The new culture of electronic publishing

Peter Schirmbacher
Humboldt-Universität zu Berlin
Computer- und Medienservice
Institut für Bibliotheks- und Informationswissenschaft
E-mail: schirmbacher@cms.hu-berlin.de

Abstract

Local conditions in the past often limited opportunities for scholarly exchange. But now these limits are gone and the global workplace has replaced them. It is important to react to these changes. Every academic department must now adopt new methods and rethink processes. Another is the intense national and international debate about open access to scholarly knowledge. The Open Access Initiative shows that a change is taking place in the communication process. This change is also important for service departments within research institutions. Libraries, computer centers and related units have to ask themselves how to react appropriately to the new conditions. What services must be changed or redeveloped, and in what quality and quantity should they be offered? This article focuses on changes in the scholarly publication process. It describes both technological changes and the changes needed in people’s attitudes.

Keywords


CHANGE IN THE SCHOLARLY COMMUNICATION PROCESS

The process of scholarly communication has experienced a major change as a result of the new opportunities in information and communication technology. The quantity and quality of this change is already well known*. Recent data shows, for example, that Humboldt-Universität zu Berlin has more than 9,400 computers in offices and each machine has Internet access via local networks. There is no operation in the whole university that is not computer-aided in some way. The number of people with computer accounts corresponds to the total number of students and staff members at the university.

The status of information technology inside the university has reached new dimensions. While in the past computers were more or less add-ons to support individual study and research, the situation has changed to such an extent that present-day operations depend on the reliability, efficiency, stability and security of computer technology. Accordingly, IT has become one of the most important elements inside the university, and a major factor in determining the speed and direction of both research and teaching.

But this fact represents not only the number of computer-aided operations, the convenience of network-based services, or the variety of computer applications found on different classes of computers — it also indicates how thoroughly the university’s working processes have been steeped in computer technology.

At the official opening of the Erwin Schrödinger-Zentrum at Humboldt-Universität, today’s workplace for scholars and students was characterized in a speech as follows**:

– It is computer-based, with a technological capacity that ranges from an ordinary PC to High-Performance Computing.

---


** Schirmbacher, Vortrag zur Einweihung des Erwin Schrödinger-Zentrums der Humboldt-Universität zu Berlin, 19.05.2003; http://www.cms.hu-berlin.de/schirmbacher/ika-einweihung030519/index.html

The workplace is characterized by multimedia-based technology. Graphics, moving images, speech and computer simulations, etc., are used – in addition to classic paper methods.

The workplace is increasingly mobile. This means not only supporting university offices, but making all services available via networks, including availability at homes or remote conference locations.

The scholars and students take it for granted that they will have access to knowledge worldwide – not only to literature in locally accessible libraries.

Electronic communication is seen as natural.

The text of the grant program of Bundesministerium für Bildung und Forschung (BMBF) dated 28.10.2004 states the following under the heading e-Science and networked knowledge management: “The BMBF has referred in its policy paper “Information vernetzen - Wissen aktivieren”* [i.e. “network Information - activate knowledge”] to the need for innovative information management for the future of research and scholarship. In the course of the globalization and continuous networking of scholarly activities, international cooperation has become a key paradigm.” One can see “a completely new dimension of scholarly communication and research. For this vision the international term “e-Science” was established.”

Today the well-established exchange of scholarly knowledge via books, journals and conference presentations, which later appear as conference proceedings, is supplemented and partly replaced by electronic media. Email now often replaces letters to colleagues. Preprint servers, mailing lists, and electronic discussion forums gives a whole new quality to the process of sharing article drafts.

Local conditions in the past often limited opportunities for scholarly exchange. But now these limits are gone and the global workplace has replaced them. It is important to react to these changes. Every academic department must now adopt new methods and rethink processes. The preprint server of the Institute of Physics is one example** (see also*** or****). Another is the intense national and international debate about open access to scholarly knowledge*. The Open Access Initiative** shows that a change is taking place in the communication process. This change is also important for service departments within research institutions. Libraries, computer centers and related units have to ask themselves how to react appropriately to the new conditions. What services must be changed or redeveloped, and in which quality and quantity should they be offered? This article focuses on changes in scholarly publication process. It describes both technological changes and the changes needed in people’s attitudes.

