

# AUTOMATING RESEARCH SOFTWARE PUBLICATION WITH RICH METADATA

What can we learn for research data publication?

Stephan Druskat

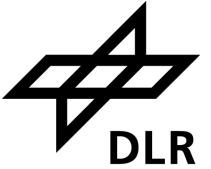
Institute of Software Technology, German Aerospace Center (DLR), Berlin

License [CC-BY-4.0 International](#)

DOI [10.5281/zenodo.14265448](https://doi.org/10.5281/zenodo.14265448)

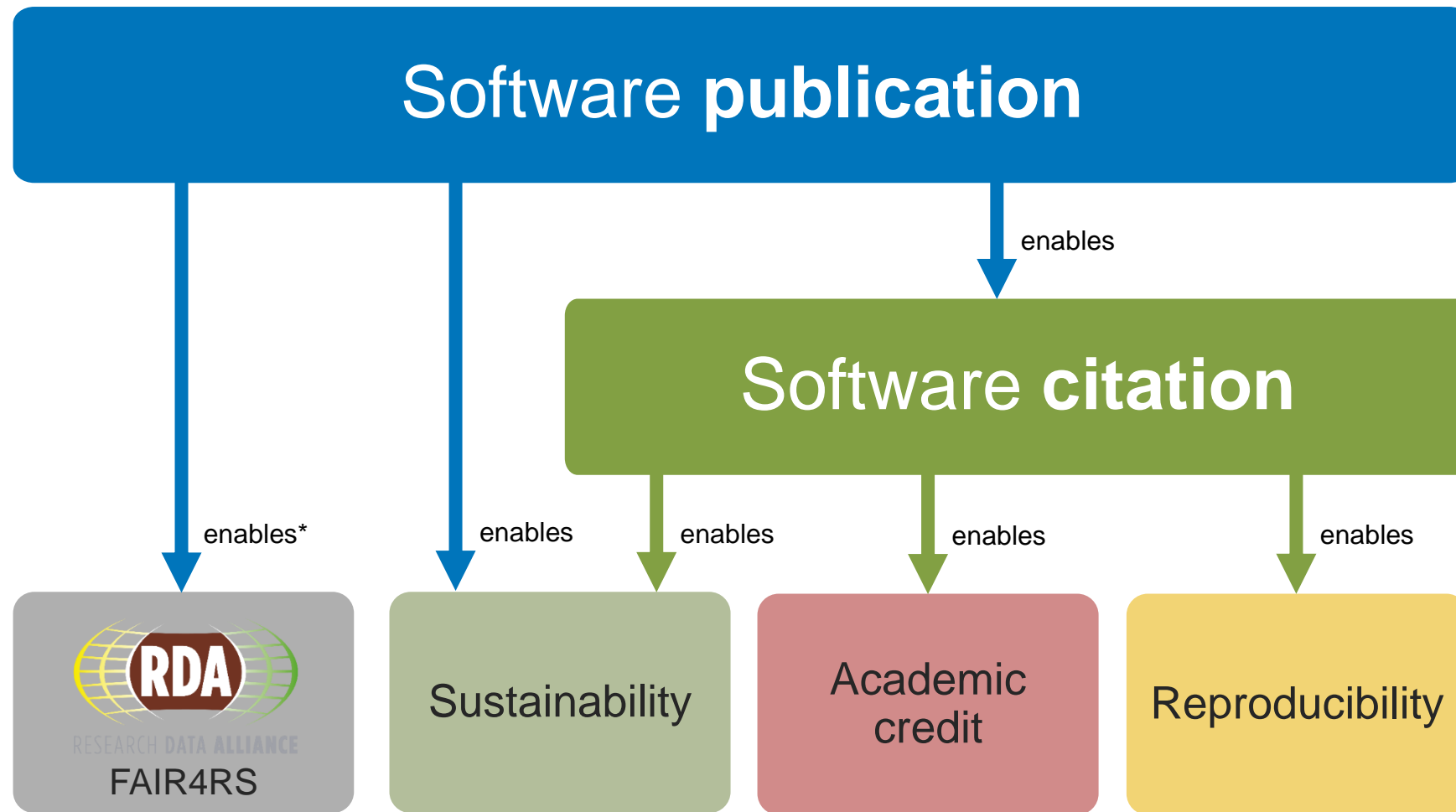


# Overview



- Software publication & FAIR research software
- HERMES concept
- HERMES implementation
- Current & future work
- Automating research *data* publication with rich metadata?
- [optional] Example

