

How would you like to have your DDI today?

IASSIST 2013-05-31

Olof Olsson, Swedish National Data Service, olof.olsson@snd.gu.se

Jannik Jensen, Danish Data Archive, jvj@dda.dk

Johan Fihn, Swedish National Data Service, johan.fihn@snd.gu.se

Stefan Jakobsson, Swedish National Data Service, stefan.jakobsson@snd.gu.se

Akira Olsbänning, Swedish National Data Service, akira.olsbanning@snd.gu.se



STATENS ARKIVER

DANSK DATA ARKIV



DDI Lifecycle as a base format

DDI lifecycle has good support for describing

- coverage
- citation
- global and local identifiers
- documentation of data (well... obviously)

And almost all of it is multilingual

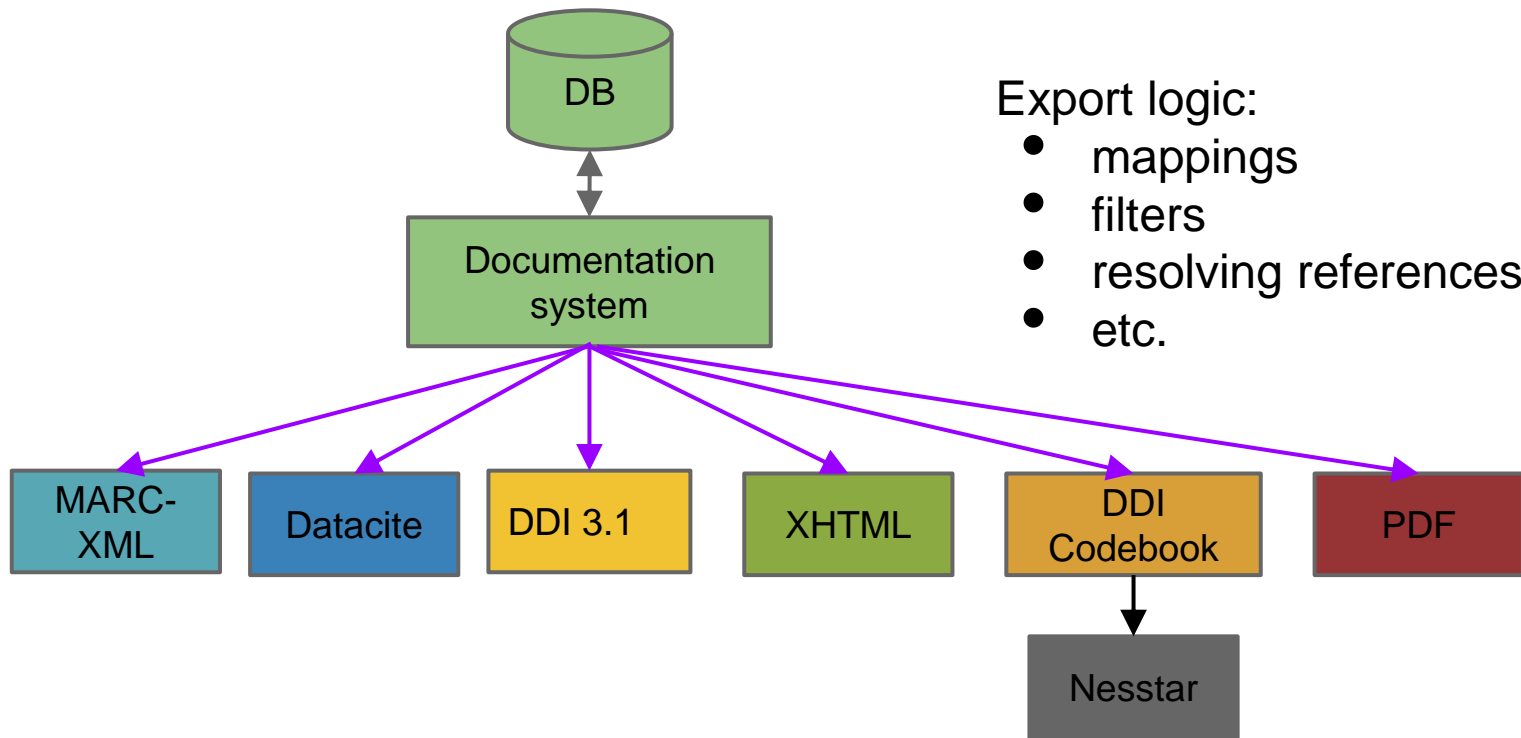
Why would anyone need information about our datasets in another format?

Does everyone need to know DDI?

