

# ARC and elabFTW Introduction IMET

January, 2025

Sabrina Zander (MibiNet)  
Stella Eggels (DataPLANT)

# Tentative agenda

## Morning

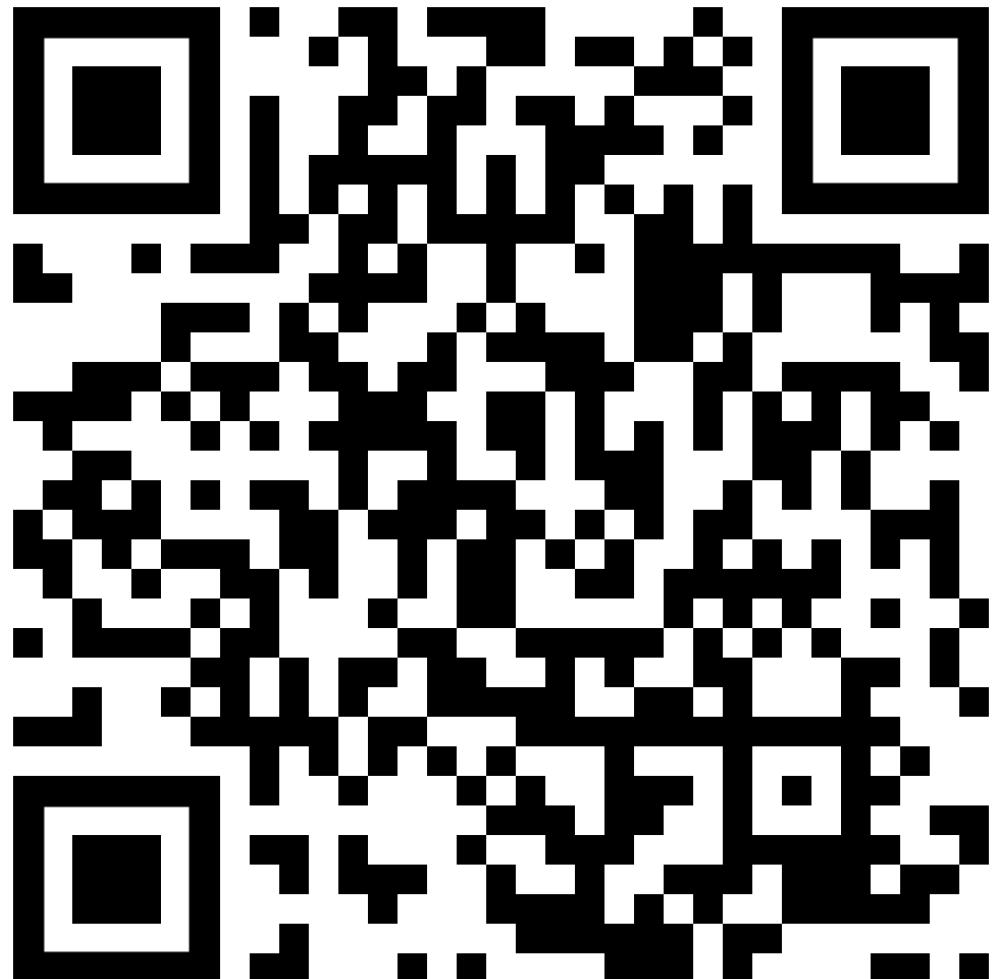
Time	Topics
09:30 - 09:45	Welcome, Intro RDM and NFDI
09:45 - 10:15	Overview DataPLANT, and DataPLANT Tool chain
10:15 - 10:45	Hands-on Start Here Guide Part I
10:45 - 11:00	<i>Short break</i> ☕
11:00 - 11:30	Metadata Intro
11:30 - 12:15	Hands-on Start Here Guide Part II

## Afternoon

Time	Topics
12:15 - 13:15	<i>Lunch break</i> 🍕
13:15 - 13:45	DataHUB Intro
13:45 - 14:15	Hands-on Start Here Guide Part III
14:15 - 15:00	Q & A ARC
14:15 - 15:00	<i>Short break</i> ☕
15:00 - 16:00	elabFTW Intro
16:00 - 16:30	elab2ARC tool
16:30 - 17:30	Q & A and wrap-up

# House-keeping

- Use the [pad](#) to raise questions and feedback
- Copy / paste links (hands-on)



# Training Materials

Slides are shared via [DataPLANT Training Material](#)

# Resources – join the open source movement

## DataPLANT (nfdi4plants)

DataPLANT Website: <https://nfdi4plants.org/>

Knowledge Base: <https://nfdi4plants.org/nfdi4plants.knowledgebase/>

DataHUB: <https://git.nfdi4plants.org>

GitHub: <https://github.com/nfdi4plants>

HelpDesk: <https://helpdesk.nfdi4plants.org>

 You can help us by raising issues, bugs, ideas...

 NEW! ARC website: <https://arc-rdm.org>

# Continuous support

Data managers in Düsseldorf, Cologne, Jülich and close by (CEPLAS, MibiNet, TRR341) offer support.

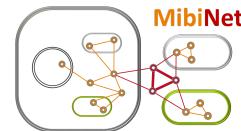
1. Slack Workspace for ad hoc support
2. Monthly user meeting (2nd Friday of the month)

→ [Details](#).



trr\_341

plant  
ecological  
genetics



# Introduce yourself

- Lab
- MibiNet / CEPLAS / TRR / ?
- Used code / programming language before
- Experience with Git / GitLab / GitHub?
- Have an ORCID
- My motivation / expectation
- My favorite lab assay

# Let's draw a typical lab workflow



# BYOD – Goals

- First few steps into ARC ecosystem
- Move existing datasets into ARCs
- Share them via the DataHUB
- Annotate with metadata
- Introduction into elabFTW
- elab2ARC tool

# FAIR data stewardship

- Findable
- Accessible
- Interoperable
- Reusable

<https://doi.org/10.1038/sdata.2016.18>

[nature](#) > [scientific data](#) > [comment](#) > [article](#)

[Open Access](#) | [Published: 15 March 2016](#)

## The FAIR Guiding Principles for scientific data management and stewardship

[Mark D. Wilkinson](#), [Michel Dumontier](#), [IJsbrand Jan Aalbersberg](#), [Gabrielle Appleton](#), [Myles Axton](#), [Arie Baak](#), [Niklas Blomberg](#), [Jan-Willem Boiten](#), [Luiz Bonino da Silva Santos](#), [Philip E. Bourne](#), [Jildau Bouwman](#), [Anthony J. Brookes](#), [Tim Clark](#), [Mercè Crosas](#), [Ingrid Dillo](#), [Olivier Dumon](#), [Scott Edmunds](#), [Chris T. Evelo](#), [Richard Finkers](#), [Alejandra Gonzalez-Beltran](#), [Alasdair J.G. Gray](#), [Paul Groth](#), [Carole Goble](#), [Jeffrey S. Grethe](#), [Jaap Heringa](#), [Peter A.C. 't Hoen](#), [Rob Hooft](#), [Tobias Kuhn](#), [Ruben Kok](#), [Joost Kok](#), [Scott J. Lusher](#), [Maryann E. Martone](#), [Albert Mons](#), [Abel L. Packer](#), [Bengt Persson](#), [Philippe Rocca-Serra](#), [Marco Roos](#), [Rene van Schaik](#), [Susanna-Assunta Sansone](#), [Erik Schultes](#), [Thierry Sengstag](#), [Ted Slater](#), [George Strawn](#), [Morris A. Swertz](#), [Mark Thompson](#), [Johan van der Lei](#), [Erik van Mulligen](#), [Jan Velterop](#), [Andra Waagmeester](#), [Peter Wittenburg](#), [Katherine Wolstencroft](#), [Jun Zhao](#) & [Barend Mons](#) 

— Show fewer authors

[Scientific Data](#) 3, Article number: 160018 (2016) | [Cite this article](#)

# The FAIR principles

<p><b>Findable</b> <b>Accessible</b> <b>Interoperable</b> <b>Reusable</b></p>	<p><b>Easier collaboration &amp; sharing</b></p> <pre>graph TD; A(( )) --&gt; B(( )); A(( )) --&gt; C(( )); A(( )) --&gt; D(( )); B(( )) --&gt; E(( )); C(( )) --&gt; E(( )); D(( )) --&gt; E(( ));</pre>	<p><b>Increased findability and visibility</b></p>	<p><b>Reproducibility</b></p>
<p><b>Added-value to the research community</b></p> <p>nfdi      NCBI EMBL-EBI</p>	<p><b>Compliance with funding policies</b></p> <pre>graph LR; A[Checkmark Document] --- B[DFG]; A --- C[EU Flag]</pre>	<p><b>Receive due credit</b></p> <pre>graph LR; A[Reuse] --- B[Citations]; A --- C[FAIR]</pre>	<p><b>Saves time &amp; workload</b></p> <pre>graph LR; A[FAIR] --- B[Time wasted]</pre>

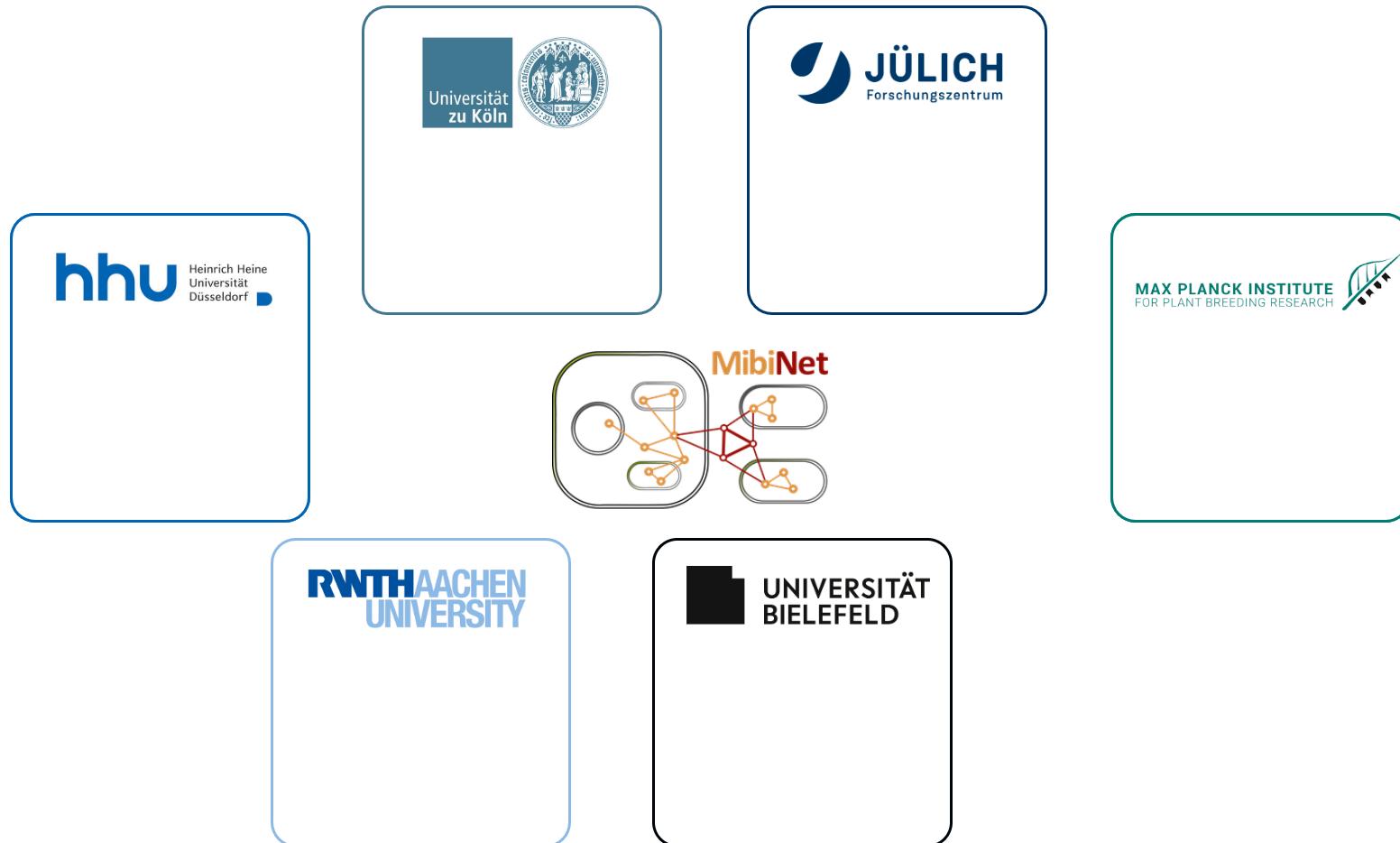
# Is your data FAIR?

Findable | Accessible | Interoperable | Reusable

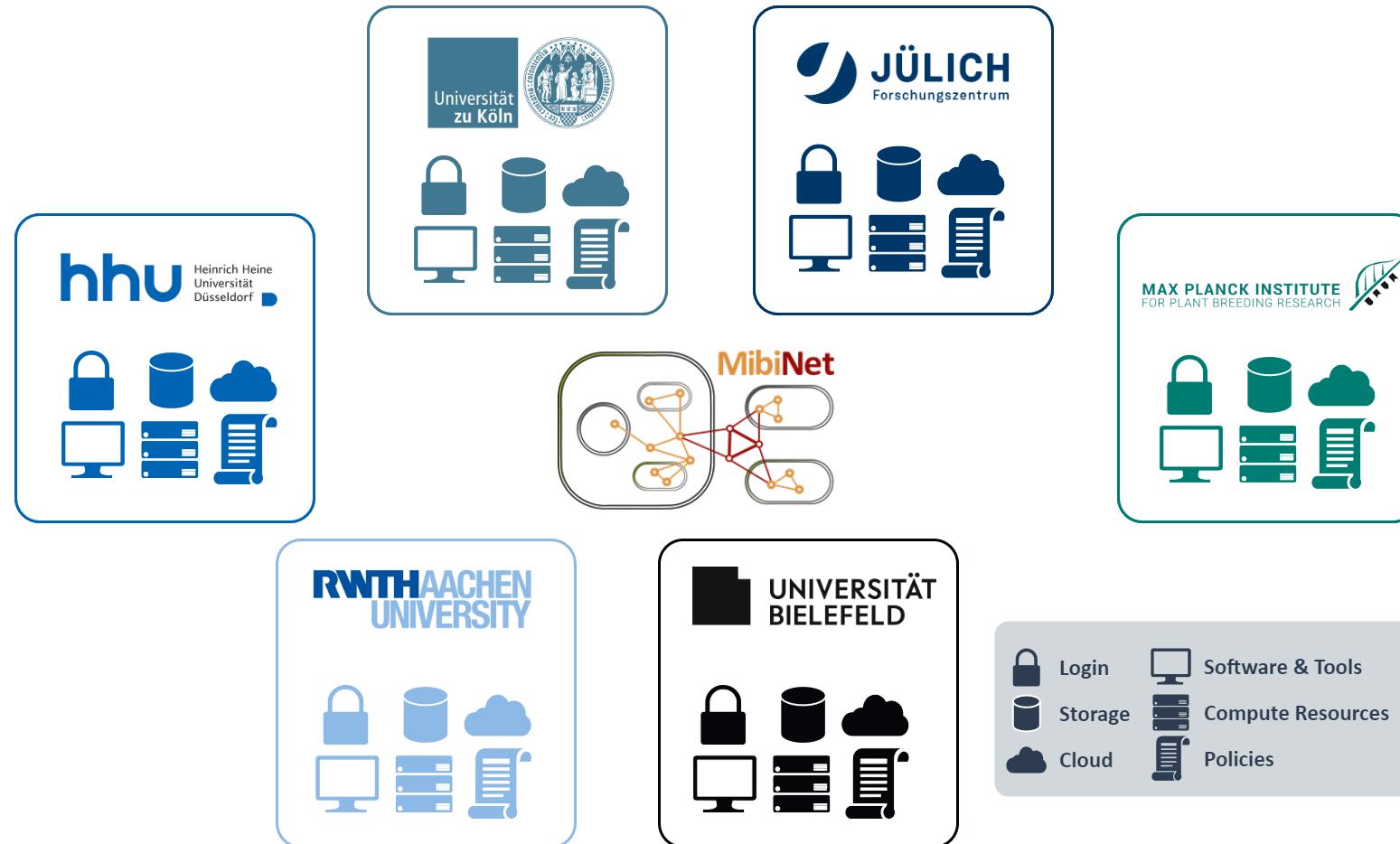
- Where do you store your data?
- How do you annotate your data?
- How do you share your data?
- What tools do you use to analyse your data?
- How do you reuse other people's data?



# MibiNet – One SFB, six locations

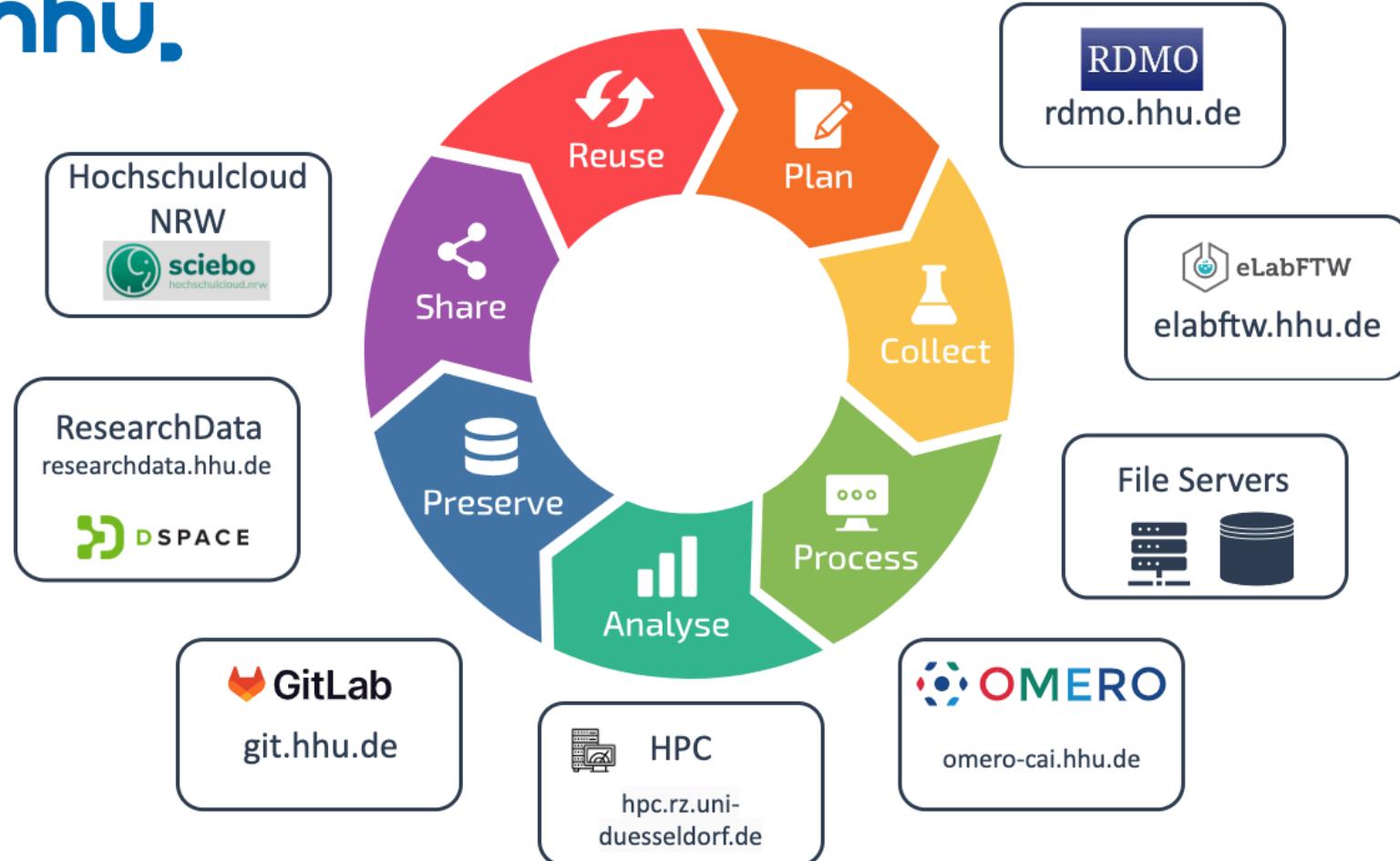


# Data silos impede collaboration

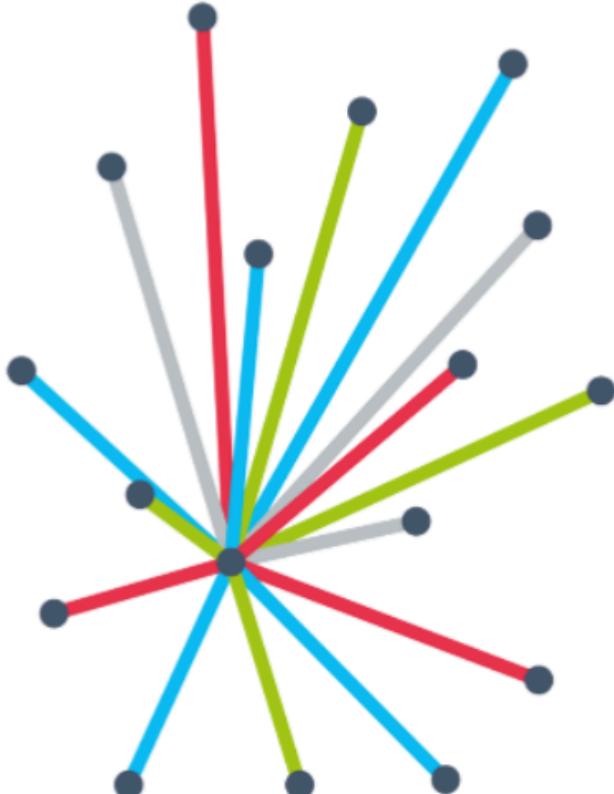


# Missing interfaces impede collaboration

hhu.



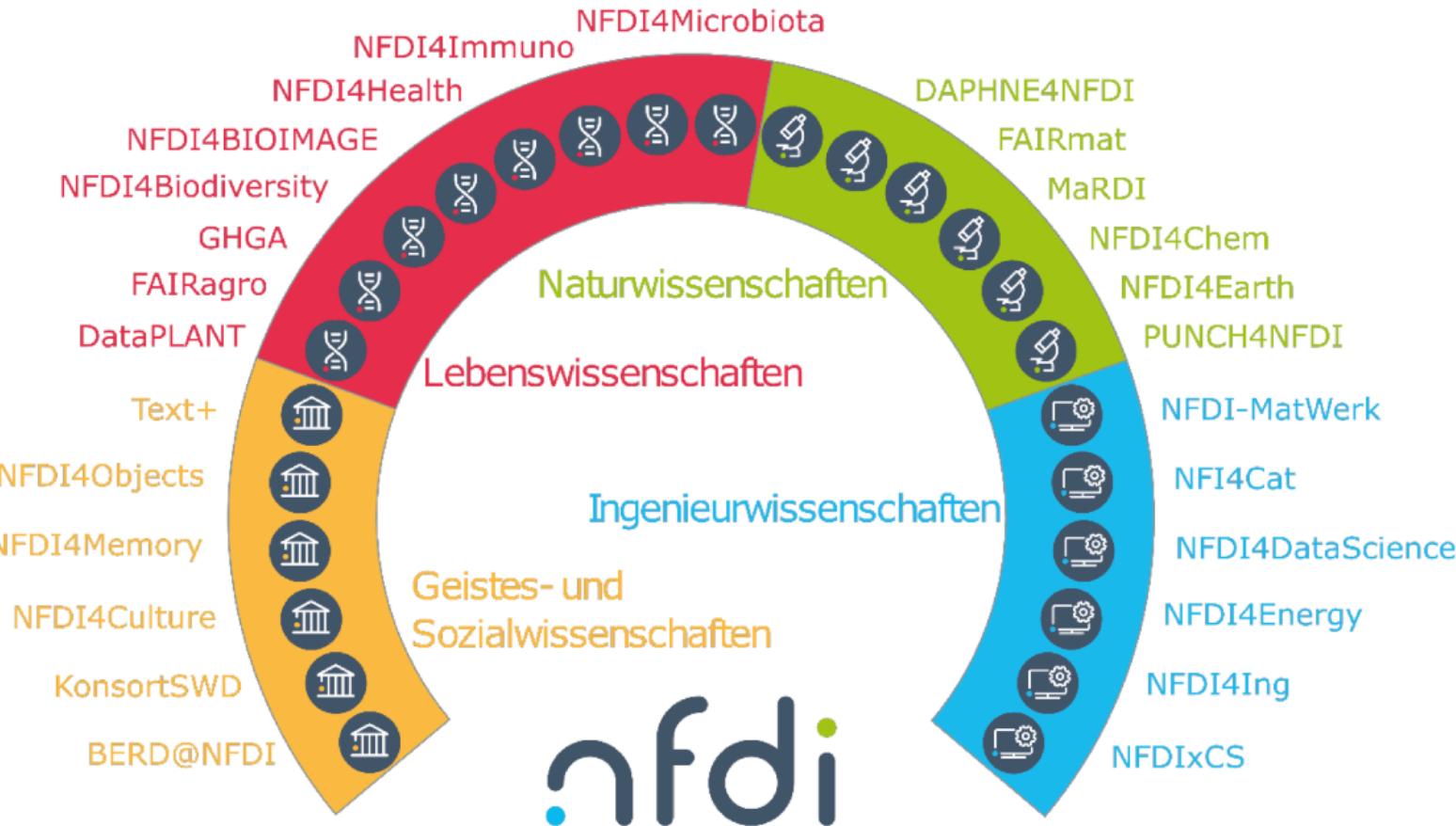
# NFDI – Nationale Forschungsdaten Infrastruktur



The aim of the national research data infrastructure (NFDI) is to systematically manage scientific and research data, provide long-term data storage, backup and accessibility, and network the data both nationally and internationally.

