gesis Leibniz Institute for the Social Sciences

Lessons Learned with Additional Mappings into DDI-FlatDB

Claus-Peter Klas, Oliver Hopt, Alexander Mühlbauer

EDDI 2017 – License CC-BY 4.0

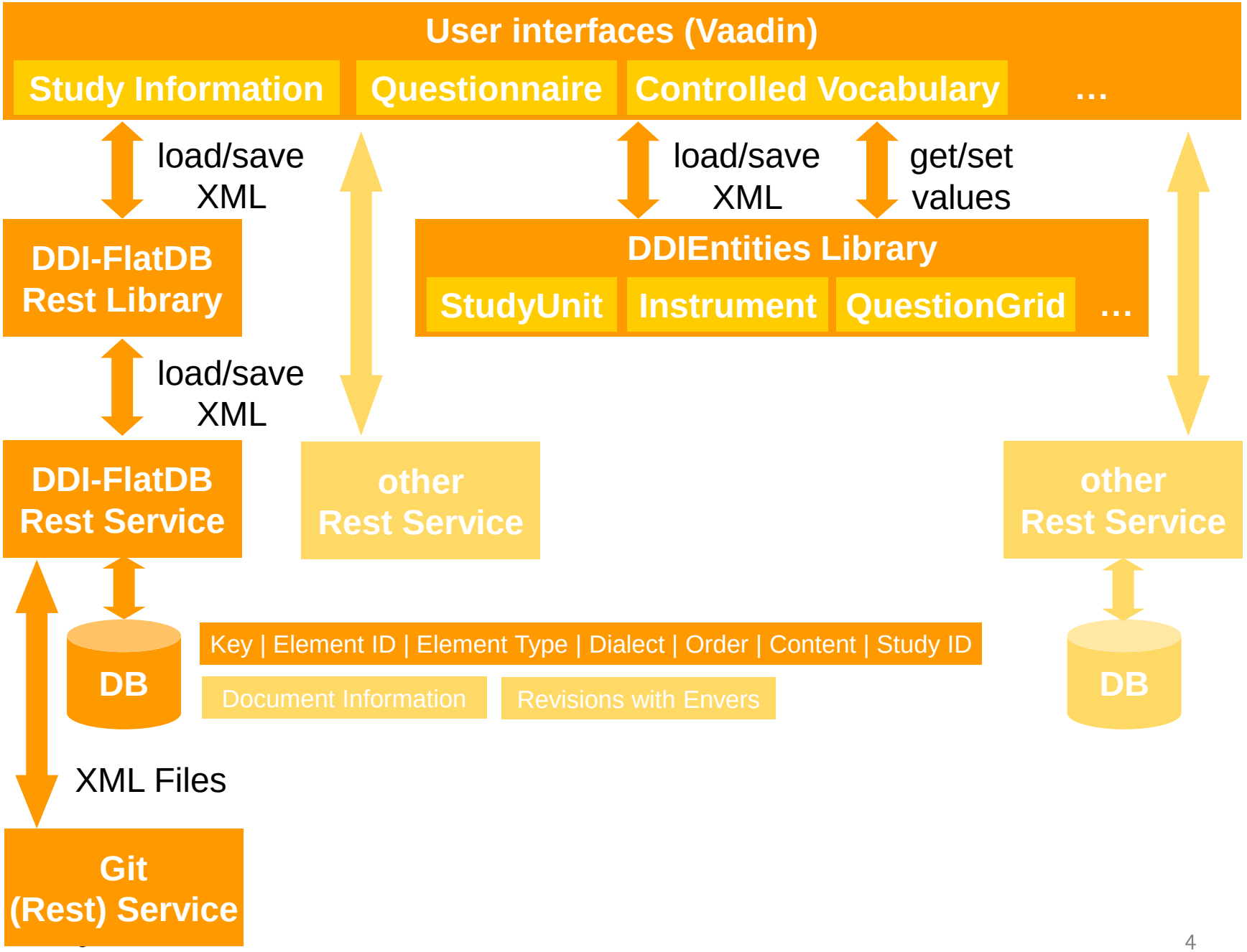
Member of Leibniz Association

Agenda

- Short technical overview of DDI-FlatDB
- Use cases covered
  - Content
  - Development performace
- Discussion

