

User-driven content selection for digitisation - the eBooks on Demand Network

Silvia Gstrein / Günter Mühlberger, University of Innsbruck, Library,
silvia.gstrein@uibk.ac.at

Abstract

Following EC's i2010 strategy European libraries are currently systematically digitising and making available their cultural heritage. Very often, however, researchers, readers and users demand access to historical books „here and now“. This is exactly where the eBooks-on-Demand (EOD) network starts from, providing a trans-European digital document delivery service for end-users from all over the world. Currently the EOD network comprises over 20 libraries from 10 European countries. Since 2007 several thousand PDF eBooks have been produced, delivered to users from over 30 countries and made available to the public through participating digital libraries. User reactions are very positive and more and more libraries are interested in the service. This paper is about the following topics: (1) the user-driven approach of content selection for digitisation, (2) financing such a service, (3) results from a user survey, (4) future perspectives within supplying web 2.0 platforms.

Keywords: Digitisation, eBooks on Demand, user selection

1. Introduction

In line with EC's i2010 strategy European libraries and other cultural institutions are currently systematically digitising and making available their cultural heritage to a wider public. Due to the enormous amount of material, it will take some decades until all books, journals and other library material will be available in digital form. Ideally, all the works in a collection or holding should be considered for digitisation. In practice, however, this is rarely feasible and choices must be made. According to the MINERVA Good Practices Handbook (Clissmann 2004) the following criteria may be suggested for setting preferences: (1) Access to material which would otherwise be unavailable or of limited availability, (2) wider and easier access to very popular material, (3) condition of the originals, (4) preservation of delicate originals by making digital versions available as an alternative, (5) project theme, (6) copyright and IPR, (7) availability of existing digital versions, (8) cost of digitisation and (9) appropriateness of the source material for online viewing.

In many cases the chosen approach is determined by the source of funding, preferably in combination with the institution's digitisation policy. This way of processing digitisation only very rarely acknowledges interests of users apart from point (2) above which highlights popularity as a criterion of selection. This raises the following questions: (1) how can individual users' needs be considered in digitisation?, (2) what happens to and who will take care of those materials not covered by all those projects?, (3) more precisely, who will digitise books of minor languages or those from special or smaller collections or institutions?

All those considerations formed the starting point for a service called "eBooks on demand": an electronic document delivery service instigated and co-funded by the user. What was needed was a model for ongoing, permanent digitisation, a

coordinated European initiative where selection of material necessarily works bottom-up rather than top-down and which is instigated by the individual reader.

Experience has shown that such a project needs co-financing by the reader requesting a certain book to be digitised. By saying so I only refer to the co-financing of scanning itself. The important but often neglected issue of subsequent costs such as storage, access, long term preservation and migration costs will not be touched here.

2. EOD network and service

Of course, such an additional library service can't be maintained within the context of usual (mass) digitisation only, but needs extra resources for order management, customer communication, payment procedures and so on. Therefore, a structure allowing for efficient processing of orders was designed. At the end a central-decentral service network was implemented, where some processes such as OCR or online payment are hosted centrally. From October 2006 to June 2008 a pilot project and evaluation was carried out within the framework of "Digitisation on Demand", co-funded by the eTEN programme, consisting of 13 libraries from 8 European countries. By July 2008, a self-sustained network was up and running with these 13 libraries as founding members.

In the following I will briefly describe how EOD works. The starting point is the online catalogue of any participating library. There, the EOD button is placed with all items available for digitisation. At the moment, these are books which fulfil at least one of the following criteria: (1) items not yet digitised, and (2) public domain books. Any user interested in a certain book tagged with the EOD button simply needs to click on this button. The library then receives the order, scans the requested book and transfers the replica via FTP to the central server located at the network coordinator Innsbruck University. Each library manages and processes the orders in a central database accessible via a normal web browser. Here, also the digital object is created. After the completion of the payment process, the user downloads the PDF from his personal tracking page. After some time the library eventually integrates the images into its digital library or repository and thus makes them available to the general public.

From a library's point-of-view the EOD service is a mixture of central and peripheral services sharing two core components: The order data manager and the digital object generator. For its own orders only, each library has access to the Order Data Manager, a central database with web-interface. According to its needs and objectives any participant is able to adapt and customize the website texts, automatically generated email texts, etc. Core part of the Order Data Manager is the Digital Object Generator, a central service for generating eBooks. It allows OCR recognition (antiqua and gothic), automated cover generation, PDF & RTF delivery, creation of Abbyy XML and the generation of the streaming link for downloading. Online payment in the form of credit cards also forms part of the central services – so the individual library doesn't need to grapple with business like this.