[nfdi.de](http://nfdi.de)

# NFDI – A community-targeted approach for RDM

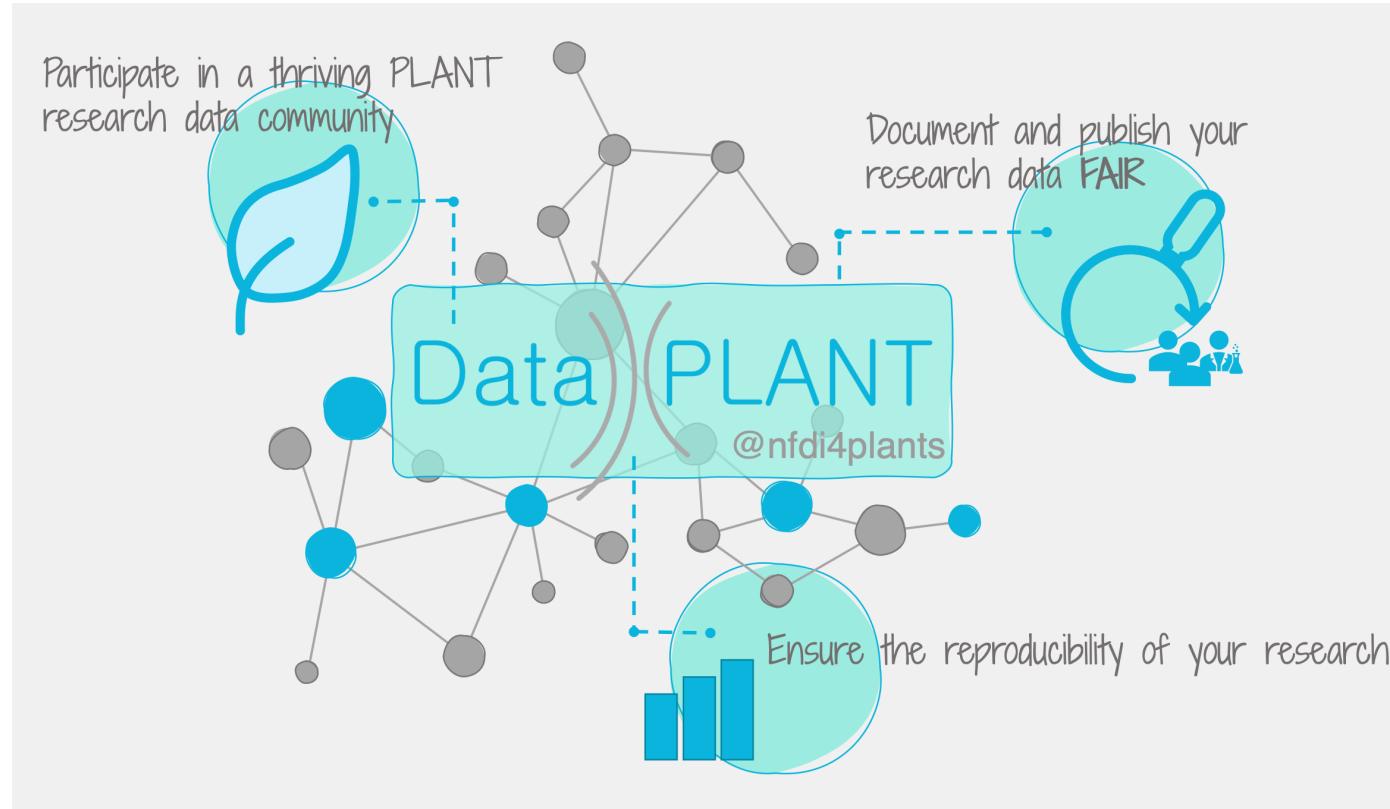


[nfdi.de](http://nfdi.de)

# MibiNet connection to NFDIs

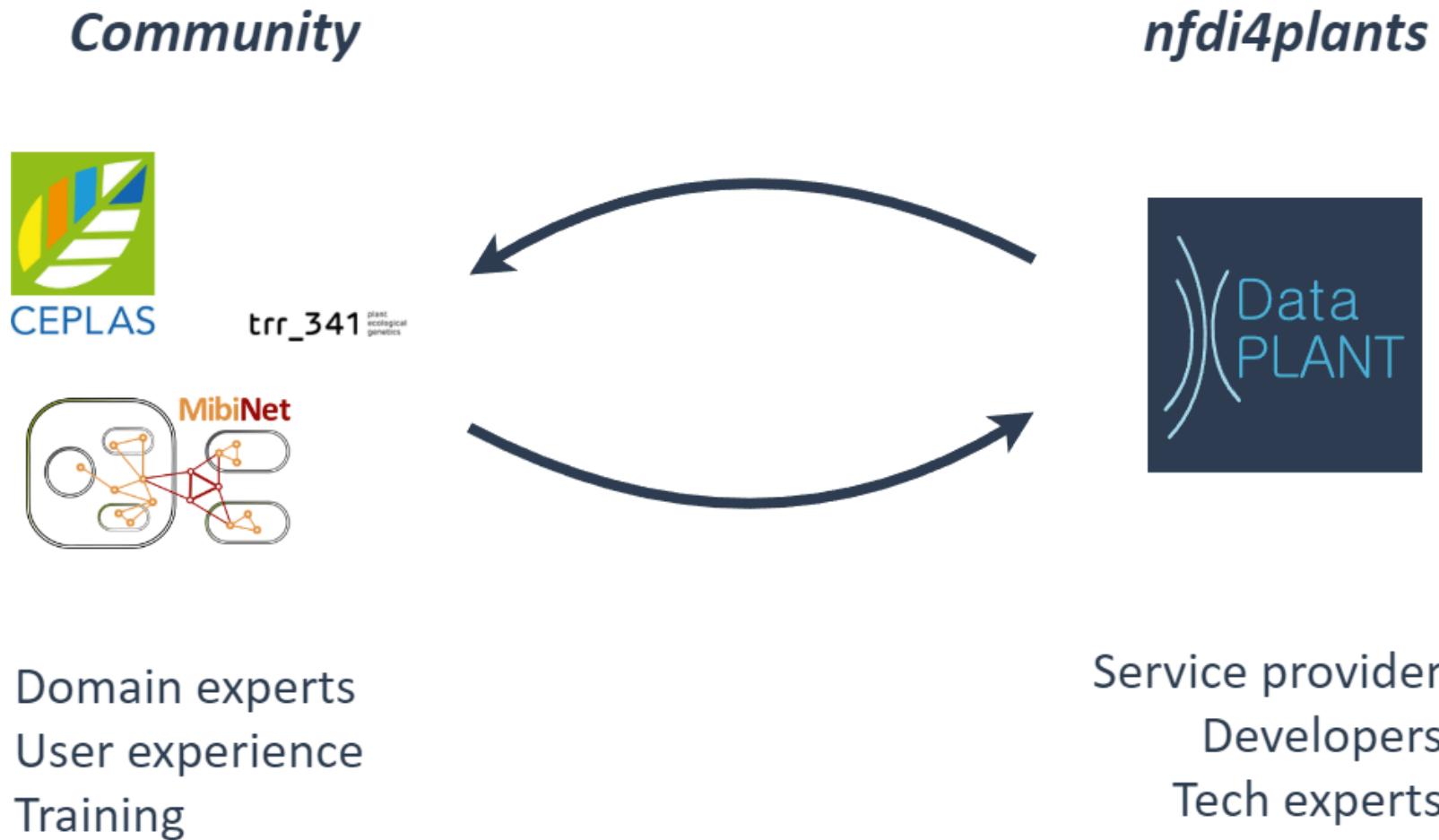


# DataPLANT – NFDI4plants

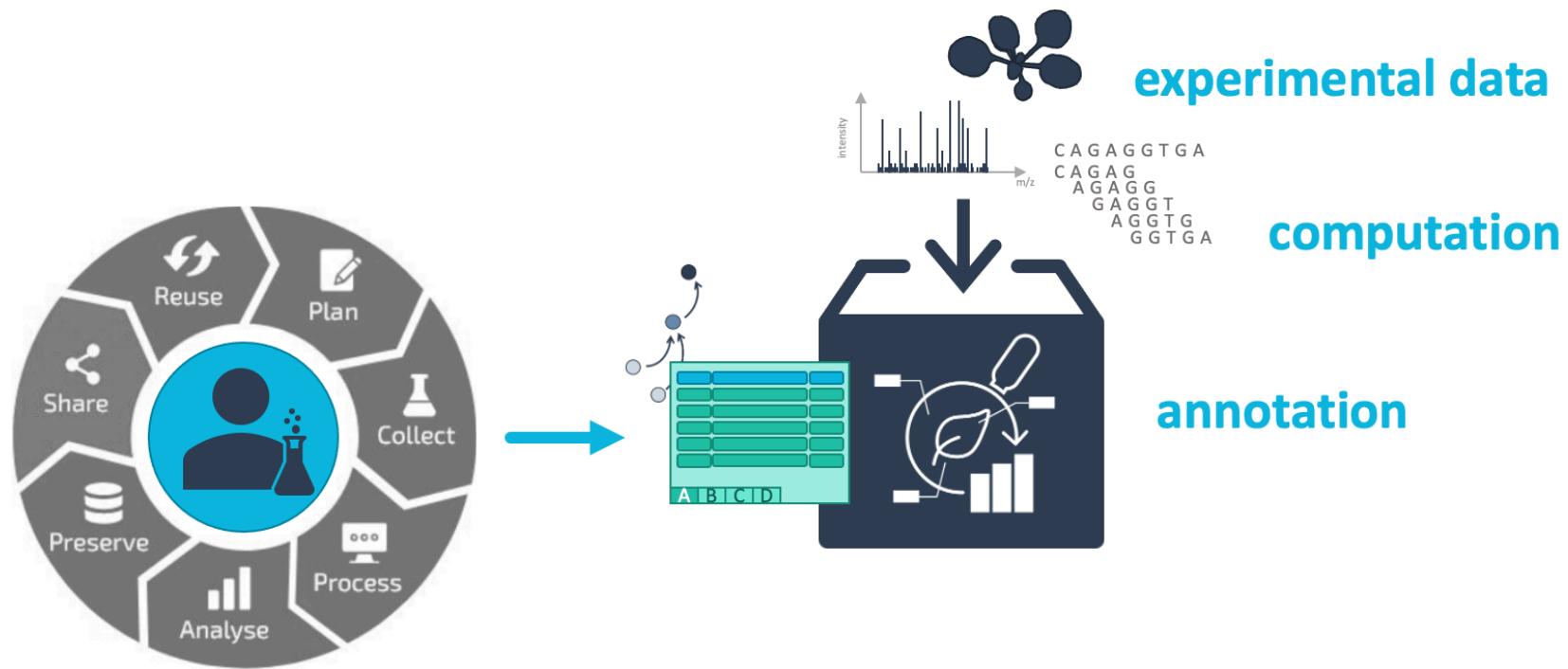


- <https://nfdi4plants.org>
- <https://arc-rdm.org>

# Data Stewardship between DataPLANT and the community

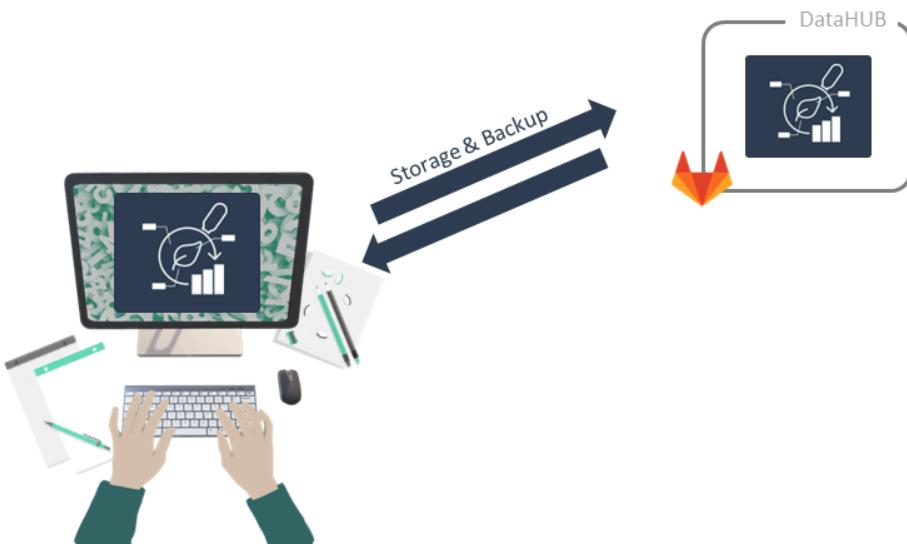


# Annotated Research Context (ARC)

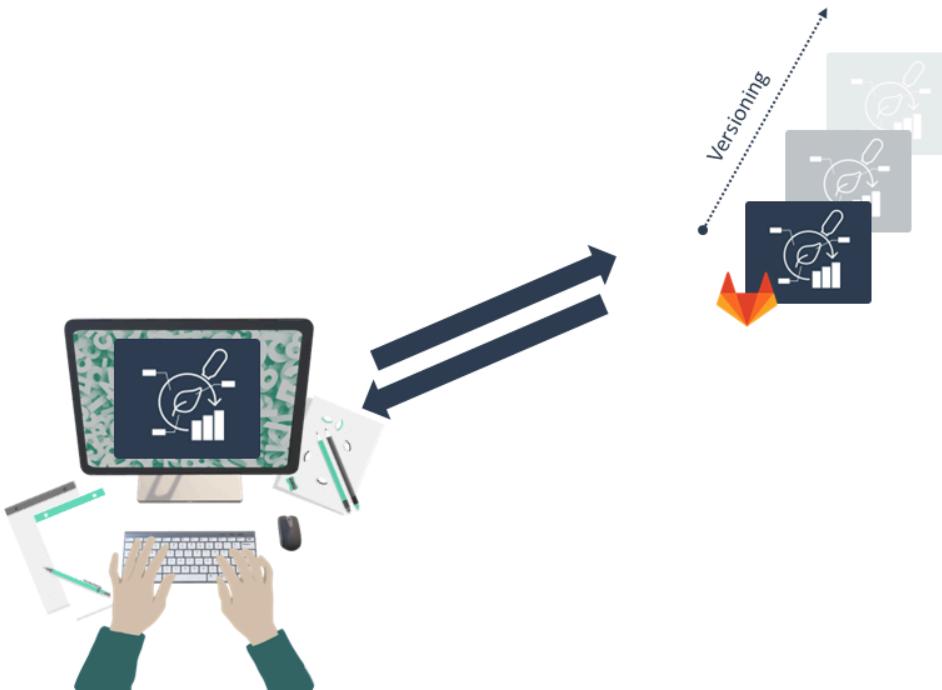


Your entire investigation in a single unified bag

# You can store your ARC in the DataHUB



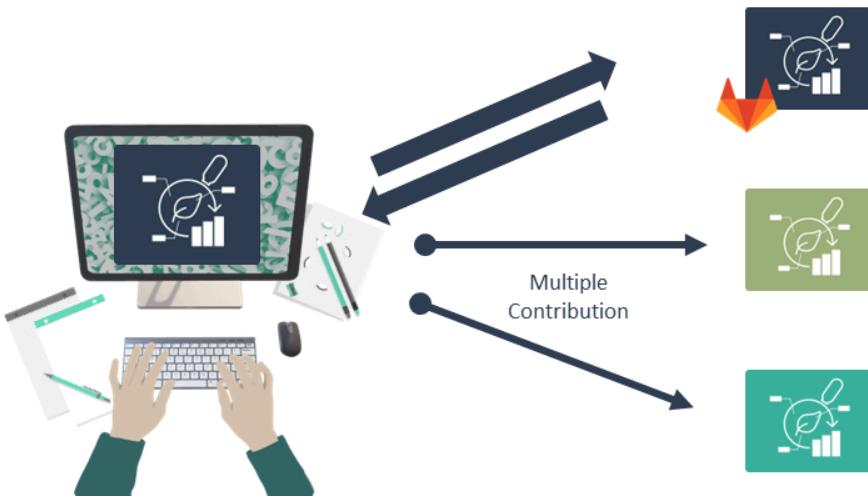
# ARCs are versioned



# You can invite collaborators



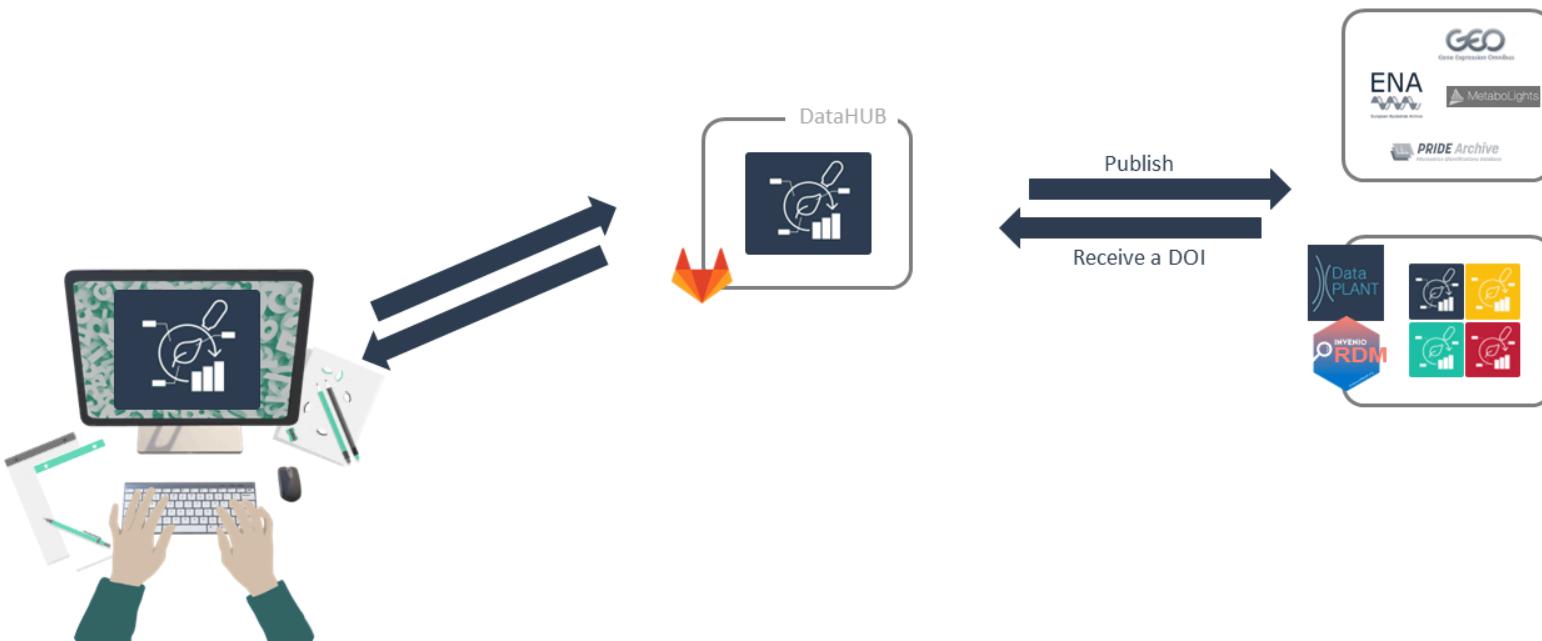
# Collaborate and contribute



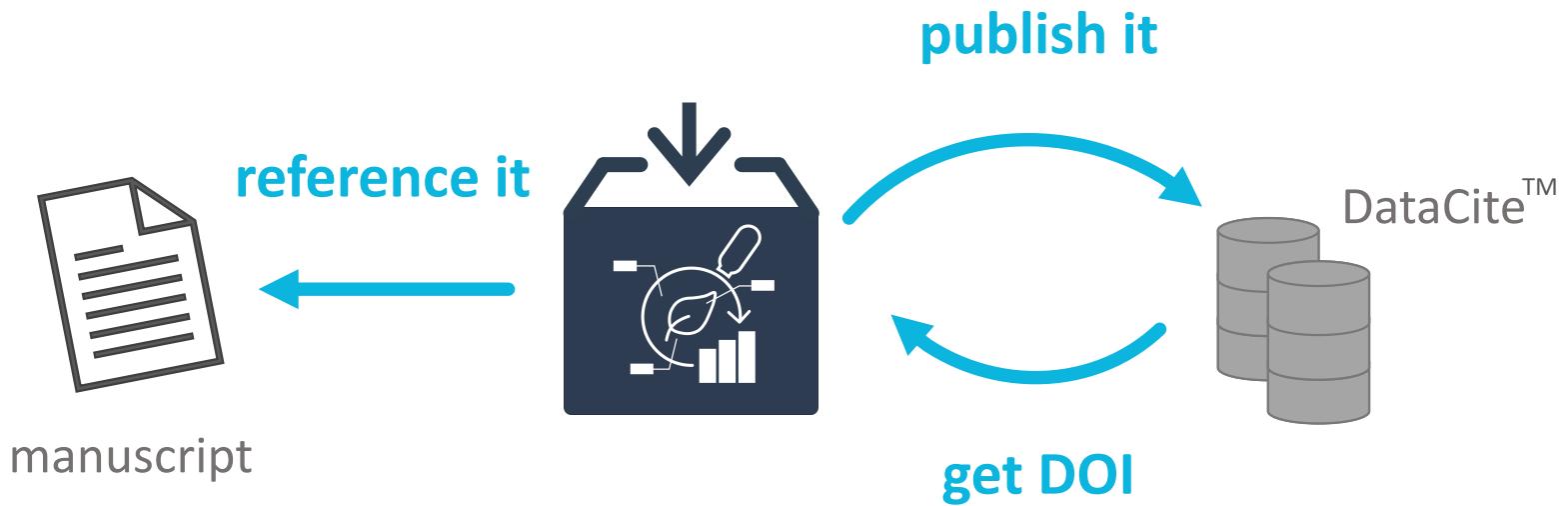
# Reuse data in ARCs



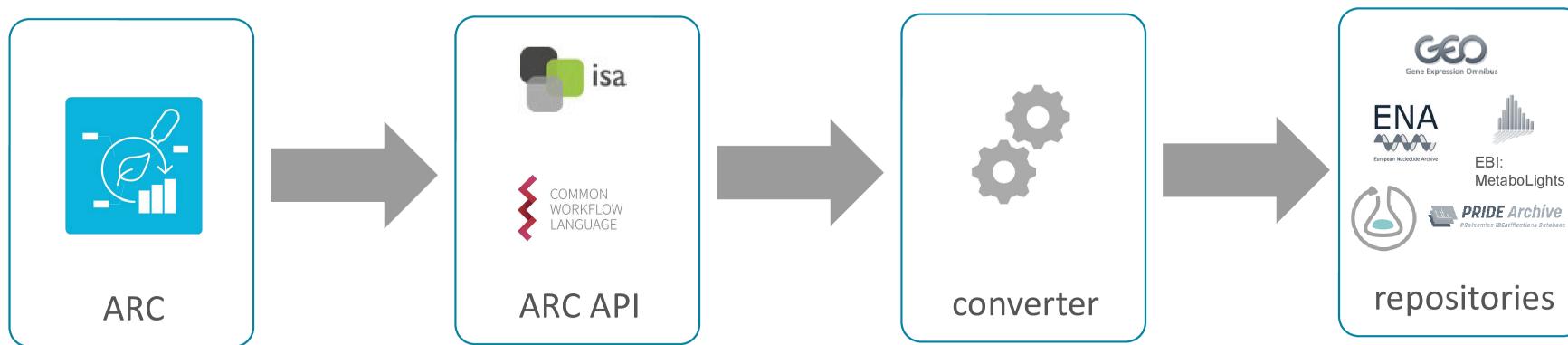
# Publish your ARC



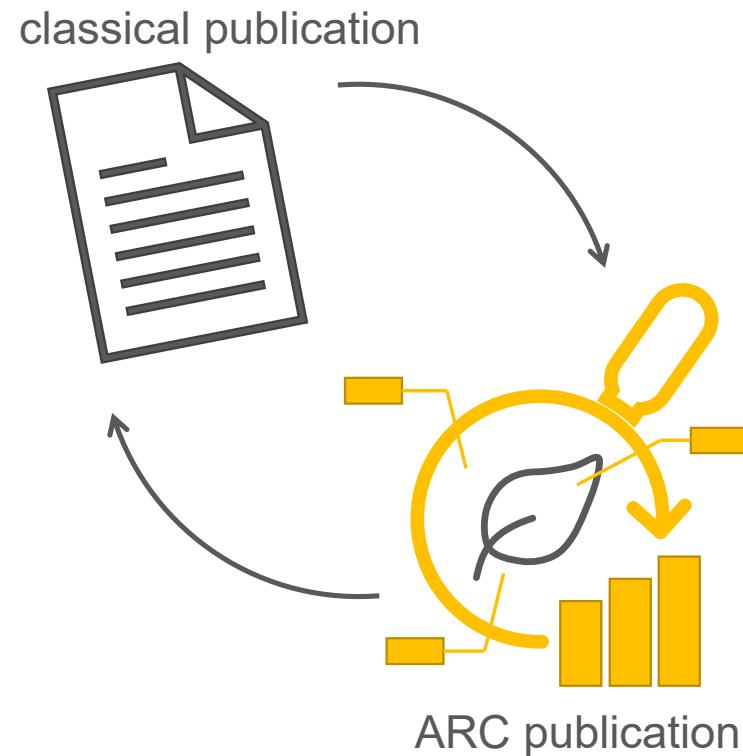
# Publish your ARC, get a DOI



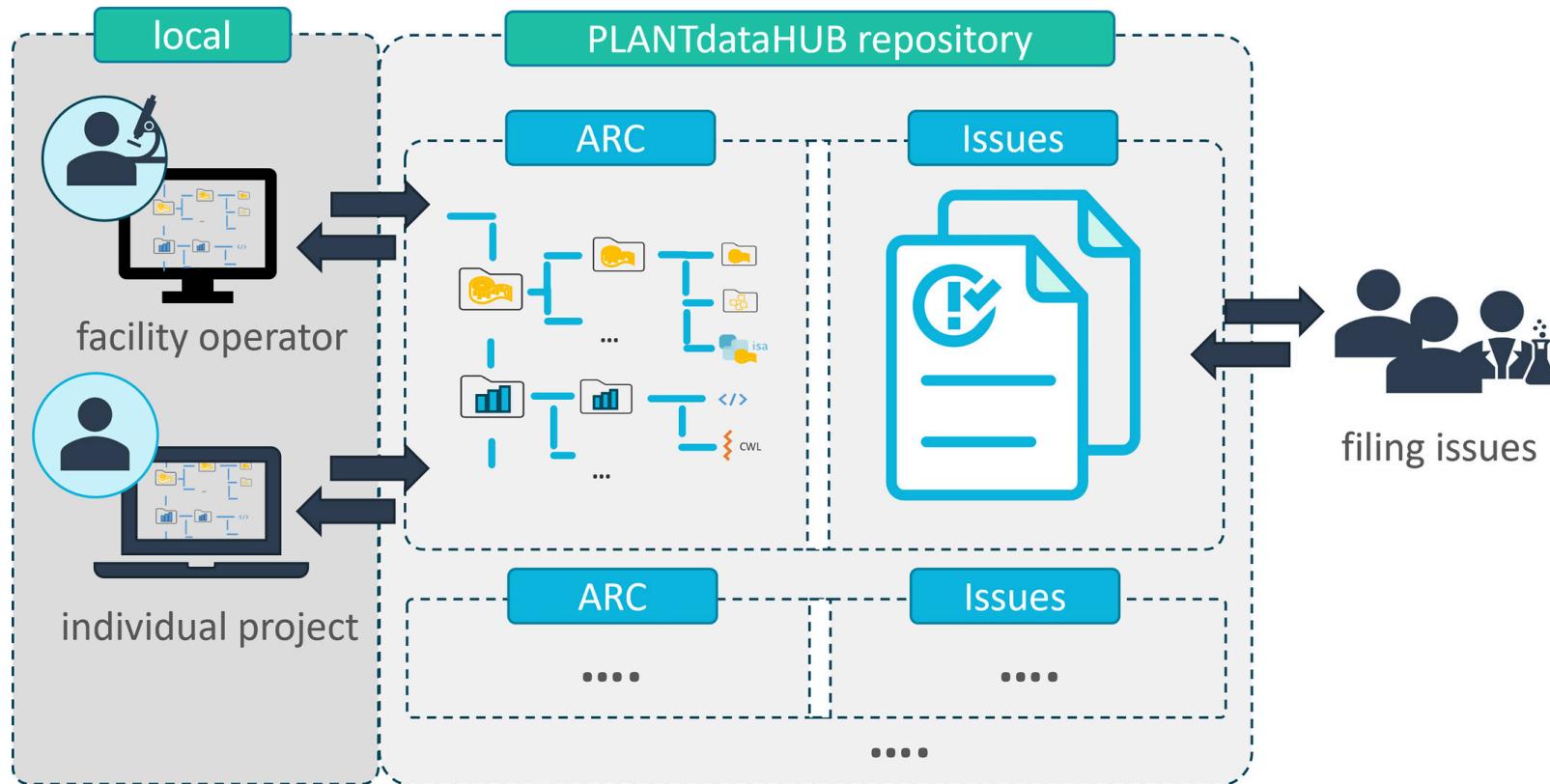
# From ARC to repositories



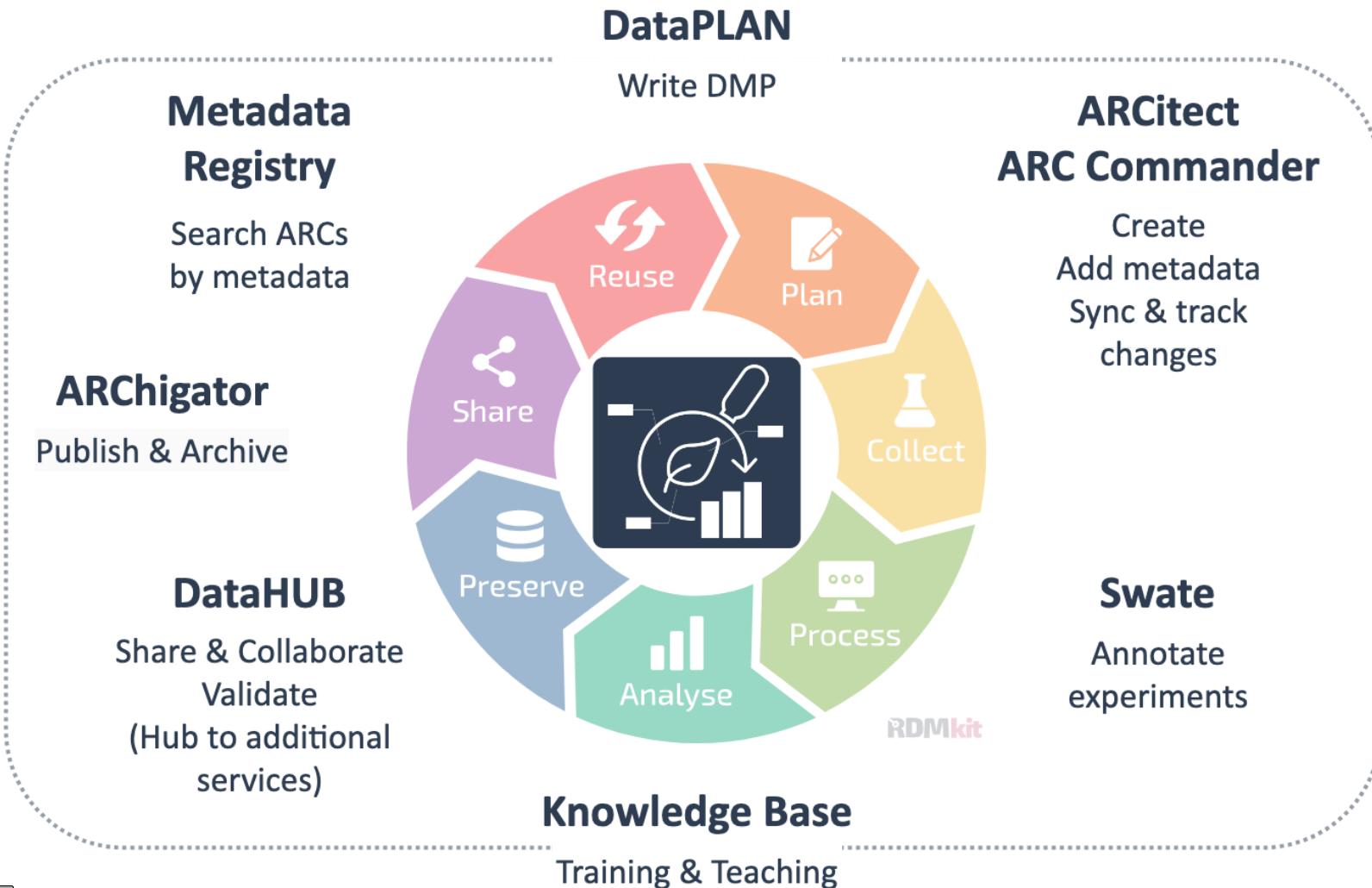
# Moving from paper to data publications



