Faculty of Mechanical Science and Engineering
Institute of Mechatronic Engineering
Chair of Machine Tools Development and Adaptive Controls
Prof. Dr.-Ing. S. Ihlenfeldt

A Metadata Model for Harmonising Engineering Research Data Across Process and Laboratory Boundaries

Martin Zinner[a], **Felix Conrad[a]**, Kim Feldhoff[a], Hajo Wiemer[a], Jens Weller[b], Steffen Ihlenfeldt[a]
[a] TUD Dresden University of Technology
[b] Symate GmbH Dresden, Germany

Venezia // 17.04.2024

Presenter: Dipl. – Ing. Felix Conrad

Professional Experience

July 2019	Diploma degree, Mechanical Engineering
August 2020	Doctoral Researcher, TU Dresden, Germany

Publications

• Benchmarking AutoML in materials design	npj scientific reports
• AutoML Applied in Production Engineering	Proc. Comp. Science
• Need for UAI – Anatomy of Usable Artificial Intelligence	MDPI MTI
• Comparative Analysis of Small Data Acquisition Strategies	IARIA Software
• Impact of Data Sampling on Performance of ML-Models	WGP 2022

Research Interest

- Machine-learning for Materials Design
- Automated Machine-learning
- Usable AI

Dipl.-Ing.Felix Conrad,

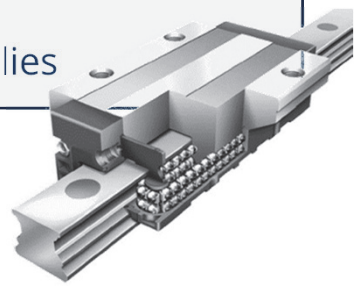
felix.conrad@tu-dresden.de
TUD Dresden University of Technology
Institute of Mechatronic Engineering
Department Machine data utilisation

TUD - Department "Machine data utilisation"

Using data for production technology

Machine health

- Failure detection
- Prediction of anomalies

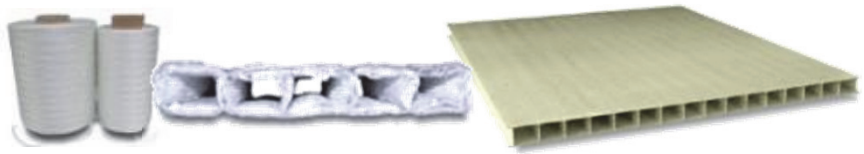


Machine operation/start-up

- Process/station monitoring
- Monitoring operating accuracy

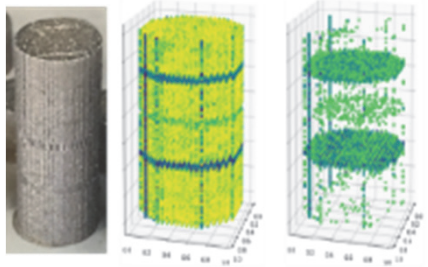
Manufacturing process chains

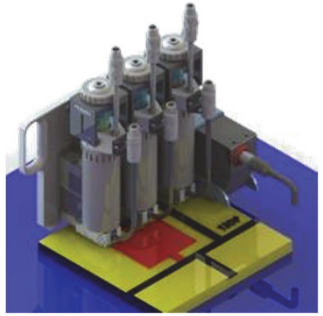
- Exploring effects
- Overarching correlations



