

## **Modernes Forschungsdaten-Management: Aktuelle Strategien, Nutzerunterstützung und neue Entwicklungen im FZJ und bei NFDI4Ing**

PD Dr. Hartmut Schlenz (IEK-1)

Unterstützt durch NFDI4Ing (Task Area CADEN)

# Forschungsdaten-Management (FDM)

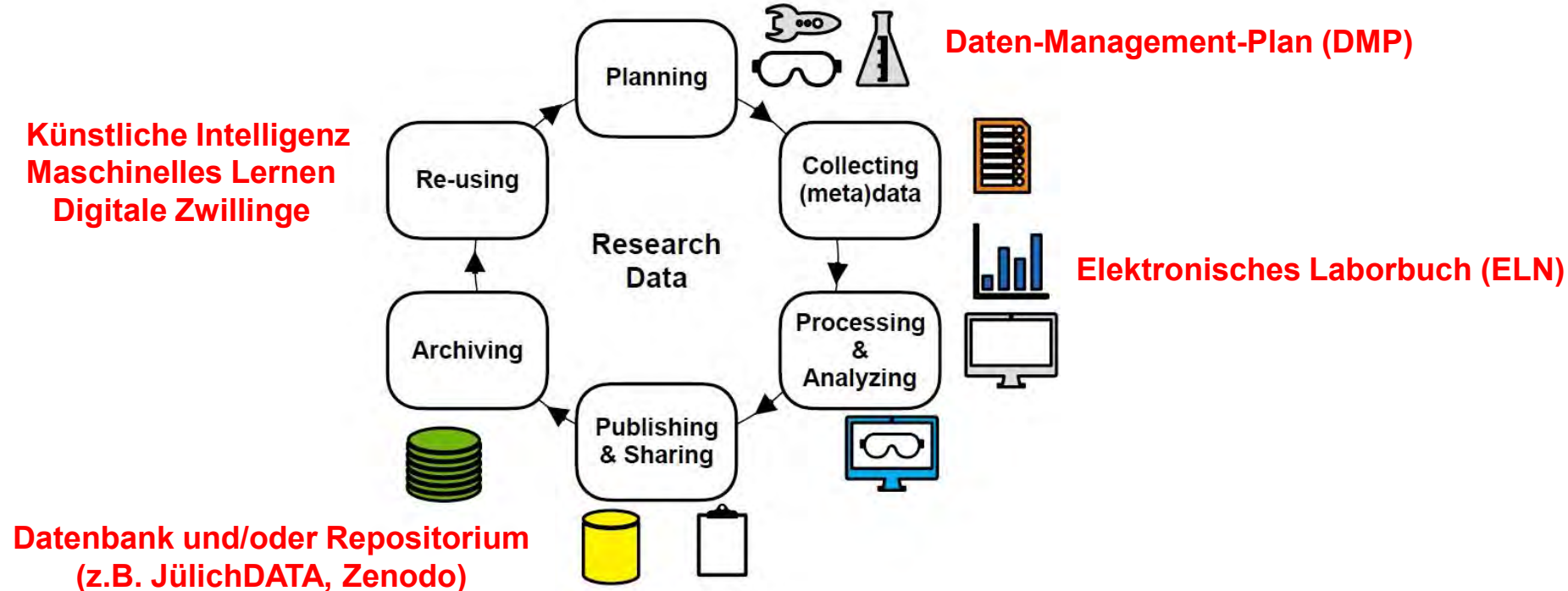
## Fragen

---

- Warum Forschungsdaten-Management (pro und contra)?
- Wie sehen die Bedürfnisse der Nutzer aus?
- Wo kann ich Hilfe bekommen (Infrastruktur, Software,...)?
- Woraus besteht ein modernes Forschungsdaten-Management?
  - Daten-Management-Plan (DMP)
  - Elektronisches Laborbuch (ELN)
  - Datenaustausch zwischen verschiedenen ELN's.
  - Daten-Repository (z.B. für Datenpublikationen)
- Was sind die aktuellen Entwicklungen?