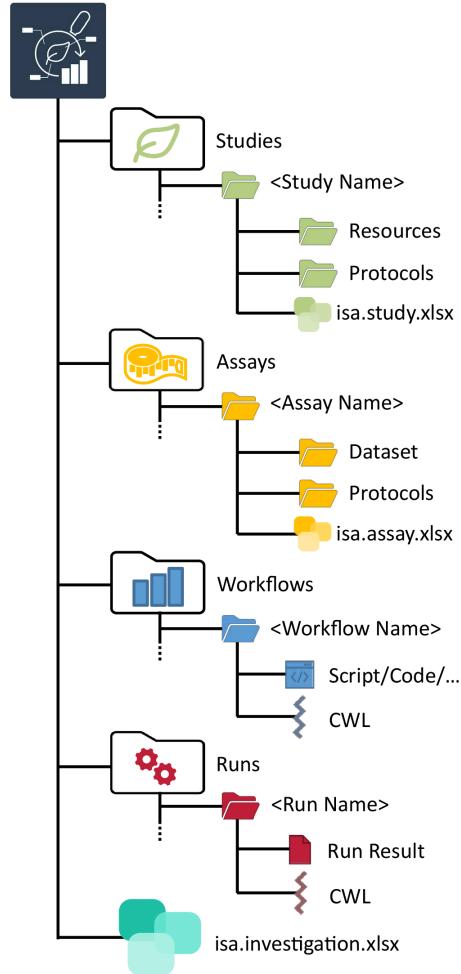
# Project management



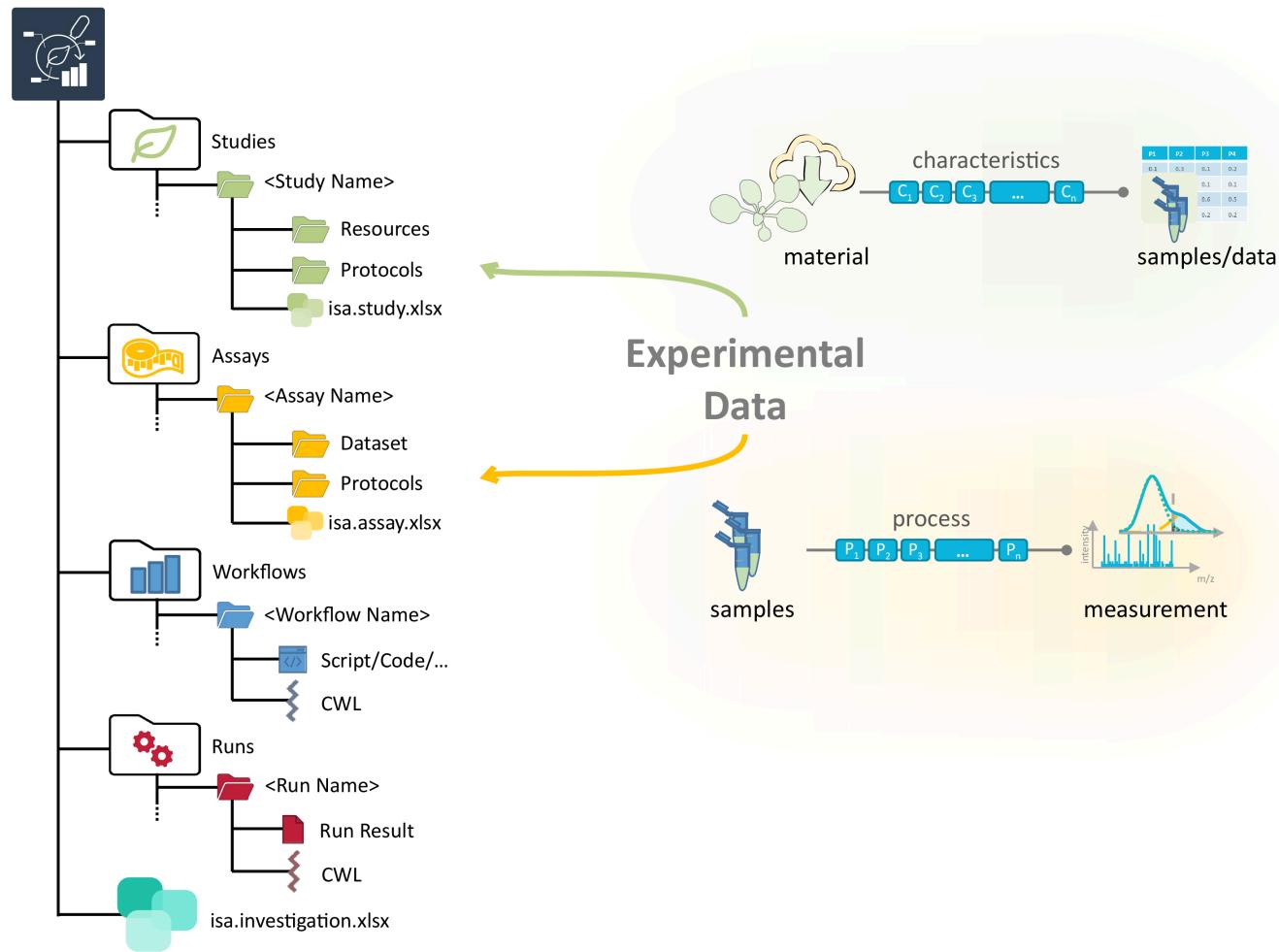
# The ARC ecosystem



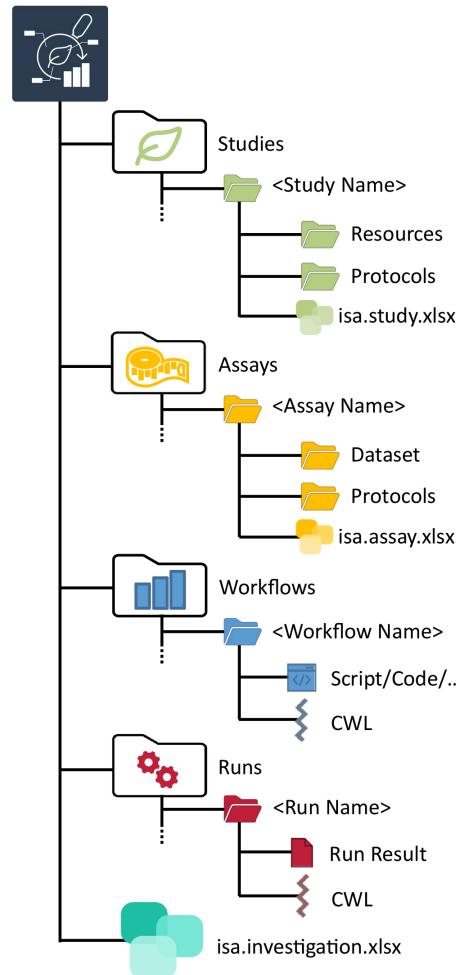
# What does an ARC look like?



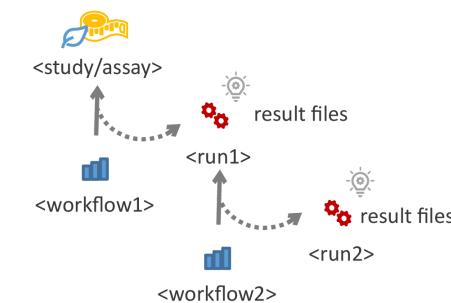
# ARCs store experimental data



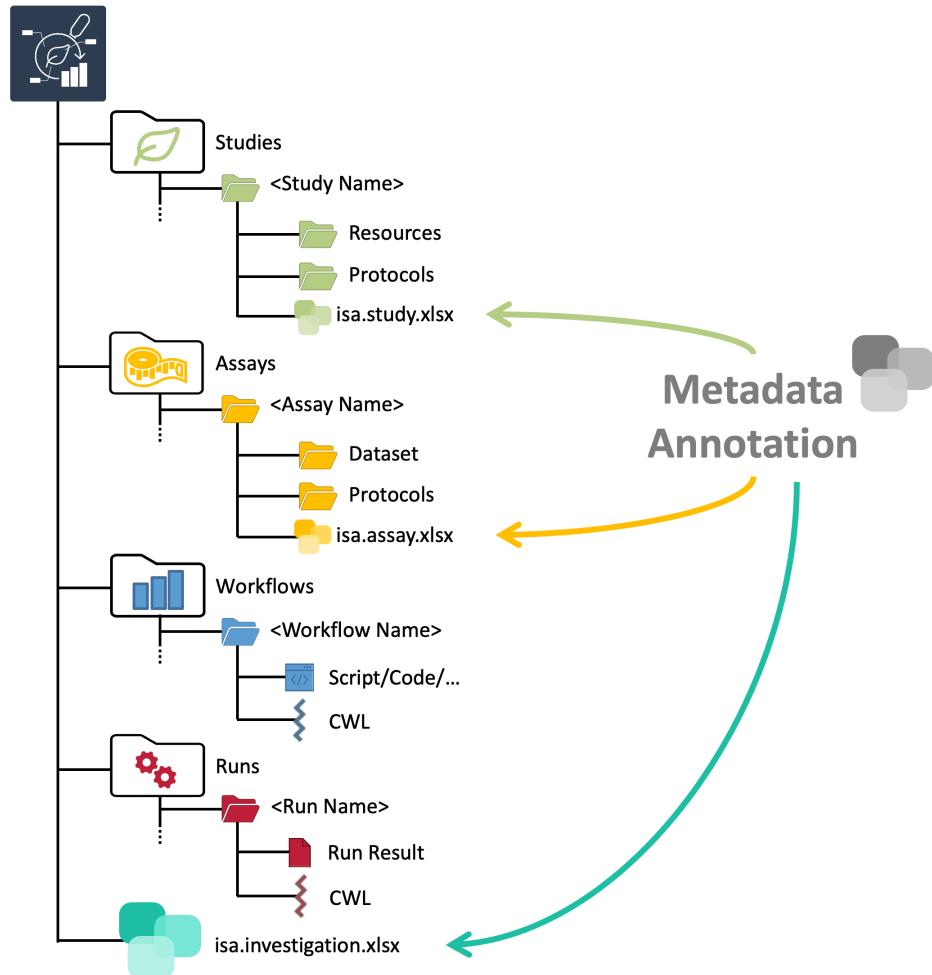
# Computations can be run inside ARCs



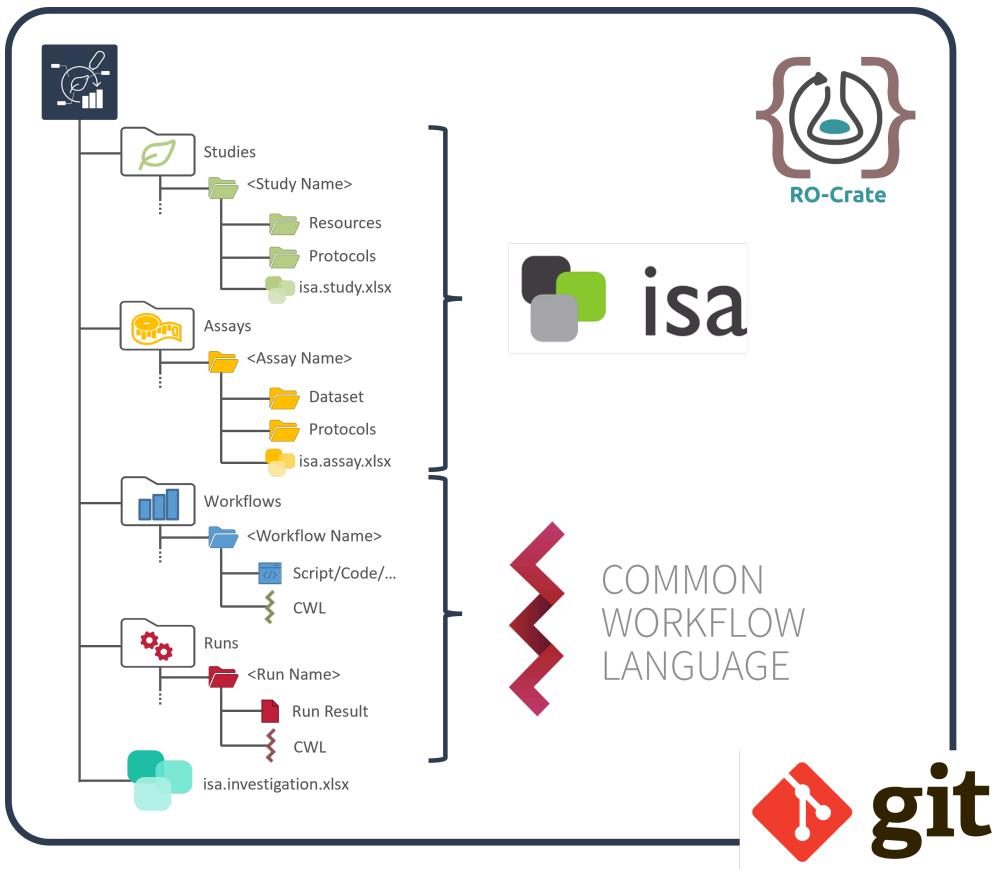
Data analysis  
Computation



# ARCs come with comprehensive metadata



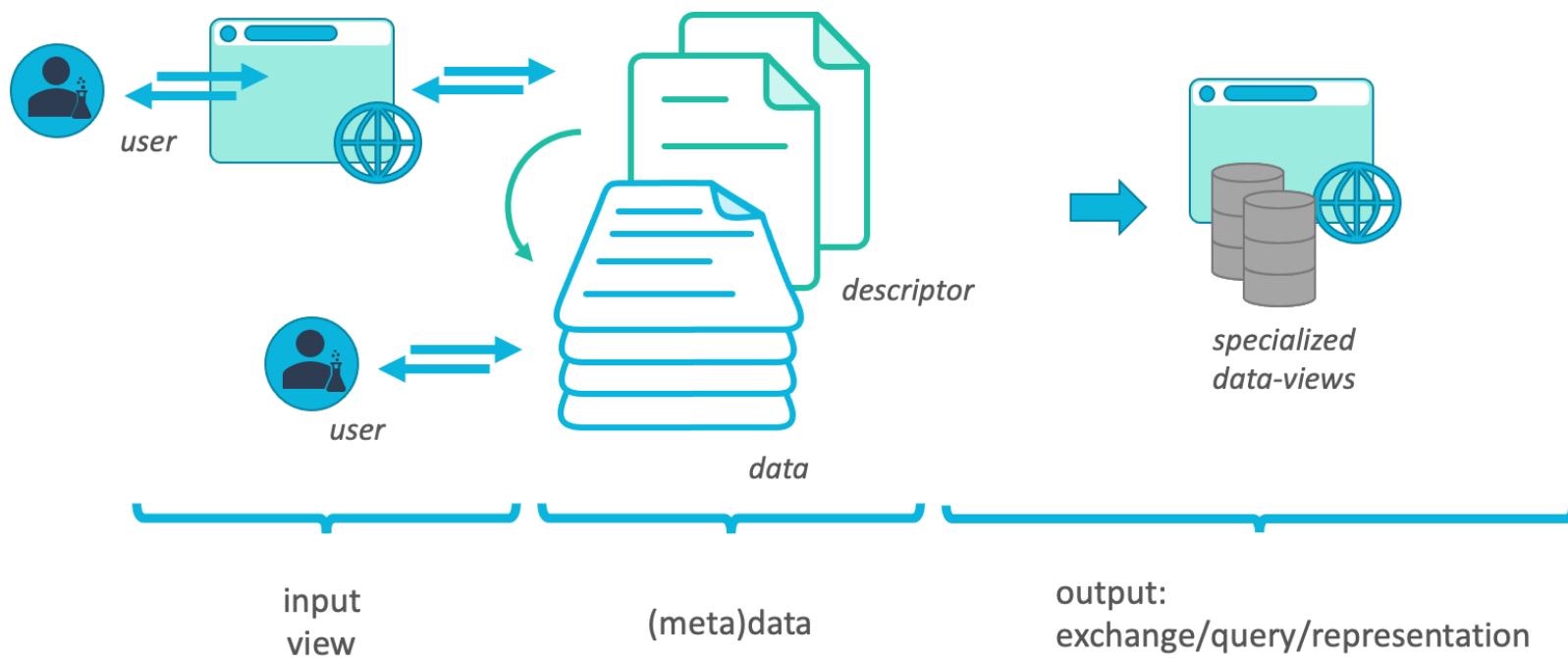
# ARC builds on standards



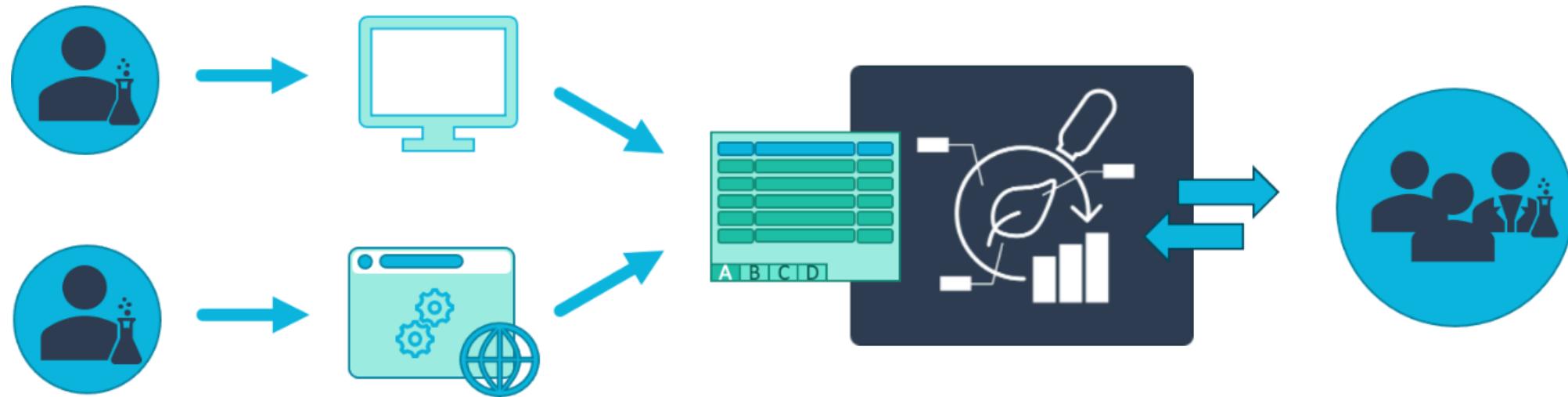
- RO-Crate: standardized exchange
- ISA: structured, machine-readable metadata
- CWL: reproducible, re-usable data analysis
- Git: version control

# Everything is a file

DataPLANT employs a **data-centric** approach to RDM



# No technical lock-in



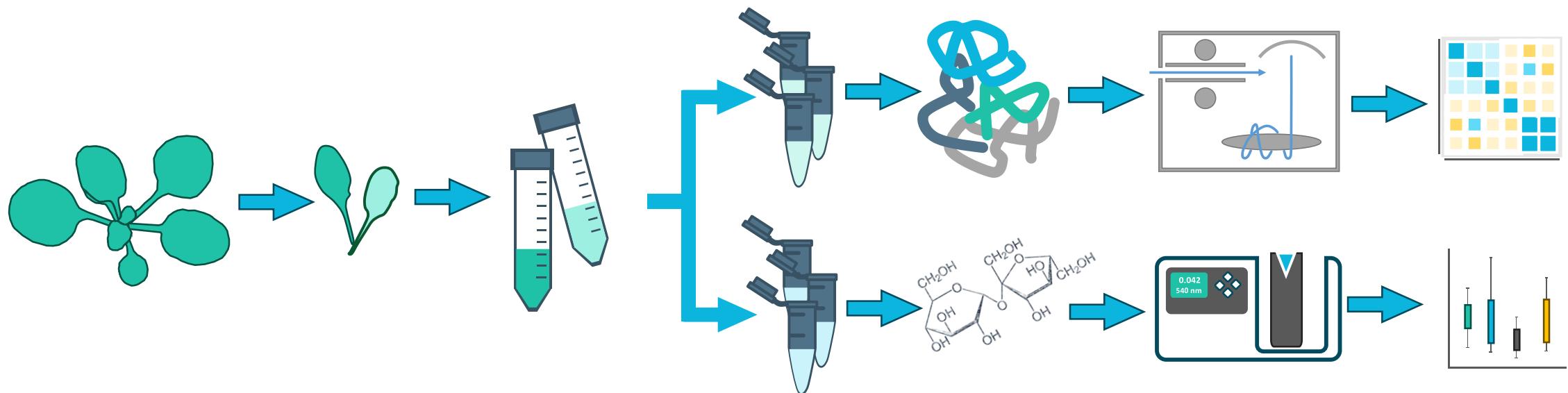
(Meta)data transparency with tool assistance but **no technical lock-in**

## Hands-on part 1: Start Here guide

[\*\*Start Here\*\*](#) guide in the DataPLANT knowledge base.

- follow the fictional researcher Viola on her journey to create her first ARC

# A small prototypic project



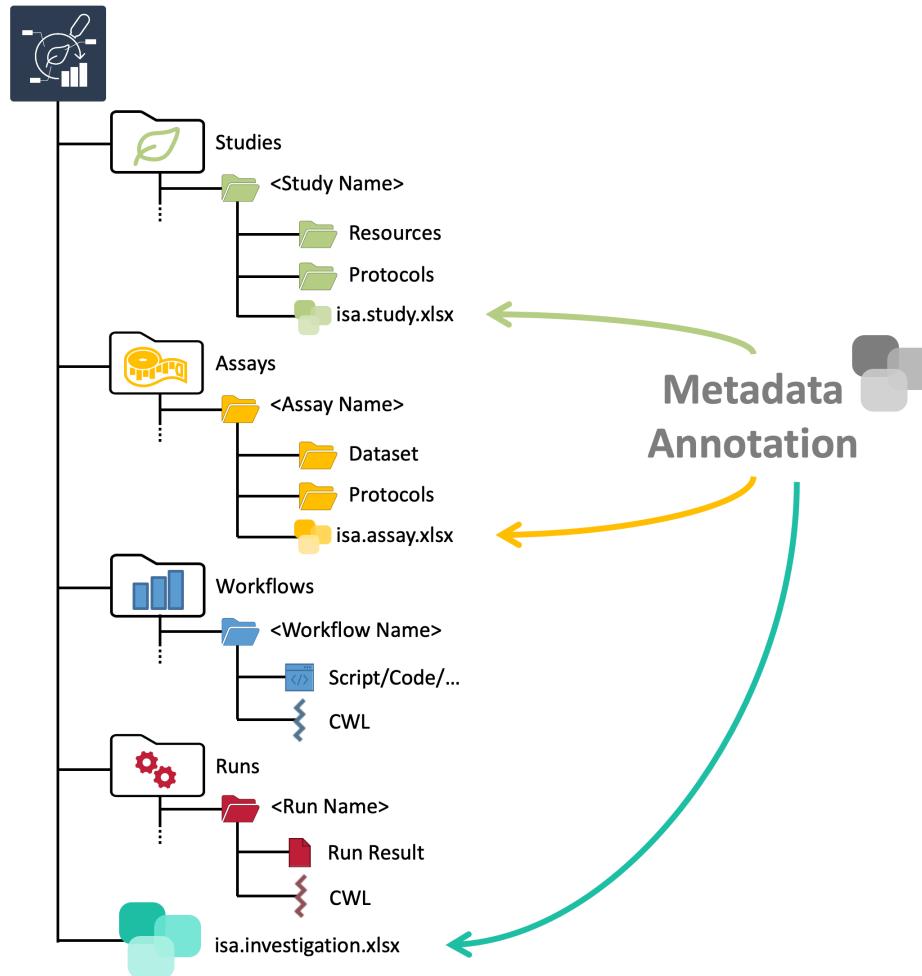
# Hands-on part 1: Setup and Investigation

Follow the [Start Here guide](#) in the DataPLANT knowledge base.

