



**NFDI<sub>4</sub>Chem**

ENHANCE  
YOUR  
DATA.

## Basics of Research Data Management (RDM): Crash Course



CC BY 4.0 ©NFDI4Chem



# Workshop topics

---



- Introduction
- Research data management basics
- FAIR principles
- Data life cycle
- NFDI4Chem





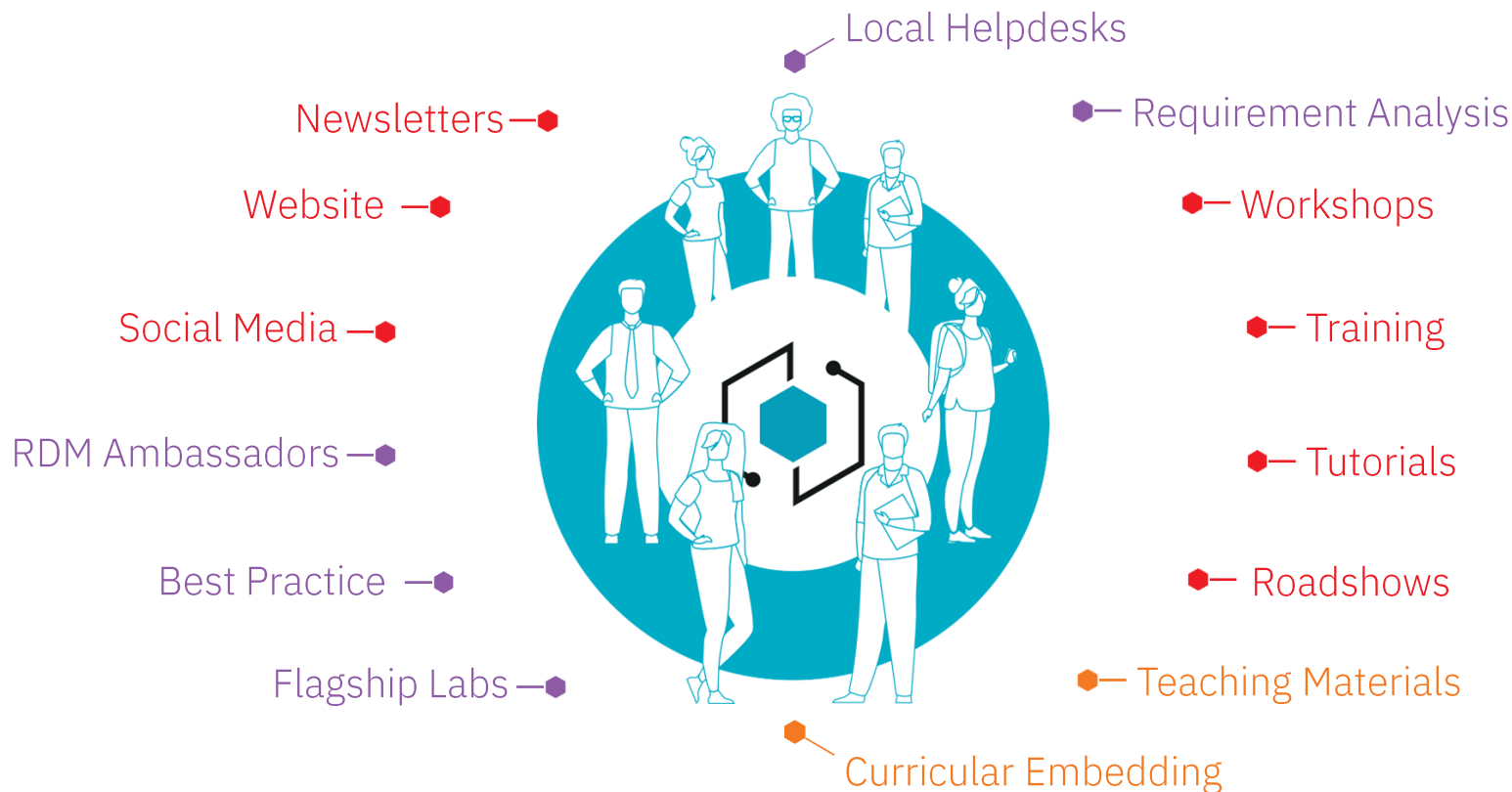
**NFDI<sub>4</sub>Chem**

ENHANCE  
YOUR  
DATA.

# Introduction



# NFDI4Chem: Involvement of the community





# Motivation



- NFDI4Chem - Data is available upon reasonable request - RDM in Chemistry SNAFU by UB RWTH Aachen University



<https://www.youtube.com/watch?v=OA0GcvacjgI>



 **NFDI<sub>4</sub>Chem**

**Data is available  
upon reasonable  
request - RDM in  
Chemistry SNAFU**



**ub**  
University  
Library

**RWTHAACHEN**  
UNIVERSITY



# Workshop motivation

---



**I am attending this workshop to learn more about**



- a) Research data management in general
- b) A specific research data management topic
- c) How to start with research data management
- d) NFDI4Chem services and support
- e) Other reasons

[www.menti.com](https://www.menti.com)

Code: 8134 3598



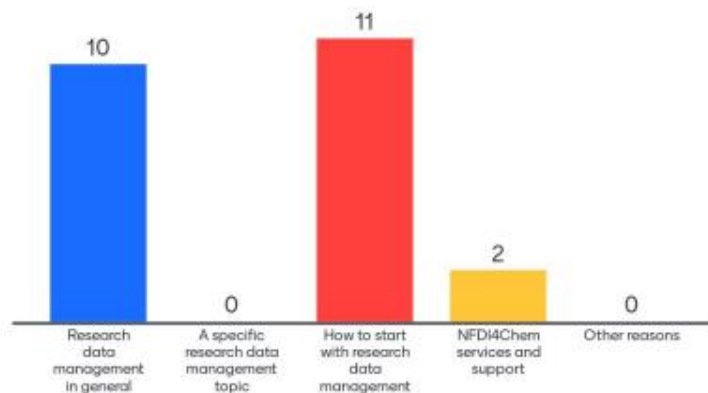
# Workshop motivation

---



Mentimeter

I am attending this workshop to learn more about





**RDM basics**



**NFDI<sub>4</sub>Chem**

ENHANCE  
YOUR  
DATA.



# Research data

---



- Difficult to define
- Discipline-specific definitions

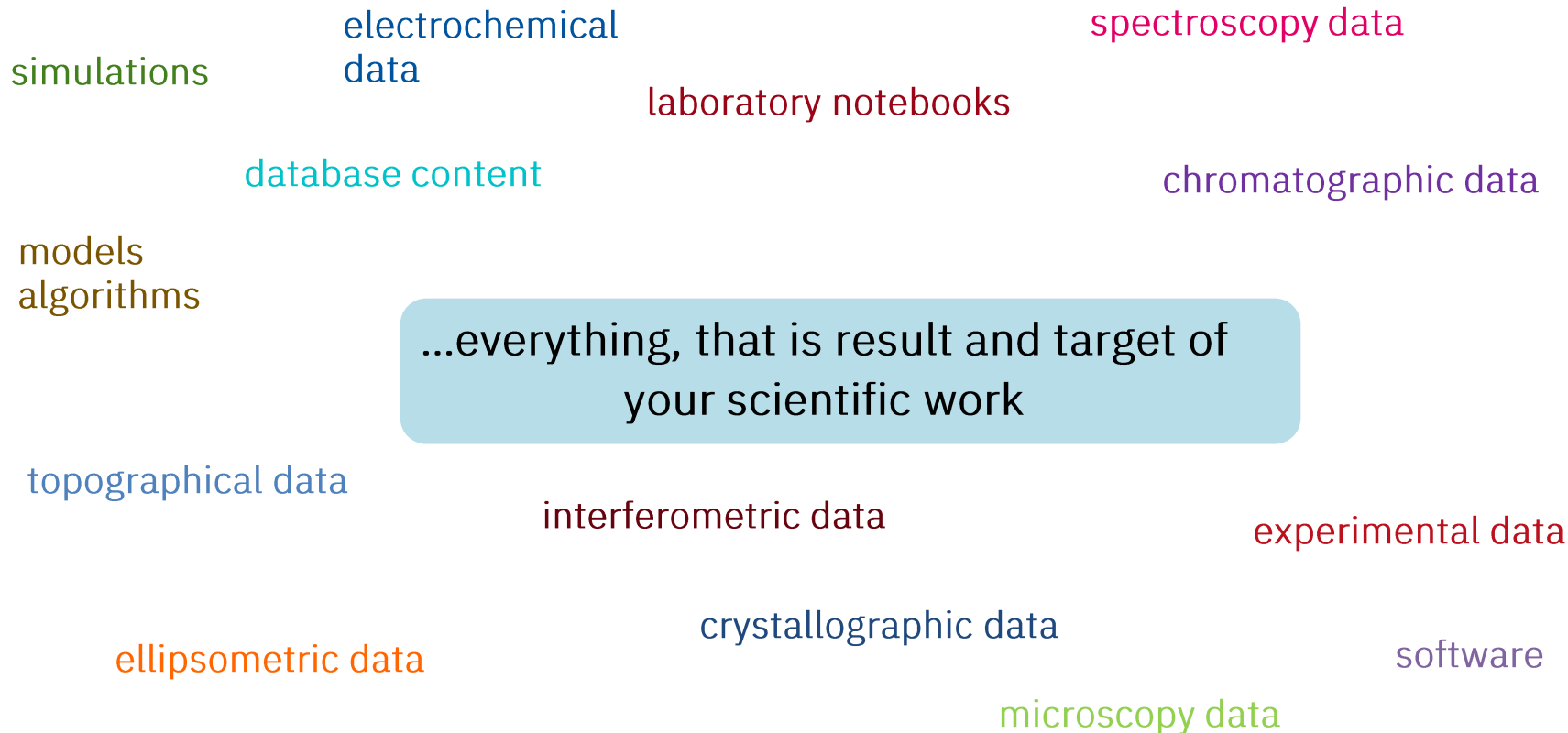
## DFG:

"Research data include **measurement data**, **laboratory values**, **audiovisual information**, **texts**, **survey data**, **objects from collections or samples** that are **created**, **developed** or **evaluated** in scientific work. **Methodological test procedures** such as **questionnaires**, **software** and **simulations** can also represent central results of scientific research and should therefore also be included under the term research data. "



# Research data in chemistry

---





# Research data management

---



- Activity of working with research data throughout the research process
- Including all aspects from data collection, to data storage and backup, through to data sharing
- One of the essential areas of responsible conduct of research

"Although handling research data can be challenging, managing your data effectively will not only help your **research to be robust and replicable**, but can help you to **anticipate potential problems** that can occur during the research process, and will **ensure that your research meets the requirements set out by research funders and publishers.**"



# Research data management – Why?



- Data security and prevention of data loss
- Verifiability, reproducibility and transparency of research results
- Reduction of scientific errors
- Faster retrieval of data and information
- Long-term availability of research data
- Data re-use in new research projects
- Required in guidelines and institutional policies on handling research data
- Requirement of third-party funders and science organisations



RDM basics

# FAIR principles



**NFDI**<sub>4</sub>Chem

ENHANCE  
YOUR  
DATA.

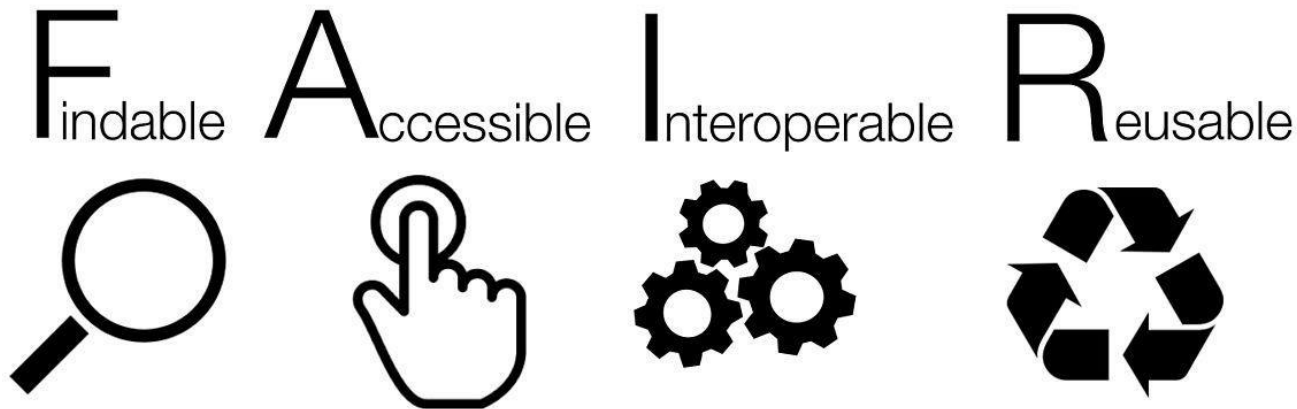


# FAIR principles

---



- According to the DFG's new guidelines on good research practice, research data must be FAIR!





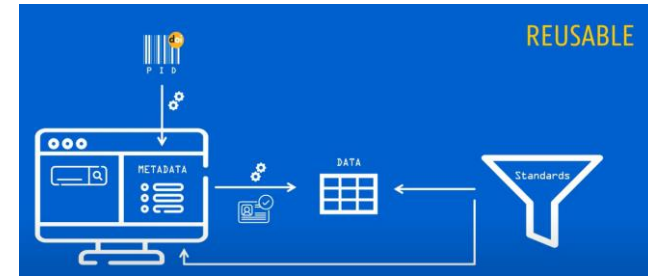
# FAIR principles



FAIR data principles by Ugent Data Stewards



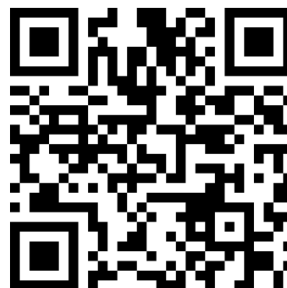
FAIR GAME







## FAIR principles



[www.menti.com](https://www.menti.com)

