elpub2002 Proceedings — J. A. Carvalho et. al. (Eds.) — VWF Berlin, 2002

Re-Engineering the Scientific Publishing Process for the "Internetworked" Global Academic Community

Bo-Christer Björk[1], Ziga Turk[2], and Bob Martens[3]

[1] Swedish school of Economics and Business Administration
Helsinki, Finland
[2] University of Ljubljana, Slovenia
[3] Vienna University of Technology, Austria

Abstract. The SciX (Open, self organising repository for scientific information exchange) project is funded by the European Commission in order to demonstrate the feasibility of new alternative models of scientific publishing made possible by the Internet. The project builds upon the previous experience of some of the partners in running an electronic peer reviewed journal and in setting up an e-prints archive. The project includes both theoretical work in making a formal model of the scientific publishing process, to be used as a basis for studying the life-cycle costs of alternative business models, and a demonstrator of a functioning e-prints archive.

1 Introduction

The Internet has dramatically changed both the technical and economic infrastructure of publishing. One area in which this has caused a lot of tension is the publishing of scientific publications, where the material for all practical purposes is a public good. This means that it would be in the best interest of both the researchers and the organisations that fund the research to make the publications available for free. Nevertheless for historic reasons the actual publication process has to a large extent been outsourced to commercial publishers, who currently take advantage of the substantial lock-in to their brand journals, in selling electronic versions at prices which do not differ much from the subscriptions to the paper versions. In dealing with digital information they can effectively use all sorts of strategies typical for e-commerce with information goods, such as bundling and differential pricing [Shapiro and Varian 1999].

This situation has over the last decade led to a "subversive movement" among scientists to start their own journals and e-print servers, which offer the material for free over the Internet. This movement, which lacks a commonly used label, can in many respects be compared to the "Open Source Movement" for the production of software applications of global interest [Raymond 2001], which also was made possible by the Internet. There are however still many psychological, legal and institutional barriers to change the process and the behaviour of academics and these have been underestimated by the pioneers and enthusiasts. Consequently only a small part of the overall volume of the scientific communication process has so far been affected such efforts. One reason for this is that the majority of such efforts haven't been based on sound and sustainable "business" models. Thus the mortality rate of such journals has been rather high [Wells 1999].

Quite a lot has been written about experiences with free electronic journals and archives, but most of this is anecdotal evidence based on individual cases. A crucial unanswered question is whether the cost of producing a free Internet based peer reviewed journal is substantially lower than for a traditional paper based one. Depending on which camp the respondent belongs to the answer could be anything between a small fraction and as much as for the paper alternative [cf. Tenopir and King, 2000 and Odlyzko, 1998]. The answer to this question could be a major parameter for determining weather the public authorities financing research and library activities should start financing free scientific publishing efforts, for instance using a business model were a hub for running several refereed journals and or e-prints archives is offering its services for free for non-commerial use. The authorities could also change their attitude towards the researchers that they fund surrendering their copyright rather unconditionally to commercial publishers.

The SciX (Open, self organising repository for scientific information exchange) project, which is financed through the IST programme of the European Commission, aims at demonstrating that the Internet enables new business models for the scientific publishing process which are much more cost and time efficient to the scientific community than the current practice. The SciX project group will create services on the Web that will enable scientists as well as practitioners from the fields of architecture and construction easy and free access to relevant research publications. In addition existing publishing practices will be analysed systematically and business models for re-engineering the scientific publishing process will be developed, taking into account also the legal, social and psychological barriers to change.

2 The scientific publishing life cycle model

In one of the subtasks of the SciX project a formal process model of the scientific publishing process is being developed. The aim of the modelling effort is to understand the scientific publishing process and how it is affected by the Internet, in

order to provide a basis for a cost and performance analysis of various alternative ways of organizing it. The modelling is done using a graphical modelling language called IDEF0 [IDEF0, 1993]. The main concepts are the activity and the flow. The flow can be used as input, output, control or mechanism. The presentation of the IDEF0 diagrams is hierarchical in a way that diagrams on lower levels are more detailed than those at top, Table 1.