THE CULTURE OF PUBLISHING

Without doubt there are many definitions of the term “culture”****.****. First in terms of the cultivation and reclamation of land, culture today may be regarded as the totality of human creations and relationships. In the course of evolution people have made explicit or implicit rules for living together, including language, tools and rules of conduct. O. Wegge[*****] writes: “Kultur ist die Gesamtheit der erlernten Verhaltensweisen und der übernommenen Einstellungen, Wertesysteme und Kenntnisse (…), die von Mitgliedern einer Großgruppe geteilt und geerbt wird. Kultur ist sowohl Ausdruck als auch Bedingungsstruktur für das Verhalten der Mitglieder einer bestimmten Gesellschaft.” (Culture is the totality of learned rules of conduct and of adopted attitudes, value systems and knowledge (…) that is shared and bequeathed by the members of a group of persons. Culture is both the expression of as well as the rules for the behavior of members in a specific society. – translation by the author) If one accepts this approach to the term culture, then it is surely legitimate to speak of a culture of scholarly communication and scholarly publishing. The “culture of publishing” embraces the whole set of rules and of human relationships involved with the publication process.

Such a definition may not help initially to grasp the process of publishing in all of its complexity - starting

---

* Förderprogramm des BMBF; Bonn, 28.10.2004; http://www.bmbf.de/foerderungen/3179.php
** http://www.physnet.de/PhysNet/
*** http://www.mathnet.de/
**** American-scientist-open-access-forum@listserver.sigmaxi.org
with the creation of a manuscript, and continuing though it’s editing, release, publication, distribution, preparation, indexing, third-party processing and possibly some guarantee of long term preservation. But this definition is helpful when the object under examination is changing the process itself. It should be understood that changes in culture result in changes in conduct as well as changes in the operating rules. Scholarly publishing is without doubt the main way in which scholarly results are put forward. Whether in monographs, textbooks, journal articles or conference proceedings, scholarly publishing on paper has a centuries-long tradition. While this article does not intend to present a history of publishing, the year 1452 certainly has an important position because of Gutenberg’s invention of printing.

Since this time the academic community has worked methodically with the publishing world to establish basic rules for those involved as well as rules for how forthcoming works are handled. Academic authors regard publishing - especially in a peer-reviewed journal - as a decisive goal. Publications are the products of scholars’ research and represent their value-added contribution to society*. Publication encourages academic debates and should contribute to knowledge. Academic publication also means peer approval. Authors aim at being read by as many people as possible and, ideally, at having their opinions quoted. To reach this aim the authors accept rules that often come at a price. In many cases authors have to adhere strictly to the specifications of the publishing houses. For example, they may have to use a specific word processor, a given font size and a particular line spacing, as well as a strictly defined citation system and detailed instructions for tables and figures. Of course other groups are also involved in this publication process and are subject to their own sets of rules. These include the reviewers who ensure the quality of a publication’s contents, the in-house editors, the publishers, the librarians and last but not least the reader. All these jobs have changed over the years, such as author specifications that optimize internal publisher workflows, publishing house specialization to respond to market conditions, or changes in the indexing of works in libraries. A sophisticated if often-regrettable ranking system for citation frequency has been developed that universities use to evaluate candidates for professorships.

The different groups of people have influenced the publication process and inevitably are influenced by it. They made rules for the varied actions in this process and – depending on circumstances – improved them. Scarcely anyone thinks about the fact that DIN A4 paper is the established size for publication or at least is used for drafts. For quotations it’s reasonable that we look for a page number and that we number tables sequentially within a manuscript. The ISBN is used for marketing and for distribution, and is involved in an ongoing debate concerning the best procedures for cataloguing of scholarly works in libraries. The number of rules is impressive because they form a nested structure where the effects of any possible change are far from clear. Also aggravating are local and linguistic anomalies that came about over the centuries. Of course new rules were established to overcome these barriers and to enable interlibrary loan between cities or even countries. This established “a totality of learned rules of conduct and of adopted attitudes, value systems and knowledge (…) that is shared and bequeathed by the members of a group.”

WHAT IS ELECTRONIC PUBLISHING?

As I already mentioned, this article is limited to scholarly electronic publishing in order to exclude the many other types of web publishing. Statistics on the number of stored theses and dissertations in the German National Library* offers a good example for the growth of electronic publications. In 1998 we had 100 publications. This grew to about 25,000 in 2004, and the real number is probably higher because this statistic includes only the papers which were send to the National Library electronically. Often university libraries have both the paper-based and electronic version of a thesis or dissertation, so in most cases the number of available electronic version is higher than the statistic shows.