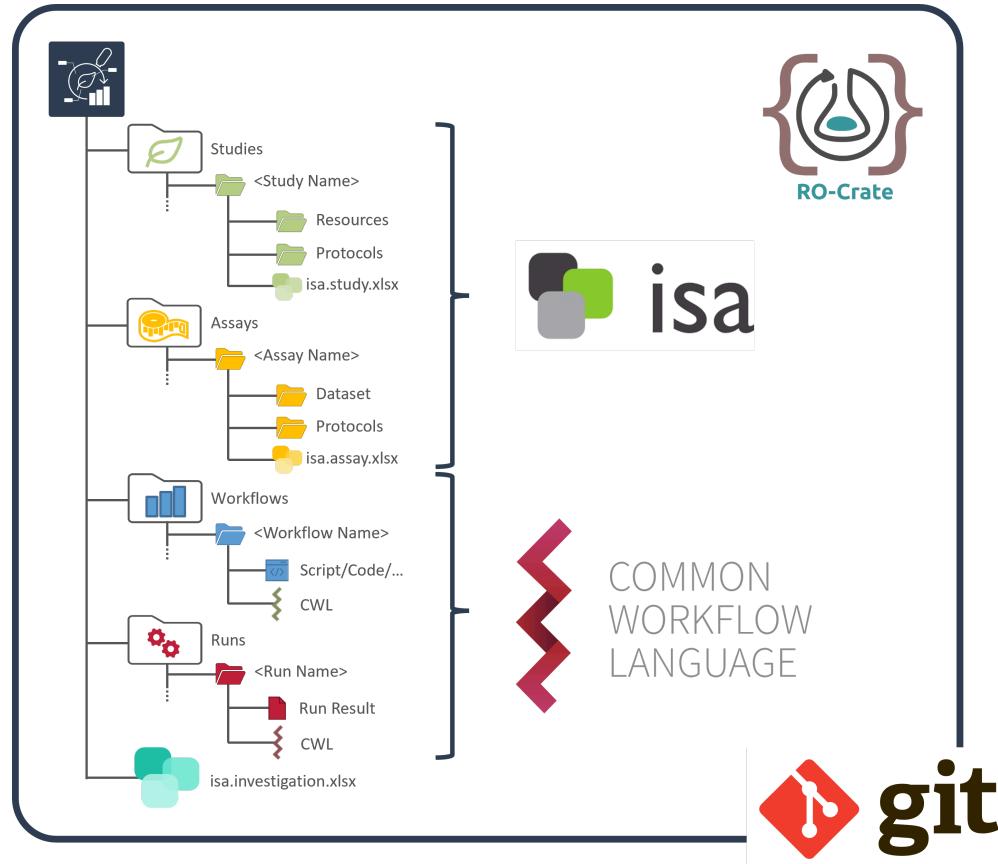
 Stop after step **Describing the Investigation**

**?** Please ask if you have any question/problems **?**

# ARCs come with comprehensive metadata

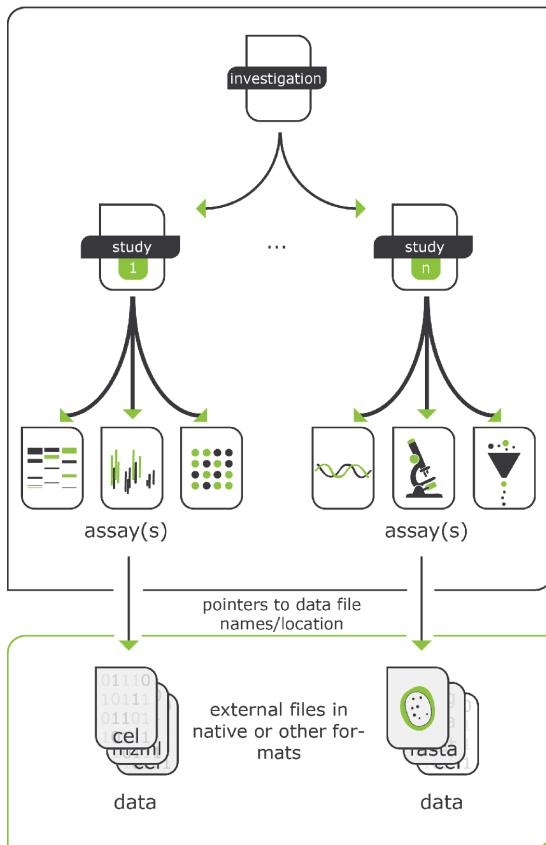


# ARC builds on standards



<https://isa-tools.org/> | <https://www.commonwl.org/>  
<https://www.researchobject.org/ro-crate/> | <https://git-scm.com>

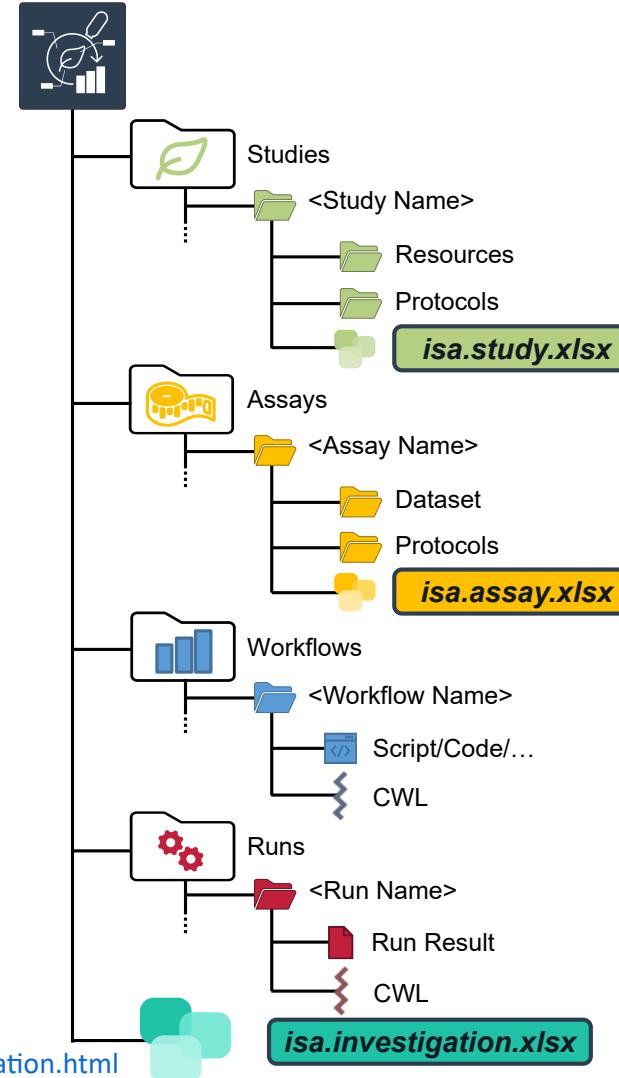
# ARC builds on ISA



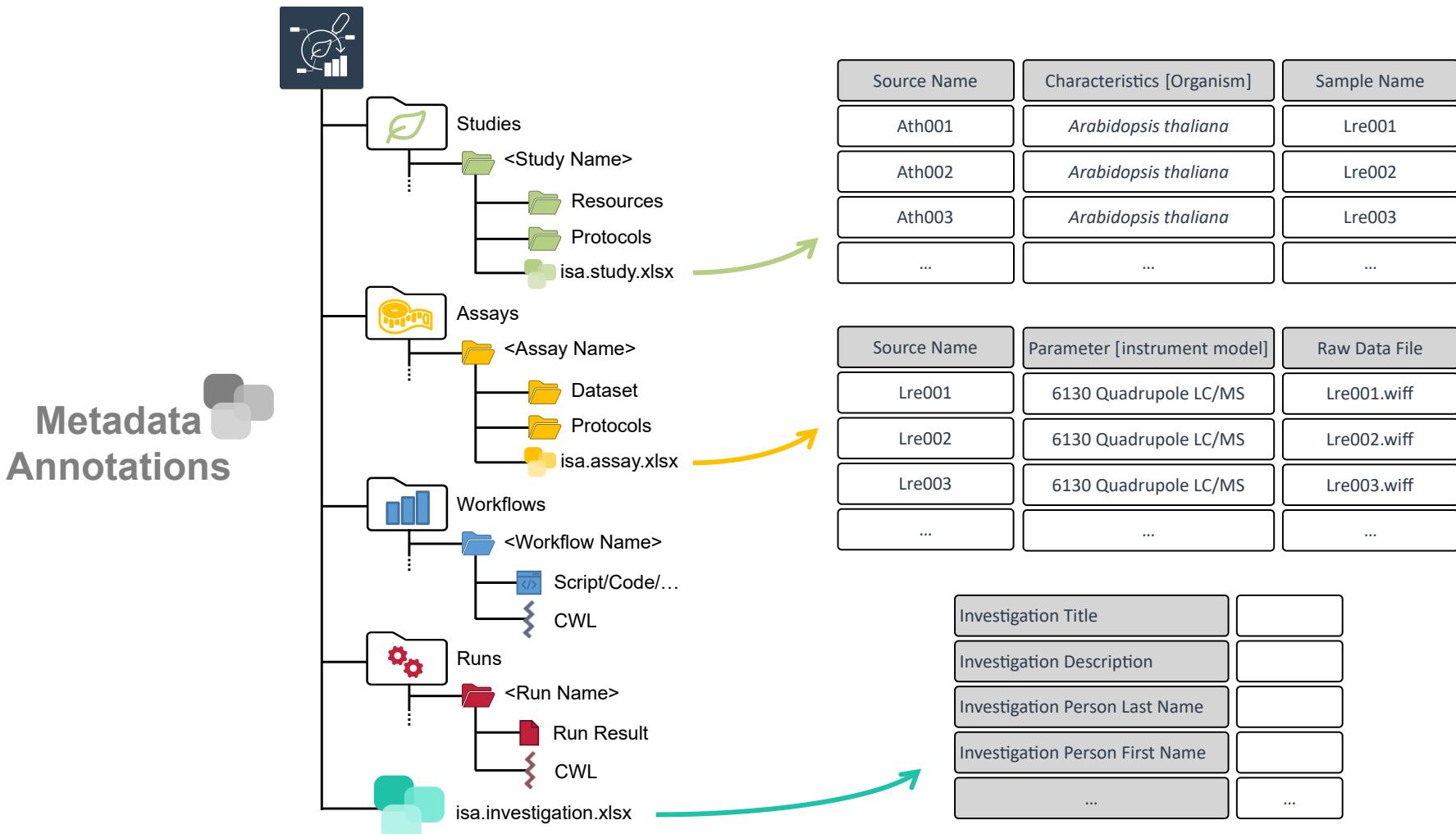
**Investigation**  
Overall goals  
Scientific context

**Study**  
Experimental design

**Assay**  
Leading to (raw) data

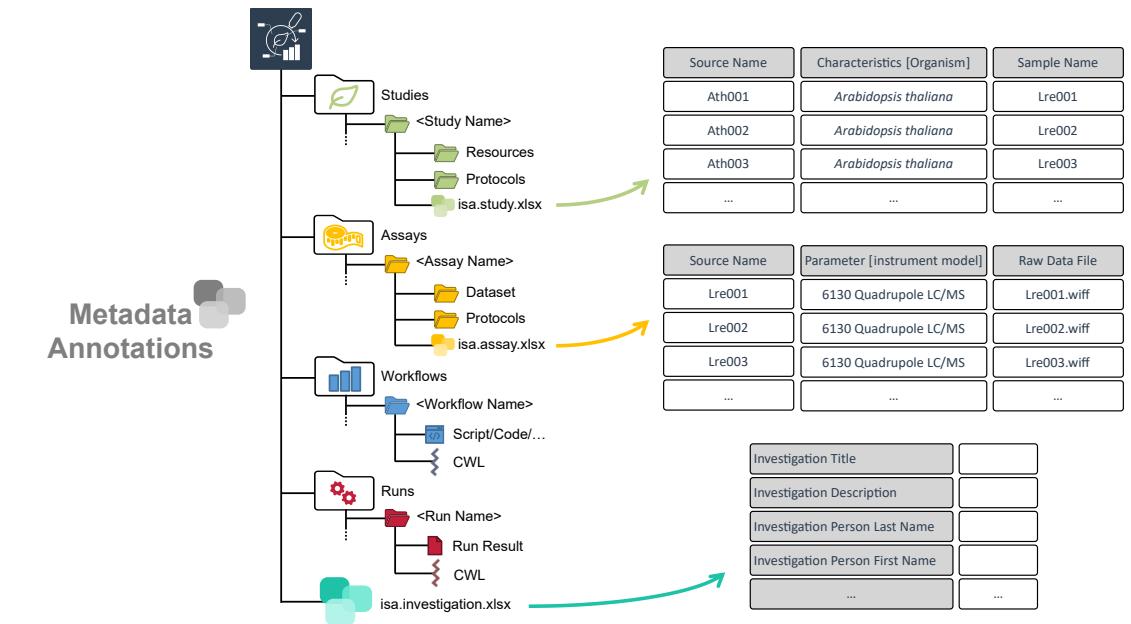


# ARC builds on ISA to connect data

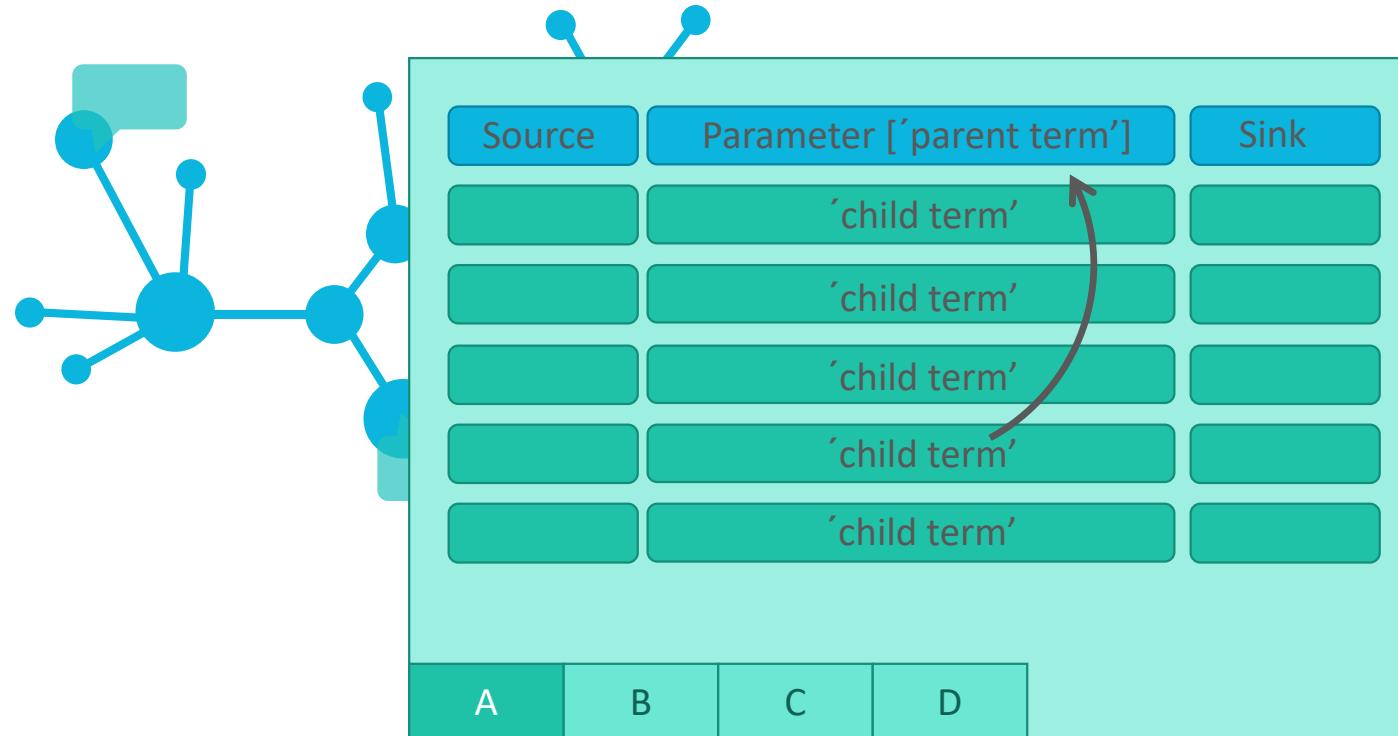


# ARC builds on ISA to link data

- Samples are linked study-to-assay, assay-to-assay
- Raw data is linked to assays
- Protocols can be referenced
- ...

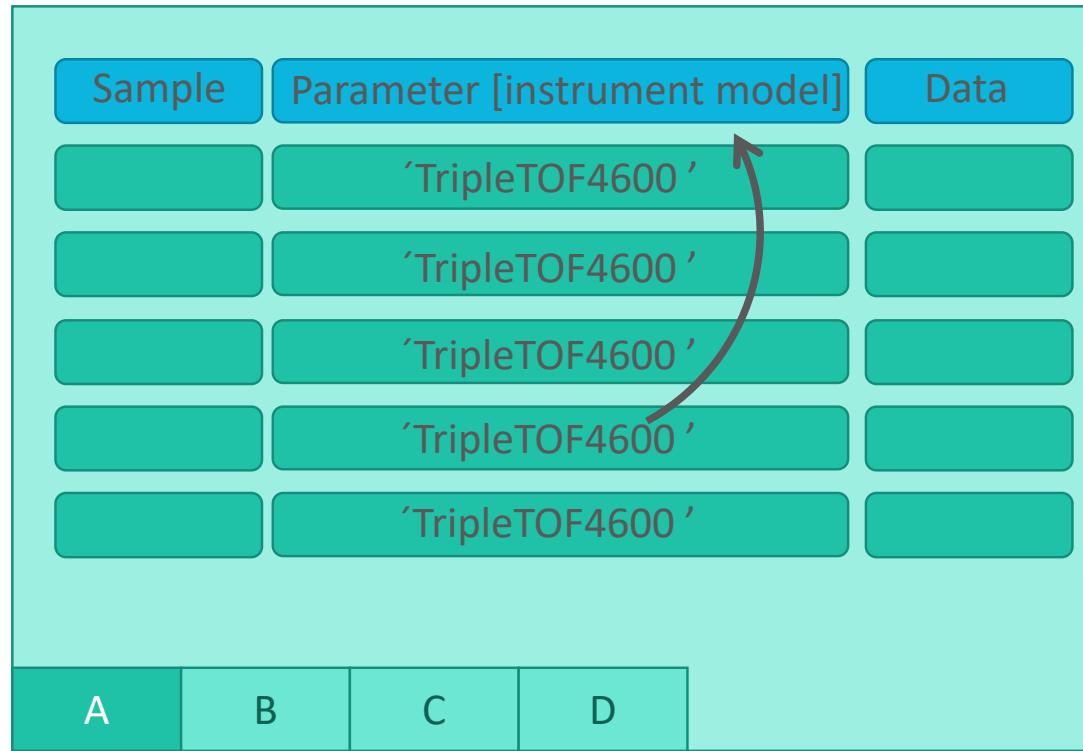


# Annotation by flattening the knowledge graph



- Low-friction metadata annotation
- Familiar spreadsheet, row/column-based environment

# Annotation principle



- Low-friction metadata annotation
- Familiar spreadsheet, row/column-based environment

# Adding new building blocks

The screenshot shows the Swate interface. On the left, there is a table with six rows of experimental data:

	Input [Source Name]	Characteristic [organism]	Factor [watering exposure]	Output [Sample Name]
1	DB_097	Talinum fruticosum	✓ 12 days drought	CAM_01
2	DB_099	Talinum fruticosum	✓ 12 days drought	CAM_02
3	DB_103	Talinum fruticosum	✓ 12 days drought	CAM_03
4	DB_161	Talinum fruticosum	✓ 12 days drought + 2 days rewatered	reC3_01
5	DB_163	Talinum fruticosum	✓ 12 days drought + 2 days rewatered	reC3_02
6	DB_165	Talinum fruticosum	✓ 12 days drought + 2 days rewatered	reC3_03

On the right, a modal dialog box titled "New building blocks" is open. It contains a search bar with "instrument model" and a list of categories: Input, Parameter, Factor, Characteristic, Component, More, and Output. The "Parameter" category is currently selected. A green button labeled "Add Column" is prominently displayed. A tooltip "Help" is visible near the "Add Column" button.

At the bottom of the interface, there are navigation buttons: "Metadata", "plant\_material", "+", and "<".

Swate (now integrated in ARCitect) can be used for the annotation of **isa.study.xlsx** and **isa.assay.xlsx** files

# Annotation Building Block types

- Input (e.g. Source Name, Sample Name)
- Protocol columns
- Characteristic // Parameter // Factor //
- Component
- Output (e.g. Sample Name, Raw Data File, Derived Data File)

The screenshot shows the MibitNet software interface with the 'Widgets' view open. The main table displays experimental data with columns for Input [Source Name], Characteristic [organism], Factor [watering exposure], and Output [Sample Name]. A 'Factor' column is highlighted in blue. Arrows point from labels 'Input', 'Characteristic', 'Factor', 'Output', and 'Sidebar' to their respective corresponding columns in the table. A 'New Parameter' dialog box is also visible on the right side.

Input [Source Name]	Characteristic [organism]	Factor [watering exposure]	Output [Sample Name]
DB_097	Talinum fruticosum	✓ 12 days drought	CAM_01
DB_099	Talinum fruticosum	✓ 12 days drought	CAM_02
DB_103	Talinum fruticosum	✓ 12 days drought	CAM_03
DB_161	Talinum fruticosum	✓ 12 days drought + 2 days rewetted	reC3_01
DB_163	Talinum fruticosum	✓ 12 days drought + 2 days rewetted	reC3_02
DB_165	Talinum fruticosum	✓ 12 days drought + 2 days rewetted	reC3_03

Let's take a detour on [Annotation Principles](#)

# Ontology term search

The screenshot shows the Swate interface with a table of experimental data and an ontology search sidebar.

**Table Data:**

	Input [Source Name]	Characteristic [Organism]	»	Factor [watering exposure]	»	Output [Sample Name]
1	DB_097		Q	12 days drought		CAM_01
2	DB_099			12 days drought		CAM_02
3	DB_103			12 days drought		CAM_03
4	DB_161			12 days drought + 2 days rewated		reC3_01
5	DB_163			12 days drought + 2 days rewated		reC3_02
6	DB_165			12 days drought + 2 days rewated		reC3_03

**Search Sidebar:**

Ontology term search  
Search for an ontology term to fill into the selected field(s)

Parent: Organism, MIAPPE:0041   [Use advanced search](#)

[Fill selected cells with this term](#)

Metadata   plant\_material   +

Swate Release Version v1.0.0-beta.03 Host Browser

Enable related term directed search to directly fill cells with child terms

# Fill your table with ontology terms

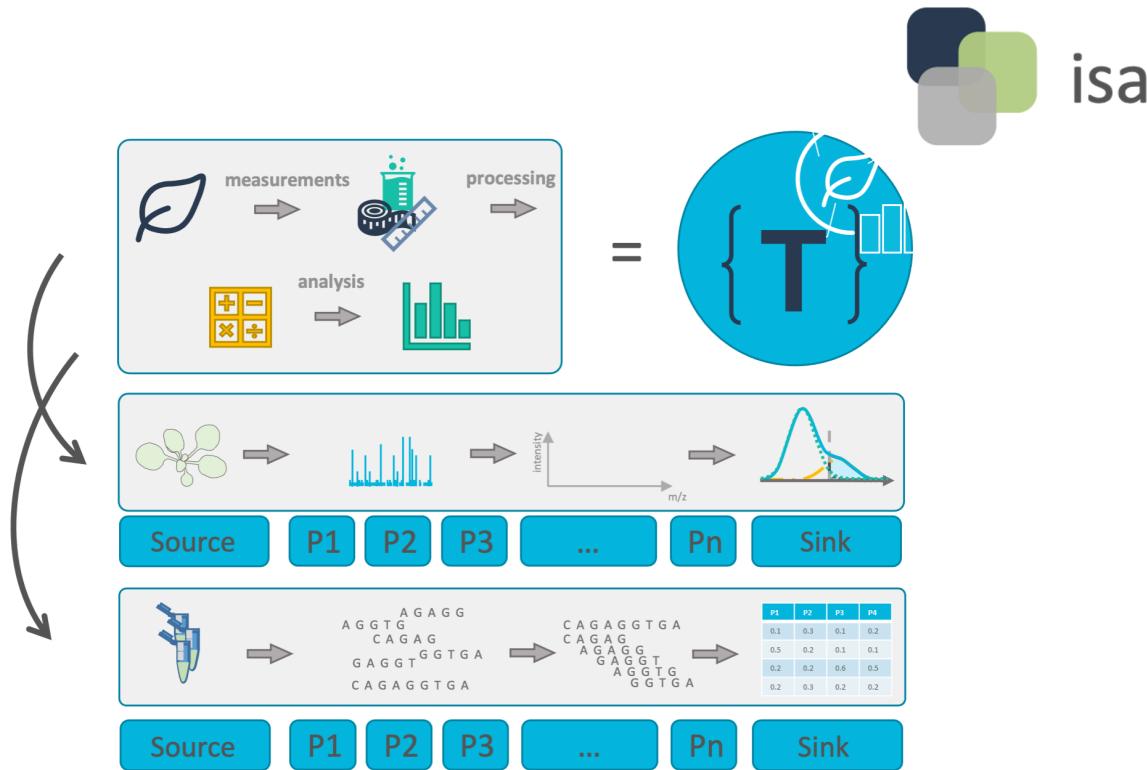
The screenshot shows the Swate interface. On the left, there is a table with six rows of data:

	Input [Source Name]	Characteristic [Organism]	Factor [watering exposure]	Output [Sample Name]
1	DB_097	Talinum fruticosum	✓ 12 days drought	CAM_01
2	DB_099	Talinum fruticosum	✓ 12 days drought	CAM_02
3	DB_103	Talinum fruticosum	✓ 12 days drought	CAM_03
4	DB_161	Talinum fruticosum	✓ 12 days drought + 2 days rewatered	reC3_01
5	DB_163	Talinum fruticosum	✓ 12 days drought + 2 days rewatered	reC3_02
6	DB_165	Talinum fruticosum	✓ 12 days drought + 2 days rewatered	reC3_03

On the right, there is an "Ontology term search" panel:

- Search bar: Talinum fruticosum
- Parent: Organism, MIAPPE:0041
- Buttons: Fill selected cells with this term, Use advanced search

