content, and the Library added value through cataloguing, searchability and dissemination. The project is still a joint initiative and commercially sustainable; the grants finished at the end of 1997. The project is also starting to provide access to journals that do not come from the JHU Press.

Project Muse provides access to the full text of over 40 of the Press's scholarly journals (but not complete runs of each journal) in the humanities, the social sciences, and mathematics. About 3 to 4 outside journals will be added each year from now on. Most of these journals are

⁵ Available online at <<http://muse.jhu.edu/muse.html>>

published in print as well but some are electronic only. All online titles provide a range of features that are not available in print:

- Hypertext links in the tables of contents, endnotes, author biographies, and illustrations
- Text designed for on-screen reading and easy printing
- Boolean searches of either the full text or author/title/keywords from the tables of contents
- Subject indexing of each article using Library of Congress subject headings
- Illustrations that are both larger than those in the print version and often in color
- Subject, title, and author indexes
- Option to create online reference lists or electronic syllabi

Scholarly Communications Project⁶

This project, based in the library at Virginia Polytechnic Institute and State University (Virginia Tech), has been working on a range of publishing activities since 1989. It was selected because of the diversity of its activities, and because it operates in a service rather than cost-recovery environment. The project

“assists primarily Virginia Tech faculty who are editors of professional journals when they want to also make their publications available to their colleagues in distributed academic communities via the Internet. It assists traditional academic publishers adapt their publications to the Internet and access by the worldwide academic communities. It has also works with a variety of units within the university to extend access to their clients locally, regionally, nationally, and internationally.”⁷

Over time, the project has added more and more online publishing activities, and is seen by the university community as the experts in this domain. The funding for most project activities is provided as part of the library budget. Some funding for specialist activities comes from outside agencies. The project co-ordinator sees her role as a service-oriented librarian to say yes to any reasonable request from within the university and then work out how to resource it.

The digital products produced through the project now include:

- 12 electronic journals - (a number of these are electronic only);
- Virginia Tech newsletters and magazines
- some digital image collections (images from the Special Collections department, American Civil War Resources, and a History of Architecture Catalogue)
- the Electronic Theses and Dissertations initiative (providing access to PDF versions of graduate theses which will no longer be accepted by Virginia Tech in paper form)
- regional (the *Roanoke Times* and the *Virginian Pilot*) and international news
- electronic reserve system (also using PDF).

⁶ Available online at <<http://scholar.lib.vt.edu/>>

⁷ *About Scholarly Communications Project VPI&SU*. Available online at <<http://scholar.lib.vt.edu/about/scpabout.html#history>>

This fairly eclectic portfolio reflects the willingness of the project to respond to requests for assistance from their user community.

Changing ecological niches

As part of the detailed interviews carried out with members of these projects, two clusters of issues emerged that related directly to libraries and their changing roles in an evolving publishing ecology. These can be broadly defined as the nature of publishing and the core business of libraries. I have chosen to discuss themes arising from my interviews in an aggregated form without identifying the specific source of each comment. This is because it is the total picture that is important, not the source of each colour. I do not pretend that these responses are representative of the full diversity of views in the library community. They do, however, serve to provide useful pointers to some future directions.

Publishing

The questions that I asked under this topic were:

- How do you define publishing?
- Do you regard this just as the initial release of content or does it include keeping the content available long-term?

The key themes running through people's responses to these questions were the importance of access and the need for archiving.

Typical comments on the theme of access were that publishing is 'making information publicly available' or is 'the act of making information public'. One respondent with a commercial publishing background preferred to talk about shaping raw material into something that the market wanted, and then disseminating that shaped material.

Two of the respondents were reluctant to identify what they did as 'publishing', although the question did not require them to. One felt that publishing required some sort of distribution (which their project did not do), but acknowledged that this was an increasingly problematic distinction to make if one was referring to electronic documents on servers. Another respondent preferred to define the library function as providing access to material that might (or might not) be owned by the library. In both of these respondents I detected a disquiet about using a term like publishing that evoked a range of no longer relevant print associations. Both stated the need for a new word or model to describe what they did. One suggested the term 'digital press', but acknowledged that press had inappropriate connotations also.

The other key theme was the need to have a changed view of one's commitment to the content in an electronic world. Once print has been produced and sent out, there is no necessary ongoing commitment from the publisher. In an electronic content environment, there is now the need to provide ongoing access and maintenance. Such maintenance might just be keeping the servers up and connected or making sure that URLs still worked. It might also include adding forward references to existing content or improving the interface. In the longer term, it might extend to migration of the content as the presentation and delivery technologies changed. Most of the respondents expressed scepticism that traditional

publishers would take on this responsibility, particularly without an ongoing revenue stream. Most also explicitly identified the library as having had this archiving role in the past and needing to now evolve an equivalent as the technologies change. However, a number stated that there should not be an automatic assumption that the library would perform this archiving function, particularly given the possible costs.

Libraries role

The interview questions that related specifically to the libraries role were:

- What is a libraries 'core business'?
- Is this sort of initiative (i.e. the specific project) part of that core business?
- Can electronic publishing (in general) become part of a libraries core business?

In the answers to these questions, the key theme was the role of electronic publishing in fulfilling the university library's mission to its community. As part of this, once again the questions of preservation and access loomed large.