Endres and Fellner** write: “Ein digitales Dokument ist eine in sich abgeschlossene Informationseinheit, deren Inhalt digital codiert und auf einem elektronischen Datenträger gespeichert ist, sodass er mittels eines Rechners benutzt werden kann.” (A digital document is a closed unit of information, whose content encoded in digital form on electronic media for use by a computer.)

If the publication of a digital document is electronic publishing, then the following example describes electronic publishing’s variety and inadequacy.

** Endres /Fellner; dpunkt.verlag 2000, Heidelberg, Seite 15.

---

* http://deposit.ddb.de/netzpub/statistik/

---
If you look back twenty years, you find yourself in the time of mostly stand alone personal computers. The PC was used mainly as typewriter with a motherboard, a 20 to 100 Megabyte hard disc, and one or two 8 or 5.25 inch floppy drives. Today are a lot of questions about the usability and availability of any document such as a thesis that was created on such a system. They start with the question of finding a floppy drive and go on to the readability of the disc and the hope that the file is accessible without a loss of information. If the answers are all positive, the next problem to be faced is finding Wordstar, which was the most common text processing system from these years and had a proprietary file format. Closely connected to the word-processor is the eight-bit the operating system. These superficial problems show that the fundamental issues for electronic publishing should be safe storage and long-term accessibility.

In recent years there have been lots of articles about electronic publishing that describe its advantages and expectations*, **. Roosendahl, Geurts and van der Vet*** write: "...dass IT dem Autor die Möglichkeit bietet, die Sichtbarkeit seiner Arbeit zu steigern, oder dass sie dem Leser das digitale Durchsuchen einer Datenbank erleichtert. Die Anwendung der IT schafft die Möglichkeit, die Umlaufgeschwindigkeit der Information zu verkürzen." (…that IT gives authors the ability to increase the visibility of their works or makes it easier for readers to do a database search. The use of IT shortens the information cycle.) The publishing process is clearly a circle, started by the author and continued by the publishing house, the library and eventually closed by the reader, who may become the next author. These various actors have different expectations of the electronic publishing. Because today there are no standards or common rules, the expectations can be contradictory. IT offers new possibilities that are able to produce new techniques within the publication process. One of the best examples is the chance to incorporate multimedia documents within scholarly works.

Authors expect user-friendly and inexpensive text processors. As a rule they also expect systems that are easily able to exchange data with other systems and scholars. Authors want publication in the shortest possible time with widest possible dissemination. And they want to ensure the copyright, the authenticity and integrity of their work.

The key technological player in the publication cycle is the publishing company, which has an interest in optimizing the process via electronic publishing. Its aim is to reduce both publication times and publication costs. Modern text processors give publishers an opportunity (which they use) to require authors to produce their scholarly works in a more finished format. Publishers transfer significant amounts of work of this sort to the authors in order to reduce the costs and to increase the profit.

Of course the scholarly libraries also confront this new situation. The number of documents that exist only in digital formats will increase. Thus libraries have an additional task. They have to handle both traditional processes and the new forms of scholarly output. In contrast to the authors and to the publishers, they are in an awkward position, because they have acquired extra work without much chance to influence the process. Every publisher gives its authors a set of guidelines that details the kind of electronic document they want to get. Society as a whole instructs libraries to collect these electronic documents, but, in contrast to the publishers, they have no chance to formulate requirements. As service institutions, libraries have to develop new technological processes in order to keep their collections under control, especially for digital materials, for which only few rules and procedures exist.

In order to close the circle let us look to the reader, in whom we can recognize additional expectations for electronic publishing. The first is to get works quickly from all over the world, at any location in the world, and in the qualities similar to the past. They want to use the new potential of electronic retrieval in order to find more interesting information and not only to find it in their local library. And of course they want materials of the same quality as when libraries ensured their authenticity and integrity.

