FAIR research data integration in CRIS at FAU Erlangen-Nürnberg
# Table of content

1. FAU at a glance
2. FAU CRIS
3. Motivation and Challenges
4. FAIR Data
5. Solution & Workflow
6. Summary
One of Germany’s Top 10 Research Universities

- Founded in 1743
- More than 3,600 members of academic staff
- More than 2,300 employees in administration
- 624 Professors
One of Germany’s Top 10 Research Universities

- More than 38,000 students in
- 267 degree programs, among them
  - 81 Bachelor’s degree programs and
  - 94 Master’s degree programs
FAU CRIS
CRIS Timeline

Start
Decision to use Converis as a CRIS

2011

Launch
Introduction of Publications and Projects

2015

Expansion
Addition of various other research output

2016

Workflows
Integration of administrative workflows

2018+

Automation
Automated import from Web of Science & Scopus

2019

Research Data Integration
supporting scientists in FAIR RDM

2022+

Current content: 125k publications, 9k research projects, 2k awards, … - in total 1m datasets and 2.5m relations.
Motivation and Challenges
“In the future, the handling of research data will receive more attention than before in assessment and evaluation.”
Motivation and Challenges

- … DFG (German Research Foundation) stated on March 14th 2022*
- Data publications is research outcome and so part of research information

* https://www.dfg.de/foerderung/info_wissenschaft/info_wissenschaft_22_25/index.html
Motivation and Challenges

- … DFG (German Research Foundation) stated on March 14th 2022*
- Data publications is research outcome and so part of research information

- Should we build up a new meta-data system for meta-data or just use our CRIS?
- Can we support FAIR research data management by using our CRIS?

* https://www.dfg.de/foerderung/info_wissenschaft/info_wissenschaft_22_25/index.html
Motivation and Challenges

- … DFG (German Research Foundation) stated on March 14th 2022*
- Data publications is research outcome and so part of research information

- Should we build up a new meta-data system for meta-data or just use our CRIS?
- Can we support FAIR research data management by using our CRIS?

- OF COURSE we will use our CRIS!
- FAIR principle support is vital!

* https://www.dfg.de/foerderung/info_wissenschaft/info_wissenschaft_22_25/index.html
- Research meta data standard: DataCite Metadata Schema (4.4 currently)
  - Schema gives detailed description of research data
- Three levels of obligation for attributes:
  - Mandatory (e.g. title, identifier)
  - Recommended (e.g. subject, description)
  - Optional (e.g. language, version)
- By using a common meta-data schema, we are able to use different repositories at the end
- We think that some recommended or optional attributes should be mandatory for better FAIR meta-data (e.g. description)
- Because we control the data collection with our CRIS, we can do this
Solution & Workflow
Solution & Workflow: Systems

- CRIS: research information system 😊
- FAUL: meta data generation & upload software
Solution & Workflow: Core Steps

CRIS

FAUL

Research dataset definition

Metadata generation

Data: upload to repository

Metadata: Research information & data portal
FAUL tool
### FAUL tool

#### 1. Select Folder

#### 2. Generate Meta-Data for CRIS

<table>
<thead>
<tr>
<th>filename</th>
<th>status</th>
<th>rename</th>
<th>remove from list</th>
<th>filepath</th>
<th>mimetype</th>
<th>size</th>
<th>checksum</th>
<th>modification date</th>
</tr>
</thead>
<tbody>
<tr>
<td>data.csv</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>C:/Users/muuq04uwuv/Des</td>
<td>application/vnd.ms-excel</td>
<td>133545532 byte</td>
<td>✗</td>
<td>2022-06-06T12:10:27</td>
</tr>
<tr>
<td>image.png</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>C:/Users/muuq04uwuv/Des</td>
<td>image/png</td>
<td>118428 byte</td>
<td>✓</td>
<td>2022-06-06T12:10:27</td>
</tr>
<tr>
<td>script.py</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>C:/Users/muuq04uwuv/Des</td>
<td>text/x-python</td>
<td>13194 byte</td>
<td>✓</td>
<td>2022-06-06T12:10:27</td>
</tr>
</tbody>
</table>

#### Metadata generation

- Research dataset definition
- Metadata generation
- Data: upload to repository
- Metadata: Research information & data portal