# Software *publication* is important, because research software is important!



Adapted from Druskat et al. (2023)

# FAIR for Research Software (FAIR4RS)



- **F:** Software, and its associated metadata, is **easy for both humans and machines to find**.
- **A:** Software, and its metadata, **is retrievable via standardized protocols**.
- **I:** Software **interoperates with other software** by exchanging data and/or metadata, and/or through interaction via application programming interfaces (APIs), described through standards.
- **R:** Software is both usable (**can be executed**) and reusable (**can be understood**, **modified, built upon, or incorporated** into other software).

Chue Hong et al. (2022)

**Publication of software with metadata**  
(Research) Software Engineering

# What is software publication?



Publication of metadata and artifacts\* for software versions in publication repositories

Persistent identifier for each version

**Not** software publication:

- Software available on a source code platform (*GitHub, GitLab, etc.*)
- A paper *about* the software

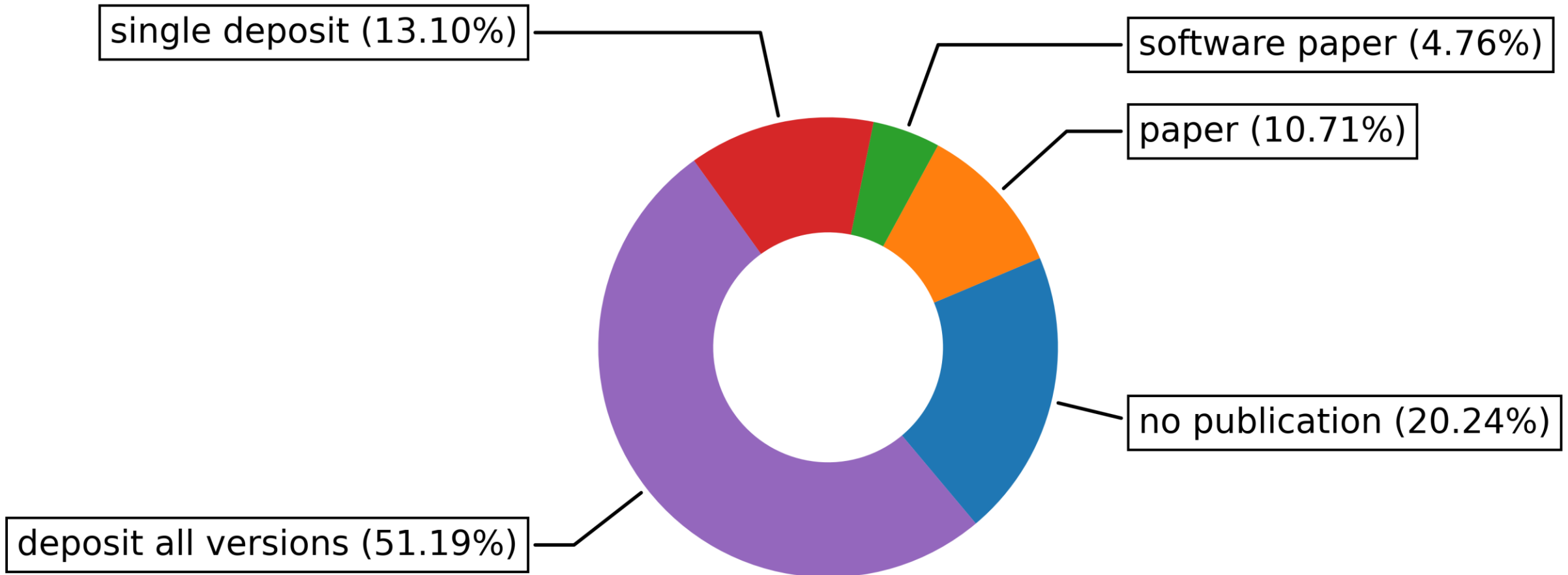
Interim solution:

- software journals ([JOSS](#), [JORS](#), etc.)

