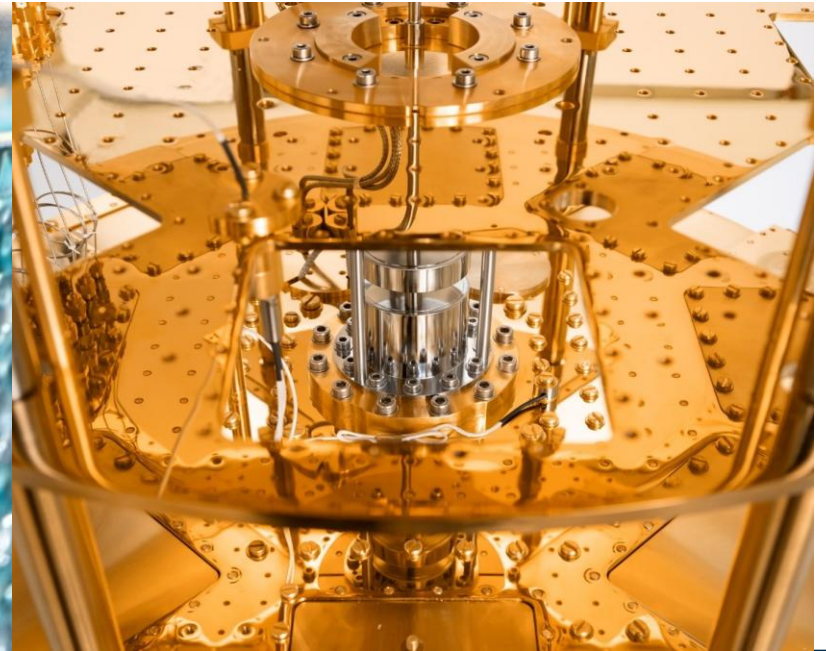
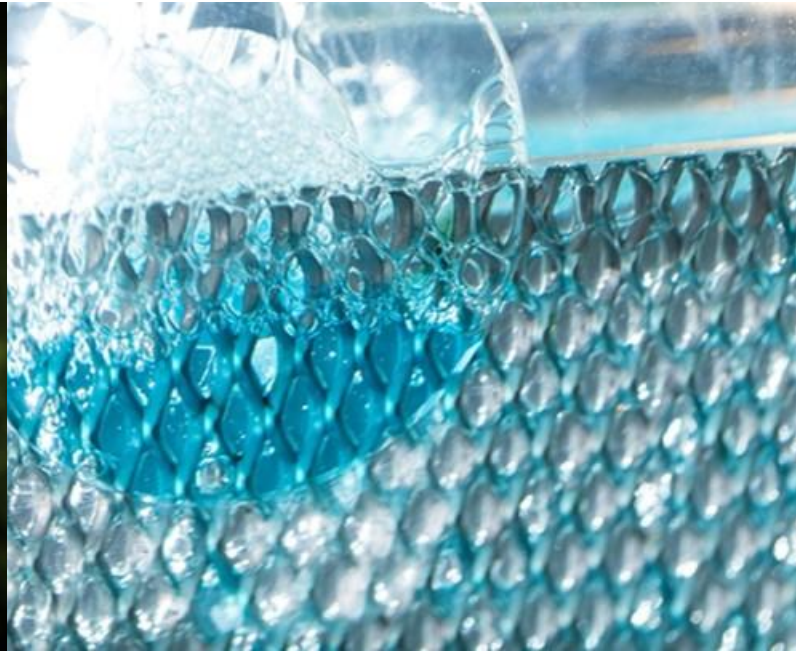


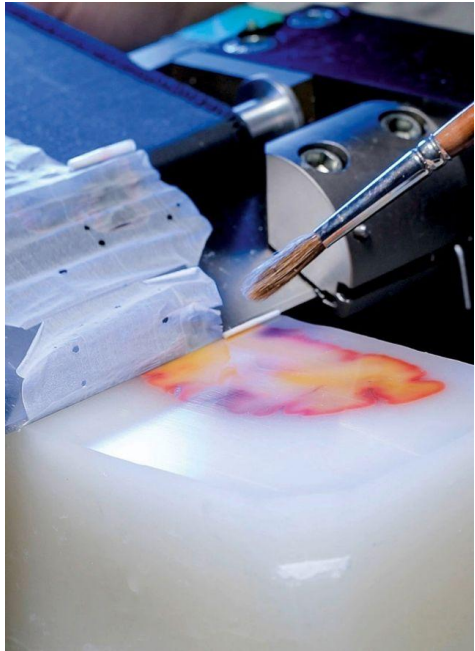


# RESEARCH DATA MANAGEMENT

## A PRACTICAL INTRODUCTION

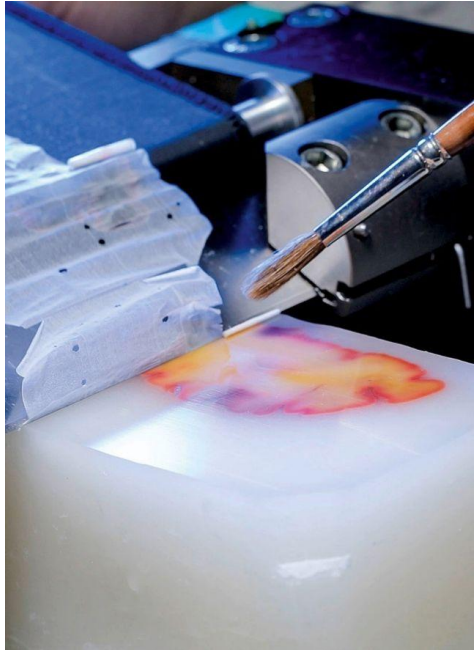
OCTOBER 2026 | H. SCHLENZ, T. BRONGER, M. SELZER, B. NESTLER, L. RIEM, S. ENAHORO

# FORSCHUNGSZENTRUM JÜLICH AT A GLANCE



Member of the  
Helmholtz-  
Community

# FORSCHUNGSZENTRUM JÜLICH AT A GLANCE

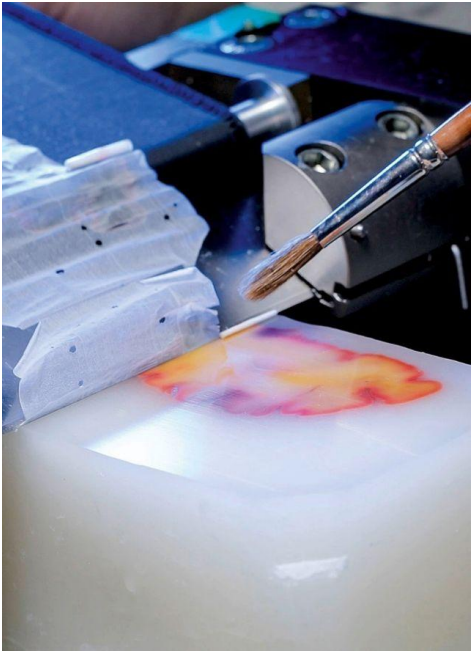


About  
7700  
Employees

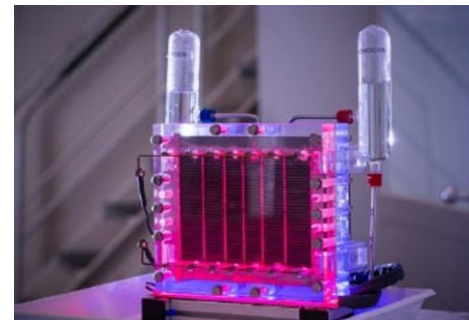
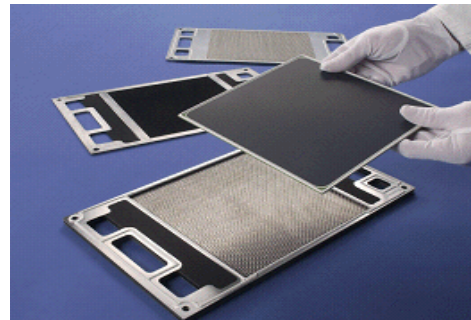
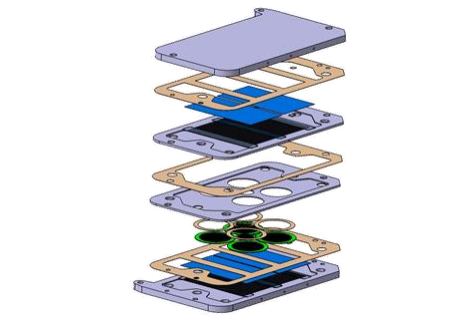
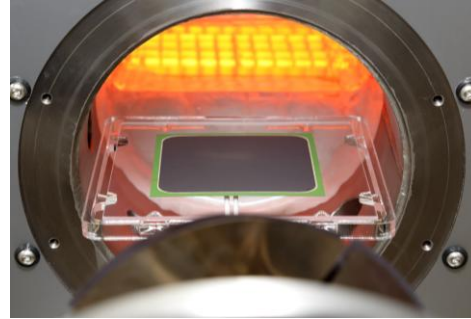
# FORSCHUNGSZENTRUM JÜLICH AUF EINEN BLICK



**Research Topics:  
Information, Energy and  
Bioeconomy**



# RESEARCH AREA ENERGY



# IT-INFRASTRUCTURE



## Exascale-Supercomputer and Quantum-Computer

# INSTITUTE OF ENERGY MATERIALS AND DEVICES (IMD-2)



## Research Topics:

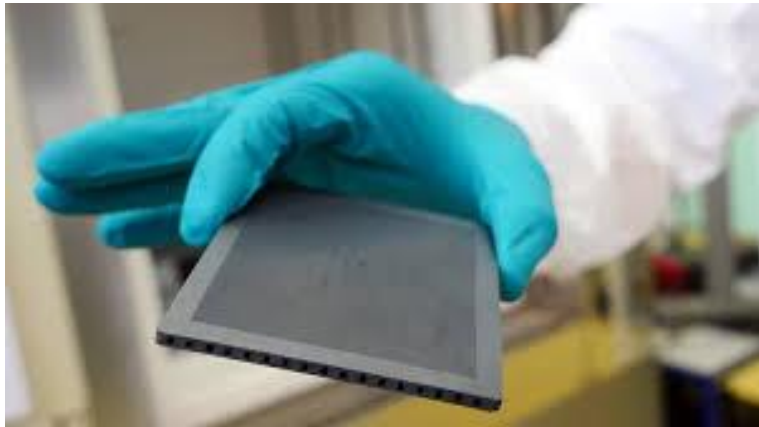
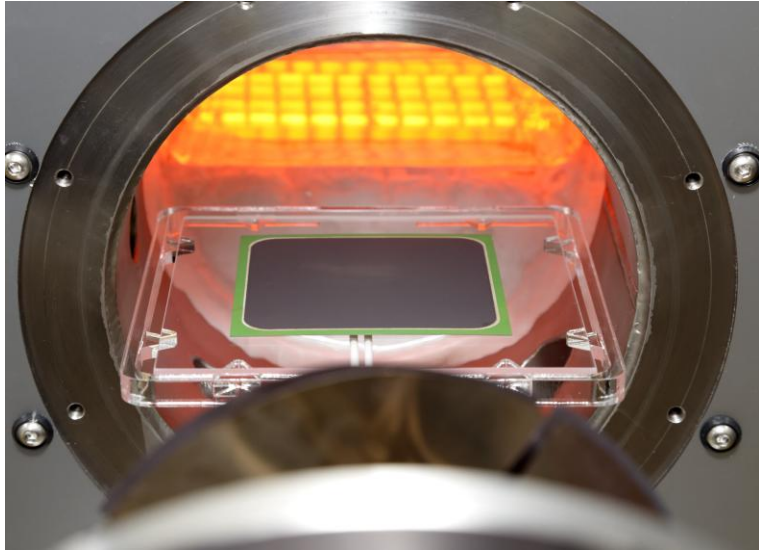
1. Electrochemical storage devices (batteries)
2. Solid-oxide fuel cells and electrolysis cells
3. Gas separation membranes
4. Materials for high-temperature technologies

## Materials:

Ceramics, metals, alloys, etc.