# Forschungsdaten-Management (FDM)

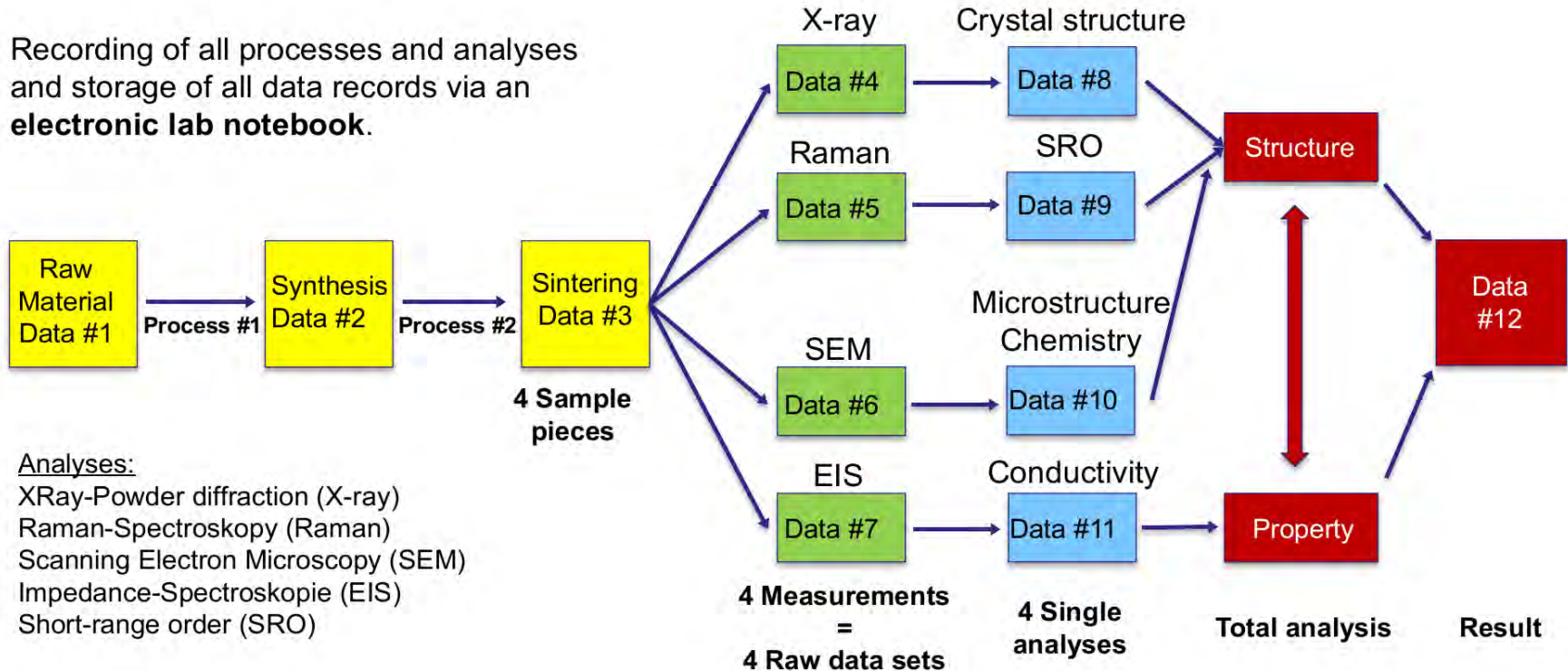
## Lebenszyklus von Forschungsdaten



# Forschungsdaten-Management (FDM)

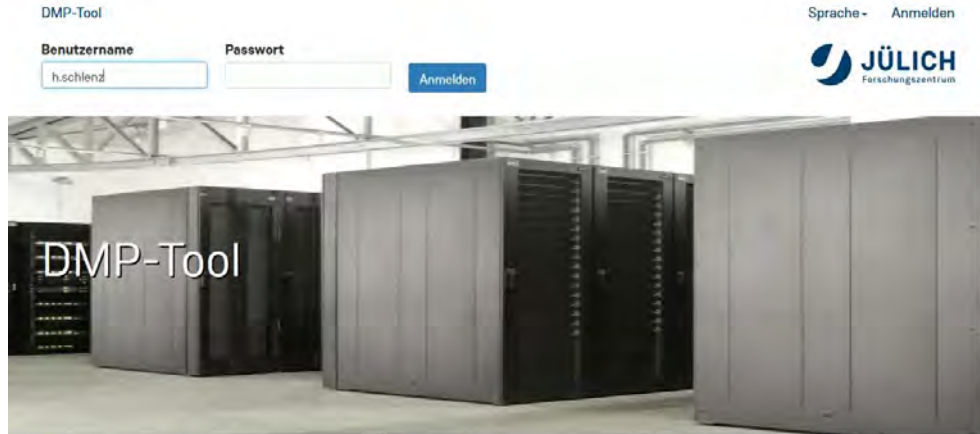
## Analyse einer Elektrokераmik (einfaches Beispiel)

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



# Forschungsdaten-Management (FDM)

## Portal für Daten-Management-Pläne des FZ Jülich



### 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.

# Forschungsdaten-Management (FDM)

## DMP-Werkzeug des FZ Jülich (DMP-Tool)

DMP-Tool

Sprache - Hartmut Schlenz -

### KI-gestützte Entwicklung von Wasserstoff-Membranen

**Beschreibung**  
Optimierung des Produktionsprozesses von Keramiken für Wasserstoff-Membranen mit Hilfe von Verfahren des Maschinellen Lernens.