Quality and process

- Quality prediction
- Finding optimal process parameters

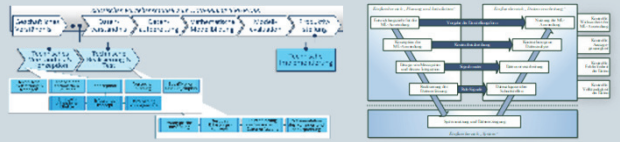




Method development

AI algorithm selection and testing

Abstraction in workflow models



User-friendly provision - UAI

Research data management

Content

Motivation

Use case GRK (Research training group) 2250

- Introduction of GRK 2250
- Status of the Research Data Management

Useable FAIR - Solution approaches for Research Data Management

- Thesaurus for Metadata
- Process data model

Results

Conclusion and outlook

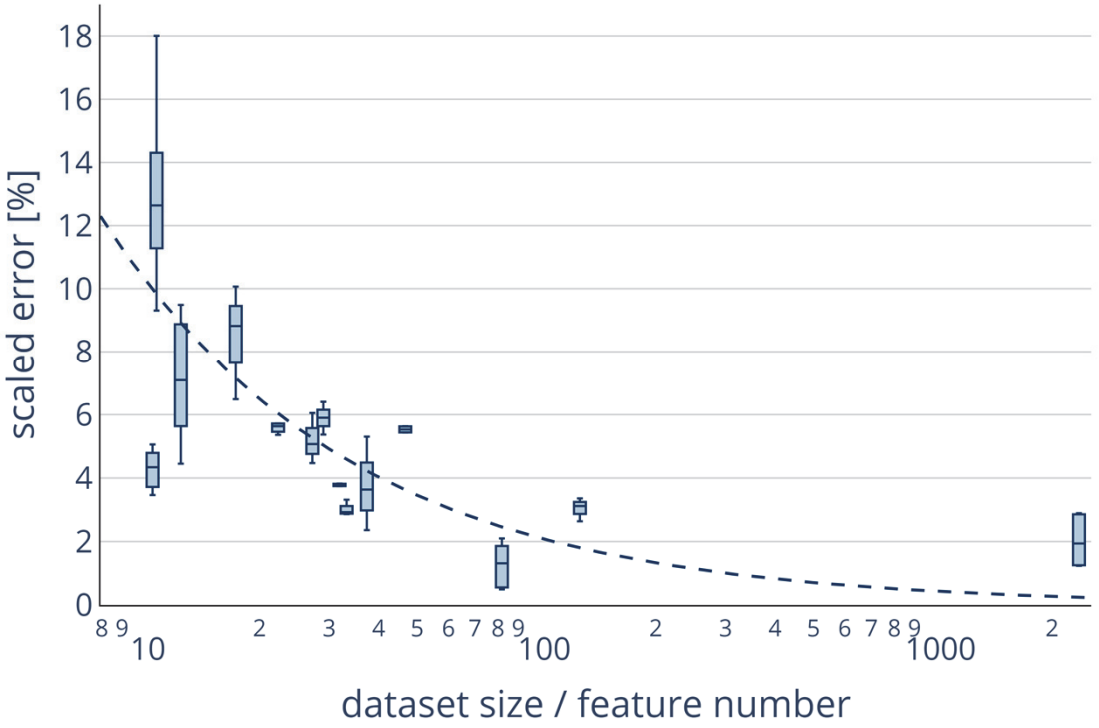
Motivation

Research Data Management perspective

- Fulfilment of the FAIR principles
- Sustainable data management
  → Ensure reusability of data
- Costly experiments in terms of time and resources
  - Drop-Tower-Experiment:
    - Costs for 1 test: ~ € 2,000
    - preparation time: 7 months

Machine learning perspective

- More data leads to better models
- 12 datasets from materials design domain:

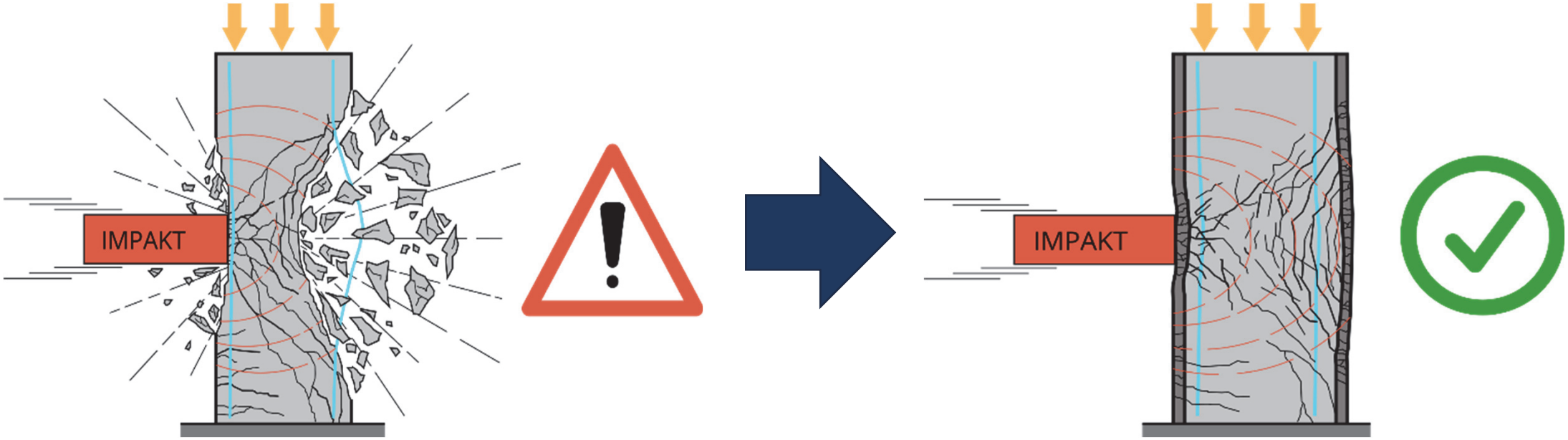


Wiemer, H.; Conrad, F.; et al.: *Illustration of the usable AI paradigm [...]* (2023)

Project GRK 2250:
Mineral-bonded composites for enhanced structural impact safety

Overall objective of GRK (Research training group) 2250:

- New generation of textile reinforced mineral-bonded composites
- Theoretical, numerical and data-driven modelling
- Application for the reinforcement of existing structures



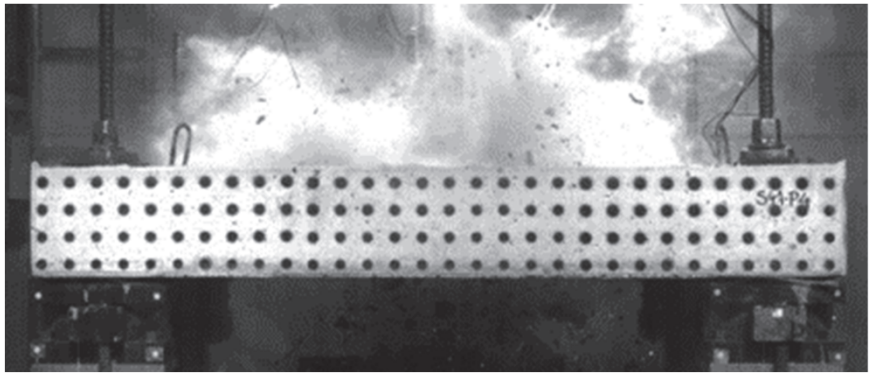
graphic: **Curosu et al.** (2019),
DOI: 10.21012/FC10.235408

