Flexible DDI Storage
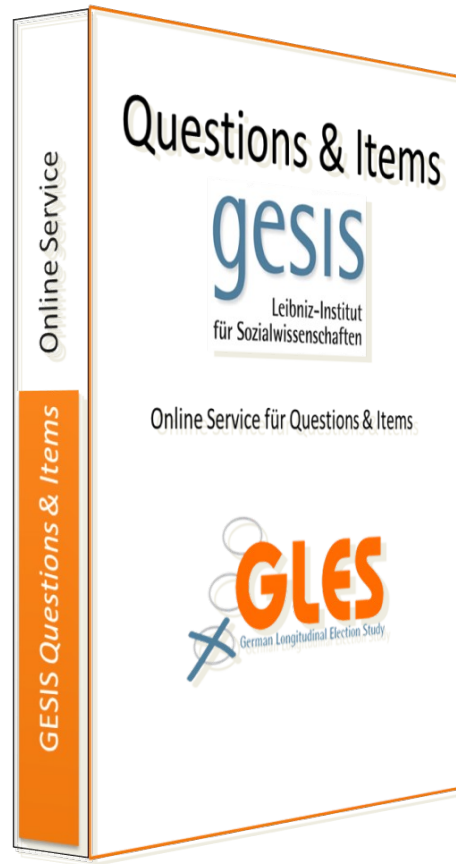
Oliver Hopt, Claus-Peter Klas,
Alexander Mühlbauer, Wolfgang Zenk-Möltgen
GESIS – Leibniz Institute for the Social Sciences
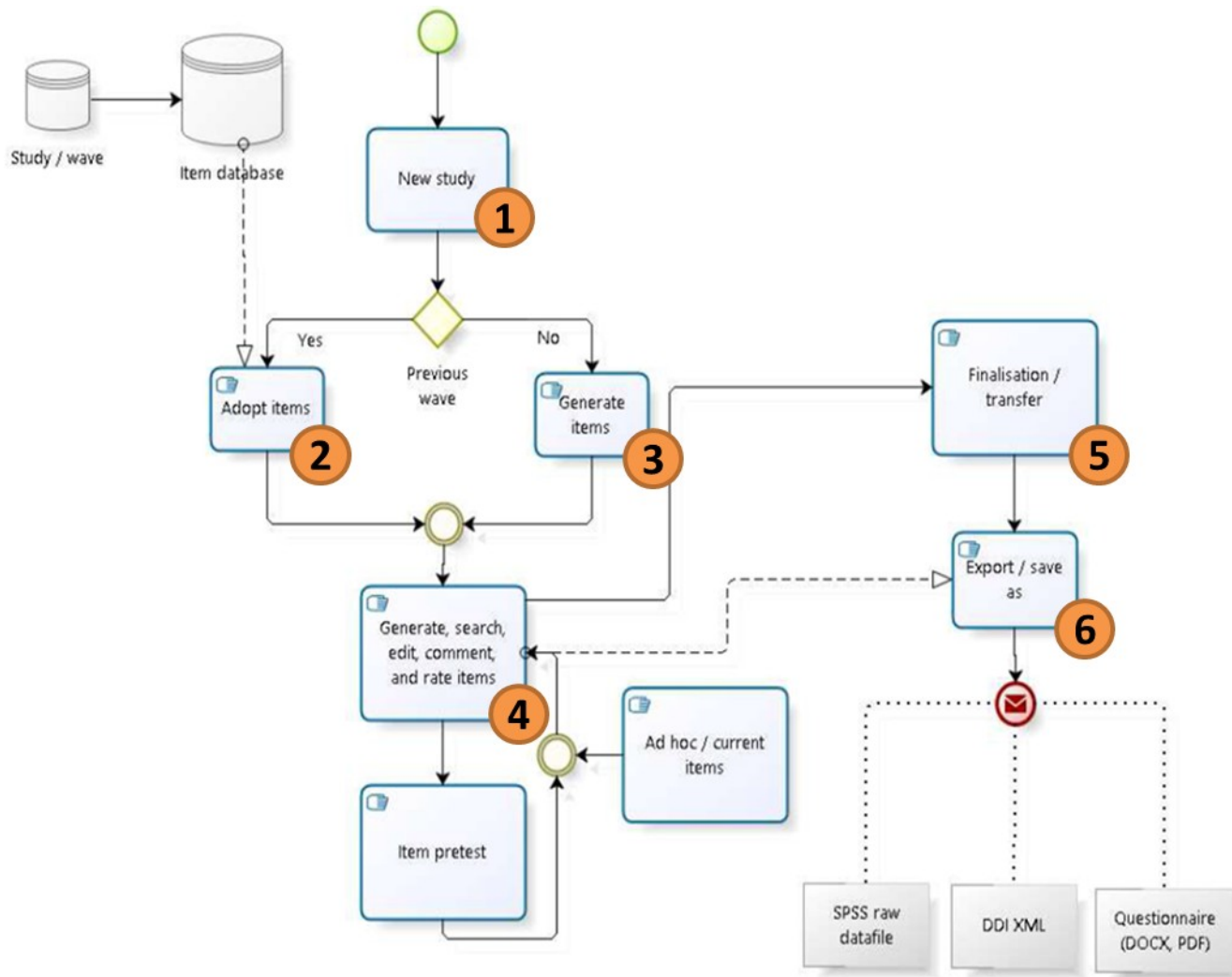
GESIS *DDI-FlatDB*

