Software Engineering 2024: deRSE 24 – Research Software Engineering

Automating your FAIR software publications with HERMES – a hands-on workshop

Overview

- **Presentation**
  - **Who is the HERMES Team?**
  - **Why publish software?**
  - **How does HERMES help?**
- **Hands-On**
  - **Setting up Project**
  - **Using HERMES in the Command-Line**
  - **Using HERMES in Github actions**
  - **Writing a Plugin for HERMES**
- **Discussion**

Who is HERMES?

Stephan Druskat, Sophie Kernchen, Michael Meinel, Tobias Schlauch

Deutsches Zentrum für Luft- und Raumfahrt

Jeffrey Kelling, Oliver Knodel, David Pape, Guido Juckeland

Helmholtz-Zentrum Dresden-Rossendorf

Oliver Bertuch

Forschungszentrum Jülich

Software publication
The good news

- **Software is an important research output**
- **Publishing research software supports**
  - **Sustainability**
  - **Reproducibility**
  - **academic credit for RSEs**

Software publication
The less good news

To be FAIR:

1. **It's a lot of work**

2. **It's a lot of manual work**

   - **Identifers (DOIs)**
   - **Rich metadata**
   - **Accessibility**
   - **Machine-readable metadata**
   - **Documenting dependencies**
   - **Licenses**
   - **Provenance**
   - **…**
   - **Versions!**

Montag 13:30 - Raum Schniecke - Workshop: Let's talk FAIR for research software

Hosts: Martin Stoffers, Tobias Schlauch, Alexander Struck, René Caspart

The session will cover practices in and handling of research software development and how they relate to the FAIR principles [0]. We will introduce current challenges we have faced and/or identified in our daily work and research.

Using the format World Cafe, we like to discuss such challenges with you to gather your perspectives and ideas. Existing material, e.g. [1], software publication is discussed and (ideally) improved. The Findability aspect is further investigated when we discuss how/where it may be touched upon.

The situation in HPC environments may differ from common research labs and we will look at user-oriented and/or ... sible and interoperable. The audience is free to focus on any aspect or cover all.

[0] https://www.rd-alliance.org/groups/fair-research-software-fair-rs-wg
[1] https://gitlab.com/hifis/hifis-workshops

Data Science 3 (2020) 37–59
DOI 10.3233/DS-190026
IOS Press

Towards FAIR principles for research software

Anna-Lena Lamprecht[a,*], Leyla Garcia[b], Mateusz Kuzak[c,d], Carlos Martinez[e], Ricardo Arcila[f], Eva Martin Del Pico[g], Victoria Dominguez Del Angel[h], Stephanie van de Sandt[i], Jon Ison[j], Paula Andrea Martinez[k], Peter McQuilton[l], Alfonso Valencia[m,n], Jennifer Harrow[o], Fotis Psomopoulos[p], Josep Ll. Gelpi[q,r], Neil Chue Hong[s], Carole Goble[t] and Salvador Capella-Gutierrez[u,**]

[a] Utrecht University, The Netherlands
E-mail: a.l.lamprecht@uu.nl; ORCID: https://orcid.org/0000-0003-1953-5606
[b] ZBMED Information Centre for Life Sciences, Germany
E-mail: ljgarcia@zbmed.de; ORCID: https://orcid.org/0000-0003-3986-0510
[c] Netherlands eScience Center, The Netherlands
[d] Dutch Techcentre for Life Sciences, The Netherlands
E-mail: m.kuzak@esciencecenter.nl; ORCID: https://orcid.org/0000-0003-0087-6021
[e] Netherlands eScience Center, The Netherlands
E-mail: c.martinez@esciencecenter.nl; ORCID: https://orcid.org/0000-0001-5565-7577
[f] EMBL-EBI, UK
E-mail: arcila@ebi.ac.uk; ORCID: https://orcid.org/0000-0002-8253-7375
[g] Barcelona Supercomputing Center (BSC), Spain
E-mail: eva.martin@bsc.es; ORCID: https://orcid.org/0000-0001-8324-2897
[h] 1. Institut Français de Bioinformatique (IFB), France
E-mail: victoria.dominguez@france-bioinformatique.fr; ORCID: https://orcid.org/0000-0002-5514-6651
[i] CERN, Switzerland
E-mail: stephanie.van.de.sandt@cern.ch; ORCID: https://orcid.org/0000-0002-9576-1974
[j] National Life Science Supercomputing Center, Technical University of Denmark, Denmark
E-mail: jison@cbs.dtu.dk; ORCID: https://orcid.org/0000-0001-6666-1520
[k] National Imaging Facility, Australia
E-mail: p.martinez@uq.edu.au; ORCID: https://orcid.org/0000-0002-8990-1985
[l] Oxford e-Research Centre, UK
E-mail: peter.mcquilton@oerc.ox.ac.uk; ORCID: https://orcid.org/0000-0003-2687-1982
[m] Barcelona Supercomputing Center (BSC), Spain
[n] Institució Catalana de Recerca i Estudis Avançats (ICREA), Spain
E-mail: alfonso.valencia@bsc.es; ORCID: https://orcid.org/0000-0002-8937-6789
[o] ELIXIR Hub, UK
E-mail: harrow@ebi.ac.uk; ORCID: https://orcid.org/0000-0003-0338-3070
[p] Institute of Applied Biosciences, CERTH, Greece