# Realization of lab-specific metadata templates



Facilities can define their most common workflows as templates

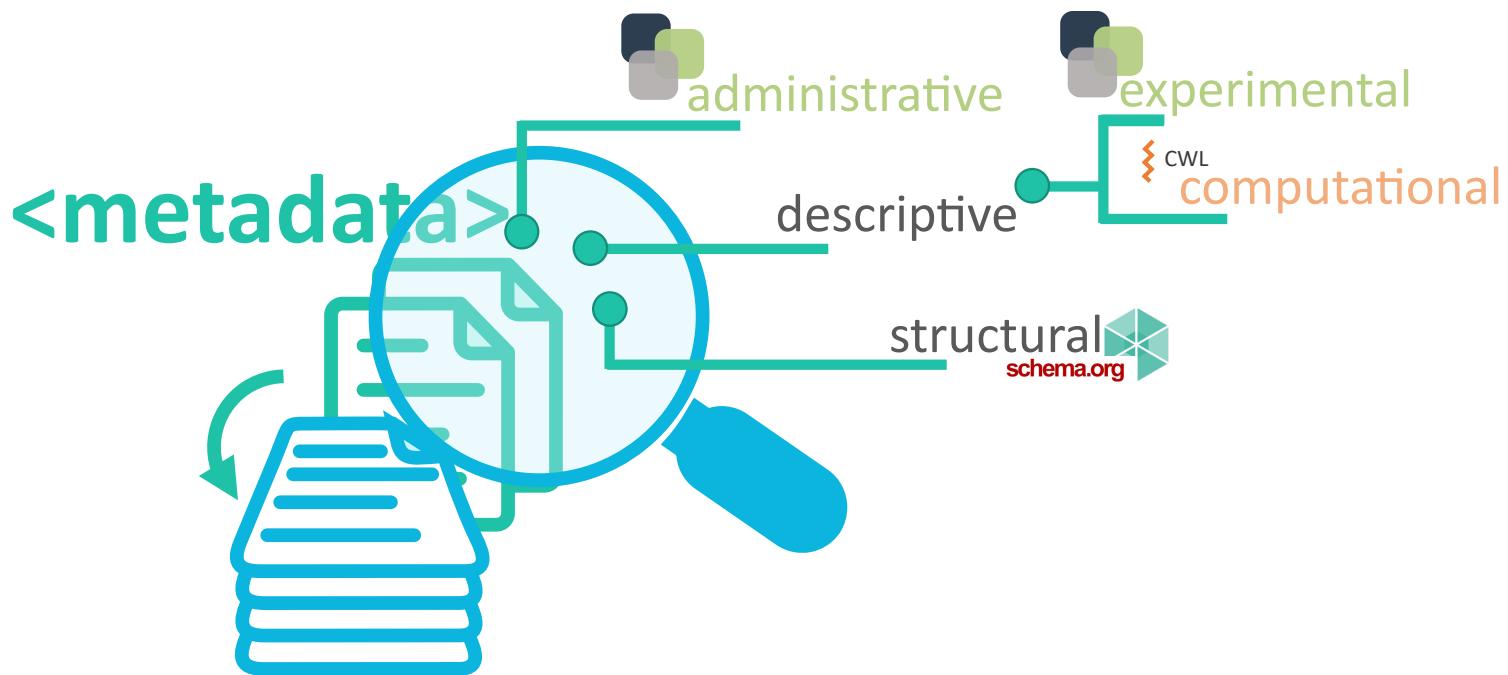
# Directly import templates via Swate

- DataPLANT curated
- Community templates

The screenshot shows the Swate web application interface. At the top, there is a dark header bar with several small icons. Below the header is a search bar with two input fields: "Search by template name" and "Search for tags". Underneath the search bar is a dropdown menu labeled "Select community" with "DataPLANT official" selected. The main content area is a table listing various templates. The columns are "Template Name", "Community", and "Template Version". Each row contains a link icon at the end. The table lists the following templates:

Template Name	Community	Template Version
DNA extraction	curated	1.1.7
Data Processing (PRIDE minimal)	curated	1.0.0
GEO - Minimal information RNA assays	curated	1.0.1
GEO - Minimal information RNA extraction	curated	1.0.0
GEO - Minimal information computational analysis	curated	1.0.0
GEO - Minimal information plant growth	curated	1.0.0
Genome assembly	curated	1.1.7

# Build on existing well-established standards



Integration of established (meta)data standards ensures compatibility by design

# ISA abstract model in a nutshell



isa

Investigation  
administrative (meta)data

- Summary
  - Titel
  - Description
- Person
- Organisation
- Publication reference

Study  
descriptive (meta)data  
information on the subject

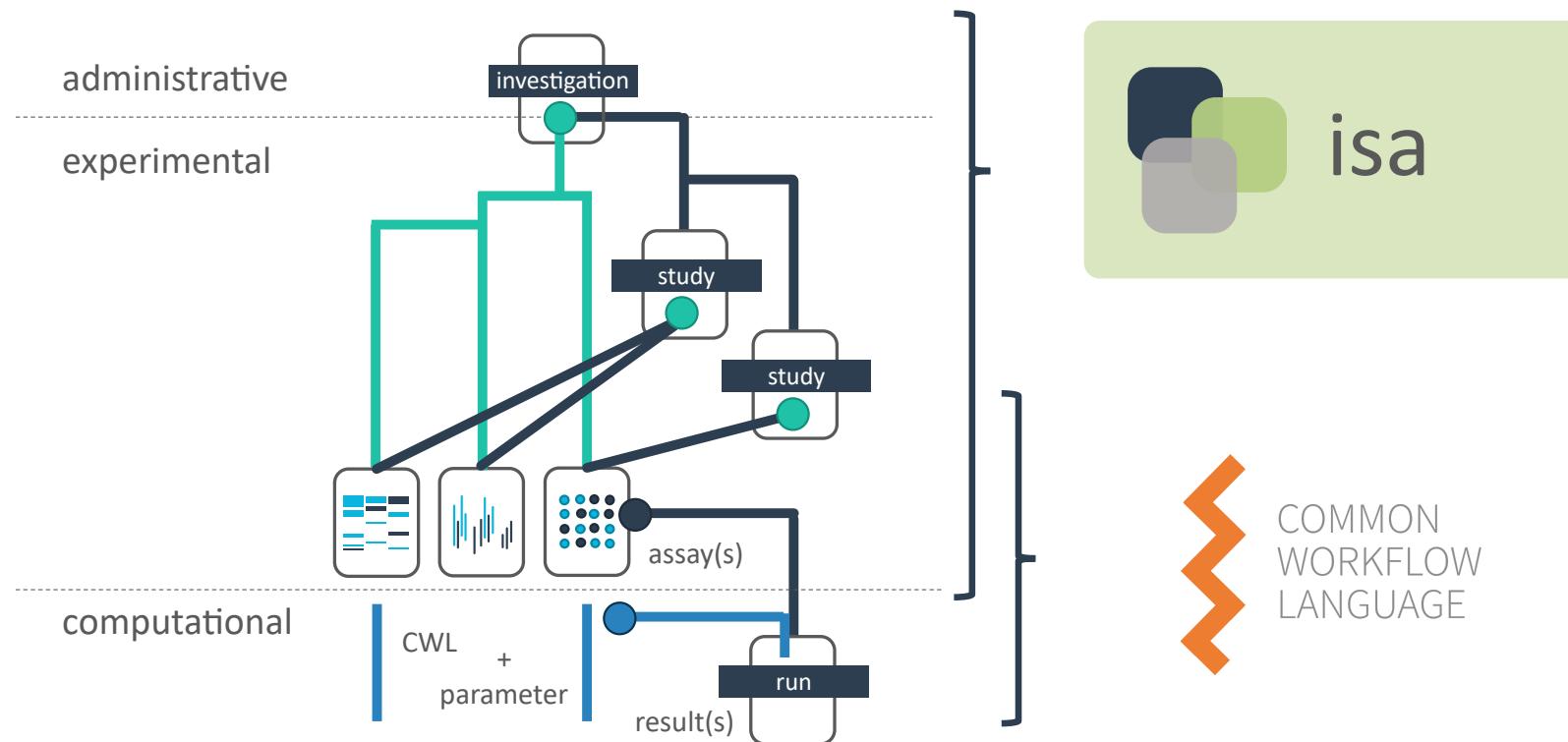
- Characteristics
- Parameters
- Components
- Factors

## Assay

descriptive (meta)data  
information on the measurement

- Characteristics
- Parameters
- Components
- Factors

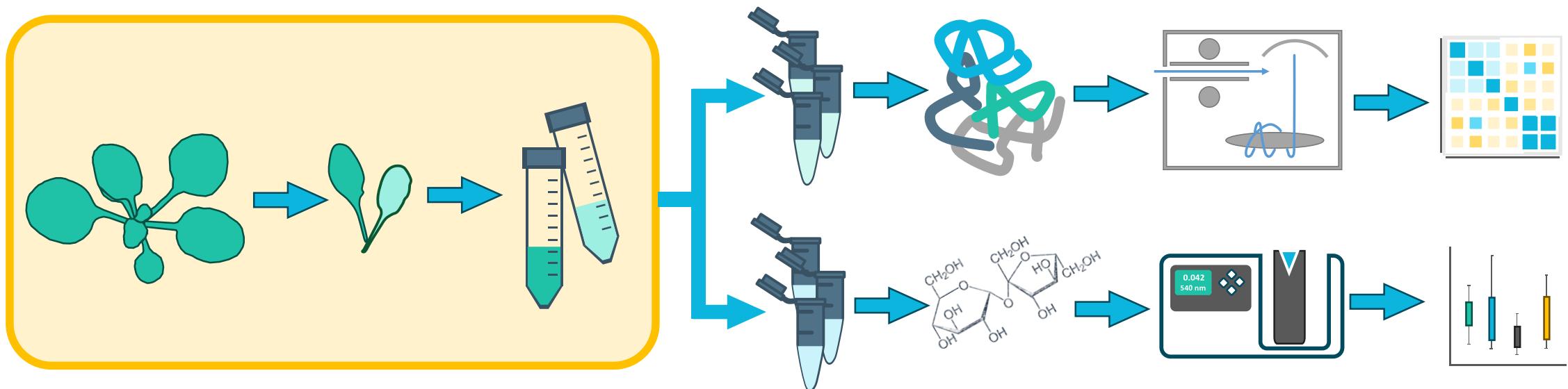
# ISA and CWL – Connected by similarity



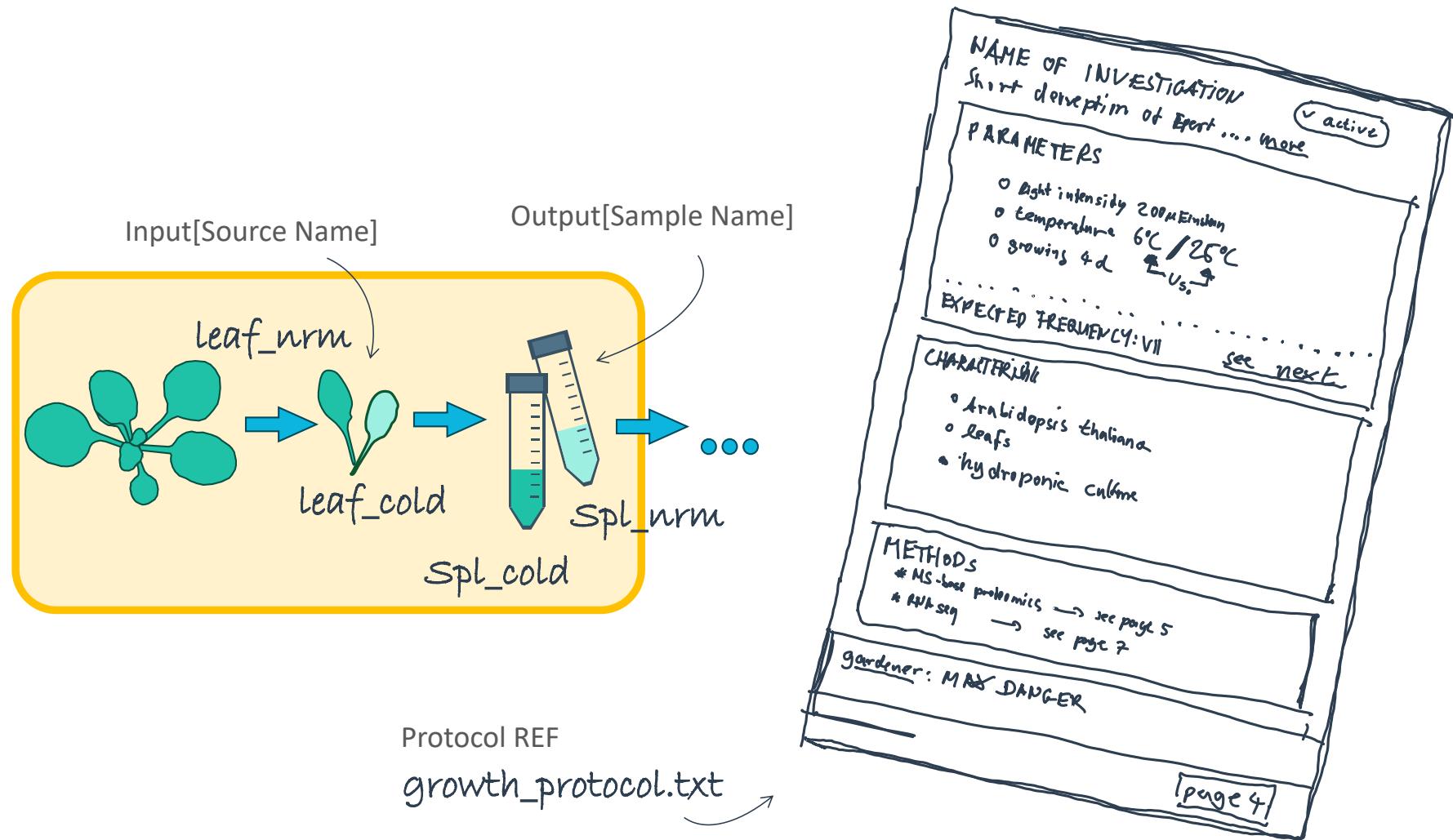
## Hands-on part 2: Start Here guide, study & assay part

[\*\*Start Here\*\*](#) guide in the DataPLANT knowledge base.

# Divide and conquer for reproducibility



# Identifying the 'study' part



# A table-based organization schema

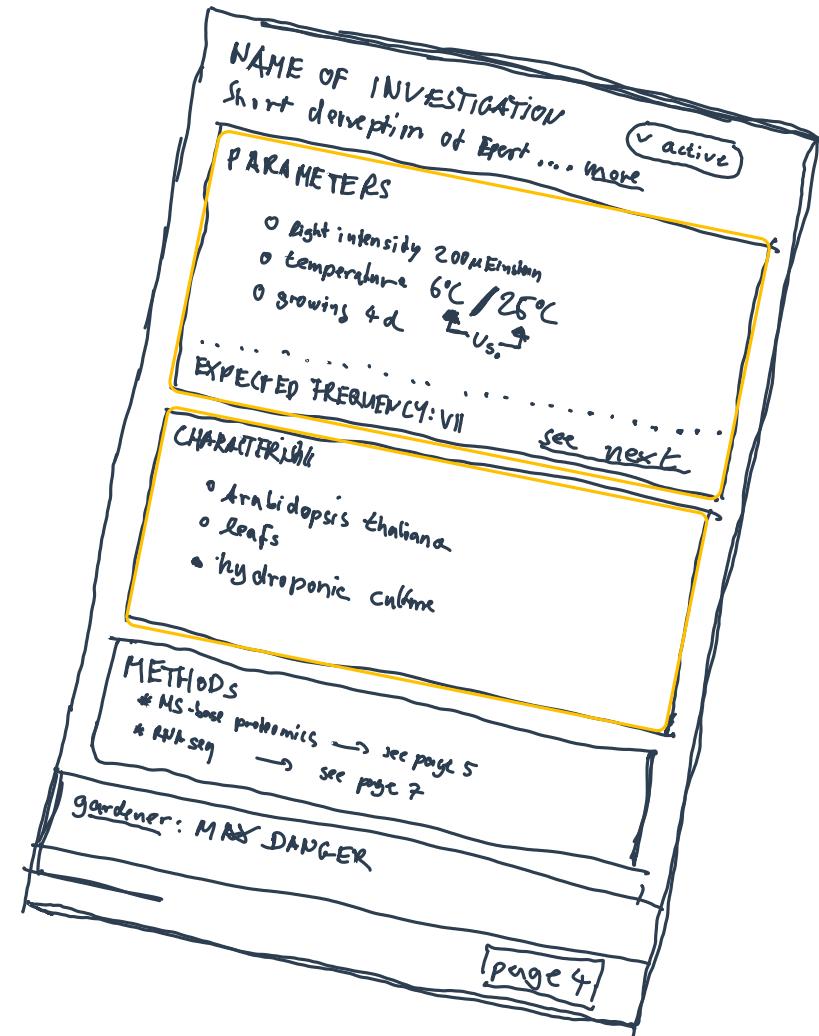
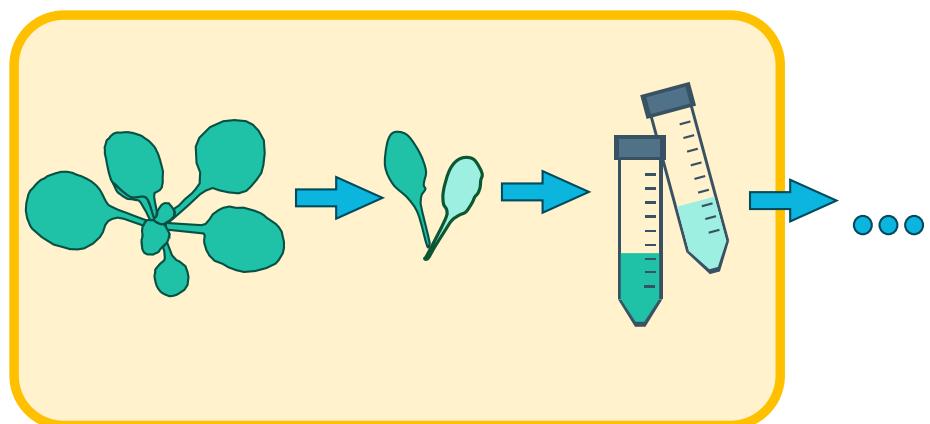
Input[Source Name]		Output[Sample Name]
leaf_nrm		spl_nrm
leaf_cold		spl_cold
A	B	C
D		

## Referencing a protocol

This allows you to reference the free-text, human-readable protocol.

- 💡 It is recommended that the protocol is in an open format (.md|.txt|.docx|...)
- 💡 But everything is possible also an URI to an electronic lab notebook

# Parameterizing the 'study'



# Finding the right metadata vocabulary

## Parameters []

- Light intensity 200 µEinstein
- Temperature 6°C / 25°C
- Growing 4d

## Characteristics []

- *Arabidopsis thaliana*
- Leaf
- Hydroponic culture
- Columbia

# OLS: Finding the right metadata vocabulary

Temperature Dependence [Temperature:Dependence\\_Annotation](#)

[http://purl.uniprot.org/core/Temperature\\_Dependence\\_Annotation](http://purl.uniprot.org/core/Temperature_Dependence_Annotation)

Indicates the optimum temperature for enzyme activity and/or the variation of enzyme activity with temperature variation; the thermostability/thermolability of the enzyme is also mentioned when it is known.

Ontology: [UNIPROT RDFS](#)

temperature [AFO:/result#AFR\\_0001584](#)

[http://purl.allotrope.org/ontologies/result#AFR\\_0001584](http://purl.allotrope.org/ontologies/result#AFR_0001584)

A temperature (datum) is a quantity facet that quantifies some temperature. [Allotrope]

Ontology: [AFO](#)

temperature [FBcv:0000466](#)

[http://purl.obolibrary.org/obo/FBcv\\_0000466](http://purl.obolibrary.org/obo/FBcv_0000466)

Mutation caused by exposure to a temperature that is higher or lower than 25 degrees Celsius.

Ontology: [FBCV](#)

temperature [PATO:0000146](#)

[http://purl.obolibrary.org/obo/PATO\\_0000146](http://purl.obolibrary.org/obo/PATO_0000146)

A physical quality of the thermal energy of a system.

Ontology: [PATO](#)

Also appears in: [NGBO](#) [HTN](#) [CAO](#) [ZP](#) [AGRO](#) [OMIABIS](#) [OBIB](#) [MONDO](#) [TXPO](#) [MCO](#) +

Welcome to the EMBL-EBI Ontology Lookup Service

temperature

Exact match  Include obsolete terms  Include imported terms

Examples: diabetes, GO:0098743

Looking for a particular ontology?

About OLS

The Ontology Lookup Service (OLS) is a repository for biomedical ontologies that aims to provide a single point of access to the latest ontology versions. You can browse the ontologies through the website as well as programmatically via the OLS API. OLS is developed and maintained by the Samples, Phenotypes and Ontologies Team (SPOT) at EMBL-EBI.

Related Tools

In addition to OLS the SPOT team also provides the Oxo and ZOOMA services. Oxo provides cross-ontology mappings between terms from different ontologies. ZOOMA is a service to assist in mapping data to ontologies in OLS.

Report an Issue

For feedback, enquiries or suggestion about OLS or to request a new ontology please use our GitHub issue tracker. For announcements relating to OLS, such as new releases and new features sign up to the OLS announce mailing list.

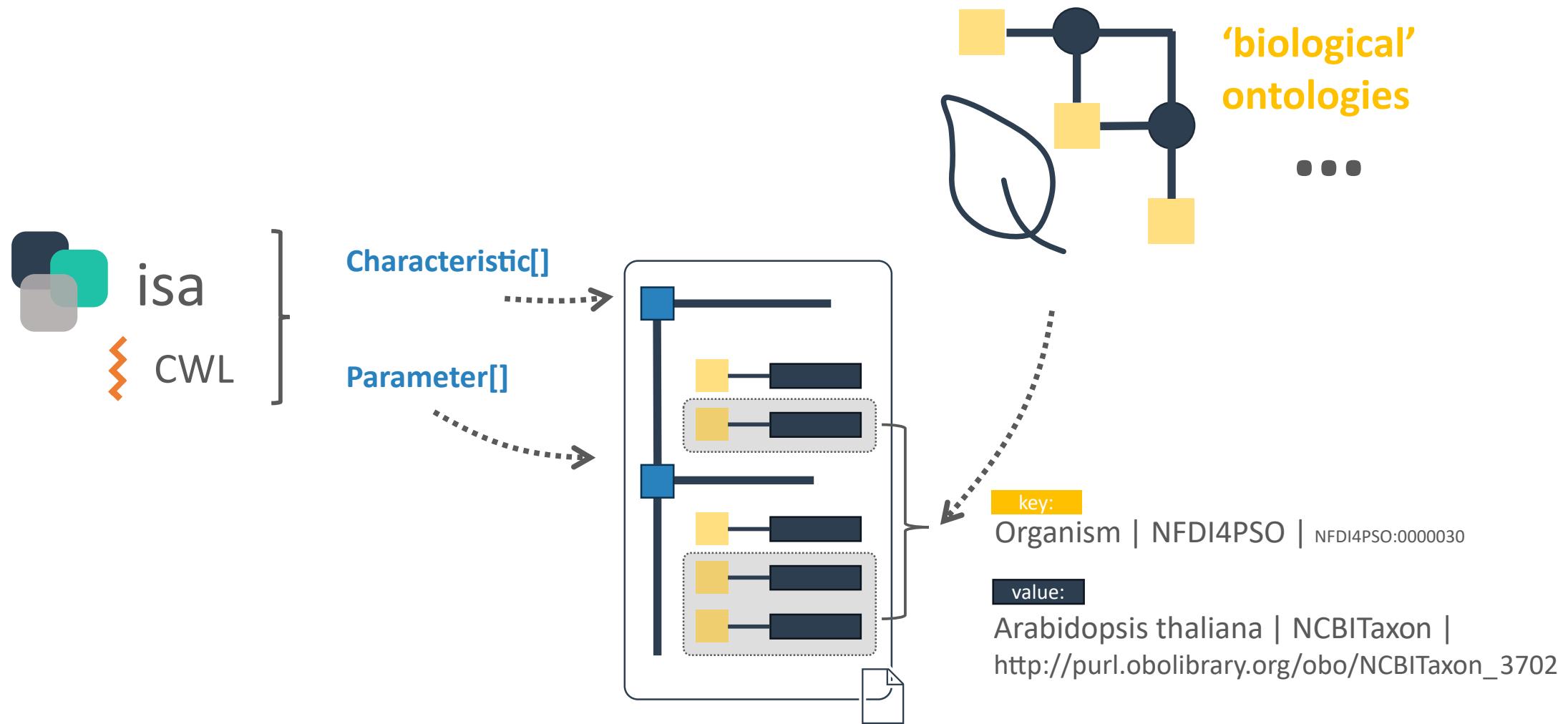
FOLLOW US

X

EMBL-EBI 2023 Licensing

Ontology Lookup Service (OLS): <https://www.ebi.ac.uk/ols4/>

# Finding the metadata vocabulary and descriptors



# Finding the metadata vocabulary and descriptors

## Parameters []

- 
- 
- 
-  Light intensity 200  $\mu$ Einstein
-  Temperature 6°C / 25°C
-  Growing 4d

## Characteristics []

- 
- 
- 
-  *Arabidopsis thaliana*
-  Leaf
-  Hydroponic culture
-  Columbia

# Finding the metadata vocabulary and descriptors

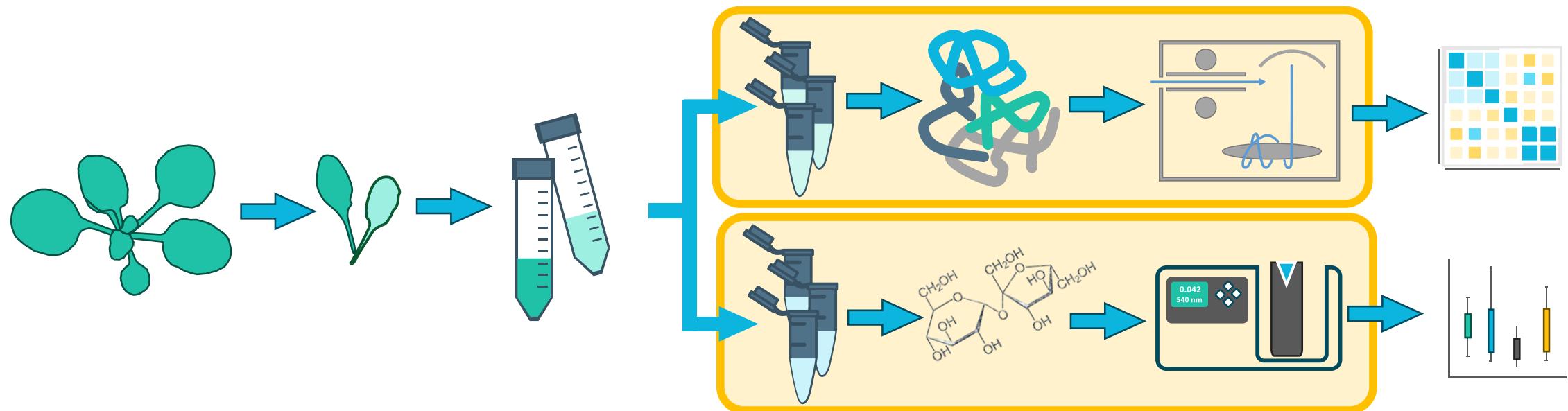
## Parameters []

- █ Light intensity
  - █ 200 µEinstein
- █ Temperature
  - █ 6°C / 25°C
- █ Growth time
  - █ 4d