Outline

- Use Case: Questionnaire Editor
- Problem description
- Concept
- Using DDI-FlatDB
- Portal tour
- Steps taken and to be taken

Use case – Questionnaire editor

Use case process

DDI 3.2 covered

- Group

- StudyUnit

- ConceptScheme
- Concept

- Instrument
- Sequence
- QuestionConstruct
- Universe
- InterviewerInstruction
- QuestionItem
- QuestionGrid
- CodeList
- CategoryScheme
- Variable

Problem description

- ## DDI-based
  - Versions and how they are understood
  - Complexity

- ## XML-based
  - Performance of processing
  - Ordering according to schema

- ## Use cases are a moving target
  - Issues with one storage for different applications

GESIS *DDI-FlatDB*

Concept

- Make the XML document the central source of truth
  - Managable revisions
  - Reuse of transformations
- Devide for performance reasons
- Repetition and linking vs. reuse
- Micro Service Architecture

DDI-FlatDB architecture
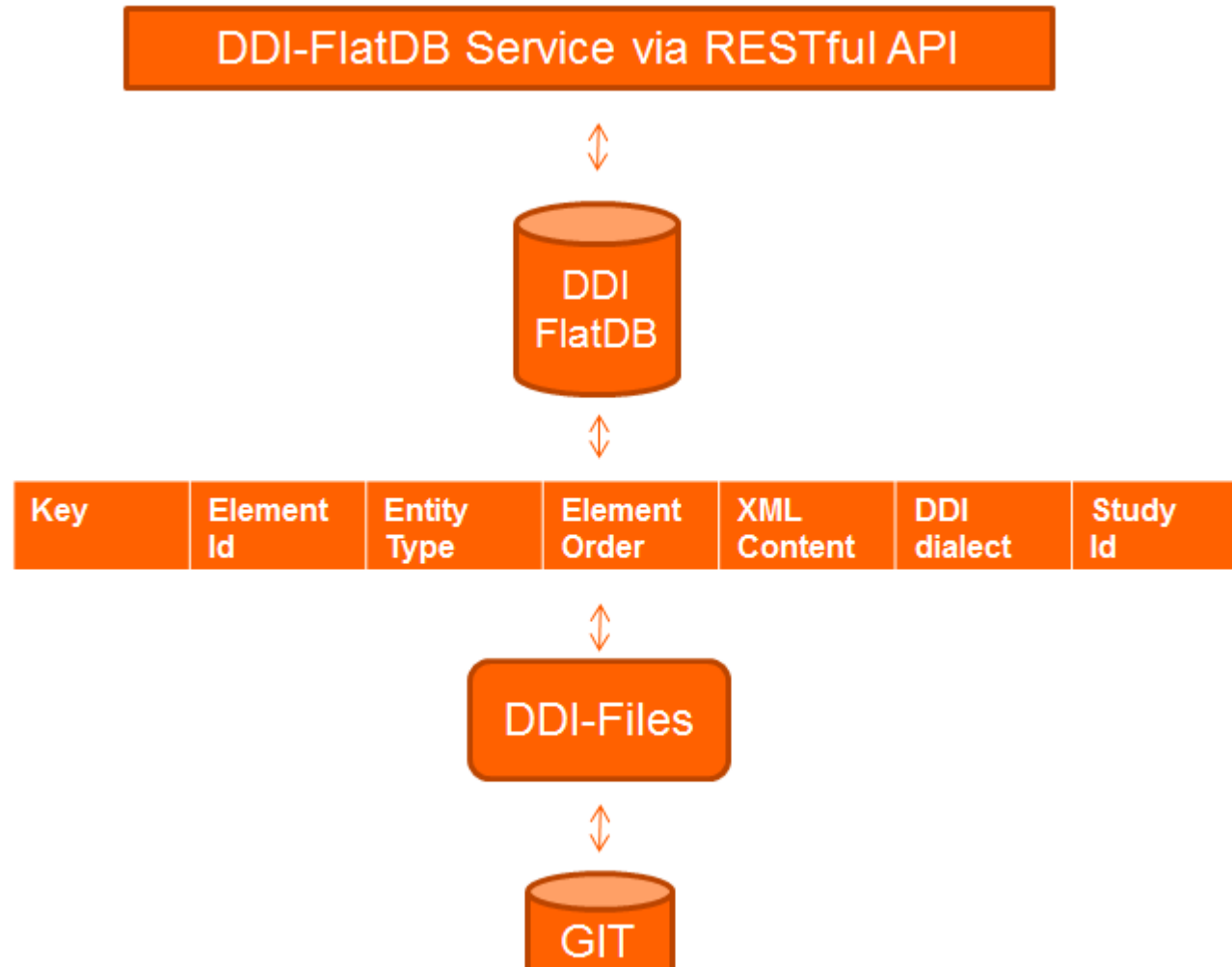
Usage – first steps

Choose document identification node (one per dialect)

- Select entities to cut
  - Need to have identifier
  - Need to form one fragment of XML
- Create a property file with cutting/reassambly information

Example config

```
Concept.path =
      /ResourcePackage/ConceptScheme/Concept
Concept.identifierPath = ./ID
Concept.parentIdentifierPath = ../ID

ConceptPackage.path = /ResourcePackage
ConceptPackage.identifierPath =
      ConceptScheme/ID
```

Working with properties

- Implement a data bean
- Define entitiy properties
  - name, isComplex, isMultiple, isLanguageDependant
- Map properties to paths
- Define a skeleton („unfilled" XML-snippet)

