



28.06.2012.



Euroopa ajalehed jõuavad Europeanasse

Rahvusraamatukogu osaleb projektis “Euroopa ajalehed”, mille raames tehakse kolme aasta jooksul Europeana vahendusel kättesaadavaks u 18 miljonit lehekülge eri riikides ja eri keeltes ilmunud ajalehti. Ühtekokku lööb projektis kaasa 17 Euroopa institutsiooni.

Europeana kaudu on kättesaadavad miljonid digiteeritud raamatud, maalid, filmid ja muuseumide ning arhiivide varad üle Euroopa. Euroopa ajalehtede projekti rahastatakse Euroopa Komisjoni konkurentsivõime ja uuendustegevuse raamprogrammist (2007-2013) ning selle eesmärk on ajalehtede sisu töötlemine ja koondamine Euroopa Raamatukokku.

Iga projektis osalev raamatukogu annab oma digiteeritud ajalehed ja täistekstid Europeanasse, mille kaudu tehakse need lõppkasutajatele kättesaadavaks. Europeanasse integreeritava spetsiaalse ajalehtede veebiliidese kaudu saab kasutaja võimaluse lehitseda ajalehti nii kalendripõhiselt kui ka otsida ajalehekstidest sõna ja fraasi järgi, mida enamik raamatukogude katalooge perioodika osas seni ei võimalda.

Projekt hõlmab ajalehtede digiteerimisega kaasnevat tööd nagu tekstituvastus Optical Character Recognition (OCR), artiklite eristamine Optical Layout Recognition (OLR), nimede määramine Named Entity Recognition (NER) ja küljenduselementide tuvastamine. Kõiki neid tehnoloogiaid kasutatakse, et konverteerida abstraktne andmemudel töötlemist võimaldavateks andmestruktuurideks.

OCRi ehk tekstituvastuse abil eraldatakse digitaalsest pildifailist tekst ja luuakse sellest redigeeritav tekstifail. OLR eristab leheküljel paiknevad artiklid, juhul kui neid on rohkem kui üks. NER määratleb olemasolevate registreeritud baasil teksti väiksemad elemendid (isiku-organisatsiooni-, kohanimed jm.) eelnevalt kindlaks määratud kategooriatesse.

Projekti käigus hinnatakse tekstituvastusmeetodite kvaliteeti ja viiakse kohalikud metaandmed vastavusse Europeana andmemudeli standardiga, tehes seejuures tihedalt koostööd huvirühmadega nii avalikus kui ka erasektoris. Projekti juhib Berliini Riigiraamatukogu (Staatsbibliothek zu Berlin – Preussischer Kulturbesitz).

Lähem teave: Hans-Jörg Lieder ja Thorsten Siegmann, e-post: info@europeana-newspapers.eu, www.europeana-newspapers.eu

Eesti Rahvusraamatukogus juhib projekti „Euroopa ajalehed” Krista Kiisa (e-post: Krista.Kiisa@nlib.ee)

Teate edastas Eesti Rahvusraamatukogu avalike suhete osakond
Annika Koppel tel. 630 7260

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.



Taustainfo

Projektis osalevad:

1. Berliini Riigiraamatukogu (projektijuht)
2. Hollandi Rahvusraamatukogu
3. Eesti Rahvusraamatukogu
4. Austria Rahvusraamatukogu
5. Helsingi Ülikool, Soome Rahvusraamatukogu
6. Hamburgi Riigi- ja Ülikooliraamatukogu
7. Prantsusmaa Rahvusraamatukogu
8. Poola Rahvusraamatukogu
9. Salfordi Ülikool
10. Firma CCS Content Conversion Specialists GmbH
11. Sihtasutus LIBER
12. Läti Rahvusraamatukogu
13. Türgi Rahvusraamatukogu
14. Belgradi Ülikooli Raamatukogu
15. Innsbrucki Ülikool
16. Landesbibliothek Dr. Friedrich Tessmann
17. Briti Raamatukogu

Europeana on mitmekeelne veebipõhine kollektsioon, mis sisaldab miljoneid digiteeritud objekte Euroopa muuseumidest, raamatukogudest, arhiividest ja audiovisuaalsetest kollektsioonidest. Europeana kaudu on hetkel kättesaadavad ligi 23 miljonit raamatut, filmi, maali, museaalset objekti ja arhiividokumente enam kui 2200 sisupakkujalt üle Euroopa.

www.europeana.eu

