Europeana Newspapers: A Gateway to European Newspapers Online

The Europeana Newspapers Project aims at the aggregation and refinement of newspapers for The European Library and Europeana. In addition, the proposed BPN addresses challenges particularly linked with digitized newspapers: • use of refinement methods for OCR, OLR/article segmentation, and named entity recognition (NER), and page class recognition to enhance search and presentation functionalities for Europeana customers, • quality evaluation for automatic refinement technologies, • transformation of local metadata to the Europeana Data Model (EDM) • metadata standardization in close collaboration with stakeholders from the public and private sector. Each library participating in the project will distribute digitized newspapers and full-texts free of any legal restrictions to Europeana. There will be a special focus on newspapers published during the First World War, thus providing a meaningful addition to the resources aggregated by the current Europeana 1914-1918 project. Newspapers provided within the proposed project will also cover a wider time period, so that numerous historical events from before and after 1914-1918 linking directly to events within that time period will be accessible for customers of The European Library and Europeana. This will allow people to easily research important local, national, or European events in a broad European context; something that has so far not been possible. The newspapers will be of as much interest to the general public as researchers, since the material is understandable without specialized historical knowledge and concerns virtually every European family. The refinement technologies to be applied encompass Optical Character Recognition (OCR), Optical Layout Recognition (OLR)/article segmentation, Named Entity Recognition (NER), and page class recognition. Quality assurance and quality prediction mechanisms will be optimized to monitor and control all refinement steps. Best practice recommendations for newspape

Project ID card

- Funded under: The Information and Communication Technologies Policy Support Programme
- Area: CIP-ICT-PSP.2011.2.1 - Aggregating content in Europeana
- Total cost: €5.16m
- EU contribution: €4.12m
- Project reference: 297380
- Execution: From 01/02/2012 to 31/01/2015
- Project status: Running
- Contract type: The Information and Communication Technologies Policy Support Programme BPN Best Practice Network

Contact information

LIEDER Hans-Jörg
VON DER HEYDT STRASSE 16-18
10785 BERLIN
GERMANY
Tel. +4930266434000
Fax.
E-mail: **hans-joerg.lieder@sbb.spk-berlin.de**

Participants

Coordinator: STIFTUNG PREUSSISCHER KULTURBESITZ, GERMANY

Participating partners:

Organisation	Country
STIFTUNG PREUSSISCHER KULTURBESITZ	GERMANY
UNIVERSITAET INNSBRUCK	AUSTRIA
OESTERREICHISCHE NATIONALBIBLIOTHEK	AUSTRIA
HELSINGIN YLIOPISTO	FINLAND
BIBLIOTHEQUE NATIONALE DE FRANCE	FRANCE
FREIE UND HANSESTADT HAMBURG	GERMANY
CCS CONTENT CONVERSION SPECIALISTSGMBH	GERMANY
Landesbibliothek Dr. Friedrich Tessmann	ITALY
Biblioteka Narodowa	POLAND
EESTI RAHVUSRAAMATUKOGU	REPUBLIC OF ESTONIA
NATIONAL LIBRARY OF LATVIA	REPUBLIC OF LATVIA
UNIVERZITET U BEOGRADU	SERBIA
STICHTING LIBER	THE NETHERLANDS

KONINKLIJKE BIBLIOTHEEK	THE NETHERLANDS
MILLI KUTUPHANE BASKANLIGI	TURKEY
THE UNIVERSITY OF SALFORD	UNITED KINGDOM
THE BRITISH LIBRARY	UNITED KINGDOM

See this factsheet online:
http://ec.europa.eu/information_society/apps/projects/factsheet/index.cfm?project_ref=297380