- Reading/writing XML is implemented in a general way

Example bean

```
public class CategoryScheme extends DDIElement
implements Serializable {
  protected String label;
  private ArrayList<Category> category = new
                ArrayList<Category>();
  public CategoryScheme(){
      setType( DDIElement.CATEGORYSCHEME );
  }
  public CategoryScheme( DDIStore ddiStore ){
      setType( DDIElement.CATEGORYSCHEME );
      load( ddiStore );
  }
  public CategoryScheme( String content ){
      setType( DDIElement.CATEGORYSCHEME );
      load( content, getDefaultDialect() );
  }
  public String getLabel(){
      return label;
  }
  public void setLabel( String label ){
      this.label = label;
  }
  public ArrayList<Category> getCategory(){
      return category;
  }
  public void setCategory(
          ArrayList<Category> category ){
      this.category = category;
  }
  public void addCategory( Category c ){
      if ( c.getId() != null && c.getId() != "" )
          category.add( c );
  }
}
```

```
CategoryScheme.fieldNames=Label,Category
CategoryScheme.defaultDialect=gesisquestionnaire32


CategoryScheme.Label.ldep=true
CategoryScheme.Label.multi=false
CategoryScheme.Label.complex=false


CategoryScheme.Category.ldep=false
CategoryScheme.Category.multi=true
CategoryScheme.Category.complex=true
CategoryScheme.Category.class=org.gesis.stardat.enti
ty.Category
CategoryScheme.Category.components=Id,Text
CategoryScheme.Category.Id.ldep=false
CategoryScheme.Category.Id.multi=false
CategoryScheme.Category.Text.ldep=true
CategoryScheme.Category.Text.multi=false


gesisquestionnaire32.CategoryScheme.skeleton=<?xml
version="1.0" encoding="UTF-8"?...
gesisquestionnaire32.CategoryScheme.ID.path=/l:Categ
oryScheme/r:ID
gesisquestionnaire32.CategoryScheme.Label.path=/l:Ca
tegoryScheme/l:CategorySchemeName/r:String
gesisquestionnaire32.CategoryScheme.Description.path
=/l:CategoryScheme/r:Description/r:Content
gesisquestionnaire32.CategoryScheme.Category.multipl
yPath=.
gesisquestionnaire32.CategoryScheme.Category.path=/l
:CategoryScheme/l:Category
gesisquestionnaire32.CategoryScheme.Category.Id.path
=r:ID
gesisquestionnaire32.CategoryScheme.Category.Text.pa
th=l:CategoryName/r:String
```

Select a questionnaire (+ management)

GESIS *DDI-FlatDB*

Manage a questionnaire

Manage a questionnaire

GESIS *DDI-FlatDB*

Edit a question grid

Export result in MS Word

Questionnaire (for fieldwork institute)

vStudytitle : Questionnaire vdocType
vStudyNo, Version 2.0.0, doi:10.4232/1.12000

Page No.

Vw234	Wahlentscheidung Briefwahl
Variable name(s)	
Filter	
Instruction	Musterstimmzettel vorlegen
Question	Sie konnten ja bei der Briefwahl zwei Stimmen vergeben. Die Erststimme für einen Kandidaten aus Ihrem Wahlkreis und die Zweitstimme für eine Partei. Hier ist ein Musterstimmzettel, ähnlich wie Sie ihn bei der Briefwahl zur Bundestagswahl erhalten haben. Was haben Sie auf Ihrem Stimmzettel angekreuzt? Bitte nennen Sie mir jeweils die Kennziffer für Ihre Erst- und Zweitstimme.
Items	[A] Erststimme [B] Zweitstimme
Response scale	[97] trifft nicht zu [98] weiß nicht [99] keine Angabe

