

## Doctoral Dissertations in the Internet: A Good Start in Finland

by ESKO HÄKLI & HENNA LEHTINEN

### THE MENTAL CHANGE FROM PRINTED TO ELECTRONIC PUBLICATIONS

At the present stage of the IT-applications the change from print publishing into electronic publishing of doctoral dissertations is not only a question of technology but it is also, and possibly even to a greater extent, a question of thinking. Although there is a large selection of technologies and publication formats available, none of them seem to be mature enough to be a self-evident choice. But obviously a greater obstacle can be found in the minds of people. In many cases legal regulations also have to be revised. Academic traditions still prefer properly printed dissertations and the attitudes of the academic community are not necessarily favourable towards electronic publishing in general. Young post-graduate students are obviously not able to appreciate a virtual dissertation. They seem to need a real tangible thing, which they also can show to their parents! And finally, the whole organization and financing of publishing may need to be revised, an exercise which is far more difficult than only starting to use new technology.

All these reasons which exist mainly - or only? - in the minds of people have made the transition from paper into electronic format rather slow also in Finland. Since autumn 1999, when this text for the first time was drafted, a remarkable change has, however, taken place. A number of universities have made the decision to start publishing all their doctoral dissertations in electronic form. Of course, even all other universities are in the process of developing and organizing their publishing activities in order to make their scientific output available in the Internet.

This step towards network publishing is only natural in Finland where the use of electronic means is widespread and where the academic datatransmission network is among to the most advanced in the whole world. At the same time the academic community has been given access to an unforeseen large amount of scientific publications in digital format, which in a practical way has shown

the great advantages of online access to the world's scholarly output. Within the framework of the National Electronic Library Programme, launched in 1998, about 3,000 journals published by well-known scientific publishers are available at the Finnish universities, free on the desk of every individual scholar. A growing number of university libraries are cancelling subscriptions to the printed versions of these same journals, which is accelerating the transition from print to electronic media. If renowned journals can be used in the Internet, why not publish doctoral dissertations in the same way?

And finally, economical reasons are speaking for the change. Due to the present policy of the government, the number of doctoral dissertations has grown substantially. More than one thousand dissertations are being published annually and the number will obviously still grow because according to the plans of the Ministry of Education for the period of 2001-2003 the amount of doctoral examinations should still increase with one third. The share of the biggest university, the University of Helsinki, is about one third of the total amount. Ten years ago about 400 - 500 dissertations were published annually. Therefore, it is simply no longer possible to finance the publication of all dissertations as properly printed books. And there is no longer any need for it, because the role of the dissertations has also changed.

#### TOWARDS ELECTRONIC PUBLISHING

Publishing of dissertations in electronic form has already had a good start but it can by no means be said to be a well-established pattern of publishing yet. So far, electronic publishing is in most cases a parallel way because the main part of dissertations is also being published as printed books. According to the Finnish regulations a dissertation can also be a combination of the author's earlier printed articles together with an introductory text summarizing the scientific results. Dissertations of this type are rather numerous in medicine and natural sciences where journal articles are the main form of scientific publishing. It will hardly be possible to publish dissertations of this type as separate and coherent electronic publications.

In Finland doctoral dissertations have to be defended publicly. To be able to do so dissertations are traditionally printed. The majority of the dissertations are published in the publication series of the universities or learned societies but in some disciplines trade publishers also have a role to play. In most cases the author receives a modest financial support from his university but basically he is himself responsible for the publication and its financing. Universities as organizations can promote the publication of dissertations but

no mechanism exists which automatically would make the university to act as the publisher. The only obvious link between the university and the dissertation is that the dissertation is compulsory and that the doctoral candidate has to hand over to the university a certain number of his dissertation, about 100 - 300 copies depending upon the decision of the individual universities. Publications are needed in the approval process, a certain number of them is made available to the academic community at large and they are also sent to other libraries within the framework of the book exchange.

In 1995 the Ministry of Education set up a Working Group to prepare recommendations for publishing dissertations in electronic form. The Working Group concluded its recommendations as follows:

- dissertations which are being defended for the doctoral degree should not be printed; depending upon the case and the discipline dissertations can afterwards be printed as monographs, or the main results can be published as journal articles;
- each university can decide how many copies are needed for the process of approval;
- libraries should abandon the use of dissertations in exchange of literature because the character of the dissertations will change;
- dissertations have to be archived either as paper based originals, on microfilm or in electronic form; additional copies may be produced e.g. by using the print-on-demand-procedure;
- university libraries are responsible for cataloguing of the doctoral theses in their own catalogues as well as in international bibliographic databases for dissertations.