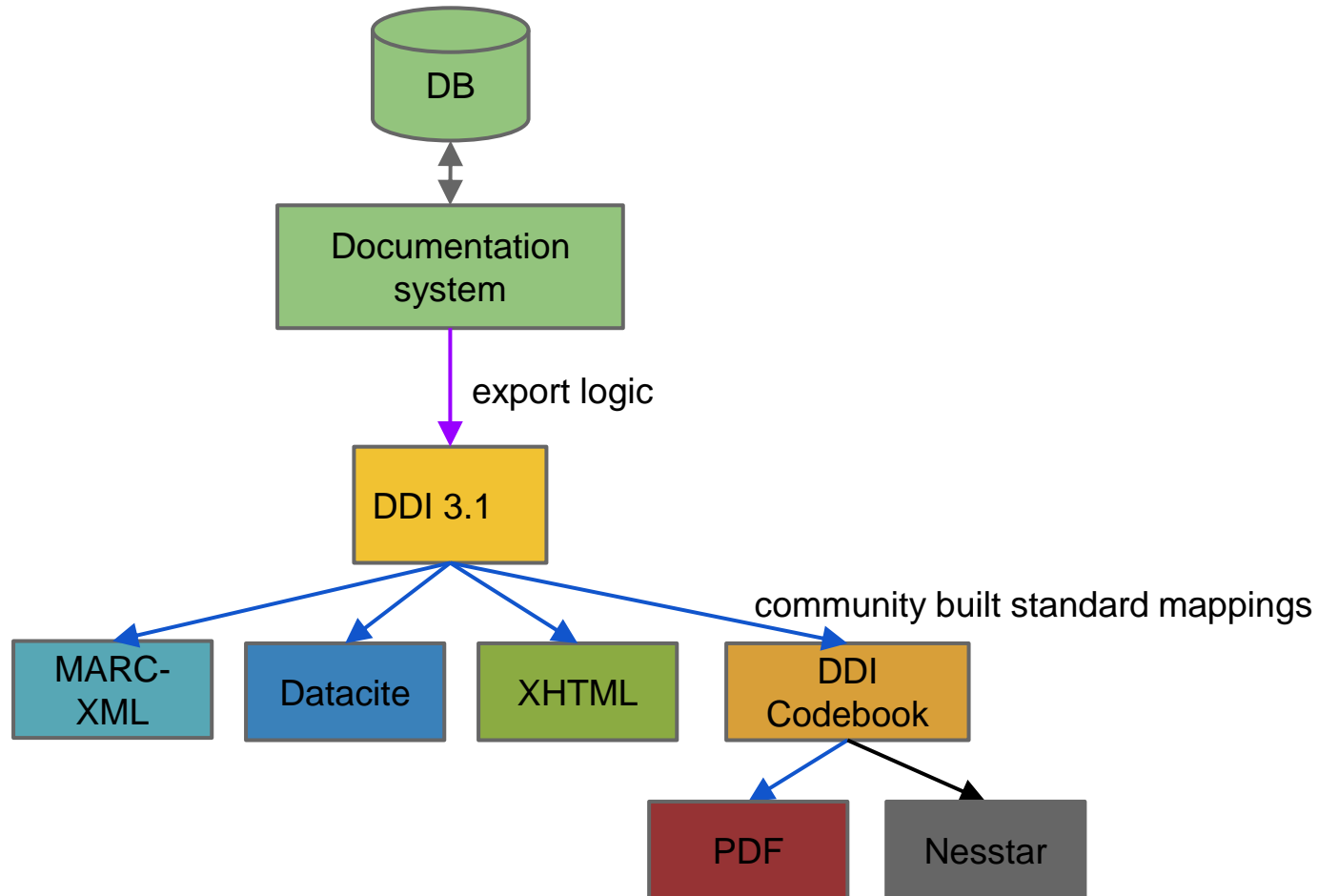
- DDI is complex
- "I just want a citation"
- MARC is in a lot of systems
- Datasets is just another resource

- Solution
 - Create a complete fine grained DDI-L documents with data documentation
 - Transform to other formats via XSLT on request

Building exports from a documentation system



Export to one format, transform to multiple formats



DDI Lifecycle to MARCXML

Used for the description of bibliographic items
Developed by the Library of Congress