**Katalog**  
Detaillierter Fragenkatalog

#### Aufgaben

Aufgaben werden automatisch aus den im Projekt gegebenen Antworten generiert. Auf der Seite jeder Aufgabe können Sie sehen, welche Ihrer Antworten zur Aktivierung der Aufgabe führen.

Keine aktiven Aufgaben gefunden.

#### Ansichten

Ansichten werden anhand der im Projekt gegebenen Antworten erstellt und können anschließend in verschiedenen Formaten exportiert werden. Zu Beginn sind alle Ansichten leer. Bitte beantworten Sie einige Fragen, indem Sie **Fragen beantworten** (oben in der Seitenleiste) besuchen.

| Ansehen                                         | Beschreibung                                                                                                                                                                                                                                                                                                                            |
|-------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Horizon 2020 FAIR Data Management Plan template | Vorlage für Horizon 2020, aus "Guidelines on FAIR Data Management in Horizon 2020", Online: <a href="http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf">http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf</a> |

#### Mitgliedschaften

Hier können Sie sehen, wer auf das Projekt zugreifen kann und weitere Mitglieder einladen. Über die Benutzerrollen können Sie verwalten, welche Rechte die Nutzenden haben. Sofern Sie nicht der letzte Besitzer sind, können Sie das Projekt mit der Schaltfläche neben Ihrem Namen verlassen.

| Benutzer        | E-Mail                                                               | Rolle    |
|-----------------|----------------------------------------------------------------------|----------|
| Hartmut Schlenz | <a href="mailto:h.schlenz@fz-juelich.de">h.schlenz@fz-juelich.de</a> | Besitzer |

#### Snapshots

Mit Snapshots können Sie alle Antworten zu einem bestimmten Zeitpunkt speichern und eine bestimmte Phase des Projekts beibehalten. Später kann der Snapshot zum Erstellen von Ansichten verwendet werden.

#### Optionen

- Fragen beantworten
- Antworten anzeigen
- Projektinformationen bearbeiten
- Projektkatalog bearbeiten
- Übergeordnetes Projekt bearbeiten
- Projekt Aufgaben bearbeiten
- Projekt Ansichten bearbeiten
- Projekt entfernen
- Mitglied hinzufügen
- Snapshot erstellen
- Zurück zu den Projekten

#### Export

- RDMO XML
- CSV durch Komma getrennt
- CSV durch Semikolon getrennt

#### Werte importieren

Importieren aus Datei

# Forschungsdaten-Management (FDM)

## DMP-Checkliste

---

1. Beschreibung der Daten (Messdaten, Texte, Bilder,...)
2. Dokumentation und Datenqualität (Metadaten)
3. Speicherung und Backups
4. Rechtliche Fragen (Patente, Publikationen)
5. Datenaustausch und permanente Verfügbarkeit (Infrastruktur)
6. Verantwortlichkeiten und Ressourcen (Kosten, Zeit)

Das Ziel ist die Erzeugung **FAIRer Daten!**  
(**F**indability, **A**ccessibility, **I**nteroperability, and **R**eusability).



# Forschungsdaten-Management (FDM)

## Interviews: Verwenden Sie ein Elektronisches Laborbuch (ELN)?

| System      | Open Source | Entwicklungsstand                                                               |
|-------------|-------------|---------------------------------------------------------------------------------|
| eLabFTW     | Ja          | Vollständig nutzbar. Ständige Weiterentwicklung. Sehr flexibel.                 |
| Kadi4Mat    | Ja          | Vollständig nutzbar. Ständige Weiterentwicklung. Schwerpunkt Materialforschung. |
| Chemotion   | Ja          | Vollständig nutzbar. Ständige Weiterentwicklung. Schwerpunkt Chemie.            |
| Sample DB   | Ja          | Vollständig nutzbar. Ständige Weiterentwicklung. Sehr flexibel.                 |
| JuliaBase   | Ja          | Vollständig nutzbar. Ständige Weiterentwicklung. Schwerpunkt Materialforschung. |
| PASTA DB    | Ja          | Generisch (momentan auf experimentelle Materialwissenschaft angewendet).        |
| elog        | Ja          | Weit verbreitet im Bereich Photon Science (u.a. am CERN).                       |
| Indigo ELN  | Ja          | Vollständig nutzbar. Ständige Weiterentwicklung. Sehr flexibel.                 |
| eLabNext    | Nein        | Kommerziell. Wird u.a. bei Bayer eingesetzt.                                    |
| Labfolder   | Nein        | Kommerziell. Probenpräparation und Mikroskopie.                                 |
| RSpace      | Nein        | Kommerziell. U.a. in der Physik, Geowissenschaften.                             |
| LabArchives | Nein        | Kommerziell. Sehr flexibel. Alle wiss. Bereiche. Integration mit LIMS möglich.  |
| CERF 5.0    | Nein        | Kommerziell. Sehr flexibel. Alle wiss. Bereiche. Integration mit LIMS möglich.  |