The recommendations were received with mixed feelings. The academic community was far from being enthusiastic about the new ideas and some libraries were afraid of losing an important source of acquisitions if the international exchange would have been cut down. A number of years have elapsed before the first steps were taken.

Today, several universities are either publishing dissertations in electronic form or have made the decision to do so. The introduction of electronic publishing of dissertations has changed the role of the universities insofar that universities have to assume the responsibility for organizing the business of electronic publishing and they also have to offer the service to all their graduate students who no longer have the possibility to choose between

publishers and printers. This is a major organizational change which has caused quite a lot discussion and planning.

#### THE PUBLISHING STARTS: SOME PRACTICAL EXAMPLES

The *University of Oulu* started testing electronic publishing of dissertations in 1997. Initially, the project was a part of the larger project ELEKTRA funded by the Ministry of Education and co-ordinated by Helsinki University Library. ELEKTRA is a collaborative project aiming at creating a mechanism for publishing copyrighted texts in the Internet and organizing a controlled network access to them. One of the main goals of the project is to create a solution to the management of authors' rights. Oulu University Library was responsible for producing dissertations in a test environment.

From the beginning of 1999, 80 per cent of dissertations printed in the series *Acta Universitas Ouluensis* have been published simultaneously in print and electronic format. The authors of the remaining 20 per cent have not given their permission to publish their dissertations in the Internet. Electronic versions are accessible free of charge. Since the vast majority of the university's dissertations, during 1999 totalling 122 titles, are published in the Acta series, the coverage of the dissertation database is quite extensive. The number of dissertations in the database, at present about 100, is increasing rapidly. HTML and PDF formats are used for distribution, and SGML format for long-term archiving. In the future XML format will possibly be used for distribution.

In 1999 the University of Oulu set up a Working Group to prepare a plan for the University's electronic publishing activities. The recommendations of the Working Group have been adopted and the University is proceeding accordingly. From the beginning of 2000 the Acta-series will be published parallelly in a printed and an electronic version. The Library is responsible for cataloguing, distribution and archiving of the electronic publications. URN is used as identifier and the metadata is produced in the Dublic Core format. The publications are accessible via the Library's homepage in such a way that they can be indexed by the various search engines.

Simultaneously with the University of Oulu a number of dissertations of the *University of Helsinki* were published within the same ELEKTRA project. This particular database contains at present 27 dissertations. The original titles were dissertations which the Finnish Historical Society was publishing as

printed books but which now were made available simultaneously in electronic form by Helsinki University Library.

Because the University of Helsinki does not have any central organization for publishing and because many of the dissertations are published by learned societies or commercial publishers, the decision making, contracts and copyright issues are more complicated than at other universities. Since the beginning of 1998 Helsinki University Library together with the University's computing services, faculties and the University Printing House has developed a technical and organizational model for publishing and distributing the University's publications in the Internet, not only doctoral dissertations but also masters' thesis and publication series of individual departments and institutions. The service was opened in autumn 1999 and received the name *E-thesis*. About two hundred publications have already been published, more than one hundred doctoral dissertations among them. The responsibility for running the service will now go over to the University's Printing House whereas the University Library will take care of the distribution and archiving.

The E-thesis programme is not aiming at publishing printed versions simultaneously. According to the basic concept the electronic version will be the primary and only version which the E-thesis service is responsible for. With the help of the print-on-demand-technology 30 - 50 copies which are needed in the process of approval are being produced, but it is on the discretion of the author to publish the dissertation afterwards either as a printed book or in other channels in the Internet. The E-thesis service is only offering an answer to the question of how to publish dissertations in the Internet. Thus, compared to the decisions elsewhere to introduce parallel publishing, the concept of the E-thesis has already gone one step further.