**Challenge: quality assurance**

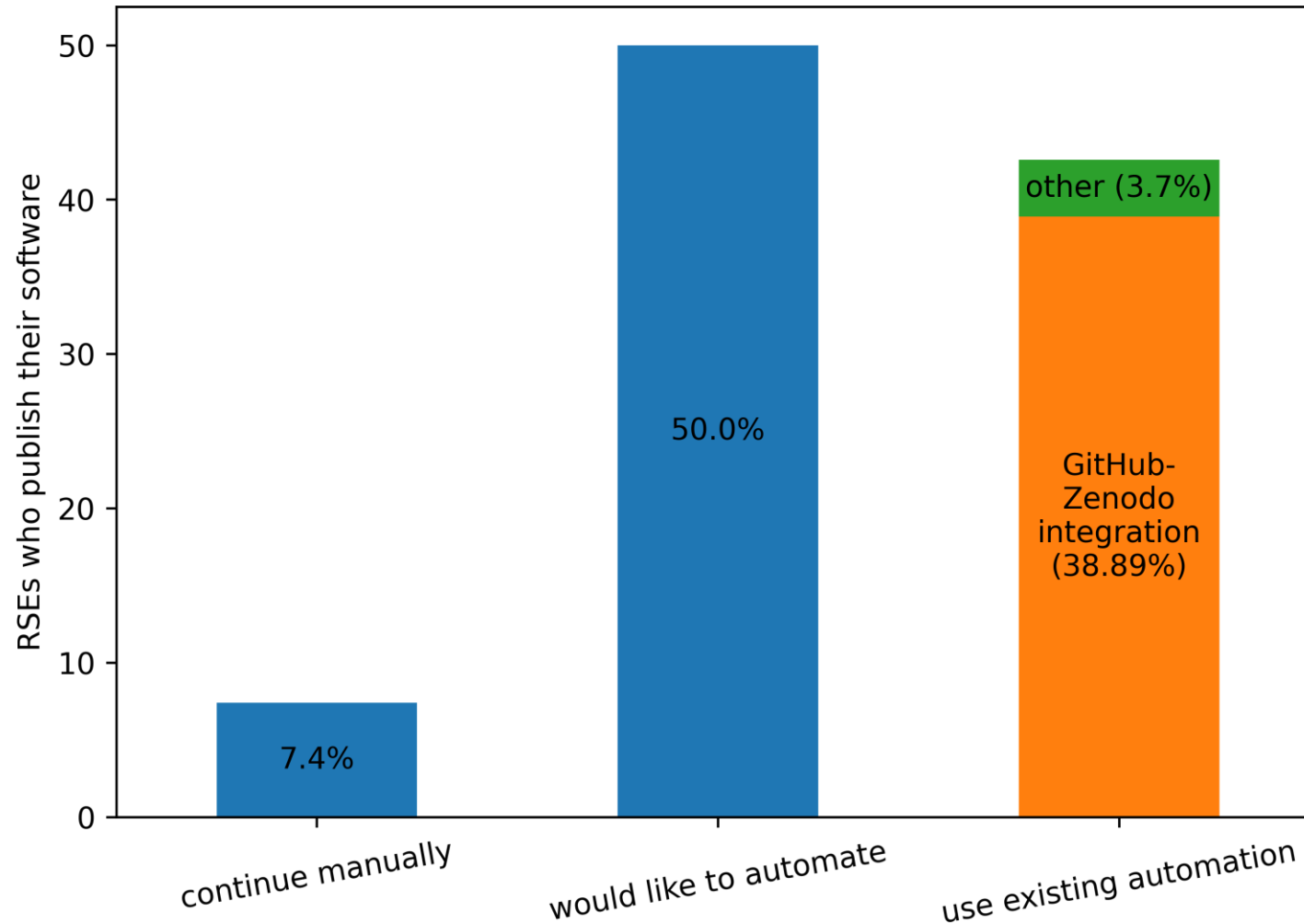
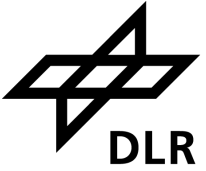
The screenshot shows the Zenodo page for the software 'Hexatomic'. The page includes a search bar, navigation links for 'Upload' and 'Communities', and a 'Log in' / 'Sign up' button. The main content area displays the software title 'Hexatomic', the authors 'Druskat, Stephan; Krause, Thomas; Lachenmaier, Clara; Bunzeck, Bastian', and a description of the software as an extensible, OS-independent platform for deep multi-layer linguistic annotation. It also shows a list of files for the 'hexatomic-v1.0.1.zip' artifact, including a directory 'hexatomic-hexatomic-4fa5704' with sub-files like 'all-contributorsrc', 'checkstyle', 'github', 'ISSUE\_TEMPLATE', 'workflows', 'githignore', and 'mvn'. The 'Files' section shows the main artifact 'hexatomic/hexatomic-v1.0.1.zip' with a size of 8.7 MB. The 'Citations' section is currently empty. On the right side, there are statistics for views (44) and downloads (0), and a list of where the software is available, including GitHub and OpenAIRE. The 'Publication date' is August 30, 2022, and the DOI is 10.5281/zenodo.7034163. The 'License' is Apache License 2.0. The 'Versions' section lists several versions, including the current one (1.0.1) and a snapshot (0.13.0-SNAPSHOT).

