



Extend research data repositories to software publication: FAIR and ready?



project.software-metadata.pub
github.com/hermes-hmc
team@software-metadata.pub

S. Druskat, M. Meinel, T. Schlauch
Deutsches Zentrum für Luft- und Raumfahrt



D. Pape, J. Kelling, O. Knodel, G. Juckeland
Helmholtz-Zentrum Dresden-Rossendorf



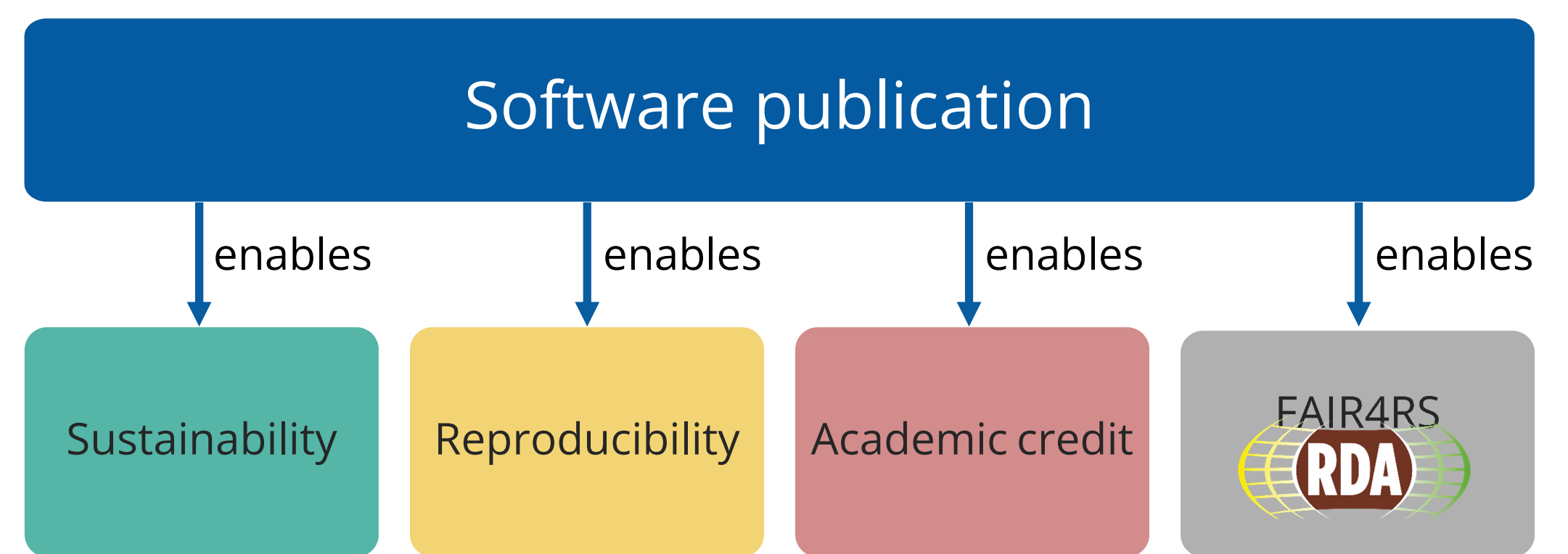
O. Bertuch
Forschungszentrum Jülich



Software Publication

“ To satisfy the principles of [FAIR research software](#), software sustainability and software citation, research software must be formally published. Publication repositories make this possible and provide published software versions with unique and persistent identifiers. However, software publication is still a tedious, mostly manual process and impedes promoting software to first class research citizenship.

S. Druskat, O. Bertuch, G. Juckeland, O. Knodel, and T. Schlauch, *Software publications with rich metadata: State of the art, automated workflows and HERMES concept*, arXiv:2201.09015 [cs], Jan. 2022. doi: 10.48550/arXiv.2201.09015.



Software vs. Data vs. Mixed

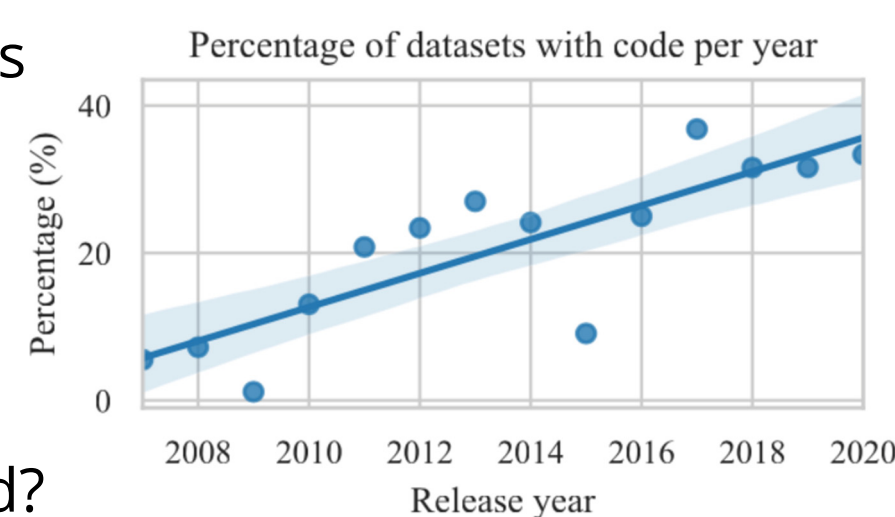
“ Software is data, but it is not just data.

