

Почетак пројекта Europeana Newspapers

Београд, 26. јули 2012.



Група од 17 европских институција учествује у пројекту “Europeana Newspapers” у оквиру ког ће у наредне три године обезбедити више од 18 милиона новинских страница за онлајн сервис Европеана. Европеана представља приступну тачку за милионе дигитализованих књига, слика, филмова, музејских објеката и архивских записа из Европе.

Пројекат “Europeana Newspapers” је финансиран од стране Европске комисије CIP 2007-2013 и има за циљ агрегацију и пречишћавање новина за The European Library (TEL).

Свака библиотека која учествује у пројекту ће дистрибуирати дигитализоване новине у пуном тексту путем Европеане. Циљ пројекта је да новинске садржаје учини директно доступним за кориснике путем посебног интерфејса који ће бити интегрисан у портал Европеане и омогућиће упите у виду фразе или појединачних речи у самом тексту новина. Овакав вид претраге представља далеко више од стандардног библиотечног каталога који обично омогућава само претрагу по датуму или наслову.

Пројекат се бави питањима карактеристичним за дигитализоване новине као што су коришћење метода за пречишћавање текста за OCR, OLR/сегментацију чланка и препознавање именских ентитета (NER), и препознавање класе странице. Методи пречишћавања се користе за конвертовање апстрактних модела података у применљиве структуре података. OCR је електронско претварање скенираних слика рукописних, куцаних или штампаних текстова у машински кодиран текст. OLR је електронска сегментација чланака из скениране странице која садржи више од једног чланка. NER лоцира и класификује атомске елементе у тексту у предефинисане категорије као што су име особе, локације и организације.

Пројекат ће вредновати квалитет технологија које се користе за пречишћавање текста и претварати локалне метаподатке у EDM стандард (Europeana Data Model) и то у блиској сарадњи са интересним групама (стејкхолдерима) из јавног и приватног сектора.

Координатор пројекта “Europeana Newspapers” је Државна библиотека у Берлину (Staatsbibliothek zu Berlin). Будите у току са развојем пројекта “Europeana Newspapers” на сајту www.europeana-newspapers.eu. За додатне информације контактирајте Ханс Јорг Лидера или Торстена Зигмана путем и-мејла info@europeana-newspapers.eu

Партнери у пројекту:

Државна библиотека у Берлину
Национална библиотека Естоније
Универзитет у Хелсинкију
Национална библиотека Француске
CCS Content Conversion Specialists GmbH
Национална библиотека Латвије
Универзитет у Београду
Библиотека др Фридрих Тесман
Универзитет у Салфорду

Национална библиотека Холандије
Аустријска национална библиотека
Универзитет у Хамбургу
Национална библиотека Пољске
LIBER фондација
Национална библиотека Турске
Универзитет у Инсбруку
Британска библиотека
The European Library (TEL)

Укратко о Европеани

Еuropeана је вишејезична онлајн колекција неколико милиона дигитализованих објеката из европских музеја, библиотека, архива и аудиовизуелних колекција. Европеана тренутно пружа интегрисани приступ преко 23 милиона књига, филмова, слика, музејских објеката и архивских докумената за око 2200 институција – давалаца садржаја из целе Европе.

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.