If we review the expectations of the various players in the publication process, we should recognize that everyone involved should change their behaviors and that we need a lot of new rules and procedures. The next chapter will describe the changing process and some of the rules and approaches for electronic publishing process.
The new culture of electronic publishing

THE CULTURE OF ELECTRONIC PUBLISHING OR THE CHANGES OF THE PUBLICATION PROCESS

Expectations for electronic publishing are often very complex. If adequate modern information technology methods can be introduced and the change-process itself succeeds, many advantages will result. But changing the attitudes of the people involved is also necessary. As is true today, the authors, publishers, libraries and users must either accept the changes or (preferably) make them deliberately. This is a process that should be developed. If one equates the commencement of electronic publishing with the availability of World Wide Web, it concerns an extremely short historical period of 12 - 14 years. The shortage of rules, standards and established behaviors is understandable. Traditional publishing can rely on centuries of tradition. Authors naturally deliver their manuscripts in the form that publishers require in order to guarantee that further work can be done with as little effort as possible. Publishers should pay attention to their experience with the quality and durability of the paper used in the printing process. Libraries are also involved in an intense discussion about the rules for cataloguing scholarly works. These are examples for the present culture of publishing.

When comparing electronic and traditional publishing, the questions are: who and what has to change, and how? A decade is definitely not much of a developmental period, so the following comments in some sense can only represent a trend. Certainly details will change in the following years. Remember that the present article mainly deals with text-oriented publishing. Multimedia - which contemporary methods also can handle decently for scholarly purposes - is still in a fledgling stage and needs more research to establish standards and procedures, as well as attitudes towards them.

Authors should be aware that their attitudes can support the benefits of electronic publishing in significant ways, particularly by using it to meet their own aims. As with conventional publishing, authors should query the institution publishing their works digitally to establish guidelines for the manuscripts. That could be a publisher or a publication service offered by the author’s institution. This topic will not be discussed further here. It is of particular importance that, if possible, authors use a file format suitable for long-term archiving. Authors who use proprietary formats, such as Wordstar in the earlier example, face the likelihood that their scholarly works will be unreadable in 20 or fewer years. The “Electronic Publishing team” of Deutsche Initiative für Netzwerkinformation has made specific recommendations in two reports.*,**

In addition to avoiding proprietary file formats authors should make sure that their works provide ways in which the contents can be retrieved effectively. In traditional publishing libraries provide access to scholarly works, such as journal articles and monographs, via card catalogues or today via the OPAC (Online Public Access Catalogue). Cataloguing literature always was and still is a slow process because this information first has to be verified. This information lets researchers gain access to the papers they want and then to the full text. With modern search engines this process certainly becomes easier.

However, even at this point a researcher only reaches the full text and not necessarily the required location within the text. Search results tend to be too inexact and too broad to evaluate in detail. Significantly improved hit rates could be achieved if the texts were available in structured form. Then search results would improve because the indexing could be automated.

What would that mean?

In most cases journal authors first develop their ideas in conjunction with the outline of its structure. Next they fill in their points systematically, perhaps reorganizing one or another passage during the final compilation. Authors think in terms of structures that are organized in chapters, sections, subsection etc. People can understand this structure if they see it on paper or on screen. But search engines read it without recognizing the organizational elements purely as part of the “so called unstructured full text”. Automated analysis could be an option if authors created a structure while writing that machines could recognize. Marking the structure of the article with modern word processors is not particularly difficult. These systems offer a way to emphasize a chapter heading that is structurally characteristic, and likewise they can distinguish a subchapter from the main heading. An introduction can also be marked to differentiate it from the legend of an illustration, pictures or tables. In the same way citations and references can be distinguished from the body of the text. These structural labels can be described in the form of a Document Type Definition (DTD) and can be reproduced in markup languages such as Standardized General Markup Language (SGML) or Extensible Markup

Language (XML). In this way they become recognizable elements for modern search tools.

Theoretically every text could have its own unique structure, but that is uncommon. Journal articles have similar structures, just as do dissertations within fields (see Diss DTD* on EDOC server of Humboldt-Universität zu Berlin). The amount of repetition in the structure of scholarly texts certainly varies. It should be clear that the more the structural elements are searchable, the better the hit rate and the higher the quality of the information retrieval.

As with traditional publishing the author should come to terms with the publisher, whether it is a commercial firm or an academic publishing service. DTDs, style sheets (for Microsoft Word) or templates (for Open Office) exist that could be used as models. These examples show how authors can influence the electronic publishing by changing their habits when writing scholarly works.

The intense debate about the publisher’s role should not be discussed at this point because this article is focused on changes in conduct at the institutions that have responsibility for scholarly publication. In this context “responsibility” means:

– publication and distribution,
– ensuring authenticity,
– ensuring integrity,
– guaranteeing long term preservation.