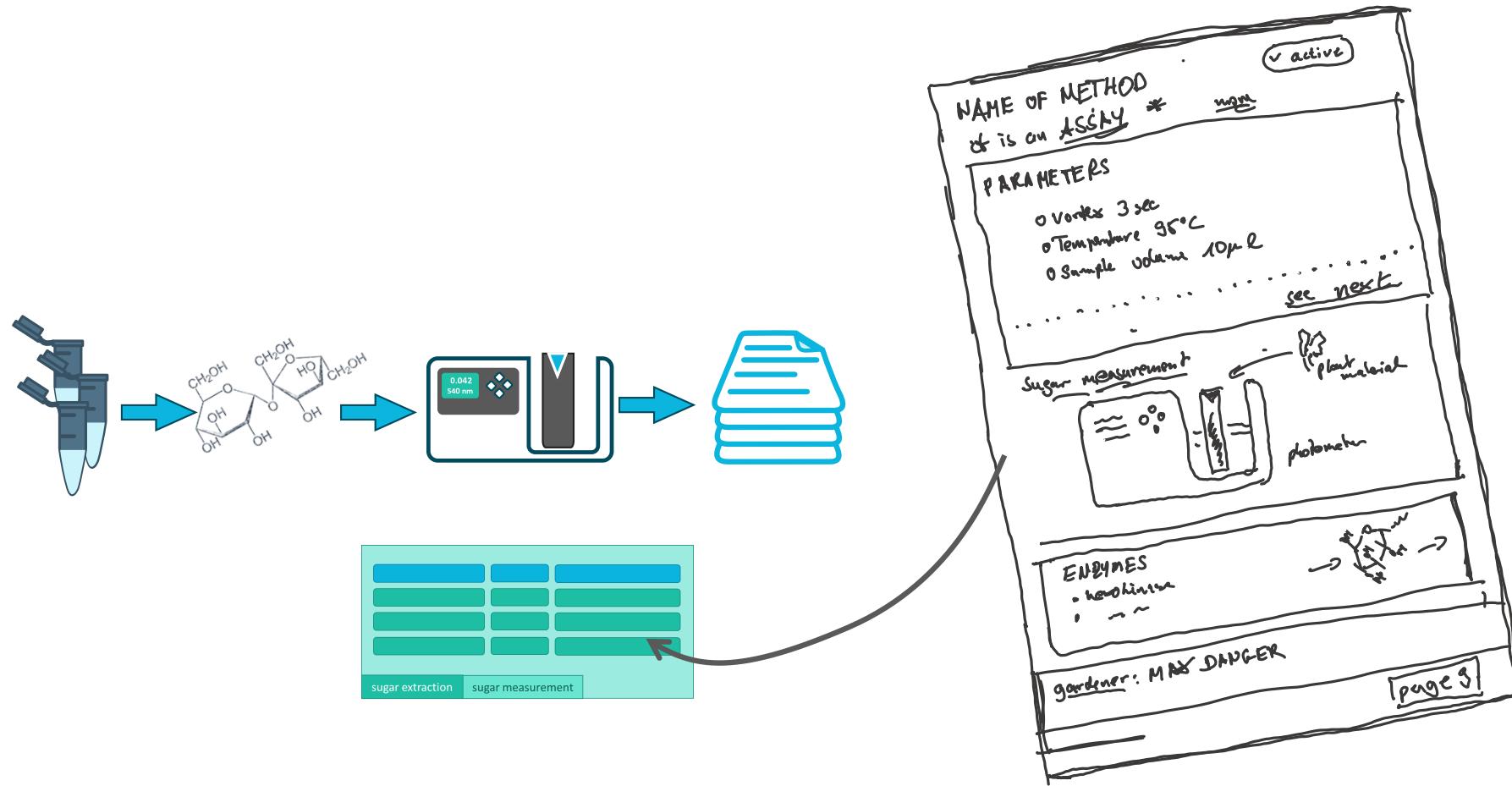
## Characteristics []

- █ Organism
  - █ *Arabidopsis thaliana*
- █ Tissue
  - █ Leaf
- █ Growth medium
  - █ Hydroponic culture
- █ Ecotype
  - █ Columbia

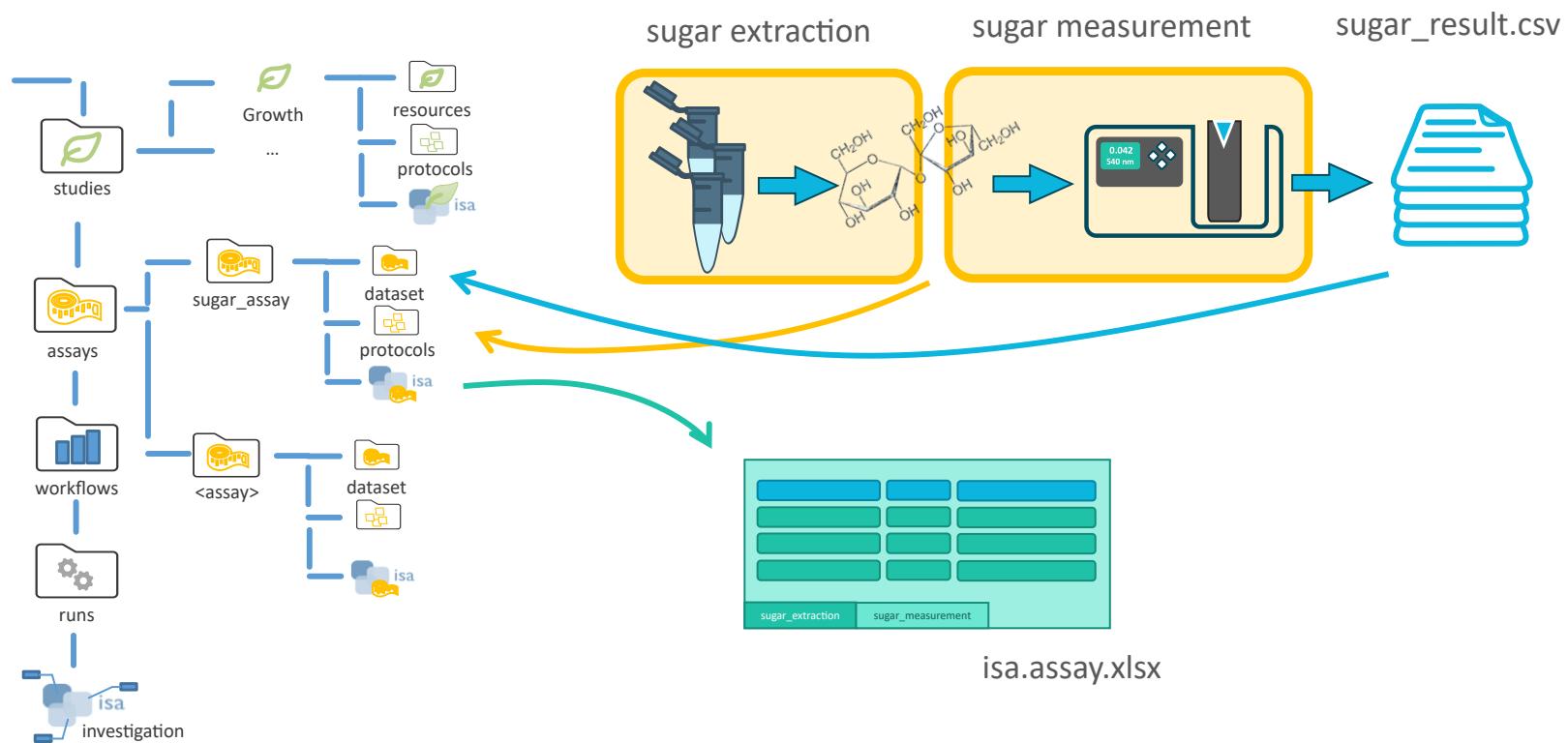
# Identifying assays



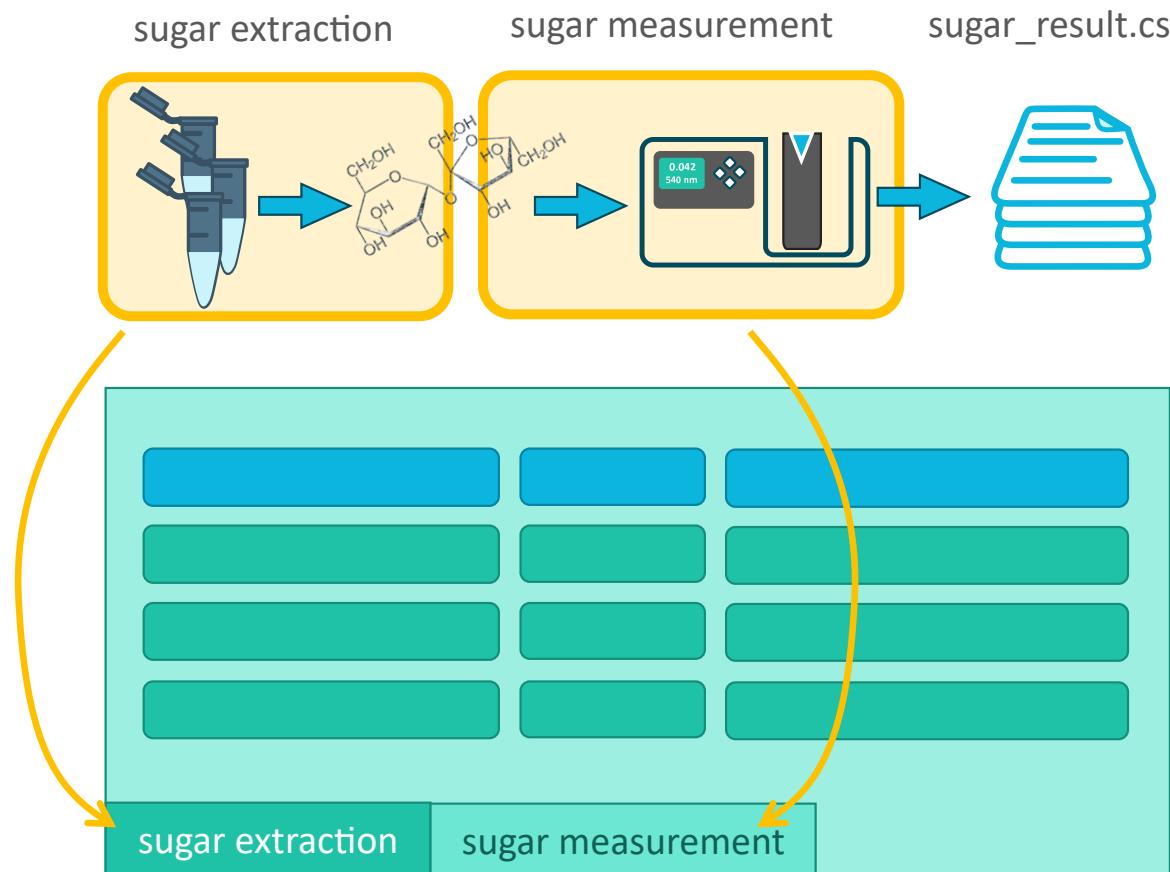
# Assay for sugar measurement



# Separating different assay elements

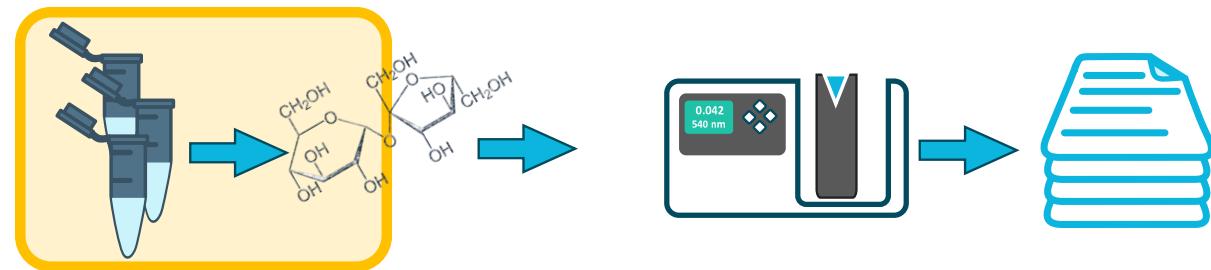


# Isolating the lab processes in an assay



# Parameteterization: sugar extraction

- Vortex Mixer
  - 3 seconds
- Temperature
  - 95 degree celsius



# Parameteterization: sugar measurement

- technical replicate
  - 1,2,3,...
- sample volume
  - 10 microliter
- buffer volume
  - 190 microliter
- cycle count
  - 5



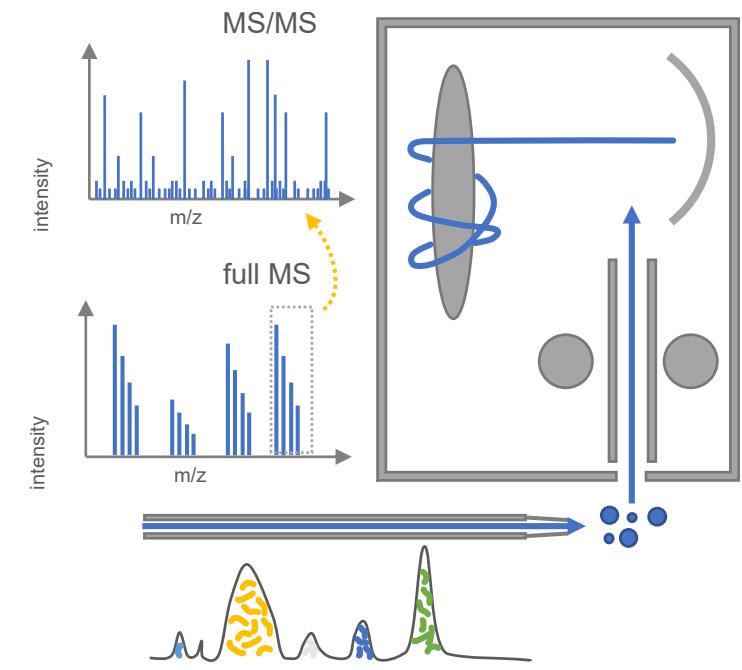
# Save time using standard methods and SOPs

## Parameter []

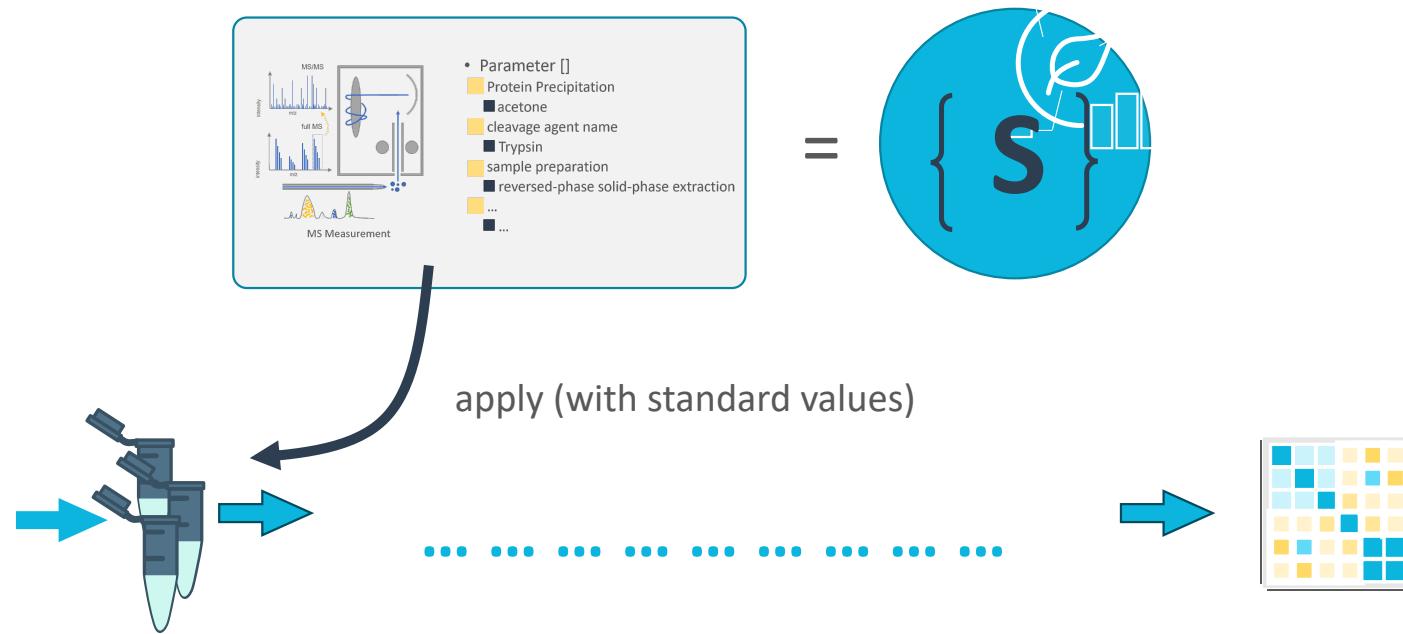
- █ Protein Precipitation
  - █ acetone
- █ cleavage agent name
  - █ Trypsin
- █ sample preparation
  - █ reversed-phase solid-phase extraction
- ...

## Component []

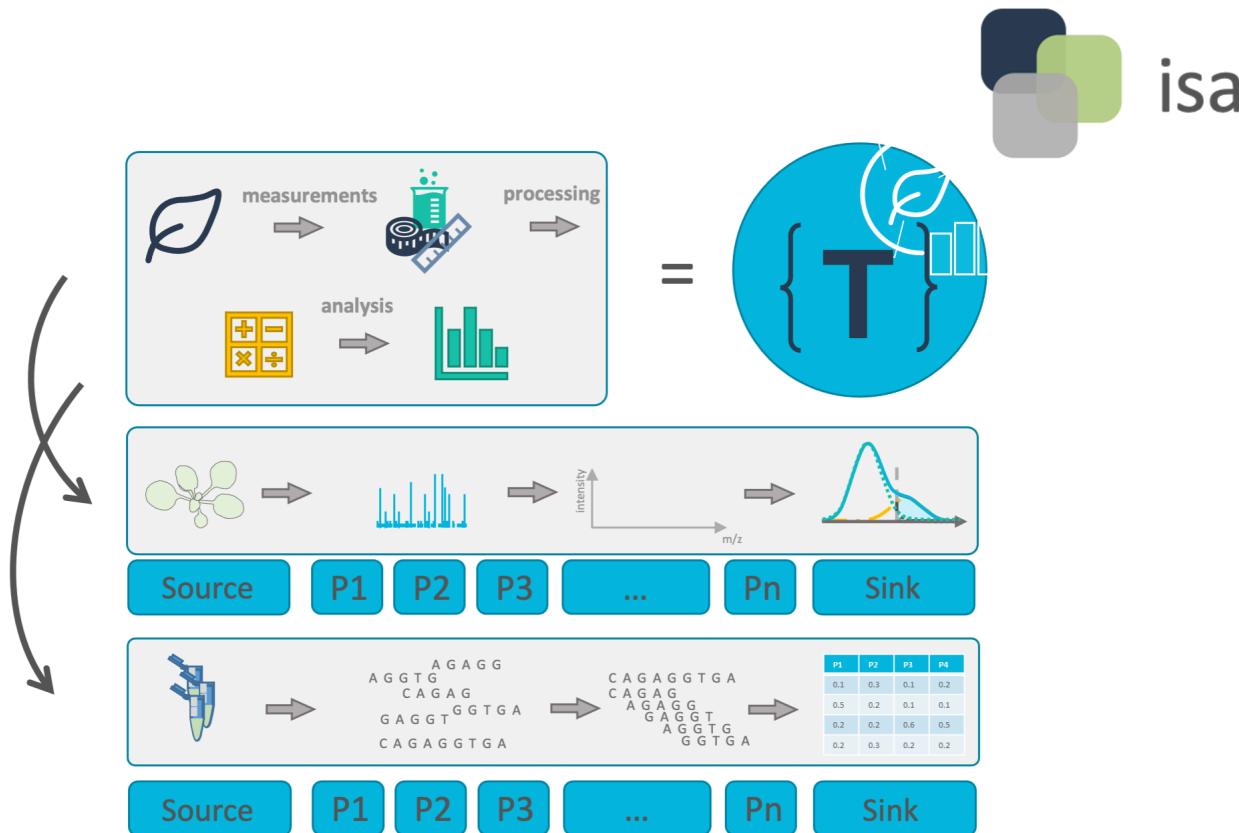
- █ chromatography instrument model
  - █ nanoElute2
- █ chromatography column model
  - █ PepSep C18 1.9 $\mu$ , 25cm x 75 $\mu$ m
- ...  
...



# Applying standard procedures to sample record



# Realization of lab-specific metadata with templates



Facilities can define their most common workflows as templates

## Hands-on part 2: ARCitect (and Swate)

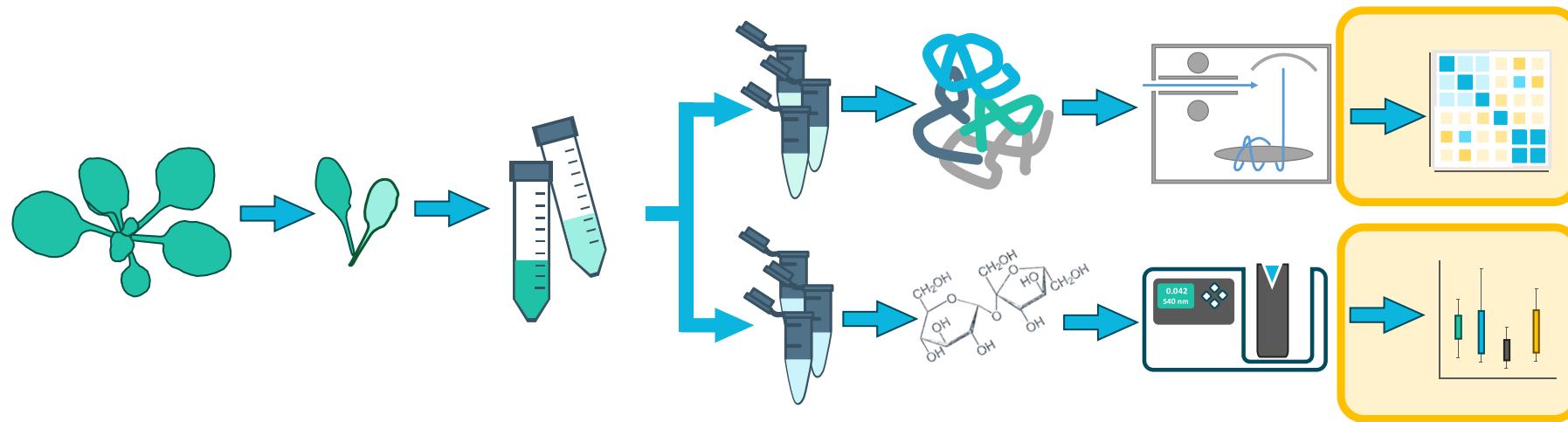
Continue the [Start Here](#) guide in the DataPLANT knowledge base.



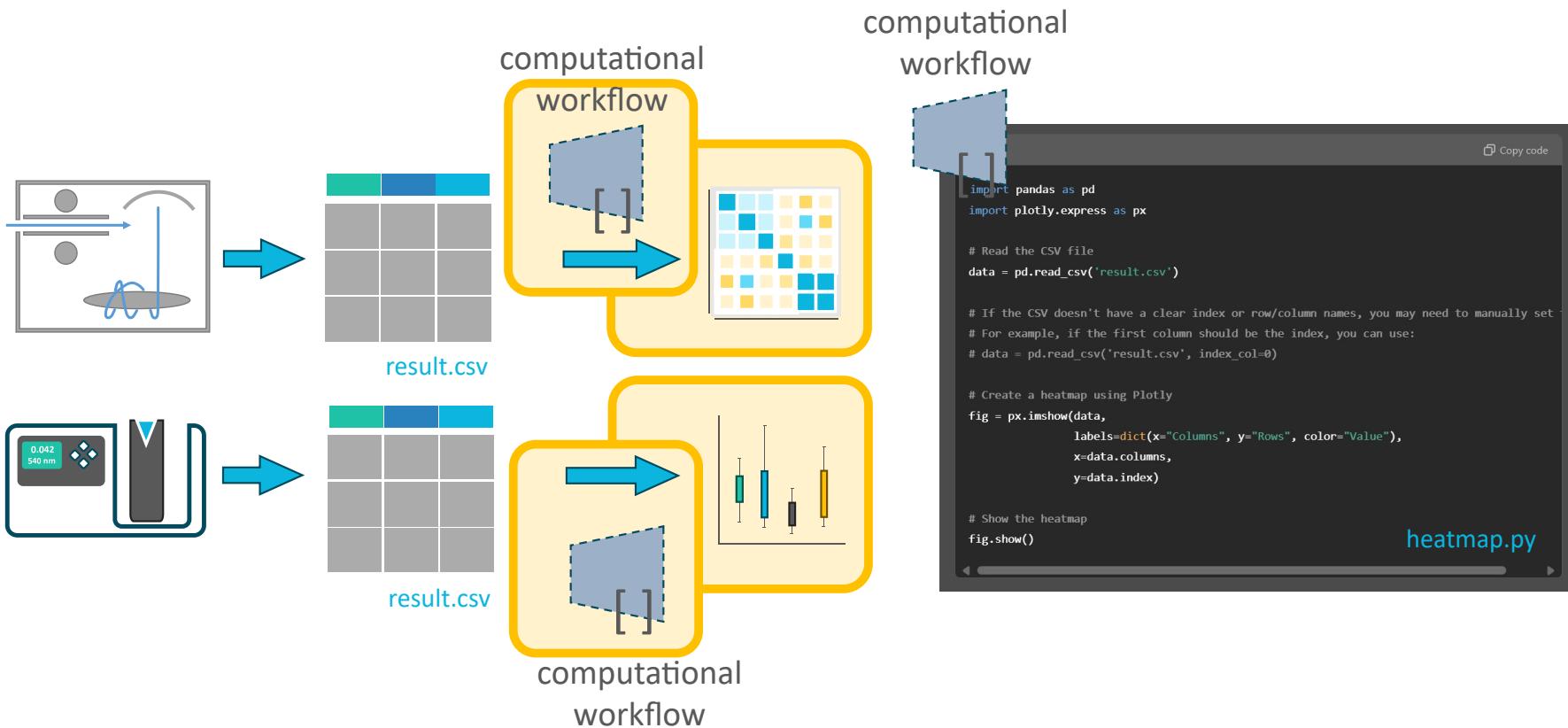
Stop after step [Add assay data](#)

## Hands-on part 3: Data

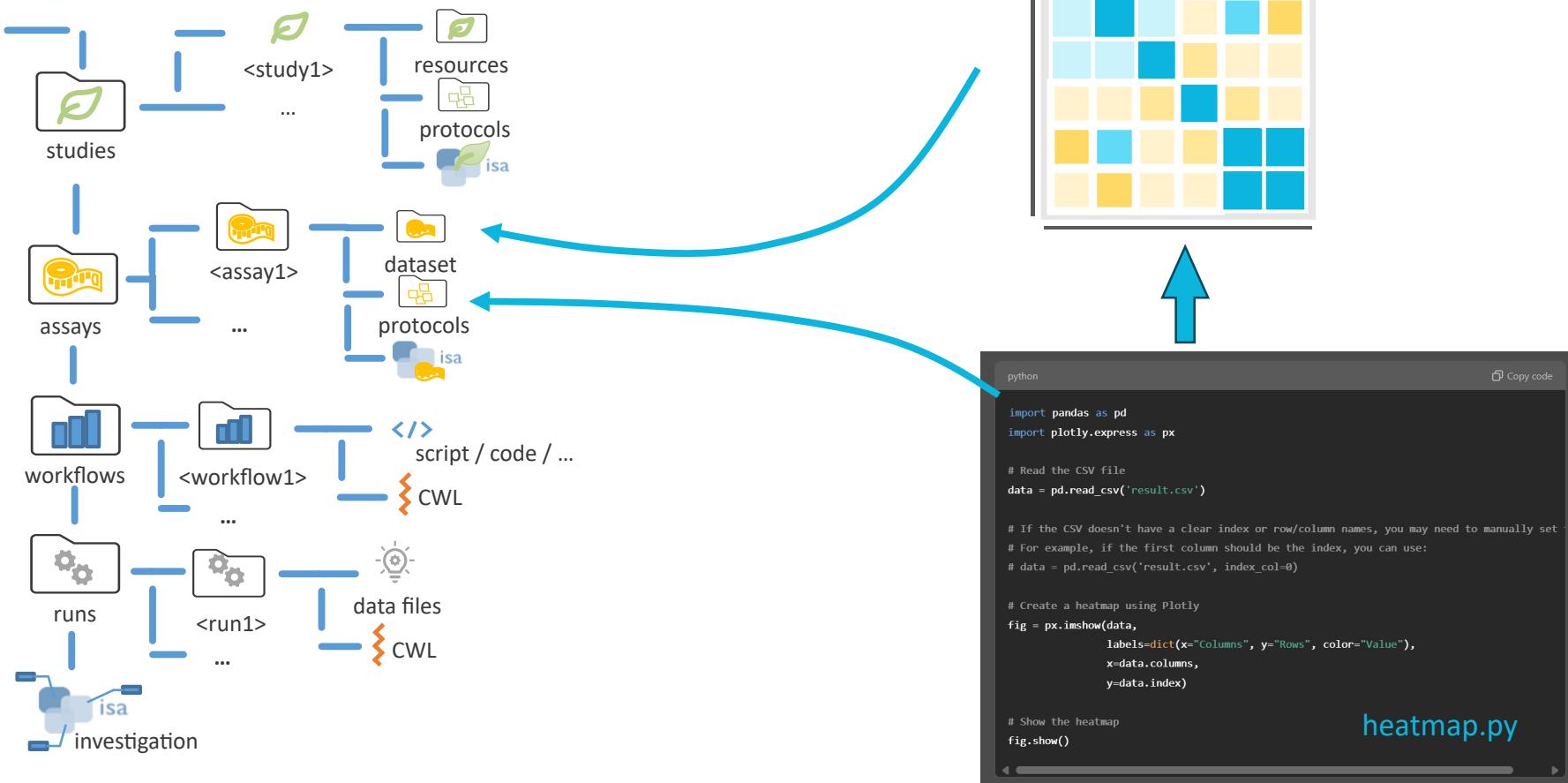
## Options to annotate the data analysis



