

Lecture by Ms. Cornelia Diebel from the Deutsche Nationalbibliothek,
"Acquiring Digital Information and Metadata concerning E-legal Deposit"

On March 6, 2013, Ms. Cornelia Diebel of the German National Library, "Deutsche Nationalbibliothek" or DNB in German gave a lecture "Acquiring Digital Information and Metadata concerning E-legal Deposit" in the Tokyo Main Library of the National Diet Library (NDL).

Ms. Diebel joined the DNB in 1992 and was appointed a team manager of the IT-Department in 2007. She is also responsible for the "Taskforce Online Resources" which aims to increase the online resources of the library and to develop automated processes concerning the collection and ingest of online resources.



<<Ms. Diebel giving her lecture>>

Introduction

The DNB, which celebrated its 100th anniversary last year, was established with two sites as the result of the German reunification in 1990 through a merger of its predecessor institutions: the

National Diet Library Newsletter

"Deutsche Bücherei Leipzig" founded in 1912 and the "Deutsche Bibliothek Frankfurt am Main" founded in 1946. The DNB has holdings of 27 million items (on the basis of bibliographic records) and the number of employees is about 600.



<<Two buildings of the DNB in Leipzig and Frankfurt, in pictures>>

Legal Deposit Legislation in Germany

The mission of the DNB is to collect, catalog, and index all materials published in Germany from 1913, preserve them for the long term, and provide the general public with access to them.

By the "Law regarding the German National Library" of June 2006 and the "Legal Deposit Regulation" of October 2008, this mission has been extended to publications in "immaterial form", or in general, "online publications." All commercial and non-commercial publishers in Germany are obliged to submit two mandatory copies of their works, free of charge, to the DNB.

The DNB is not the only library in Germany which has the mandate to collect publication by the legal deposit system. Germany comprises sixteen states and each state has its own state constitution and is largely autonomous in regard to its organization, especially in the field of cultural and educational policy. Therefore some states have their own deposit library. The DNB tries to harmonize the workflows among libraries so that the publishers and other distributors do not need to double their investments in order to fulfill the duty to provide the online resources to the libraries. In fact this is not that easy because of some regulations in the Copyright Law which limit the scope of possible actions.

Automated workflow for online publications

The DNB has established an automated workflow for acquisition, cataloging and preservation of online publications under the Legal Deposit Regulation revised in 2008, based on the previous experience of acquiring a part of online publications such as online dissertations (e-thesis), which started in 1998 on a voluntary basis.

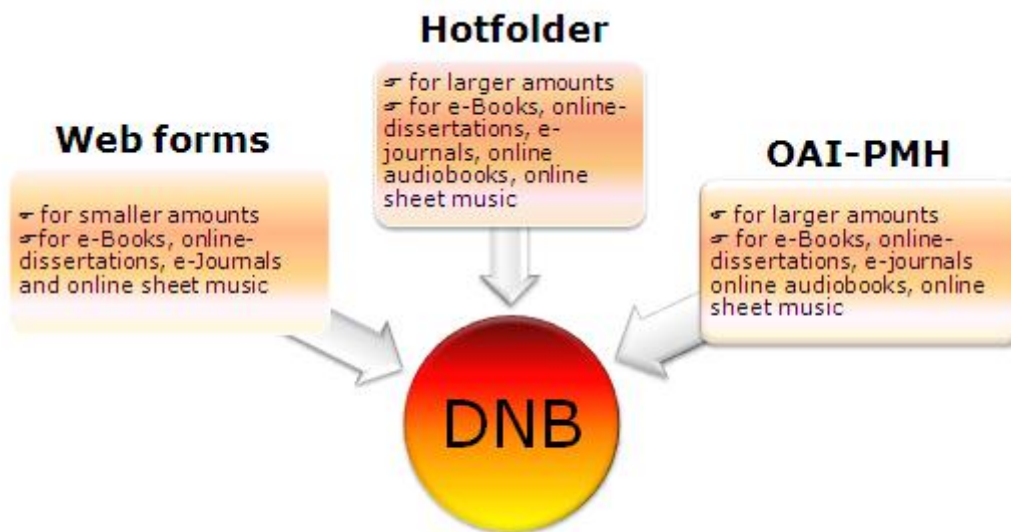
In order to simplify the process of deposit by publishers, the following four conditions need to be fulfilled.

