"Electronic Theses and Dissertations and the Necessity of a New Culture of Electronic Publishing"

ETD's at Humboldt-University at Berlin

Peter Schirmbacher

Humboldt-University at Berlin, Computer- and Mediaservice

Schirmbacher@cms.hu-berlin.de

Humboldt-University at Berlin,

Computer- and Mediaservice, Unter den Linden 6,

10099 Berlin

www.hu-berlin.de/cms/schirmbacher/

Keywords: ETD, electronic publishing, new culture

Abstract

The creating, using and supply of scholarly documents in a digital way by the internet is a new, additional and improved kind of scholarly communication in comparison to traditional paper-based procedure of publication. In this sense you can speak of a "culture of electronic publishing". That means that all participants in this process should work together and change their attitudes of minds. That starts with the authors, who should structure their texts in order to support the retrieval and also the long term archiving. We need a special collaboration and responsibility between university libraries and computing centres.

The electronic publishing of theses and dissertations is a very good example for this changing situation. In many countries worldwide we can notice the alteration. In Germany where we have some special projects in this field too. The German initiative for networked information (DINI) tried to summarize the different experiences and worked out a special paper called "Recommendations of electronic publishing in universities".

In my presentation I will speak about the present situation in Germany, report on how to build up a document server, a policy for a university digital library and some special points of maintenance of integrity and authenticity.

Electronic Theses and Dissertations in Germany

In Germany there is a duty to publish every dissertation and in most universities it is allowed to realize this duty also in an electronic version. University libraries have the duty to collect the dissertations from their own students and the German National Library is obliged by law to collect all paper based German dissertations. I believe that most of you have already heart from German project "Dissertation Online" which was finished two years before. Beside the technical results of this project the German National Library decided together with the people who are involved in this and similar German projects to built up a German office for coordination and support for electronic theses and dissertations. This activity got an additional grant by the German Research foundation (DFG). This office, located in Frankfurt/Main at German National library, has an so called advisory board, which consists of member from libraries, computing centres, scientific communities, the German learned societies, publishing companies and the DFG. It is the aim of this board, I am as a representative of computing centres also a member of this board to advice the office, to give recommendations to universities, to discuss problems of present developments and to initiate new projects.

Year of publication	Number of electronic dissertations
1995	18
1996	31
1997	134
1998	477
1999	1.414
2000	3.235
2001	4.483
2002	4.712
2003 (up to 01.04.)	381
Sum	10.173

Number of Electronic Dissertations a year in Germany based on Web-Statistic of German National Library

Scientific field	Percentage of electronic dissertations (in 2001)
All	16%
Humanities	up to 10%
Natural sciences - physics - mathematics - computer science	about 25% 37% 33% 23%
Agriculture	21%
Medicine	13%

Percentage of electronic dissertations depended on the scientific field

The two tables give a statistical overview of electronic dissertations in Germany. You can see that the number is growing rapidly. Counted are only such dissertations, which the German National Library got electronically. A student in Germany can publish in most universities electronically and additional he or she gives four paper based copies of the dissertation to the library. The library can decided if it send a paper based or electronic copy to the National Library. So the real number of electronic dissertations should be a little bit higher.

But more important, I believe, then the absolute number of electronic dissertations is the question how many percentage of the students publish their dissertation electronically. In average at present 16 percent of the students publish electronically. In my university that are about 25 %. Of course it is an average and it depended on the branch of scholarly discipline.

So in humanities we have an average of electronic dissertations from all dissertations from 10 % and 25 % in natural sciences. Inside the natural sciences about 38 % of physicists and 32 % of mathematicians publish electronically and only 17 % of the student for computer sciences. In agriculture it is 13 % an in medicine it is also 13 %.

After eight years that are impressive results but at the other side they show that a lot of things are to do.

What is a culture of scholarly publishing

If I write about the present situation so I mentioned that we are at the very beginning in the field of electronic publishing. Especially we are trying to find out what is new and what have to change in comparison with the traditional way. What is that the traditional way or what is the culture of publishing.

In a lot of years the scholarly community developed some rules and standards, some technologies, some be-

haviour and patterns. All these things together are for me the culture of publishing.

We have a long tradition in scholarly publishing. It started with the invention of paper 2200 years before. The human used the papyrus roll and Gutenberg produced the bible in a new printing technology 1452. So we have several hundred years' experiences organizing public and scientific libraries. We have experiences in acquisition, exploitation, cataloguing, long term archiving and so on. I believe every body knows it. These established rules and procedures were developed in a process about a lot of years. They changed a lot of time and there are in some cases not consistent rules all over the world.

As a part of the culture of publishing we have an established convention of publishing. The authors are interested in publishing. A referee SYSTEM exists. Publishing houses have a lot of experiences in publishing, dissemination and marketing. Libraries have the duty to collect and of course the scholarly community is interested in to read the latest findings. Some other examples for the culture of publishing are the citation SYSTEMs, rules of scholarly communication, protection of intellectual properties and so on.

It is not the place to describe the culture of electronic publishing in all details. In my presentation during the conference I will give an overview.

Changes in publishing process

So what is changing or what are at first the expectations of electronic publishing?

