Demonstration of the Infrastructure Use Case 02
Framework for curation and distribution of reference datasets

Contributors:
Thomas Hammerschmidt, Mariano Forti, Luis Ávila, Jürgen Olbricht, Birgit Skrotzki, Anika Lenze, Angelika Gedsun, Tilmann Hickel, Hanna Tsybenko, Mirek Chmielowski, Sirieam Hunke, Yusra Shakeel, Rainer Stotzka
Introduction and Motivation

“The aim of IUC2 is to develop a framework for reference material data sets using creep properties of single- & polycrystalline Ni-based superalloy as example.”

https://nfdi-matwerk.de/infrastructure-use-cases/iuc02-framework-for-curation-and-distribution-of-reference-datasets

- **User/User Journey**: Research task could be alloy design, optimization of mechanical properties (chemical composition, heat treatment strategy)
- **Data available**: Normal’ data set (DS) e.g. from SFB/TRR 103 vs. reference data set (RDS) e.g. from BAM
- A **RDS** would be a DS used
  - as high quality data to correlate material properties and serve as baseline for the optimization process
  - to calibrate the experiments in the own lab
- **Requirements (contents, shape, FAIR, accessibility)**

Demonstrator at the booth at marketplace: GitLab (Notebook with data)
**User Journey**: [https://app.conceptboard.com/board/x7zn-2yg0-eabn-piic-dbiu](https://app.conceptboard.com/board/x7zn-2yg0-eabn-piic-dbiu)

creep data for different superalloys (Horst et al, Superalloys 2020)
“develop a framework for reference material data sets”

Reference Data Set

- Definition, criteria for reference data (material, equipment, procedures, results)
- ... Community process needed → Poster
- Curation process

User/Usage Analytics

- Testing laboratories
- Researchers

Possible uses
- Comparison of data for interpretation of individual measurement results
- Calibration/verification of measurement devices
- ... Community process needed → Poster

→ Different user cases derive different requirement profiles (for content and shape of data)

Creating the reference dataset

Content:
- MSE perspective or “Fit for purpose”¹
- Agreement on content based on input from domain experts

Shape:
- RDM perspective or “Fit for use”¹
- Creation, structuring and distribution of (creep) (meta)data: FAIR DOs

Distribution/Publishing of the Reference Data Set
- Ontology, Licensing, DOI

¹ Ariza, Angela et al., Datennachnutzung in der Praxis; DOI: 10.5281/zenodo.7568266
Agreement and Data Structuring

- **Agreement on:**
  - (meta)data and documentation, considering the existing datasets (BAM, RUB, SFB/TR 103)
  - data structure and vocabulary (ISO 204-standard on creep testing)
  - a “rating” of the (meta)data to define quality requirements in terms of contents

- **Vision:** Creation (a) data schema template(s) or check lists to be used by the community (to generate RDS)
Shaping Exemplary Data Set: FAIR Digital Object

Data object
- Link to the location e.g. storage location of the data file
- Link(s) to metadata object(s)
- PID
- Type
- ....

Metadata object
- Link to the location e.g. storage location of the metadata document
- PID
- Type
- ....

FAIR Digital Object:
- Representation of data
- Contains all information towards FAIR
- Conception: technology agnostic
- Bridges between data repositories, disciplines, etc.
- Implementation
  - Handle PID
  - Information Record

<table>
<thead>
<tr>
<th>PID</th>
<th>PID Profile</th>
<th>Type</th>
<th>Location</th>
<th>URL</th>
</tr>
</thead>
</table>

28.06.2023 Creep Reference Data
Exemplary RDM Reference Data Set: Technical Process Overview

Data Examination & Structuring

Requirement Analysis

JSON Schema Registration

Metadata Document Registration

Kernel Information Profile Adoption

FAIR DO Creation

FAIR DO Interpretation

28.06.2023

Creep Reference Data
Towards Ontology for Creep Data (Reference Data Set)

Ongoing work

- Domain: MSE (High temperature application/ creep testing)

- Requirements:
  - Representation of creep testing in a conceptual way for a reference material and based on a testing standard/norm → user can transfer the concepts in their own work

- Status: Metadata annotation /agreement on controlled vocabulary (with BAM/RUB)

- Reuse:
  - NFDIcore
  - PMDcore Ontology on tensile testing
Data is produced

Define criteria for RDS and verify

Agreement on content

Data Structuring

Creating FDO(s)

Publish/Distribute data

User Journey

USER

Key Takeaway

28.06.2023

Creep Reference Data
Back-Up slides
A reference dataset is available. It provides:

- Data schema
- Example data

A group wants to publish data related to previous publication. They follow the next steps:

- get the data schema
- Prepare their data for schema compatibility
- Fill the data schema with their data
- compare the data with reference dataset to check completeness
Application for data schema and control vocabulary