# INSTITUTE OF ENERGY MATERIALS AND DEVICES (IMD-2)



# NATIONAL RESEARCH DATA INFRASTRUCTURE



Nationale  
Forschungsdaten  
Infrastruktur

**NFDi4ing**



Members of NFDI4ING since 2018  
Funding by the DFG since 2020  
Archetype CADEN

# CONTENT

- The needs of users in the academic community
- The electronic laboratory notebook as a key tool
- RDM support at FZ Jülich: RDM Toolbox, training courses, funding
- Best Practice Handbook for RDM (Zenodo)
- RDM training
- Data management plan (DMP): Theory and practical application
- Hard- and Software for data sampling (interfaces)
- The comprehensive collection of data and metadata
- How ELNs can chat with one another
- Data flows and communication in autonomous laboratories
- Providing data for machine learning and other AI applications
- Summary

# TYPICAL USER QUESTIONS

Even after 8 years of the NFDI, I still frequently hear the following comments:

## LEVEL 1:

- What is research data management (RDM) and why do I need it?
- What are the benefits of RDM for me?
- That's just extra work!
- Why should I bother with software monsters (quote!) like an ELN?
- My data belongs to me! (I don't want to share it at all.)
- That's too complicated for me.

## LEVEL 2:

- What are Metadata, how do I get them, and what can I do with them?
- What is Provenance-Tracking?
- What is an Ontology?
- What is semantic RDM?
- What is a Knowledge Graph?

# USER EXPECTATIONS TO PRACTICAL RDM



# SURVEYS AT FZ JÜLICH, HGF, AND ALL OVER GERMANY

1. Time savings
2. Documentation
3. Metadata
4. Data analysis
5. Reproducibility
6. Findability
7. Data security
8. Location-independent
9. Automation
10. Data exchange / Export

This is already largely in line with the **FAIR principles**:  
**F**indability, **A**ccessibility, **I**nteroperability, and **R**eusability

# LIFE CYCLE OF RESEARCH DATA

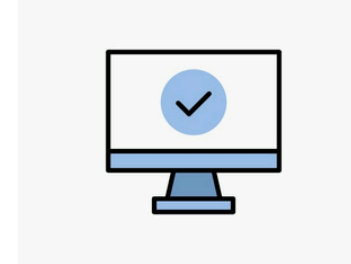


## ELN as a central tool

Today, almost every stage of the data cycle can be covered by ELNs:

- Planning (projects, teams, labs)
- Data and metadata collection
- Documentation
- Data analysis, charts
- AI applications (including intrinsic data analysis)
- Publication (direct transfer to other platforms, e.g. Zenodo)
- Archiving (in-house databases)
- Reuse (including AI modules)

# RDM-TOOLS AT FZ JÜLICH



ACCESS



DATA AND LAB MANAGEMENT



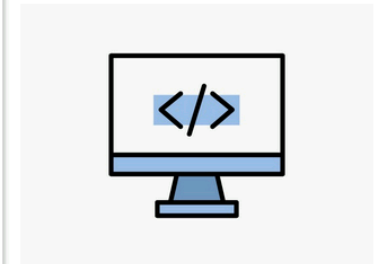
DATA DISCOVERY



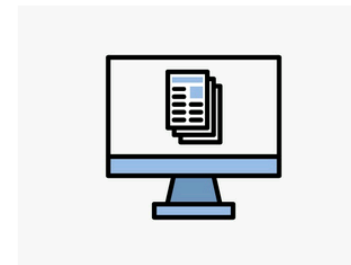
DATA MANAGEMENT PLANS



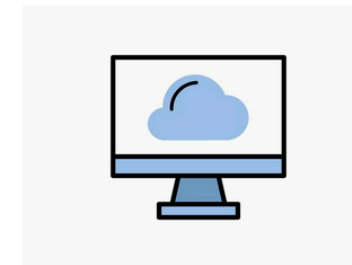
DATA SHARING



METADATA MANAGEMENT



PUBLICATION



STORAGE



VERSION CONTROL

# RDM-TOOLS AT FZ JÜLICH



## METADATA MANAGEMENT

### Metadata Management



B2SHARE



DirSchema



FAIR Python Cookiecutter



Forgejo-aneksajo

HELIPORT

Heliport



JülichDATA



Metador



odMLtables



SampleDB

Somesy

# RDM-TOOLS AT FZ JÜLICH



## METADATA MANAGEMENT

### Metadata Management



B2SHARE



DirSchema



FAIR Python Cookiecutter



Forgejo-aneksajo



A directory structure and metadata linter based on JSON Schema.

JSON Schema is great for validating (files containing) JSON objects that e.g. contain metadata, but these are only the smallest pieces in the organization of a whole directory structure, e.g. of some dataset of project. When working on datasets of a certain kind, they might contain various types of data, each different file requiring different accompanying metadata, based on its file type and/or location.

**DirSchema** combines JSON Schemas and regexes into a solution to enforce structural dependencies and metadata requirements in directories and directory-like archives. With it you can for example check that:

- only files of a certain type are in a location (e.g. only `jpg` files in directory `img` )
- for each data file there exists a metadata file (e.g. `test.jpg` has `test.jpg_meta.json` )
- each metadata file is valid according to some JSON Schema

If validating these kinds of constraints looks appealing to you, this tool is for you!

#### Dirschema features:

- Built-in support for schemas and metadata stored as JSON or YAML
- Built-in support for checking contents of ZIP and HDF5 archives
- Extensible validation interface for advanced needs beyond JSON Schema
- Both a Python library and a CLI tool to perform the validation

# BEST-PRACTICE-HANDBOOK: NEW VERSION 1.1

Constructing tomorrow with data

 **NFDI4ING**  
Research data management  
in engineering sciences

 **Research Data Management**  
**A practical introduction**

H. Schlenz<sup>1</sup>, T. Bronger<sup>1</sup>, M. Selzer<sup>2</sup>, B. Nestler<sup>2</sup>, S. Enahoro<sup>1</sup>, L. Riem<sup>1</sup>  
1. FZJ Forschungszentrum Jülich GmbH, Germany  
2. KIT Karlsruhe Institute of Technology, Germany

**Abstract**

The motivation for this **Best Practice Manual** for research data management resulted from many discussions and interviews that we conducted with users from the engineering and natural sciences communities. It became clear that a **practice-oriented guide** for real research data management is missing. The available, more theoretically oriented presentations are generally only of limited help in practical application. For example, they do not explain how to specifically write a data management plan and what third-party funding bodies look out for. Even for the practical operation of an electronic laboratory notebook, users are almost exclusively dependent on the manuals of the respective system, which represents a major hurdle for many users. Legal issues are also often a problem, for example when it comes to granting licenses or using external data.

We have therefore decided to write this **Best Practice Manual for RDM** in order to close these gaps and provide free practical help for scientists. We go into detail about how to write a **DMP**, explain the **legal aspects of RDM** that need to be taken into account in everyday life, show examples of how to use and operate lab notebooks like **JuliaBase**, **eLabFTW** and **Kadi4Mat**, and discuss aspects of **data quality and analysis**. We show how data can be exchanged and tracked between different ELNs and we demonstrate the **publication of data**. We also explain how research data can be used and **reused for AI** and what requirements must be met for this. Finally, we explain secure ways of permanently **storing data** and provide helpful tips and references for finding further information.

**Table of Contents**

1. Introduction
2. The life cycle of research data
3. Legal aspects
4. The data management plan
5. Data collection, data storage and documentation
6. Data quality
7. Data exchange and data tracking
8. Data publication
9. Permanent data storage

Appendix  
Literature  
Index

**Book Publications**

103 pages, freely available at Zenodo (PDF)  
German: <https://doi.org/10.5281/zenodo.16870414>  
English: <https://doi.org/10.5281/zenodo.16870779>

  This work is licensed under a Creative Commons Attribution 4.0 International License.

**ORCID-Numbers**

H. Schlenz: 0000-0002-2991-4846  
T. Bronger: 0000-0002-5174-4484  
M. Selzer: 0000-0002-9758-646X  
B. Nestler: 0000-0002-9768-5277

**Applicant & co-applicant institutions**



## Content:

1. Introduction
2. The life cycle of research data
3. Legal aspects
4. The data management plan
5. Data collection, data storage and documentation
6. Data quality
7. Data exchange and data tracking
8. Publication of data
9. Permanent data storage
10. Appendix, Bibliography, Index

First published at the **CoRDI Conference** (2025).

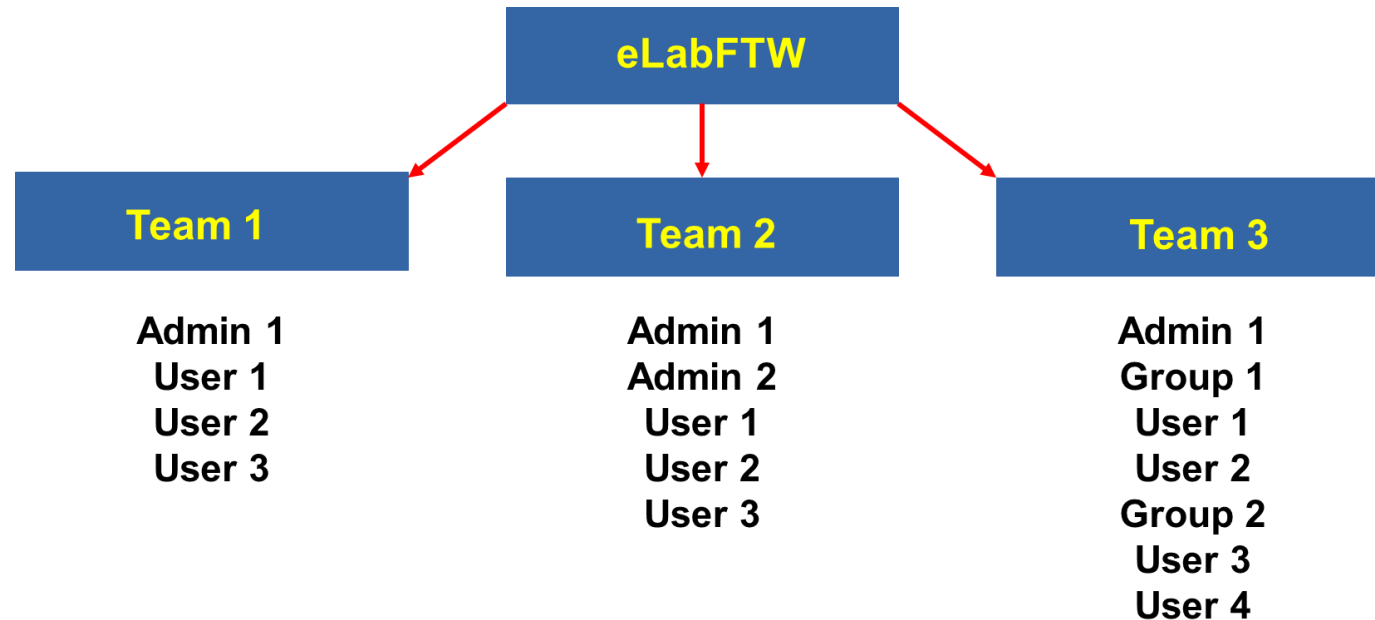
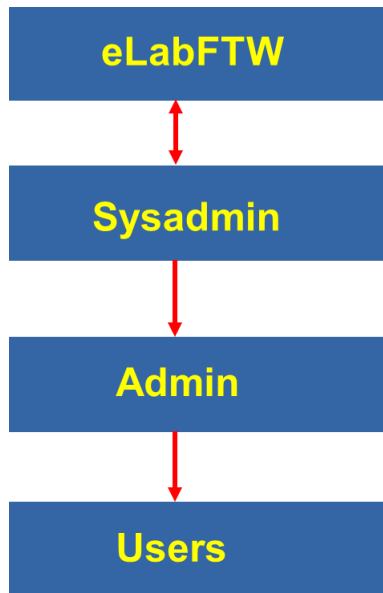
**German:** <https://doi.org/10.5281/zenodo.18468239>  
**English:** <https://doi.org/10.5281/zenodo.18468308>

