

## Marketing for participation: How can Electronic Dissertation Services win authors?

Bettina Berendt<sup>1</sup>

Elke Brenstein<sup>2</sup>

Yunfan Li<sup>1</sup>

Bert Wendland<sup>3</sup>

<sup>1</sup>Humboldt University Berlin, Institute of Information Systems

<sup>2</sup>Humboldt University Berlin, Institute for Computer Uses in Education / Institute of Work Psychology

<sup>3</sup>Humboldt University Berlin, Electronic Publishing Group, Computer and Media Services [Computing Centre]

[berendt@wiwi.hu-berlin.de](mailto:berendt@wiwi.hu-berlin.de)

[brenstein@educat.hu-berlin.de](mailto:brenstein@educat.hu-berlin.de)

[yunfan.li@student.hu-berlin.de](mailto:yunfan.li@student.hu-berlin.de)

[bwendland@cms.hu-berlin.de](mailto:bwendland@cms.hu-berlin.de)

<sup>1</sup>Spandauer Str. 1, D-10178 Berlin, Germany

**Keywords:** electronic theses and dissertations, training and author education issues, empirical studies

### Abstract

*Technical progress in electronic publishing affords increasingly sophisticated archiving and retrieval options for authors as well as readers of ETD and other university document publishing services. The "marketing" of these services has received less attention, despite its great importance for success. In particular, an ETD service must market to its potential authors - their work is what constitutes the service's content. Providing such content may require additional time and effort: The creation of an adequate electronic document involves the intellectual and technical (re-)structuring and markup of a text. What forms of information policy and author support are needed to meet these new challenges for authors? This paper describes a large-scale, questionnaire-based study of the marketing strategies of the Document and Publication Server of Humboldt University Berlin (<http://edoc.hu-berlin.de>). This service emphasises long-term archiving and structured retrieval, which requires authors to prepare their texts for conversion into the relatively complex, SGML-based format DiML (Dissertation Markup Language). In this study, the experiences, assessments, plans, and wishes of two central target groups of authors were investigated: people working on their doctoral and "Habilitation" dissertations. The results show that a large percentage of authors do not know the service, and if they do, start using it very late. The paper concludes with recommendations for improving the information flow between ETD service and authors.*

### Introduction

Technical progress in electronic publishing affords increasingly sophisticated archiving and retrieval options for authors as well as readers of university document publishing services. However, an ETD service's success depends not only on technical features, but also on successful relationships with its "market", in particular, its potential authors - the creators of the service's content.

To create documents that not only look good today, but are also retrievable and will remain so tomorrow, authors must invest a sizeable amount of intellectual and

manual work to annotate their documents adequately, and these skills have to be acquired and practised. University ETD services play a central role in teaching these skills and in motivating authors to meet this challenge. To be able to do this, an ETD service needs an in-depth knowledge of its "market" to be able to supply authors with the information and service they need.

Acquiring such market knowledge requires answers to a number of questions: (a) What is the goal of the service? For example, should all authors be convinced to publish electronically, or should only those be supported who explicitly express an interest in publishing their dissertation online? (b) Does the service have other goals, for example university image-building; what financial constraints have to be considered? (c) Who are the target groups of the service? (d) Do they use the service, and why or why not; why do they contribute (or not); which of its features do or don't they like; and what do they expect or wish? (e) Which methods are adequate for finding answers to these questions? Empirical studies that address such issues are still comparatively rare, but their importance is increasingly being realised (cf. Jewell, 2000; Zhang, Lee, & You, 2001).

We investigated these questions in a case study of the Electronic Document and Publication Service of Humboldt University Berlin. After an overview of this service and the questions and goals of the study in section 2, we will describe the study and its results in sections 3-5, and conclude by proposing recommendations for improving the service in section 6.