MARCXML is a schema based on MARC21

DDI Lifecycle to MARC-XML

Supported fields

leader

001 Control Number

005 Date and Time of Latest Transaction

007 Physical Description Fixed Field

008 Fixed-Length Data Elements

010 Library of Congress Control Number

020 International Standard Book Number

022 International Standard Serial Number

024 Other Standard Identifier

034 Coded Cartographic Mathematical Data

041 Language Code

045 Time Period of Content

245 Title Statement

246 Varying Form of Title

260 Publication, Distribution, etc. (Imprint)

490 Series Statement

506 Restrictions on Access Note

520 Summary, etc.

522 Geographic Coverage Note

542 Information Relating to Copyright Status

650 Subject Added Entry - Topical Term

720 Added Entry - Uncontrolled Name

850 Holding Institution

856 Electronic Location and Access

DDI Lifecycle to DataCite metadata

DataCite is a not-for-profit organization

22 members/associate members

DataCite aims to:

- establish easier access to research data on the Internet
- increase acceptance of research data as legitimate, citable contributions to the scholarly record
- support data archiving that will permit results to be verified and re-purposed for future study

Transform supports 16 out of 17 fields

DDI Lifecycle to XHTML

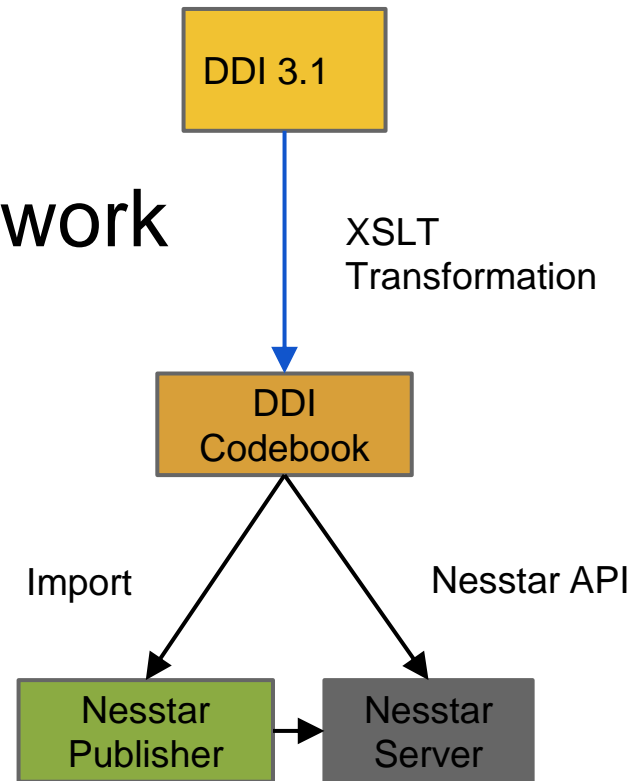
- Interactive codebook with filters, links and graphics
 - JavaScript/ JQuery
- Generic version
 - Overview of the study
 - Citations
 - Universes, concepts, questions and variables
- Specific DDA version
 - Descriptive statistics
 - Pie charts
 - Survey design -flow statements
- Parameters to control the output and layout

DDI Lifecycle to DDI Codebook

- Supports citation fields
- Merges variables and questions into <var>
- Target one or multilingual DDI Codebook

Import into nesstar

- Important for CESSDA archives
- Enables DDI-L tools to work with Nesstar



DDI Codebook to PDF

- Transforms Codebook -> FO -> PDF
- Original version written by Pascal Heus in 2006, and updated for FOP 0.93 by Oistein Kristiansen in 2010
- Now split into (somewhat) more manageable cleaned up parts. Still quite a massive chunk of XSLT...
- Documentation added in files about the XSLT templates and their resulting FO

DDI Codebook to PDF

- Render on server through web interface (eXist), or stand alone via a shell script (FOP)
- Works as intended, but still needs more pruning of left-over variables, and better looking PDF layout.

DDI 3.1 converter via XSLT

File Parameters

DDI 3.1 XML-file

Output format

xQuery transformation
using eXist xml-db



<http://xml.snd.gu.se/app/transform/index.html>

The future of the project

- A Lot of improvements to be done
 - additional fields to transform
 - control more behavior via params
- Hackathon for DDI-RDF transformation?
<https://github.com/linked-statistics/DDI-RDF-tools>
- New target formats?
- Need of users, testers and contributors in the community. join us!

Questions & Answers



<https://code.google.com/p/ddixslt>

Olof Olsson, Swedish National Data Service, olof.olsson@snd.gu.se

Jannik Jensen, Danish Data Archive, jvj@dda.dk

Johan Fihn, Swedish National Data Service, johan.fihn@snd.gu.se

Stefan Jakobsson, Swedish National Data Service, stefan.jakobsson@snd.gu.se

Akira Olsbänning, Swedish National Data Service, akira.olsbanning@snd.gu.se



STATENS ARKIVER

DANSK DATA ARKIV

SND

Swedish National Data Service