# How is software currently published/publicized?



RSE software publication practices, N=84 (Druskat & Theis (submitted))

# How do RSEs publish their software?



RSEs' software publication process, N=54 (Druskat & Theis, submitted)

# Automated software publication: Zenodo-GitHub integration

DOIs for versions & projects, metadata, persistence

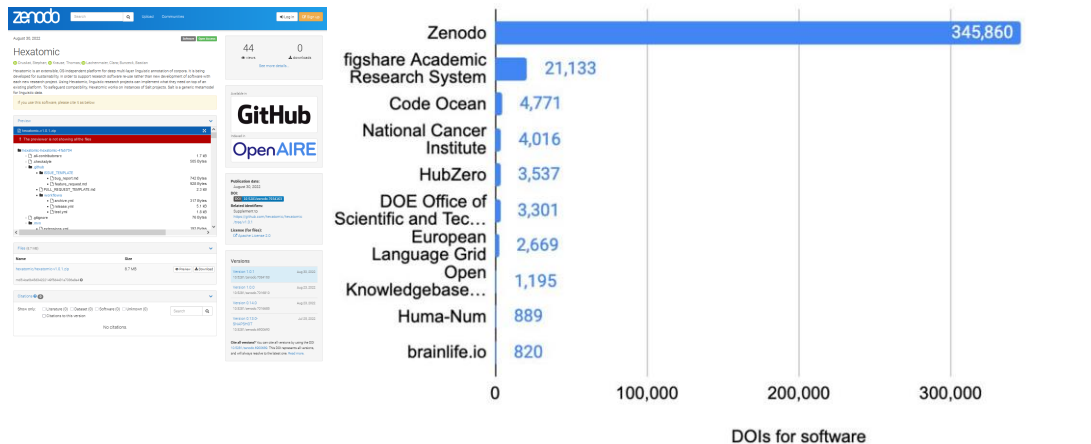
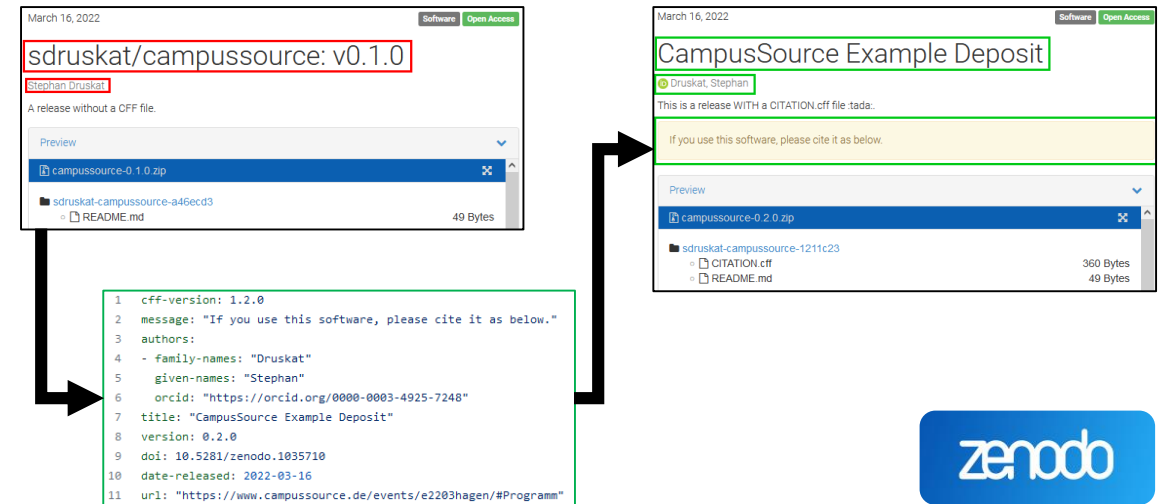


Figure from Mejias (2023).

