EUROPEANA NEWSPAPERS

A Gateway to European Newspapers Online

In this ICT project funded under the EC's Competitiveness and Innovation Framework Programme (CIP 2007–2013) a group of **17 European public and private partner institutions** have joined forces and will – over the next 3 years – provide more than 18 million newspaper pages to The European Library (TEL) and Europeana.

europeana
newspapers

PROJECT ACTIVITIES

Aim
The project aims at improving access to digitized newspapers in Europe. Each participating library will deliver digitized newspapers (images, fulltext and metadata) to Europeana where full text cross-searches through the materials will be possible. The project is co-ordinated by the Staatsbibliothek zu Berlin – Preußischer Kulturbesitz (Berlin State Library – Prussian Cultural Heritage Foundation, Germany).

Refinement of digitized newspapers
The main purpose of this work area is the enrichment of already available digitized newspaper images with the results of OCR (Optical Character Recognition), OLR (Optical Layout Recognition), article segmentation and page class recognition, and – on a smaller scale – Named Entity Recognition (NER).

Evaluation and Quality Assessment
The quality of the refined and aggregated data (images, texts, metadata), as well as the entire refinement process and used technologies will be evaluated. Quality prediction tools will be tested and optimised.

Aggregation and presentation
Eventually the project will collect some 18 million newspaper pages for **Europeana / The European Library**: 8 million pages already available at partner institutions (images, images with full texts, images with full texts by article), 10 million pages refined with OCR within the project, and 2 million pages refined with OCR, OLR and article segmentation, again to be done within the project. Additionally a survey of European newspaper holdings in non-partner libraries will be undertaken.

Metadata best practice recommendations
The project will design a **Europeana Data Model** for newspapers that will be derived from the analysis of all metadata models used both in- and outside the project consortium. Based upon the results achieved best practice recommendations concerning digitization, refinement, workflows and metadata for newspapers will be given.

Dissemination and Exploitation
All information about the project will be published on the project website at: www.europeana-newspapers.eu. Workshops and other events (e.g. information days) will be organised during the project life-cycle with the objective to increasing the visibility of digitized newspaper content in Europe. Please check the project website regularly.

PROJECT TIMELINE

2012
Project website
Workshop on refinement

2013
National Information Days
Workshop aggregation & presentation

2014
Workshop European newspapers & digital agenda
Final Press Conference

PROJECT PARTNERS

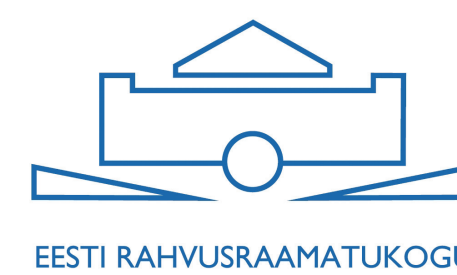- LIBER – Ligue des Bibliothèques Européennes de Recherche / Association of European Research Libraries
- BnF Bibliothèque nationale de France
- KB Koninklijke Bibliotheek / National Library of the Netherlands
- STAATS- UND UNIVERSITÄTSBIBLIOTHEK HAMBURG CARL VON OSSIETZKY
- EESTI RAHVUSRAAMATUKOGU / NATIONAL LIBRARY OF ESTONIA
- Landesbibliothek Dr. Friedrich Teßmann / Bibliotech Provinziela / Biblioteca Provinciale
- Staatsbibliothek zu Berlin Preußischer Kulturbesitz
- University of Salford MANCHESTER
- Österreichische Nationalbibliothek
- The European Library / Connecting knowledge
- MILLI KÜTÜPHANE / NATIONAL LIBRARY, ANKARA – 1946
- universität innsbruck
- BRITISH LIBRARY
- УНИВЕРЗИТЕТСКА БИБЛИОТЕКА "СВЕТОЗАР МАРКОВИЋ" У БЕОГРАДУ
- BIBLIOTEKA NARODOWA
- LNB LATVIJAS NACIONĀLĀ BIBLIOTĒKA
- CCS
- 1640 THE NATIONAL LIBRARY OF FINLAND

Staatsbibliothek zu Berlin
Preußischer Kulturbesitz