General idea of DDI-FlatDB

- XML document as source of information
- Cut into pieces for „client“ performance
- Pieces cached in a database
- XML processing is configurable
  - How to cut/recombine
  - How to read/write data
  - Supporting several versions/dialects


User interfaces (Vaadin)
Study Information
Questionnaire
Controlled Vocabulary
…
load/save XML
load/save XML
get/set values
DDI-FlatDB Rest Library
DDIEntities Library
StudyUnit
Instrument
QuestionGrid
…
load/save XML
DDI-FlatDB Rest Service
other Rest Service
other Rest Service
DB
Key | Element ID | Element Type | Dialect | Order | Content | Study ID
Document Information
Revisions with Envers
DB
XML Files
Git (Rest) Service

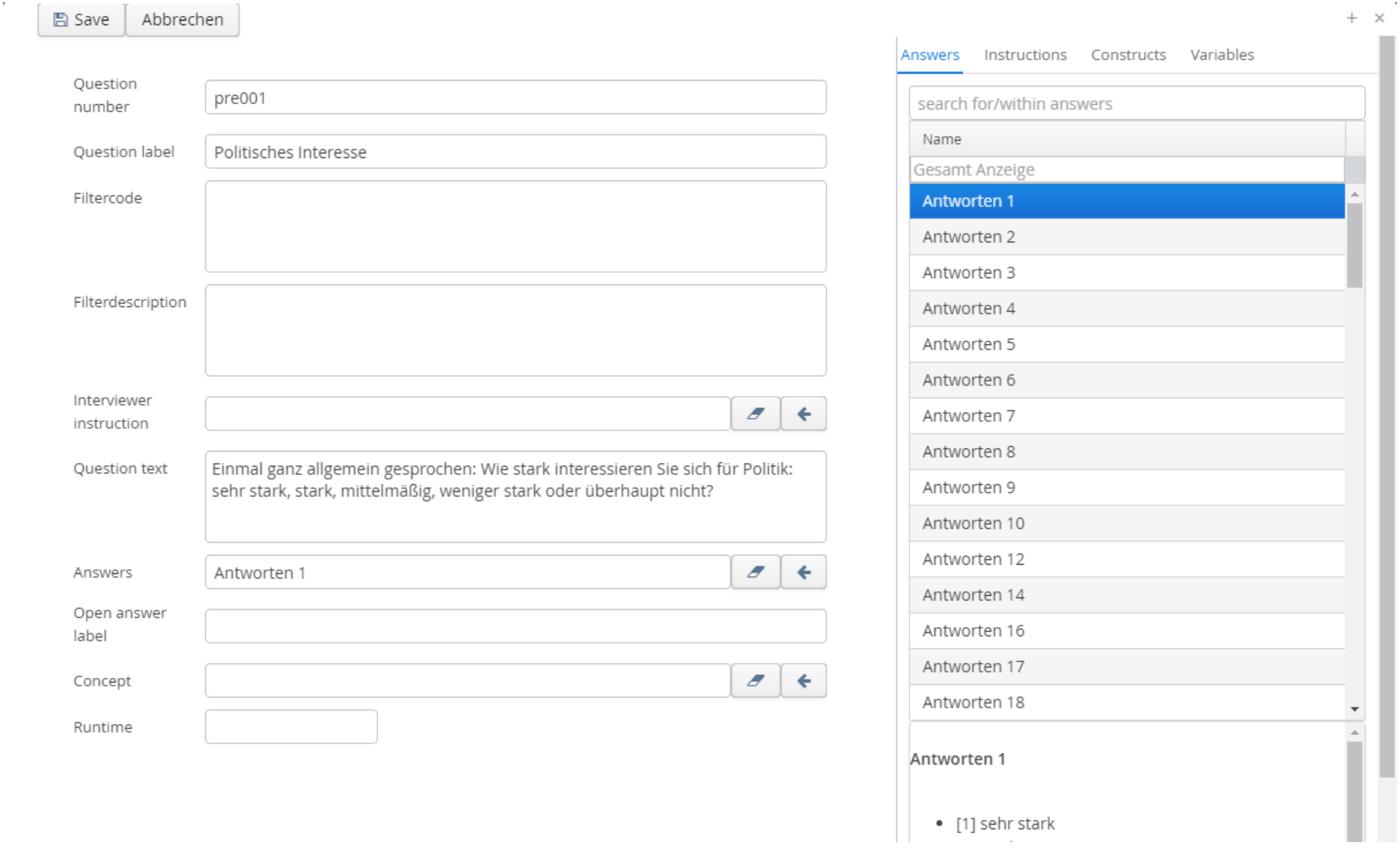
Use cases