Code: 8134 3598



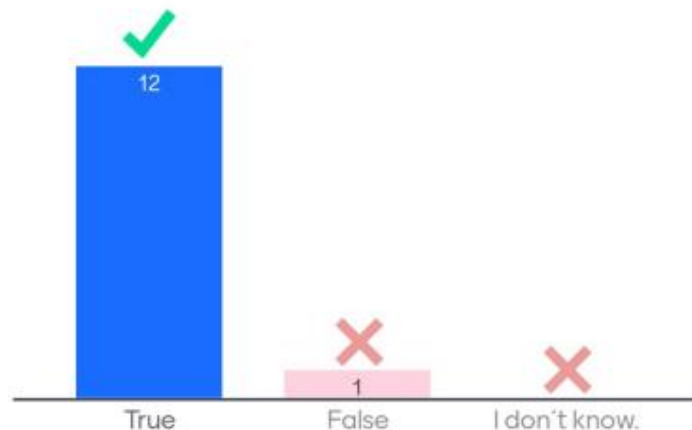
## Quiz - Results

---



Mentimeter

FAIR is an acronym for findable, accessible, interoperable and reusable





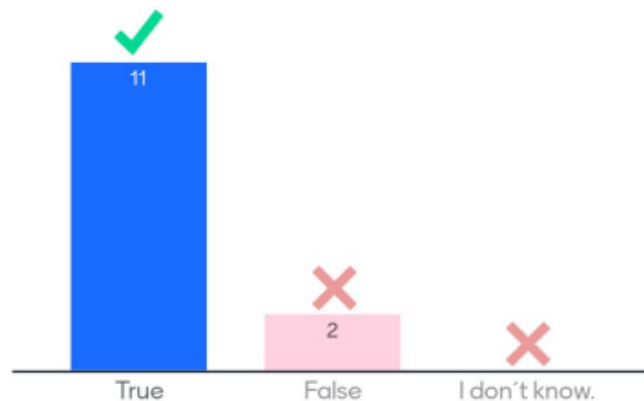
## Quiz - Results

---



Mentimeter

FAIR is a requirement by many research funders





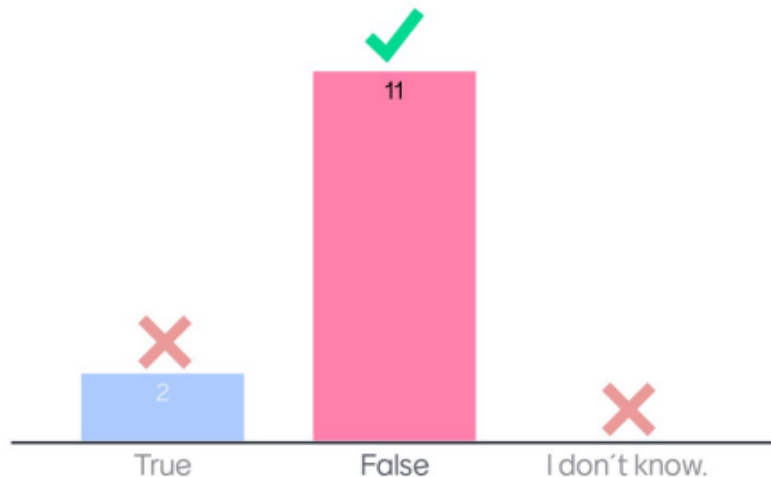
# Quiz - Results

---



Mentimeter

## Data must be open access to be FAIR





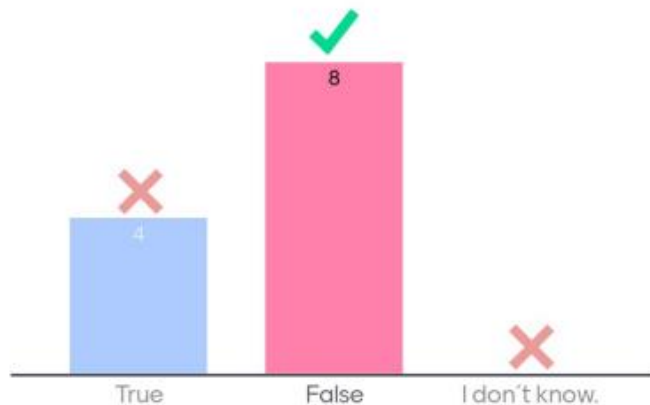
# Quiz - Results

---



Mentimeter

If the data are no longer available, the metadata become meaningless





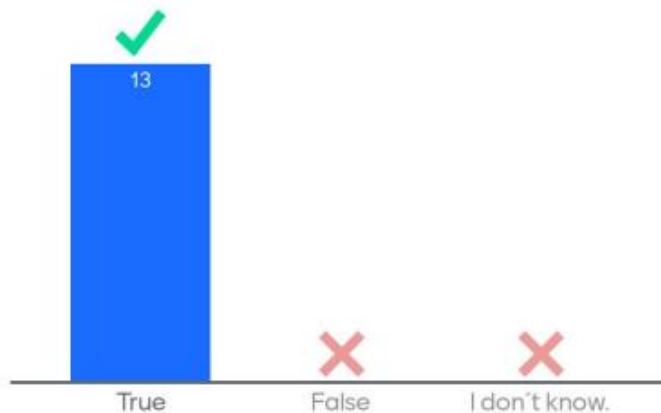
# Quiz - Results

---



Mentimeter

Data repositories are a key infrastructure enabling FAIR data





RDM basics

# Data life cycle

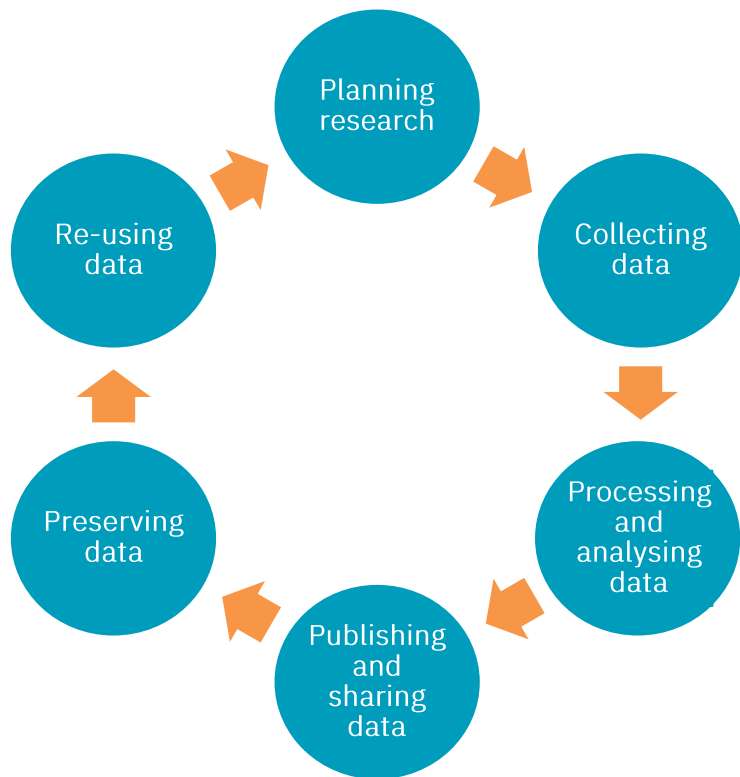


**NFDI**<sub>4</sub>Chem

ENHANCE  
YOUR  
DATA.



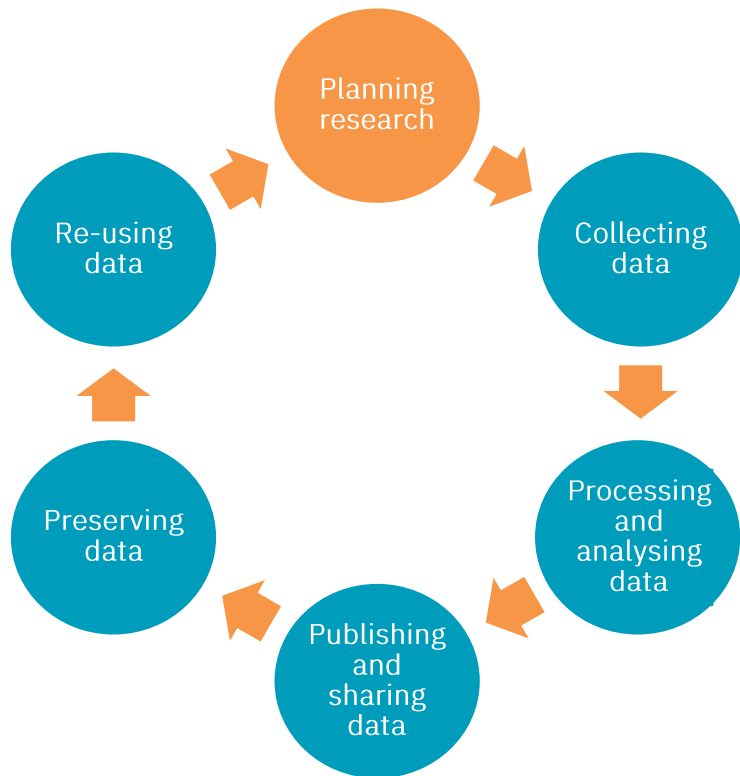
# Data life cycle



- A key concept in research data management
- Describes the lifespan of the data and beyond
- Based on various phases
- Different approaches to the same model depending on the institution, the funder, ...



# Planning research



- Research design
- Planning data management (formats, storage locations, ...)
- Create an initial **data management plan**
- Determine responsibilities
- Locate existing data
- Clarify authorship and data ownership
- Coordinate access conditions, prepare consent procedures



Planning research

## Data management plan



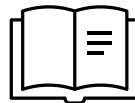
# What is a Data Management Plan (DMP)?

---



Description of the handling of research data during and after a research project

A DMP is a formal and at the same time a living document

- **WHAT** data goes into a project (reuse) and comes out of it (potential reuse)?
- **HOW** does the team take care of the data? 
- **WHO** is allowed to do **WHAT** with the data **WHEN**?



# Benefits of a DMP

---





# Requirements of Research Funders – Example: DFG



## Checklist



A checklist must be submitted as part of the proposal

1. Data description
2. Documentation and data quality
3. Storage and technical archiving the project
4. Legal obligations and conditions
5. Data exchange and long-term accessibility
6. Responsibility and resources

Released December 2021

Link to DFG checklist: [de/en](#)





**NFDI**<sub>4</sub>Chem

ENHANCE  
YOUR  
DATA.

Planning research

# Good Research Practice and Funders



# Funding organisation guidelines

---



There are different formal requirements depending on the individual research funding organisation! Before you apply, find out exactly what is relevant in your case.

Overview:

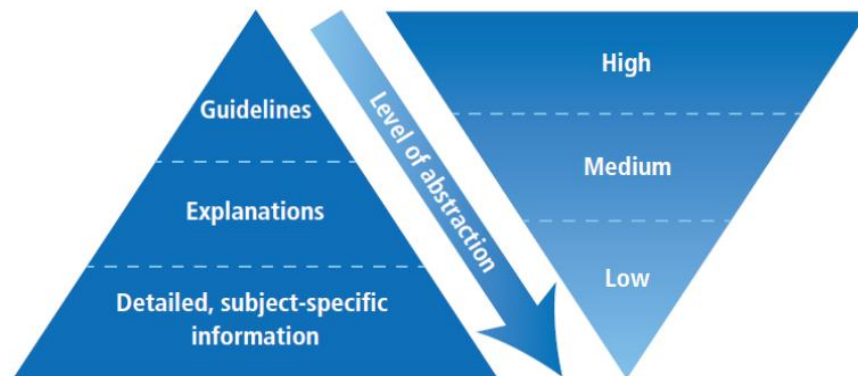
<https://www.forschungsdaten.info/themen/informieren-und-planen/foerderrichtlinien/#c492536>



# DFG Guidelines for Safeguarding Good Research Practice (2019)



- Fundamental revision of the recommendations from 1998
- Transition extended until 31st July 2023
- Modifications:
  - Recommendations
  - Multidimensional approach
  - Codex with 19 guidelines
  - 11 guidelines on the research process
  - RDM is relevant in 8 of these 11 guidelines



**All research institutes must implement these guidelines in a legally-binding manner in order to be eligible to receive DFG funding.**



# General GRP-Guidelines (DFG)

---



- Guideline 7: Cross-phase quality assurance
- Guideline 10: Legal and ethical frameworks, usage rights
- Guideline 11: Methods and standards
- Guideline 12: Documentation
- Guideline 13: Providing public access to research results
- Guideline 14: Authorship
- Guideline 15: Publication medium
- Guideline 17: Archiving



# Example: Guideline 12 Documentation

---



- Documentation of all information relevant to the production of a research result (in accordance with existing recommendations and guidelines)
- Selection of results must be avoided!
- Documentation and research results must not be manipulated; they are protected as effectively as possible against manipulation.

## Relevance in terms of research data:

- Documentation of necessary information to understand the research (results)
- Information on research data used or generated, the methods, evaluation and analysis steps, the development of the hypothesis and citations
- Possibility of replication
- Documentation of the source code in the development of research software



# Research Integrity in Chemistry

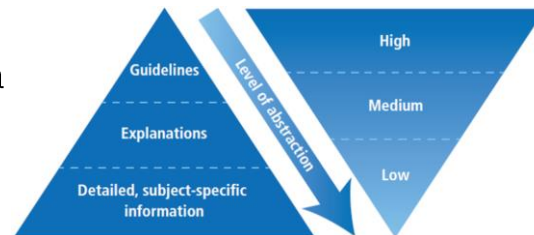


Portal: <https://wissenschaftliche-integritaet.de/en>  
for subject-specific information

- Articles available in German and English

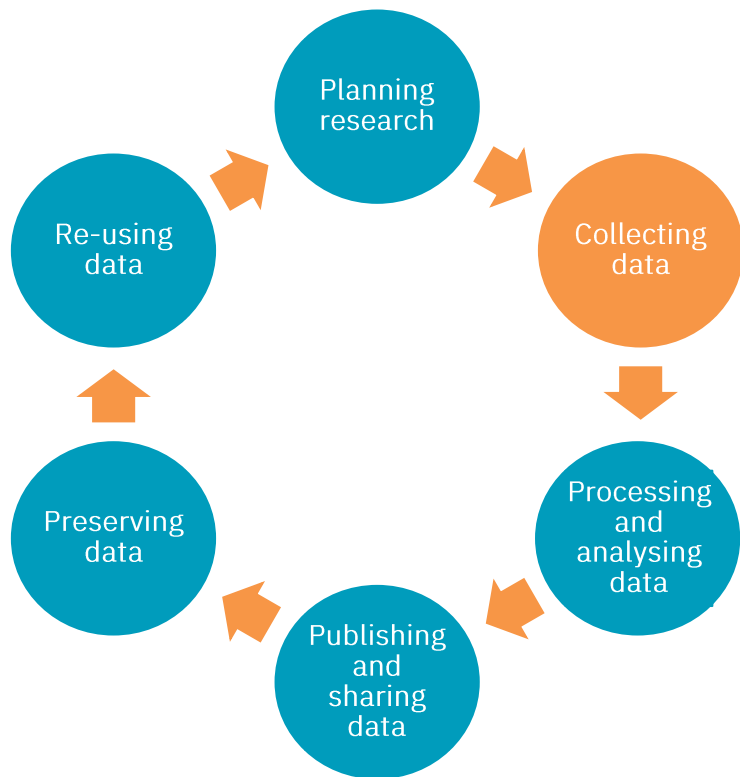
## Exemplary results for chemistry:

- Documentation of research results in experimental chemistry
- Quality assurance in experimental chemistry
- Use of chemistry-specific repositories
- Handling research software - Case studies
- Further links to performance dimensions and evaluation criteria
- Author order in physics and chemistry
- Electronic laboratory journal and repository in chemistry
- Ethical principles in chemistry





# Collecting data



- Perform experiments, measurements, simulations, observations...
- Collect and create **metadata**
- Document and describe data
- Enter, digitize, transcribe and translate data
- Check, validate and clean data
- Save and manage data



Collecting data

# Metadata





## What is metadata and why is it important?

- Data that describes data
- Makes datasets searchable (and findable)
- Makes datasets understandable and FAIR
- Machine and human-readable
- Standardization is ongoing





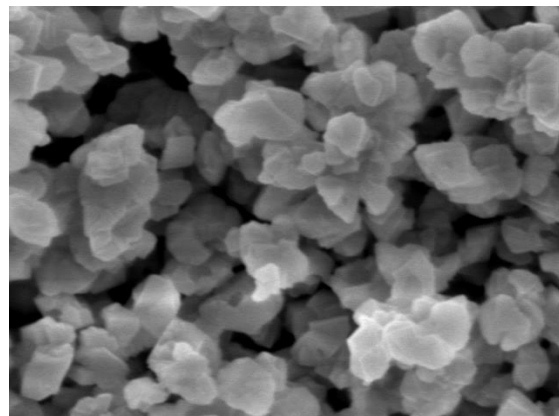
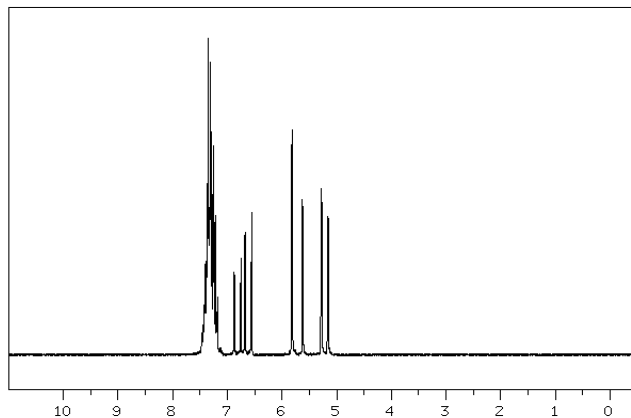
# Metadata and FAIR



FINDABLE	ACCESSIBLE
<p>F1. <b>(Meta)</b>data are assigned a globally unique and persistent identifier</p> <p>F2. Data are described with rich <b>metadata</b> (defined by R1 below)</p> <p>F3. <b>Metadata</b> clearly and explicitly include the identifier of the data they describe</p> <p>F4. <b>(Meta)</b>data are registered or indexed in a searchable resource</p>	<p>A1. <b>(Meta)</b>data are retrievable by their identifier using a standardised communications protocol</p> <p>A1.1. The protocol is open, free, and universally implementable</p> <p>A1.2. The protocol allows for an authentication and authorisation procedure, where necessary</p> <p>A2. <b>Metadata</b> are accessible, even when the data are no longer available</p>
INTEROPERABLE	REUSABLE
<p>I1. <b>(Meta)</b>data use a formal, accessible, shared, and broadly applicable language for knowledge representation.</p> <p>I2. <b>(Meta)</b>data use vocabularies that follow FAIR principles</p> <p>I3. <b>(Meta)</b>data include qualified references to other <b>(meta)</b>data</p>	<p>R1. <b>(Meta)</b>data are richly described with a plurality of accurate and relevant attributes</p> <p>R1.1. <b>(Meta)</b>data are released with a clear and accessible data usage license</p> <p>R1.2. <b>(Meta)</b>data are associated with detailed provenance</p> <p>R1.3. <b>(Meta)</b>data meet domain-relevant community standards</p>



# Exercise - What metadata is needed?







What metadata is needed?



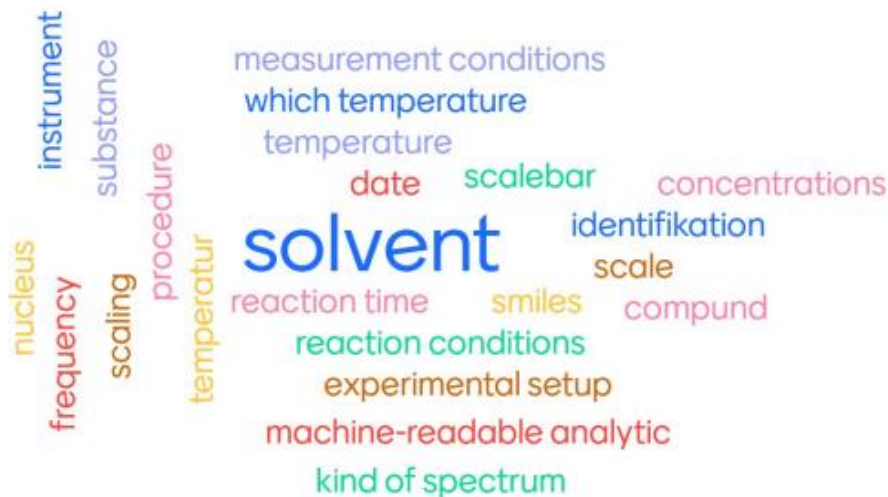
[www.menti.com](https://www.menti.com)

Code: 8134 3598



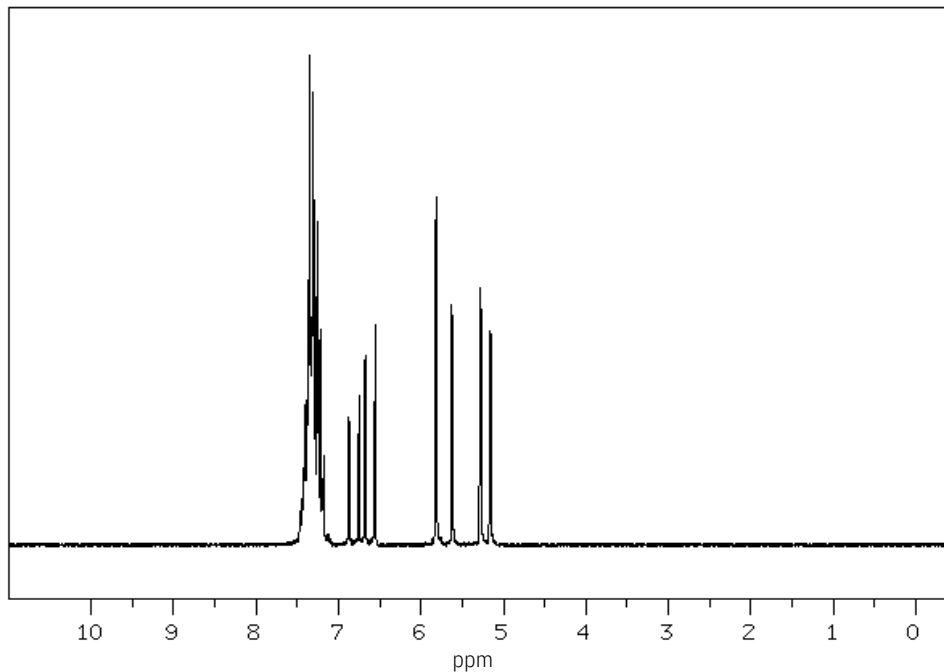


## What metadata is needed?





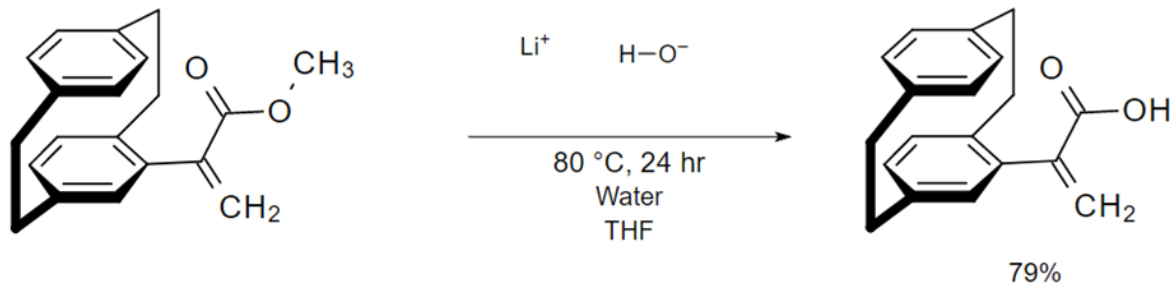
# Descriptive Metadata



Molecule  
Solvent  
Pulse frequency  
Concentration  
Scale with units  
Peaks  
Shifting



