Our Institute

Rare Book Collection, Max Planck Institute for the History of Science
What kind of Research Data are we dealing with in the Humanities?
What kind of Research Data are we dealing with in the Humanities?
Challenges

Doing Research in the Digital Age

• Research data: new formats of research output
• Lack of use of standards
• Incompatible data
• Data tied to old software
• Vanishing data
• Disruption of the research lifecycle
Our Philosophy

• Separate data from software
• Design for end-of-life
• Use dedicated systems and add harmonised access
Tech Stack

REPOSITORY
- Fedora 5
- CDSTAR
- Solr
- Cantaloupe Image Server
- Backend API (Java)
- Admin Frontend (React)

INGEST
- Goobi

DATA ARCHIVE
- DARIAH Repository

KNOWLEDGE GRAPH
- ResearchSpace
- Blazegraph

MAPPING
- X3ML Mapping Engine
- 3M Mapping Tool

“GLUE”
- GitLab (code base and deployment)
- RabbitMQ (message broker)
- KeyCloak (authorisation)
Workflow
Digitisation

Goobi

Mets/Mods XML

Media Files
X3ML Mapping Engine
Workflow
Research Data

Research Dataset
Research Dataset → XML Version → 3M Mapping Tool → Data Archive
Knowledge Graph
### All Things

<table>
<thead>
<tr>
<th>by</th>
<th>from</th>
<th>from</th>
<th>has type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sou 1 V848a -3.2</td>
<td>Rara R718t</td>
<td>Rara B717o -5</td>
<td>Rara E58bod -1</td>
</tr>
</tbody>
</table>
Thank You!

Esther Chen
@EstherDuesberg
www.mpiwg-berlin.mpg.de

Florian Kräutli
@fkraeutli
www.mpiwg-berlin.mpg.de