### Questions and goals of the study

The Document and Publication Server at Humboldt University (HU) Berlin (<http://edoc.hu-berlin.de>), in the following referred to as the *edoc* service or *edoc*, has been

operated as a document server for the publication of scientific work since 1997. Doctoral dissertations have so far accounted for the largest proportion of published works. In two development projects, all preconditions for the complete publication process were created.<sup>1</sup> This included (a) providing formatting templates (in particular, for use in Microsoft Word, which is by far the most popular text processing SYSTEM among doctoral students) and instructions for authors, (b) converting documents into the SGML-based DiML format (Dissertation Markup Language), (c) assigning electronic signatures and timestamps that secure the authenticity and integrity of all documents, and (d) establishing a workflow for all employees involved in the publication process. At the beginning of 2001, edoc became a permanent service offered jointly by the HU computing centre (now computer and media service) and the university library. To publish on edoc, authors have to use the provided templates and structure their dissertations accordingly. In the preparation of their documents, authors are accompanied by intensive and qualified support, comprising information materials, training courses on the use of the Word template, and user support via telephone, email, or in person.

The edoc service has two main target groups: *authors* and *readers* of academic documents, in particular, doctoral and "*Habilitation*" dissertations. The objective of the service is to encourage authors to create content, and to encourage readers to access and utilize this content. In order to serve both groups well, the service team must know about the opinions, experiences, and intentions of both actual and potential users. In contrast to commercial services, the aim of edoc is not to convert every potential user into an actual user, but to make users able to decide whether to become "customers" or not. In other words, to put authors as well as readers into the position to freely decide whether to become *edoc authors* (*edoc readers*) or not (see also Wendland, 2002).

As is the case for most organisations, the edoc team knew more about people who use the service (authors in particular) than about those who don't. Members of the Digital Publishing Group interact closely with authors who publish their dissertation on edoc. From informally collected feedback, the edoc service team knew that most of the authors consider digital publication as a fast and cheap way to fulfill the German university publication requirement. However, many complain about an apparent lack in information flow since they had learned about the possibility of publishing online in a very late stage of

writing, or even after they had finished it. They comment that it would have been better to use the required template from the beginning of their writing, and regret to not have known about the possibility of online publishing earlier. Beside this, some authors find the rules they have to follow to format their document too strict and want the process to be easier to follow.

However, the service team did not know much about the remaining dissertation authors. For example, they did not know why these authors do not take advantage of the opportunity to publish online. Since the beginning of the edoc projects, a relatively stable (and arguably small) percentage of dissertations has been published online: Between 1998 and 2001 (data are only available for this period), 21% of all dissertations submitted at HU have been published on edoc (13% if the medical faculty is included in the statistics).<sup>2</sup> How should this proportion be interpreted? In particular, does it indicate that the remaining doctoral students and doctors do not *wish* to publish online, that they do not *feel capable* of publishing online and/or adequately supported in their attempts, or that they are simply *unaware* of this possibility?

What could be the factors that determine such outcomes? Authors' decisions whether to become an edoc author or not are likely to be affected by their knowledge about publishing opportunities, by their opinions about publishing their dissertation, by their views about the edoc service and digital publishing in general, and by the service's characteristics relevant to authors, e.g. information, public relations, training, and other support.

General attitudes towards digital publishing also play an important role in a *reader's* decision whether to become an edoc reader or not (in fact, in the decision whether or to what extent to use *any* digital document archives). In addition, the service's characteristics relevant to readers are important, in particular, the search functions and features of the edoc Web site.

To find out more about authors and readers, we developed two surveys. Since we<sup>3</sup> knew something about the edoc authors and wanted to find out more about them as well as about the edoc non-authors (or not-yet authors), we developed the *Digital Dissertation Questionnaire*. Since we did not know much about the edoc readers (except for the occasional informal feedback), and hardly anything about the edoc non-readers, we developed the *Document and Publication Server Website Questionnaire*. For reasons of space, we will only describe the first survey in detail, and briefly report relevant evidence from the second survey in the Conclusions.

1 "Digitale Dissertationen" of the HU central library and computing centre, [http://edoc.hu-berlin.de/epdiss/index\\_en.html](http://edoc.hu-berlin.de/epdiss/index_en.html), and "Dissertationen Online", a national project involving seven universities and libraries, <http://www.dissonline.de/>; s.a. [http://edoc.hu-berlin.de/e\\_projekte\\_en/](http://edoc.hu-berlin.de/e_projekte_en/)

2 Statistics are based on <http://amor.cms.hu-berlin.de/~h1775bvo/HTML/kfl-studstat.html>, [http://edoc.hu-berlin.de/statistik\\_PhDThesis.html](http://edoc.hu-berlin.de/statistik_PhDThesis.html), [http://edoc.hu-berlin.de/statistik\\_Habilitation.html](http://edoc.hu-berlin.de/statistik_Habilitation.html). The percentages 21% and 13% are the online ratios of doctoral dissertations.