# Forschungsdaen-Management (FDM)



# Forschungsdaten-Management (FDM)

## Interviews: Zusammenfassung (pro und contra)

### Vorteile / Erwartungen

- + Zeitersparnis (häufigste Antwort)
- + Dokumentation von Metadaten
- + Ortsunabhängiger Zugriff
- + Alle Daten an einem Ort
- + Reproduzierbarkeit und Vergleichbarkeit von Experimenten
- + Datensicherung (Vermeidung von Datenverlust)
- + Automatisierung / Standardisierung von Workflows und wiederkehrender Prozesse
- + Auffindbarkeit von Daten nach Weggang der an Messungen beteiligten Personen
- + Erleichterung bei der Suche von Daten
- + Vereinfachter Datenaustausch / Kooperationen
- + Bereitstellung von Daten zur Datenanalyse

### Nachteile / Bedenken

- Einarbeitung erfordert (hohe) Ressourcen
- Digitale Systeme erfordern Maintenance
- Je nach Laborumgebung nicht praktikabel (z.B. bei Arbeit mit Handschuhen)
- Flexibilität von ELNs bisher unzureichend für manche Experimente
- Datenschutz muss gewährleistet sein



Bedarf an Beratung und Information

# Forschungsdaten-Management (FDM)

## FDM-Portal des FZ Jülich



HOME > THEMEN > FDM

### News:



BETEILIGEN SIE SICH AM FDM  
ROCKET-CHAT!



HMC FAIR FRIDAY MIT SÖREN  
AUER



NICOLAS CARPI ZU GAST IN  
JÜLICH



NEVER LEHRGANG:  
FUNDAMENTALS OF SCIENTIFIC  
METADATA - WHY CONTEXT  
MATTERS!

### Top Themen:



AUF FDM SPEZIALISIERTE  
EINHEITEN



FAIR-PRINZIPIEN



KARRIERESCHUB -  
DATENPUBLIKATION!



FDM CHALLENGE-PROJEKTE

Zusätzlich neue  
**FDM/ELN Taskforce des FZJ**,  
für die Entwicklung neuer FDM-  
Strategien und -Lösungen.

# Forschungsdaten-Management (FDM)

## FDM-Portal des FZ Jülich

### Services nach Funktionen sortiert

- Access
- Data and Lab Management
- Data Discovery
- Data Management Plans
- Data Sharing
- Metadata Management
- Publication
- Persistent Identifier Assignment
- Storage
- Short-term Storage
- Archival Storage
- Backup
- Cloud Services
- Synchronization
- Version Control

### Services nach Instituten sortiert

- IAS-9
- IBG-3
- IEK-5
- INM-6
- INM-7
- ITS
- JSC
- PGI/CNS-TA
- ZB

### Toolbox

Bei Forschungsprojekten in den verschiedensten wissenschaftlichen Disziplinen fallen Daten an, die gemanagt werden müssen. Unter Forschungsdatenmanagement versteht man den Umgang mit Forschungsdaten über den gesamten Datenlebenszyklus hinweg. Diese Toolbox soll einen **Überblick über Werkzeuge des Forschungsdatenmanagements** geben, die am Forschungszentrum zur Verfügung stehen und über deren Nutzung Jülicher Forscherinnen und Forscher Auskunft geben können.

Im Sinn der Toolbox umfasst das Forschungsdatenmanagement all jene Prozesse, die (mindestens) eine Datenmanagementfunktion erfüllen. Wir unterscheiden dabei zwischen den folgenden Funktionen: Access, Discovery, Storage, Publication, Sharing, Metadatenmanagement und Version Control. Zwei weitere Kategorien geben einen Überblick über Tools zur Erstellung von Datenmanagementplänen und zum Daten- und Labormanagement. Die **Werkzeuge sind filterbar** nach den funktionalen Kategorien oder auch nach den involvierten Organisationseinheiten.

*Die Toolbox richtet sich vor allem an die Community der Jülicher FDM-Anwenderinnen. Beschreibungen von Tools sind deshalb ausschließlich auf Englisch verfügbar.*

### SERVICES NACH DATENMANAGEMENTFUNKTIONEN SORTIERT

### SERVICES NACH INSTITUTEN SORTIERT



Haben Sie oder Ihr Team ein Tool entwickelt haben, das eine (oder mehrere) der genannten Datenmanagementfunktionen erfüllt, und Sie möchten dieses in die Toolbox aufnehmen lassen, damit es von anderen genutzt werden kann? Dann füllen Sie bitte das **Template** aus und senden Sie es an eine der unten genannten Kontaktpersonen.