# ELN-TRAINING: EXAMPLE



eLabFTW

- What is eLabFTW?
- Why should I use eLabFTW?
- How does the structure of eLabFTW look like?
- Which possibilities do I have as a user?
- Which rights do I have as user?
- How can I store my data?
- How can I export my data?
- Which data formats are available?
- How do I protect my data?
- Is it possible to encrypt my data?



# INTERFACE: EXAMPLE eLabFTW

eLabFTW

EXPERIMENTS

RESOURCES

TEAM

SEARCH

DOCUMENTATION

## Empyrean

Create

←























 Owned by Leo Riem

Started on07.08.2024

★★★★★

ID109

Custom ID

Get next

Category

Not set

Status

Running

Tags

empyrean

FZJ

Korund\_PSD\_1D\_test - Kopie

Leo\_Riem

XRD

Yoo Jung Sohn

Add a tag

► Suggested tags

Visibility

+

Everyone with an account

Can write

+

Everyone with an account

# TEMPLATES: EXAMPLE eLabFTW

HAUPTTEXT

Datei Bearbeiten Ansicht Einfügen Format Werkzeuge Tabelle

↶ ↷ Fett 14pt B I U ↺

≡ ≡ ≡ ≡ x² x₂

≡ ≡ ≡ ≡

Ω ⌚ (i) 🔗 ↴ ↵

Sample:

Al<sub>2</sub>O<sub>3</sub> (corundum standard)

Additional information to the CIF-File (optional):

|                                 |                     |
|---------------------------------|---------------------|
| _diffrn_radiation_type          | 'Cu Kα~1+2~'        |
| _diffrn_source                  | 'sealed X-ray tube' |
| _diffrn_source_target           | Cu                  |
| _diffrn_measurement_dvice_type  | 'Scintag XDS2000'   |
| _diffrn_detector                | 'Ge solid-state'    |
| _diffrn_radiation_monochromator | 'graphite'          |
| _pd_meas_number_of_points       | 3667                |
| _pd_meas_counts_total           | 0                   |
| _pd_meas_counts_background      | 0                   |
| _pd_meas_counts_container       | 0                   |
| _pd_meas_counts_monitor         | 0                   |
| _pd_meas_rocking_angle          | 0:360.0             |
| _pd_instr_monochr_post_spec     | 'graphite (0001)'   |

ZUSÄTZLICHE FELDER

Powder X-ray diffraction experiment

Heute 21.02.2024

\_diffrn\_ambient\_temperature295 K

\_diffrn\_radiation\_wavelength1.540598 Å

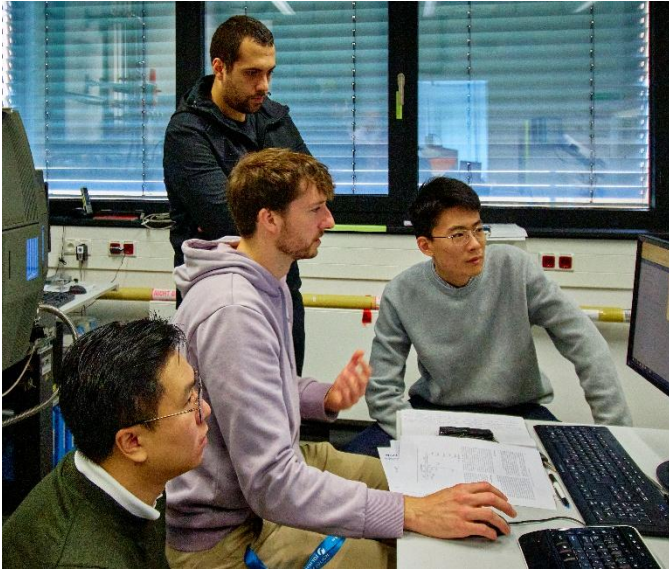
\_pd\_meas\_info\_author\_nameHartmut Schlenz

\_pd\_meas\_info\_author\_mailh.schlenz@fz-juelich.de

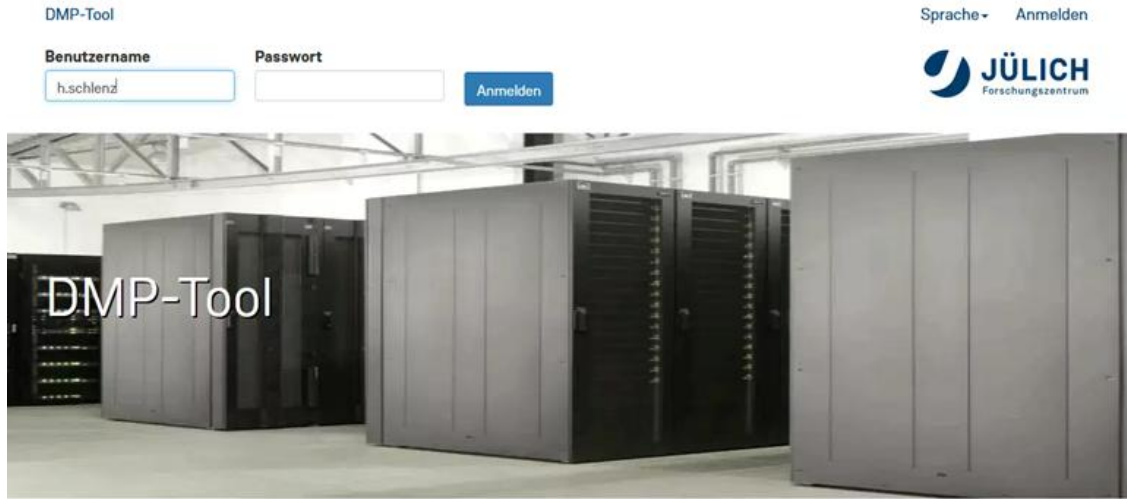
\_pd\_meas\_info\_author\_phone+924616196956

\_pd\_meas\_datetime\_initiated

# RDM TRAINING: SUMMER SCHOOL AT IMD-2 (13-15 OCT'26)



# DATA MANAGEMENT PLAN



The image shows a screenshot of the DMP-Tool login interface at the top, featuring fields for 'Benutzername' (username) and 'Passwort' (password), with a 'Sprache' dropdown and an 'Anmelden' button. Below the login form is a photograph of a server room with several large black server racks. The text 'DMP-Tool' is overlaid on the left side of the server room image. Below the image is a welcome message in German, followed by detailed instructions about the purpose and use of the DMP tool, and a footer note stating it is based on the RDMO of the AIPs Potsdam.

DMP-Tool

Sprache Anmelden

JÜLICH  
Forschungszentrum

Benutzername Passwort Anmelden

h.schlentz

DMP-Tool

Willkommen im DMP-Werkzeug des FZ Jülich

Hier können Forschende des Forschungszentrums Datenmanagementpläne (DMPs) erstellen und verwalten.

Ein DMP begleitet ein Forschungsvorhaben. In ihm können die am Vorhaben beteiligten Personen für sich und andere dokumentieren, wie mit den Daten im Vorhaben umgegangen wird. Darüber hinaus geht der DMP auf die Urheber-, Nutzungs- und Zugangsrechte sowie die Verantwortlichkeiten während und nach Ende des Forschungsvorhabens ein.

Bei Projektantragsstellung ist ein DMP ein hilfreiches – und zunehmend verlangtes – Dokument, das die Chancen auf Bewilligung erhöht.

Ein DMP besteht aus einem Fragenkatalog, den die Forschenden ausfüllen. Er ist ein lebendes Dokument, d.h. er wird während eines Forschungsvorhabens stets an die aktuelle Situation angepasst.

Jedes Forschungsprojekt sollte einen eigenen DMP bekommen.

Basierend auf RDMO des AIPs Potsdam.

## RDMO:

*Research Data Management Organiser:*  
Server installation and administration required. Incurs additional costs.  
<https://rdmorganiser.github.io/>