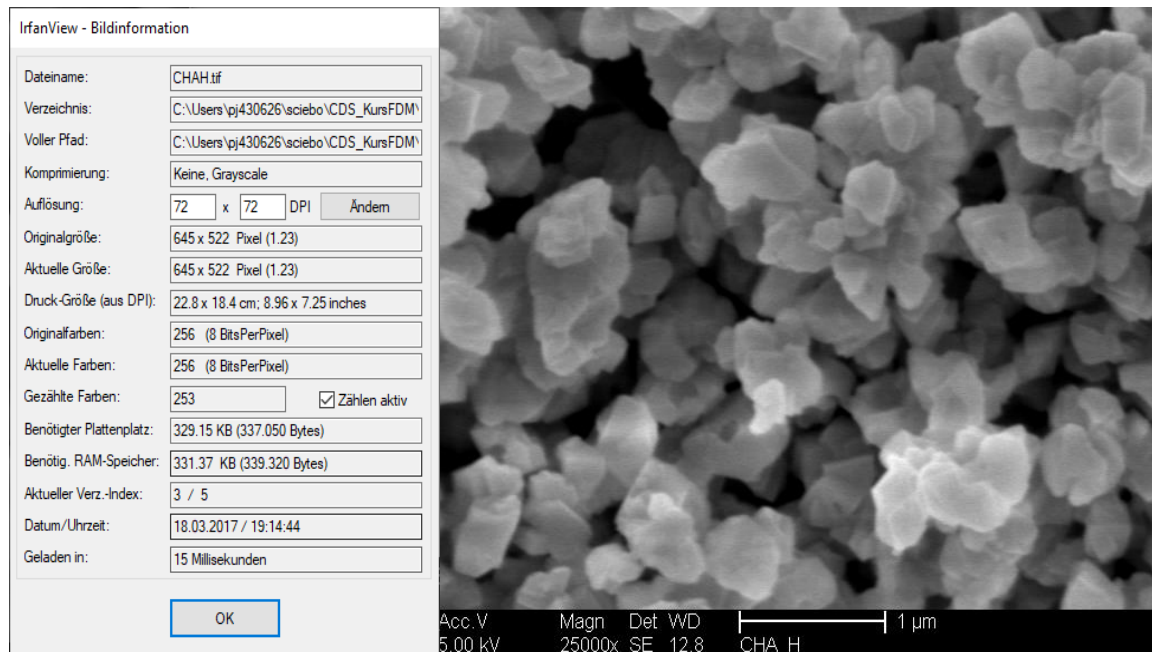
# Descriptive Metadata



Educt Name/ sum  
formula  
Educt structure  
Reactants  
Solvent  
Conditions  
Product Name/ sum  
formula  
Product structure  
Yield



# Structural/ Technical Metadata



Size  
Form  
Specifications of  
objects  
File name  
File path  
Microscope  
Enlargement  
Resolution



# Administrative Metadata



Published on 12-6-2021

Contributor: Christoph Zipfel  
1. Institute of Organic Chemistry, Karlsruhe Institute of Technology, Germany

Author: Christoph Zipfel<sup>1</sup>  
1. Institute of Organic Chemistry, Karlsruhe Institute of Technology, Germany

Reaction DOI: 10.14272/reaction/SA-FUHFF-UHFFADPSC-RKXSPVTUW-UHFFADPSC-NUHFF-NUHFF-NUHFF-ZZZ

ID: CRR-17632

Reference:

79%

Reaction Table	Formula	Density/Molarity	Amount(g)	Volume(ml)	Amount(mmol)	Equiv/Yield
<b>Starting materials</b>						
C20H20O2		-/-	1.800	-	6.157	1.000
<b>Reactants</b>						
Lithium hydroxide	LiOH	-/-	0.737	-	30.783	5.000
<b>Products</b>						
C19H18O2		-/-	1.356	-	4.872	79%
<b>Solvents</b>						
Water	H2O	-	30.000	-	-	50%
THF	C4H8O	-	30.000	-	-	50%

**Description:**  
2-(4-(2-phenylcyclohexyl)acrylic acid methyl ester) (1.80 g, 6.16 mmol, 1.00 equiv) and lithium hydroxide (737 mg, 30.8 mmol, 5.00 equiv) were dissolved in a water/THF mixture (1:1, 60 mL). The mixture was stirred at 80 °C for 24 h. The mixture was acidified with hydrochloric acid (1 M aq. solution) and diluted with dichloromethane (50 mL). The phases were separated, and the aqueous layer was extracted with dichloromethane (2 x 15 mL).

**Additional information for publication and purification details:**

Publication Date  
License  
Contributer  
Author  
Identifier  
Database ID  
Reference  
Reaction Table  
Description  
Additional Information



# Metadata Schema

---



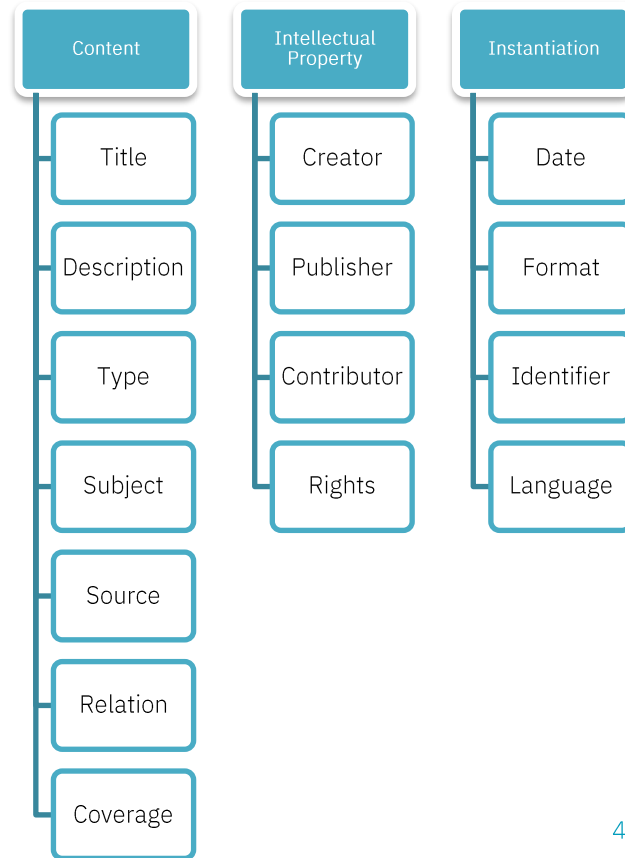
- A Metadata scheme determines and structures the metadata elements.
- Pairs of **metadata element - value**
- Defines input type
- Sets restrictions, such as the use of (controlled) vocabulary or required fields





## International Data Exchange format

- 22 elements –  
15 with an ISO certificate
- Refinements and encoding schemes for subject specific applications
  - <http://www.dublincore.org/>
  - [http://wiki.dublincore.org/index.php/User\\_Guide](http://wiki.dublincore.org/index.php/User_Guide)







Defines core metadata for research data and it is community driven

DataCite term	SRD 78 (Atomic Spec)	comments
Identifier		lacking - but required - looking for a DOI or
Creator	Alexander Kramida, Yuri Ralchenko, and Joseph Reader, Edward B. Saloman	took the 'active' names from website - can also add affiliation and identifier for each name - DataCite allows more than one and recognizes ordering, DC wants "an entity"
Title	NIST Atomic Spectra Database - SRD 78	took name and appended SRD #, like EDI
Publisher	Atomic Spectroscopy Group, Physical Measurement Laboratory, National Institute of Standards and Technology	If it needs to be reduced - could take group and lab away and leave NIST
PublicationYear	2015	
Subject	Atomic spectra, atomic ground state, atomic ionization energy, atomic transition probability, atomic energy levels	EDI entry lists all element names - might be overkill for this purpose
Contributor		Could list "past contributors" or "students contributing to data entry" - how far to go?
Date		Could give more specific date here if that makes sense
Language	en	
ResourceType	Dataset/Atomic Spectra	"Dataset" comes from list, "Atomic Spectra" is free text
AlternateIdentifier	<a href="http://www.nist.gov/pml/data/asd.cfm">http://www.nist.gov/pml/data/asd.cfm</a>	Should add property alternateIdentifierType with value URI?
GeoLocation		
RelatedIdentifier	<a href="http://www.nist.gov/pml/pubs/atspec/index.cfm">http://www.nist.gov/pml/pubs/atspec/index.cfm</a> , <a href="http://www.nist.gov/pml/data/asd_contents.cfm">http://www.nist.gov/pml/data/asd_contents.cfm</a> , <a href="http://physics.nist.gov/PhysRefData/ASD/Html/help.html">http://physics.nist.gov/PhysRefData/ASD/Html/help.html</a> , <a href="http://www.nist.gov/pml/data/asbib/index.cfm">http://www.nist.gov/pml/data/asbib/index.cfm</a>	These are resources listed on the homepage (Intro to atomic spectroscopy, intro to ASD contents, help, bibliography)
Size		
Format		
Version	5	
Rights	<a href="http://www.nist.gov/data/license.cfm">http://www.nist.gov/data/license.cfm</a>	
Description	This database provides access and search capability for NIST critically evaluated data on atomic energy levels, wavelengths, and transition probabilities that are reasonably up-to-date. The Atomic Spectroscopy Data Center has carried out these critical compilations. The Data Center is located in the Physical Measurement Laboratory at the National Institute of Standards and Technology (NIST).	



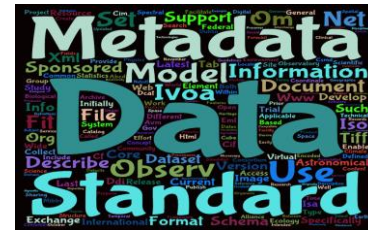
# Find standards, controlled vocabulary, ...