GitHub-Zenodo integration,  
improved metadata via `CITATION.cff`



Left: GitHub release for `sdruskat/campussource: v0.1.0` by `Stephan Druskat`. Metadata is minimal, showing only the release name and a zip file.

```
1 cff-version: 1.2.0
2 message: "If you use this software, please cite it as below."
3 authors:
4 - family-names: "Druskat"
5   given-names: "Stephan"
6   orcid: "https://orcid.org/0000-0003-4925-7248"
7 title: "CampusSource Example Deposit"
8 version: 0.2.0
9 doi: 10.5281/zenodo.1035710
10 date-released: 2022-03-16
11 url: "https://www.campussource.de/events/e2203hagen/#Programm"
```

Right: GitHub release for `CampusSource Example Deposit` by `Druskat, Stephan`. Metadata is significantly improved, including a clear citation instruction and a `CITATION.cff` file.



- Most often used by far: [Zenodo](#) (developed & run at CERN), GitHub integration
- Persisted records, persistent identifiers ([DataCite](#) DOIs w/ metadata), landing pages
- Integration only for specific platform combination, insufficient metadata in basic usage

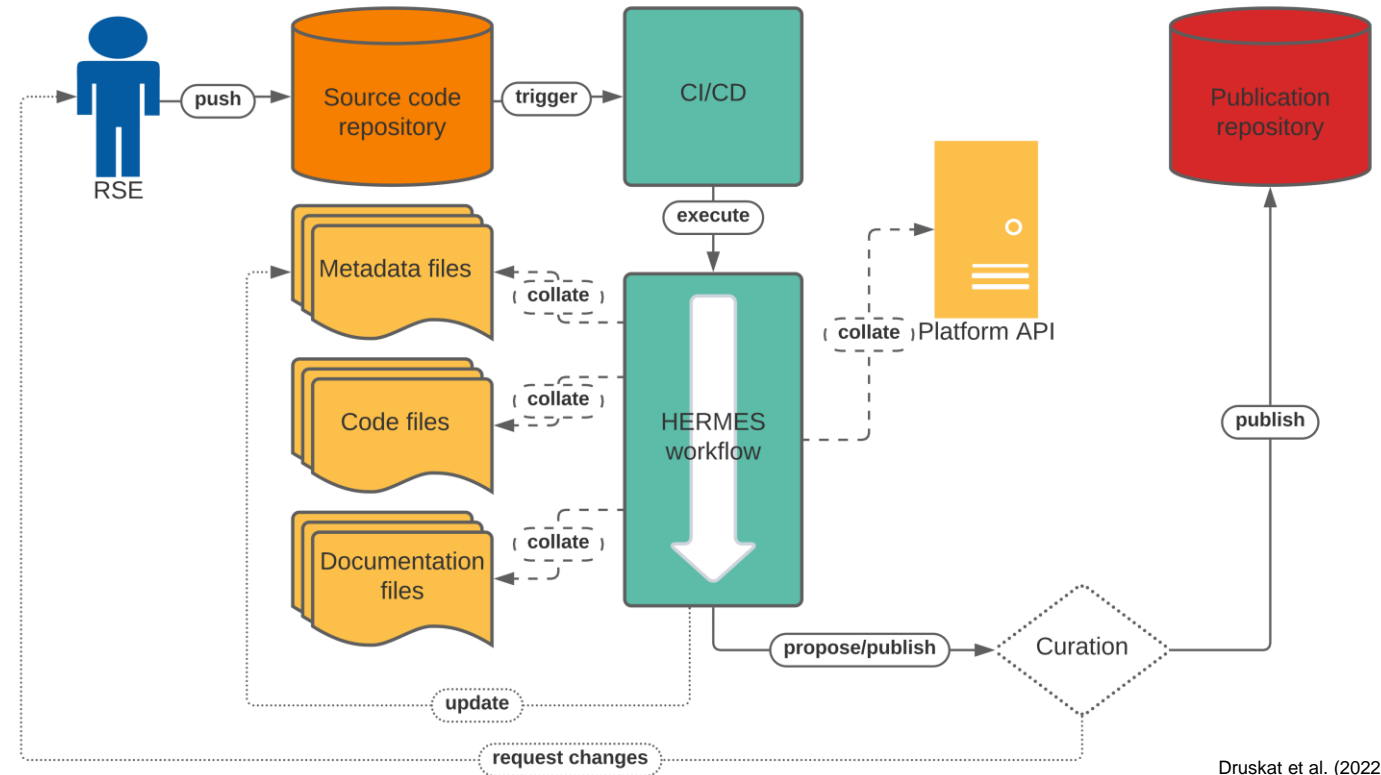


# HERMES: Helmholtz Rich Metadata Software Publication

(Helmholtz Metadata Collaboration project)



- Automated software publication for all platform combinations
- Use existing metadata to enrich records/improve FAIRness
- Enable:
  - closed source publication,
  - curation & sign-off processes,
  - updating metadata records



Druskat et al. (2022)

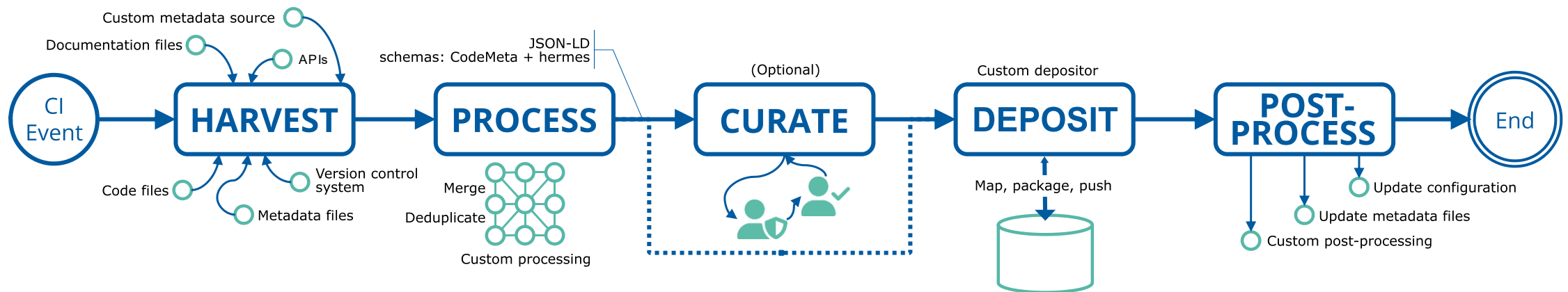
▪ [software-metadata.pub](https://software-metadata.pub)



