CEPLAS research data management (RDM)

Steering Committee Meeting

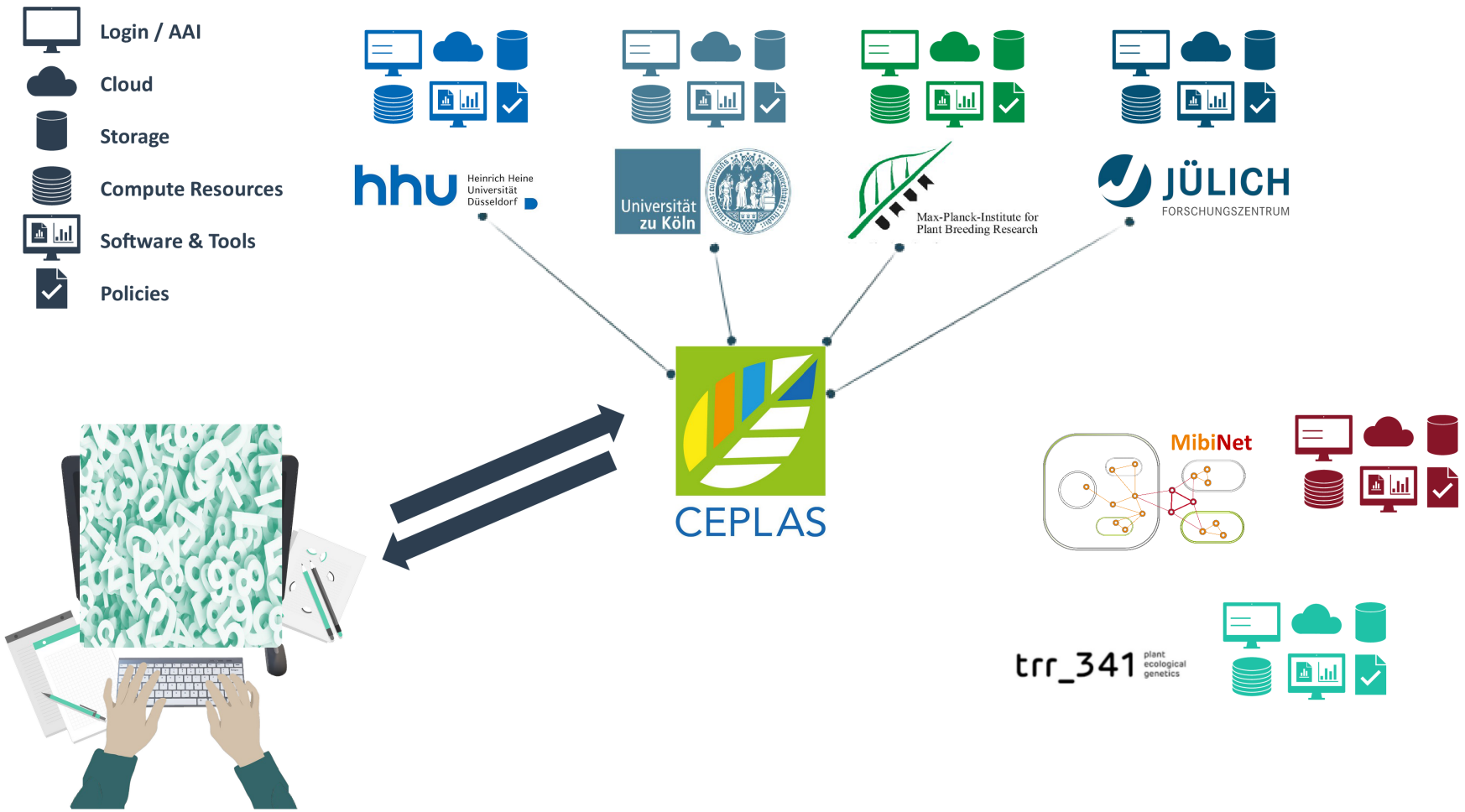
October 27th, 2023

Dominik Brilhaus, CEPLAS Data Science

Goals of CEPLAS RDM

- share data
- share resources
- collaborate

Avoid Scattered Data Silos

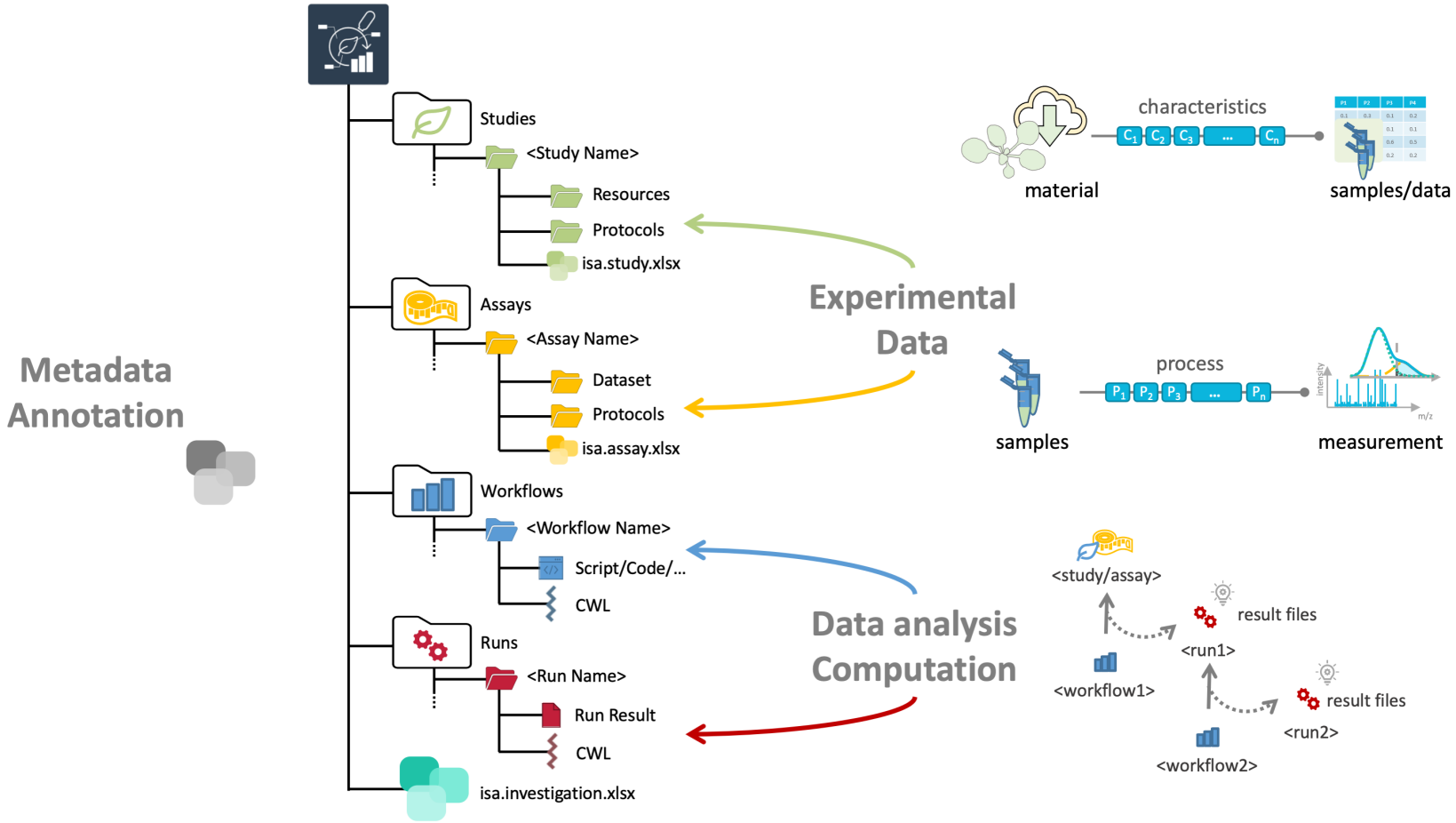


