

The extent of Newspaper Digitisation in European Libraries



The Hague, 15th of March 2013

What is the extent of newspaper digitisation in European libraries? As part of the Europeana Newspapers project (www.europeana-newspapers.eu), which aims to aggregate and refine newspaper content to the freely accessible online service Europeana (www.europeana.eu), The European Library undertook a survey to identify and analyse all newspaper collections digitized by national, research and public libraries in Europe by 2012. A series of twelve questions was circulated with a focus on newspaper titles and time range, metadata in use, data distribution capabilities, and quality of digitisation including the technologies used for refinement. The survey also helped identify new potential partners. The most important findings of the survey are set out below:

Access to digitised newspapers is nearly always free of charge.

Of the 47 respondents, at least 40 (85%) offered free access to their digitised newspapers. Only one library had pay per view, whilst another three offered subscription services for users (i.e. paid access per day or per month). Only four libraries licensed their newspaper contents to other groups (e.g. schools and universities).

Access to twentieth-century content remains problematic.

Over half of the surveyed libraries (27 out of 47; 57%) have a cut off date beyond which they will not publish digitised newspapers on the web. Most frequently, this is based on a 70 year sliding scale, meaning that content after 1942 is inaccessible in digital form. 23% (eleven out of 47) had an agreement with a rights organisation so that in-copyright digitised newspapers could be published. However, this tended to be restricted to individual titles rather than collective agreements for complete collections.

There is still much to be done to exploit the richness of digitised newspaper content

36% (17 out of 47) of the surveyed libraries have not used any form of Optical Character Recognition (OCR) on their digitised newspaper content, meaning that searching through the full text of newspaper content is not possible. And while 64% have used OCR, only 17 of the libraries (36%) exposed the resulting full text to the viewer, indicating that they had reservations about the quality of the OCR'd text. There were also low numbers (36%) for those that had undertaken zoning and segmentation and only six libraries (13%) had included features such as faceted browsing or extracting entities such as place or name.

Associated Partners

The Europeana Newspaper project will extend its network. Associated partners and network partners will be able to attend project workshops, information days and relevant meetings and are asked to provide content to the Europeana Foundation. During the survey, 35 libraries indicated they wished to be considered as Associate Partners. Guided by a series of principles from the survey, the General Assembly approved of the following eleven Associate Partners:

- National Library of Czech Republic
- National and University Library Ljubljana, Slovenia
- National and University Library of Iceland
- National and University Library Zagreb, Croatia
- St. Cyril and Methodius National Library, Bulgaria
- Lucian Blaga Central University Library, Cluj-Napoca, Romania
- National Library of Wales
- National Library of Portugal
- National Library of Spain
- National Library of Belgium
- National Library of Luxembourg

Europeana Newspapers Survey Report available at www.europeana-newspapers.eu.

For any further information please contact Ulrike Kölsch at Staatsbibliothek zu Berlin via ulrike.koelsch@europeana-newspapers.eu.

Project Partners:

Berlin State Library

National Library of the Netherlands

The Europeana Newspapers project is partially funded under the ICT Policy Support Programme (ICT PSP, http://ec.europa.eu/ict_psp) as part of the Competitiveness and Innovation Framework Programme by the European Community. This document reflects only authors' views. The EC is not liable for any use that may be done of the information contained therein.



National Library of Estonia
University of Helsinki, National Library of Finland
National Library of France
CCS Content Conversion Specialists GmbH
National Library of Latvia
University of Beograd
Dr. Friedrich Tessmann Library
University of Salford

Austrian National Library
Hamburg State and University Library
National Library of Poland
LIBER Foundation
National Library of Turkey
University of Innsbruck
The British Library
The European Library

The Europeana Newspapers project is partially funded under the ICT Policy Support Programme (ICT PSP, http://ec.europa.eu/ict_psp) as part of the Competitiveness and Innovation Framework Programme by the European Community. This document reflects only authors' views. The EC is not liable for any use that may be done of the information contained therein.