A number of projects identified what they were doing as explicitly part of their mission to provide access to information or supply information to their community. A number of respondents also emphasized the role of the university library in furthering teaching, learning and research. Providing a service to users was a common thread in the responses.

Based on this service orientation, all of the projects saw what they were doing as a natural outgrowth of their mission to their communities. Comments like '[this project] is a way to provide access outside the library' were typical. Despite the reluctance of a number of projects to identify what they were doing as publishing, all felt it fitted well into their libraries portfolio of activities. A number also stated the need for more such initiatives from other libraries. One respondent explicitly stated the need for libraries to provide alternative models to existing publishers in the new field of electronic scholarly journals.

All felt that there was no 'in principle' barrier to electronic publishing becoming part of the core business of all academic libraries, although this might depend on particular campus circumstances. One respondent put it best by stating that if it made sense for the library to do something and if users would expect this to live at the library, then the library should go ahead. Only one of the projects was based at a university with an existing press. This no doubt made it easier for the other projects to move into a new area without any precedent of ownership by another stakeholder.

Discussion

In his book *Wonderful Life*(Gould 1989) Stephen Gould discusses the fossil fauna of the Burgess Shale. These date back to 530 million years ago, just after the 'Cambrian explosion' - a time of dramatic diversification in evolution. The importance of the Burgess Shale is that it preserves many anatomical designs that existed for a while and then disappeared, never to be seen again. In accounting for this astonishing diversity, Gould points to three main factors:

- the first filling of the ecological barrel, providing empty ecological niches that life could rapidly colonize;
- a directional history for genetic systems, making significant change harder over time;
- early diversification and later locking as a property of all systems.

If one considers the current state of electronic scholarly communication, it displays many of the characteristics of the Cambrian explosion. The new ecological niche is that of online publishing, free from many of the constraints of the print world. I am reluctant to push the analogy too far, but occasionally in moments of whimsy I like to think of the first ejournal as being a little like the first lungfish pushing its way arduously up some primeval beach on its way to start the colonization of the land.

It is certainly true that the last decade has seen a great diversity of forms of electronic scholarly communication: ftp-based journals, mailing lists, journals on CD-ROM, Web-based journals, MUDs/MUSHs/MOOs as collaborative and publishing spaces, and proprietary SGML-based journals. We are starting to see some of this diversity being narrowed down as particular forms are abandoned and as the scholarly world standardises on a subset of these early experiments. Early proprietary experiments like the OCLC Guidon interface have been abandoned in favour of open standards like the Web. Many electronic publishers are choosing to standardise on Acrobat for parallel print/electronic delivery, or where the formatting requirements make HTML problematic.

In terms of the theory of punctuated equilibrium, I suspect that we are in the middle of a punctuatory jump after the stasis of the last century (at least) of print journal publishing. New species of communication artifacts are emerging to fill the new online niches. Existing players in the scholarly communication ecology are changing their roles and evolving within the new environment. If the analogy with punctuated equilibrium holds, then a new period of stasis should be anticipated. My (relatively safe) prediction is that the overwhelming majority of print scholarly journals will have completed their transition to an online existence by 2010 and we will then see another (although perhaps shorter, due to the pace of technological change) period of relative stasis. In the move to this new online status quo, there is the potential for new players to emerge and existing players to redefine their roles. The library projects discussed in this paper are at the forefront of redefining what it means to publish and what a libraries role in the publishing process can be.

Bibliography

Alvarez, L. W. , Alvarez W., Asaro, F. and Michel, H. V. (1980). "Extraterrestrial cause for the Cretaceous-Tertiary extinction" *Science*, v. 208, pp. 1095-1108.

Benton, Michael (1993)). "Life and Time"" in Gould, Stephen (ed.), *The Book of Life*, Random House, Sydney.

Carley, Kathleen (1995) "Communication Technologies and their Effect on Cultural Homogeneity, Consensus, and the Diffusion of New Ideas", *Sociological Perspectives* 38, no. 4 (Winter): 547-571.

Eldredge, Niles and Gould, Stephen Jay (1972). "Punctuated Equilibrium: an alternative to phyletic gradualism", pp. 82-115 in T. J. M. Schopf (ed.), *Models in Paeleobiology*, Freeman, Cooper and Company, San Francisco.

Gould, Stephen Jay (1989). *Wonderful Life: the Burgess Shale and the Nature of History*. W. W. Norton and Company, New York

Kaufer, David and Carley, Kathleen (1993). *Communication at a Distance: The Influence of Print on Socio-Cultural Organization and Change*, Lawrence Erlbaum and Associates, Hillsdale, NJ.

Kauffman, Stuart A. (1993). *The origins of order*, Oxford University Press, New York.

Newman, Michael (1997). "The Highwire Press at Stanford University: A Review of Current Features." *Issues in Science and Technology Librarianship*, Summer. Available from World Wide Web: <<http://www.library.ucsb.edu/istl/97-summer/journals.html>>

Reich, Vicky (1997). Personal Communication via email, December 4.

Treloar, Andrew (1998a). "Technology as Agent for Transformation: Five Case Studies of University Libraries as Facilitators for Electronic Scholarly Publishing", *Proceedings of VALA '98*, Melbourne, January. Available from World Wide Web: <<http://www.deakin.edu.au/~aet/Research/Conferences/VALA98/>>

Treloar, Andrew (1998b). "Five Case Studies of Libraries as Facilitators for Electronic Scholarly Publishing", *Communications of the ACM*, April.