

Početak projekta Europeana Newspapers

Beograd, 20. juli 2012.



Grupa od 17 evropskih institucija učestvuje u projektu “Europeana Newspapers” u okviru kog će u naredne tri godine obezbediti više od 18 miliona novinskih stranica za onlajn servis Europeana. Europeana predstavlja pristupnu tačku za milione digitalizovanih knjiga, slika, filmova, muzejskih objekata i arhivskih zapisa iz Evrope.

Projekat “Europeana Newspapers” je finansiran od strane Evropske komisije CIP 2007-2013 i ima za cilj agregaciju i prečišćavanje novina za The European Library (TEL).

Svaka biblioteka koja učestvuje u projektu će distribuirati digitalizovane novine u punom tekstu putem Europeane. Cilj projekta je da novinske sadržaje učini direktno dostupnim za korisnike putem posebnog interfejsa koji će biti integrisan u portal Europeane i omogućiti upite u vidu fraza ili pojedinačnih reči u samom tekstu novina. Ovakav vid pretrage predstavlja daleko više od standardnog bibliotečkog kataloga koji obično omogućava samo pretragu po datumu ili naslovu.

Projekat se bavi pitanjima karakterističnim za digitalizovane novine kao što su korišćenje metoda za prečišćavanje teksta za OCR, OLR/segmentaciju članka i prepoznavanje imenskih entiteta (NER), i prepoznavanje klase stranice. Metodi prečišćavanja se koriste za konvertovanje apstraktnih modela podataka u primenljive strukture podataka. OCR je elektronsko pretvaranje skeniranih slika rukopisnih, kucanih ili štampanih tekstova u mašinski kodiran tekst. OLR je elektronska segmentacija članaka iz skenirane stranice koja sadrži više od jednog članka. NER locira i klasifikuje atomske elemente u tekstu u predefinisane kategorije kao što su ime osobe, lokacije i organizacije.

Projekat će vrednovati kvalitet tehnologija koje se koriste za prečišćavanje teksta i pretvarati lokalne metapodatke u EDM standard (Europeana Data Model) i to u bliskoj saradnji sa interesnim grupama (stejholderima) iz javnog i privatnog sektora.

Koordinator projekta “Europeana Newspapers” je Državna biblioteka u Berlinu (Staatsbibliothek zu Berlin). Budite u toku sa razvojem projekta Europeana Newspapers na sajtu www.europeana-newspapers.eu. Za dodatne informacije kontaktirajte Hans Jorg Lidera ili Torstena Zigmara putem e-mejla info@europeana-newspapers.eu

Partneri u projektu:

Državna biblioteka u Berlinu
Nacionalna biblioteka Estonije
Univerzitet u Helsinkiju
Nacionalna biblioteka Francuske
CCS Content Conversion Specialists GmbH
Nacionalna biblioteka Latvije
Univerzitet u Beogradu
Biblioteka dr Fridrih Tesman
Univerzitet u Salfordu

Nacionalna biblioteka Holandije
Austrijska nacionalna biblioteka
Univerzitet u Hamburgu
Nacionalna biblioteka Poljske
LIBER fondacija
Nacionalna biblioteka Turske
Univerzitet u Insbruku
Britanska biblioteka
The European Library (TEL)

Ukratko o Europeani

Europeana je višejezična onlajn kolekcija nekoliko miliona digitalizovanih objekata iz evropskih muzeja, biblioteka, arhiva i audiovizuelnih kolekcija. Europeana trenutno pruža integrisani pristup preko 23 miliona knjiga, filmova, slika, muzejskih objekata i arhivskih dokumenata za oko 2200 institucija – davalaca sadržaja iz cele Evrope.

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.