### Template für die Aufnahme neuer Dienste und Service in die Toolbox

Word file  
PDF file



ACCESS



DATA AND LAB MANAGEMENT



DATA DISCOVERY



DATA MANAGEMENT PLANS



DATA SHARING



METADATA MANAGEMENT



PUBLICATION



STORAGE



VERSION CONTROL



# Elektronische Laborbücher (ELN)

## eLabFTW als ein Beispiel

---



The Free and Open Source  
Electronic Lab Notebook



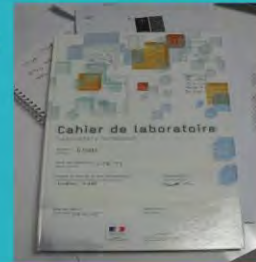
[www.elabftw.net](http://www.elabftw.net)

## Why use an ELN?

### Problem:

Need to store metadata of experiments (what, when, where, who, why, how)

### Solution:



### Limitations:

- Can be lost or burned
- No backups
- Often empty
- Not readable by others easily
- Can't search for something

# Elektronische Laborbücher (ELN)

## eLabFTW als ein Beispiel

### eLabFTW: key features

Web based

No client to install

Responsive design  
(works with all screen  
sizes)

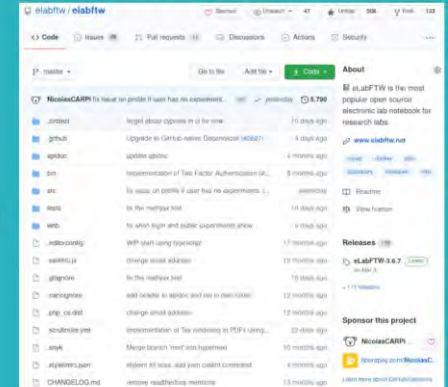


### eLabFTW: key features

Open source

Access to the source code is  
immediate

Users contributions are welcome





# Elektronische Laborbücher (ELN)

## eLabFTW als ein Beispiel

---

### eLabFTW: key features

- Trusted timestamping (RFC 3161)
- Blockchain timestamping
- API access
- SAML authentication
- Templates for experiments
- Database for products/protocols
- Export in ZIP, PDF, CSV, JSON
- Molecule editor
- Todolist
- Tex rendering
- Very customizable
- Translated in 17 languages

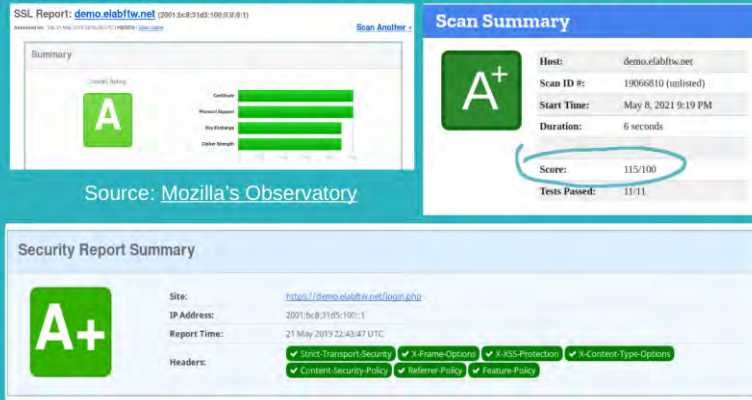
### eLabFTW: key features For research centers/universities

- Host it on your network
- Centralize your results on one server
  - No data gets lost if a laptop of a postdoc is stolen
  - Several teams can be hosted on the same instance

# Elektronische Laborbücher (ELN)

## eLabFTW als ein Beispiel

### eLabFTW: secure software



### eLabFTW: supported software

Free support on GitHub  
PRO support available from Deltablot  
Hosting available from Deltablot

<https://www.deltablot.com>

[contact@deltablot.email](mailto:contact@deltablot.email)



# Elektronische Laborbücher (ELN)

## TAPI-Projekt (2022/23)

---

**Template- und API-Entwicklung für die Elektronischen Laborbücher eLabFTW und Kadi4Mat (TAPI)**

Finanziert durch den Vorstand des FZ Jülich.

H. Schlenz & T. Bronger

- Erhöhung der **Akzeptanz** für die ELN-Nutzung, durch deutlich vereinfachte Arbeitsabläufe.
- Darstellung der **Vorteile durch die ELN-Nutzung** und die Reduzierung der Arbeitsbelastung, im Vergleich zum Papier-Laborbuch.
- **Programmierung von Vorlagen (Templates)** für die deutlich vereinfachte Nutzung von ELN's und Geräten.
- **Programmierung von Schnittstellen (API)** für die Automatisierung der Datenerfassung an Geräten durch ein ELN.

# Elektronische Laborbücher (ELN)

## eLabFTW-Templates (Vorlagen)

eLabFTW EXPERIMENTS DATABASE

### Experiments

Create new ▸

Select all

Untitled

☐ IN ARBEIT

- Experiment
- Templates
- Ellipsometrie
- Plasmabehandlung
- Manage templates

### Experiment

File Edit View Insert Format Tools Table

↶ ↷ Paragraph Link... Ctrl+K

Insert template...