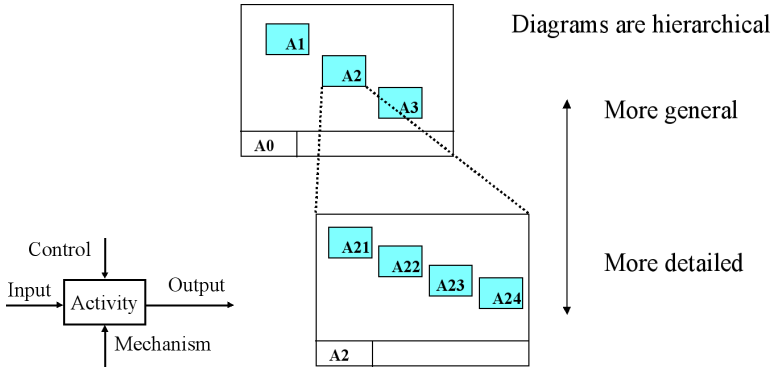

Table 1. The basic concepts of the IDEF0 method [Karhu, 2000]

The SPLC-model explicitly includes the activities of all the stakeholders in the overall process, including the activities of the:

- *Researchers* who perform the research and write the publications;
- *Publishers* who manage and carry out the actual publication process;
- *Academics* who participate in the process as editors and reviewers;
- *Libraries* who help in storage and indexing of the publications;
- *Bibliographic services* which facilitate the identification and retrieval of publications;
- *Readers* who search for, retrieve and read publications;
- *Practitioners* who implement the research results (in-) directly.

In the model the unit of observation is the single publication, how it is written, edited, printed, distributed, archived, retrieved and read, and how eventually it may affect practice. The viewpoint taken is life-cycle cost per publication. Thus at later stages all cost and time data which is collected will be translated to a per publication basis. The aim is to clarify the process and to find a way of measuring the total life-cycle cost of a publication, since the objective of the reengineering efforts in the SciX project is to try to optimise the total life-cycle costs, rather than the cost of some particular stage.

The current version of the SPLC-model includes 19 separate diagrams, arranged in a hierarchy up to seven levels deep. There are altogether 58 activity boxes and more than 100 input, control, mechanism or output arrows.

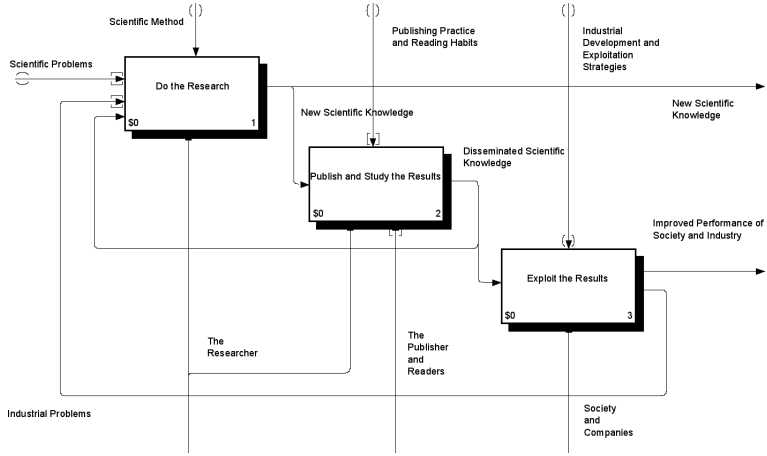

Fig. 1. *Main diagram of the Scientific Publishing Life Cycle Model (SPLC)*

The philosophy of the main diagram (fig. 1) is to show how science can help in solving problems of everyday life by a combination of research (which creates scientific knowledge) with publishing (which ensures that results help in the advance of our collective body of knowledge) and exploitation activities. The main stakeholders in the process are shown as mechanism arrows coming into the activity box from below, and the main drivers controlling the behaviour of the stakeholders are shown coming in from above.

Figure 2 tries to clarify the dual nature of the publication process. From the perspective of society it is the efficiency of the total process, including both the production, archiving and "consumption" of publications, that should be optimised. Optimising only one of these may lead to a sub optimal solution for the total process.

Figure 3 shows the main stages of the publication process, which to a large extent are carried out by different stakeholders. Based on the results of his research, the researcher writes a manuscript, which then in the next stage through a number of transformations is changed into a publication (on paper or electronic). The last activity is extremely important from a life-cycle viewpoint and involves the archival storage of the publication in research libraries all over the world, as well as value-added services through bibliographic services.