Data schema / vocabulary from experts

Reference data

Structured reference data

Inhalt/Content
Die Daten umfassen Ergebnisse Referenzmaterial der Nickel-Ba-
Data include results of ASTM E nickel-base alloy.
Folgende Dateien sind enthalten
Arc/Type
Zertifikat des Referenzmaterials Reference material certificate
Ergebnisse Kriechversuche nach ASTM
Material: NIMONI
Versuchsparameter: Probe m

<table>
<thead>
<tr>
<th>Proben- bezeichnung</th>
<th>Zeichnung</th>
<th>Versuch</th>
<th>Nr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>K2K180-08</td>
<td>K2K180-08</td>
<td>K2K180-08</td>
<td>249</td>
</tr>
<tr>
<td>K2K180-04</td>
<td>K2K180-04</td>
<td>K2K180-04</td>
<td>242</td>
</tr>
<tr>
<td>K2K180-02</td>
<td>K2K180-02</td>
<td>K2K180-02</td>
<td>238</td>
</tr>
<tr>
<td>K2K180-01</td>
<td>K2K180-01</td>
<td>K2K180-01</td>
<td>248</td>
</tr>
<tr>
<td>K2K180-00</td>
<td>K2K180-00</td>
<td>K2K180-00</td>
<td>227</td>
</tr>
</tbody>
</table>

28.06.2023
Creep Reference Data

Notebook 00_GetSchema
Notebook 01_BAM_to_schema
Application for data schema and control vocabulary

Data from working group ➔ Mappings ➔ data schema / data structure ➔ Structured user data

- Metadata:
  - Test info:
  - Test material:
  - Testing and measuring equipment:
- Primary data:
  - Test results:
    - Bibliographic information:
    - State at test start:
    - Test sequence:
    - State at end of test:
    - raw_elongations:
    - raw_times:
- Secondary data:
  - Test results:
    - Test sequence:
      - Corrected measured temperature:
      - Loading rate:
      - Unloading rate:
      - Heating speed:
      - Cooling speed:
      - Elongation and extension:

Notebook 02_GetRubTables
Notebook 03_Rub_to_schema
Application for data schema and control vocabulary

User 1 data + User 2 data → Direct comparison of data (automated)

- Metadata:
  - Test Info:
    - Bibliographic information:
      - Title: Creep Reference Data

- Test 1:
  - Value: 115.08.2023
  - Unit: ""
  - Symbol: ""

- Test ID:
  - Value: KK 970
  - Unit: ""
  - Symbol: ""

- Projects:
  - Title: "1st. Ringversuch 2017"
  - Unit: ""
  - Symbol: ""

- Operators:
  - Value: "NN1"
  - Unit: ""
  - Symbol: ""

- Specified test parameters:
  - Testing Standard:
    - Value: "EN 139"  
    - Unit: ""
    - Symbol: ""
  - Specified Temperature:
    - Value: "600°C"
    - Unit: ""
    - Symbol: ""
  - Initial stress:
    - Value: "500 MPa"
    - Unit: ""
    - Symbol: ""
  - Test type (interrupted/not interrupted):
    - Value: "MISSING"
    - Unit: ""
    - Symbol: ""
  - Test force:
    - Value: "100 kN"
    - Unit: ""
    - Symbol: ""

- Test order:
  - Operator:
    - Value: "NN1"
  - Project:
    - Value: ""
  - Symbol: ""

- Test data:

- Metadata:
  - Test Info:
    - Bibliographic information:
      - Title: Notebook 04_CompareDataSources

28.06.2023
Creep Reference Data
Demonstrator scheme

**BAM Creep reference data (NIMONIC 75, CMSX-6)**

**Jupyter to merge the diff. datasets**

**Jupyter to fill schema**

**Merging of the vocabulary (diff. data entries), Glossary**

**RUB uses/evaluates the reference data**

**Jupyter to access? + read + analyze data (multiple datasets possible)**

**User Journey**

→ User’s experience told by means of **User Journey** considering researcher’s and data stages, researcher’s goals and actions, etc

**set requirements to constrains?**

**User Journey:**

https://app.conceptboard.com/board/x7zn-2yq0-eabn-piic-dbiu

---

28.06.2023

Creep Reference Data
Summary

Data structuring
- via script or interface
- Check against template/schema
- specific new agreements might be necessary (e.g. for other material class)

Creating the FAIR DOs
- Create JSON Schema
- Define metadata document
- Adopt Kernel Information Profile
- Create & resolve FDO

Distribution of data
- license
- DOI
- FAIR DOs („metadata“)
- Ontology
- (log/processed data)

User accesses and uses RDS to solve research question

Definition of/criteria for reference data

Agreement on data schema, thesaurus, ranking, quality levels

Curation process, quality assurance