- GLES questionnaire Editor (QMD)
- Study level editor (DBK case)
- Variable [documentation and] discovery (Explore Data project)
- CESSDA Controlled Vocabulary manager

- (DDILimDAS/CESSDA CMM)

gesis Leibniz Institute for the Social Sciences

QMD



Model classes involved in QMD

- StudyUnit
- Instrument
- Sequence
- QuestionConstruct
- QuestionItem
- QuestionGrid
- CodeList
- CategoryScheme
- TextDomain
- Universe
- Instruction
- Concept
- Variable

Effort for QMD case

- High – about a year
  - XML-consuming
  - Bean mapping
  - UI mapping
  - „Automatic“ UI creation

Model classes involved in Study Information

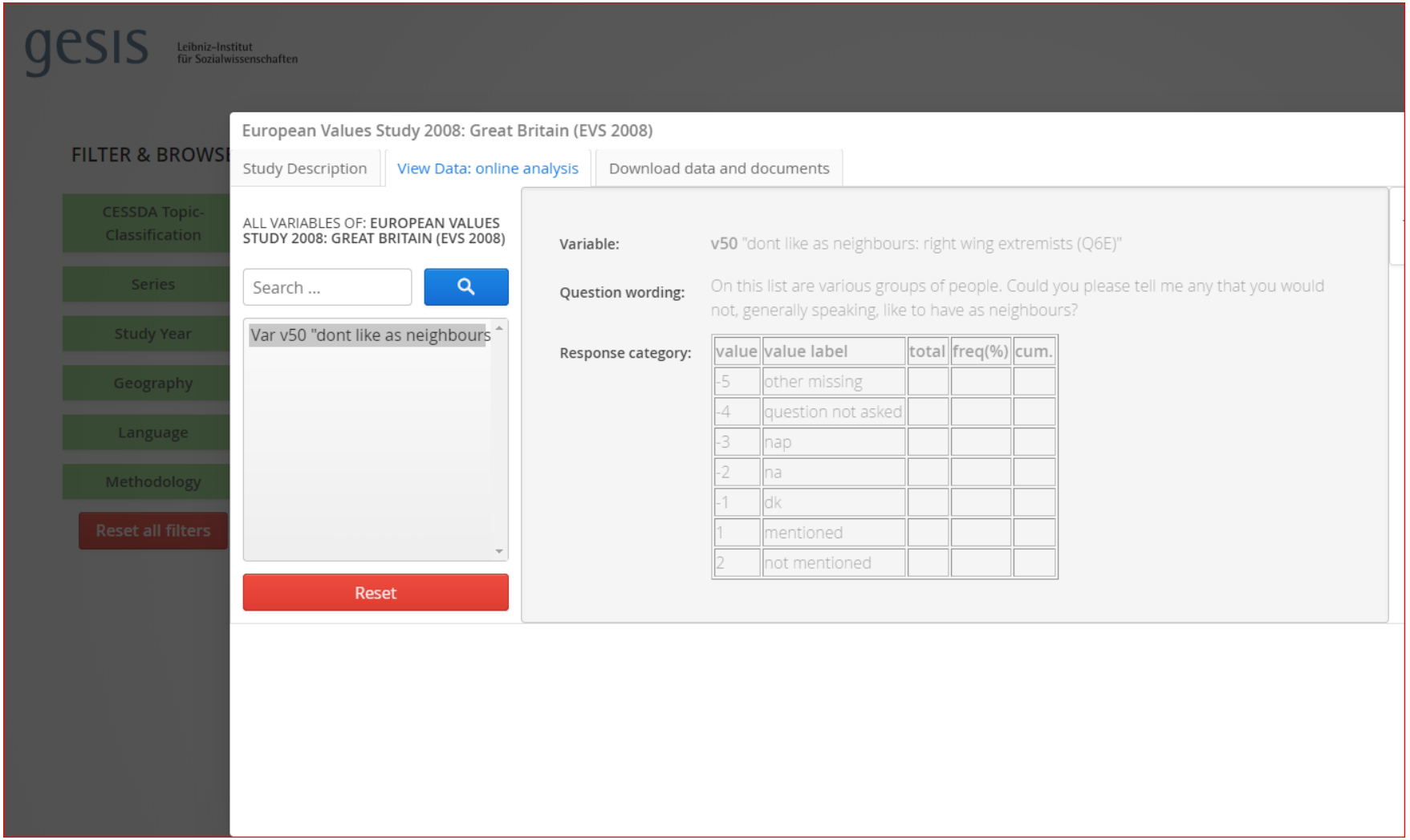
- StudyUnit
- Concept
- OtherMaterial
  - Creator
  - Contributor
  - Note

