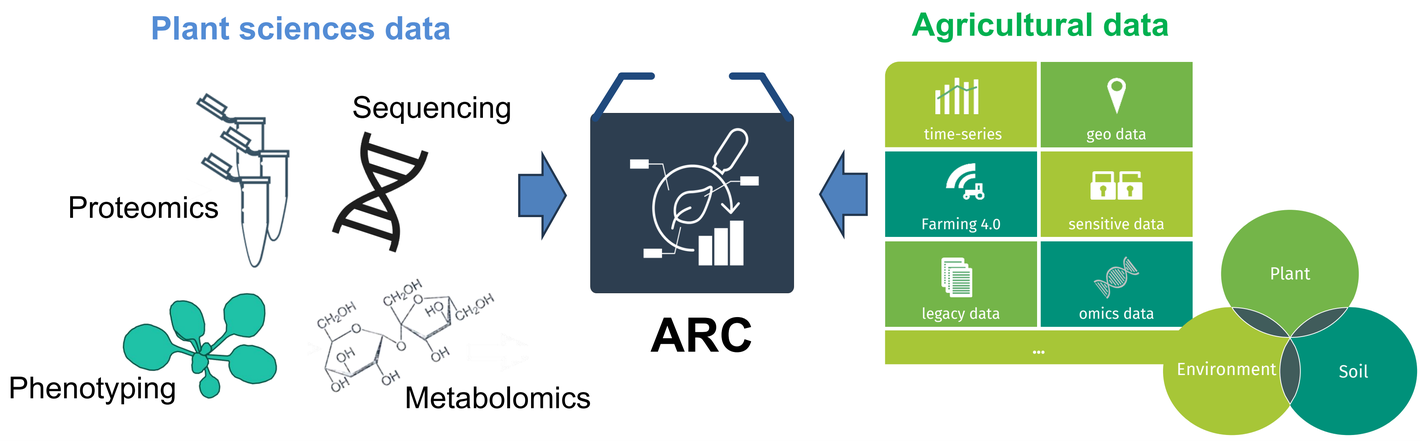
DataPLANT & FAIRagro – Connecting plant and agricultural sciences via common RDM frameworks

Stella Eggels[1], Lars Grygosch[2], Angela Kranz[1]
[1] Institute of Bio- and Geosciences (IBG-4: Bioinformatics), CEPLAS, BIOSC, Forschungszentrum Jülich, Wilhelm Johnen Straße, Jülich, Germany
[2] Institute of Bio- and Geosciences (IBG-2: Plant Sciences), Forschungszentrum Jülich, Wilhelm Johnen Straße, Jülich, Germany

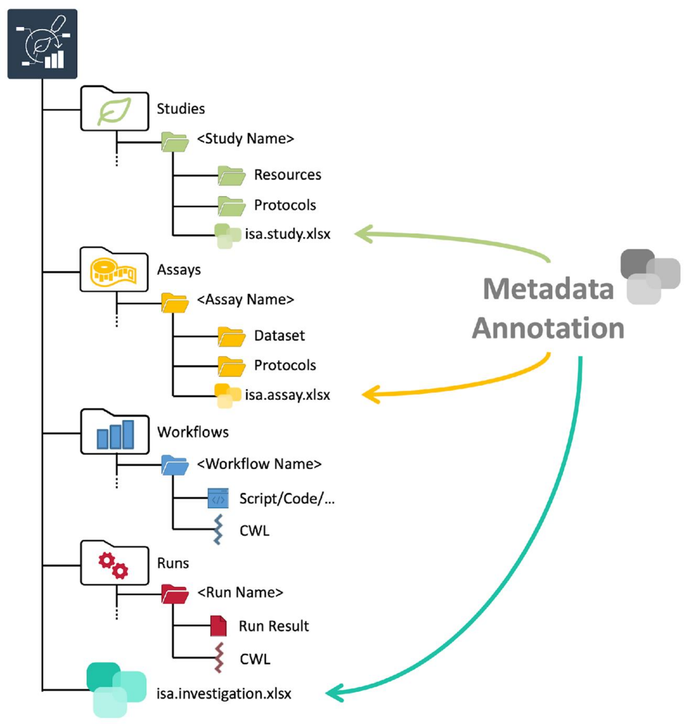
Summary

How can we connect research data from different research domains, scales and types? Using research data management (RDM) frameworks developed by DataPLANT, we demonstrate the harmonization of research data for plant and agricultural science based on actual use cases from FAIRagro. The annotated research context (ARC) serves as the central concept for data and metadata packaging [1].



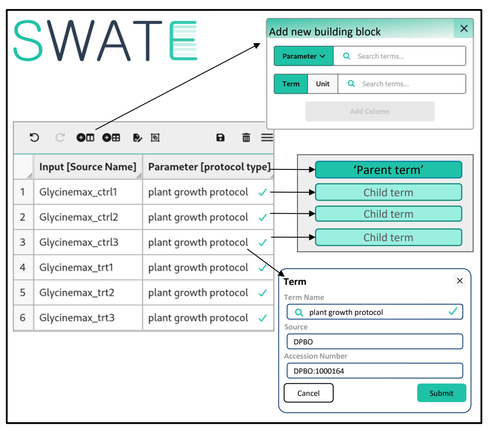
RDM frameworks: FAIR digital objects and metadata templates