The first condition is that metadata or records and the electronic item are collected together.

The second condition is that the DNB and publishers need to agree on the metadata format. The DNB has several options for metadata delivery. Presently, the DNB accepts the following formats of online publications: ONIX 2.1 for Books (the DNB is now preparing to accept ONIX 3.0 as an improvement of ONIX 2.1), MARCXML or XMetaDissPlus; further formats will be added eventually. After a testing phase metadata are stored in the catalogue database without any increments, as they are input by the publishers. The contents of the fields are checked only by control samples and cannot be manually processed because of the vast number of records.

The third condition is agreement on the accepted file formats. The DNB prefers open standards and needs to know what formats are used in order to ensure digital long term preservation and accessibility of the objects in the reading rooms of the DNB. Currently, in addition to PDF (PDF/A (an ISO-standardized version of the PDF specialized for the digital preservation of electronic documents) and all other types of PDF), the now heavily used e-book format EPUB, documents in HTML and different image and audio formats can be automatically archived. The DNB can now also collect online audiobooks which are usually produced in MP3. More formats are to follow, the amazon e-book format AZW for instance.

The last, but crucial, condition is agreement between the publishers and the DNB on the interfaces for data exchange. The use of different interfaces is the main technical challenge. It is impossible to establish the workflow without using the same interfaces. At the moment, the DNB offers three different types of interfaces: Web forms, Hotfolder (putting data in a certain folder at a DNB server by using sFTP (Secure File Transfer Protocol) or WebDAV and OAI-PMH (harvesting automatically). Before anyone can use one of the submission methods it is necessary to register with the DNB and to receive authorization.



<<Details of three different types of interfaces the DNB offers>>

The DNB is also working on the collection, archiving and indexing of websites. Regarding the collection of websites, the DNB has decided to adopt a selective approach. As the first step, the DNB is focusing on the collection of organizations and public authorities. The legal collection definition includes the digitized versions of older publications released on the Internet. To improve the archiving process, the DNB aims to obtain the master files, which allow the information to be saved without any loss.

The new metadata for online publications are ingested to the database of the DNB through the automated process. For example, checking for the existence of parallel print editions is done and if that they exist, linked to them – at this point, bibliographic information of print editions is carried over into the records of the online publication. Online publications will be indexed automatically with Dewey Decimal Classification and subject headings, especially when no parallel print version exists.

Users can access deposited online materials only on the premises of the DNB. In addition, the Copyright Law in Germany does not allow users to make digital copies or send them by e-mail.

Ms. Diebel closed her lecture by mentioning the continuous efforts to enhance e-legal deposit in Germany, saying that the DNB will encourage more and more publishers to deposit their publications through development of additional ways of delivery including preparation for new mapping.



<<View from the audience>

The following Q&A session that lasted 60 minutes dealt with a wide range of questions from the floor.

Before starting the Q&A session, Mr. Toru Toyoda, the then Deputy Director General of the Acquisitions and Bibliography Department, gave a brief presentation about e-Legal deposit of online publications in Japan. After that, questions from the audience were posed. The topics included the DNB's permanent identifier for online publications, how to deal with publishers who are not willing to deposit their publications, how to establish good relations with publishers, etc.

The lecture meeting was full of suggestions as to how we should deal with e-Legal deposit of online publications in Japan, which is going to start in July 2013.

For the content of her lecture, please also refer to the slides below:

[Slides](#) [PDF Format, 1.49MB]