Code sample...

Table >

Special character...

Horizontal line

Page break

Date/time >

|                     |  |
|---------------------|--|
| Probe               |  |
| Präparation         |  |
| Experimentator/in   |  |
| Datum und Uhrzeit   |  |
| Microspot oder para |  |
| Wellenlängenbereich |  |
| Winkelbereich       |  |
| Link zu Rohdaten    |  |
| Bemerkungen:        |  |



# Elektronische Laborbücher (ELN)

## Röntgenbeugungs-Template für eLabFTW

**Experimente**

Zurück zur Übersicht Anzeigemodus

---

Begonnen am 08 / 16 / 2021

Status Running

Titel  
Emptyrean-Template.tpl

Schlagwörter  
[Schlagwort hinzufügen](#)

---

Datei Bearbeiten Ansicht Einfügen Format Werkzeuge Tabelle

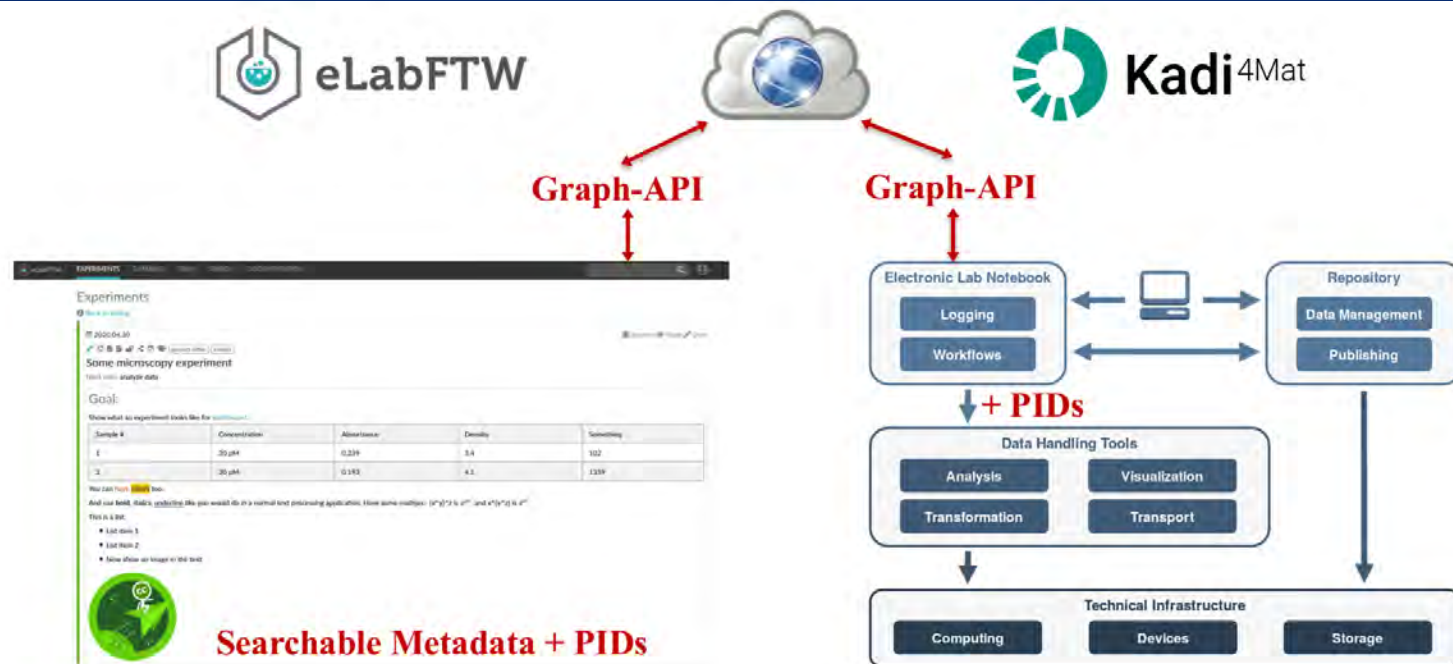
↶ ↷ Absatz B I U ☰ ☷ ☹ ☺ ☻ ☼ ✕ × ⌵ ⌴ ⌵ ⌴ ⌵ ⌴ A ▼ 🔍 Ω ⌚ {} 🔗 📄

| PXRD EXPERIMENTAL DATA - Emptyrean-Template (CIF) |                                |                                      |
|---------------------------------------------------|--------------------------------|--------------------------------------|
| <b>RAW filename:</b>                              |                                |                                      |
| chemical_formula_sum                              | Al <sub>2</sub> O <sub>3</sub> |                                      |
| chemical_name_common                              | Corundum                       |                                      |
| diffrn_ambient_temperature                        | 295                            |                                      |
| diffrn_radiation_probe                            | x-ray                          |                                      |
| diffrn_radiation_wavelength                       | 1.540598                       | # CuKα-1~ 1.540588, CuKα-2~ 1.544426 |
| diffrn_radiation_type                             | 'Cu Kα-1+2~'                   |                                      |
| diffrn_source                                     | 'sealed X-ray tube'            |                                      |
| diffrn_source_target                              | Cu                             |                                      |
| diffrn_source_type                                | 'Malvernpanalytical Emptyrean' |                                      |
| diffrn_source_power                               | 45                             | # kV                                 |
| diffrn_source_current                             | 10                             | # mA                                 |