ARC

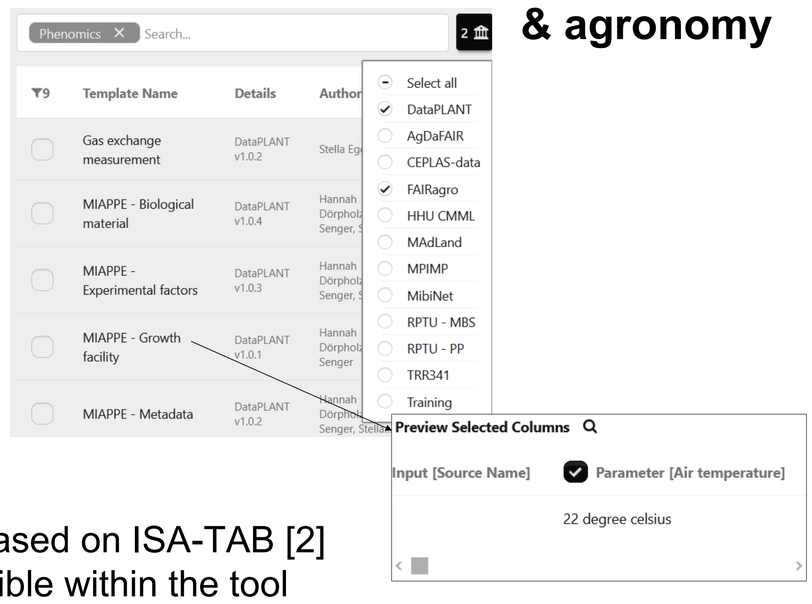


Data and metadata are stored in a FAIR digital object that is based on the ISA (investigation, study, assay) standard [2]

Metadata annotation tool



Templates for plant sciences & agronomy



- Swate [3] allows metadata annotation based on ISA-TAB [2]
- Extensive ontology collection [4] accessible within the tool
- Ready-to-fill templates [5] for minimum information metadata standards (e.g. MIAPPE), data repository checklists and commonly used protocols with a focus on plant sciences and agronomic data
- Community templates can be contributed by any user
- Templates can be flexibly modified or combined

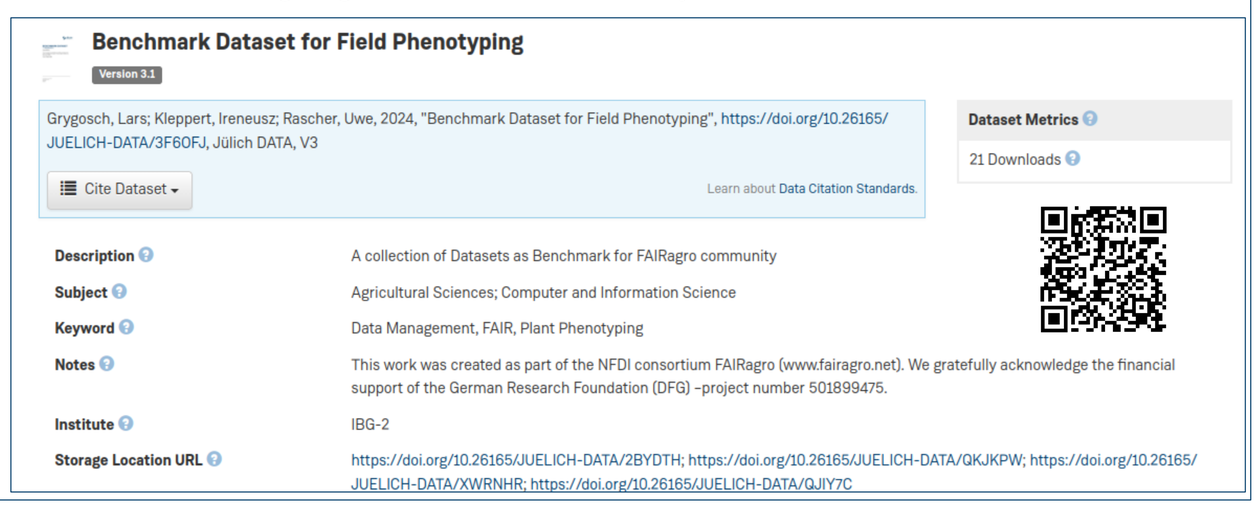
Use case: FAIRagro ARC

Sustainable crop production requires knowledge on the structural and functional status of the crops in the field to plan and execute precise field interventions. To quantify complex plant traits such as stress resilience or disease resistance, the combination of noninvasive phenotyping with sensors, advanced data analytics methods and machine-learning algorithms are nowadays state of the art.

Aims:

- Development of structured metadata formats to enable heterogeneous cross-discipline data to move towards FAIR data.
- Consolidated benchmark dataset across spatial, temporal, and thematic dimensions as a foundational resource for the research community.

⇒ Benchmark dataset and metadata are packaged in the ARC format



Acknowledgements and resources

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation
Project no. 442077441 & 501899475

[1] Weil et al. (2023), The Plant Journal 116 (4), 974-988. DOI: 10.1111/tpj.16474
[2] Sansone et al. (2008), OMICS A Journal of Integrative Biology 12.2, 143-149.DOI: 10.1089/omi.2008.0019
[3] https://github.com/nfdi4plants/Swate
[4] https://github.com/nfdi4plants/nfdi4plants_ontology
[5] https://github.com/nfdi4plants/swate-template-registry

The whole DataPLANT consortium has contributed to the development of the mentioned tools, RDM concepts and illustrations.