# HERMES: Implementation



- Continuous integration workflow: on <event> run `hermes` as configured
- Tutorials for GitHub/GitLab: [docs.software-metadata.pub](https://docs.software-metadata.pub)



- `hermes` Python package (Meinel et al. 2024) + CI templates (GitHub, GitLab)
- Plugins via Python Extension Point mechanism for each step
- Details: Kernchen et al. (2024)

# HERMES: Use and extensibility



- Dogfooding:
  - <https://github.com/softwarepub/hermes/releases/tag/v0.8.1>
  - <https://zenodo.org/records/13311079>
- User testing at live coding workshop (deRSE24)
- `hermes-plugin-python` (Fritzsche & Meinel 2024) created by high school intern within a few days
- Details: Kernchen et al. (2024)

Published August 13, 2024 | Version 0.8.1

Software Open

hermes

Meinel, Michael<sup>1</sup>; Druskat, Stephan<sup>1</sup>; Kelling, Jeffrey<sup>2</sup>; Bertuch, Oliver<sup>3</sup>; Knodel, Oliver<sup>2</sup>; Pape, David<sup>2</sup>; Kernchen, Sophie<sup>1</sup>

Show affiliations

hermes

Files

hermes.zip

Name	Size	Download all
hermes.zip md5:40a1e0da0e6e7761e1b53091b708914	42.2 KB	Preview Download
README.md md5:8f06e42aa23604ebcd910f05dccc0f6	3.2 KB	Preview Download

59 VIEWS 6 DOWNLOADS

Show more details

Versions

Version	Date
Version 0.8.1 10.5281/zenodo.13311079	Aug 13, 2024
Version 0.8.1b1 10.5281/zenodo.13221384	Aug 5, 2024

View all 2 versions

Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.13221383. This DOI represents all versions, and will always resolve to the latest one. Read more.

