

Preservation and Access of Digitally Deposited Newspapers

The Comellus Project

Name: Matti Hosio, Project Manager and Minna Kaukonen, Senior Planning Officer

Employer: The National Library of Finland, Centre for Preservation and Digitisation

Contact information:

matti.hosio(at)helsinki.fi, tel. +358 50 3182 320 (until 7th June 2013)

Minna.Kaukonen(at)helsinki.fi, tel. +358 50 4155 450

Biographical statement of presenter:

Matti Hosio

Matti Hosio works as Project Manager at the Centre for Preservation and Digitisation of the National Library of Finland until June 2013. He graduated as Master of Science in Technology in 2008 from the University of Oulu, Finland. His current activities include the development of digital newspaper handling process for the National Library of Finland.

Minna Kaukonen

Minna Kaukonen is Senior Planning Officer and Deputy Director of the Centre for Preservation and Digitisation at the National Library of Finland. Minna Kaukonen has been involved in several national and international digitisation projects on the European level since 2002. She has been Secretary of the LIBER Steering Committee on Digitisation and Resource Discovery in 2009 - 2011.

Language of the presentation: English

The Comellus Project (2011 – 2014) of The National Library of Finland is improving the accessibility of newspapers by electronic deposition. The project is a pilot which renews the newspaper handling process towards a fully digitalised mode: from digital deposition to reception and handling. This means both developing the knowledge base and exploring new ways of cooperation. Two newspaper publishers are involved in the project. Computer Output Microfilm -printing ensures the long-term preservation of the digitally deposited newspapers.

The National Library has comprehensive digital newspaper collections until 1910 in its Historical Newspaper Library (<http://digi.nationallibrary.fi>) which is available for all the users, free of charge. The volume of the newspapers is about two million pages. The Library has had legal deposit since 1707. There are eight free deposit libraries around the country. The digitally deposited newspapers will eventually be available in all of them.

The Development towards the Digital Newspaper Handling Process

The National Library of Finland and its Centre for Preservation and Digitisation receives all the newspapers published in Finland, in paper format. At the moment the Centre microfilms all this material and delivers the microfilms to libraries both in Finland and abroad.

Because the production of newspapers is a completely digital process nowadays, the development of a digital newspaper process also at The National Library of Finland has been chosen as one of the key development areas of the Library. At the same time this development contributes to expanding knowledge in the South Savo Area where the Centre for Preservation and Digitisation is located. The region has active public and private organizations and cooperation in the field of digitization and digital archiving.

The newspaper process development has benefited from benchmarking other library actors having digital deposit, e.g. Estonia and Norway.

The Different areas of development

The new process accelerates the handling of newspapers considerably and enables even more high-quality long-term preservation both on microfilm and in a digital format. The digital deposition format includes more metadata, does not require OCR or suffer from print-related flaws. A great deal of the work required to prepare the material for long-term archival is done automatically by the software. The result of the development project is a new digital newspaper process including the technical production environment, software and servers.

The process was piloted in spring 2013. The new digital process was developed and tested parallel to the traditional analog process. This helped in creating an efficient process that still is in many ways familiar to the operators. In connection with the pilot, the permanent staff – mainly from the current analog newspaper process – was trained to use the new workflow and software.

The digital deposit system

- The newspaper publishers

The module for the digital deposit of newspapers was created by a company that has delivered the majority of Finnish newspapers their digital editorial systems. The deposit format was carefully described and the functional requirements were established for the tender, in order to have the demanded content in the correct format.

The newspaper publishers deliver the print-PDF files and the associated metadata daily to the National Library. The print-PDF files include the exact

images of the printed pages and therefore act as digital surrogates. The provided metadata is delivered along with the print-PDFs and enables automatic processing of the deposits. The editorial metadata opens up new perspectives for the construction of more usable interfaces with varied search possibilities. Data about e.g. page structures, articles, authors or geography can be valuable for the end user.

The module for the digital deposit can be easily reused by other newspapers in the future.

Reception

- The National Library of Finland

The electronic reception control system was built from scratch by us in the project. This made an effect in designing the process because the features could be implemented fully based on the process workflow.

All the work was carefully documented, in the form of user manuals, maintenance manuals and the documentation of the newspaper process.

The process used agile methodology and gained from the experiences, comments and feedback of the permanent staff.

Computer Output Microfilm

The digitally deposited newspapers are printed on microfilm by using Computer Output Microfilm (COM) technology.

The long-term preservation is secured, since the microfilms produced are estimated to last at least 500 years. The microfilming of printed newspapers is replaced by the COM conversion in the process.

The optimization of quality as well as the workflow were considerably developed in the project. The COM process is a critical part of the whole.

The interface

The newspaper history can be discovered via the same interface: digi.nationallibrary.fi . Some of the older newspapers have often either continued until today or have been predecessors to modern newspapers. Eventually the old digitized newspapers and the digitally deposited newspapers form a digital entity without comparison.

Agreements needed

The Act on Collecting and Preserving Cultural Material does not include digital deposit at the moment and only printed newspapers are deposited in Finland.

Digital deposit requires separate agreements with The National Library and newspaper publishers.

The copyright issues must be taken into account and the digital access to deposited newspapers is to be negotiated.

Kopioisto is the Finnish copyright organisation for authors, publishers and performing artists and is an important counterpart in this process. The National Library has a joint project with Kopioisto going on, regarding expanding the use of the 20th century newspapers for research purposes and for the general public. There have been preliminary negotiations with some newspapers regarding joining the digital deposit system. The talks will continue in the autumn and next spring and a special effort will be allocated to this.

The Finances

The project has been financed by the European Commission and its European Social Funds (ESF – Leverage from the EU), the City of Mikkeli, The National Library of Finland and two newspaper publishers. The project will last until the end of May 2014.

The number of personnel has varied during the project. In addition to the Project Manager there have been IT experts and operators working for the project. The overall budget is a little over 650 000 euros.