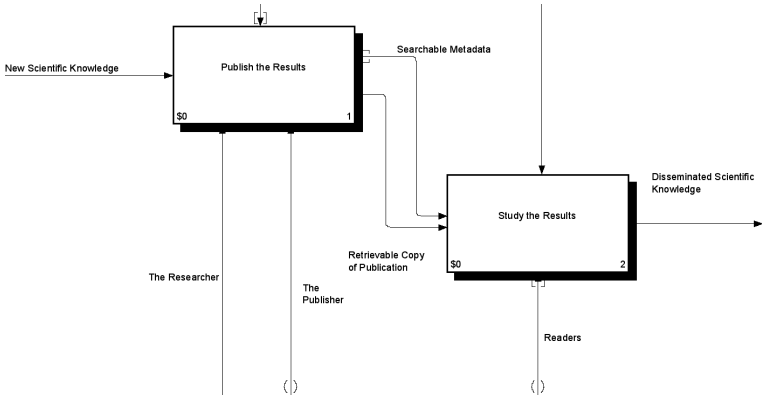

Fig. 2. *Publish and study the results*

Figure 4 structures the activities of the readers of scientific activities. Note that from a cost per publication perspective the activities of individual readers all over the world and in different time periods should be summed up. The "Find out about Publication" activity results in the output metadata of interesting publication (including the location from which a paper or electronic version can be retrieved). This output is used as the control of the retrieve publication activity.

Unfortunately the space limitations of this paper do not allow a detailed presentation of the overall model. The interested reader is advised to go to the website of the SciX project (www.scix.net), where the current version of the model will be

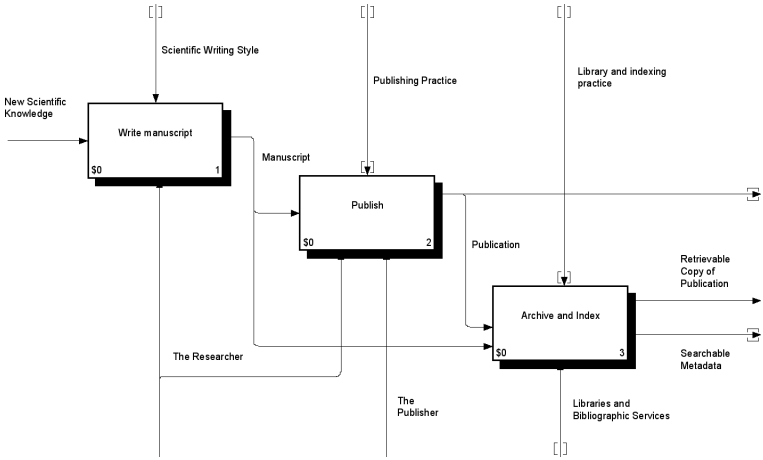

Fig. 3. *The main stages of the publication process*

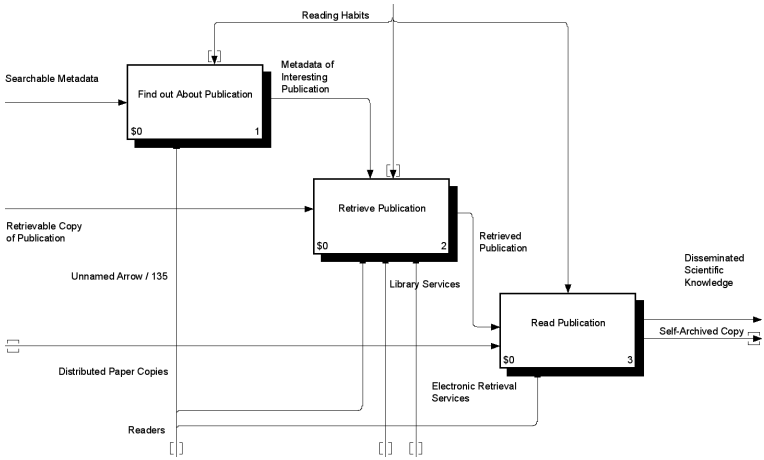

Fig. 4. *Study the results*

made available. Figure 5 is presented below to give a flair of a diagram from one of the more detailed levels.

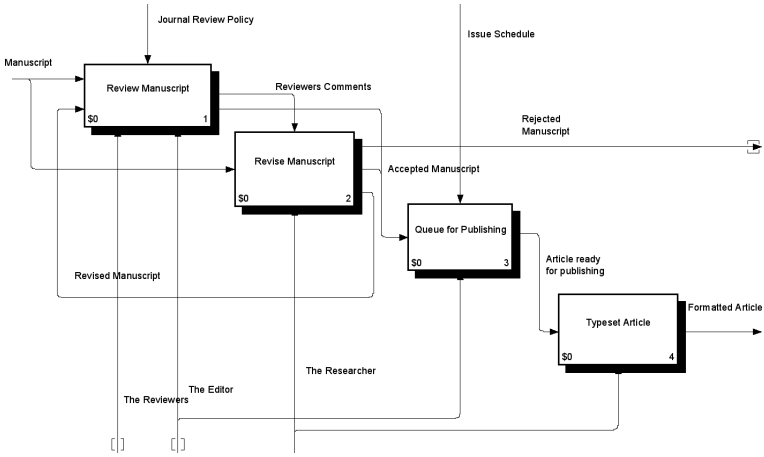

Fig. 5. *Article-Specific Activities*

Work on this model continues, and it will be validated through interviews with experts such as editors of journals. The modelling tool allows the inclusion of cost and duration data for activities. An attempt will be made to assemble cost data (on a per publication basis) from various sources, including surveys with editors

of electronic journals. This will help in determining the life-cycle cost of various publishing business models (paper, electronic, free, subscription etc.).

