Electronic Publishing '01 –
2001 in the Digital Publishing Odyssey
A. Hübler et al. (Eds.)
IOS Press, 2001

Human and Economic Impacts of Electronic Publishing at the NRC Research Press : A Case Study

Scott Mellon and Aldyth Holmes

NRC Research Press
Canada Institute for Scientific and Technical Information
National Research Council of Canada
Ottawa, Ontario K1A 0S2 Canada

Abstract
The NRC Research Press has a seventy-year history of provision of high-quality scientific journals, currently totalling 14 titles in various areas of science and technology. These titles were made available to our clients electronically in 1996, [see http://researchpress.nrc.ca] and we are currently engaged in a second phase of what will likely be an ongoing process of technical evolution. This paper focuses on the impacts of such rapid change on our organization and people, and discusses change management principals, which have been found to be effective. Such impacts include the recruitment, and retention of appropriately qualified staff and the costs associated with each; and the design of space to facilitate new working styles. We will explore the "build-or-buy" conundrum and present our current philosophy on this issue. Strategies such as matrix management and out sourcing for managing technology workers in a hitherto non-technical operation; and techniques for learning and transferring technology to and among staff will likewise be addressed. This paper will also discuss our co-evolution with our clients, the journals' readers, authors and editors, and ways in which their development and changing expectations have informed our products, systems and processes.

0. Introduction

The National Research Council of Canada (NRC) began publishing journals in 1929 in response to a perception that the work of Canadian scientists was not being adequately disseminated because Canadians had limited access to the scientific journals of the time. The situation of a country new to original scientific research was much like the situation that new disciplines find themselves in today, viz. lack of an appropriate venue for publication. The major journals have never had more than very limited interest in new disciplines or in the science that is perceived to be only of regional interest. From those early beginnings, the NRC has grown in stature and become a recognised major research institution on the world stage. As NRC grew in international prestige so did the journals being published by them. Today the NRC publishes fourteen internationally recognised journals and is launching a fifteenth this fall. These journals attract authors from around the world and are distributed both in paper and electronically to over 12,000 subscribers world wide.

One of the factors that has given us a unique view of the world for a not-for-profit publisher is that we publish journals in a wide variety of disciplines. As we will show, the different disciplines have exhibited quite different behaviour patterns in their acceptance of electronic journals.

The decision to become an electronic publisher was made in 1994 and implemented over a three year period. It was recognised at the time of the initial decision to produce electronic versions of the journals, that the technology available at the time would not meet all the suggested possibilities for enhancing the journals and that any choices of technology would necessarily be transitional to new better technologies as they became available. The decisions made had to be justified from a business perspective and the implications for staff had to be taken into consideration.

0.1 An outline of the changes (background)

In 1994 our production process was to receive hard copy of a manuscript previously accepted and peer reviewed by our external editorial offices, and then assign a scientific copyeditor who marked up the changes to improve the quality of the English, removed any ambiguities in the text (querying authors as necessary), and changed the style to conform to the house style. This would include for example the use of SI units and the use of italics in conformance to the standards for the various disciplines we publish in. (Council of Biology Editors standards for expression of taxonomies, IUPAC for naming of chemical compounds etc.) The marked up paper was then sent to a typesetting house, and a galley proof was generated and sent to the author. The author and the copyeditor then proofed the galley. The authors' changes were reviewed by the copyeditor and transferred with the copyeditors corrections to the typesetter. This marked up document was sent to the typesetting company and final page proofs produced. These were then sent to the printer who used camera (or in some cases, electronic files) to produce the plates for the press and the blue lines for the copyeditors to review.

The first electronic versions of the journals were produced from PDF files created in-house. That is to say we replaced the contracted out typesetting function by an in-house group. When we established this group we asked authors to send us diskettes of the accepted paper to reduce our need to re-key whole papers. We were surprised to find that we were able to get the disks from nearly all authors, wherever they were in the world. We did have some interesting conversion challenges in the early days. But we could nearly always at least get the ASCII code for the text of the paper if not for the mathematics. Today we have far fewer problems. There are better conversion programs available and authors are tending to use a limited range of major software packages to prepare their material.

Our choice of software was pragmatic, based largely on what staff would find easiest to work with. WordPerfect was selected for on screen editing because it was the NRC office standard at that time. In fact that changed less than a year after we began the project when our parent organisation switched to Microsoft Word. WordPerfect had the added advantage that the tags added by the copyeditor could be displayed on screen and if necessary deleted, leaving the copyeditor with a sense of control over the text.