As a part of the planning process the University of Helsinki revised its internal regulations. In stead of 160 printed copies a dissertation can be submitted in 30 copies if the text is delivered in electronic format. The author signs an agreement on publication assuring that the contents of the electronic version is identical to that of the print version if the printed version is not produced by using the print-on-demand-procedure. In addition, the author gives the permission to store the dissertation in the Internet and to make it available freely without any requirements for remuneration. The network editors provide the dissertation with a standard number such as URN and metadata (Dublin Core) if the author himself has not taken care of it. The University Library is responsible for the cataloguing as well as the distribution and long-term archiving. For that purpose it has the right to convert the publications into a new format if that proves necessary.

Due to the lack of a central publications office and the size of the University the network editors of the new E-thesis service have to be flexible in accepting texts for publications. A number of word processing programmes are accepted as well as RTF and PostScript formats and templates are offered for several word processing programmes as well as a style sheet. Dissertations are distributed in both PDF and HTML formats the latter being so far also the archival format. It is anticipated that XML may replace HTML.

The *University of Kuopio* has recently made the decision to publish all the University's dissertations also in the electronic form provided that the authors give their permission. The texts are given to University's printing office either on disk or transmitted as files online. The printing office converts them into PDF format and stores them on the University's server. The Publication Committee of the University is responsible for electronic publishing also but the library has a visible role in the process. It maintains the Internet page of the dissertations and gives advice on publishing. The secretary of the University's Publication Committee also comes from the University Library. Decisions on long-term preservation have not yet been made.

In Kuopio the University and the local Polytechnic are launching close co-operation by using the services provided by the University. In addition to the doctoral dissertations also other types of thesis will be published in the Internet and made available free of charge.

*Tampere University Library* launched its database of electronic dissertations in October 1999 and 13 full-text dissertations have been added to it during the first three months. The total number of dissertations at the University during 1999 was 104. Until now all dissertations have also been published as printed books, mainly in the University's *Acta* -series. The Working Group who was planning the change proposed that from August 1, 2000 the switch from print to electronic should be completed, and that the publishing of printed dissertations should no longer be supported. A part of the financial subsidies which were granted for printing are now on the University's decision used to finance the production of the electronic versions of the dissertations. The author assigns a copyright agreement with the library giving it the right to copy and distribute the dissertation in the Internet.

The dissertations database contains links to the University's virtual bookstore Granum which is used to distribute the dissertations. The distribution format is PDF, and the long-time archiving format RTF. For preservation purposes the dissertations are stored on CD or DVD discs as RTF files, a procedure which is quite unique. The abstracts of the dissertations will be available in HTML format to allow searching. This applies also to dissertations not

published in the University's Acta-series as well as in cases where articles constituting a combined dissertation have been previously published.

*Jyväskylä University Library* has, so far, chosen a different approach which strictly speaking is something else than electronic publishing although it is resulting in a database of full-text publications. The Library is scanning printed publications and storing them on the Library's server. The Library has a database containing both master's and doctoral theses. The University is using Xerox hardware and software. The paper version is scanned and the electronic dissertation is distributed in low-resolution image format (JPEG). The advantages of this procedure are e.g. a high speed. The database contains more than one thousand titles in full-text. The Library has emphasized copyright issues in connection with storing and distributing publications through Internet. By using an image format the Library originally wanted to avoid plagiarism. Another issue is the hardware. At the moment there are only three dissertations that have not been published as a printed book. They have been scanned from the manuscript and only a limited number of copies necessary for the public defence have been printed.

The present solution chosen by *Jyväskylä University Library* does not represent the mainstream of electronic publishing in Finland. It is too heavily tied to one proprietary software and hardware. The JPEG image files can not be indexed by the search engines, nor can they be integrated into a larger national context. The low resolution of the images makes the text difficult to read on the display. Therefore, the Library has prepared a new plan aiming at electronic publishing in a large scale using the same standards as used by the other libraries. No decisions have yet been made because the University is still considering how to arrange its publishing activities.

Some of the other universities are making preparations to start publication of electronic dissertations. The publication of electronic dissertations have been investigated in *Lappeenranta University of Technology Library* and *Helsinki University of Technology Library*, but no electronic dissertations have yet been published. At Lappeenranta the Library and the University's Computing Centre started in September 1998 a collaborative project LUTPub (Lappeenranta University of Technology Publications) aiming at publishing all the abstracts of the theses. The follow-up project LUTPubII aims to publish full-text theses and dissertations in electronic format. Helsinki University of Technology has investigated the publication of structured documents in its project called HUTpubl (Helsinki University of Technology electronic publishing). The project aims to establish a set of technical procedures for electronic publishing, archiving and distribution of HUT scientific publication series. The documents are distributed over network using HTML format, and

SGML is used for long-term preservation. The prototype is planned to be used for dissertations as well. The production of electronic dissertation will start in september 2000 when two departments will start a pilot. Printed dissertations will still be published but only in a reduced number of copies.