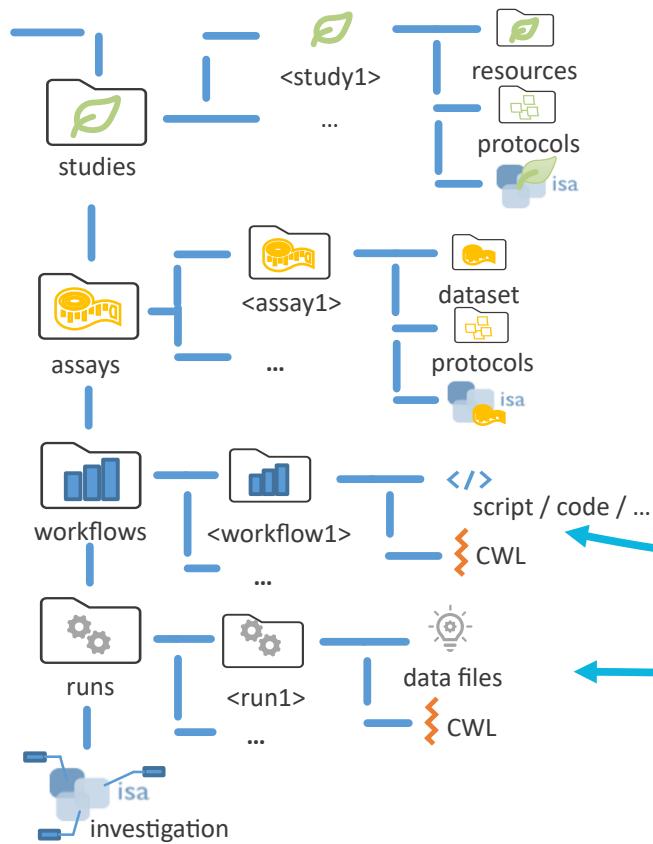
# A computational workflow is like a protocol



# Option I: Create a virtual assay



# Option II: Create a workflow and run



A screenshot of a code editor window titled "heatmap.py" containing the following Python script:

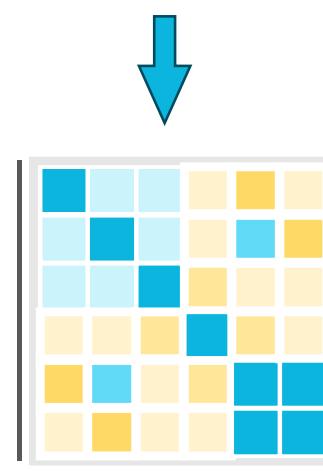
```
python
import pandas as pd
import plotly.express as px

# Read the CSV file
data = pd.read_csv('result.csv')

# If the CSV doesn't have a clear index or row/column names, you may need to manually set -
# For example, if the first column should be the index, you can use:
# data = pd.read_csv('result.csv', index_col=0)

# Create a heatmap using Plotly
fig = px.imshow(data,
                 labels=dict(x="Columns", y="Rows", color="Value"),
                 x=data.columns,
                 y=data.index)

# Show the heatmap
fig.show()
```



# Use CWL to wrap your workflow

CWL workflow document (\*.cwl)

```
graph LR; A["CWL workflow document (*.cwl)"] --> B["CWL job parameter (*.yaml)"]; B --> C["CWL runner"]; C --> D["output folder"]
```

1 → 2 → 3

```
#!/usr/bin/env cwl-runner

cwlVersion: v1.0
class: Workflow
inputs:
  tarball: File
  name_of_file_to_extract: string

outputs:
  compiled_class:
    type: File
    outputSource: compile/classfile

steps:
  untar:
    run: tar-param.cwl
    in:
      tarfile: tarball
      extractfile: name_of_file_to_extract
      out: [extracted_file]

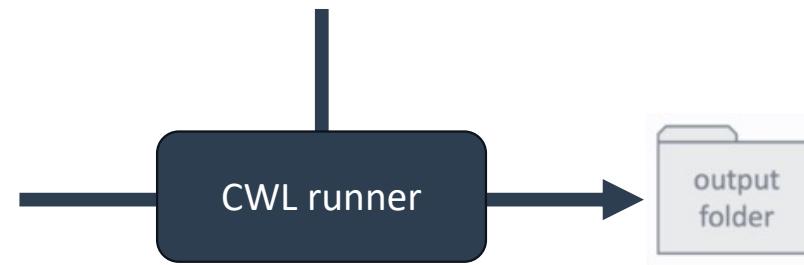
    compile:
      run: arguments.cwl
      in:
        src: untar/extracted_file
        out: [classfile]
```

CWL  
tool descriptors (\*.cwl)

CWL job parameter (\*.yaml)

```
job
yaml

file: fastq
param: 5
workflow: wf.cwl
output_folder:
  /temp
```



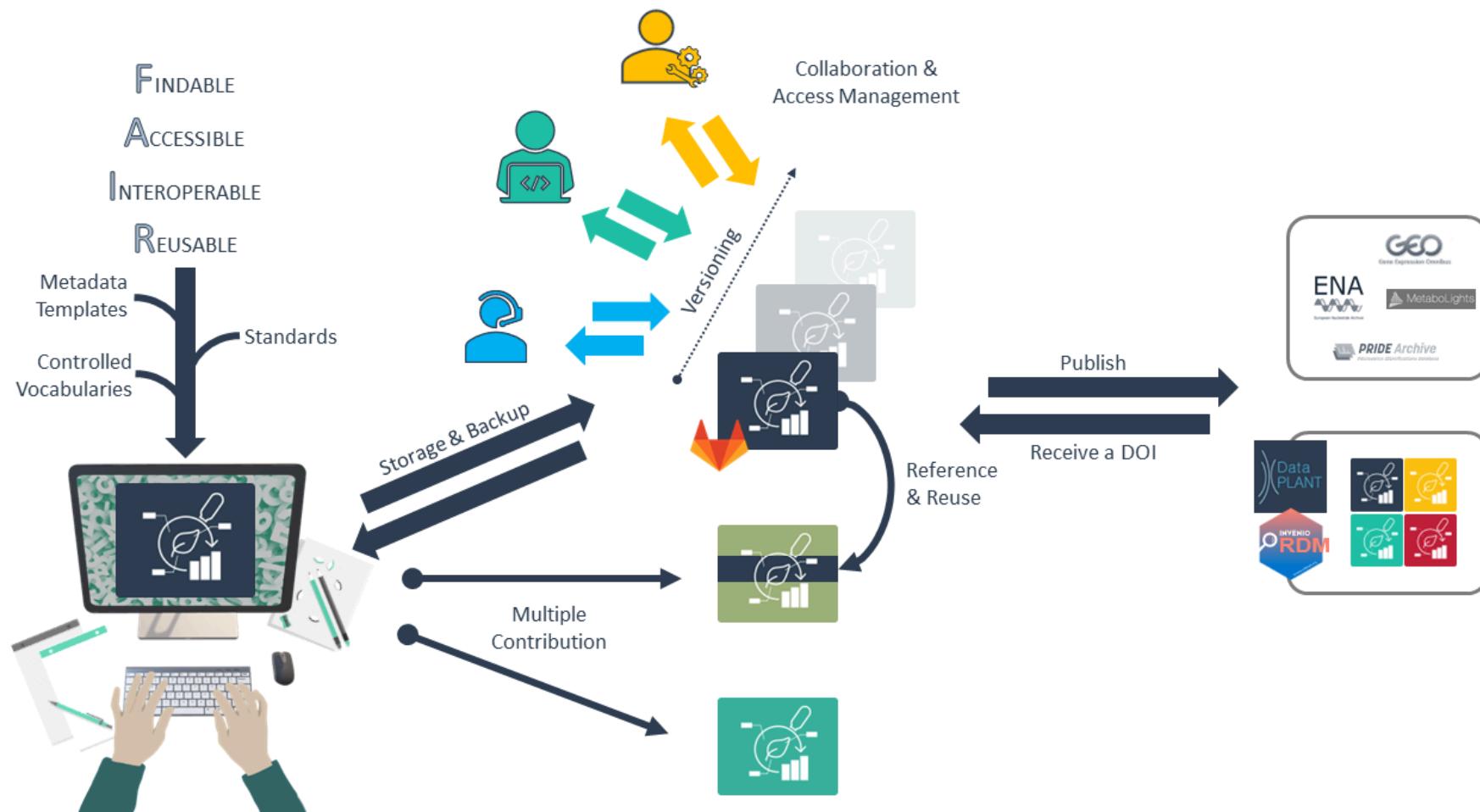
## Hands-on part 3: Data

Continue the [Start Here](#) guide in the DataPLANT knowledge base.

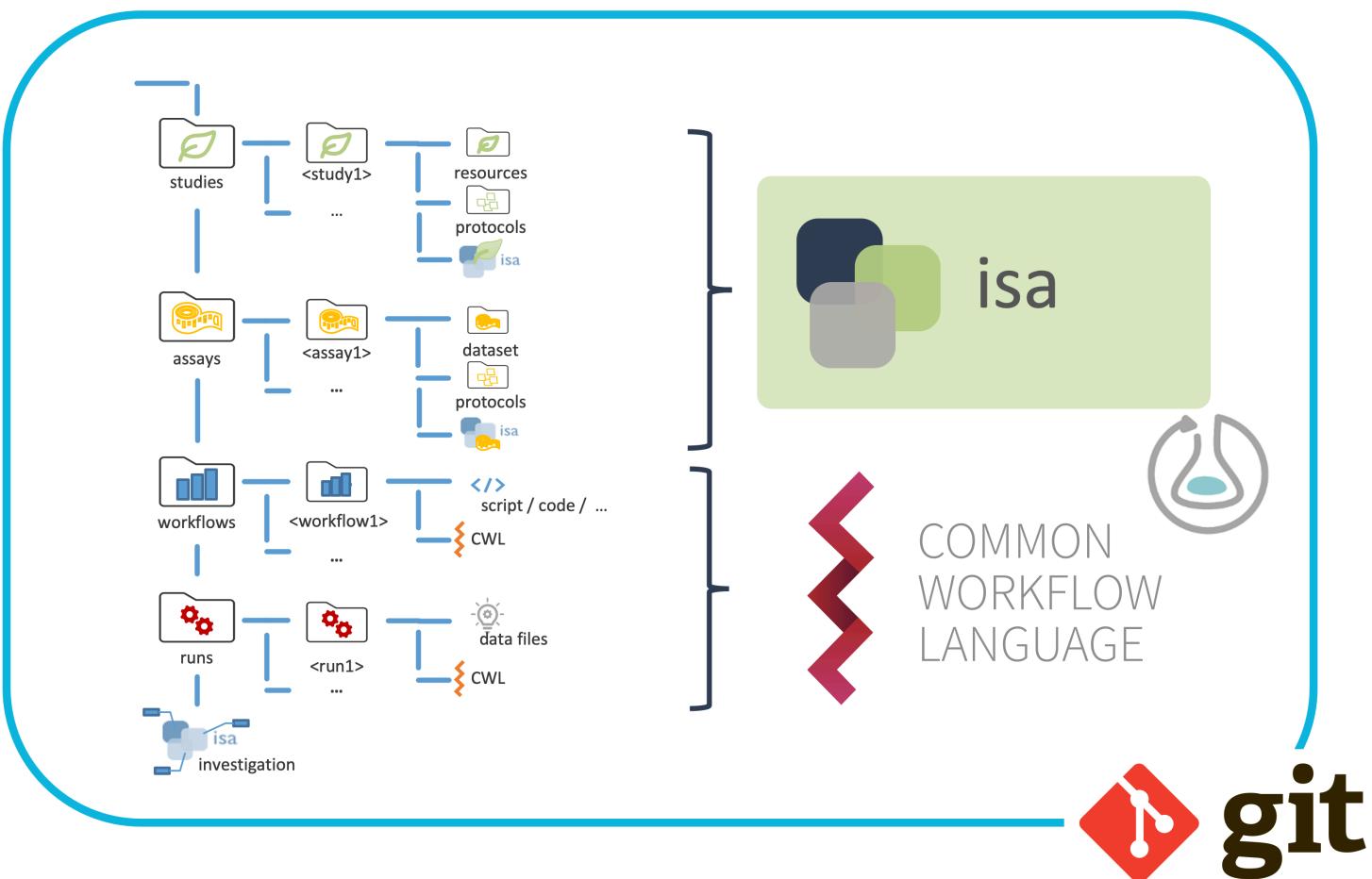
-  We will skip the **Add a datemap** step
-  Do the **Data analysis** step and stop

# Hands-on part 4: DataHUB

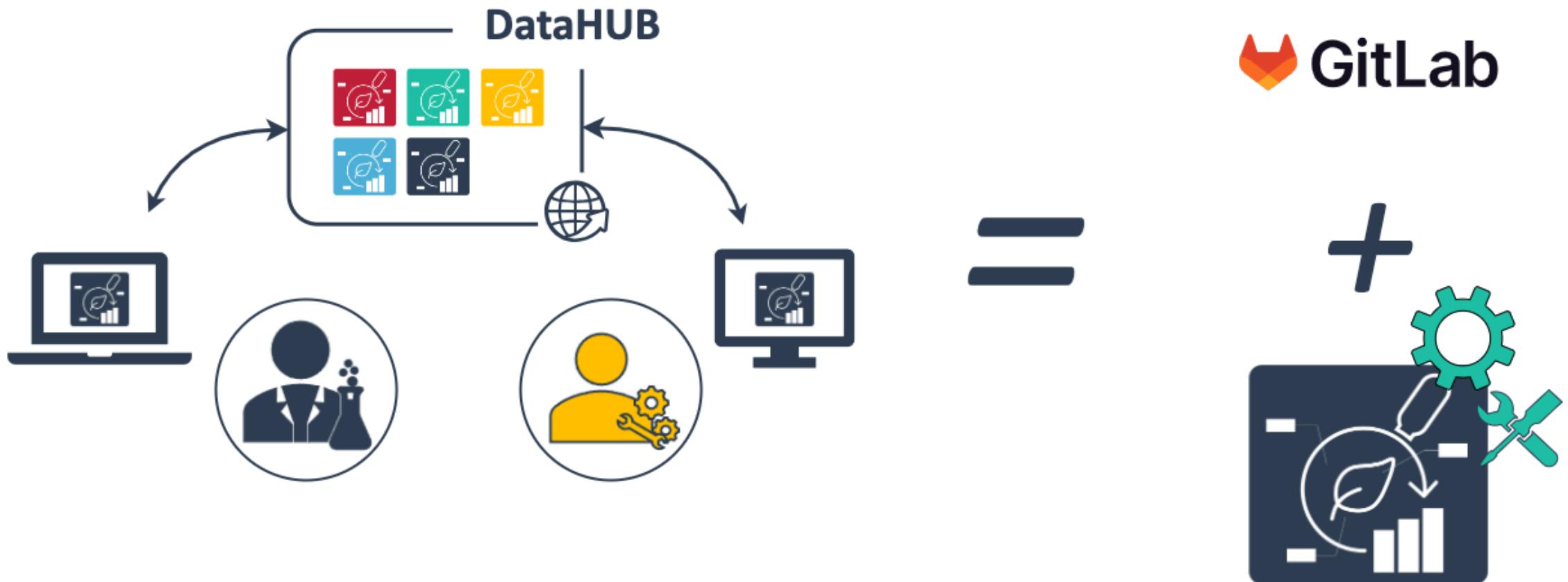
# DataHUB overview



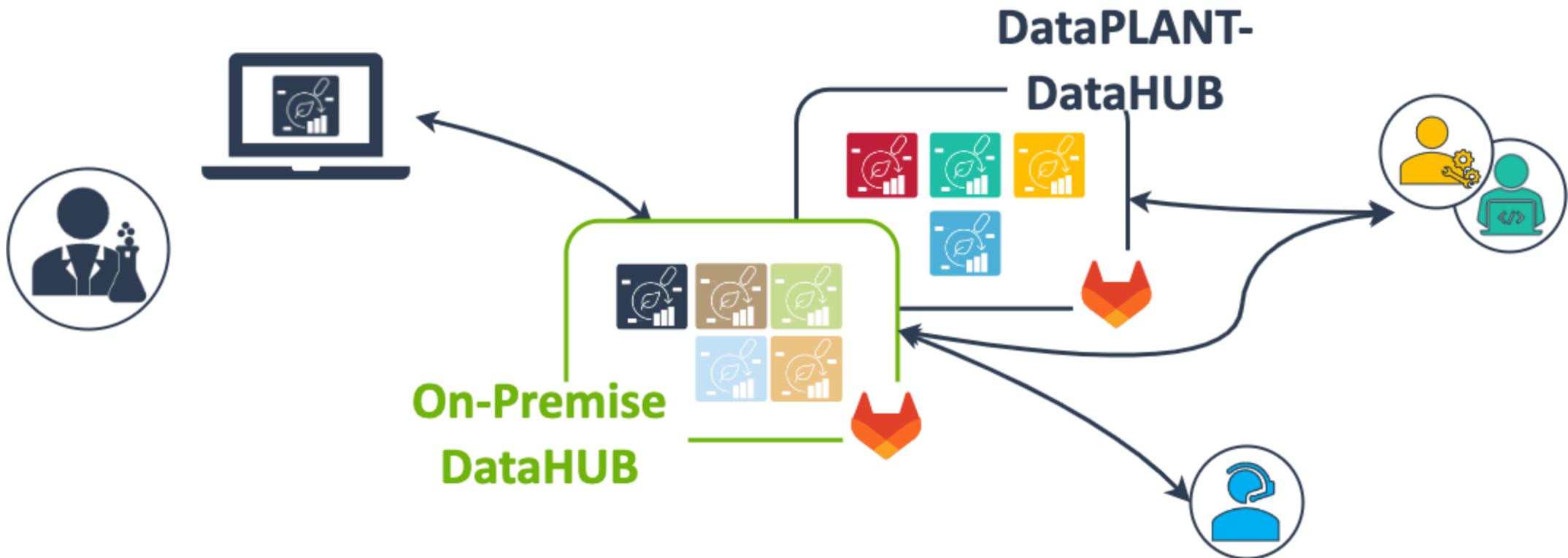
# ARC builds on standards + Git



# The DataPLANT DataHUB – a GitLab *Plus*

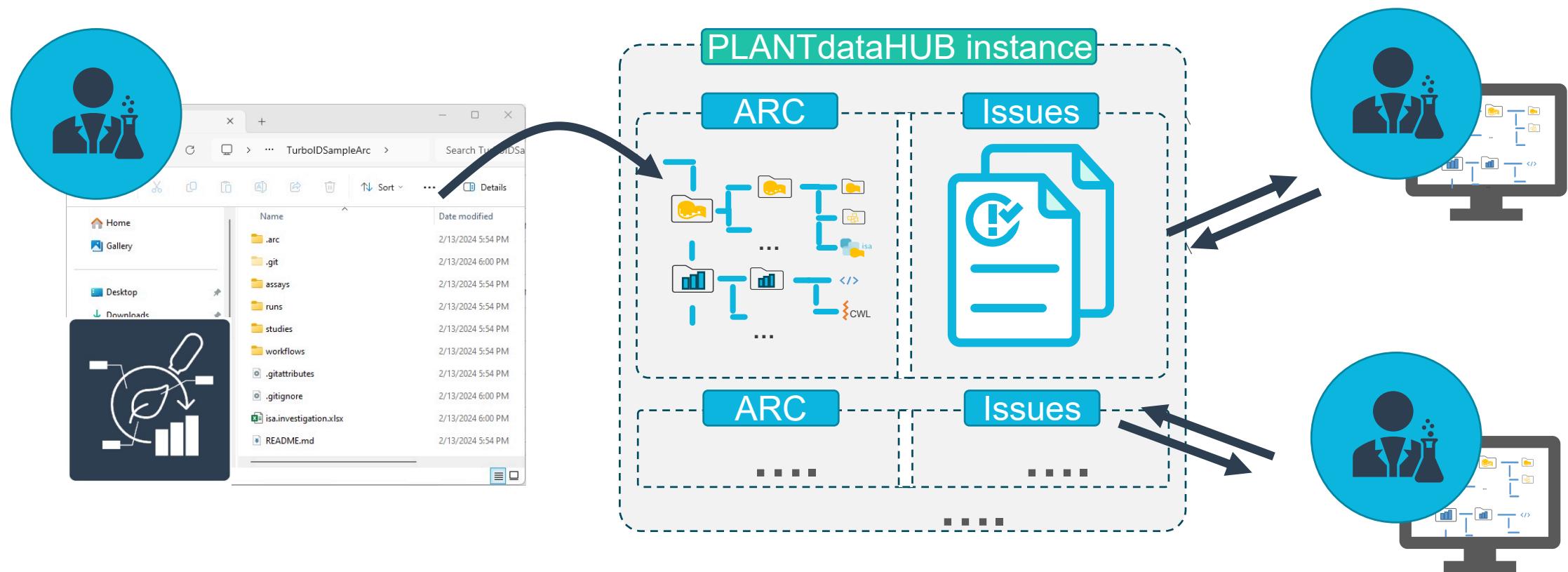


# On-premise DataHUBs

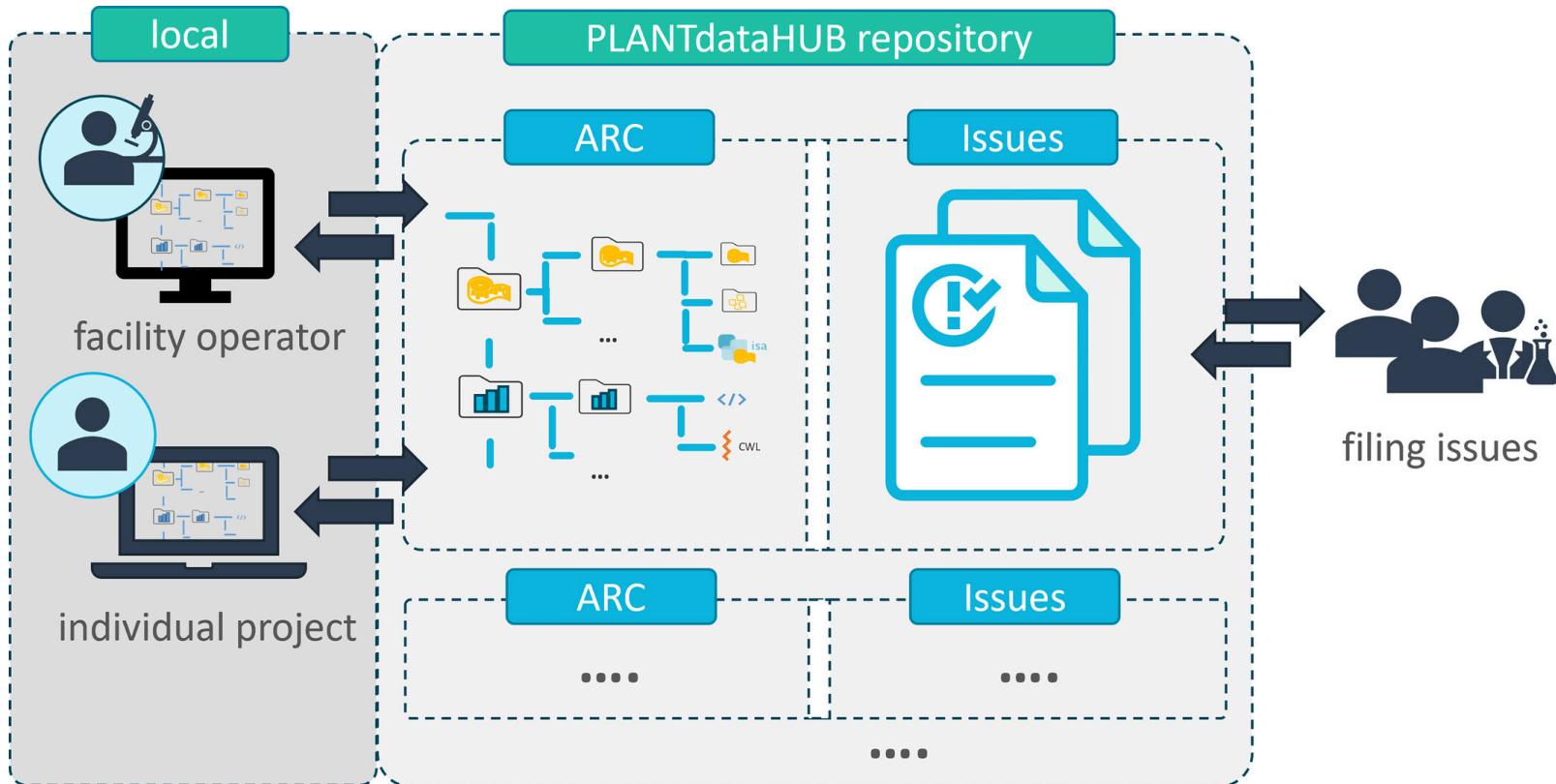


ARC services are available as on-premise option

# Using the DataHUB to collaborate



# Project management



# ARCs come with their own wiki space

- directly associated to your ARC
- same access rights as your ARC
- share meeting minutes or ideas with collaboration partners
- keep ARC clean of files that are not considered "research data"

The screenshot shows a DataHub interface for a 'Demo\_ARC' project. The left sidebar includes options for 'Pinned', 'Manage', 'Plan', 'Issues' (0), 'Issue boards', 'Milestones', and 'Wiki'. The main content area is titled 'Home' and contains a message about CQC pipelines, a 'Last edited by Demo User just now' notice, and a placeholder 'This is the wiki to Demo\_ARC. We will announce meeting schedules here.' Below this are sections for 'Meeting Schedule' (listing events like Kick-off, Proposal discussion, and RNA-seq pipeline) and 'Ideas and drafts' (listing a Golden Gate protocol).

Demo User / Demo\_ARC / Wiki / Home

CQC pipelines on DataHUB received major updates and are now opt-in. If you used the DataHUB publication service before, head [here](#) to learn how to opt-in.

> Pages 2

Home

Last edited by Demo User just now

This is the wiki to Demo\_ARC. We will announce meeting schedules here.