BETRIEBEN VON TINY

# Datenaustausch zwischen verschiedenen ELN's

## eLabFTW und Kadi4Mat



**NFDi4ing**



**TA CADEN:** T. Bronger, H. Schlenz, M. Selzer, B. Nestler, M. Flemisch, M. Jayavarapu

# Datenaustausch zwischen verschiedenen ELN's

## SciMesh (Probenbasierte Nachverfolgung, sog. Provenance Tracking)

### 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.

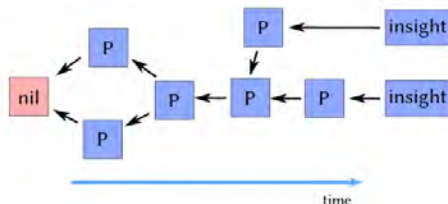


Fig. 1 Example graph of processes.

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 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.

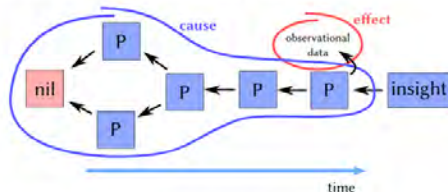


Fig. 2 Example graph of processes, showing cause and effect.

### Workflow

#### Institute A with ELN A

create sample #1  
with URL `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 Institute A

open the data sheet for sample #1

ELN B makes an HTTP GET request  
against the URL to that URL, 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>

NFDi4ing



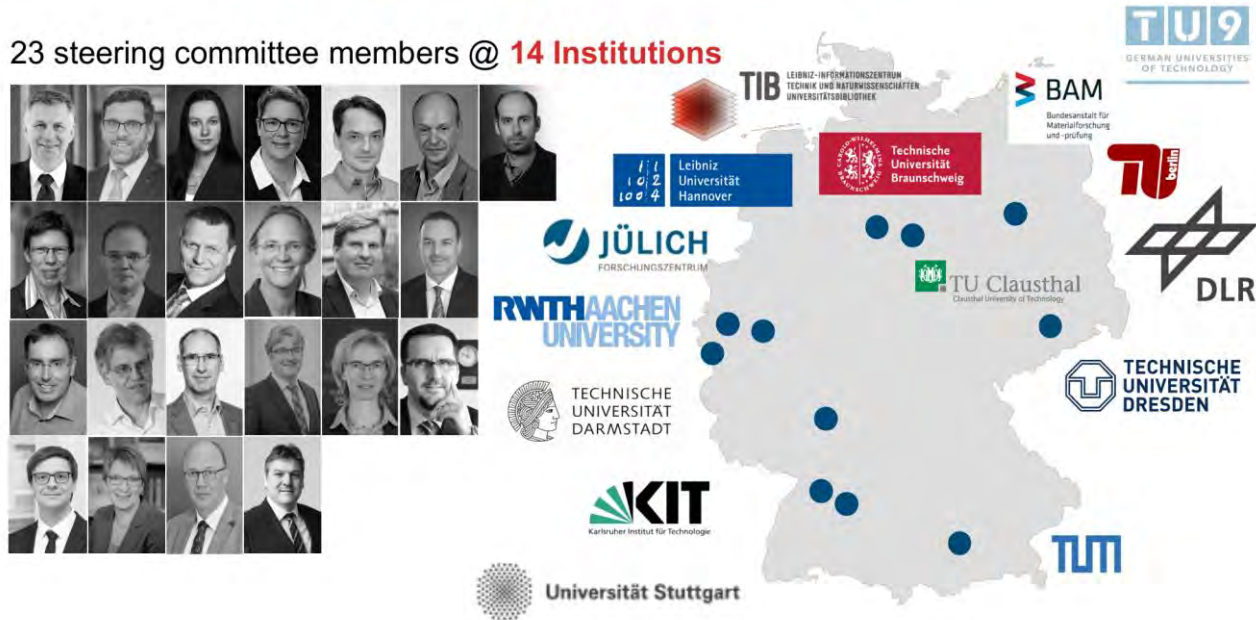
# Nationale-Forschungsdaten-Infrastruktur (NFDI)

## Konsortium NFDI4ing (<https://nfdi4ing.de>)

Relevance   Community Needs   RDM Solution   Operating Model   Scientific Services   Community Embedding   Structure