Roadmap

1. Agree on an RDM system
2. Build use-cases and templates
3. Train researchers
4. Start sharing

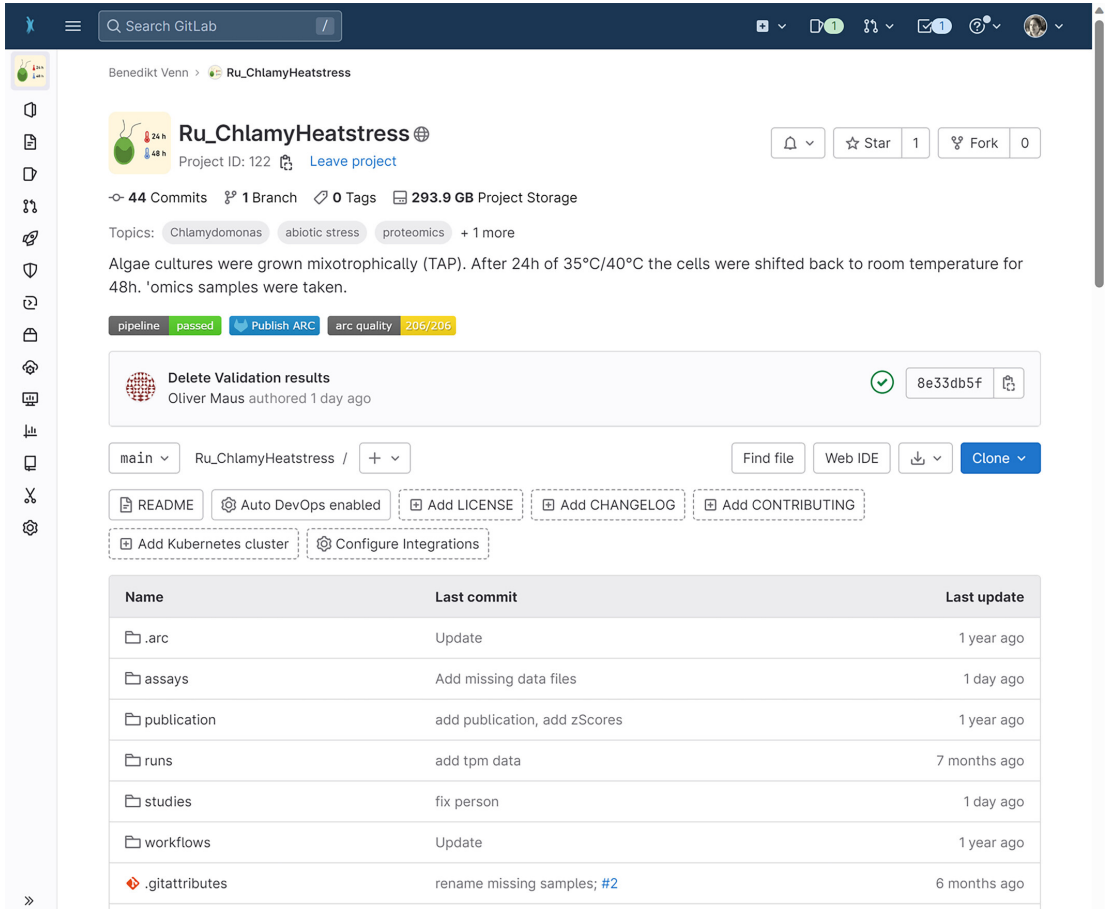
How to store data – ARCs

ARC = Annotated Research Context



Where to share data – DataHUB

The DataHUB is a federated platform for ARCs



Weil, H.L., Schneider, K., et al. (2023), PLANTdataHUB: a collaborative platform for continuous FAIR data sharing in plant research. Plant J. https://doi.org/10.1111/tpj.16474

ARC ecosystem demo @ *Tag der Forschungsdaten*

November 14th, 2023

Nachmittagsprogrammn an der HHU

Wann: ab 13:00 Uhr

Wo: Hörsaal 6C (Gebäude 26.11)