The Swedish-speaking university *Åbo Akademi* in Turku runs a publishing house of its own. The publishing house has recently announced its plans to start publishing all the University's dissertations in electronic form. For the approval process only a limited number of copies will be printed on paper. The publishing house may publish attractive dissertations also as printed books but the decision will be made separately in each case and the texts shall be edited to better correspond to the needs of a printed book which will be on sale on the book market.

#### LEGAL AND ORGANIZATIONAL ISSUES

Requirements which the students have to fulfill in order to pass the examinations for various university degrees are regulated by the ordinances of the government. Each ordinance for a specific subject field applies to examinations in all faculties of the country in this field. For a doctoral degree a written dissertation is required. According to the oldest ordinances from the 1970's (medicin, odontology, business and commerce) the dissertation has not only to be written but it also has to be published. Nowhere has, however, been mentioned that a dissertation should be printed although that has been the understanding until now. The new ordinances only say that a dissertation has to be prepared (written) and that it has to be defended publicly.

In other words, the present government regulations do not prevent universities from accepting dissertations in electronic form. Because a dissertation has to be defended publicly, universities require that dissertations have to be available in due time before the act takes place. For that purpose the candidates have to deliver a certain number of copies of their dissertations to be distributed among the faculty members and those involved in the process. In principle the dissertation should be available to all interested citizens but not necessarily free of charge. Both the ELEKTRA and E-thesis projects have shown that the electronic dissertations can easily be put on display by using WWW pages. In the ELEKTRA project where the access is limited the electronic full-text versions were displayed ten days in advance in PDF format without a possibility of printing them. After the period of ten days an abstract remains. A great number of visits have been counted on these pages.

Universities still have their individual internal rules which may require printing of the dissertations. A special procedure also exists for granting the permission to print (imprimatur). These rules, of course, have to be revised if the practice will be changed. Here, traditions and attitudes may still play an important role.

The experience gained in the ELEKTRA and E-thesis projects has shown that the copyright issues have to be taken seriously. In the ELEKTRA project the copyright issues were solved through written contracts which are signed by the publishers of the printed versions and the authors. The Finnish Joint Copyright Organisation KOPIOSTO is handling the drafting of the contracts and procuring of copyrights. During the test period a decision was made that the protected texts could be read free of charge, but that a charge was added for printing. The principles as well as the practice is being developed further in order to find a feasible procedure which can be applied also to a great number of publications and rights holders. Experiences from the project have also been applied to the EU-projects TECUP and DIEPER. In the E-thesis project another approach was chosen. Authors agree that texts can be used freely in the Internet. We still need a longer experience to be able to see, which line will receive more support in the future.

Organizational issues have created great differences between universities. One could have imagined that small universities with a limited annual amount of dissertations would have found it easier to change from one pattern of publishing to another. Small universities usually also have centralized publication activities. The fact, however, is that the biggest universities have shown the way, in spite of all organizational obstacles. The University of Helsinki e.g. has no organized patterns for publishing, neither a common publications committee nor any common publication series. Traditionally the publication series of the learned societies have been the main vehicle for the University's publications and they can not be governed by the University. Therefore, the University had to create a new arrangement for its electronic publishing which seems to require that the University has to create a publication policy of its own.

Organizational decisions are related to many practical issues such as:

- who will be made responsible for the service?
- what shall be published in electronic form: dissertations only or also other types of publications such as masters' theses and publications of the individual departments and institutions?

- who is going to select the items to be published if they are not only dissertations?
- how will the editorial policy be formulated: will there be a real publishing house with competent editors who prepare all the texts for publication or is one inclined to try to persuade individual authors to produce ready-to-be-published texts?
- which copyright-solution will be chosen?
- how will the texts be made available, where and according to which principles? Shall the users approach every university's publications separately?
- who will take care of the long-term preservation of the publications?
- who will finance the activities? Can resources be re-allocated from supporting printed publications to meet the needs of electronic publishing?