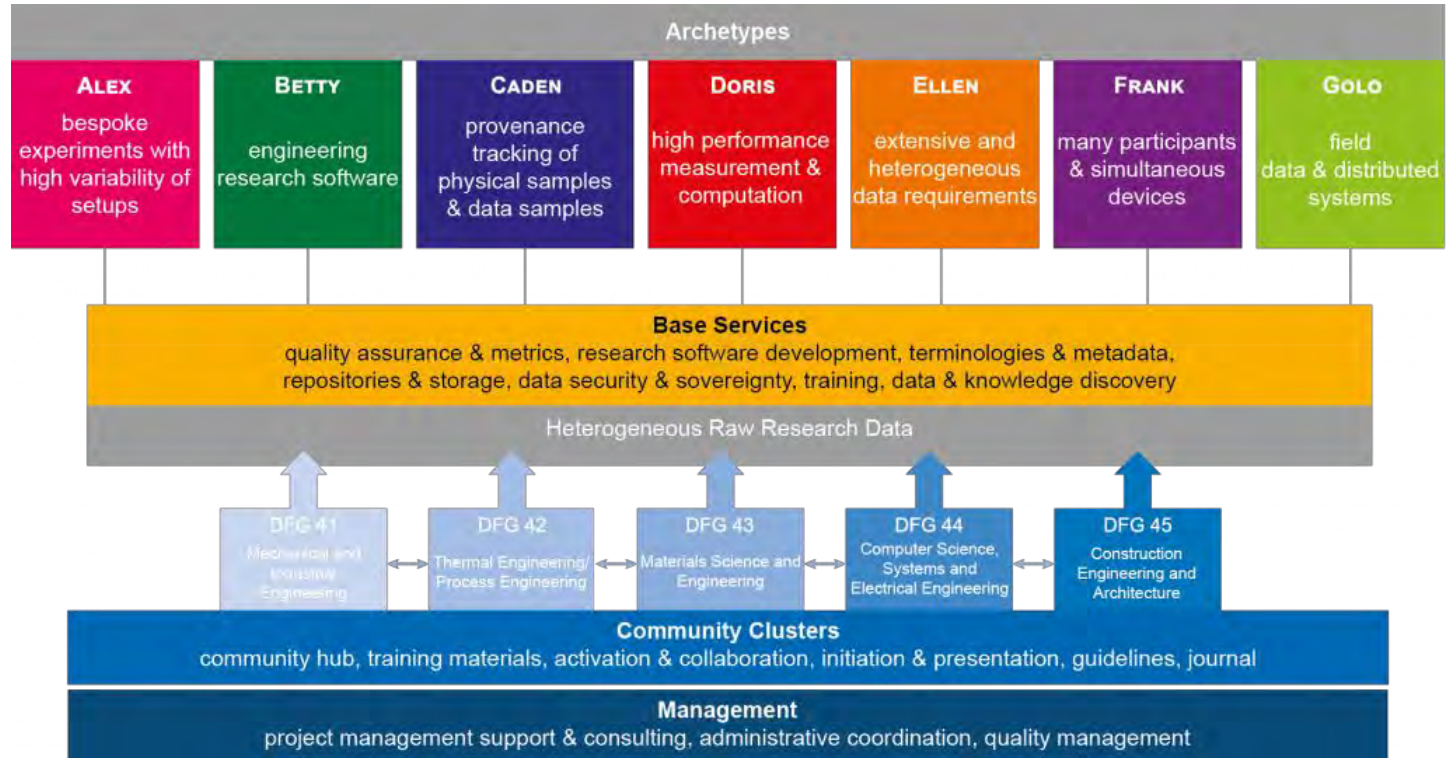
### NFDI4ing: a strongly integrated and coherent team

23 steering committee members @ 14 Institutions



# Nationale-Forschungsdaten-Infrastruktur (NFDI)

Konsortium NFDI4ing (<https://nfdi4ing.de>)





# Forschungsdaten-Management (FDM)

## Schlussbemerkungen

---

- Die Nutzung von **DMP's** and **ELN's** wird für neue Projekte (u.a. Master- und Doktorarbeiten) dringend empfohlen.
- Die Erzeugung **FAIRer Daten** sollte ein gemeinsames Ziel sein.
- Ein gut strukturiertes FDM ermöglicht den Einsatz von **Maschinellern** und die Konstruktion **Digitaler Zwillinge** (Nachnutzung).
- **Elektronische Laborbücher (ELN)** sind ein wichtiger und zentraler Baustein in einer modernen FDM-Strategie.
- **Templates** für Experimente und die Automatisierung durch intelligente **Schnittstellen (API)** wird die Verwendung von ELN's deutlich vereinfachen.
- Die **Nationale-Forschungsdaten-Infrastruktur (NFDI)** wird in den nächsten Jahren (weitere) umfangreiche FDM-Lösungen schaffen.