ohne Anmeldung

Vorträge und Talks

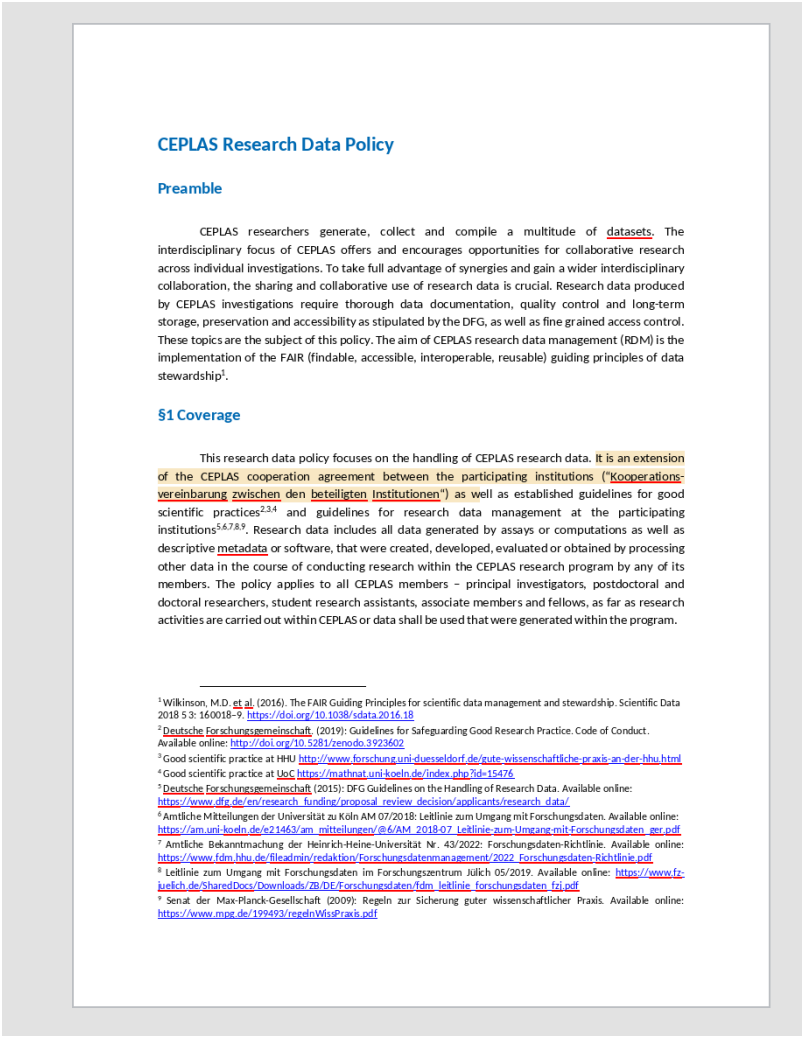
13:00–13:10	Begrüßung		
13:10–13:40	Dr. Dominik Brilhaus	CEPLAS, Data Science and Management	Annotated Research Contexts (ARCs) – A FAIR research data management journey along a mutable data life cycle

https://www.fdm.hhu.de/veranstaltungen/tag-der-forschungsdaten-in-nrw-2023

CEPLAS research data policy

To be circulated in SC soon

- §1 Coverage
- §2 Legal framework conditions
- §3 Handling of research data
- §4 Responsibilities
  - Data delivery
  - Data access
  - Data selection
  - Data use
  - Training
- §5 Validity and governance



CEPLAS Research Data Policy

Preamble

CEPLAS researchers generate, collect and compile a multitude of datasets. The interdisciplinary focus of CEPLAS offers and encourages opportunities for collaborative research across individual investigations. To take full advantage of synergies and gain a wider interdisciplinary collaboration, the sharing and collaborative use of research data is crucial. Research data produced by CEPLAS investigations require thorough data documentation, quality control and long-term storage, preservation and accessibility as stipulated by the DFG, as well as fine grained access control. These topics are the subject of this policy. The aim of CEPLAS research data management (RDM) is the implementation of the FAIR (findable, accessible, interoperable, reusable) guiding principles of data stewardship[1].

§1 Coverage

This research data policy focuses on the handling of CEPLAS research data. It is an extension of the CEPLAS cooperation agreement between the participating institutions ("Kooperations-vereinbarung zwischen den beteiligten Institutionen") as well as established guidelines for good scientific practices[2,3,4] and guidelines for research data management at the participating institutions[5,6,7,8,9]. Research data includes all data generated by assays or computations as well as descriptive metadata or software, that were created, developed, evaluated or obtained by processing other data in the course of conducting research within the CEPLAS research program by any of its members. The policy applies to all CEPLAS members – principal investigators, postdoctoral and doctoral researchers, student research assistants, associate members and fellows, as far as research activities are carried out within CEPLAS or data shall be used that were generated within the program.