Currently, more than 20 libraries from 10 European countries are offering the service, ranging from Portugal in the far west to Estonia in the East. The University of Innsbruck library is both, the coordinator and central service provider of this network at the same time. Concerning the types of libraries involved - there are 6 national libraries, several university and state libraries as well as research and academy of sciences` libraries.

| | |
|--|---|
| <p>Austria</p> <ul style="list-style-type: none"> • University and Regional Library of Tyrol (co-ordinator) • University Libraries of Graz and Vienna, Library of the Medical University of Vienna • Vienna City Library <p>Germany</p> <ul style="list-style-type: none"> • University Libraries of Regensburg, Greifswald and Humboldt-Universität zu Berlin • Bavarian State Library • Saxon State and University Library (Dresden) <p>Denmark</p> <ul style="list-style-type: none"> • The Royal Library <p>Estonia</p> <ul style="list-style-type: none"> • National Library • University Library of Tartu <p>France</p> <ul style="list-style-type: none"> • Medical and Dental Academic Library of Paris | <p>Czech Republic</p> <ul style="list-style-type: none"> • Moravian Library in Brno • Research Library in Olomouc • Library of the Academy of Sciences in Prague • National Technical Library <p>Hungary</p> <ul style="list-style-type: none"> • National Széchényi Library of Hungary • Library of the Hungarian Academy of Sciences <p>Portugal</p> <ul style="list-style-type: none"> • National Library <p>Slovakia</p> <ul style="list-style-type: none"> • University Library in Bratislava <p>Slovenia</p> <ul style="list-style-type: none"> • National and University Library |
|--|---|

Table 1: Overview over EOD libraries

3. Pricing and user experience

Up to now some 3200 books have been digitised and delivered to customers - books that wouldn't have been otherwise digitised. This makes up some 840.000 scanned pages for nearly 2000 users worldwide. The top 3 libraries receive about one request per working day and deliver 250 to 350 books per year. The average delivery time is one week.

The average price of an order is about 50 EUR. The price of an eBook is set by the respective library and is calculated from a base price plus the number of the pages. There is still a quite heterogeneous pricing system as at the moment each library sets up its own prices. In average, digitisation via EOD includes a basic fee of 10 Euro plus 0.15 to 0.30 Euros per page. In view of this price range, a certain degree of harmonization certainly needs to be agreed upon in the near future.

Experience has shown that the revenue gained from digitisation merely co-funds digitisation – to be more precise: the actual personnel costs for scanning including additional costs for the service as such. Quite clearly then a customer-oriented service leads to higher costs than mass digitisation, which lays huge emphasis on efficiency and economisation. Given, though, that the master files remain with the library and is later made freely accessible to the public, the user by no means can be made to pay the actual total cost of digitisation. The user must be seen as the instigator and co-funder of digitisation; in turn the library profits by building up its digital library in the long run.

In 2008 we carried out a user survey among EOD customers to receive feedback to the service. We also wanted to find out what users thought about our pricing policy. The findings were that 30 % of our customers thought the price was high or very high, but still the overall price-value relation was found acceptable by the majority.

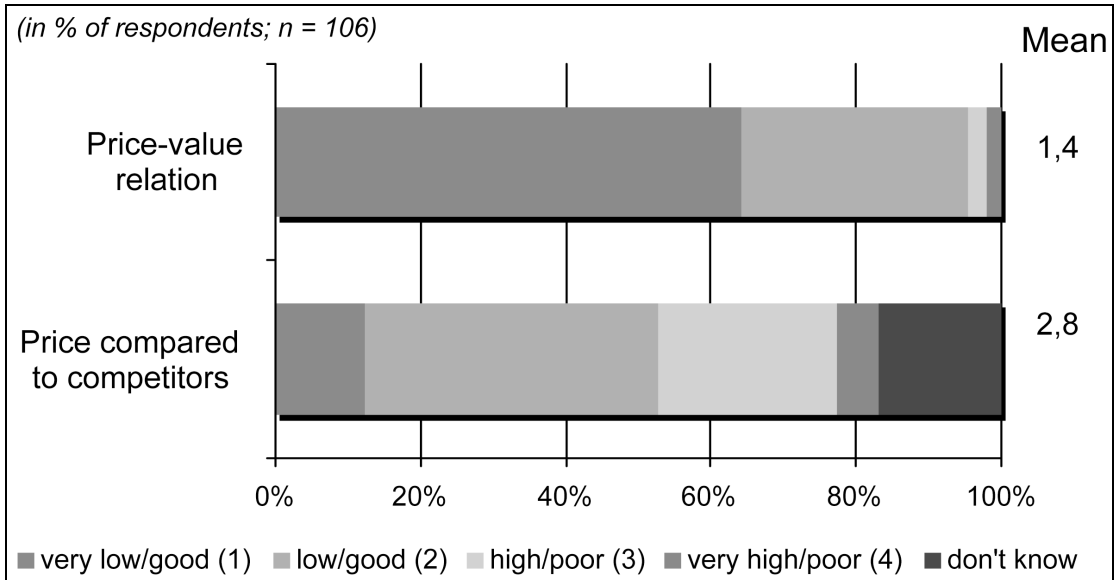


Figure 1: Users' evaluation of price-value relation

It's noteworthy that the overwhelming majority of users are either researchers or people who require eBooks for "professional or scientific use" (over 60 percent). Second place (16 percent) is book collectors and people who could be said to count among special interest groups such as amateur historians, collectors, or ethnographers for example.