More information: https://www.grk2250.de/

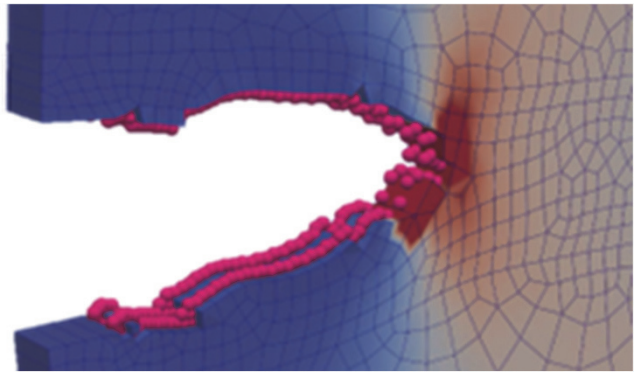
Project GRK 2250:
Mineral-bonded composites for enhanced structural impact safety

Experimental: material development

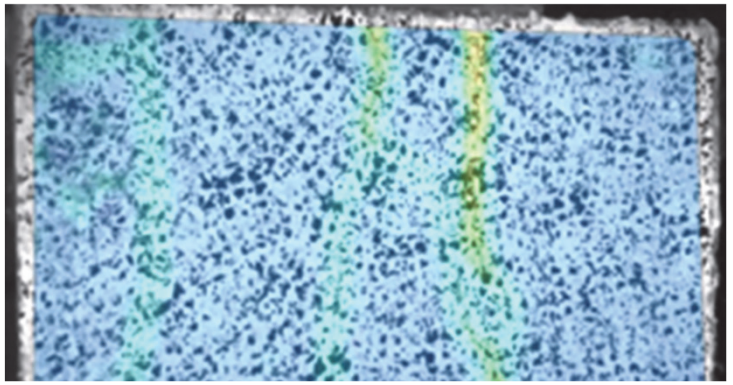
Experimental: structural assessment



Numerical simulation



Data driven and cross-disciplinary



Status quo: Research Data Management in GRK 2250

What:

- Three periods with changing researchers, totalling ~40 researchers over 9 years
- Cumulative data volume of 10 terabytes from ~15 test systems across six laboratories
- Each test system conducted 20 to 300 experiments

How:

- Infrastructure:
  - Shared drive accessible to all partners
  - Mostly Excel spreadsheets or .csv files for (meta)data storage
- Highly manual workflow

Result:

- Unstructured storage of research (meta)data

Approach: FAIR – extension to Useable FAIR

Findable

- Unique identifiers
- Rich metadata
- (Meta)data in a searchable resource

Accessible

- open access
- long-term accessibility

Usable

The data management system:

- Easily adapted by data curators to research network
- users can store, edit and use (meta-)data that complies with FAIR with minimal effort.

Interoperable

- Meta(data) using common standards + vocabularies
- Integration with other meta(data) enabled.

Reusable

- Meta(data) is prepared for easy reuse, with domain-relevant community standards