---

Friedrich-Alexander-Universität Erlangen-Nürnberg
### FAUL tool

#### 1. Select Folder

<table>
<thead>
<tr>
<th>filename</th>
<th>status</th>
<th>rename</th>
<th>remove from list</th>
<th>filepath</th>
<th>mimetype</th>
<th>size</th>
<th>checksum</th>
<th>modification date</th>
</tr>
</thead>
<tbody>
<tr>
<td>data.csv</td>
<td>✓</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td>13354553 byte</td>
<td>2022-05-06T11:56:27</td>
<td></td>
</tr>
<tr>
<td>image.png</td>
<td>✓</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td>118428 byte</td>
<td>2022-05-06T11:56:27</td>
<td></td>
</tr>
<tr>
<td>script.py</td>
<td>✓</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td>13194 byte</td>
<td>2022-05-06T11:56:27</td>
<td></td>
</tr>
</tbody>
</table>

- Interface locked: generating checksums for your dataset

---

**Research dataset definition**

**Metadata generation**

**Data: upload to repository**

**Metadata: Research information & data portal**

---

**CRIS**

**FAUL**
CRIS workflow

CRIS: step 1a

- Add title
- Add authors
- Upload FAUL meta data file
CRIS workflow

**CRIS: step 1b**

- CRIS analyses file content
- Content of dataset is listed
CRIS: step 2

- Add meta data for each file
CRIS step 3

- Relate the dataset to other research information already in the system (optionally, but strongly recommended)
  - Publications
  - other datasets
  - Projects
  - Awards
  - Research fields
  - Research infrastructure
  - ...
- No need to enter author details
CRIS workflow

CRIS: step 4

- CRIS checks storage limit for chair
- Approval workflow
- CRIS generates upload token
- Permissions granted on upload system
- User enters CRIS' id to get the token into FAUL
### FAUL tool

#### 1. Select Folder

<table>
<thead>
<tr>
<th>filename</th>
<th>status</th>
<th>rename</th>
<th>remove from list</th>
<th>filepath</th>
<th>mimetype</th>
<th>size</th>
<th>checksum</th>
<th>modification date</th>
</tr>
</thead>
<tbody>
<tr>
<td>data.csv</td>
<td>✔</td>
<td></td>
<td>✗</td>
<td>C:/Users/muuq04uwuvu/Des</td>
<td>application/vnd.ms-excel</td>
<td>13354553 bytes</td>
<td>67bd6004c</td>
<td>2022-05-06T12:10:27</td>
</tr>
<tr>
<td>image.png</td>
<td>✔</td>
<td></td>
<td>✗</td>
<td>C:/Users/muuq04uwuvu/Des</td>
<td>image/png</td>
<td>118428 bytes</td>
<td>14b7b1a0b</td>
<td>2022-05-06T12:10:27</td>
</tr>
<tr>
<td>script.py</td>
<td>✔</td>
<td></td>
<td>✗</td>
<td>C:/Users/muuq04uwuvu/Des</td>
<td>text/x-python</td>
<td>13194 bytes</td>
<td>3e7e090f</td>
<td>2022-05-06T12:10:27</td>
</tr>
</tbody>
</table>

#### 2. Generate Meta-Data for CRIS

#### 3. Upload Dataset to FAUdataCloud

---

### Research Dataset Definition

- Metadata generation
- Data: upload to repository
- Metadata: Research information & data portal

---

Friedrich-Alexander-Universität Erlangen-Nürnberg

12. Mai 2022

38
Back to FAIR data…
Addressing FA in FAIR

- Dataset visibility is controlled inside CRIS
- Internal access: neither meta-data nor data is publically available, but both can be shared using private URIs
- Public Meta-Data: Only the data is not publically available. DOI is assigned automatically
- Open Data: meta data and data are available to the public
- Data portal supports OAI Protocol for Metadata Harvesting
Addressing FA in FAIR
Research Information & Data Portal

under construction
Addressing FA in FAIR
Research Information & Data Portal
Summary
Summary

- No additional system for research meta-data handling
- DataCite 4.4. implemented in our CRIS
  - Not the schema itself, but all relevant attributes
- Workflow and software to support the whole process of data set management
- Portal for integrated research information and research data driven by CRIS