Our decision on a layout package was equally pragmatic. We had researched various packages including Ventura and Quark Express, but our final decision was driven by the arrival of an NRC employee whose job in another part of the organisation had disappeared; they were available and they were expert in Ventura, which we felt was adequate. The choice of Ventura saved us training dollars and when Corel acquired WordPerfect we thought we would probably have two packages that would work together well.

1. Impacts of Change

1.1 Human Impacts

Staff working in our current production system have a much greater variety of tasks in each job than they did in the manual system. Essentially the tasks of pre-marking, and page layout merged. We have worked to reduce the number of times a paper goes around the review/author approval cycle - it had been as many as four, but is now seldom more than one. We have also shortened the approval process by FTP'ing PDF files to authors instead of relying on surface mail.

Copyediting is still a separate function because of the requirement for scientific knowledge. However due to the multi-disciplinary nature of our journals, and inevitable staff changes, the level of such knowledge is not high enough to allow for substantive editing. (Our process assumes that any substantive editing required has already taken place within our volunteer community editorial office as an adjunct of the peer review process.) Still there was a major shift in workstyle by copyeditors, for whom working on screen was new, and to some unsettling.

We struggled initially with the correct supervisory structure. We introduced the program with a Head of Journals and three Senior Copyeditors who were responsible for the production of clusters of journals in related subjects. As we added staff to replace the contracted out typesetting, they reported to the senior copyeditor whose journals they worked on. Initially we added a layout person who handled the final page layout and the graphic images and another position called pre-marking. The pre-marker converted the author's version of the accepted paper to WordPerfect, stripping out all formatting and began re-formatting to our house style before passing it to the copyeditor. That is to say senior copyeditors supervised staff who did purely technical tagging and layout. This was felt to be ineffective use of their time, and has since evolved into a separate Production Unit with its own Head and structure.

We have also struggled with the correct structure for technical support. NRC Research Press is fortunate in being part of the National Research Council of Canada, and of the Canada Institute for Scientific and Technical Information, so that access to leading-edge technologies is readily available. But how to manage it in this particular situation? After some experimentation, we have determined that our computer support is best managed outside the Press, which while still within the Council, effectively means, outsourced. This has several advantages: we are able to obtain access to fractions of people with a wide variety of skill sets, and the people involved are professionally managed by other computer professionals. This latter issue implies that they have a career path, a succession plan, and hopes for promotion.

Our ready access to technology can sometimes be a two-edged sword however, as the technology has occasionally threatened to become our driver. That is to say, we sometimes need to fight off a "Technology for the sake of technology" syndrome, and in particular, the infamous "Not-Invented-Here" syndrome. We constantly need to remind ourselves and others that we are in the business of publishing, and not the business of software-for-publishing. In general, our bias is to buy technical solutions whenever possible.

1.2 Economic Impact

Electronic publishing was introduced for 2 journals in 1996. Five further journals were converted to this process in 1997 and the remaining 7 journals in 1998. The following table shows our cost breakdowns by function from 1994 to 2000. The most interesting changes are in the indices for copyediting, printing and overhead.

Table 1
Index of Average Publishing Costs per Page for 14 Journals

	1994	1995	1996	1997	1998	1999	2000
Copyediting	100	82.55	88.77	118.80	128.21	102.43	87.42
Peer Review	100	112.14	115.43	127.23	137.12	135.98	135.53
Printing	100	120.07	92.16	110.23	68.35	48.95	61.38
Overhead	100	83.83	141.25	107.96	181.77	184.06	166.31
pages	100	103.24	102.84	96.40	92.47	104.04	109.95

Copyediting: The cost of copyediting actually increased in the first two years of electronic publishing and is now sinking back to previous levels, even though more of the actual production process is now done at this stage, and even though no adjustment was made for salary increases. We believe that this is evidence of the cost of the "learning curve". The transition can be clearly seen in 1998 when all journals were finally being produced electronically.
Printing: Printing costs appear to have dropped dramatically; however this is in part a statistical illusion. Prior to the introduction of electronic publication the typesetting function was contracted to the printers and appears as part of the printing index. Once the process was re-engineered and the production of the electronic file was done in-house the costs became part of the copyediting budget. The increase in printing cost in 2000 is most probably attributable to two factors: the reduction in print runs as electronic subscriptions began to replace paper subscriptions which tended to increase the average cost per page; and the change in policy for two journals which began to offer free colour for authors. This had the result of significantly increasing the amount of colour published with a corresponding impact on the average cost.
Overhead: Increases in overhead costs in 1998 and 1999 are attributable to technology upgrades.

1.3 Impact on Production Time

Similarly, as the following table shows, our average production time also typically increased in the first year or two of our new system, before any real long term improved could be detected.