3 One of the authors (Bert Wendland) is a member of the edoc team and, among others, responsible for author support and training courses.

## Method

**Participants.** Participants in this study were students working towards their doctoral dissertation or doctors engaged in post-doctoral research or qualification (usually, the "Habilitation") at Humboldt University Berlin.<sup>4</sup> The sample ( $n = 1180$ ) was obtained from a central HU email database which includes all members of HU (except the Medical Faculty) unless they explicitly requested to be excluded. Thus, the majority of potential respondents could be reached in this way.

**Materials.** The questionnaire was constructed to examine patterns of use and the degree of satisfaction with the services offered by edoc, as well as to assess general opinions and experiences regarding digital publishing. Questions addressed (a) how respondents had learned about the edoc service, (b) whether they intended to use edoc to publish the dissertation they are currently working on, (c) how they used and judged specific aspects of the edoc service, (d) their general attitudes towards digital publication, and (e) demographic information.

The questionnaire was developed and administered with the Rogator Online Survey Software.<sup>5</sup> Questions were implemented as multiple-choice and open-ended questions, with branching where applicable. Each question appeared on a new Web page, with a progress bar indicating the number of completed questions. The questionnaire was hosted on the Rogator server.

**Procedure.** The study was conducted over a 49-day period in spring 2003, initialized by a bulk emailing and the activation of a hyperlink on the edoc home page. Doctoral students and doctors were informed about the purpose of the study via email. They were asked to fill out the online version of the questionnaire. Alternatively, participants could choose to return a paper copy of the questionnaire (which was enclosed as .pdf attachment) via campus mail. They were given a 3-week-period to complete either the online or the paper questionnaire.

Visitors to the edoc Web site were informed about the purpose of both the Digital Dissertation survey and the Document and Publication Server Website Survey on the Edoc Survey Intro Page, which was placed on the server and hyperlinked from the home page. Visitors were encouraged to answer both surveys or only the second depending on whether they were HU doctoral students / doctors or not.

**Analysis.** Two data sources were used: the data collected by the Rogator software, and the logs of the edoc Web server. The qualitative and quantitative questionnaire responses were exported into SPSS for further

analysis. The requests for the Edoc Survey Intro Page as well as for all actions starting on it were logged - in particular, requests for the two questionnaires (redirected to the actual page hosted at the Rogator Web site). The server also logged the "referrer" for all requests, which, if non-empty, is the Web page from which a visitor came to a requested page by clicking on a hyperlink. Accesses from known robots were filtered out.

Visitors to the Intro Page could have been alerted to the survey (a) by our email, (b) by the hyperlink on the edoc home page, and/or (c) by a search engine. Since we did not publish the URL elsewhere, other options can be neglected. Therefore, it can also be assumed that a person who typed in the Intro Page URL, had learned about it from our email. A typed-in URL results in an empty referrer in the server log. We therefore concluded that all those visitors whose request for the Intro Page had an empty referrer were in fact from our email sample.

## Results

**Response rates and completion rates.** 360 people visited the Edoc Survey Intro Page. The data indicate that at least 325 of them came to that page as a direct result of receiving the email (nearly all visitors requested the Intro Page with an empty referrer, except for 5 visitors who were referred from a search engine, and 30 who reached it via the home page). So nearly one third of the sample was attracted to visit the Intro Page. However, the interest to follow through was lower: 159 people started the questionnaire, and 101 finished it. In addition, 3 people filled out hardcopy questionnaires. This indicates a response rate of 13.7% and a completion rate of 64.2%.

**Respondent characteristics.** 98 respondents specified their faculty or institute.<sup>6</sup> They came from a broad range of disciplines: the Natural Sciences (Biology: 13.3%, Computer Science: 8.2%, Physics: 8.2%, Mathematics: 7.1%, Geography: 7.1%, Chemistry: 4.1%, Psychology: 3.1%), the Social Sciences (Economics / Business: 18.4%, the Social Sciences Institute: 4.1%, Law: 2.0%), the Arts (Education, Languages: 5.1% each, History: 4.1%), and the Agricultural Faculty (8.2%).

