

Personalization based on users' requirements in MANUSCRIPTORIUM (European Digital Library of Manuscripts)

Tomáš Psohlavec¹, Jakub Heller², Roberto Caldelli³, Tomasz Parkoła⁴

¹ AiP Beroun s.r.o., Talichova 807, Beroun, 266 01, Czech Republic

² Cross Czech a.s., Růžová 17, Praha 1, 11000, Czech Republic

³ Media Integration and Communication Centre Firenze, Viale Morgagni 65, 50134
Florence, Italy

⁴ Poznan Supercomputing and Networking Center, Noskowskiego 12/14, 61704 Poznan,
Poland

tp@aipberoun.cz; hellerj@crossczech.cz; caldelli@lci.det.unifi.it, tparkola@man.poznan.pl

Abstract: Manuscriptorium (<http://www.manuscriptorium.eu>) is a resource provided by the National Library of the Czech Republic (<http://www.nkp.cz>) as a strategic leader and content coordinator as well as by the AIP Beroun Ltd. (<http://www.aipeberoun.cz>) as a technical provider and system administrator. Manuscriptorium became a major resource at the European level due to realization of the ENRICH project (<http://enrich.manuscriptorium.com>) funded under eContent+ programme. The main results of the project are now available and therefore a short Manuscriptorium case study can be presented, including demonstration of new end-user features such as personalized collections and virtual documents.

Keywords: Manuscripts, Incunabula, Digital Library, Interoperability, Shared repositories, User requirements

Introduction

The aim of the ENRICH project is the creation of a base for the European digital library research environment for study of specific historical cultural heritage consisting of manuscripts, incunabula, early printed books, historical archival materials, etc. The main innovation of ENRICH lies in a common easy-to-use interface which enables concentration of dispersed resources into a unique research environment and retrieval of data from distant servers. The project allows the users to search and access documents which would otherwise be hardly accessible by providing free access to almost all digitized manuscripts in Europe. During the demo session we will present selected partners documents/collections within the end-users interface in order to demonstrate aggregation results and at the same moment we will demonstrate the Manuscriptorium end-users features with special focus on the newly available "Personalized Virtual Library" feature. Four typical ways of cooperation will be shortly mentioned using the real-life examples.

Manuscriptorium Platform

ENRICH builds upon the existing Manuscriptorium platform (<http://www.manuscriptorium.com>) adapted to needs of organizations holding repositories of manuscripts. Manuscriptorium is a resource provided on-line by the National Library of the Czech Republic (<http://www.nkp.cz>) as a coordinator as well as by the AiP Beroun Ltd. (<http://www.aipeberoun.cz>) as a technical provider and system administrator. A searchable Open catalogue of historical documents is an important part of the service along with the Digital Library that contains all digital documents aggregated so far. Various front-end interfaces are implemented into the system including the OAI-PMH and Z39.50 interfaces which ensures inclusion to partners portals (Manuscriptorium DL is searchable via OAI-PMH from the TEL portal, i.e. any ENRICH partner contributing to Manuscriptorium automatically enriches the European Digital Library) as well as OAI-PMH harvester which aggregates partner's documents into Manuscriptorium, set of converters for performing various metadata formats conversions etc.

Data Aggregation

The ENRICH groups together the richest owners of digitized manuscripts among national libraries in Europe; ENRICH partner libraries possess almost 85% currently digitized manuscripts in the national libraries in Europe, which is enhanced by substantial amount of data from university libraries and other types of institutions. The consortium will make available more than 5 076 000 of digitized pages by the end of November 2009.

The principle of integration is centralization of metadata within the Manuscriptorium digital library and distribution of data among other resources within the virtual net environment. The project creates conditions that enable the partners bringing together appropriate mass of digital content. Depending on level of readiness of a particular partner following general ways of cooperation are possible:

- For Advanced digitization projects operating digital libraries equipped with the OAI-PMH interface: partners provide Manuscriptorium with descriptive and structural metadata via one of their available OAI profiles. Metadata are harvested and processed within Manuscriptorium using individual input interface, so called Connector, which is prepared according to the metadata properties.
- For Advanced digitization without the OAI-PMH interface: available metadata are processed using the individually prepared Connectors for partners with larger number of documents (where Connector creation is reasonable). The difference against OAI-PMH enabled partners is in the method of transferring the metadata to the input of Manuscriptorium.
- For Starting or smaller scale digitization projects: it is possible to use Manuscriptorium dedicated tools to create and transfer the metadata content. These tools are:
 - M-Tool: an application which enables to create the descriptive and structural metadata

- M-Can: on-line application which enables upload of the metadata into the Manuscriptorium environment, check of correctness and subsequently transfer for import
- For Large-scale digitization projects without structural metadata: it is possible to create the metadata in an automated way (using some of the commonly available generation tools) and process the results as structural metadata within dedicated individual Connector.