The expectations are not consistent. A lot of groups or different players are involved in the publishing process and all of them have some other interests.

Let me start with the authors:

From my point of view the authors are interested in:
- Modern text processors available for different platforms and free of charge or very cheep;
- It should be a prompt publishing and a wide dissemination;
- They attach importance to protection of intellectual properties, authenticity an integrity;

Also the users have some special demands to electronic publishing:
- Because they know that is possible they want the free access to information resources worldwide;
- They expect new possibilities for retrieval;
- And on the other side they are a little bit worried about the authenticity of the document which they find for instance in the web.

The most problems for the libraries are that a lot of things are changing. There are procedures, rules and standards but in most cases electronic publishing does not fit in it. One of the biggest problems should be the

long term archiving of documents. How can they fulfil their collecting standing order? What can they really guarantee the users?

Relative new players in this process are the computing centres. In the past they were not involved. Nowadays they are necessary to realize the computing network, especially the appropriate bandwidth and so on. But they are also involved building up the so called digital library, because they have experiences in computer based archiving with tape robots or other equipment.

Of course also the publisher or publishing companies are looking to this new development. They have to find their new or old position, how you like it. What are the tasks of publisher in this process of electronic publishing? Where are the chances to produce surplus value?

These different expectations should show you the difficulties for new rules and standards. Sometimes the interests of the different groups are antagonistic.

I want try to explain which things are changing in the time from traditional publishing to electronic publishing. I will do it with the help of the steps of publishing how I see their.

In the process of creation an electronic document I see at first the interests or demands of the authors. Partly I already mentioned it.

They expect
- Modern text processors
- Supporting tools for multimedia applications
- Guarantee of integrity and authenticity
- Short period up to publishing
- Worldwide availability
- Some time they are thinking about long term archiving but in most cases this is not their problem

At the other side there are some requirements to the authors, which they should take in consideration. In most cases these are points which are a part of a really new culture of publishing. That means that also the authors have some responsibilities for the publishing process. Only if he or she is willing to respect some facts the advantages of electronic publishing can be realized.

Requirements:
- In order to realize an long term archiving an authors should not use proprietary file formats.
- They should use standard file formats or common rules which make it possible to convert files to standard file formats like XML or SGML.
- The authors should structure their documents in order to support a qualified structured retrieval.
- That are only examples also like a citation SYSTEM and so on, important is that the authors undertake responsibility for the publishing procedure.

Especially in the field of availability and publishing or presentation you can see the different level, consider the pros and cons between traditional publishing and electronic publishing.

The traditional way is described by
- Stabile technology,

- Long publishing cycle,
- Quite small availability
- An established referee SYSTEM for journal and some other publications
- Immutable paper based documents and so on.

In different to that the process of electronic publishing is described at present by:
- No established technology,
- Short publishing cycle,
- A comparable referee SYSTEM is to built up in most cases,
- New procedure are necessary in order to guarantee authenticity and integrity like time stamps and digital signatures

The exploitation is a typical example or result of the years of discussions about rules and standards between librarians. In order to describe electronic documents a new discussion started few years before and I do not know when they will really be finished.

The community tries to find new ways:
- They developed the Dublin core metadata set and try to integrate this in the online public access catalogs (OPAC)
- They developed rules for the authors in order to collect metadata and discussed if this really the right way to success.
- Some institution produced Templates in order to structure documents
  (Document type definitions)
- They discuss standard file formats for different media types
- They organized the Open Archives Initiative in order to realize the technical possibilities to access
- They established the Budapest Open Access Initiative in order to make it possible for everybody to access

All these are examples for the present time which is characterized by discussions, attempts, descriptions, different projects. We can say in the middle of development a new technology which will be characterized by some rules, standards and organizational guidelines.

In the field of archiving we have to register the most disadvantages of electronic publishing over the traditional way. The traditional way is clear with really very long experiences. It is determined by paper based formats and such a paper based document is immutable.

Inside the electronic publishing technology we find a lot of unsolved problems. It begins with the storage media. There are no experiences about a long term. What is the right media, tape or CD-Rom or some other. But if this problem is solvable by periodical conversions.

An additional problem is the file format. We recommend to SGML or XML. But there are not enough tools to support the authors and the service staff in libraries and computing centres to convert to XML from the different file formats like Microsoft Word, WordPerfect, Star office or LateX. We are on the way to find some

new technologies to guarantee authenticity and integrity but also here we are at the very beginning.

The only point in the field of archiving which the traditional publishing procedure could have some disadvantages are cost for rooms and boards for the literature. But nobody can say today how much are the cost for periodical converting and so on.

If you look to retrieval there you can find another situation. Successful retrieval is one of the main objectives in publishing. Electronic publishing gives a chance of higher quality and quantity of retrieval.

Quality that means
- the Chance of better hits
- to find it easier what you are looking for
- And also the integration of multimedia documents

Quantity of retrieval that means
- Basically worldwide access to literature. I know that this is only an aim and not the reality.
- Process of search is quicker.
- And the availability is extremely broader

All of us know that these are possibilities which are only could realize if we come together to new rules, standards and technologies. That is that what I mean, when I speak from a "new culture of electronic publishing".

Berlin, 23.05.2003