---



- FAIRSharing.org

<https://fairsharing.org/>



- NFDI4Chem – Knowledge Base

[https://knowledgebase.nfdi4chem.de/knowledge\\_base/docs/format\\_standards/](https://knowledgebase.nfdi4chem.de/knowledge_base/docs/format_standards/)

- Metadata Directory of the Research Data Alliance

<http://rd-alliance.github.io/metadata-directory/>



Collecting data

## Electronic lab notebook (ELN)





**Do you use an electronic lab notebook (ELN)?**



- a) Yes, I use an ELN
- b) No, I use a paper notebook
- c) I tried several ELNs but none of them meet my requirements
- d) I use other options to document my research

[www.menti.com](https://www.menti.com)

Code: 8134 3598





## Do you use an electronic lab notebook (ELN)?





# What is an ELN?



## Simple System: Blank Sheet

- Enter text
- Add notes
- Add files as attachments (e.g. images, tables)
- Sharing
- Searching

e.g. Evernote, GoogleDrive, Dropbox, MS Sharepoint

## Electronic Lab Notebook (ELN)

- + Structured metadata in human and machine-readable formats
- + Discipline-specific functions/ editors
- + Rights management
- + Audit trail
- + API (Application Programming Interface)

e.g. Labfolder, RSpace, eLabFTW, Labguru

## Laboratory Information Management System

- + Sample management
- + Instrument integration
- + Electronic signatures
- + Reporting or statistics modules

e.g. Benchling, Starlims, Limesophy



# Advantages of an ELN



## Avoid Data Loss

- Linking experimental descriptions to collected data (analog and digital)
- Secure data storage, backups

## Standardised Documentation

- Structured and standardised collection of metadata
- Generation of interoperable (meta)data



F<sub>indable</sub>



A<sub>ccessible</sub>



I<sub>nteroperable</sub>



R<sub>eusable</sub>

## Knowledge Management

- Data are findable
- Data are accessible
- Data are available, even after change of personnel!

## Publication

- Data provision for publication of research results
- Simple transfer of data to repositories



# A plethora of available ELNs

---



scif<sup>o</sup>rmation



SciNote

... any many more





# How to introduce an ELN



## Needs assessment:

- Analysing current situation (budget, IT resources, software environment)
- Definition of important features
- ELN concept (generic, discipline-specific)
- Drawing on experiences of other research institutions



## Testing the selected products:

- Demo versions or free trial access for individual users
- Testing no more than 2 – 3 ELNs
- In-depth testing using real-life use cases from the lab



## Introducing the chosen ELN:

- Run training courses, training material
- Designate contact persons from the test team
- Continuous mentoring



# How to introduce an ELN



## Needs assessment:

- Analysing current situation (budget, IT resources, software environment)
- Definition of important features
- ELN concept (generic, discipline-specific)
- Drawing on experience



## Testing

- De
- Te
- In-




## Introducing the chosen ELN:

- Training material, run training courses
- Designate contact persons from the test team
- Continuous mentoring

Motivate by demonstrating the benefits!





## ELN Guide


Electronic laboratory notebooks in the context of research data management and good research practice – a guide for the life sciences

**PUBLISSO** an eDIP-Publication Portal for Life Sciences

PUBLISSO ELN filter


Or check out the eln-finder: <https://eln-finder.ulb.tu-darmstadt.de/home>

Product	Last update	Values from provider query	Values from external sources	Cloud usage (Software as a Service)														
				Local use (on-premise)	Adm. Tool	EDR/CFP 2 year 11	Electronic signature	Winning functions	Electronic time stamps	PDF	Mail	IP	MS Office	Data Base	Copy	Isbn		
Axlab	2019-12	Y	N	Y	N	Y	Y	Y	Y	Y	Y	N	N	N	N	Y	Y	
Benchling	2018-11	N	Y	Y	N	Y	N	Y	Y	Y	N	N	N	N	N	N	Y	
Biovia	2019-12	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	
eLABJournal	2019-12	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
eLabFTW	2019-11	Y	N	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	N	
IDBS	2018-11	N	Y	Y	Y	Y	N	Y	Y	N	N	N	N	N	N	N	Y	
LabArchives	2019-12	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	
LabCollector	2019-12	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	N	N	Y	
Labfolder	2019-11	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	Y	Y	
Labguru	2018-11	N	Y	Y	Y	N	Y	N	Y	N	N	N	N	N	N	N	N	
LabWare	2019-11	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
Limnophy	2019-12	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	
NuGenesis	2019-12	Y	N	N	Y	Y	Y	Y	Y	N	N	N	N	N	N	N	N	
openBIS	2019-11	Y	N	N	Y	N	N	Y	N	N	N	Y	N	Y	N	Y	N	
RSpace	2019-11	Y	N	Y	Y	Y	Y	Y	Y	N	N	Y	N	Y	N	Y	Y	
				Pro viding	Security of evidence				File-Export									
Kategorie																		



**Tobac**

**Needs assessment for an electronic lab notebook**



## 1. Lab requirements

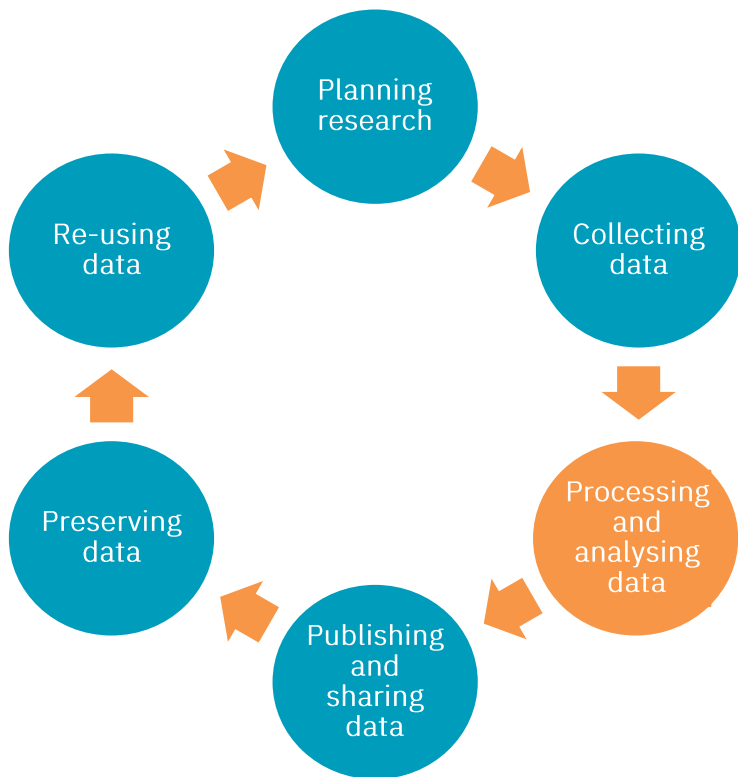
Requirements	yes	no	Additional comments
<p>1. User can create documents and save them for others, both, absolute (similar to a paper) and relative</p>	<input type="checkbox"/>	<input type="checkbox"/>	
<p>Ability to import data as a file (like csv, excel, graphs, JSA, etc. available, JSA, xlsx, vba, and formatted)</p>	<input type="checkbox"/>	<input type="checkbox"/>	Not clear if this feature is required
<p>Ability to import and process standard formats, formatting (ASTA, xlsx, Comcat, jls, xls, etc.), data generated by measuring instruments</p>	<input type="checkbox"/>	<input type="checkbox"/>	Not clear if this feature is required
<p>Special data entry and processing editing, scientific calculation (like graph, tables, etc.), ability to add pictures, formulas, attributes (mobile, chemical formula, graphs and data included)</p>	<input type="checkbox"/>	<input type="checkbox"/>	Not clear if this feature is required
<p>Ability to make your own templates (e.g. for standardised test, experiments, SOPs or projects)</p>	<input type="checkbox"/>	<input type="checkbox"/>	Possible future feature update
<p>Ability to import templates or use pre-generated templates for experiments and data entry (e.g. for standardised test, experiments, SOPs, chemistry, analytical chemistry, materials research)</p>	<input type="checkbox"/>	<input type="checkbox"/>	Possible future feature update, might be required in the future
<p>Ability to export data as a pdf, excel, format, etc. (xls, doc)</p>	<input type="checkbox"/>	<input type="checkbox"/>	Not clear if this feature is required
<p>Graphical representation of attributes, processes, graphs</p>	<input type="checkbox"/>	<input type="checkbox"/>	
<p>Ability to create links to files outside the E-LN, e.g. to avoid duplication of stored data, to show connections</p>	<input type="checkbox"/>	<input type="checkbox"/>	
<p>Sample management, creating material databases within the E-LN</p>	<input type="checkbox"/>	<input type="checkbox"/>	
<p>Standard interfaces, e.g. connection to a laboratory information management system (LIMS), data exchange with LIMS, data import from LIMS (cloud)</p>	<input type="checkbox"/>	<input type="checkbox"/>	Not clear if this feature is required
<p>Calculation functions, individual units and groups, calculation with units (addition, subtraction, multiplication, division, etc.)</p>	<input type="checkbox"/>	<input type="checkbox"/>	Not clear if this feature is required
<p>Results, time management, annotation and rights</p>	<input type="checkbox"/>	<input type="checkbox"/>	

The E-LN must at least document the following functions, SOPs and workflows:



# Processing and analyzing data

---

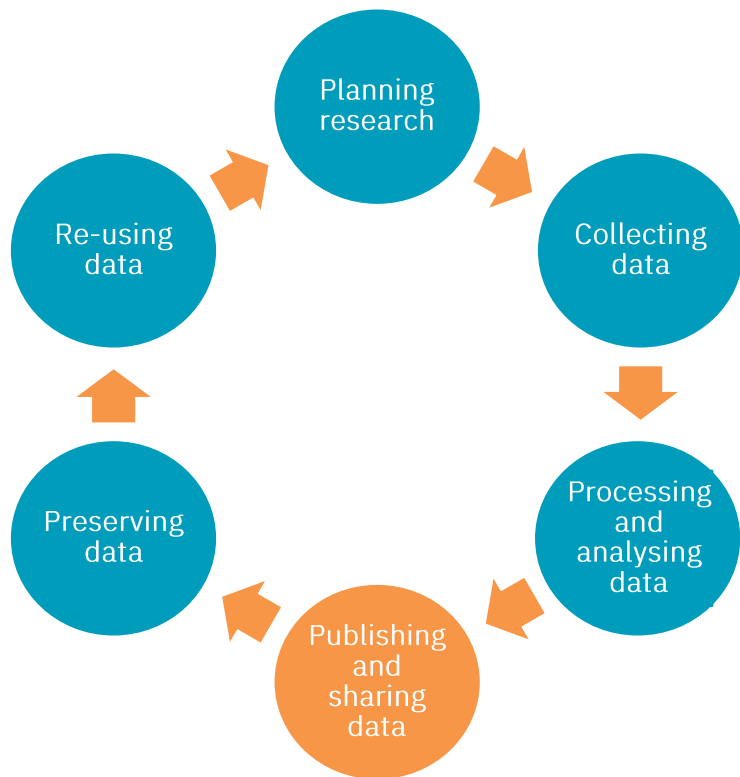


- Interpret data
- Use the data in scientific publications
- Backing-up data and preparing it for data storage
- Enabling data exchange during the project



# Publishing and sharing data

---



- Share, disseminate, publish data
- Making data known and findable (catalogues)
- Making data citable (DOI)
- Issue licenses
- Access control (if necessary)



# Reasons for Data Publication

---



- Advantages

- Better (re-)use of your research data
- Increased visibility, openness, transparency and accountability

*“As open as possible, as closed as necessary”*

- Clearly citable with persistent identifiers (e.g. DOI)
- Long-term availability
- Data production as an independent scientific result (e.g. Data Citation Index)



- Requirements

- Of funders
- Of scientific journals







## How do you publish your data?



[www.menti.com](https://www.menti.com)  
Code: 8134 3598

- a) Research article + supplement
- b) Article in data journal
- c) Repository
- d) Website (personal or institutional)
- e) I only ever publish data if specifically required by journal or reviewers
- f) I never publish data
- g) Other options



# Mentimeter: Data Publication



Mentimeter

## How do you publish your data?





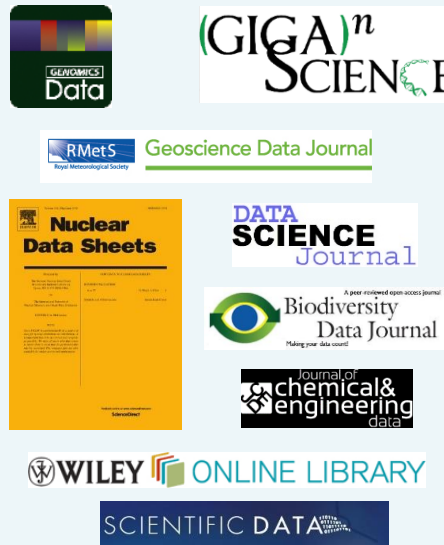
# How can I publish my data?



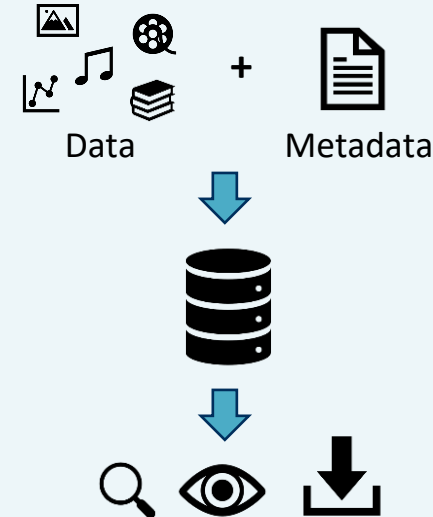
## Supplement of an article in scientific journal



## Data Journal



## Data Repository



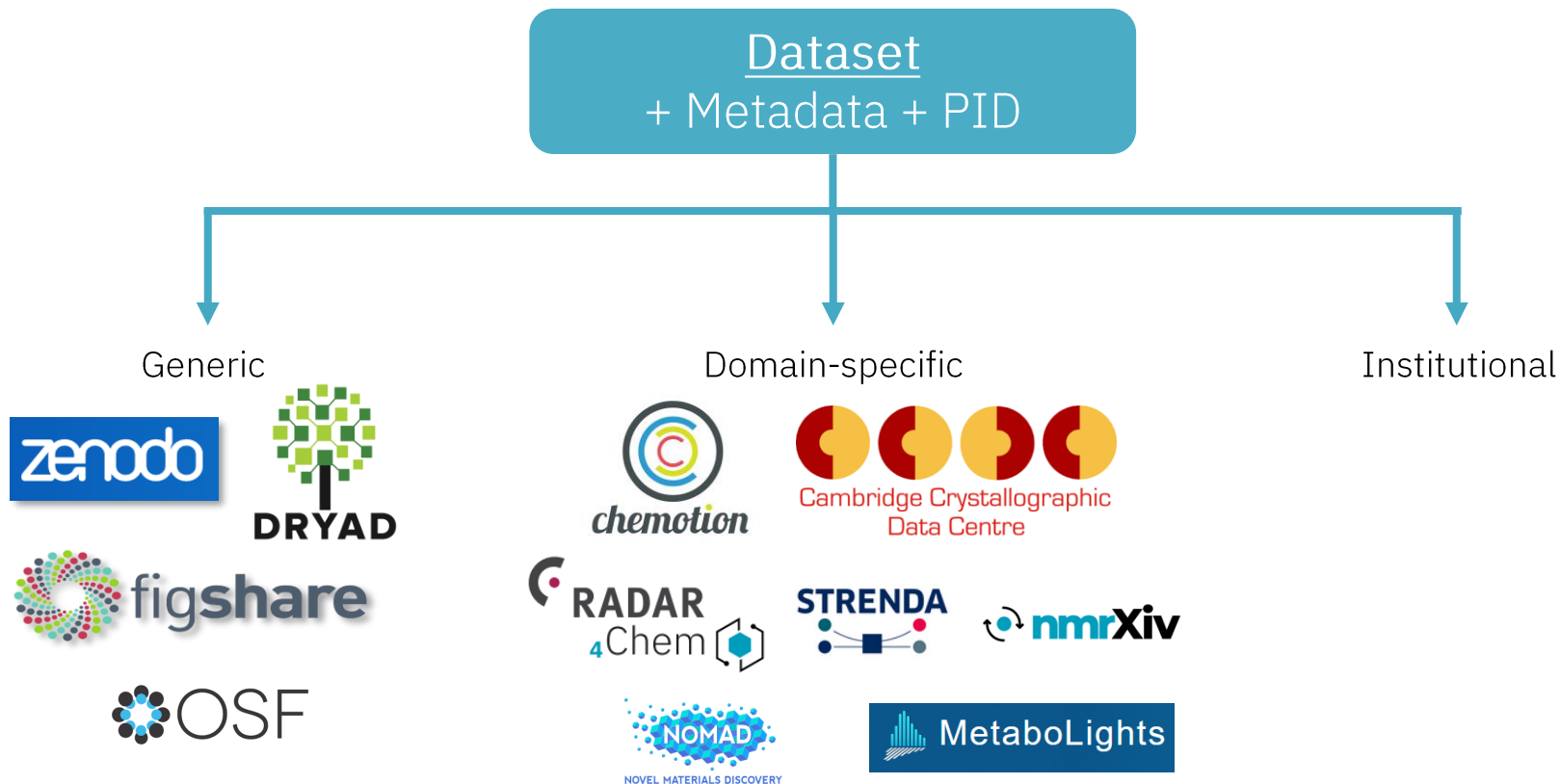


Publishing and sharing data

# Repositories



# Data Repositories







## Further reading: Publishing

---

- Persistent identifiers (PIDs)

[https://knowledgebase.nfdi4chem.de/knowledge\\_base/docs/pid/](https://knowledgebase.nfdi4chem.de/knowledge_base/docs/pid/)

- Data availability statements

[https://knowledgebase.nfdi4chem.de/knowledge\\_base/docs/data\\_availability\\_statement/](https://knowledgebase.nfdi4chem.de/knowledge_base/docs/data_availability_statement/)

- Best Practices

[https://knowledgebase.nfdi4chem.de/knowledge\\_base/docs/best\\_practice/](https://knowledgebase.nfdi4chem.de/knowledge_base/docs/best_practice/)





## Further reading: Repositories

---

- Repositories

[https://knowledgebase.nfdi4chem.de/knowledge\\_base/docs/repositories/](https://knowledgebase.nfdi4chem.de/knowledge_base/docs/repositories/)

- How to choose the right repository

[https://knowledgebase.nfdi4chem.de/knowledge\\_base/docs/choose\\_repository/](https://knowledgebase.nfdi4chem.de/knowledge_base/docs/choose_repository/)