Effort for Study Information case

- High for concept – one month
  - Modeling the use case
  - Functional specialties in UI
- Low for the data management (CRUD) – one month
  - Uncovered cases within XML consuming

gesis Leibniz Institute for the Social Sciences

Variable discovery



Model classes involved in Variable Information

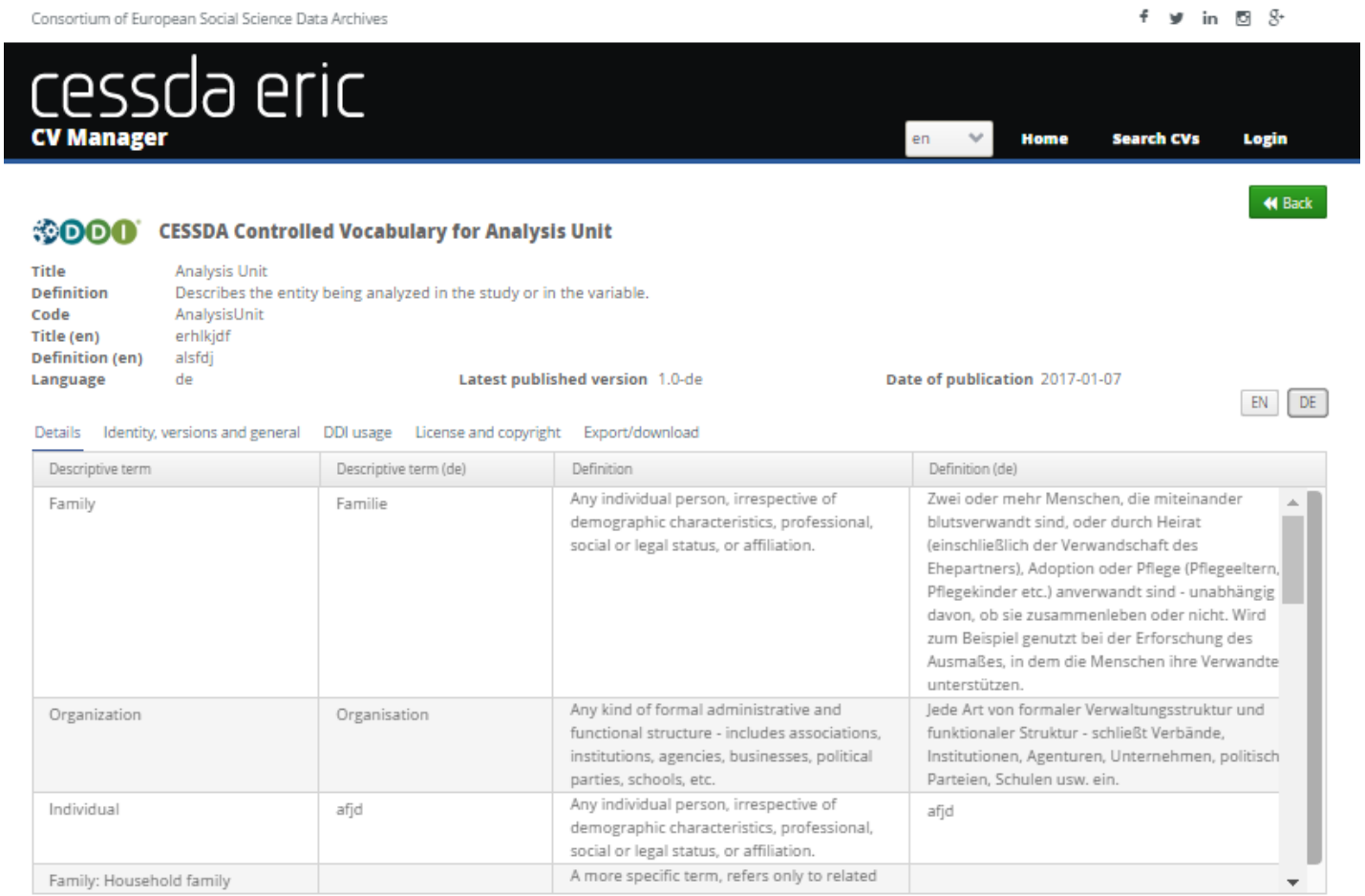
- StudyUnit
- QuestionConstruct
- QuestionItem
- QuestionGrid
- CodeList
- CategoryScheme
- TextDomain
- Universe
- Instruction
- Concept
- Variable
- (Statistic)