#### TECHNICAL ISSUES: STANDARDIZATION, DISTRIBUTION, PRESERVATION

A great number of technical issues have to be solved before electronic publishing of dissertations starts. The selection of standards e.g. can have a direct impact on the resources needed to provide the service as well as on the speed of the production. In order to save efforts at individual universities a development project has been launched in Finland to collect the experiences from the most advanced universities and to prepare recommendations for the best practice. The National Electronic Library Programme is financing a project, where the university libraries of Helsinki, Jyväskylä and Oulu together are preparing a model for Internet publishing. Such a model has to concentrate on standards and other technical issues while organizational and policy issues have to be left to the individual universities.

For the time being Finnish universities are making their electronic dissertations available on their Internet pages. Usually the university library is responsible for the distribution. No common service exists, so far, but electronic dissertations are catalogued or will be catalogued in the union catalogue Linda, as well as in the local library catalogues. It is, however, obvious already now that users want to have one single service where they can find all the dissertations and other electronic publications produced at the country's universities. It does not make any difference if the use of the publications is free or fee-based. In both cases the access should be as simple

as possible. This view has received strong support from the experiences of using electronic journals online.

These questions of access have not yet been discussed. The University of Tampere has offered other universities its services. It is prepared to take care of publishing and distributing dissertations through its virtual bookstore Granum. This arrangement could very well be one channel for distributing the dissertations both free-of-charge and as a fee-based service. Another possibility for fee-based access could be the ELEKTRA service. And finally, the National Electronic Library Programme is developing a common user interface, a common portal, for all electronic resources which will be made available for Finnish libraries and their users. The same portal could very well be used as a gateway to the electronic publications of the Finnish universities. For the time being all alternatives are open and have to remain open for some time in order to allow the development the necessary freedom.

One of the biggest unsolved issues concerning electronic publications is the long-term preservation. There are no standards available securing the permanent archiving. The only way to proceed is to apply open standards which allow further development and possibly also migration to new standards. Of course, permanent archiving is depending upon several other factors as well, such as the development of software and hardware. But these problems are common to all types of electronic publications.

In the academic community the long-term preservation is closely connected with an other issue, the authentication of the publications. The authentication includes several parts starting from the unambiguous identification of the publication. It shall not be able to destroy or change the documents. In principle documents like doctoral dissertations have to be protected also against their authors who shall not have a chance to make any changes in their texts afterwards. And finally, there must exist a storage with the help of which it will be possible also in the distant future to control that publications really are what they are maintained to be. A well protected and well kept master copy must exist somewhere.

One can ask, whether every university has to set up a storage which meets all these requirements. Our answer is no. At least in Finland the Legal Deposit Legislation is being revised and the new law will cover also electronic publications in all their forms. The legal deposit collection of electronic publications will easily cater for all the needs mentioned above and free the individual universities from maintaining there own archives for long-term preservation. If trade publishers are already interested in using the services of

the National Library, why should the same service not also meet the needs of the universities?

#### ABOUT THE FUTURE

The most recent experiences suggest that the universities will quicker than anticipated go over to publishing their doctoral dissertations in the Internet. In spite of the fact that it is rather easy to mount publications on a server all decisions to start publishing scholarly and scientific texts in electronic form should be considered carefully. Technical and organizational decisions should support the best possible access to the publications both in a short and long term perspective. Decisions shall take into account the needs of the long-term preservation in particular.

The second point is that during the age of Internet libraries are becoming part of a common virtual space where libraries should be able to work together. Therefore, a greater uniformity e.g. in adoption of standards will be required in the future and every country should strive at a common policy as well as common arrangements in all important areas. Electronic publishing obviously belongs to those areas. We are happy to say that until now research libraries in Finland have been able to formulate a common policy in their important issues and it is to be hoped that this favourable development will continue.

Esko Häkli  
Director  
Helsinki University Library  
Tel: + 358 9 1912 2721  
Fax: +358 9 1912 2719  
esko.hakli@helsinki.fi

Henna Lehtinen  
Planning Officer  
Secretariat for National Planning and Co-ordination  
Helsinki University Library  
Tel: +358 9 1914 4033  
Fax: +358 9 753 9514  
henna.lehtinen@helsinki.fi