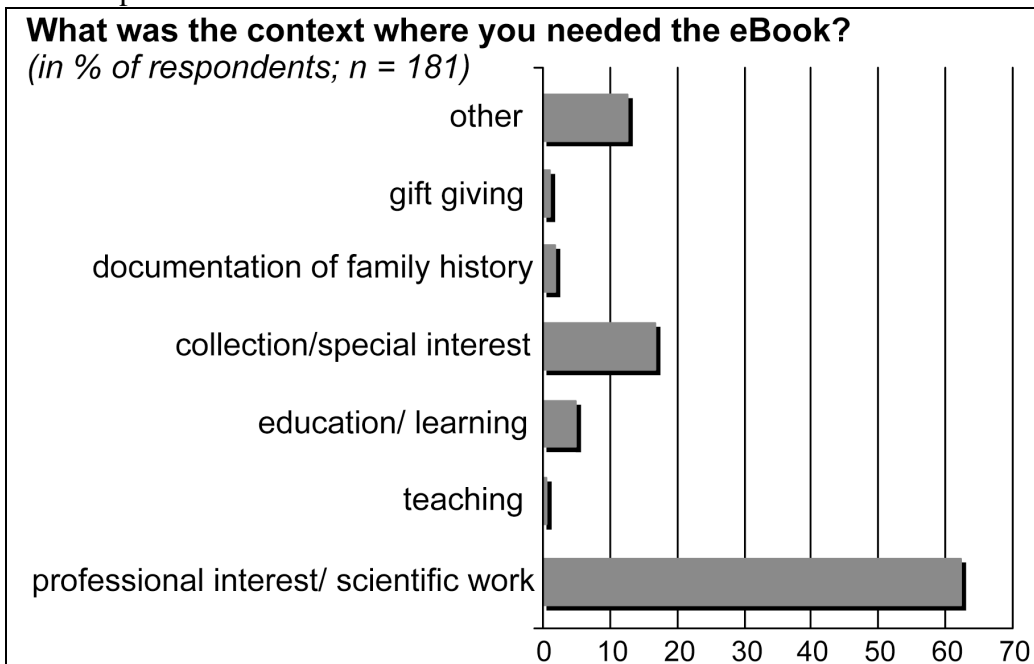


Figure 2: Users' domain of interest

Asked why they chose the EOD service, almost half of the interviewees answered that without EOD the book would have been "impossible or difficult to access". This shows that EOD achieved one of its main goals of being an additional alternative for accessing books.

