

Proposal for Fedora User Group presentation at Open Repositories, Atlanta, Georgia, May 2009.

Fedora and GSearch in a Research Project about Integrated Search

By Gert Schmeltz Pedersen, DTU Library¹, Technical University of Denmark.

The Royal School of Library and Information Science² in Denmark is performing a research project about integrated search, partly funded by DEFF³, Denmark's Electronic Research Library. DTU Library provides assistance in the form of a Fedora and GSearch installation.

The research project involves empirical investigations following the principles of polyrepresentation in information retrieval⁴:

“The polyrepresentation principle suggests that cognitively and functionally different representations of information objects may be used in information retrieval to enhance quality of results. In the paper, several empirical studies that intentionally or unintentionally have tested the principle are introduced and discussed. The continuum proposed by Larsen (2004; Ingwersen & Larsen, 2005) showing the structural dimension of the retrieval techniques involved in polyrepresentation is further elaborated by adding a novel second dimension consisting of query structure and modus. The new two-dimensional continuum can be seen as a constructive framework for further investigations of polyrepresentative principles in IR.”

The Fedora and GSearch installation provides a backend for the empirical investigations, containing an index of metadata records harvested from various sources, thus allowing polyrepresentation of literature.

The installation so far harvests from two types of sources, one using OAI-PMH⁵, and one using the SRU protocol⁶.

The presentation will focus on the technical challenges involved in the setup and indexing of the various sources, facilitating the integrated search.

¹ DTU Library, <http://www.dtic.dtu.dk>

² The Royal School of Librarianship, <http://db.dk>

³ Denmark's Electronic Research Library, <http://deff.dk>

⁴ B.Larsen, P.Ingwersen, J.Kekäläinen. The Polyrepresentation Continuum in IR, <http://portal.acm.org/citation.cfm?id=1164840>

⁵ Protocol for Metadata Harvesting, <http://www.openarchives.org/pmh/>

⁶ Search/Retrieval via Url, <http://www.loc.gov/standards/sru/>