„Suchkiste":
A discovery interface for DFG Nationallizenzen

Till Kinstler, kinstler@gbv.de
Gerald Steilen, steilen@gbv.de

Lightning Talk, ELAG 2009, Bratislava, 24.04.2009

DFG Nationallizenzen

- DFG (German Research Foundation) spends 80 million Euro for commercial digital content

- so far: about 100 products bought, > 80 million metadata records, > 500 million pages, e.g. Nature Archives 1865-2007, Springer Online Journal Archives 1860-2001, China Ancient Books 中國基本古籍庫 , English Language Women's Literature of the 18. & 19. Centuries, ...

- accessible for everyone with domicile in Germany

Suchkiste: general conditions

- publishers must deliver metadata records of licensed content to us (Verbundzentrale des GBV, http://www.gbv.de/), part of them delivers also fulltext content

- we convert metadata (which comes in a zoo of formats, most not standardized or even well defined) to Picaplus format, store it in our central metadata management system (OCLC CBS) and provide interfaces to it (Z39.50, SRU, unAPI, download, ...)

- the „Suchkiste project" aims at developing a discovery interface centered on users' needs for this content (motivation: enhance access and usage of DFG Nationallizenzen)

Suchkiste: how?

basic infrastructure:

- take data from CBS, currently in (customized) USMARC format

- index it with Solr using solrmarc

- currently: discover it through customized VuFind user interface

- because we must control access to the metadata of some products, but also for personalised functions in the future: authentication and authorization using Shibboleth

this basic infrastructure is ready and running in a prototype installation with some 20 million metadata records and about 370000 fulltexts at the moment

Suchkiste: what's next?

We have strong interest in some questions around discovery interfaces for libraries, so we will focus on the following points in the sequel of the project:

- so called relevance ranking: in a phd project methods for ranking discovery results will be evaluated doing (inter alia) user tests. Maybe we'll have kind of „pertinence ranking" or „usefulness ranking" in the end?

- redesign of user interface, idea: away with the „search box" -> „short title list" -> „full title list" paradigm (which comes from technical restrictions in the past). We imagine a workbench where users can operate with different tools on the somhow visualized data space. Tools are things like a (text) search box, facets, „simliar items", recommendations, personal preferences, social network data, cut&paste, sorting, [your ideas here]...)

- important (as seen at elag09): mashupability, linkability of data

Questions? Ideas? Contact us:

http://blog.nationallizenz.de/ (project blog, in German)
http://www.nationallizenzen.de/ (about DFG Nationallizenzen, Registration ...)
Till Kinstler, kinstler@gbv.de, Skype: tillgbv
Gerald Steilen, steilen@gbv.de, Skype: gs-gbv.de
On Twitter: http://twitter.com/suchkiste (Suchkiste twitters what it does),
http://twitter.com/tillk, http://twitter.com/nesnu (Hashtag #suchkiste)