3 The SciX repository

The main output of the SciX project will be a demonstrator of an e-prints archive for a particular community of practice (researchers and practitioners interested in the application of IT in architecture and construction). A survey made by Björk and Turk [2000] showed that at that time the researchers in this domain accessed 50% of the scientific publications they read over the Internet, and that as readers they strongly prefer material that is freely accessible to subscription based.

Since 1999, two of the partners in the project (Bob Martens and Ziga Turk) have been managing CUMINCAD – Cumulative index of CAD – the largest freely available database of papers related to computer-aided architectural design, particularly related to the education in this area [Martens et al, 2001]. In the framework of annual conferences organized by regional CAAD-Associations (ACADIA in North America, eCAADe in Europe, Sigradi in South America and CAADRIA in Australasia) thousands of papers have been published. Rarely were the proceedings published by a professional publisher, therefore, the texts were neither entered into commercial indexes, nor were they sold commercially. The full texts were not broadly available; only conference attendees had copies. On the other hand, the associations retained in most cases the copyright to this work and could therefore allow its publication/archiving in the CUMINCAD. Thus this work is available on the net and rescued from oblivion. At the time of writing, CUMINCAD includes 4007 papers with abstracts. 883 papers are available in full text as well (`http://itc.fgg.uni-lj.si/cumincad/index.htm`).

The work in SciX will build on the database of papers assembled in CUMINCAD as well as the experiences of it as well as with the Electronic Journal of Information Technology in Construction (`http://itcon.org/`). The aim is to create a repository which on one hand is very cheap to run and maintain, due to a high level of automation, and at the same time offers potential readers added value, firstly by containing a critical mass of papers in the research domain as well as added services.

The aim is to make the SciX repository available either as an open source application or a low cost service which other communities of practice could use for outsourcing the IT-infrastructure they need.

Acknowledgements

The presented work has been conducted in the context of the SciX project, funded by the European commission under the contract IST – 2001 – 33127. SciX is a 24

month project with an EU funding of 1.000.000 euro. Co-ordinated by the University of Ljubljana (Slovenia), the partners include Swedish Business School of Finland, Icelandic Building Research Institute, an eBusiness company Indra/Atlante (Spain), Vienna University of Technology (Austria), FGG Institute (Slovenia) and the University of Salford (UK).

The home page of the SciX project is at `http://www.scix.net/`.

References

Björk, Bo-Christer, Turk, Ziga: 2000, How Scientists Retrieve Publications: An Empirical Study of How the Internet Is Overtaking Paper Media, in *Journal of Electronic Publishing*, Michigan University Press, Vol. 6/2. 2000, `http://www.press.umich.edu/jep/06-02/bjork.html`

Martens, B., Turk, Z. and Cerovsek, T.: 2001, Digital Proceedings: Experiences regarding Creating and Using, in *Architectural Information Management* (eCAADe Conference Proceedings), Helsinki, pp. 25-29.

Karhu, Vesa: 2001, A generic construction process modelling method – A model based approach for process description. Ph.D. thesis, Royal Institute of Technology, Stockholm, Sweden.

IDEF0: 1993, *Integration definition for function modelling (IDEF0)*. Federal Information Processing Standards Publication 183.

Raymond, E.: 2001, *The Cathedral and the Bazaar*, O'Reilly, Sebastopol, California `http://www.tuxedo.org/~esr/writings/homesteading/cathedral-bazaar/html`

Shapiro, C., Varian, H.: 1999, *Information Rules A strategic guide to the Network Economy*, Harvard Business School Press, Boston

Tenopir, C. and King, D.: 2000, *Towards Electronic Journals – Realities for Scientists, Librarians and Publishers*, Special Libraries Association, Washington D.C.

Odlyzko, A.: 1998, The Economics of Electronic Journals, in *Journal of Electronic Publishing*, Vol. 4/1 `http://www.press.umich.edu/jep/04-01/odlyzko.html`

Wells, A.: 1999, *Exploring the development of the independent, electronic, scholarly journal*, M.Sc. in Information Management, Electronic Dissertations Library, Department of Information Studies, University of Sheffield, UK `http://panizzi.shef.ac.uk/elecdiss/edl0001/index.html`