[1] Wilkinson, M.D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 2018 5 3: 160018–9. https://doi.org/10.1038/sdata.2016.18

[2] Deutsche Forschungsgemeinschaft. (2019): Guidelines for Safeguarding Good Research Practice. Code of Conduct. Available online: http://doi.org/10.5281/zenodo.3923602

[3] Good scientific practice at HHU http://www.forschung.uni-duesseldorf.de/gute-wissenschaftliche-praxis-an-der-hhu.html

[4] Good scientific practice at UoC https://mathnat.uni-koeln.de/index.php?id=15476

[5] Deutsche Forschungsgemeinschaft (2015): DFG Guidelines on the Handling of Research Data. Available online: https://www.dfg.de/en/research_funding/proposal_review_decision/applicants/research_data/

[6] Amtliche Mitteilungen der Universität zu Köln AM 07/2018: Leitlinie zum Umgang mit Forschungsdaten. Available online: https://am.uni-koeln.de/e21463/am_mitteilungen/@6/AM_2018-07_Leitlinie-zum-Umgang-mit-Forschungsdaten_ger.pdf

[7] Amtliche Bekanntmachung der Heinrich-Heine-Universität Nr. 43/2022: Forschungsdaten-Richtlinie. Available online: https://www.fdm.hhu.de/fileadmin/redaktion/Forschungsdatenmanagement/2022_Forschungsdaten-Richtlinie.pdf

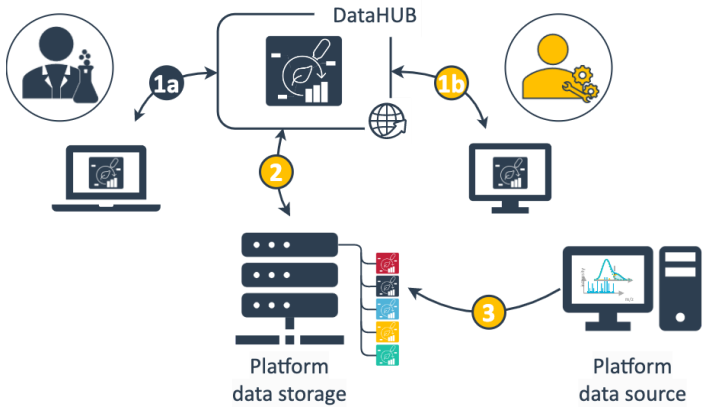
[8] Leitlinie zum Umgang mit Forschungsdaten im Forschungszentrum Jülich 05/2019. Available online: https://www.fz-juelich.de/SharedDocs/Downloads/ZB/DE/Forschungsdaten/fdm_leitlinie_forschungsdaten_fzj.pdf

[9] Senat der Max-Planck-Gesellschaft (2009): Regeln zur Sicherung guter wissenschaftlicher Praxis. Available online: https://www.mpg.de/199493/regelnWissPraxis.pdf