## Simple alternative:

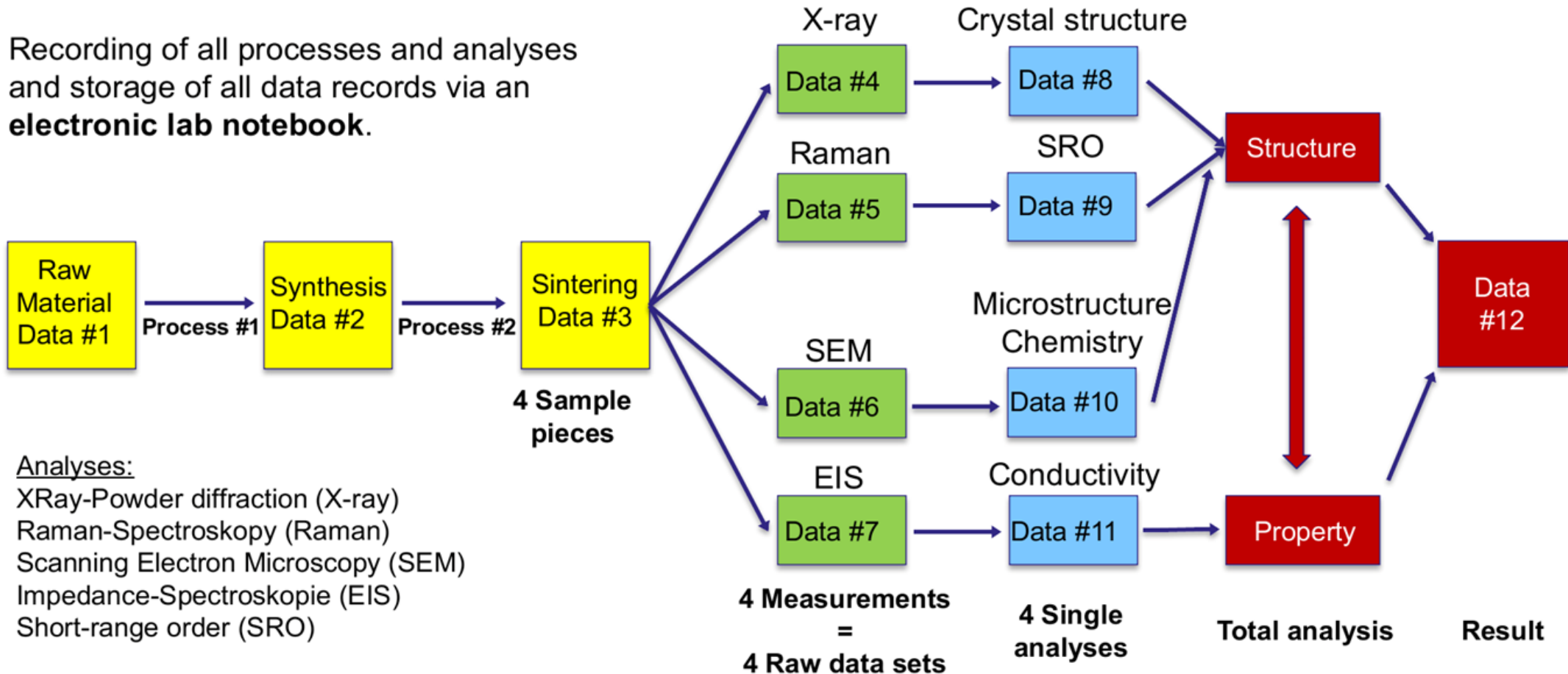
Text document, in accordance with the DFG guidelines.

## New at FZ Jülich (under development):

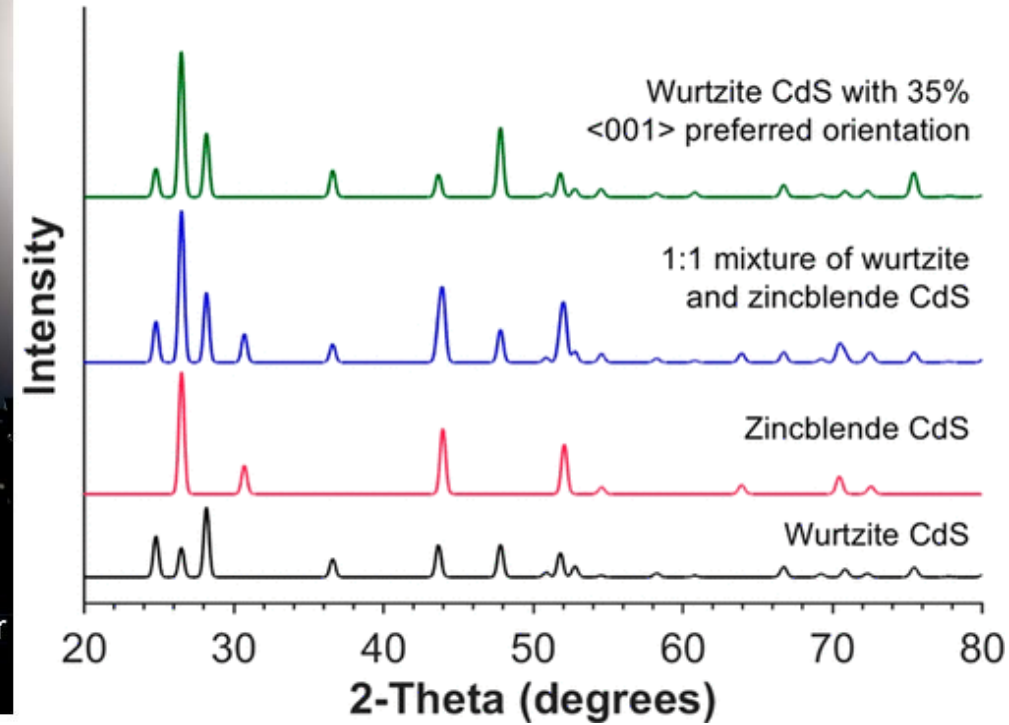
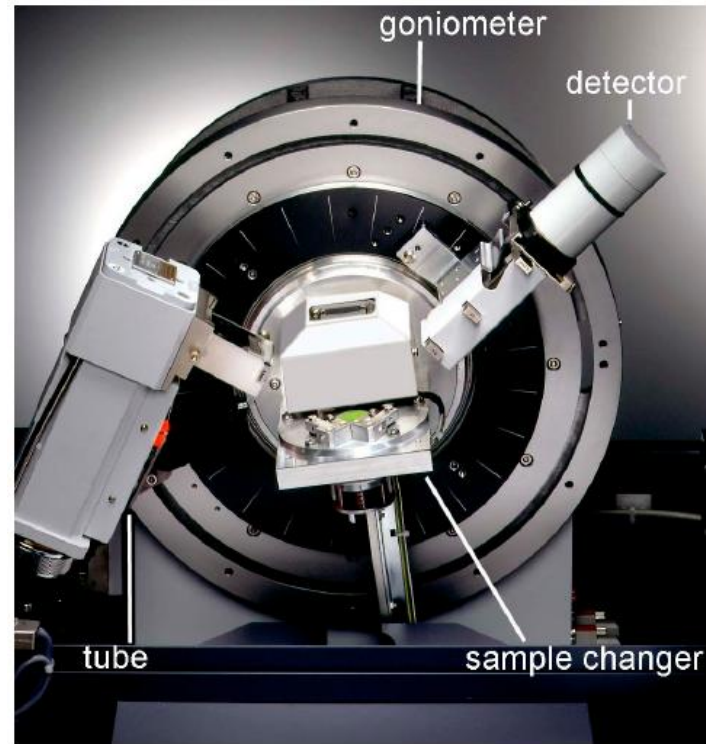
A Research Data Planning Workbook.  
A quite simple step-by-step guide.  
(S. Rank, D. Jerger, A. Sprenger (UE))

# TYPICAL EXPERIMENTS AT IMD-2

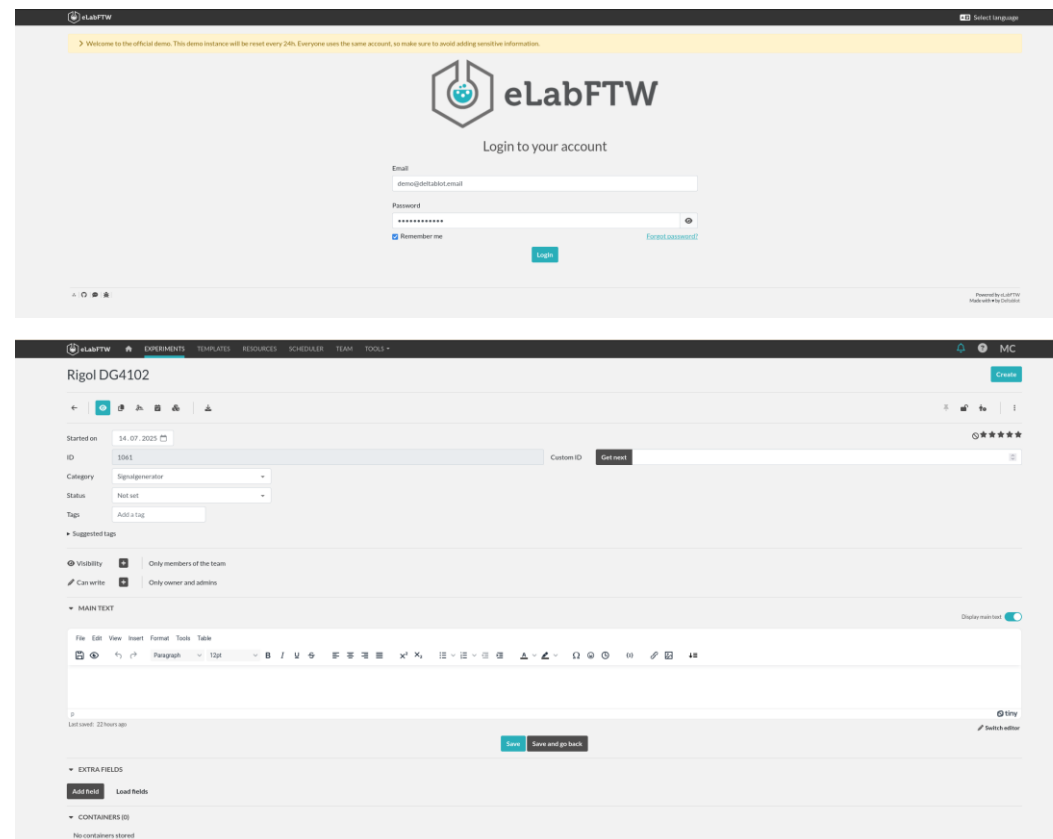
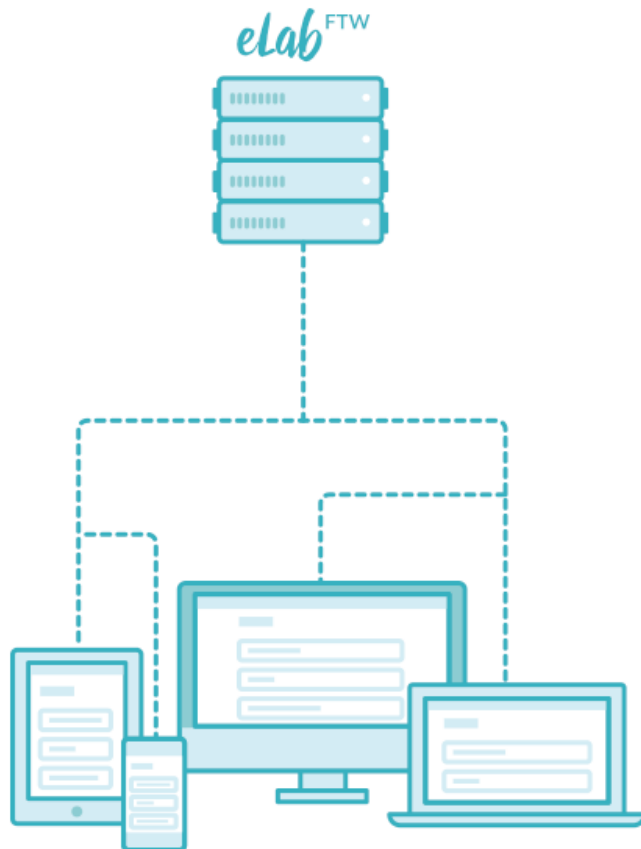
Recording of all processes and analyses and storage of all data records via an **electronic lab notebook**.



# POWDER X-RAY DIFFRACTION



# ELECTRONIC LAB NOTEBOOKS



## Problem

Automatic Transfer  
of the **complete** data  
and metadata  
from all analytic  
devices.