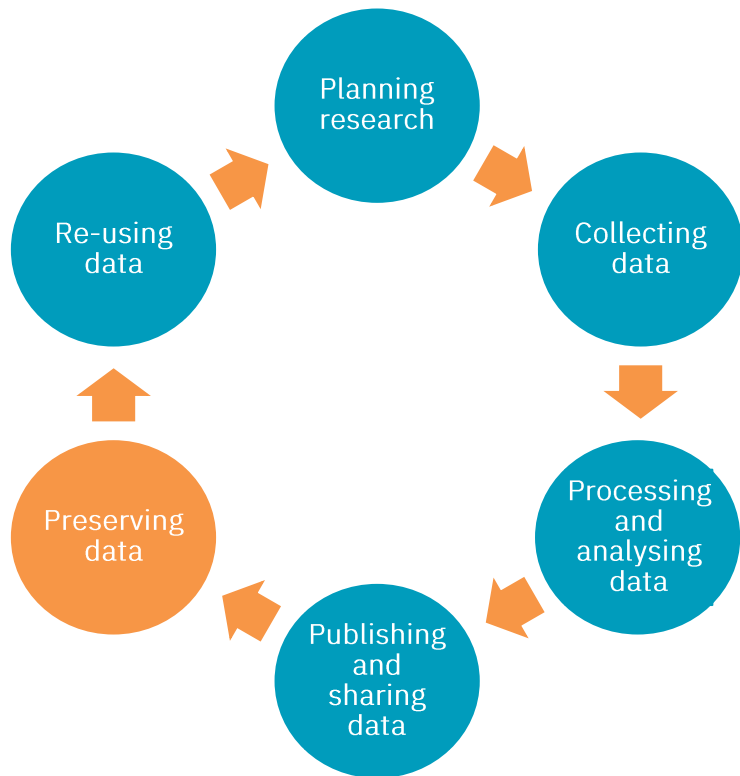
- Repository choosing tools

[re3data.org](https://re3data.org)

[fairsharing.org](https://fairsharing.org)



# Preserving data



Preserving ≠ Not deleting

**Aim:** Keep data, metadata and documentation safe, available and re-usable in the long term

**Risks:**

- Corruption of data or storage medium
- Outdated file formats
- Lack of metadata and/or documentation
- Data not findable or accessible





# Steps of data preservation

---

- What to keep? – Data selection
  - Define selection criteria
- Where to preserve data? - Suitable location or medium
  - External data repository or archive versus institutional infrastructure
- Prepare data and files for the preservation
  - Organized files and suitable file formats
  - Including metadata, documentation, access rights and conditions
- Perform periodic checks of the preserved data



# Different types of data preservation

---



- Access for the data producer
- **Active data**
- Data kept short-term
- Purpose: Protection and discovery

## Backup

- Access for the data producer
- **Final Data**
- Data kept long-term
- Purpose: Preservation of information

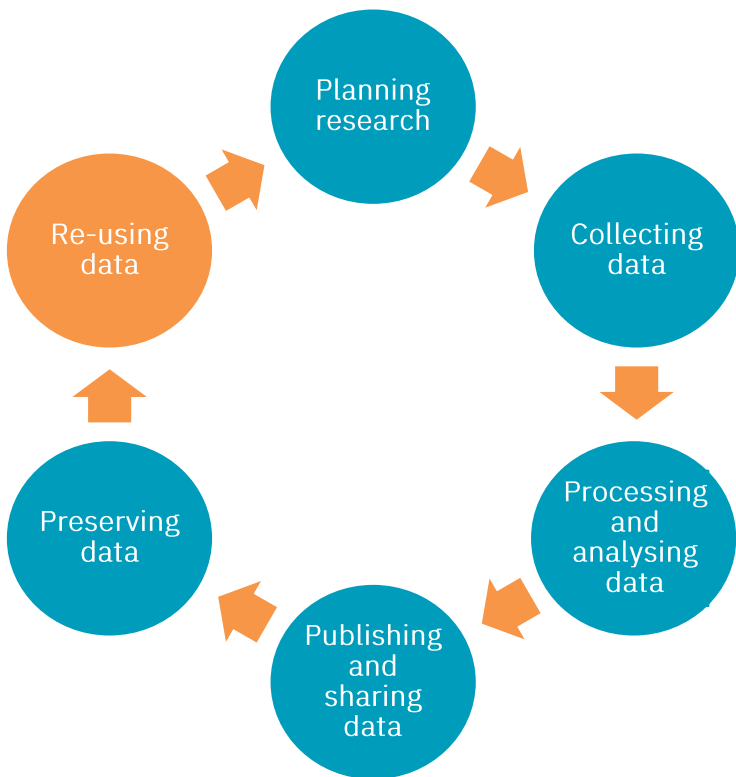
## Archiving

- Open access
- **Final Data**
- Data persistence depends on the publishing institution
- Purpose: Provision of data for reuse

## Publication



# Re-using data



What is the potential of your research data?

- Conduct further research with the data
- Put data into new contexts, using data in an interdisciplinary way
- Big Data applications
- Review, critique and discuss research findings
- Teaching and learning
- Citing research data



**NFDI4Chem**

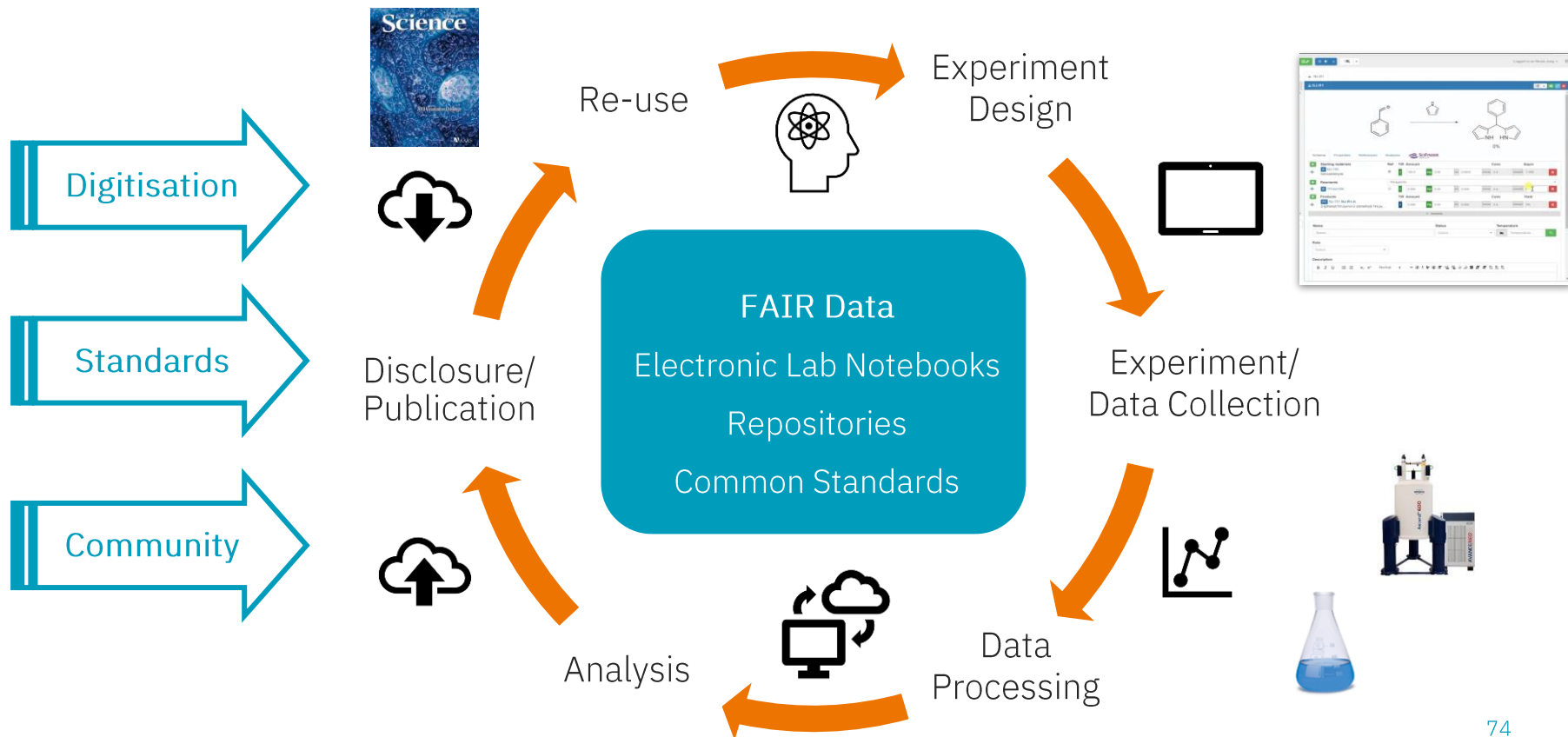


**NFDI<sub>4</sub>Chem**

ENHANCE  
YOUR  
DATA.



# NFDI4Chem: Vision





# NFDI4Chem: Strategy



## Portal

Helpdesk / Support

Knowledge Base

Teaching / Training

Search

### SmartLab

ELN

SW Tools

Devices/API

Data

### Repositories

Publication

Archiving

Software

Standard

Legal / Policies

Terminology





- Helpdesk  
[helpdesk@nfdi4chem.de](mailto:helpdesk@nfdi4chem.de)
- Knowledge Base for RDM in chemistry  
[https://knowledgebase.nfdi4chem.de/knowledge\\_base/](https://knowledgebase.nfdi4chem.de/knowledge_base/)
- Events such as workshops, Q&A, Stammtisch  
<https://www.nfdi4chem.de/index.php/events/>





**NFDI<sub>4</sub>Chem**

ENHANCE  
YOUR  
DATA.

**Thank you for your attention!**



# Question & Answer Session