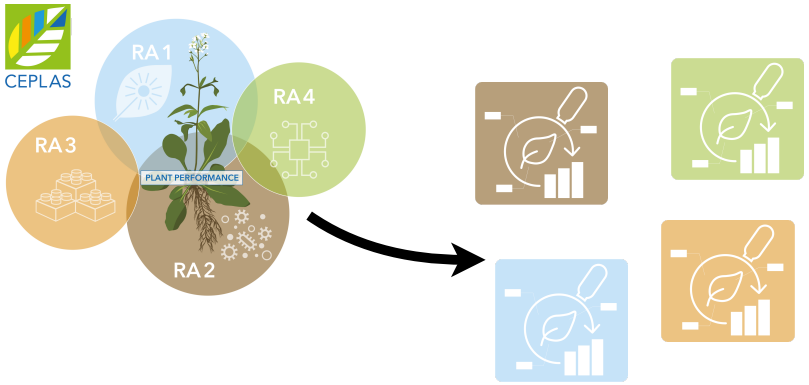
ARC entry points

Workshops and lab hackathons

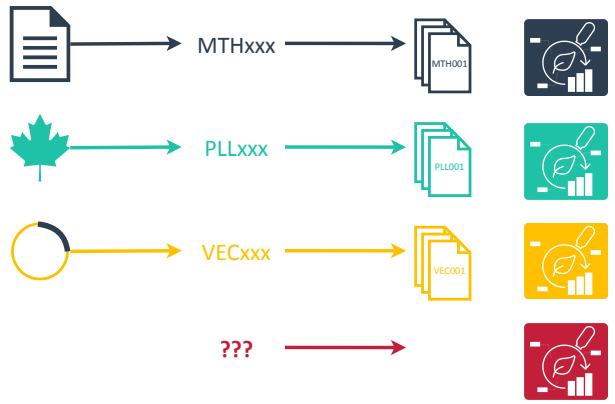
Core facilities



Case studies



Resource ARCs



What I need from you

1. Lighthouse projects
    i. Mainly CEPLAS-funded
    ii. Cross-lab / Cross-RA collaborations
    iii. CEPLAS researchers still present, active and willing to polish ARC
2. Appointments for lab Hackathons
    i. Sit together two days to create ARCs
3. Your backing "My boss said..."

Available dates

for lab hackathons

- Nov 15th & 16th (Wed + Thu)
- Nov 22nd & 23rd (Wed + Thu)
- Nov 30th & Dec 1st (Thu + Fri)
- Nov 7th & Dec 8th (Thu + Fri)

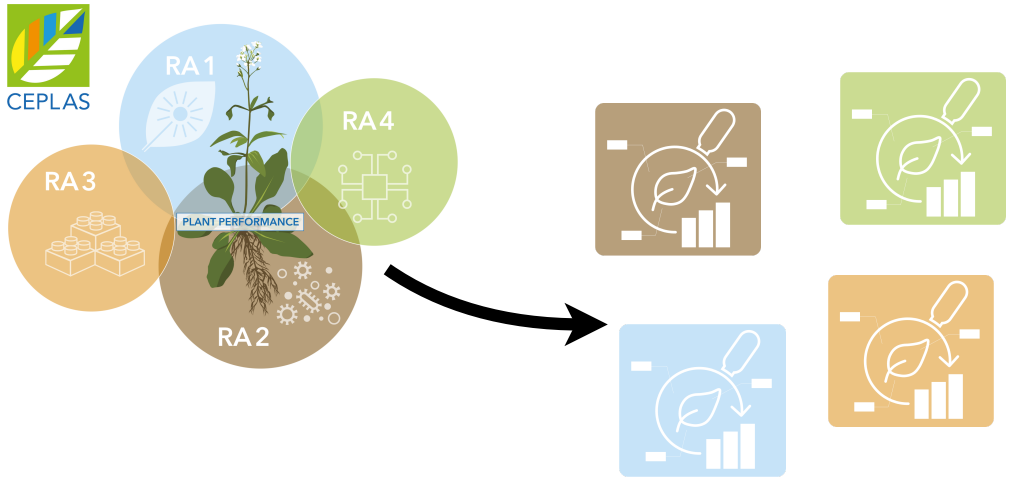
Lab Hackathons

- Good internet connection
- Isolated from lab // office // daily duties
- Access to the data to be ARCified
  (file share, hard disks)
- Agile, communicative, collaborative, positive
- Technical *ad-hoc* support by DataPLANT

💡 ARC Club (@HHU Plant Biochemistry)