2451-8484 © 2020 – IOS Press and the authors.

DFG

Leitlinien zur Sicherung guter wissenschaftlicher Praxis

Kodex

New upload

Drag and drop files here

Choose Files

What problem does HERMES solve?

Advancing the state of the art in software publication

- **Enable automated publication of research software according to the FAIR principles** (even for non-open source software)

- Automate collation of metadata from different sources

- Automated synchronization of software metadata between publication and repository

- Differentiated treatment of software, documentation and data components
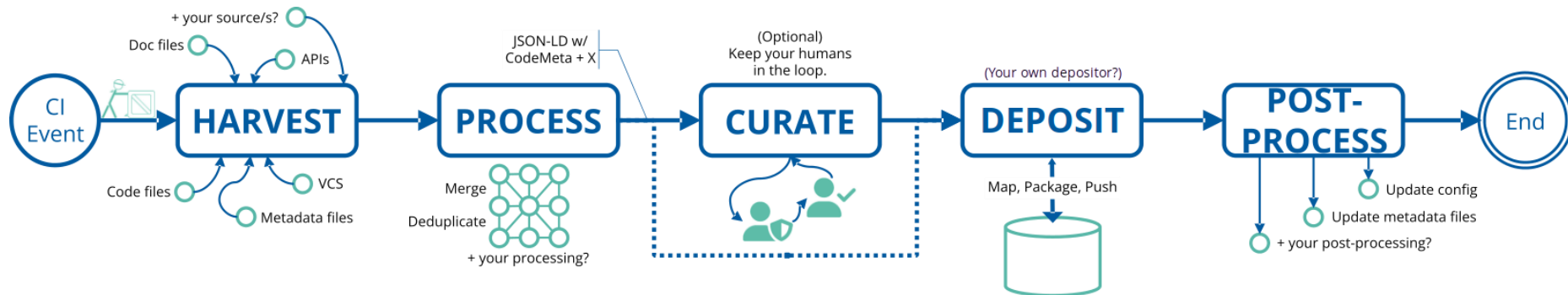
Software publication with HERMES