Users' Requirements for Community Building Features

As the resulting digital library service should be of high quality and user-centered, a survey was carried out investigating preferences of the digital library users in respect to static virtual collections, dynamic virtual collections and individual virtual documents which were the major functions planned for implementation in the user's personalization area. Individual collections allow any reader to create and maintain personal set of documents within reader's profile.

Virtual documents are documents created manually with parts of other digital library documents. For example, a reader can build a virtual document demonstrating the art of illumination for a selected period (e.g. showing all illuminations from one scriptorium in a virtual document in spite of the fact that they are from various originals owned by different institutions in different countries).

The survey was filled in by over 450 respondents from 12 European countries participating in the ENRICH project. Full analysis of the survey performed in frame of the ENRICH project is available in form of the ENRICH project deliverable D 4.1 [7]. The survey investigated necessity of the personalization features and it appeared that all of them were positively evaluated and therefore needed by the users. Third part of questions queried about the features such as possibility to attach a file or add notes to particular individual item (collection or document) were also positively evaluated. Functionality for sharing individual collections and documents with other users was welcome by most of respondents. The survey displayed that most of respondents prefer to choose which individual collections and documents will be available to other users and which will be hidden from them, that majority of respondents prefers to have the full configuration possibilities giving them full control over the shared content and that users would also appreciate possibility of giving a copy of personal collection or document to other users. The purpose of copying of an element is to create a new instance of the element and to enable selected user so he/she can start to use it as his/her own. It appears that for respondents the functionality to give a copy of an individual collection or document is more important than the functionality which allows to simply sharing it.

Another example of end-users interface extension is the Multilingual service module which above all enables to perform translation of selected records and also a pilot simple search query translation. The features enable the end-users to directly improve the translation results by enhancing the embedded historical documents specific translation dictionary.

References

1. Knoll, Adolf – Mayer, Tomáš – Psohlavec, Stanislav – Vomlel, Jan: Digitization of Rare Library Materials. Storage of and Access to Data: The Solution for the Compound Document, Manuscripts and Old Printed Books [CD-ROM]. Praha: Národní knihovna České republiky, 1997.
2. Giesecke, Michael: Der Buchdruck in der frühen Neuzeit: Eine historische Fallstudie über die Durchsetzung neuer Informations- und Kommunikationstechnologien. Frankfurt am Main: Suhrkamp, 1998. 957 pp. ISBN 3-518-28957-8;
3. Uhlíř, Zdeněk: Teorie a metodologie elektronicko-digitálního zpracování rukopisů a hybridní knihovna. [The theory and methodology of electronic-digital processing of manuscripts and the hybrid library.] Praha: Národní knihovna České republiky, 2002. 324 pp. ISBN 80-7050-410-2.
4. Foster, I., Kesselman, C., Nick, J., Tuecke, S.: The Physiology of the Grid: an Open Grid Services Architecture for Distributed Systems Integration. Technical report, Global Grid Forum (2002)
5. Knoll, A., Digital Access to Old Manuscripts, *Linguistica Computazionale, Digital Technology and Philological Disciplines*, 277 – 286 (2004).
6. Mazurek, C., Stroiński, M., Węglarz, J., Werla, M., (2006). Metadata harvesting in regional digital libraries in PIONIER Network, *Campus-Wide Information Systems*, Vol. 23, No. 4, 241–253 (2006).
7. ENRICH project deliverable D4.1 – Definition of requirements for the creation of personalised virtual digital libraries, http://enrich.manuscriptorium.com/files/ENRICH_WP4_D_4_1_final.pdf
8. Roberto Caldelli, Cezary Mazurek, Paolo Mazzanti, Tomasz Parkoła, Marcin Werla: *Users requirements for personalised virtual digital libraries*, QQML2009 - Qualitative and Quantitative Methods in Libraries, International Conference, Chania Crete Greece, May 2009 http://www.isast.org/proceedingsQQML2009/PAPERS_PDF/Caldelli_et_al-Users_requirements_for_personalised_virtual_digital_libraries_PAPER-QQML2009.pdf