Vw002	Beabsichtigte Stimmabgabe

Application architecture
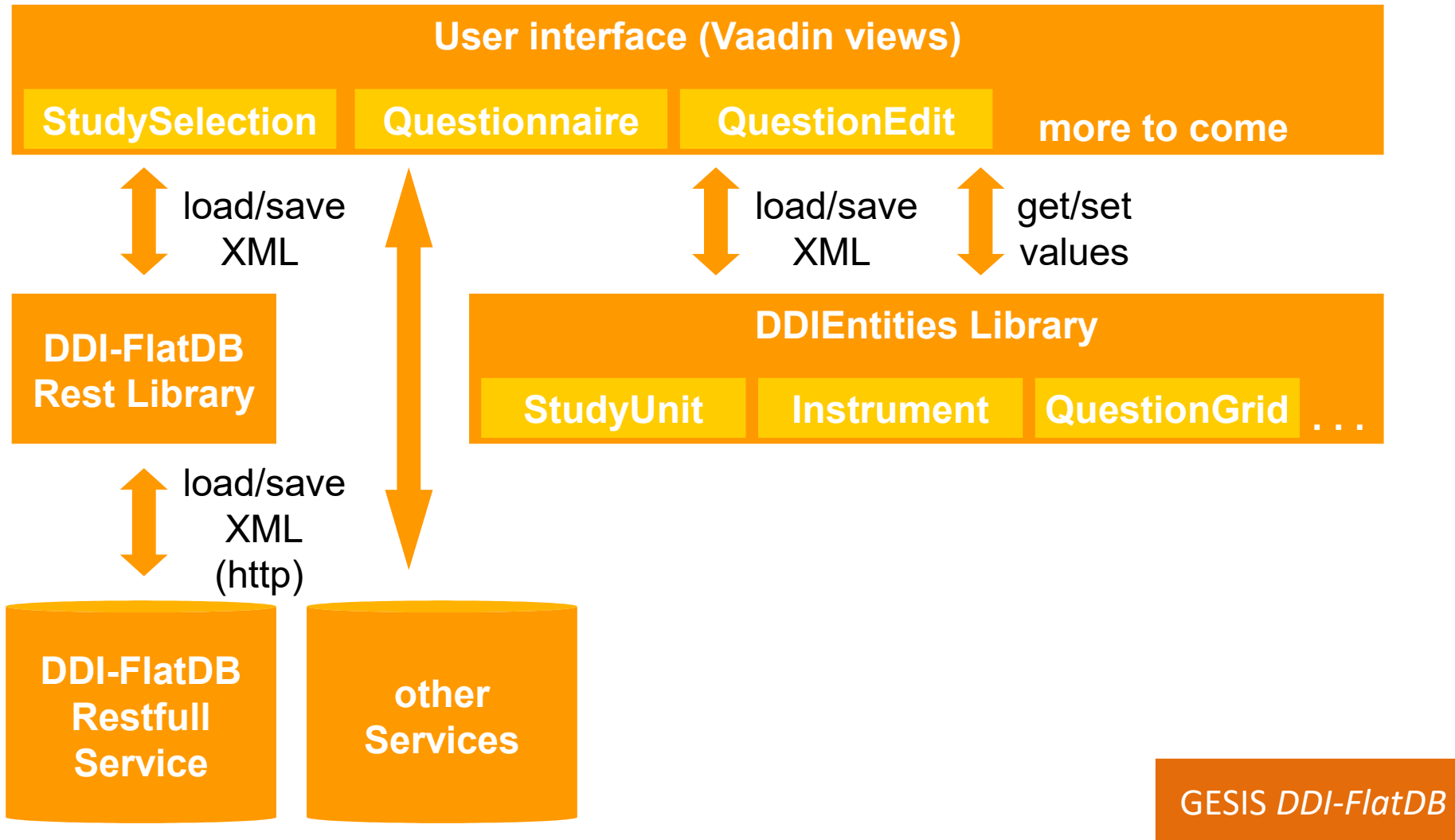
Results so far

- XML Storage
    - Simple and fast DB
    - Accessable as service
    - Revisions through Envers and GIT
    - Relyable recreation of valid DDI

- Questionnaire management
    - Creating/changing Questions and QuestionGrids
    - Arranging/structuring Questionnaires
    - Tagging of items
    - Exporting items to MS Word
    - Linking between items
    - (Deleting items)
    - (Creating/cloning questionnaires/studies)

Outlook on services arround DDI-FlatDB

- Completing the required features
- Collaborative features
- Complete suite approach together with „Explore Data" project
- Study level editor
- General categorization of entities
- Hirarchical search index with ElasticSearch
- Link-DB
- Open source

Contact

The software is planed to be open source. We did not decide for a licence yet so if you are interested, please to contact us.

Oliver Hopt (Oliver.Hopt@gesis.org)
Claus-Peter Klas (Claus-Peter.Klas@gesis.org)
Alexander Mühlbauer (Alexander.Muehlbauer@gesis.org)
Wolfgang Zenk-Möltgen (Wolfgang.Zenk-Moeltgen@gesis.org)

GESIS – Leibniz-Institute for the Social Sciences

GESIS *DDI-FlatDB*