Approach: FAIR – extension to Useable FAIR

Requirements for cross-process research data management

Making metadata understandable

Record all necessary metadata

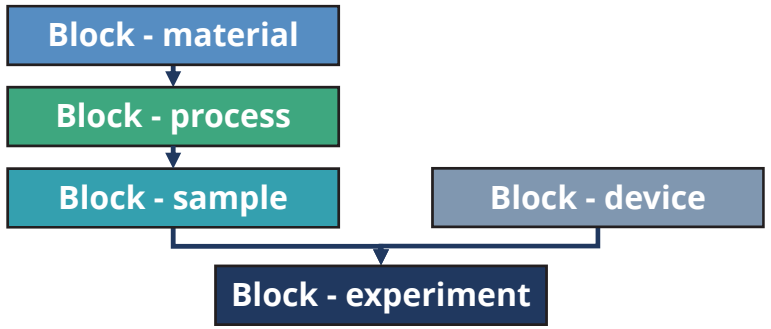
Thesaurus for metadata

Domain ontology

Translation via thesaurus for domain ontology

Researcher ontology

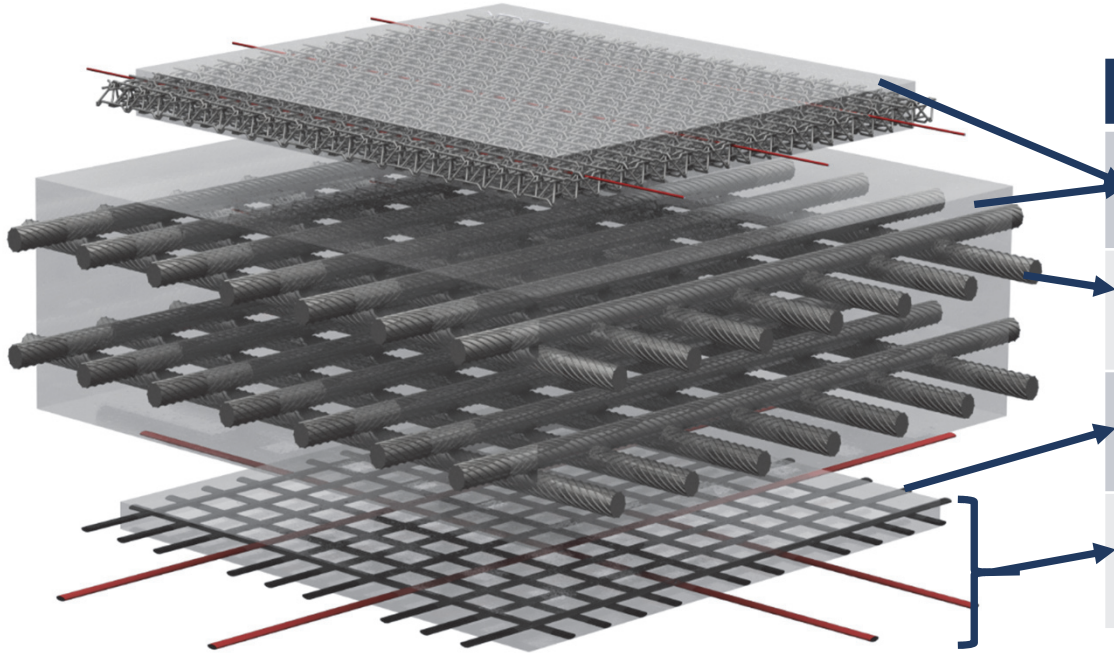
Process data model



Approach - 1 : Problem - Diverse Standards in Materials Testing

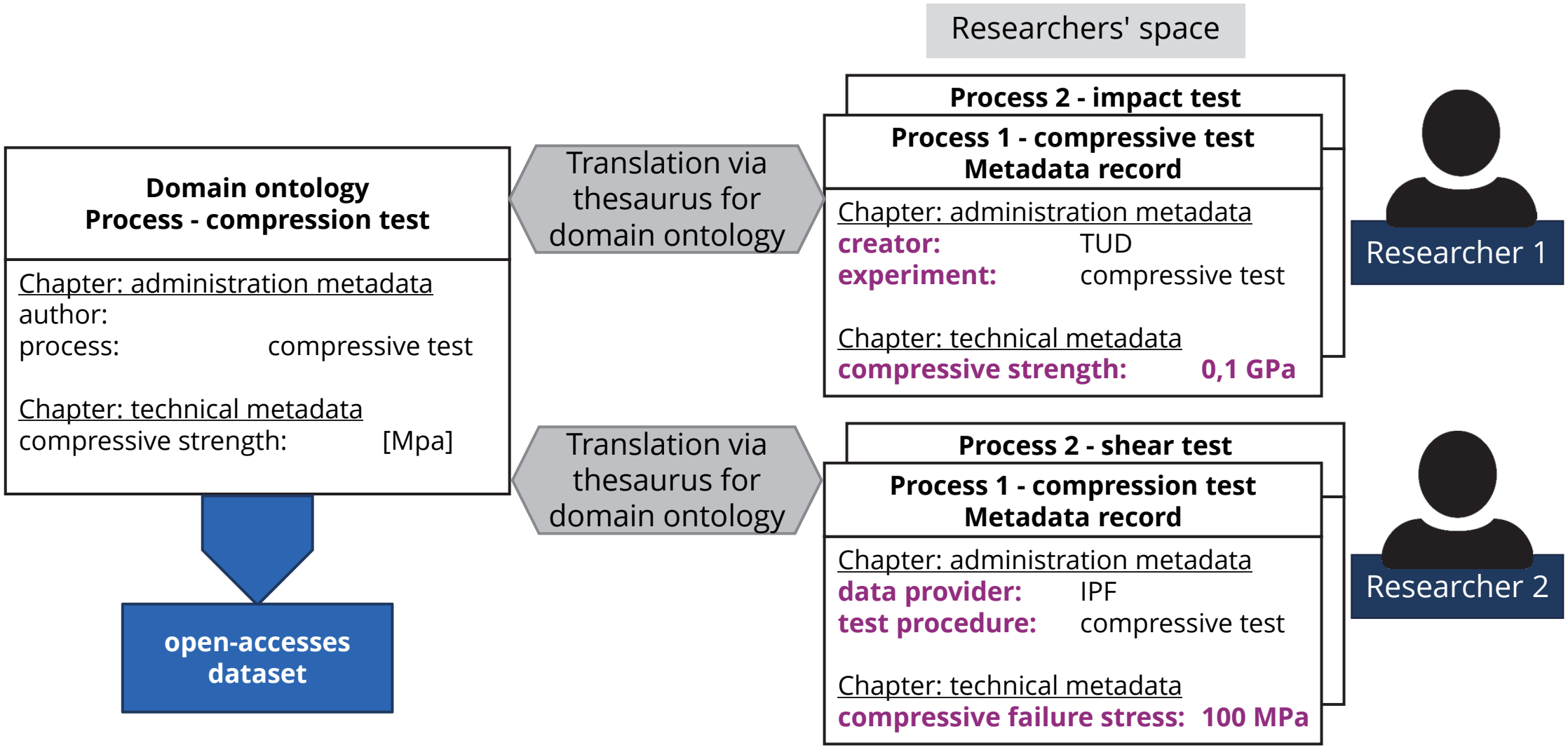
Different conventions in international standards for the same physical properties

- **Example compression test:**



Standard	Material	Symbols	Naming
ISO 1920-4	Concrete	f_c	compressive strength
DIN 50106	Metals	R_{dB}	compressive strength
ISO 604	Plastic	σ_m	compressive strength
ISO 14126	Fibre-reinforced plastic	σ_{cM}	compressive failure stress