# METADATA IN COMMERCIAL BINARY DATA

## Problem:

- The APIs of many commercial laboratory instruments often allow only partial data to be retrieved, whilst metadata is frequently not available at all!
- This means that FAIR data cannot be generated, and automation is prevented.
- ELNs are often unable to import measurement data automatically.
- Interface programming (API) and the use of templates in ELNs are made considerably more difficult.

## Possible solutions:

- Manufacturers are collaborating and making their APIs available, or will produce new devices with open APIs in future.
- Use of the FZJ software **MARBLE** (S. Brinkmann, V. Hofmann, FZJ).
- **Legal position:** Researchers own all the data and metadata they generate!

<https://jugit.fz-juelich.de/marble>



# METADATA IN COMMERCIAL BINARY DATA

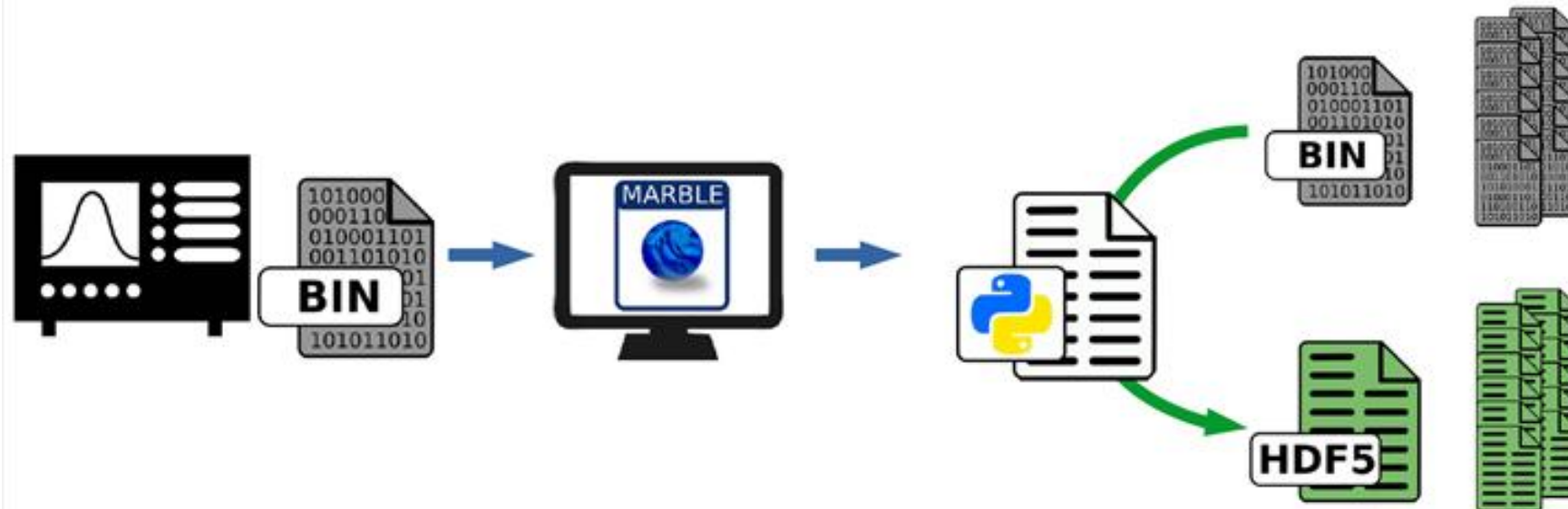


Measurement data is only partially exported. Incomplete metadata.  
Automation is prevented.

Users identifies  
data & metadata

Python  
translator

Translates files  
into FAIR format



Software decrypts  
binary files.  
NO Reverse-Engineering!

# METADATA IN COMMERCIAL BINARY DATA

GitLab: **MARBLE**

[jugit.fz-juelich.de/marble](https://jugit.fz-juelich.de/marble)

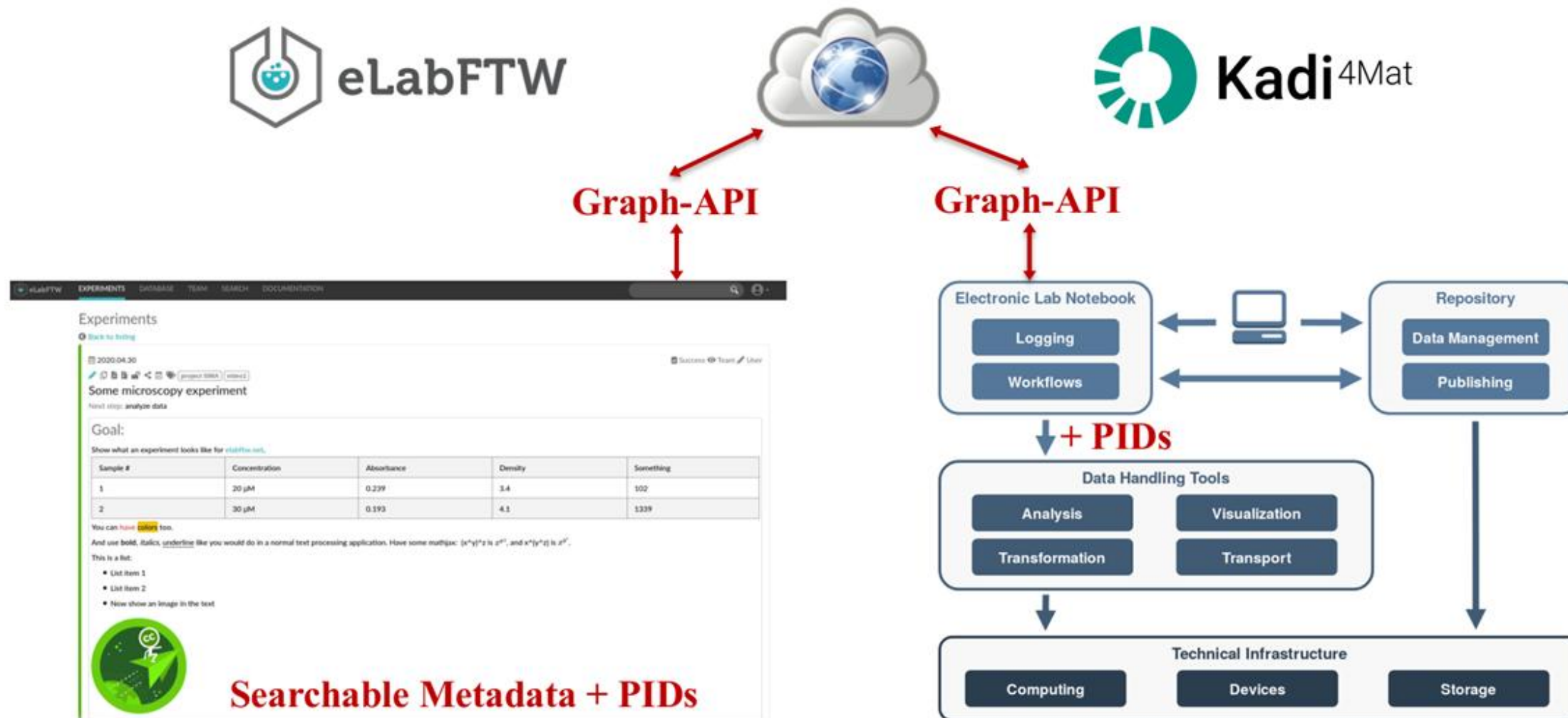
Legal rulings / legislation::

1. OLG Naumburg (27.08.2014 – 6U 3/14)  
<https://openjur.de/u/768245.html>
2. Geschäftsgeheimnisse §3 GeschGehG (2019)
3. Legal departments of FZ Jülich (Helmholtz)
4. Legal departments of the Max-Planck-Society



Software decrypts  
binary files.  
NO Reverse-Engineering!

# HOW ELN'S CAN CHAT WITH EACH OTHER



# HOW ELN'S CAN CHAT WITH EACH OTHER

## SciMesh at a glance

### Introduction

SciMesh represents scientific insights as a knowledge graph. In its current realisation, it focuses on sample-based workflows, in which samples (physical ones or data artefacts) undergo a sequence of processing and measurement steps. However, it is not limited to that. Most generally, it represents scientific insight by declaring a relationship between cause and effect. In other words, if certain prerequisites are set, then certain observations will be made.

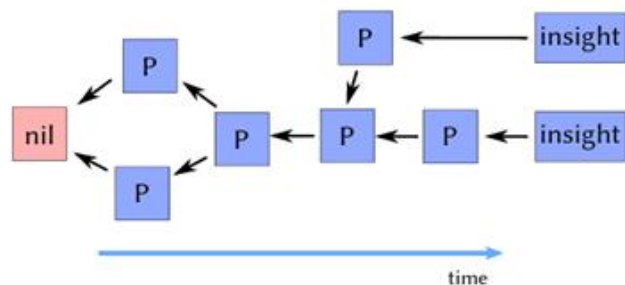


Fig. 1 Example graph of processes.

Fig. 1 shows a graph illustrating this line of thought. Starting from an initial state called "nil", processes change this state. There is a monotonic increase in time from left to right. Consequently, this time axis is also inherent in a chain of processes. A process may be "take glass substrate out of rack", "heat sample", "mix two substances together", or "wait for solar eclipse". It is really that general. Every process has got one or more parents. If it is a single one, it may be the initial state. The initial state is totally void. It bears no information at all, which is why the chain of processes must define the states in-between as completely as necessary to be useful for scientific conclusions. Conversely, one process may be the parent of multiple others. This way, chains can be branched off, possibly by different scientists years after the work on the main trunk.

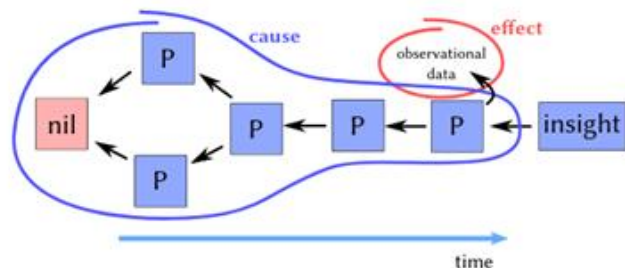


Fig. 2 Example graph of processes, showing cause and effect.

