Use of DDI-Lifecycle at the GESIS Data Archive

Catharina Wasner
GESIS – Leibniz Institute for the Social Sciences

EDDI 2015, December 2, Copenhagen

"DDI-LimDAS"

- means: "Use of DDI-Lifecycle at the GESIS Data Archive"

- is: an internal project at the GESIS Data Archive

- consists of: Catharina Wasner, Wolfgang Zenk-Möltgen

- is supported by: Alexander Mühlbauer, Martin Friedrichs

- lasts: two years (01/2015-12/2016)

The GESIS Data Archive (DAS)

What we do:

Providing quantitative data of Social Sciences for reuse and replication. That includes

- data acquisition,
- data processing,
- data documentation,
- longterm archiving,
- consultancy,
- software design and development.

Goals

Change products, projects and processes at DAS from DDI-C as documentation and archiving format to DDI-L:

- develop a metadata schema for all use cases in the data archive in the DDI-L format

- develop recommendations for the modification of metadata editing workflows in DAS to meet the demands of DDI-L

- start coordinating the metadata management of existing and future projects of DAS and whole GESIS

Basic Principles

- consider preliminary work done so far in different DAS/GESIS projects
- consider prospective DAS/GESIS developments and future plans
- focus on reusability in GESIS
- get in touch with DDI-L using partners from CESSDA and DDI-Alliance to benefit from existing experiences
- include best practices from external partners to facilitate interoperability, if possible

Approach

1. Which projects/tools are using DDI at DAS or should use it?
2. Who is using which DDI version?
3. Which projects/tools should participate in defining the metadata schema?
4. Development and pretest of a survey instrument:
   - For which purposes is DDI used in these selected projects and what are their requirements regarding the metadata?
   - Where can documentation and preliminary work of these projects be found?

Approach cont.

5. Synopsis of the survey results
6. Analysis of the results in consideration of the following questions:
    – Are there overlaps between the requirements of the projects?
    – How can the requirements „translated" into DDI-L?
    – Which version of DDI-L should the schema use in the long run?
7. Search for and evaluation of DDI-L best practices outside DAS

Which projects/tools are using DDI at DAS or should use it?

Which projects/tools should participate in defining the metadata schema?

- pragmatic decisions: which projects/tools have done most preliminary work on DDI so far?
- political decisions: which projects/tools in DAS do we like to support with DDI-L first?

CharmStats Codebook Explorer
STARDAT Reporting Tool Data Catalog
STARDAT GLES Question Search
Dataset Documentation Manager

Survey Instrument

- What is the piece of information to be documented in the project/tool?
- Which metadata is produced by means of the project/tool?
- What is the semantic definition of these metadata (fields)?
- Is DDI used and if yes which version?
- Where can I find documentation on these questions?

Analysis of the Results

- Pieces of information: studies, variables, questions, answers, interviewer instructions, concepts, tables for frequency/descriptives
- Changes/differences of questions regarding time/space
- Groups of variables relating to concepts, time, space
- The first draft is planned for January/February 2016

Further Plans

To have in the end:

- a good metadata schema
- a good DDI 3.2 mapping
- proposals of metadata schemas and mappings for all selected projects
- in the best case DDI profiles
- recommendations for the modification of metadata editing workflows
- a starting point for GESIS internal DDI management: this would include the whole Lifecycle (study planning and data collection as well)
- good documentation of all results

The End
(hopefully ☺)