

Start des Europeana Newspapers Projekts

Berlin, 1. Juli 2012



Im Rahmen des Europeana Newspapers Projekts machen siebzehn europäische Partnerinstitutionen in den nächsten drei Jahren mehr als 18 Millionen Zeitungsseiten über das Online-Portal Europeana zugänglich. Europeana bildet ein zentrales Verzeichnis von über 23 Millionen digitalisierten Büchern, Gemälden, Filmen, Ausstellungsstücken sowie Archivalien aus ganz Europa.

Das Europeana Newspapers Projekt wird gefördert vom Competitiveness and Innovation Framework Program 2007-2013 der Europäischen Kommission und wird über The European Library digitalisierte Zeitungsbestände zusammenstellen und diese technisch aufbereitet präsentieren.

Das Hauptziel des Projekts besteht darin, den Nutzern die Zeitungsinhalte über einen Content Browser mit spezieller Suchoberfläche direkt zugänglich zu machen. Der Content Browser wird in das Europeana Portal integriert und das Suchen nach einzelnen Wörtern oder nach Wortfolgen innerhalb der Zeitungstexte ermöglichen. Dies geht weit über die Suchmöglichkeiten herkömmlicher Bibliothekskataloge hinaus, in denen normalerweise nur anhand des Datums oder des Zeitungstitels gesucht werden kann. Der direkte Zugriff auf die digitalisierten Zeitungsinhalte erspart den Nutzern zudem den Gang in Sonderlesesäle, in denen historische Zeitungen für gewöhnlich nur unter strengen Nutzungsbedingungen bereitgestellt werden.

Das Projekt konzentriert sich auf die spezifischen technischen Herausforderungen bei der Aufbereitung digitalisierter Zeitungen wie Optical Character Recognition (OCR), Optical Layout Recognition (OLR), named entity recognition (NER), Artikelsegmentierung sowie der Klassifizierung von Seiten. OCR bezeichnet die elektronische Umwandlung gescannter Texte in maschinen-lesbare Zeichen. OLR ermöglicht die Identifizierung und Trennung einzelner Artikel auf einer gescannten Zeitungsseite mit mehreren Artikeln. Durch NER können die Zeitungstexte nach normierten Einheiten wie Personennamen, Körperschaften oder Geographika durchsucht werden.

Im Rahmen des Projekts werden außerdem die Qualität der genannten technischen Verfahren evaluiert sowie Empfehlungen für die Umwandlung lokaler Metadaten in das Europeana Data Model entwickelt, indem ein Netzwerk aus privaten und öffentlichen Einrichtung aufgebaut wird.

Das Europeana Newspapers Projekt wird von der Staatsbibliothek – Preußischer Kulturbesitz koordiniert. Aktuelle Informationen werden unter www.europeana-newspapers.eu bereitgestellt. Für weitere Informationen wenden Sie sich bitte über info@europeana-newspapers.eu an Hans-Jörg Lieder oder Thorsten Siegmann an der Staatsbibliothek zu Berlin.

Projektpartner:

Staatsbibliothek zu Berlin
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

National Library of the Netherlands
Österreichische Nationalbibliothek
Staats- und Universitätsbibliothek Hamburg
National Library of Poland
LIBER Foundation
National Library of Turkey
Universität 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.



Europeana in Kürze:

Europeana ist eine mehrsprachige online zugängliche Sammlung von mehreren Millionen digitalisierten Objekten aus europäischen Museen, Bibliotheken, Archiven und Filmsammlungen.

Aktuell ermöglicht Europeana den Zugang zu über 23 Millionen Büchern, Filmen, Gemälden, Museumsobjekten und Archivalien aus 2.200 europäischen Kulturinstitutionen.

www.europeana.eu

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.