Approach - 1 : Problem - Diverse Standards in Materials Testing



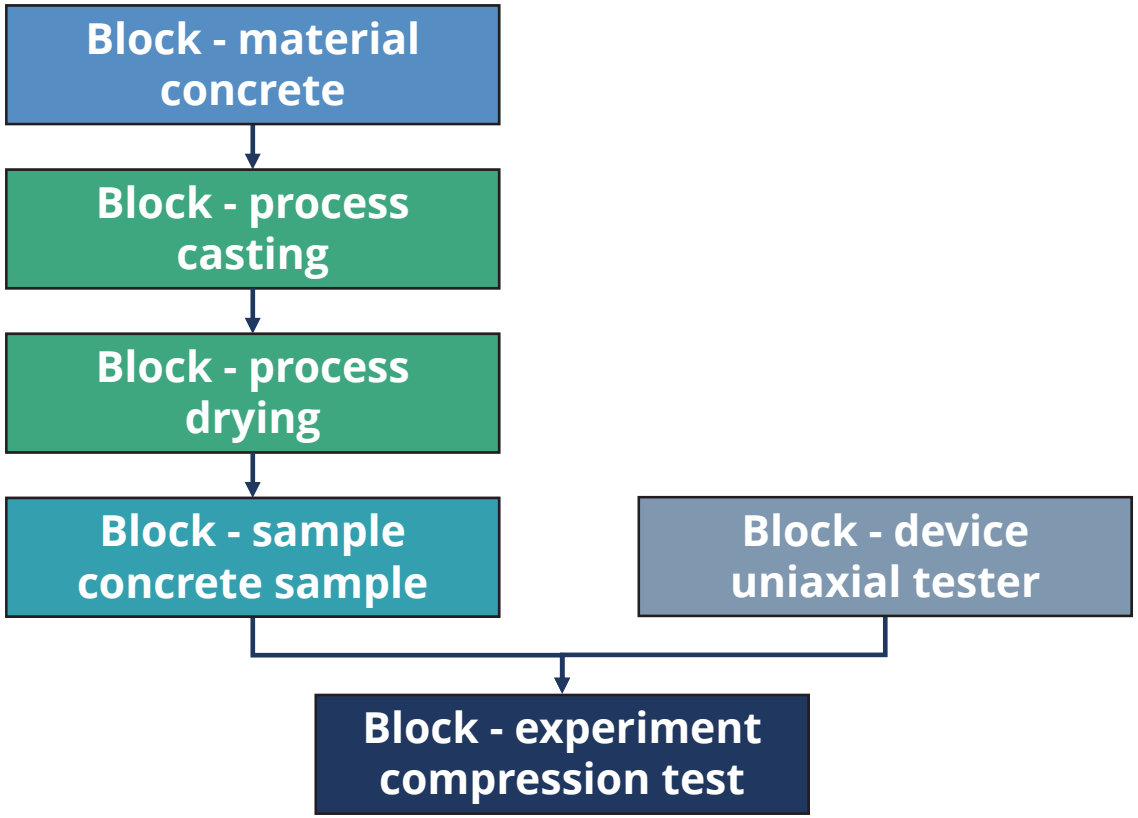
Approach - 2 : Process Data Model

Problem:

- Processes can only be stored reproducibly if the entire process chain is stored reproducibly

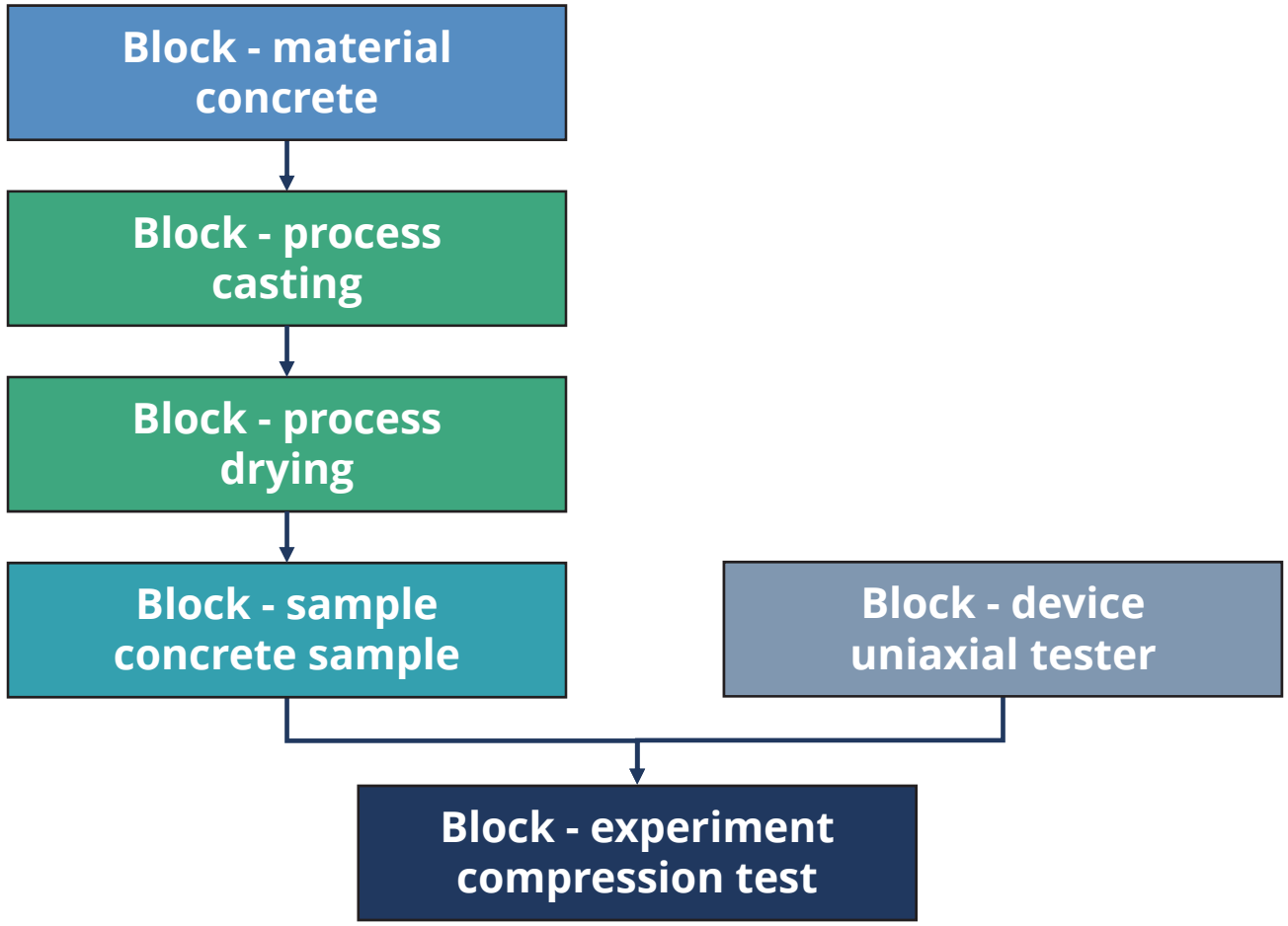
Solution:

- *Process Data Model*



Approach - 2 : Process Data Model

Block - material concrete	
Chapter: administration metadata	
author subproject	“Felix Conrad” “C3II”
Chapter: technical metadata	
water content	150 kg/m³



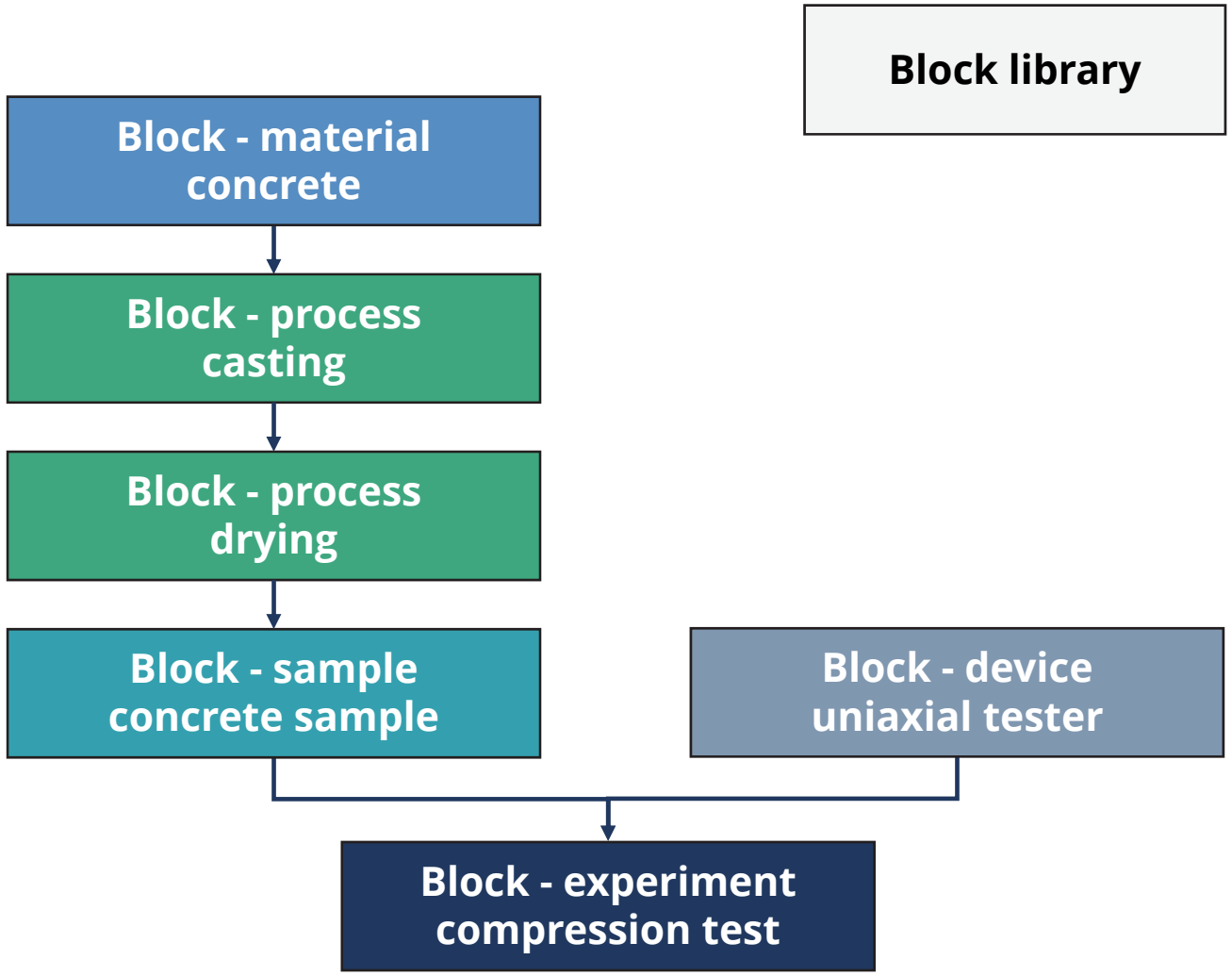
Approach - 2 : Process Data Model

Metadata library

subproject	
datatype choice	string [“A1I”;…;”C3III”]
compressive strength	
datatype unit	float [MPa; GPa]
⋮	

