DwB Discovery Portal A New CESSDA Portal for European Research Data Discovery

John Shepherdson - UKDA

Pascal Heus - Metadata Technology

Ørnulf Risnes - NSD

Overview

- Supports DwB goal "equal and easy access to official microdata for the European Research Area"
 - provides "more coherent system for resource discovery of official statistics"
 - demonstrates ability to ingest metadata from multiple sources, via multiple protocols

Scope of the portal

- Content-wise
 - Metadata from NSIs + Archives
- Technical
 - ▶ Build prototype/beta
 - Sound, future proof methods, architecture, components
 - ▶ Standards-based
 - ▶ Extensible
 - Easy to handover to 'sustainability' body

Aspects of the portal

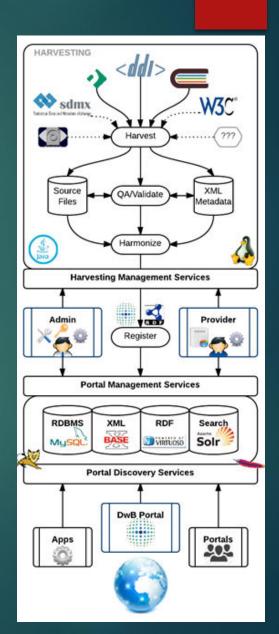
Research data discovery (obviously)

Provider portal, QA

Platform for additional services

Metadata ingest

- ► Metadata gets...
 - harvested
 - ▶ made ready for QA
 - transformed into canonical model
 - ▶ indexed
 - exposed



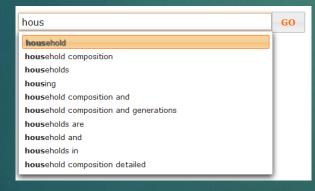
Canonical metadata model

- ▶ Harmonisation
 - ▶ DDI-C, DDI-L, MISSY, CIMES, etc

- ▶ Builds on DISCO
 - ▶ (DDI discovery RDF)

Research data discovery

- Powered by Solr
- ▶ Facets
 - producer, geography, date, data type ...
- Suggestions / autocomplete



'Did you mean?' functionality





Provider portal, QA

Quality reports on own metadata

Tools to improve quality

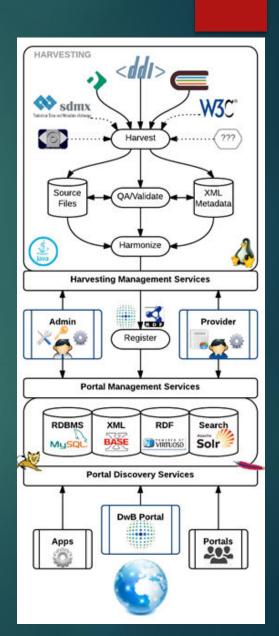
Profile and visibility management

Tools to improve quality

Version control of metadata (optional)

Platform for services

- DwB search portal is just a front-end application
- Machine-actionable interfaces for most functions



What Next?

- ► Integrate ELSST multi-lingual thesaurus
 - expansion of queries
 - ▶ 14 languages, synonyms/ broader/ narrower terms
- Pre-harvest preparation tools
 - ▶ based on e.g. <u>DataForge:SledgeHammer</u>

ELSST

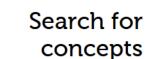
UK Data Service



ELSST search

ELSST suggestions

ELSST guide



ELSST covers the core social science disciplines

Enter your search term and select a language to browse ELSST



▶ LOGIN / ♣ REGISTER C THE SAURUS SEARCH Search and browse the ELSST Thesaurus. • English Q GO Explore the thesaurus

STOP

ELSST BLOG



Thesaurus alignment task



Mapping and matching accommodating divergence and protecting integrity



ISKO UK Conference 2013

ABOUT US

ELSST is a broad-based, multilingual thesaurus for the social sciences.

It facilitates access to data resources across Europe, independent of domain, resource, language or vocabulary.

It was originally based on the monolingual thesaurus, Humanities and Social Science Electronic Thesaurus (HASSET), of the UK Data Archive at the University of Essex.

ELSST covers the core social science disciplines: politics, sociology, economics, education, law, crime, demography, health, employment, information and communication technology and, increasingly, environmental science.

STRUCTURE

ELSST employs the usual range of terms and thesaural relationships:

- · Preferred Terms and Non-Preferred or Use For terms (UFs)
- Broader Terms (BTs)
- · Narrower Terms (NTs)
- Related Terms (RTs)

ELSST's UFs are entry level terms rather than pure synonyms.

Scope Notes (SNs) are used to define the extent of the intended meaning within the domain of the thesaurus.

Use Notes (UNs) provide information and guidance on how terms may be used for indexing.

Maintenance and support

The thesaurus is updated annually. Users are encouraged to suggest new candidate concepts to the thesaurus (in English). The suggested concepts are reviewed annually by the multinational ELSST Thesaurus Management Team, and those that are approved are added to the next version of the thesaurus.

Questions?