Table 2
Average Days per article spent in production process

(Year electronic)	1993	1994	1995	1996	1997	1998	1999	2000
Journal 1 (1996)	124	124	121	118	115	186	119	94
Journal 2 (1997)	189	176	191	157	196	256	293	220
Journal 3 (1998)	131	143	135	149	155	199	214	115

Journal 1 was a monthly biomedical journal and was one of two journals used as prototypes to develop the processes for electronic production. Additional staff were employed and trained on the layout as part of their training to work on other journals. This explains the drop in time in the year the new processes were introduced, additional staff training on this title continued the following year but the benefits of the first years leaning show clearly in the second year of electronic production. However when the staff allocated to this title were reduced to normal levels i.e. no new staff assisting in its

production as part of their training, the time peaks and then recovers as learning is consolidated.

Journal 2 was a bimonthly engineering journal with more complex graphics than the two prototype journals but again a similar pattern emerges. As additional staff have been previously trained on journal 1 are assigned and a new staff added to train for this title the time increases and when the more experienced staff member is withdrawn and reassigned the learning curve goes up sharply to begin declining the following year.

Journal 3 is a large prestigious monthly biology journal. The pattern repeats itself.

The implications of these data are that delays or increase in production time could have been reduced by acknowledging that it took not one but two years for staff to reach an acceptable speed of work with the new technology. This implies that the transition was under-resourced, and that inadequate funds were available to double staff positions for that length of time.

1.4 Technology Impacts

There was much discussion when the decision to pursue electronic publication was taken as to the appropriate software packages to consider and the desirability of using SGML. Two studies were undertaken two years apart asking the question how much would it cost to go to an SGML base workflow and what would be the benefits from it. In both cases the cost of introducing SGML did not justify the benefits. The underlying problem for any science journal was the difficulty of handling mathematics in SGML and the cost of developing DTDs and all the associated staff retraining for implementation. While we could see the potential value of having SGML tagged files in the long term in terms of the ease with which they could be converted to newer technologies and could be archived we could not find short term benefits from the perspective of the author and reader. The original decision was, therefore, to go with a PDF format recognising that this would be transitional to some other format at some future date. This responded to the clients needs at that time. Scientists five years ago did not want to read on screen. They wanted to locate the article and have it delivered to the desktop where they wanted to print it out so that they could read it and annotate it just as they had done with paper. We fully expected this behaviour to change as the nintendo generation joined the research community but we did not know what the time scale for the changed demands from authors would be. Moreover, we could not afford to be financially exposed by investing in unproven technologies. We were content to observe the major scientific societies who had deeper pocket experiment and observe the behavioural impact on the scientists of these experiments. At first we thought the need for full text search capability would force us to abandon PDF very quickly but early consultations with our colleagues in NRCs Institute for Information Technology assured us that full text searching within PDF would be available very quickly. They were correct. By the time we had the first seven titles in PDF the products were on the market.

2. Lessons Learned

It is clear that we underestimated the cost of the learning curve, and the time required for staff training and development. In fact the "learning curve" which was originally seen as mostly a matter of learning software, ultimately included a significant change in both corporate culture and required skill sets. Editors in particular have had to learn a new language, and they need much more technical knowledge of the production process, including, for example, text and graphics formatting. The change has also required an ongoing search for process and workflow improvements. The effect of that

underestimation was seen in publication delays, and consequently in dissatisfied authors and editors. Any future major changes in our production systems will need to take this into account, and appropriate resources found.

Another important lesson, we feel, is that our re-engineering has brought us benefits which we would not have been able to predict. For example, our new ability to send PDF files to authors has significantly reduced time-to-publication, even though it was invented for quite another purpose. Similarly we are seeing intangible but very real benefits which flow out of our improved position within the Canadian publishing community in particular, and hope to be able to share the benefits of our experiences with those colleagues.

3. Next Steps

Five years after the original decision to go electronic we are starting to see the behavioural changes that we predicted. Scientists are acquiring higher quality monitors which make reading on screen less demanding. Along with better desktop technology for the subscribers came increasing pressure to make more use of the electronic world. Furthermore, expectations of readers and authors have continued to evolve, and many features not possible in PDF files, from reference linking to Virtual Reality simulations, have become quite normal in some disciplines. We have therefore launched a major project to update our production system yet again. Our goals for this project are first, to continue to shorten the production cycle in response to continuing pressures from some of our communities; second to add value to our online products by giving us the capabilities available through HTML, and ultimately online XML articles; and third, to "future-proof" our archival data by storing them in formats which will have long term reliability. In so doing the lessons of our first major foray into electronic publishing will be very valuable to us, and I hope, to others.