Block - experiment compression test

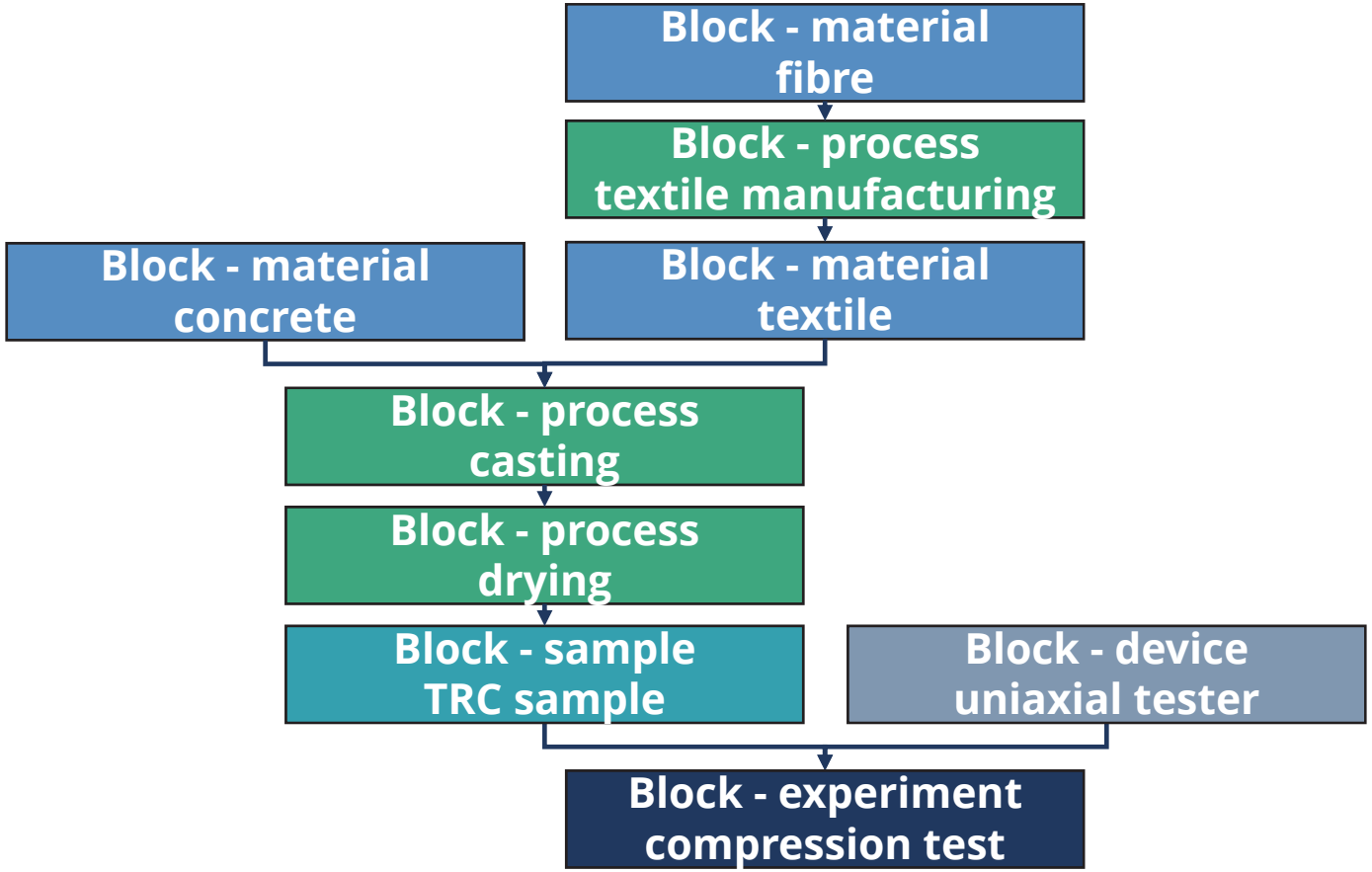
Chapter: administration metadata	
author subproject	“Felix Conrad” “C3II”
Chapter: technical metadata	
compressive strength	100 Mpa