### Workflow

#### Institute A with ELN A

create sample #1  
with URI `http://A/samples/1`

do things with sample #1 and  
record them in ELN A

send sample physically to Institute B

add URL of sample #1 in ELN B  
to ELN A

ELN A responds with the SciMesh graph  
for sample #1

open the data sheet for sample #1

ELN A makes an HTTP GET request  
against the URL, requesting  
an RDF content-type

ELN A shows the data from ELN A  
and ELN B together in one data sheet

#### Institute B with ELN B

add sample with URI `http://A/samples/1`  
to ELN B

send URL of sample #1 in ELN B to Institut A

open the data sheet for sample #1

ELN B makes an HTTP GET request  
against the URL to that URI, requesting  
an RDF content-type

ELN B shows the data from ELN A for sample #1

do things with sample #1 and  
record them in ELN B

ELN B responds with the SciMesh graph  
for the activities with sample #1 at Institute B

open the data sheet for sample #1

ELN B shows the data from ELN A  
and ELN B together in one data sheet

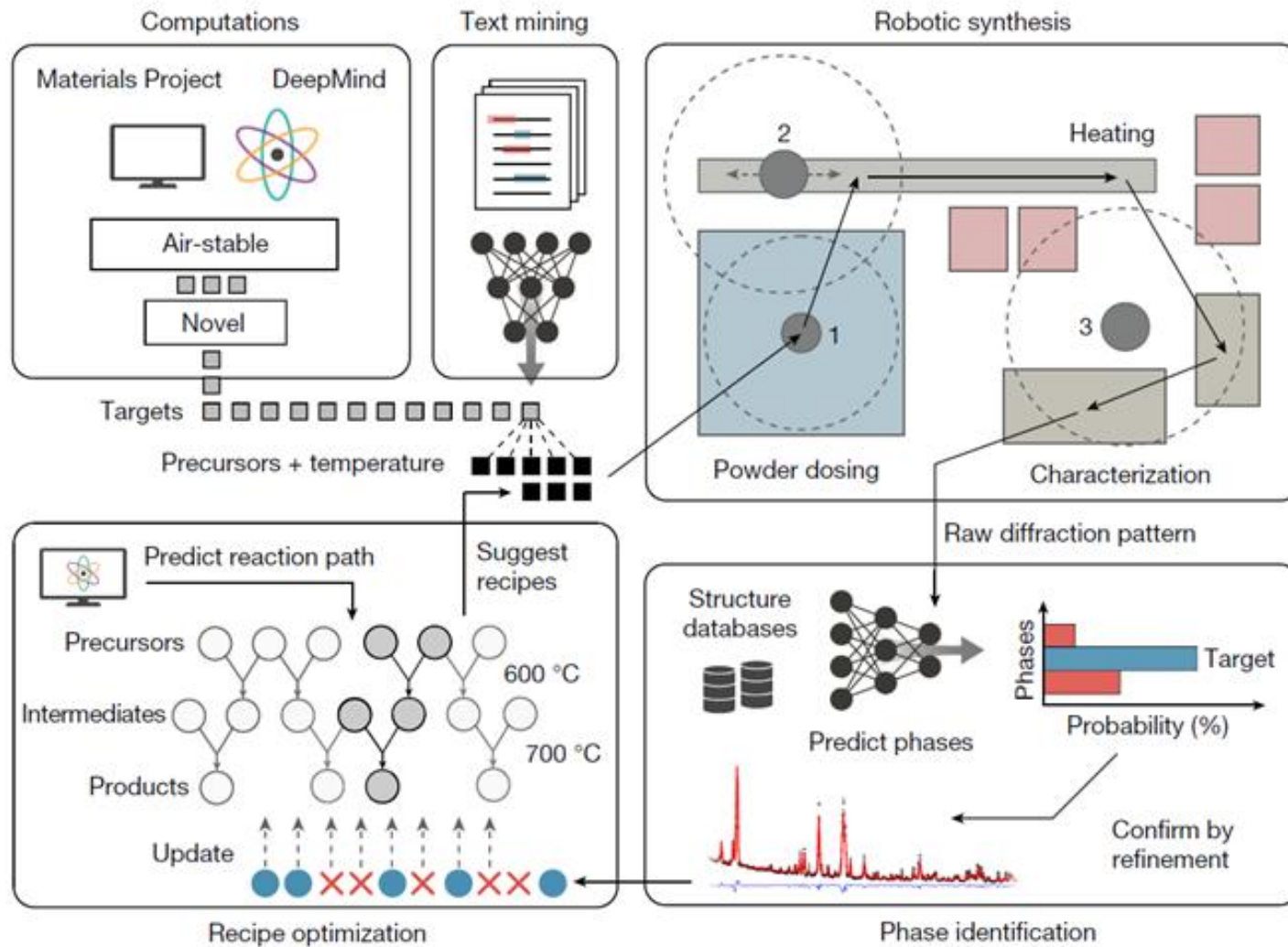


Fig. 7 Possible workflow for two institutes working together with the same samples.

T. Bronger  
M. Flemming  
H. Schlenz  
M. Selzer  
M. Jayavarapu

<https://scimesh.org>

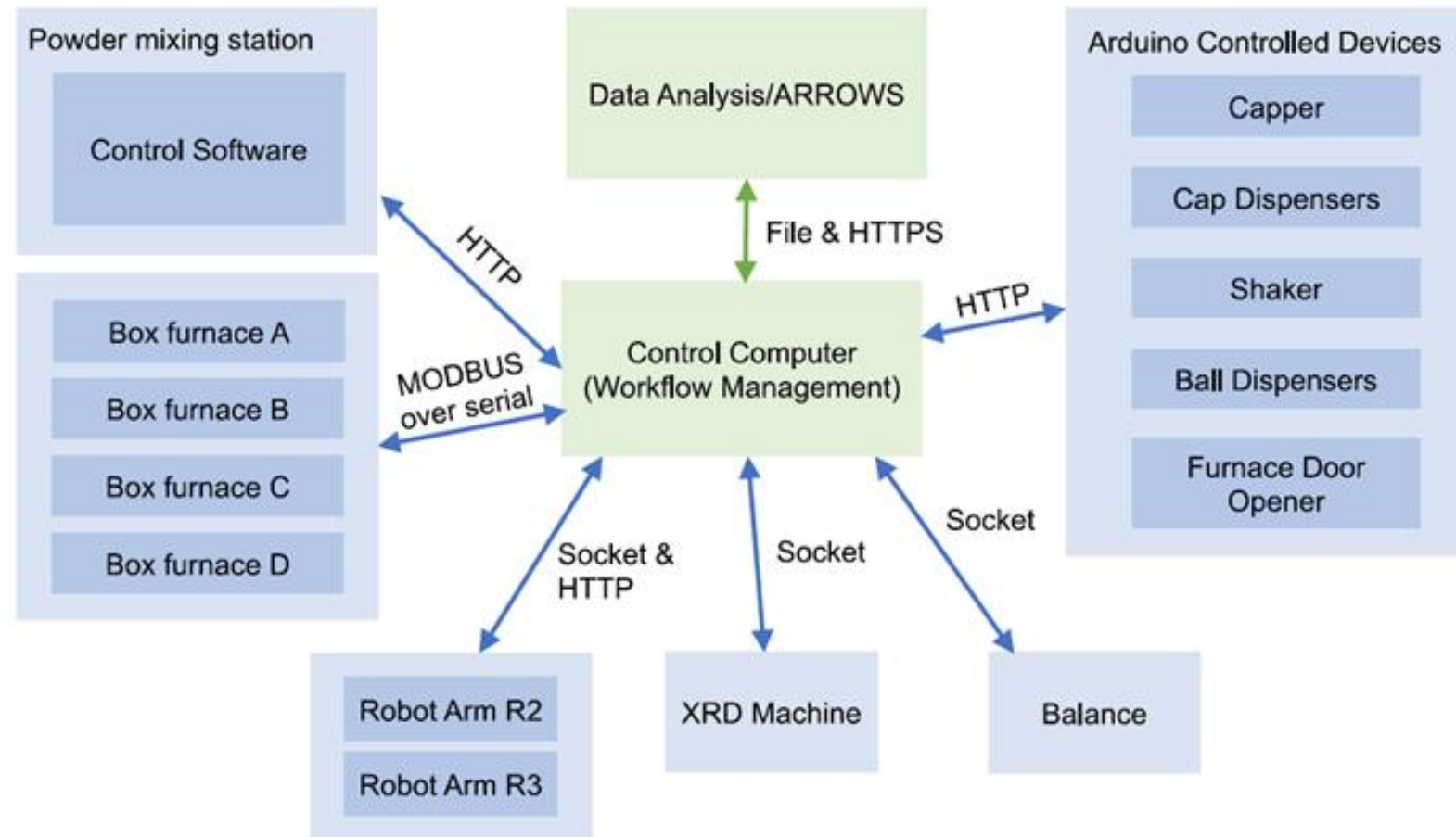
# DATA STREAMS IN SELF-DRIVEN LABS



The A-Lab

N.J. Szymanski et al., An autonomous laboratory for the accelerated synthesis of novel materials, Nature, Vol 624 (2023) 86.

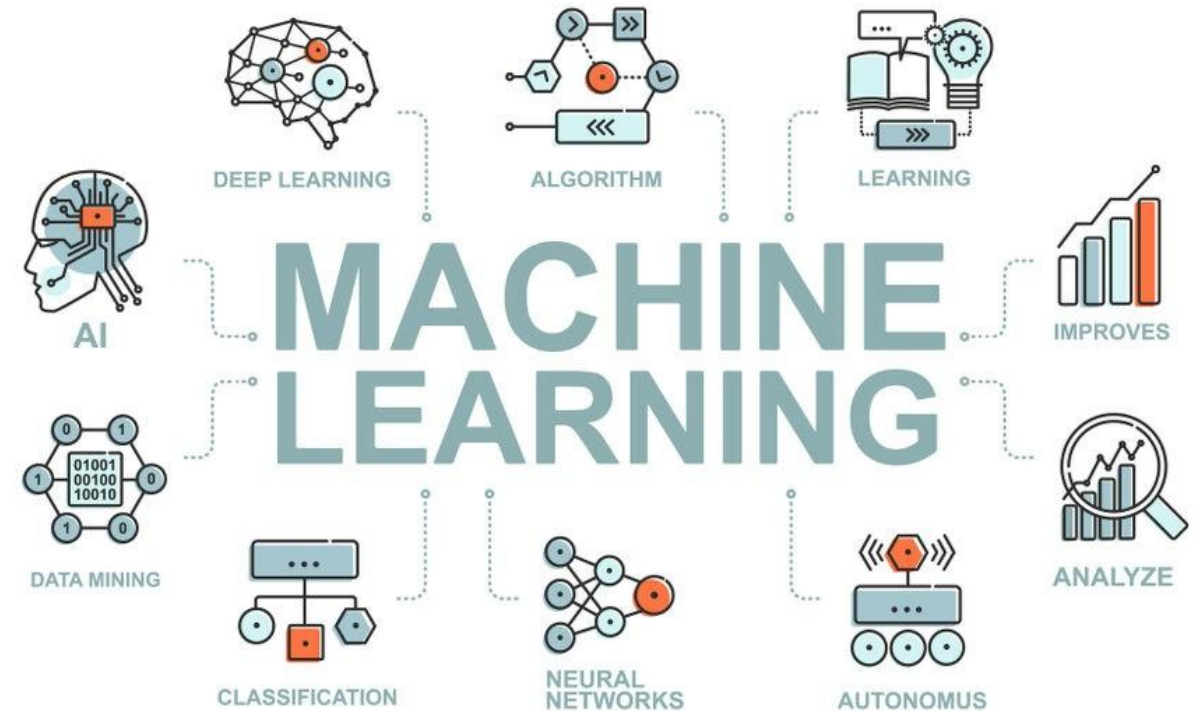
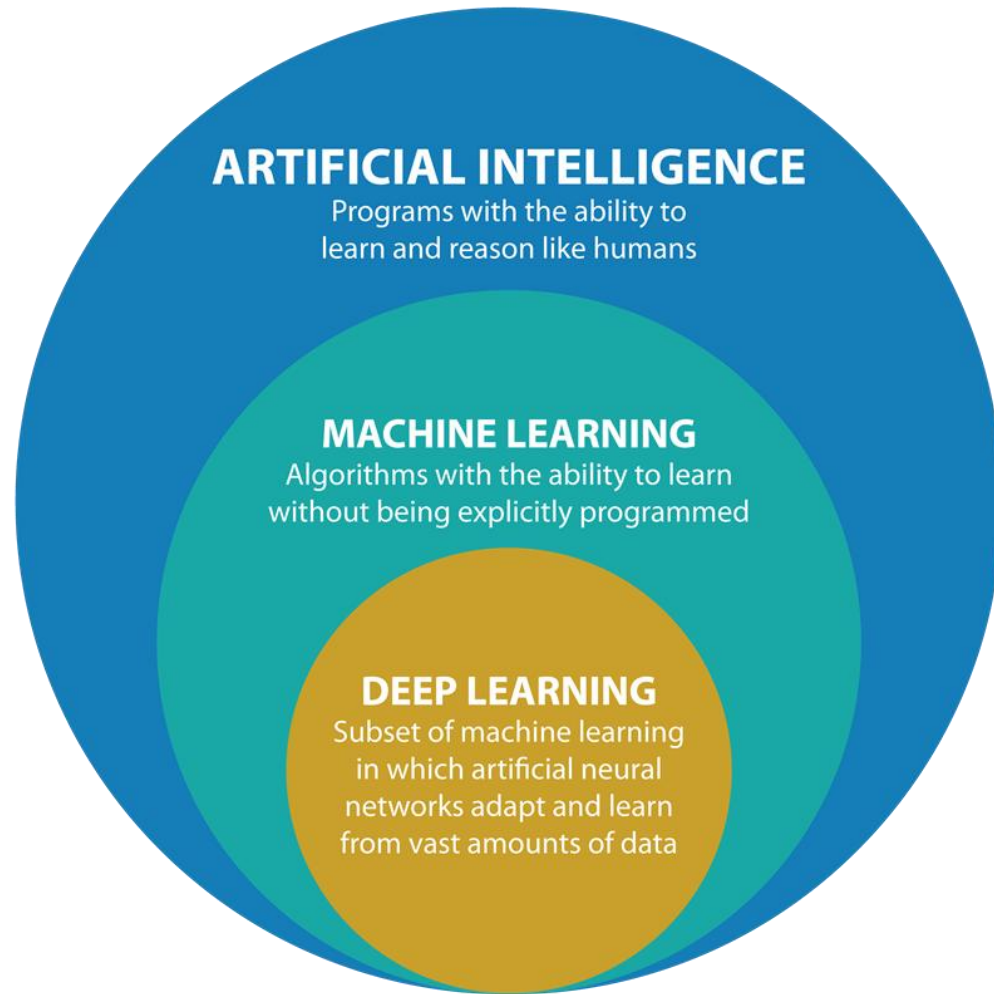
# COMMUNICATION WITHIN THE A-LAB