Meeting Schedule

- 2024-06-12 Kick-off
- 2024-06-27 Proposal discussion
- 2024-07-04 Discuss RNA-seq pipeline

Ideas and drafts

- Golden Gate protocol

## Hands-on part 4: DataHUB

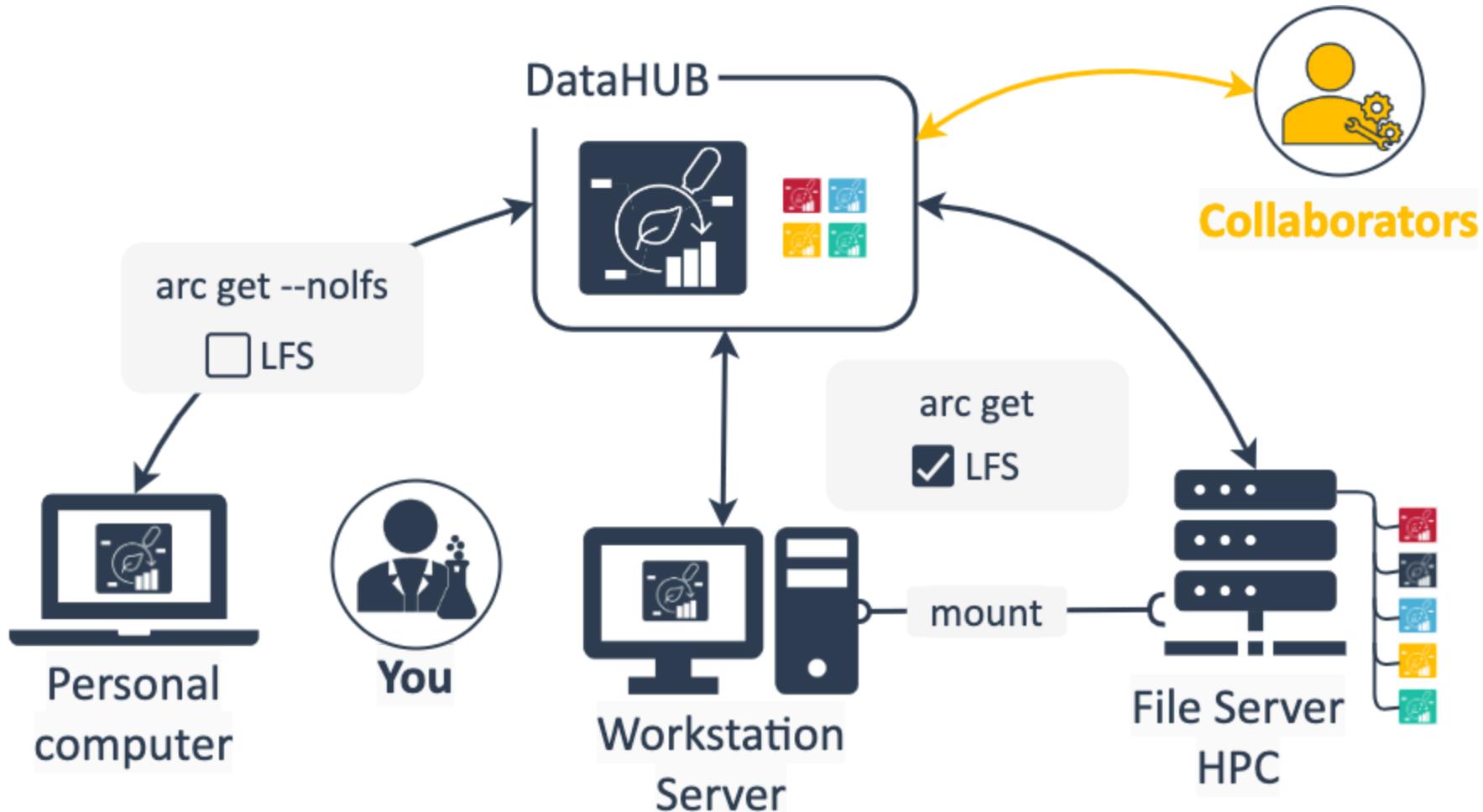
Continue the [Start Here](#) guide in the DataPLANT knowledge base.



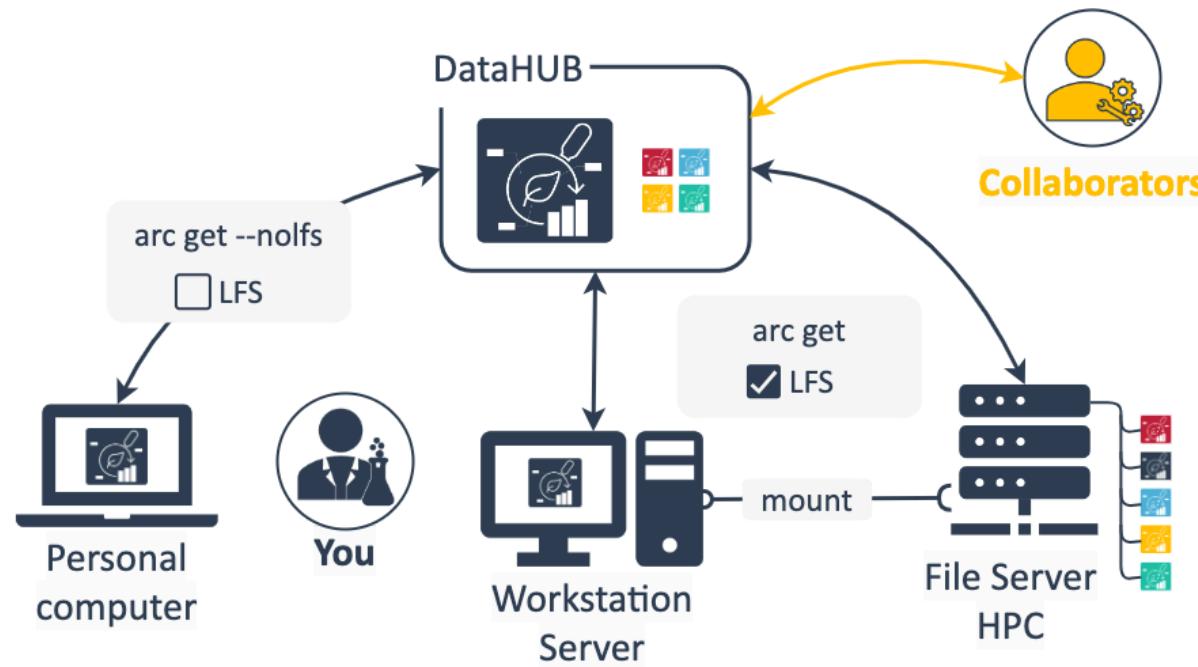
Until the final **complete** step

# DataHUB concepts and terminology

# Where do I store my ARC?



# ARC storage and sharing



- DataHUB as "ground truth" / original clone
- You can sync and communicate all changes to your ARC via the DataHUB

💡 ARCitect and ARC commander provide options to avoid syncing large files (LFS = Large file storage)

# Example setup to store and use ARCs

## Personal computer

- work on small files
- annotate metadata
- add scripts, protocols

## Workstation / Server

- work on large files
- run computations

## FileShare

- mount to local machine, sync ARC from there

## HPC

- direct connection HPC to DataHUB (depends on security settings)
- or mount to local machine and sync, ARC from there

# Projects and Groups are not the same

- "Project" = ARC
- "Groups" = Group of users

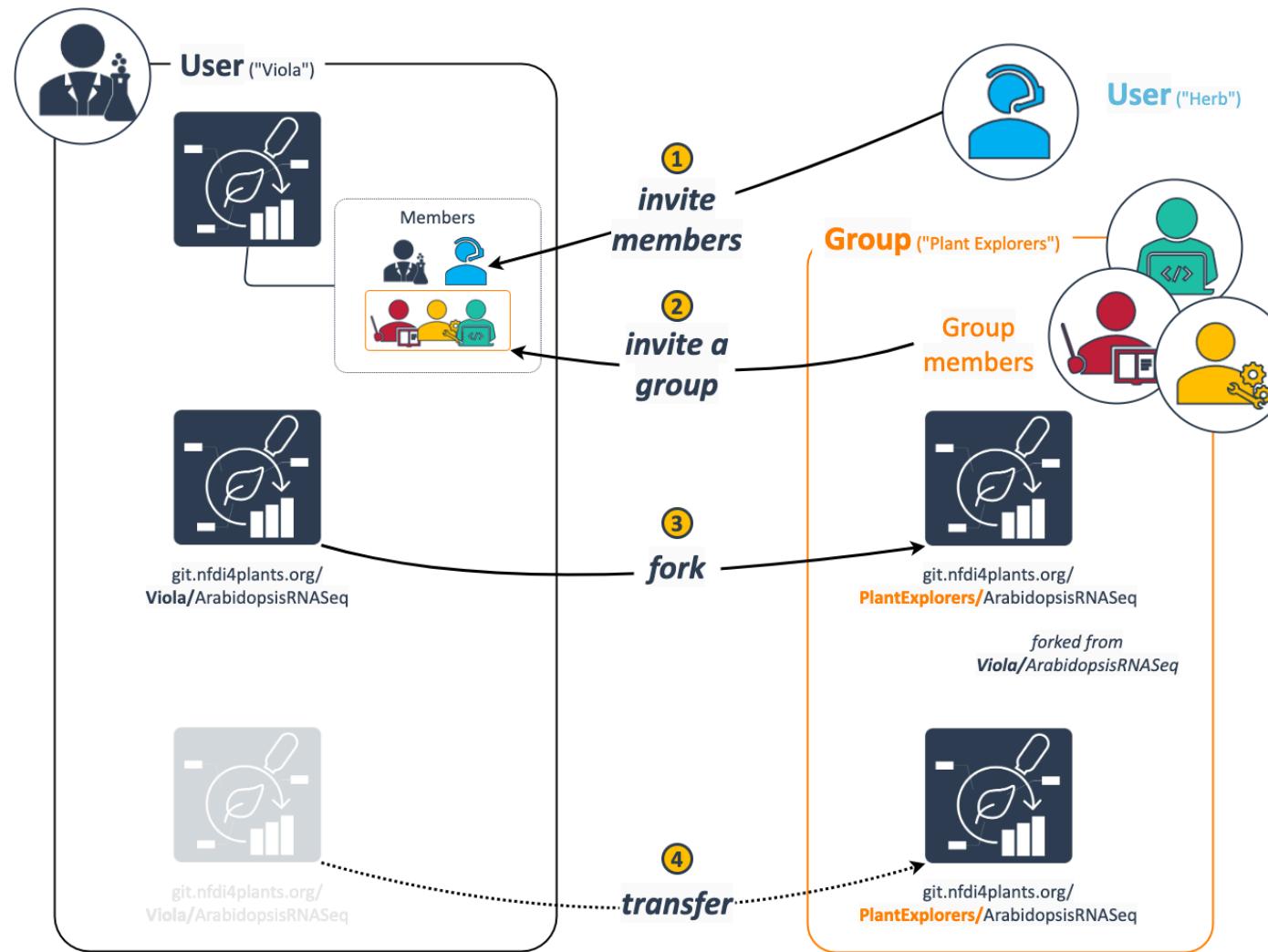
# Project = ARC

- In the DataHUB, ARCs are called "projects"; they are the same.
- An ARC can be shared with individual users (invited as "members") or a group.

# DataHUB Groups

- A "Group" is a group of users with specific permissions
- A group can share ARCs
- A group can be invited to an ARC
- Groups can have subgroups

# Options to share an ARC via the DataHUB



# Namespaces

- Every user has a personal namespace, where they can upload or create new ARCs
- Every group and subgroup has its own namespace

Type	URL	Namespace	Name
A personal ARC	<a href="https://git.nfdi4plants.org/brilator/Facultative-CAM-in-Talinum">https://git.nfdi4plants.org/brilator/Facultative-CAM-in-Talinum</a>	brilator	Dominik Brilhaus
An group-shared ARC	<a href="https://git.nfdi4plants.org/hhu-plant-biochemistry/Samuilov-2018-BOU-PSP">https://git.nfdi4plants.org/hhu-plant-biochemistry/Samuilov-2018-BOU-PSP</a>	hhu-plant-biochemistry	HHU Plant Biochemistry

 **Personal** is not the same as **private**

# Visibility

The visibility of ARCs and groups can be managed individually for each ARC or group

## Visibility

The visibility of each ARC can be managed in the settings of the ARC



**Private** – ARC access must be granted explicitly to each user or group.



**Internal** – ARC can be accessed by any logged in user.



**Public** – ARC can be accessed without authentication.



By default every ARC and every group is set to **private**.

# Permissions & Roles

When inviting new members to an ARC or group, you can choose between different levels.

## Permissions & Roles

Roles are assigned when adding a user to an ARC or to a group. This is a very simplified summary.

**Guest** – Can only see the ARC's wiki

**Reporter** – Can read, but not add or edit data

**Developer** – Reporter permissions + can read, add, and edit data

**Maintainer** – Developer permissions + can add new members

**Owner** – Maintainer + can delete ARC, manage memberships and permissions

💡 By default you are **Owner** of an ARC you create or upload to the DataHUB.

# ARC DataHUB members // ARC Investigation contacts

The screenshot shows the 'Members' section of a project named 'Samuilov-2018-BOU-PSP'. It includes a search bar, buttons for 'Import from a project', 'Invite a group', and 'Invite members', and a table with columns for Account, Source, Max role, Expiration, and Activity. The table lists four members:

Account	Source	Max role	Expiration	Activity
Adriano Nunes-Nesi @unesnesi	HHU Plant Biochemistry by Sebastian Triesch	Maintainer	Expiration date	User created: Jul 05, 2023 Access granted: Jul 10, 2023 Last activity: Aug 09, 2023
Andreas Weber @andreas.weber	HHU Plant Biochemistry by Sebastian Triesch	Maintainer	Expiration date	User created: Mar 10, 2023 Access granted: Jul 31, 2023 Last activity: Sep 11, 2023
Dominik Brilhaus @brilator	Direct member by Dominik Brilhaus	Owner	Expiration date	User created: Feb 21, 2022 Access granted: Dec 06, 2023 Last activity: Mar 26, 2024
Franziska Fichtner @franziska.fichtner	HHU Plant Biochemistry by Sebastian Triesch	Maintainer	Expiration date	User created: Aug 11, 2023 Access granted: Aug 11, 2023 Last activity: Aug 11, 2023

**DataHUB: ARC members**

[https://git.nfdi4plants.org/hhu-plant-biochemistry/Samuilov-2018-BOU-PSP/-/project\\_members](https://git.nfdi4plants.org/hhu-plant-biochemistry/Samuilov-2018-BOU-PSP/-/project_members)

The screenshot shows the 'Samuilov-2018-BOU-PSP' page in ARCIct. It includes sections for Identifier, Title, Description, and Contacts. The 'Contacts' section lists ten individuals with their names, ORCID IDs, and scores (4/10 or 3/10). The 'Identifier' section shows the project name and path.

Contact	Score
Sladjana Samuilov <orcid>	4/10
Nadine Rademacher <orcid>	3/10
Samantha Flachbart <orcid>	3/10
Leila Arab <orcid>	3/10
Saleh Alfarraj <orcid>	3/10
Franziska Kuhnert <orcid>	3/10
Stanislav Kopriva <orcid>	3/10
Andreas P. M. Weber <orcid>	4/10
Tabea Mettler-Altmann <orcid>	3/10

**ARCIct: Investigation Contacts**

Investigation contacts are not automatically invited as members to the ARC.

# elabFTW Introduction

# Login into eLabFTW HHU

The screenshot shows the login interface for eLabFTW HHU. At the top center is the eLabFTW logo, which consists of a stylized flask icon containing blue liquid and the text "eLabFTW". Below the logo is a horizontal line. Underneath the line, the text "Über Ihre Institution anmelden" is displayed. Below this text is a sub-instruction "Wählen Sie einen Identitätsanbieter aus". A dropdown menu is open, showing the option "HHU-IDM". To the right of the dropdown is a checked checkbox labeled "Anmeldung merken". Below the checkbox is a teal-colored button labeled "Anmelden". At the bottom left of the page, there are icons for social media sharing (Twitter, LinkedIn, Facebook, and a speech bubble) and a link to "DATENSCHUTZ-BESTIMMUNGEN". At the bottom right, there is a note indicating the page was generated in 0.00437 seconds.

Über Ihre Institution anmelden

Wählen Sie einen Identitätsanbieter aus

HHU-IDM

Anmeldung merken

Anmelden

Twitter LinkedIn Facebook DATENSCHUTZ-BESTIMMUNGEN

Bereitgestellt von eLabFTW  
Seite generiert in 0.00437 Sekunden

# eLab-FTW - Dashboard

Dashboard  
Welcome Sabrina

---

Experiments [Create](#) Resources [Create](#) Scheduled bookings [Scheduler](#)

---

Browse by category

[IMAGING](#) [BEHAVIOUR](#) [METABOLISM](#) [DISEASE](#) [NEURO](#) [ASSAY](#) [STUDY](#)

Browse by status

[RUNNING](#) [SUCCESS](#) [REPEAT](#) [FAIL](#) [DRAFT](#)

Last modified experiments

[INV0001\\_A1\\_Sugar measurement](#) 23 seconds ago  
[INV0001\\_A1\\_Sugar Extraction](#) 1 minute ago  
[Growth Protocol for Arabidopsis thaliana](#) 9 minutes ago

---

Browse by category

[STANDARD](#) [INVESTIGATION](#) [STUDY](#)

Last modified resources

[STUDY](#) [INV0001\\_S1\\_AthalianaColdStress](#) 15 minutes ago  
[INVESTIGATION](#) [INV0001\\_A.thaliana\\_ColdStress\\_Sugar](#) 17 minutes ago

# eLabFTW - Experiment panel

The screenshot shows the eLabFTW Experiment panel. At the top, there is a navigation bar with links for 'EXPERIMENTS' (which is highlighted in blue), 'RESOURCES', 'TEAM', 'SEARCH', and 'DOCUMENTATION'. To the right of the navigation bar are icons for search, notifications, help, and user profile.

The main area is titled 'Experiments' and contains a table of experiment entries. The table has columns for 'Date', 'Title', 'Next step', 'Category', 'Status', 'Tags', 'Rating', and 'Owner'. There are also sorting and filtering options at the top of the table.

Two experiments are listed:

- INV0001\_A1\_Sugar measurement**: Date 2025-01-09, Status SUCCESS, Tags A.thaliana, INV001, sugar. Owner: Sabrina Zander.
- INV0001\_A1\_Sugar Extraction**: Date 2025-01-09, Status RUNNING, Tags A.thaliana, INV001, sugar. Owner: Sabrina Zander.

A 'Load more' button is located at the bottom of the table.

At the bottom of the page, there are links for 'PRIVACY POLICY' and 'TERMS OF SERVICE'. On the right, it says 'Powered by eLabFTW 5.1.12' and 'Made with ❤ by Deltablot'.

Default "working area" for protocols of daily lab work

# eLabFTW - Resource panel

Screenshot of the eLabFTW Resource panel interface.

The top navigation bar includes links for EXPERIMENTS, **RESOURCES**, TEAM, SEARCH, and DOCUMENTATION, along with a search bar, notification bell, help icon, and user profile icon.

## Resources

Expand all - Select all

Date	Title	Next step	Category	Status	Tags	Rating	Owner
2025-01-09	INV0001_S1_AthalianaColdStress		STUDY	A.thaliana INV001			Sabrina Zander
2025-01-09	Growth Protocol for Arabidopsis thaliana		METHODE	A.thaliana			Sabrina Zander
2025-01-09	INV0001_A.thaliana_ColdStress_Sugar		INVESTIGATION	A.thaliana INV001			Sabrina Zander

**Create**

Load more

X | Q | M | S | PRIVACY POLICY | TERMS OF SERVICE | Powered by eLabFTW 5.1.12  
Made with ❤ by Deltaplant

Place to categorize content

# eLabFTW - Experiment information

The screenshot shows the eLabFTW experiment information page for 'INV0001\_A1\_Sugar Extraction'. The page includes fields for Started on (09.01.2025), ID (185), Category (Assay), Status (Running), Tags (A.thaliana, INV001, sugar), and MAIN TEXT (Last saved: 6 minutes ago). Annotations highlight several features:

- signature**: Points to the signature icon in the top right.
- timestamp**: Points to the timestamp icon in the top right.
- different export functions**: Points to the download icons in the top right.
- status**: Points to the status dropdown menu.
- tags**: Points to the tags input field.
- sharing**: Points to the sharing settings below the tags.
- lock**: Points to the lock icon in the top right.
- request action**: Points to the request action icon in the top right.
- pin**: Points to the pin icon in the top right.
- More options**: Points to the 'More options' button in the top right.
- Transfer ownership**: Points to the Transfer ownership option in the dropdown menu.
- See revisions**: Points to the See revisions option in the dropdown menu.
- See changelog**: Points to the See changelog option in the dropdown menu.
- Archive/Unarchive**: Points to the Archive/Unarchive option in the dropdown menu.
- Delete**: Points to the Delete option in the dropdown menu.

# eLabFTW - Tag usage

Very individual, some examples:

- Investigation ID
- Subgroup
- Organism /part (e.g. A.thaliana, leaf, root, mitochondria)
- Goal (strain generation, plasmid generation)
- Methode (e.g. transformation, western blot, extraction, biolector)
- Genes/ Proteins (e.g. mKate2, eGfp, Rrm4)

Use keywords you would search/filter for



Make rules for tag usage

# eLabFTW - Extra fields

Add an extra field X

Group +  
Undefined group

Field type

- Text
- Text
- Dropdown menu
- Radio buttons
- Date
- Date and time
- Email
- Time
- Checkbox
- Number
- Experiment
- Resource
- User
- URL

Close Save

# eLabFTW - Extra fields examples

Dropdown	URL	Number	Date
<p>Field type Dropdown menu</p> <p>Name (required)* Antibiotic used</p> <p>Description Select the antibiotic used</p> <p>Choices Kanamycin Streptomycin Ampicillin</p> <p>Add another choice</p> <p><input type="checkbox"/> Allow selection of multiple values?</p>	<p>EXTRA FIELDS</p> <p>UNDEFINED GROUP</p> <p>OMERO dataset <a href="https://omero-cai.hhu.de/webclient/?show=dataset-12881">https://omero-cai.hhu.de/webclient/?show=dataset-12881</a></p>	<p>Extra fields</p> <p>End date 06 / 09 / 2021</p> <p>Magnification 20X</p> <p>Pressure (Pa) 12</p> <p>Wavelength (nm) 488 405 647</p>	<p>Field type Number</p> <p>Name (required)* Drug concentration</p> <p>Description Indicate the drug concentration</p> <p>Available units ng/mL µg/mL mg/mL</p> <p>Add another choice</p>

# eLabFTW - Experiment categories examples

- Study
- Assay
- Discussion
- Teaching
- Demo
- R&D
- Production

# eLabFTW - Categories in Resources

Use Resources to categorize content & manage lab.

Use unique identifier & naming conventions ( e.g. Methode MTH0001, Media MED0001, Plasmid PL0001).

- List of Investigation
- Investigation
- Plant sheet
- Plasmid sheet
- Equipment
- Collections
- Methods
- Manuals
- Media & Solution
- Formblatt Z

# eLabFTW - Scheduler

SCHEDULER   MEMBERS   EMAIL   PROCUREMENT REQUESTS BETA

Filter by category   Equipment - Zeiss AXIO Observer Z1 Inverted Fluorescence Microscope   Scope

< > today   6 - 12 Jan 2025   week list month

	Mon 06/01	Tue 07/01	Wed 08/01	Thu 09/01	Fri 10/01	Sat 11/01	Sun 12/01
00							
01							
02							
03							
04							
05							
06					6:00 - 12:30 (Sabrina Zander)		
07							
08							
09							
10							
11							
12							

# eLabFTW - Booking parameter

You can modify booking parameter for reach Resource

Zeiss AXIO Observer Z1 Inverted Fluorescence Microscope Create

[←](#)

Started on	09.01.2025 <a href="#"></a>	Custom ID <a href="#" style="background-color: #333; color: white; padding: 2px 5px;">Get next</a>
ID	115	Transfer ownership
Category	Equipment	See revisions
Status	Not set	See changelog
Tags	Add a tag	Archive/Unarchive

[Delete](#)

# eLabFTW - Templates

Create templates for experiments you often use

- „Skeleton“ of a real experiment
- Select „Templates“ from User menu
- You can have different permissions for the template itself and for Experiment that will be created from that template
- You can add tags, steps, links to it (e.g. link to project)
- Can also be imported directly in your text (allows merging several templates in one experiment)

# eLabFTW - Tipps and tricks

- Use # to link experiments/resources in the text
- Use .gb files for plasmid cards ( can be open in a plasmid viewer in elabFTW)
- Use versions, possibility to upload new versions of files/methods

# eLabFTW - Support

## General

<https://github.com/elabftw/elabftw>

<https://www.elabftw.net/> (Newsletter)

## HHU specific

[elabftw@hhu.de](mailto:elabftw@hhu.de)

[https://rocketchat.hhu.de/channel/elb\\_eln](https://rocketchat.hhu.de/channel/elb_eln)

<https://wiki.hhu.de/display/ZIM/eLabFTW>

# elab2ARC Documentation

<https://nfdi4plants.github.io/nfdi4plants.knowledgebase/resources/elab2arc/>

# Working Scenario

## 1. Use ARCs to

- Organize and store (raw) data locally
- Share and communicate data with collaboration partners
- Annotate and prepare data for publication

## 2. Use elabFTW to

- Document daily lab work „on the fly“
- Organize lab routines with colleagues

# eLabFTW & ARC Connection

## Aim

- Reducing duplication/  
additional workload
- Transformation of data  
from elabFTW into ARC
- User-friendly, easy  
implementation

## Challanges

- ELNs are unstructured;  
automated translation  
and interoperation into  
well-structured ARC  
challenging

## Solution

- Start with automated  
transition of data  
without
  - metadata mapping,
  - structured  
metadata  
extraction,
  - re-structuring of  
elabFTW content

# Conversion

## ① Content Preparation

ARC



eLabFTW entry

Experiment1 DNA Extraction

MAIN TEXT

Experiment Protocol:  
Goal: Extract DNA  
Procedure: Use Kit  
Result: Gel picture

ATTACHED FILE

Abbildung-14-Agarose-Gel-Elektrophorese-Unterschiedliche Konzentration-an-plasmidischer.png

## ② eLab2ARC Conversion

eLabFTW to ARC

eLabFTW Experiment ID: 45702

eLabFTW API Key: 34-780adfbf62259888

DataHub Access Token: n8UhrOvF9g4HtVz2R

Submit Options: One Click Submission, Step by Step

ARC updated Successfully

Files Changed in ARC

eLabFTW text and files

Experiment Protocol:  
Goal: Extract DNA  
Procedure: Use Kit  
Result: Gel picture

ARC files updated

ARC file path is arc/assays/Experiment1\_DNA\_Extraction/protocols/elabFTW\_protocol.md

## ③ Updated ARC content shown in ARCitect

assays

Experiment1\_DNA\_Extraction

dataset

0\_Abbildung-14-Agarose-Ge

readme.md

protocols

eLabFTW\_protocol.md

isa\_assay

# Conversion

The elab2ARC tool will automatically convert your eLabFTW experiments into ARC format

- create a new assay folder with eLabFTW experiment name as assay name
- create the assay folder structure (dataset/protocols/isa.assay)
- convert experiment main text into a .md file and store it in the protocol folder
- add all attachments of the eLabFTW experiment into the dataset folder
- enter name/email/affiliation of the eLabFTW experiment metadata into the isa.assay sheet

# elab2ARC Tool

Tool: <https://nfdi4plants.github.io/elab2arc/>

Documentation:

<https://nfdi4plants.github.io/nfdi4plants.knowledgebase/resources/elab2arc/>

Issues/Problems:

<https://github.com/nfdi4plants/elab2arc>

# Contributors

If not referenced otherwise, figures and slides presented here were created by members of DataPLANT (<https://nfdi4plants.org>).

Additional slides were contributed by

- name: Dominik Brilhaus  
github: <https://github.com/brilator>  
orcid: <https://orcid.org/0000-0001-9021-3197>
- name: Cristina Martins Rodrigues  
github: <https://github.com/CMR248>  
orcid: <https://orcid.org/0000-0002-4849-1537>
- name: Hajira Jabeen  
github: <https://github.com/HajiraJabeen>  
orcid: <https://orcid.org/0000-0003-1476-2121>
- name: Kevin Frey  
github: <https://github.com/Freymaurer>  
orcid: <https://orcid.org/0000-0002-8493-1077>
- name: Sabrina Zander  
orcid: <https://orcid.org/0009-0000-4569-6126>