Effort for Variable Information case

- Medium for concept – some months
  - Simple modeling the use case
  - Functional specialties in UI
  - How to structure the search index
- Low for the data management (CRUD) – some days
  - Mostly covered by Questionnaire case and Study Information

gesis Leibniz Institute for the Social Sciences

CESSDA CV Manager



Model classes involved in Controlled Vocabulary

- CVScheme
- CVConcept

(both based on SKOS)

Effort for CV case (concept was ready)

- Low for model – few hours
  - Very simple and preexisting data model
  - Straight forward concept for functionality/UI
  - Becoming familiar with SKOS
- Very low for the data management (CRUD) – few hours
  - Two new beans
  - One dialect mapping

Effort for DDILimDAS/CMM case

- Study level, questionnaire, simple variable
- Very low for concept – one hour
  - Two conceptual changes
    - Externalize Persons and Institutions: three new beans, one bean extended
    - Instrument „inside“ StudyUnit: functional changes in QMD and Variable Discovery
- Very low for the data management (CRUD) – half an hour
  - Defining paths for another dialect

Discussion

- DDI-FlatDB is able to provide
  - Flexible storage across versions/dialects
  - Hides DDI complexity from end users
  - (Conversion between versions/dialects)
- Developers can focus on the service
  - Good UI integration
  - Efficient (fast) access
  - Application driven development
  - Permeable between applications
- Drawbacks
  - Lets discuss offline ;-)

DDI-FlatDB and DDI-Entities are available…

- … as Open Source:
  - https://git.gesis.org/stardat/stardat-ddiflatdb
  - https://git.gesis.org/stardat/stardat-ddi-entities
- … as docker:
  - http://play-with-docker.com/?stack=https://git.gesis.org/stardat/stardat-ddiflatdb/raw/master/doc/play-with-docker/docker-compose-v3.3.1.yml

Contact

- Oliver Hopt (Oliver.Hopt@gesis.org)
- Claus-Peter Klas (Claus-Peter.Klas@gesis.org)
- Alexander Mühlbauer (Alexander.Muehlbauer@gesis.org)

GESIS – Leibniz-Institute for the Social Sciences

License

This presentation is offered under license CC-BY 4.0

The license does not apply to the following copyrighted material used in this presentation:

- Photograph of the GESIS Main Office
- GESIS logotype