Provides an executable tool, not evidence, with a very different lifecycle, being a creative work and often used to build other software with it as a dependency.

Katz, D. et al, *Software vs. Data in the context of citation*, Dec. 2016. doi: 10.7287/peerj.preprints.2630v1

“ Harvard Dataverse, 40k datasets

- 20% contain code mixed with other types
- 1% code only



- New *ReplicaSet* type for mixed?
- Code deposition has grown over the years

Trisovic, A., *Cluster Analysis of Open Research Data: A Case for Replication Metadata*, Jan. 2023. doi: 10.2218/ijdc.v17i1.833



“ Helmholtz Open Science Policy: data + software KPIs as of 2024

- Largest research association in DE (44k empl., 60% scientific)
- KPI metrics under development
- FAIR for data and software becomes obligatory soon
- Incentive to make data and code openly available

Helmholtz Association, *Helmholtz Open Science Policy v1.0*, Sep. 2022. doi: 10.48440/os.helmholtz.056

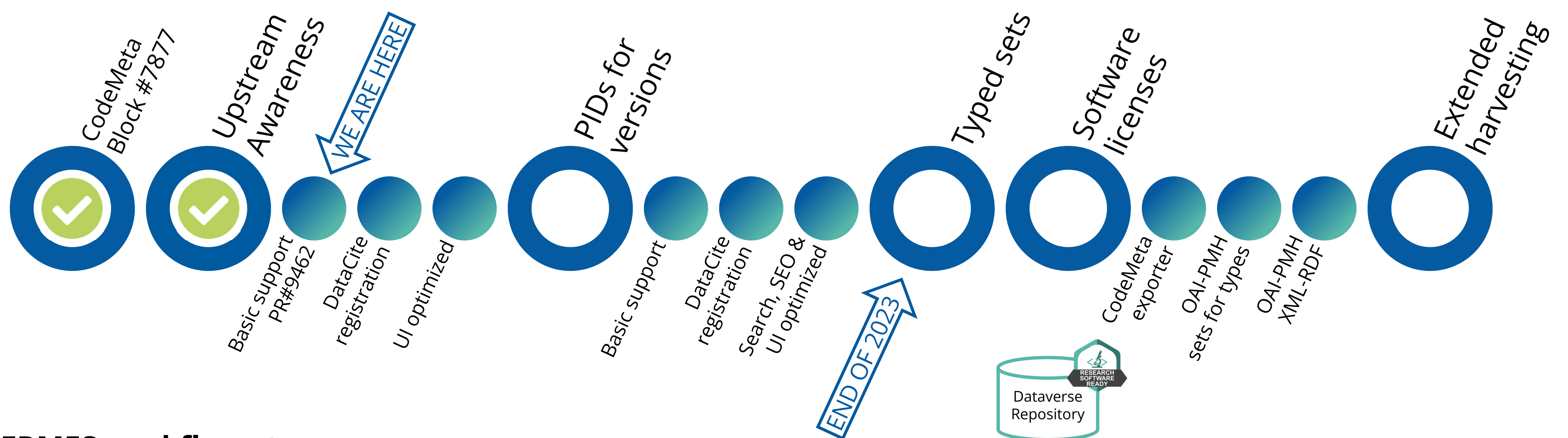
“ Fulfilling the 9 best practices¹⁾ is not sufficient. **Dataverse repositories need to become “Research Software Ready”:**

- Rapid software deposition growth expectable
- Think stereotypes: *data rel., prototype, infrastructure*²⁾
- Make it countable. No support for a mixed type (yet?)
- Vision: code detected at deposit – ask user if split useful



1) Garijo, D., *Nine best practices for research software registries and repositories*, Aug. 2022. doi: 10.7717/peerj-cs.1023
2) ARDC, *A National Agenda for Research Software*, Mar. 2022. doi: 10.5281/zenodo.6378082

Roadmap to make Dataverse research software ready



HERMES workflow steps