Overview (simplified)
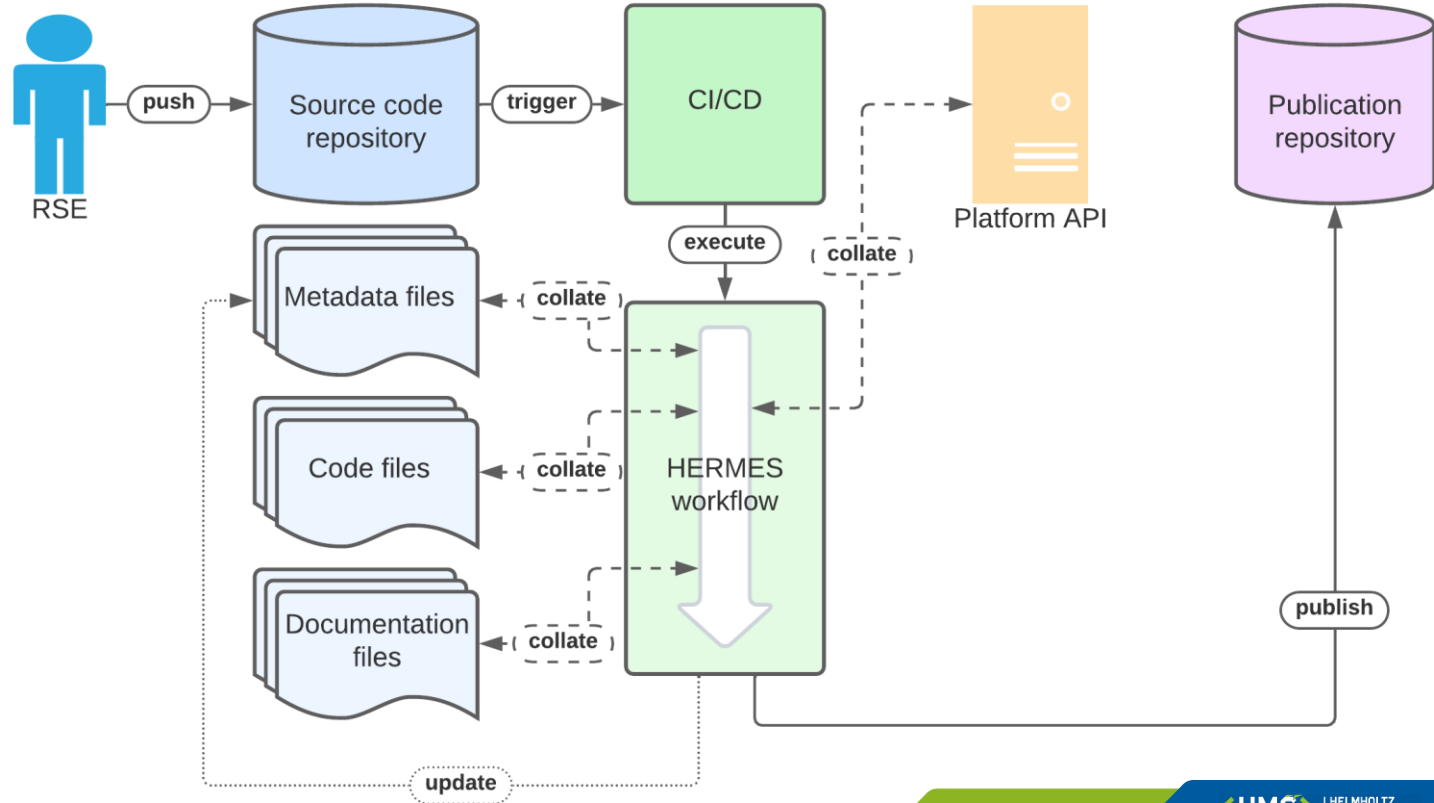
HERMES
Using a push-based CI workflow to publish software

- Push vs. Pull workflow
- CI triggered (configurable events)
- Pluggable (Python modules)
- Can keep humans in the loop

https://software-metadata.pub

https://doc.software-metadata.pub

Software publication with HERMES

Metadata: types, formats, sources for HERMES

- **Metadata**
  - Differences in generation, scope, mode, aspects
  - Generic software metadata vs. software-specific metadata
- **Metadata formats**
  - Metadata files, snippets, third-party systems, API responses
  - Structured vs. unstructured
- **Sources**
  1. Collectable structured metadata
  2. (Metadata from minable structured data)
  3. (Metadata from minable unstructured data)

Citation File Format
Structured metadata as citation hint

What is a `CITATION.cff` file?

`CITATION.cff` files are plain text files with human- and machine-readable citation information for software (and datasets). Code developers can include them in their repositories to let others know how to correctly cite their software.

This is an example of a simple `CITATION.cff` file:

```
cff-version: 1.2.0
message: "If you use this software, please cite it as below."
authors:
  - family-names: Druskat
    given-names: Stephan
    orcid: https://orcid.org/1234-5678-9101-1121
title: "My Research Software"
version: 2.0.4
doi: 10.5281/zenodo.1234
date-released: 2021-08-11
```

https://go.fzj.de/cffinit

Javascript webapp to jumpstart your CFF file!

What will HERMES give you?

Project outputs

- **Software**
  - Software for software publication workflow automation
- **CI templates**
  - GitLab CI, GitHub Action, Jenkins, Travis CI
- **Training materials**
  - Adaption of HIFIS training materials to include workflow usage
- **Project website**
  - One-stop shop for information and documentation
- **Policy proposals**
  - Proposals for updates to policies/ guidelines at Helmholtz and Allianzinitiative

Hands On?

- **Who has used HERMES already?**

- **Two Groups: Setting HERMES Workflow up vs. Writing HERMES Plugin**

- [Set up automatic software publishing — HERMES Documentation (software-metadata.pub)](#)