

Building Europeana v1.0: towards a Large-Scale Content Ingestion

Julie Verleyen

Europeana Office,
Koninklijke Bibliotheek,
National Library of the Netherlands

Europeana version 1.0 is an EU project that will develop a fully operational Europeana website. This will allow users to access over 10 million digital items from Europe's rich cultural and scientific heritage by 2010.

This - already important - challenge is only the top of the iceberg of content that will continue to flow *to* and *from* Europeana after the release of the first operational version of the service.

Meeting this challenge implies the setting up of a system that supports a large-scale workflow: at one end of the chain there are processes that smoothly ingest data submitted by Content Providers and Aggregators, while (web) services at the other end interact with components giving access to the data. We present here the preparation of such system and in particular its standard-based approach to the functional modelling of the workflow, using the Open Archival Information System (OAIS) Reference Model with a focus on what happens at the interface between the data submitters and Europeana, the receiver. Update, traceability, duplication management, conversion: Europeana addresses these functional needs at the scale of hundreds of institutions and millions of items.

We also illustrate that besides the technical solutions the organisational aspects become more and more crucial when everything increases (number and size of data sets, number of parties involved, multiplicity of sources). A sound network strategy, clear agreements, rules and workflows are to be carefully established next to performing tools for partner and data management system, tracking of information, data transfer and storage facilities.