**Extended Data Fig. 3 | Communication protocols connecting each module in the A-Lab.** A local area network (LAN) is built to connect all the pieces of the A-Lab with a control computer using an RS-485 interface (or DB25 for the box furnaces). Each module on the RS-485 interface has an Internet Protocol (IP)

assigned to enable communication with the computer. For enhanced cybersecurity, only the control computer has access to the internet, whereas the LAN is isolated from it.

# PROVISION OF DATA + METADATA FOR ML AND AI



# PROVISION OF DATA + METADATA FOR ML AND AI

1. For what type of AI application should data be provided?
  - i. Machine learning (ML), deep learning (DL), language model (LLM), ...
    - a. Structured or unstructured data?
2. Which data types/-formats shall be used?
  - i. ASCII / CSV, JSON, HDF5, ...
  - ii. Grafics, Photos, Audio-Files, Videos (Live-Stream)
  - iii. PDF (Literature), and other text formats
3. Which data sources are available?
  - i. Measuring instruments (Control computer), Cameras
  - ii. ELN's
  - iii. Data bases
  - iv. Knowledgegraphs
  - v. Internet

# PROVISION OF DATA + METADATA FOR ML AND AI



Tables (2D)

|    | x  | y1    | y2   | y3    | x4 | y4   |
|----|----|-------|------|-------|----|------|
| 1  | 10 | 8,04  | 9,14 | 7,46  | 8  | 6,58 |
| 2  | 8  | 6,95  | 8,14 | 6,77  | 8  | 5,76 |
| 3  | 13 | 7,58  | 8,74 | 12,74 | 8  | 7,71 |
| 4  | 9  | 8,81  | 8,77 | 7,11  | 8  | 8,84 |
| 5  | 11 | 8,33  | 9,26 | 7,81  | 8  | 8,47 |
| 6  | 14 | 9,96  | 8,1  | 8,84  | 8  | 7,04 |
| 7  | 6  | 7,24  | 6,13 | 6,08  | 8  | 5,25 |
| 8  | 4  | 4,26  | 3,1  | 5,39  | 19 | 12,5 |
| 9  | 12 | 10,84 | 9,13 | 8,15  | 8  | 5,56 |
| 10 | 7  | 4,82  | 7,26 | 6,42  | 8  | 7,91 |
| 11 | 5  | 5,68  | 4,74 | 5,73  | 8  | 6,89 |

JSON-Code

```
{
  "Herausgeber": "Xema",
  "Nummer": "1234-5678-9012-3456",
  "Deckung": 2e+6,
  "Waehrung": "EURO",
  "Inhaber":
  {
    "Name": "Mustermann",
    "Vorname": "Max",
    "Teilnahme am Bonusprogramm": true,
    "Hobbys": ["Reiten", "Golfen", "Lesen"],
    "Alter": 42,
    "Kinder": [],
    "Partner": null
  }
}
```

Beispiel: HDF5-Objekte abrufen und drucken

```
import hdf.object.*;    // include the common HDF object package
import hdf.object.h5.*; // include the HDF5 object package
import hdf.hdf5lib.*;   // include the Java HDF5 interface
/**
 * Retrieve and print HDF5 objects from file hdf5_test.h5
 */
public class TestH5File
{
    public static void main(String[] argv)
    {
        // create an H5File object
        H5File h5file = new H5File("hdf5_test.h5",
HDF5Constants.H5F_ACC_RDONLY);
        try
        {
            // open file and retrieve the file structure
            h5file.open();
        }
        catch (Exception ex)
        {
            System.out.println(ex);
        }
        javax.swing.tree.MutableTreeNode root = h5file.getRootNode();
        if (root != null)
        {
            printNode(root, "    ");
        }
        try { h5file.close(); }
        catch (Exception ex ) {}
    }
    // print out the data object recursively
    private static void printNode(javax.swing.tree.TreeNode node, String
indent)
    {
        System.out.println(indent+node);

        int n = node.getChildCount();
        for (int i=0; i<n; i++)
        {
            printNode(node.getChildAt(i), indent+"    ");
        }
    }
}
```

# PROVISION OF DATA + METADATA FOR ML AND AI

The screenshot displays the HDFView 3.0 application window. The left sidebar shows a tree view of the file structure, with 'arrays' expanded to show '2D int8x9'. The main panel is divided into three sections: 'General Object Info', 'Dataspace and Datatype', and 'Storage Layout'. The 'General Object Info' section shows the object name '2D int8x9', path '/arrays/', type 'HDF5 Scalar Dataset', and object reference '1784651'. The 'Dataspace and Datatype' section shows 2 dimensions, size 8x9, and data type '32-bit integer'. The 'Storage Layout' section shows 'CONTIGUOUS' layout, no compression, and no fill value. A 'Table' window is overlaid, showing a grid of 8x9 data values ranging from 11 to 85. The bottom status bar shows the current object's path and dimensions.

General Object Info

Name: 2D int8x9  
Path: /arrays/  
Type: HDF5 Scalar Dataset  
Number of Attributes: 0  
Object Ref: 1784651

Dataspace and Datatype

No. of Dimension(s): 2  
Dimension Size(s): 8 x 9  
Max Dimension Size(s): 8 x 9  
Data Type: 32-bit integer

Storage Layout: CONTIGUOUS  
Compression: NONE  
Filters: NONE  
Storage: SIZE: 288, allocation time:  
Fill value: NONE

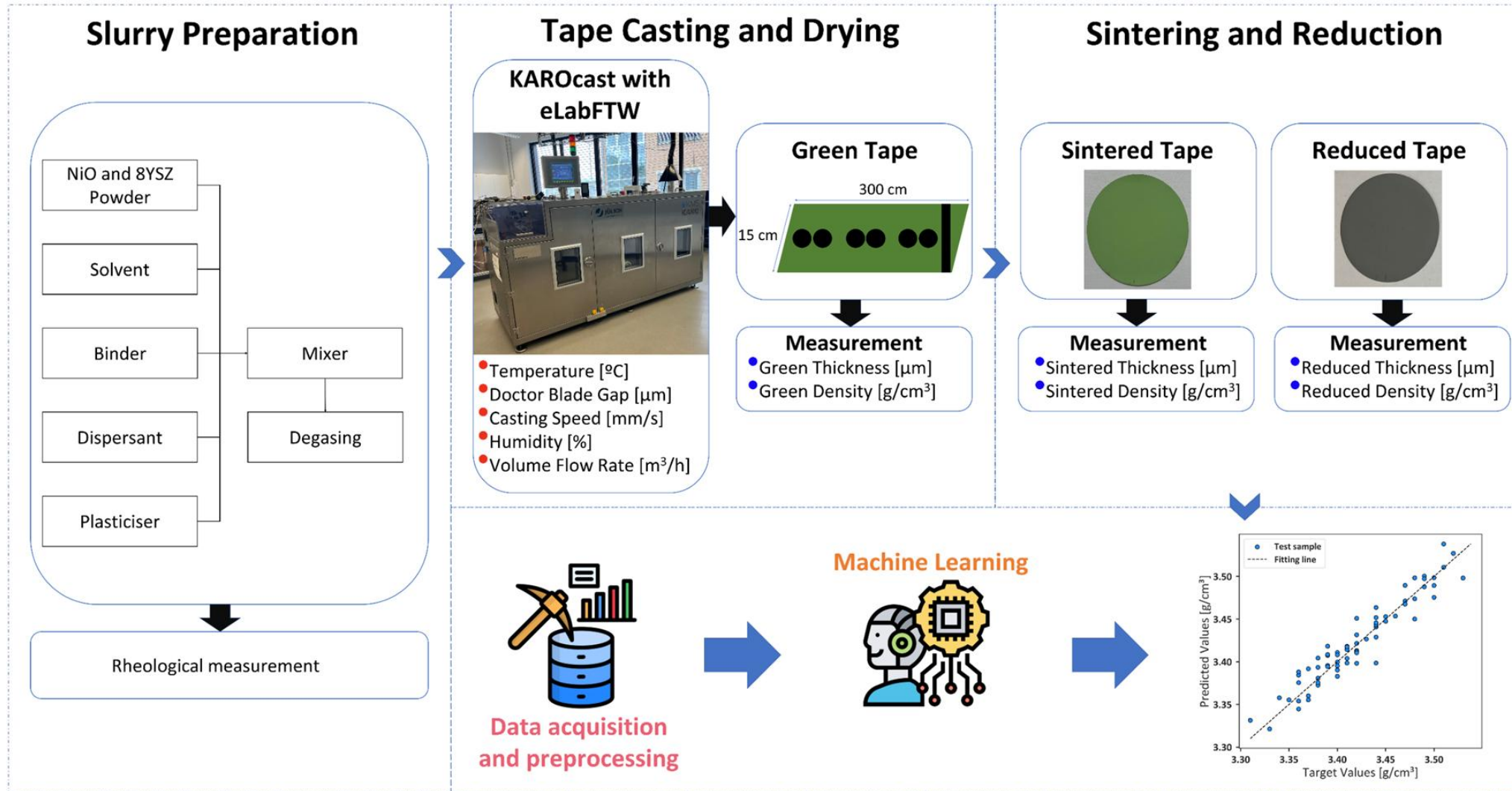
Table

0-based

|   | 0  | 1  | 2  | 3  |    |
|---|----|----|----|----|----|
| 0 | 11 | 12 | 13 | 14 | 15 |
| 1 | 21 | 22 | 23 | 24 | 25 |
| 2 | 31 | 32 | 33 | 34 | 35 |
| 3 | 41 | 42 | 43 | 44 | 45 |
| 4 | 51 | 52 | 53 | 54 | 55 |
| 5 | 61 | 62 | 63 | 64 | 65 |
| 6 | 71 | 72 | 73 | 74 | 75 |
| 7 | 81 | 82 | 83 | 84 | 85 |