69.3% were doctoral students, and the remaining were doctors (101 respondents).

Doctoral students main source of income was a job as research assistant ("Wissenschaftliche Mitarbeiter"; 73.2%), or a scholarship in a graduate programme ("Graduierertenkolleg"; 7.0%). The remainder received a differ-

4 In accordance with standard English and American terminology, we refer to the first group as "doctoral students", although technically, enrollment as a *student* is optional.

5 <http://www.rogator.de>

6 In the following, we report valid percent or percent of cases for each question, and list the total number of respondents who answered each question in parenthesis at the end of the sentence.

ent scholarship or worked outside the university (71 respondents).

Among the respondent were slightly more males (67.6%) than females. Most respondents (77.8%) were under 35 years of age (age range 25-46) and German (83.2%). Foreign students originated from 11 different countries. (Questions answered by 102, 99, and 101 respondents, resp.).<sup>7</sup>

**Knowledge about digital dissertation publishing.** Most participants said that they found out about the opportunity of publishing their doctoral or "Habilitation" dissertation on the Document and Publication Server from "this questionnaire" (44.1%), from a friend or colleague (30.7%), from the hyperlink on the HU library page (12.6%), from a professor or from a hardcopy information leaflet (3.1% each), from a search engine or from the examination office (2.4% / 1.6%) (127 respondents). 28.3% of 127 respondents wished they had learned earlier about the digital publishing opportunity.

**Attitudes towards digital dissertation publication.** Most people would prefer to publish their dissertation in book form: as a book by a publishing house (45.9%) or as a book by a copyshop (12.6%). 1.8% named microfiche. 19.8% intend to publish online only, and 62.2% to publish online in addition to a book publication (111 respondents). A chi-square analysis showed that women chose the "publish only online" option significantly less often than men (6% vs. 26%,  $p < 0.05$ ).

Half of the participants (49.5%) said that they had not decided yet whether or not to publish their dissertation on the Document and Publication Server. 34.9% (15.6%) said that they had already decided in favour of (against) electronic publishing (109 respondents). These three subgroups were then asked to specify reasons for their respective decisions.

(1) Asked about the reasons why they had not made the decision yet, 56.6% said that they would make the decision after they finish writing the dissertation. 35.8% respondents said that their dissertation submission date is far from now, so they would think about it later, and 13.2% that they still knew too little and would make a decision when they know more (53 respondents).

(2) Reasons for choosing online publishing were reported to be that "it's cheap" (60.5%), "it's a faster and easier publishing means than traditional publishing" (71.1%), and that one's documents "can be accessed by more people than with traditional publishing means" (65.8%). In addition, the increasingly wide use of the Internet was mentioned (38 respondents).

(3) Most participants who chose not to publish online (56.3%) agreed with the statement "In my discipline, online publication is considered informal and not as highly valued as a real publication", and the same number said they would "like to publish a real/physical book." Copyright concerns were another reason. Two participants said that they gave up their plan of publishing online because the format requirements were too complicated to be fulfilled (16 respondents).

**Concrete steps taken towards a digital dissertation publication.** People intending to publish on the edoc server have to fulfill certain format requirements, i.e. they must prepare their dissertations according to a "dissertation template". Of those participants who chose to publish online, nearly half said that they had not decided yet when to start using the dissertation template (47.4%). 36.8% started or intended to start using the template after they finished writing, compared to only 13.2% at the beginning and 2.6% in the middle of the writing process (38 respondents).

The "author information" on the Web site was named as most helpful for learning how to use the template by 55.6%. Beside that, a third of the respondents asked their friends or the computing centre staff. Only 2 persons named the training course offered by the Digital Publishing Group as helpful for learning how to use the template. (36 respondents).

In case of experiencing problems with the template, 50% sent email and 29.4% asked the computing centre staff personally; 26.5% phoned to get a solution. 5 participants said that so far, they had not experienced any problems (34 respondents).

