Europeana Newspapers
A Gateway to European Newspapers Online

A group of 17 European public and private partner institutions have joined forces in the Europeana Newspapers project and will, over the next 3 years, provide more than 18 million digitized newspaper pages to the Europeana service.

Each library participating in the project will distribute digitized newspapers and full-texts to Europeana.

The Europeana Newspapers project is funded under EC’s CIP 2007–2013 and aims at the aggregation and refinement of digitized newspapers for The European Library and Europeana.

The project has started on 1st February 2012. CIP-ICT-PSP-2011-5-Best Practice Network project no. 297380.
Refinement

In the Europeana Newspapers project refinement methods will be used to convert an abstract data model into implementable data structures. The Europeana Newspaper project will use refinement methods for:

- Optical Character Recognition (OCR): electronic conversion of scanned images of handwritten, typewritten or printed text into machine-encoded text.
- Optical Layout Recognition (OLR): article segmentation
- Named Entity Recognition (NER): seeks to locate and classify atomic elements in text into predefined categories such as the names of persons, organizations, locations etc.
- Page Class Recognition

Why use refinement methods?

These refinement methods will make the newspaper content easier to search and retrieve via the Europeana Portal. It will allow the user to find individual articles, perform subject searches and display the full text online.

Evaluation and Quality Assessment

The Europeana Newspapers project will build upon the successful general-purpose tools implemented in the EU-funded IMPACT project (Improving Access to Text) and develop them further into an evaluation and quality assessment infrastructure.

Project Activities

Refinement

In the Europeana Newspapers project refinement methods will be used to convert an abstract data model into implementable data structures. The Europeana Newspaper project will use refinement methods for:

- Optical Character Recognition (OCR): electronic conversion of scanned images of handwritten, typewritten or printed text into machine-encoded text.
- Optical Layout Recognition (OLR): article segmentation
- Named Entity Recognition (NER): seeks to locate and classify atomic elements in text into predefined categories such as the names of persons, organizations, locations etc.
- Page Class Recognition
Aggregation and presentation
The project will coordinate the aggregation and presentation of all digitized newspapers for Europeana:
- All metadata will be aligned to the Europeana Data Model (EDM)
- A full text index and newspaper content browser will be developed to improve the Europeana searching and retrieving experience.
- The digital newspaper collections will be analysed according to metadata in use, quality of digitization, accessibility and data distribution infrastructure.

Metadata best practice recommendations
The Europeana Newspapers project will design and release a comprehensive metadata model based on de-facto standards such as METS and ALTO. Partners will disseminate the model among stakeholders in order to find a common agreement and to make it a best practice example for newspaper digitization in Europe.

Workshops and Events
The following workshops and events will be organised with the objective of increasing the visibility of newspaper content in Europeana by raising the standard and awareness of digitized newspapers at European level:
- Workshop on refinement
- Workshop on aggregation and presentation
- Workshop on newspapers in Europe and the digital agenda for Europe
- National Information Days (in 10 European countries)
- Final Press Conference
Follow us
europeana
newspapers

Visit our website
http://www.europeana-newspapers.eu

Become a fan on Facebook
http://www.facebook.com/EuropeanaNewspapers

Follow us on
http://twitter.com/#!/eurnews

Join the LinkedIn
http://www.linkedin.com/groups?gid=4425919

Check the project presentation at
http://www.slideshare.net/Europeana_Newspapers

Contact us at info@europeana-newspapers.eu