- Programming skills (Python, Java, C, etc.) are often required to transfer complex data formats.
- Ideally, ready-made scripts can handle such tasks.
- HDFView as an example of a simple, user-friendly interface for HDF4 or HDF5 data.
- HDF5 is the standard data format used by major research facilities, such as synchrotrons (DESY, BESSY, ESRF) or neutron core facilities, like Garching, ESS (Sweden), and ILL (France).

# REAL EXPERIMENT: TAPE CASTING OF CERAMICS



# REAL EXPERIMENT: TAPE CASTING OF CERAMICS

Energy and AI 24 (2026) 100687



Data-driven machine learning modelling for the manufacturing of the fuel electrode support in solid oxide cells

Tan Le-Dinh<sup>a,b,c</sup>, Hartmut Schlenz<sup>c,d,e,\*</sup>, Norbert H. Menzler<sup>c</sup>, Alejandro A. Franco<sup>a,b,c,f</sup>, Olivier Guillon<sup>c,g</sup>

<sup>a</sup> Laboratoire de Réactivité et de Chimie des Solides, UMR CNRS 7314, Université de Picardie Jules Verne, F-80039, Amiens Cedex 1, France

<sup>b</sup> Réseau sur le Stockage Electrochimique de l'Energie (RS2E), Hub de l'Energie, FR CNRS 3459, 15 rue Baudelocque, 80039 Amiens Cedex, France

<sup>c</sup> Forschungszentrum Jülich GmbH, Institute of Energy Materials and Devices (IMD), IMD-2: Materials Synthesis and Processing, 52428 Jülich, Germany

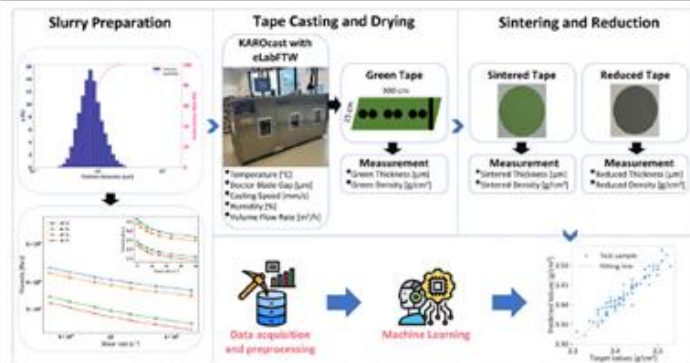
<sup>d</sup> Institute of Geoscience, Division of Geochemistry and Petrology, University of Bonn, Meckenheimer Allee 139, D-53115 Bonn, Germany

<sup>e</sup> ALISTORE-European Research Institute, Hub de l'Energie, FR CNRS 3104, 15 rue Baudelocque, 80039 Amiens Cedex, France

<sup>f</sup> Institut Universitaire de France, 103 boulevard Saint Michel, 75005 Paris, France

<sup>g</sup> Jülich Aachen Research Alliance: JARA-Energy, Jülich 52425, Germany

## GRAPHICAL ABSTRACT



## RDM processing steps:

1. Python programming of the interface (API) between the device and **eLabFTW**
2. Automatic collection of all processing data and metadata with **eLabFTW**
3. Simulations for the characterization of the raw material => Results stored in **eLabFTW**
4. Physical analysis of the samples  
=> Results stored in **eLabFTW**
5. Export of all data and metadata from **eLabFTW** in JSON format
6. Machine learning and deep learning for predicting optimum process parameters  
=> Results stored in **eLabFTW**
7. Optimised process engineering as a result.  
**NO more Trial-and-Error!**

# SCIANCE: AI IN SCIENCE

## Shaping AI in Science, by Scientists, for Europe

The EU-funded SCIANCE project supports the development of the Resource for AI Science in Europe (RAISE) and coordinates AI-enabled scientific research across Europe through a bottom-up approach.

### Workshops for AI in Science Pilot Areas



Astronomy  
& Physics



Materials  
Science



Earth  
Sciences



Life  
Sciences



Social Sciences  
& Humanities



<https://www.sciance.eu/>

**Projekt funding:  
3 million EUR over three years**

# SCIANCE: AI IN SCIENCE

## AISWG Members

10 AISWG working groups and their members.



<https://www.sciance.eu/aiswg-members>

## Materials Science

*SCIANCE co-chair: **Andrey Ustyuzhanin** (CU)*

**Román Orús** (DIPC / Multiverse Computing)

**Stephan Roche** (ICREA and ICN2)

**Sonia Conesa Boj** (Delft University of Technology)

**Nicola Marzari** (University of Cambridge)

**Karsten Reuter** (Fritz Haber Institute of the Max Planck Society)

**Gabor Csanyi** (Max Planck Institute for Polymer Research)

**Ares Argelia Gomez Gallegos** (Linnaeus University)

**Tejs Vegge** (Technical University of Denmark)

**Eleonora Ricci** (University of Edinburgh)

**Janine George** (Federal Institute for Materials Research and Testing (BAM))

**Hartmut Schlenz** (Forschungszentrum Juelich GmbH, Institute of Energy Materials and Devices (IMD), Materials Synthesis and processing (IMD-2))

**Gian-Marco Rignanese** (UCLouvain)

**Alfons Hoekstra** (University of Amsterdam)

**Maciej Haranczyk** (IMDEA Materials Institute)

**Michele Ceriotti** (EPFL)

**Gianaurelio Cuniberti** (TU Dresden)

**Planned investment in AI research (Horizon Europe): > 3 billion EUR / year.**

# SUMMARY

- **Practice-oriented RDM** opens up new opportunities in research and enables the generation of **FAIR data** and **metadata**.
- The **needs of users** must always be, and remain, the focus.
- **Automation** is made easier or even made possible in the first place.
- An efficient RDM is essential for the **application of AI** in science.
- **Automated laboratories** will be the **future of materials research**.  
RDM is an integral part of this.
- **Electronic laboratory notebooks** can now be used for most tasks, wherever their use is appropriate and feasible.
- Our **Best-Practice-Handbook** is designed to make it easier for researchers to work with research data in practice and to **clearly explain** the use of **everyday tools** and how they relate to each another.

# ACKNOWLEDGEMENT

# NFDi4ing



**Funded by  
the European Union**

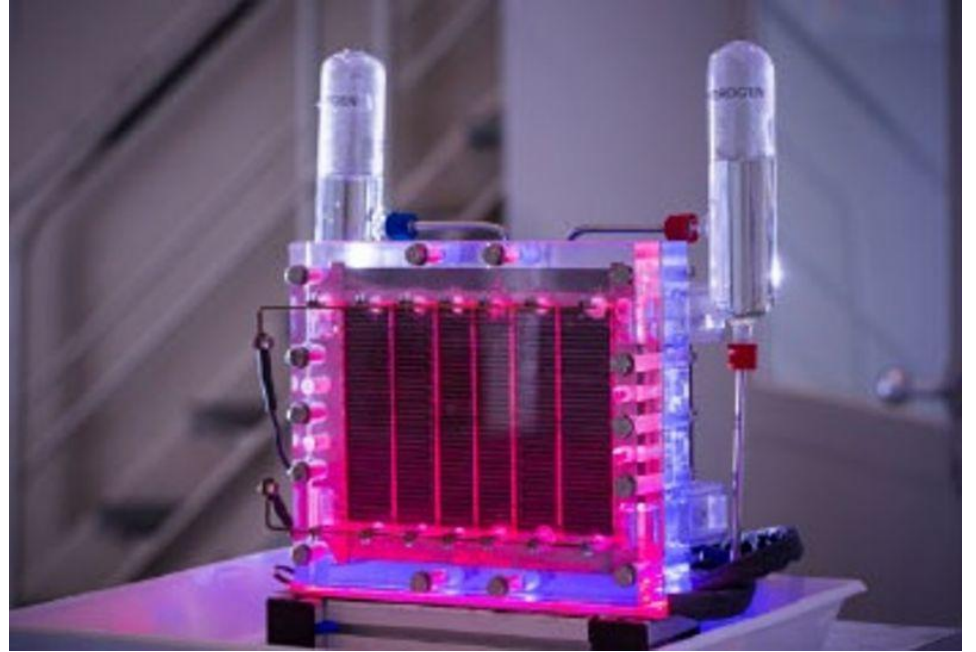
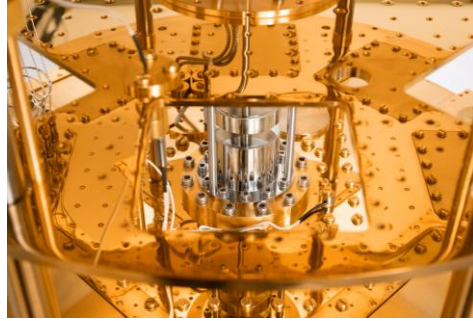
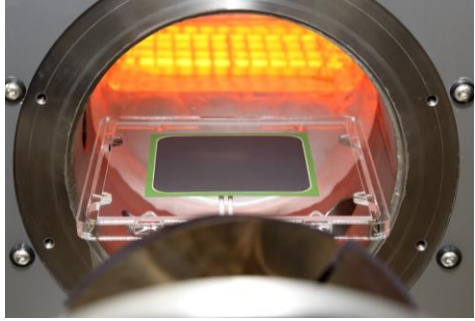


**Funding by the BoD at FZJ:  
Project TAPI**

## Board of Directors of the FZJ

**Torsten Bronger (FZJ)  
Michael Selzer (KIT)  
Britta Nestler (KIT)  
Christian Langenbach (DLR)  
Tan Le-Dinh (FZJ)  
Salome Enahoro (FZJ)  
Leo Riem (FZJ)  
Sven Rank (FZJ)  
Dorit Jerger (FZJ)  
Amber Sprenger (FZJ)  
Steffen Brinkmann (FZJ)  
Volker Hofmann (FZJ)**

# THANK YOU FOR YOUR ATTENTION



Thank you for your attention