Case studies

- Create ARCs from CEPLAS publications
- Publish CEPLAS data as ARCs



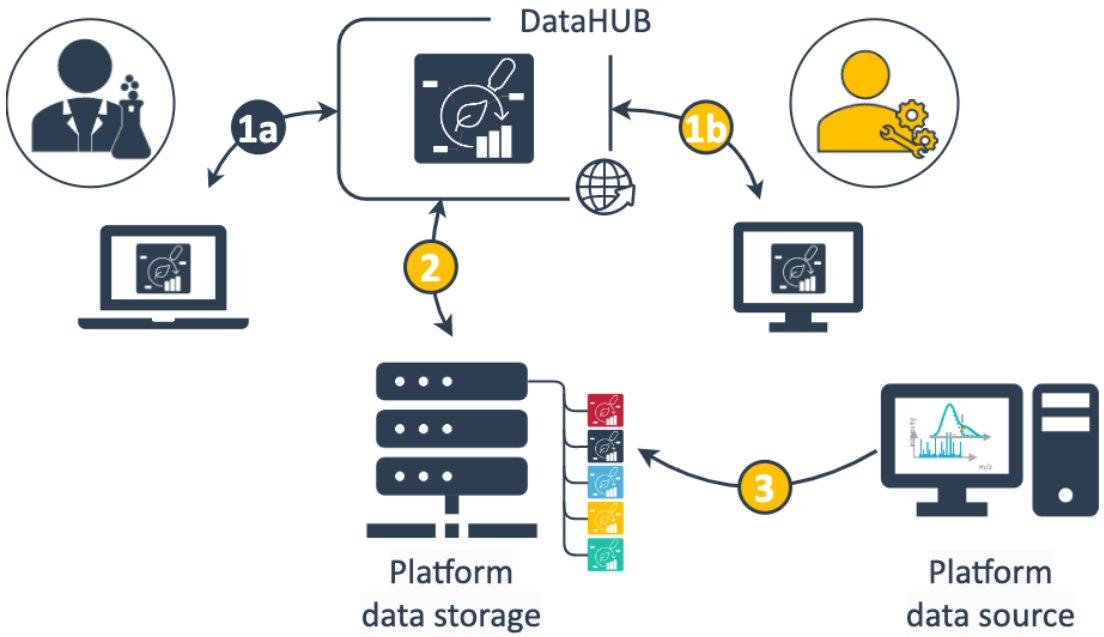
Publication to ARC

ARC Data Publications

Core facilities / enabling platforms

- Help with switch towards ARC ecosystem
- Create templates
- Create use-cases

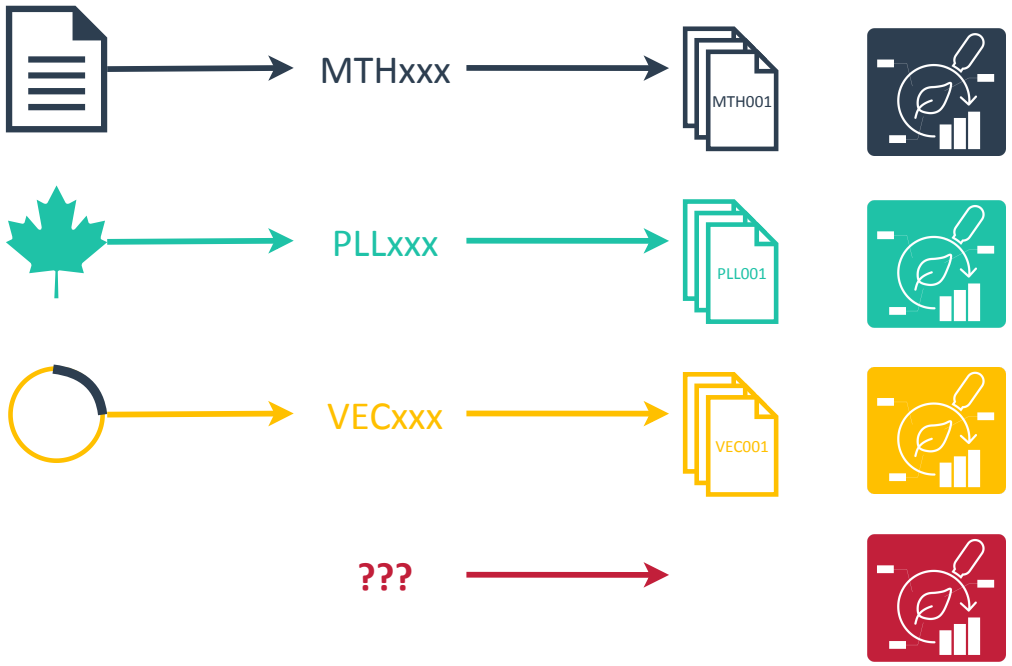
💡 ARCs in Enabling Platforms



Data of common interest: resources

In addition to ***research data***, which CEPLAS resources do we want to share and catalog?

- Identify standards
- Identify / build databases
- Create templates



Network

We're connecting to other consortia using ARCs

- TRR341, https://trr341.uni-koeln.de/
- CRC1535, https://www.sfb1535.hhu.de/
- FAIRagro, https://www.fairagro.net/
    - IPK -> TransCend
- NFDI4BioImage, https://nfdi4bioimage.de/

Resources

DataPLANT (nfdi4plants)

Website: https://nfdi4plants.org/

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

DataHUB: https://git.nfdi4plants.org

GitHub: https://github.com/nfdi4plants