

# Updates on the XSLT stylesheets for DDI

IASSIST, Toronto 2014-06-05

Olof Olsson, Swedish National Data Service, [olof.olsson@snd.gu.se](mailto:olof.olsson@snd.gu.se)

Jannik Jensen, Danish Data Archive, [jvj@dda.dk](mailto:jvj@dda.dk)

Johan Fihn, Swedish National Data Service, [johan.fihn@snd.gu.se](mailto:johan.fihn@snd.gu.se)

Akira Olsbänning, Swedish National Data Service, [akira.olsbanning@snd.gu.se](mailto:akira.olsbanning@snd.gu.se)



STATENS ARKIVER

DANSK DATA ARKIV



SND

Swedish National  
Data Service

# DDI Lifecycle as a base format

DDI lifecycle has good support for describing

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

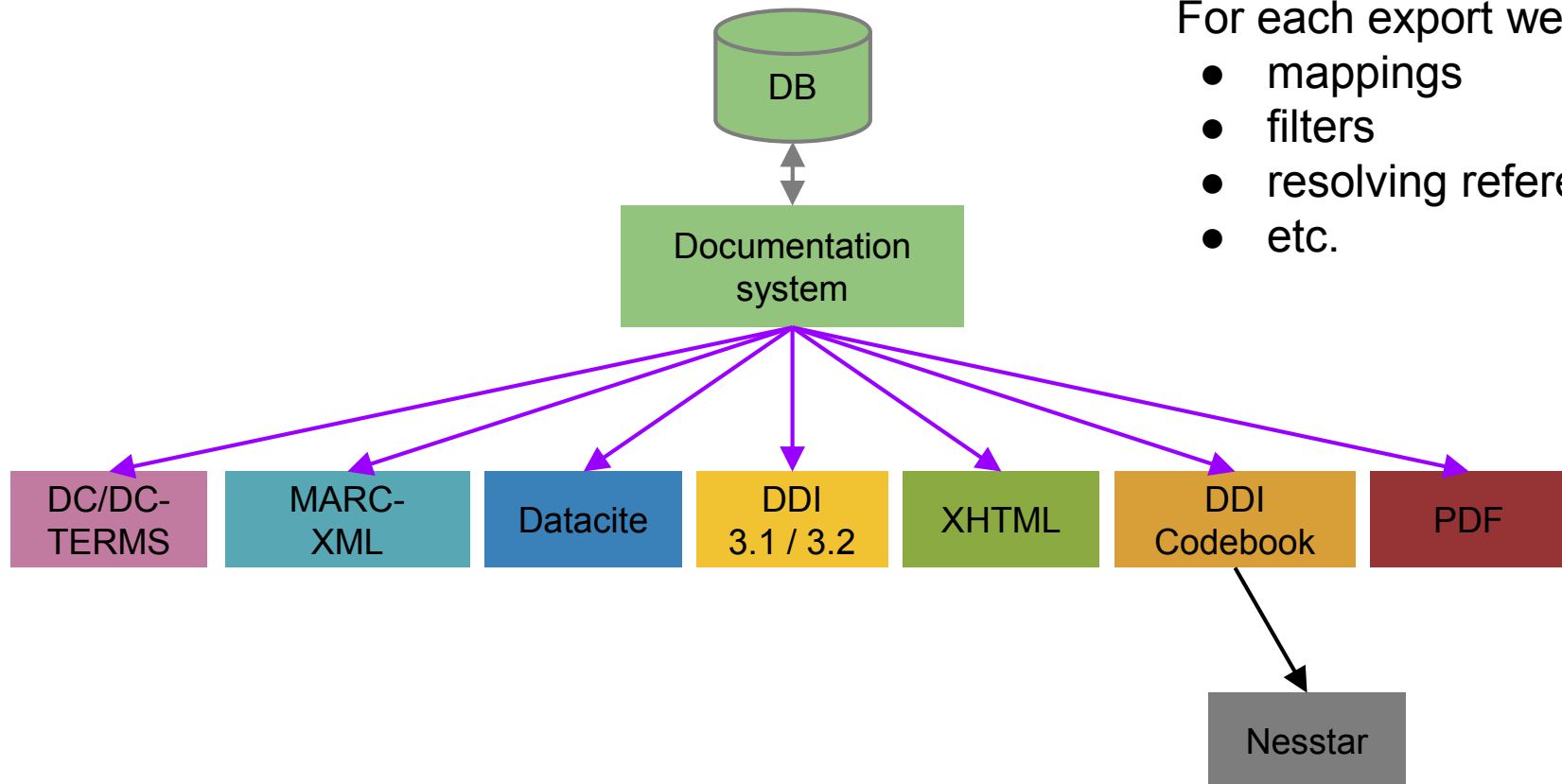
And almost all of it is multilingual

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"
- Formats like MARC is in a lot of systems
- Description of a Study is just another resource
  
- Solution
  - Create a complete fine grained DDI-L documents with data documentation
  - Transform to other formats via XSLT on request

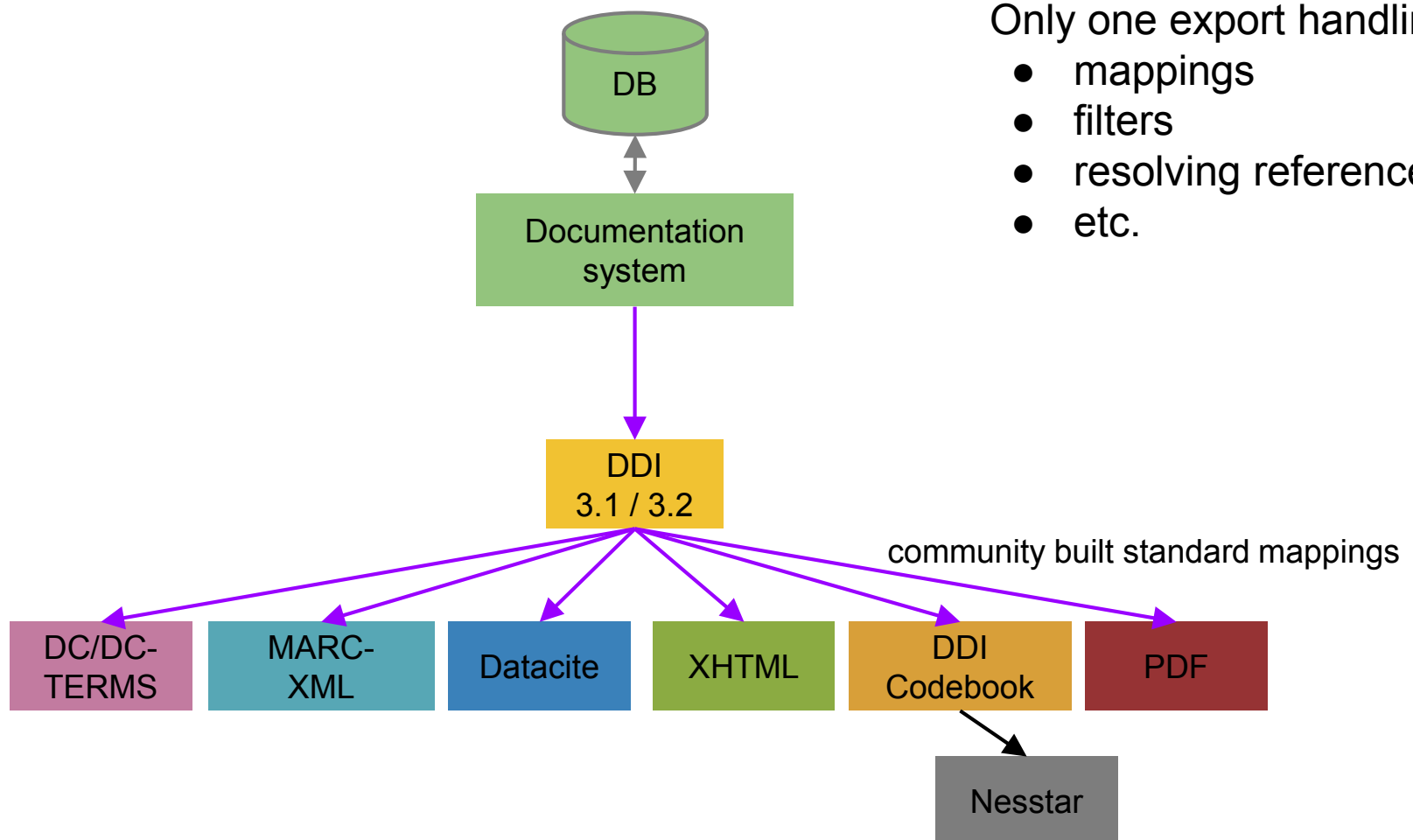
# Building exports from a documentation system



For each export we need:

- mappings
- filters
- resolving references
- etc.