Approach - 2 : Process Data Model

Process data models crucial for long and branched process chains

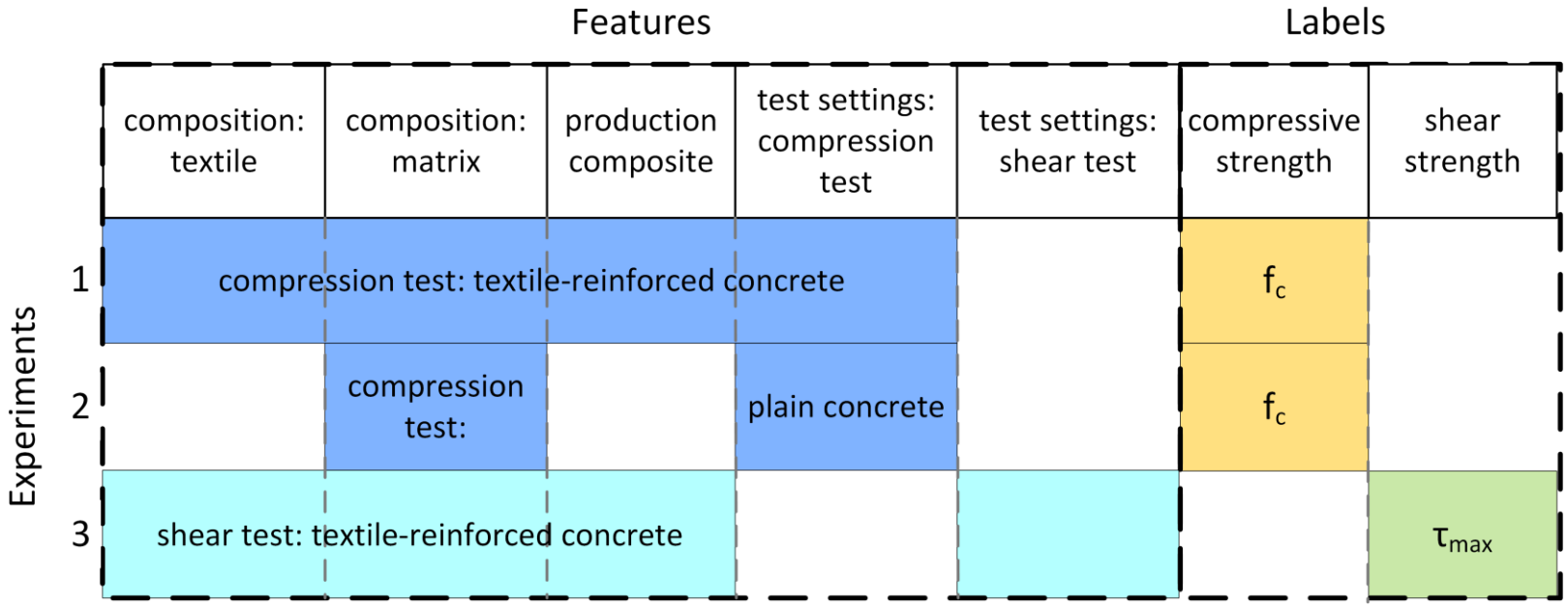
- **Example:** Compression test of textile reinforced concrete



Results: Example of a Combined Dataset

Combined Dataset: different processes provided by different institutions:

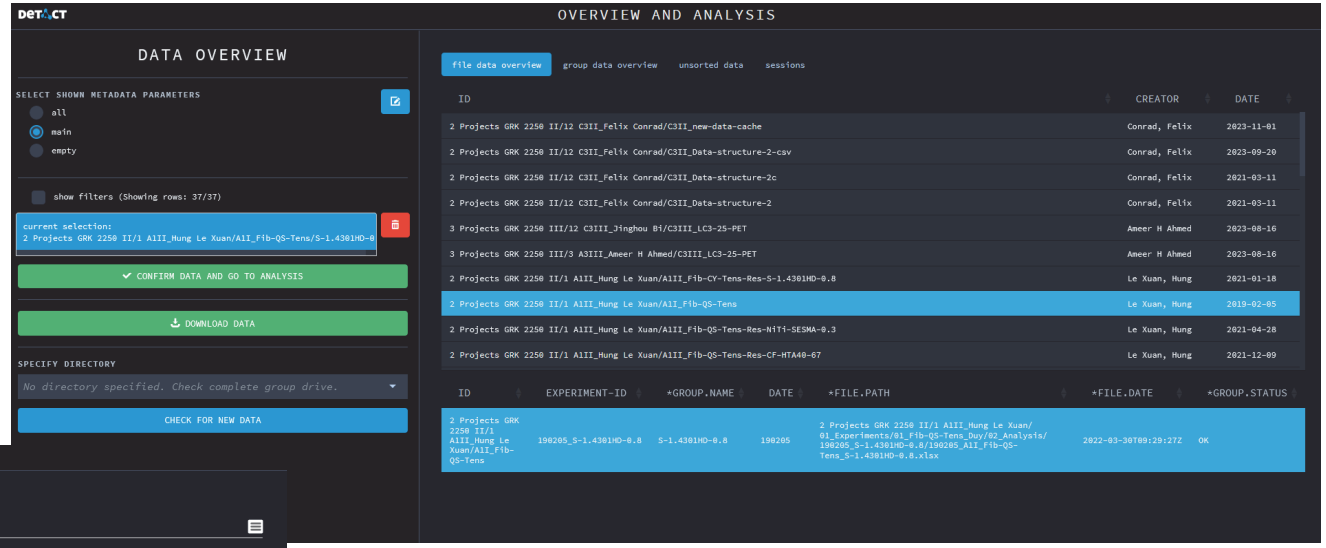
- Experiment 1: compressive test of textile reinforced concrete – Institute 1
- Experiment 2: compressive test of plain concrete – Institute 1
- Experiment 3: shear test of textile reinforced concrete – Institute 2



Results: Implementation of RDM Platform

Platform ensures FAIR principles :

- **F**indable
- **A**ccessible
- **I**nteroperable
- **R**eusable



Further tools for researchers:

- Visualisation of Data
- Analysing Data

Conclusion

- Strategy **harmonizes diverse working cultures** of researchers across domains
- Metadata model **allows merging data:**
  - from similar processes provided by different institutions or fields
  - Merging data from different processes along a process chain

Outlook

- **Completion of the implementation** to a full research data management tool
- **Usability** aspect not fully investigated:
  - **Participant engagement** remains to be analysed
  - **Easy transferability** of the implementation to other research projects to be investigated

Thank you for your attention!

Felix Conrad
TU Dresden
Institute of Mechatronic Engineering
Chair of Machine Tools Development and Adaptive Controls

☎ +49 351 463-32051
✉ felix.conrad@tu-dresden.de