Only 1 person had attended the training course.<sup>8</sup> Reasons for not attending were named by 33 participants: that they used a text-processing SYSTEM other than Microsoft Word, the SYSTEM that the course concentrates on (40.6%), or that they already knew how to use the template (28.1%). 6 persons said that they did not know the training course, and 3 that it did not fit their schedule.

**Participants' recommendations on how to improve information flow and better promote the edoc service.** 44 participants gave suggestions. Proposals included: (a) to have examination office staff hand out information leaflets, preferably at an early stage; (b) to inform professors about the possibility of digital dissertation publication, and to ask them to promote it; (c) to organise doctoral seminars, or other forms of active conversation between doctoral students, conducted within the faculty, or jointly for related scientific fields; and (d) to offer lectures and tutorials in the faculties.

7 All results reported in the following were investigated for possible relationships to these demographics; no significant differences were obtained except in one case (reported below).

8 In the earlier question, this respondent named the computing centre staff (who run the courses) as most helpful, which may be a sign that s/he also considered the course itself helpful. The 2 who had singled out the training course did not indicate that they had attended it.

## Discussion

Respondents' demographics mirrored those of the targeted population of all authors of HU doctoral and "Habilitation" dissertations. The distribution over scientific fields (grouped by Natural Sciences, Social Sciences, and Agricultural Science, and Arts as described above) was highly similar; the Arts were however underrepresented relative to the Natural Sciences (14.3% and 51.1% in the sample of respondents compared to 23.9% and 42.1% among all HU doctoral students and doctors). People working on a post-doctoral dissertation were overrepresented (30.7% vs. 10.5% in the targeted population), as were males (67.6% vs. 57.3% in the targeted population). These observations may be related to one another: The proportion of men is higher outside the Arts disciplines than within them, and the proportion of men is higher in the group of "Habilitation" authors than in the group of doctoral dissertations authors.<sup>9</sup>

Knowledge about digital dissertation publication was very limited. Up to one half of respondents only learned about this opportunity from this questionnaire investigation. This suggests that the edoc service has to be promoted strongly to support its goals - to offer every doctoral student and doctor an opportunity to publish online.

The expressed attitudes indicate that digital publication is broadly accepted. We expect that adequate promotion of the service will have a good chance of success, since every sixth respondent is already strongly in favour of digital publishing, and half are still in the decision process. These authors may change their current preference for a book publication if they learn more about the advantages of digital publishing.

Financial issues appeared to play only a minor role; only a very small number of respondents (5) said that their eventual choice of publishing means "will depend on my budget at that time". It is conceivable that the stable financial situation of most participants is a reason both for the expressed general preferences for book publication and the expressed disregard for the high expenses incurred by a book publication.

The majority of potential authors started, or intended to start, using the template at a late stage in their dissertation publishing writing process. This may be a consequence of the fact that many only learn(ed) about the edoc service quite late. It may also be an indicator of the perceived difficulty of the template and of the effort needed to use it, an effort many will not want to make during the process of content development. However, re-structuring an existing text is much more labour-intensive than the early adoption of a template, and it may also be argued that "structured writing" aids content development.

Around a fifth of all our respondents said that they cannot use or do not need the training course. The first most probably represent the "Latex users group" (around 20% of all edoc authors) who are often highly skilled at utilising templates for structuring their documents. Many of the remaining authors can profit from the training course; however, the "training course attendee" group has in the past consisted mainly of medical doctoral students and doctors, who were not part of our sample. The sample did contain many of the third group, the "training course non-attendees using MS Word". Their low level of knowledge of and interest in the training course mirrors the generally low level of knowledge and activities concerning digital dissertation publication.

## Conclusions: Recommendations for improving the edoc service

The results show that the information flow presents the largest problem. Based on the results, we propose three recommendations for a better information flow and marketing promotion.

**Inform doctoral students when they start their dissertation project.** An alarmingly large percentage of doctoral students are not even aware of the existence of the edoc service. Several participants suggest that all doctoral students should be informed about the digital publishing opportunity when they start their dissertation project. This could happen in the form of a leaflet about edoc, including some basic points about the edoc service, like its goals, the format requirements, the Web site address, and especially the advantages of digital publication.