The present publishing culture has fixed rules for these points, based on a thousand years of paper-handling experience.

There are distribution channels, marketing strategies, specialist bookshops and of course research libraries. Nothing comparable exists for electronic media. An Internet publication can be made available promptly because it requires little effort and it almost automatically includes worldwide distribution, which is one of the main advantages of electronic publishing.

But no established solutions for the problems of authenticity, integrity and preservation exist as yet. As was said earlier, this is not surprising because of electronic publishing’s short history. It is important, however, that people are aware of these problems and are working to solve these problems and to change behaviors. This seems obvious but is less common in reality.

Everyone who puts something on the Internet should be aware that, unless there are explicit restrictions, it involves worldwide distribution and worldwide access.

Unfortunately, there are many servers in academic institutions, some research teams and some scholars who seem not to have this impression. The culture of electronic publishing expects that those who operate publication servers are aware of the worldwide service they offer as part of the framework of scholarly publications. This includes responsibility for the quality of services. This does not the quality of the contents but the quality and stability of publication. For example, the server operator is responsible for ensuring that every scholarly publication that has been found once on the server can always be found again at the same place, unchanged, and will genuinely be the stated author’s work.

The previously mentioned DINI certificate* includes a set of expectations for publication services that make a distinction between minimum standards and recommendations designed to prepare for future developments. A server’s policies should be expected to have basic rules for running its current services. Both the policymaker and the reader should be able to glean how long the organization intends to take responsibility for the content, what mechanisms it has to protect the server against the alteration or loss of data, and what measures are taken to ensure the integrity of individual documents. An example of a minimum requirement for such a publication service could be the provision of persistent identifiers for each document. In other words each document on the server should receive a unique address that ensures that the document will always be reachable at that address.

In traditional publishing, the publication and distribution phases are followed by advertising, cataloging, and storage. These roles are not fundamentally different in electronic publishing, but the boundaries between the parties’ tasks are not so clearly delineated. Thus research libraries can offer publication services as an additional feature of (for example) a university, and many publishers that think that they can take on the cataloging and verification of their scholarly works, tasks that in some sense previously belonged to libraries.

In this article there is no space to take part in this discussion. At first glance it appears unrealistic to assume that a publisher would take responsibility for its whole

The new culture of electronic publishing

stock over so long period that its readers’ interest fell off – and thus also its sales and its profits. This discussion becomes sharper when one considers the hot debate over Open Access to scholarly knowledge.

For the purposes of this article it is not especially important to know who administers the various tasks in electronic publishing. It is more important that they are consciously accepted and constructed. It is more important that the new culture of electronic publishing deals with issues like standardized metadata, data exchange, the integration of electronically accessible materials in the OPAC, securing long term access, and guarantees for the authenticity and integrity of the documents.

The readers are the ones who most quickly accepted these changes in the nature of publication and have changed their behaviors. This is attested to by the fact that searches are eight times greater for electronically accessible materials than for materials exclusively in paper*. Such figures must be used cautiously and they very widely depending on the subject. But there is a clear tendency for readers to use these new opportunities and consequently they induce changes in the publication process. Naturally readers must also grapple with existing limitations. Their influence on document providers and thus their effect on the whole organizational process is a key factor. They came into being as part of the current discussion over unhindered access to information and are its strongest advocates.

CONCLUSION AND PROSPECTS

The implementation of modern methods of information and communications technology has led to a great change in scholarly communication. Electronic publishing as a form of scholarly communication stands at the beginning of its development. It is worthwhile to show the process of change and also the apparent advantages of this new technology. It is also important not just to create an electronic reproduction of the existing publication process but to change the whole process by taking advantage of the benefits of electronic communication, and thus separate it from somewhat outmoded practices. This article gives in some sense only a snapshot in time. The process requires the development of a new culture of electronic publishing that addresses specific practices and minimal standards.

To simplify this discussion, only essentially text-oriented scholarly works were addressed. When one considers contemporary developments, it becomes clear the scholarly communication cannot stop at this level, but must tackle multimedia as part of its future work. Demands for changes in practice for everyone involved in the publication process and for their conscious participation in the creation, cataloging, and preservation of these new media will certainly grow stronger, because only that way can the body of scholarly thought be preserved.

REFERENCES

5. Disponível em: <http://www.mathnet.de/>
6. American-scientist-open-access-forum@listserv.siam.org
17. See footnote 12.