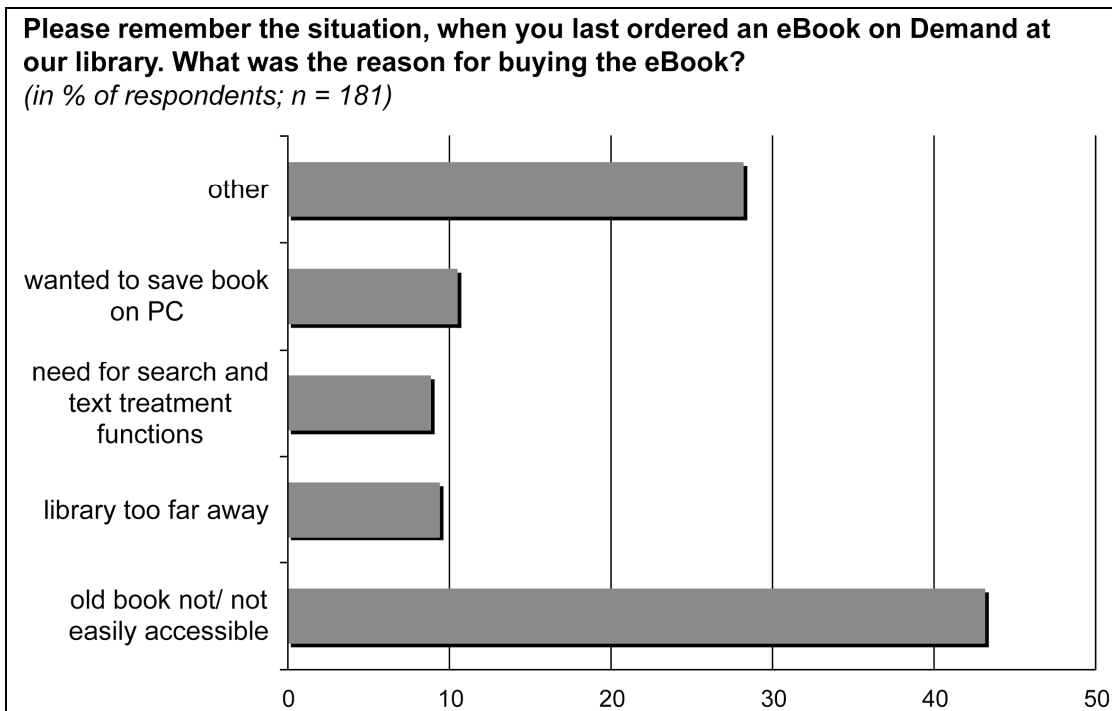


Figure 3: Reasons for ordering

In the EOD customer survey, 60% of our customers told us that they usually print out selected pages or even the whole book on paper. Thus, there was an obvious request for “re-materialization” of digital material. In response, we have lately been offering “reprints on demand”, supplementing the digital file. This service is also carried out in a centralised way. Each library only needs to take care of scanning images and some more metadata. Everything else connected with the service such as image enhancement, the creation of pre-press PDF and related files is carried out by the central coordinator at the University of Innsbruck library.

4. Future perspectives

Within the EC Culture programme a new project was initiated in May 2009 with duration of 4 years. The project will focus on larger scale involvement of three main target groups: participating libraries, requesting users and the general public. 20 libraries from 10 European countries take part in this project which mainly concentrates on the following 3 objectives: (1) to enlarge the EOD network by additional European libraries, especially those from countries not yet represented (2) to take EOD as a best practice model for any other European-wide network of this kind; furthermore to train stakeholders (libraries, museums, or other cultural operators) to run such a multinational cultural service based on state-of-the-art information technologies and (3) to support intercultural dialogue among readers and users of historical books with the help of web 2.0 technology.

Within this latter goal web 2.0 based social platforms such as Wikipedia, LibraryThing or Goodreads will be supplied with information on selected historical books. Readers all over the world interested in a specific book thus will be able to easily interact with each other, exchange information and share reading and research experiences – independently from where they are and which background they may have.

During the last years the internet has become more and more of a “social network”. Cultural organisations need not only “offer” their cultural heritage in the internet, but they need to pro-actively “supply” web 2.0 platforms with their content, since these are the sites where the majority of users are actually present. Not only do users create and edit information, but upload texts, images, and videos, they “meet” other users, they “share” content, they “cooperate”, “rate” and generally “take part” in a digitally mediated community. Whereas many sites, such as library catalogues or repositories, are only visited sporadically in the event of retrieving some clearly defined data, people have their personal accounts at web 2.0 applications and use them regularly – comparable to the way they regularly read a newspaper, meet friends at a cafe or visit a museum. First, EOD will therefore explore the technical feasibility to automatically generate Wikipedia source-code for selected books which have been ordered and digitised via the EOD service to then later on add metadata to articles about the respective authors. Other social platforms will be tackled in a second step.

5. Conclusions

Experiences made so far indicate that “on demand” services are desirable ways to make books accessible either in digital or in re-printed form. EOD has proven to be an important additional library service, making holdings all over Europe more accessible to customers and providing researchers with material they need “on demand”, here and now.

In fact more “on demand” services could be imagined, such as digitisation on demand for the blind and visually impaired and the creation of „real“ eBooks with a corrected full text approximating 99% accuracy. This would pave the way to transfer eBooks onto mobile-devices, making the written word available every time, everywhere.

Acknowledgement

This work programme has been funded with support from the European Commission. This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

References

- Clissmann, Ciaran (2004): Good practices handbook. Version 1.3. Minerva Project. Minerva Working Group 6 Identification of good practices and competence centres. Online: http://www.minervaeurope.org/structure/workinggroups/goodpract/document/goodpractices1_3.pdf [Access: 25/10/2009]
- Mező, Zoltán; Svoljsak, Sonja; Gstrein, Silvia (2007): EOD - European network of libraries for eBooks on demand. In: Research and advanced technology for digital libraries, p. 570–572.
- Mühlberger, Günter / Gstrein, Silvia (2009): eBooks on Demand (EOD): a European digitisation service. IFLA Journal (35/1), 35–43.
Online: <http://archive.ifla.org/V/iflaj/IFLA-Journal-1-2009.pdf> [Access: 25/10/2009]
- Svoljsak, Sonja; Gstrein, Silvia (2007): EOD - eBooks on demand. In: Users and use of DL & economics of digital libraries, p. 229.