# Focus the export to one rich format, transform to all other 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

Changes:

Updated for DDI 3.2

# 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
- Updated for DataCite version 3.0



# DDI Lifecycle to XHTML

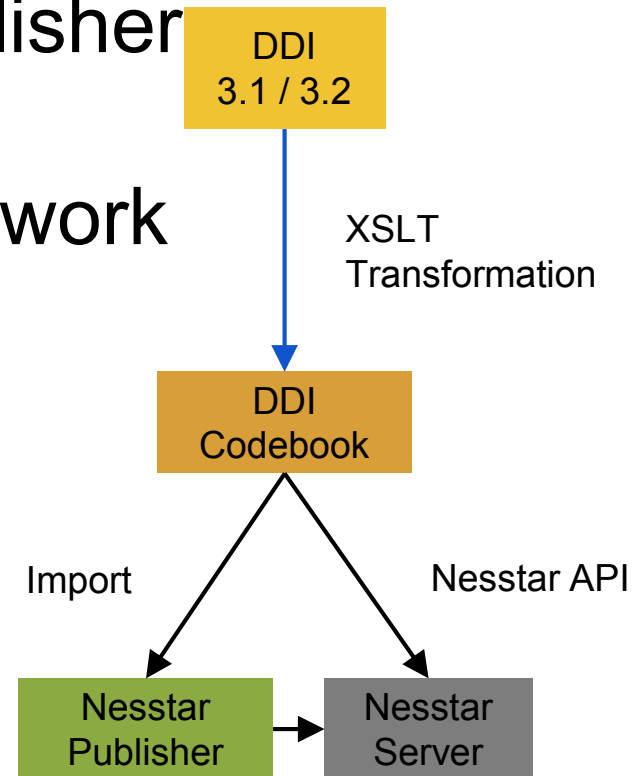
- Interactive codebook with skip statements, links and graphics
  - JavaScript/ JQuery
- Generic version
  - Overview of the study
    - Citations
    - Universes, concepts, questions and variables
- Specific DDA version
  - Descriptive statistics
    - Pie charts
  - Survey design -flow statements
-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

# DDI Lifecycle to DDI Codebook

- Supports study citation fields
- Merges variables and questions into <var>
- Target one or a multilingual DDI Codebook instance
- Added support for documentation information fields i DDI-L in the transformation to DDI-C

# Import into nesstar

- Import via Nesstar Publisher or Nesstar API
- Enables DDI-L tools to work with Nesstar



# DDI Codebook to PDF

- Original version written by Pascal Heus in 2006, and updated for FOP 0.93 by Oistein Kristiansen in 2010
- Transforms DDI Codebook -> FO -> PDF
- Now split into more manageable cleaned up parts
- 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)
- Updated with parameter documentation (dynamic configuration)



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
  - added more control more behavior via params
- Hackathon 2.0 for DDI-RDF transformation?  
<https://github.com/linked-statistics/DDI-RDF-tools>
- New target formats?
- Seeking other interested organizations
- Need of users, testers and contributors in the community, join us!
- Move to GitHub

# Questions & Answers



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

(will move soon)

Olof Olsson, Swedish National Data Service, [olof.olsson@snd.gu.se](mailto:olof.olsson@snd.gu.se)

Jannik Jensen, Danish Data Archive, [jvj@dda.dk](mailto:jvj@dda.dk)

Johan Fihn, Swedish National Data Service, [johan.fihn@snd.gu.se](mailto:johan.fihn@snd.gu.se)

Akira Olsbänning, Swedish National Data Service, [akira.olsbanning@snd.gu.se](mailto:akira.olsbanning@snd.gu.se)



STATENS ARKIVER

DANSK DATA ARKIV



SND

Swedish National  
Data Service