The advantages of digital publication should be strongly emphasized in the leaflet. Some of the respondents said that they do not want to publish online because they think online publication is informal and their copyright is not well protected. With more knowledge of the advantages of digital publishing, such as persistent identifiers and addresses, digital signatures and timestamps for protection against distortion, and guaranteed long-time archiving, they may change their mind. A digital publication also offers more options for content, such as inclusion of multimedia elements.

Due to the absence of a uniform start of dissertation projects (people do not have to enroll as doctoral students), this may require a close cooperation with a number of institutions and people, from the personnel department, to the matriculation office and grant-giving bodies. Since *everyone* starting on a dissertation project is likely to get into contact with a supervisor at an early

<sup>9</sup> population statistics based on <http://amor.cms.hu-berlin.de/~h1775bvo/HTML/kfl-studstat.html>; the distributions in 1998-2001 reported there were treated as a representative baseline.

stage, it is essential to make academic staff aware of the edoc service and of the necessity to promote it.

**Send a regular email newsletter.** Several participants suggested to send out emails to inform doctoral students about news on the edoc server. The edoc service may make use of the mailing list created for the present study to distribute a regular newsletter. The newsletter can be sent once per month or once each semester, depending on current events. It could include a list of newly published dissertations, the training course schedule, news about digital publication events, etc.

**Offer more services with the formatting.** Many participants complained about the complexity of the digital dissertation format rules. Some of them said that they gave up their plan of publishing in digital form because they found themselves unable to fulfill the formatting requirements.

The edoc service may be able to relax some of the format rules through technical improvements. In addition, more tutorials could be provided to teach the formatting.

Few doctoral students use the dissertation template at the beginning or in the middle of their dissertation writing. Most of them use the template after they finish writing. More help is needed here, which may need to be provided by a paid service. The edoc service might act as a switchboard to match dissertation authors with formatters.

**Further research.** To find out why certain marketing strategies work (or not), further work should investigate more closely the differences within the population of authors. For example, the success of edoc's marketing and support strategies varies across faculties and institutes: In the Natural Sciences, the proportion of doctoral dissertations that are published is very high, up to 67% in Mathematics. In the Medical Faculty, unlike elsewhere, the training course is a popular and helpful support measure, yet the overall proportion of dissertations published online is among the lowest in the university (7%). As a first step, we therefore intend to repeat this study in the Medical Faculty.

The issues investigated in this study go beyond the question of dissertation publication. The difficulties au-

thors experience in the use of the dissertation template may be regarded as indicating, among other things, more general problems in the structured use of metadata. Such problems can occur in authoring as well as in reading and researching documents. This was one of the issues investigated in our second survey, the Document and Publication Server Website Questionnaire. This addressed all visitors to the edoc Web site (for a detailed description, see Berendt et al., in preparation).

Its results showed that doctoral and post-doctoral dissertations are regarded as the central component of edoc, further supporting the argument that edoc is a good location for publishing a dissertation. They also confirmed that people still experience many difficulties when trying to use metadata in a structured way - they still prefer to "search" and "browse" to locate documents they need. Further research is needed on how to help students and researchers use metadata productively, in research, education, and training.

## Bibliography

- BERENDT, B., BRENSTEIN, E., LI, Y., & WENDLAND, B. ( in preparation). Using electronic thesis archives: An investigation of Web site visitors' attitudes to search features and e-publishing
- JEWELL, C. ( 2000). Campus Collaboration: The University of Waterloo Electronic Thesis Project . Paper presented at the Third International Symposium on Electronic Theses and Dissertations, March 2000, St. Petersburg, [http://www.lib.uwaterloo.ca/ETD/Campus\\_Collaboration.PDF](http://www.lib.uwaterloo.ca/ETD/Campus_Collaboration.PDF)
- WENDLAND, B. (2002). Braucht ein Dokumentenserver eine Policy? Der Dokumenten- und Publikationsserver der Humboldt-Universität zu Berlin . *Bibliotheksdienst* 36(6), 741-746 , [http://bibliotheksdienst.zlb.de/2002/02\\_06\\_07.pdf](http://bibliotheksdienst.zlb.de/2002/02_06_07.pdf)
- ZHANG, Y., LEE, K, & YOU, B. ( 2001). Usage patterns of an Electronic Theses and Dissertations (ETDs) SYSTEM . *Online Information Review* 25(6), 370-377