External resources

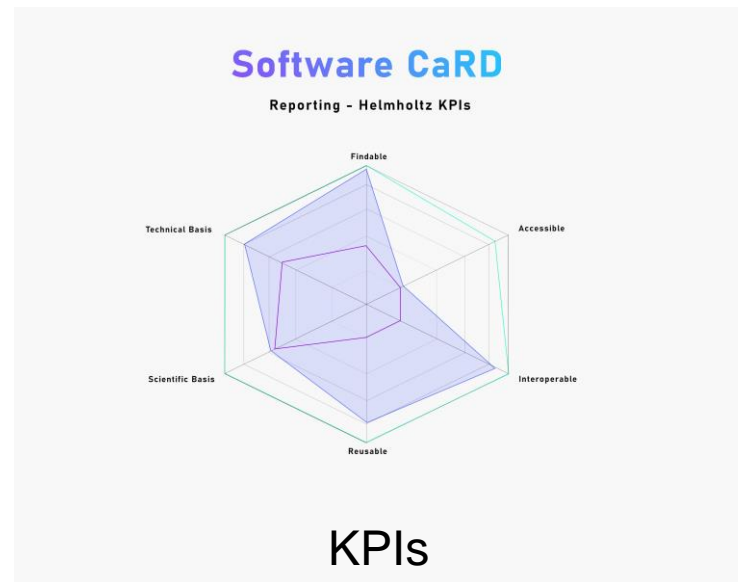
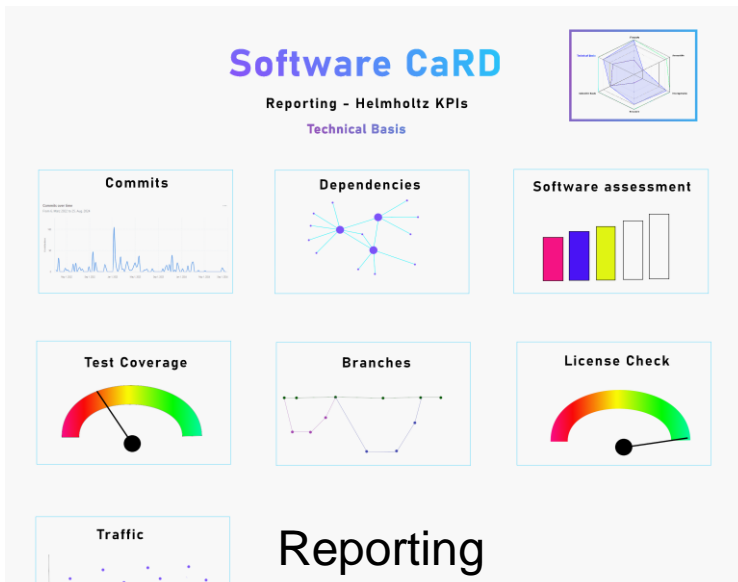
Indexed in

# Current work: hermes 1.0.0, Software CaRD



- hermes 1.0.0
- Software Curation and Reporting Dashboard (Software CaRD)
  - Input: Consistent knowledge graph produced by HERMES
  - Compliance checks against configurable policies (KPIs, curation)

```
{
  "@context": [ ... ],
  "@type": "SoftwareSourceCode",
  "name": "hermes",
  "version": "0.8.1",
  "license": "https://spdx.org/licenses/Apache-2.0",
  "author": [
    {
      "@id": "https://orcid.org/0000-0001-6372-3853",
      "@type": "Person",
      "affiliation": { ... },
      "familyName": "Meinel",
      "givenName": "Michael",
      "email": "michael.meinel@dlr.de"
    },
    ...
  ],
  "hasPart": [
    {
      "@type": "CreativeWork",
      "name": "README",
      "encoding": {
        "@type": "TextObject",
        "encodingFormat": "text/markdown",
        "url": "file:///README.md"
      }
    },
    ...
  ]
}
```

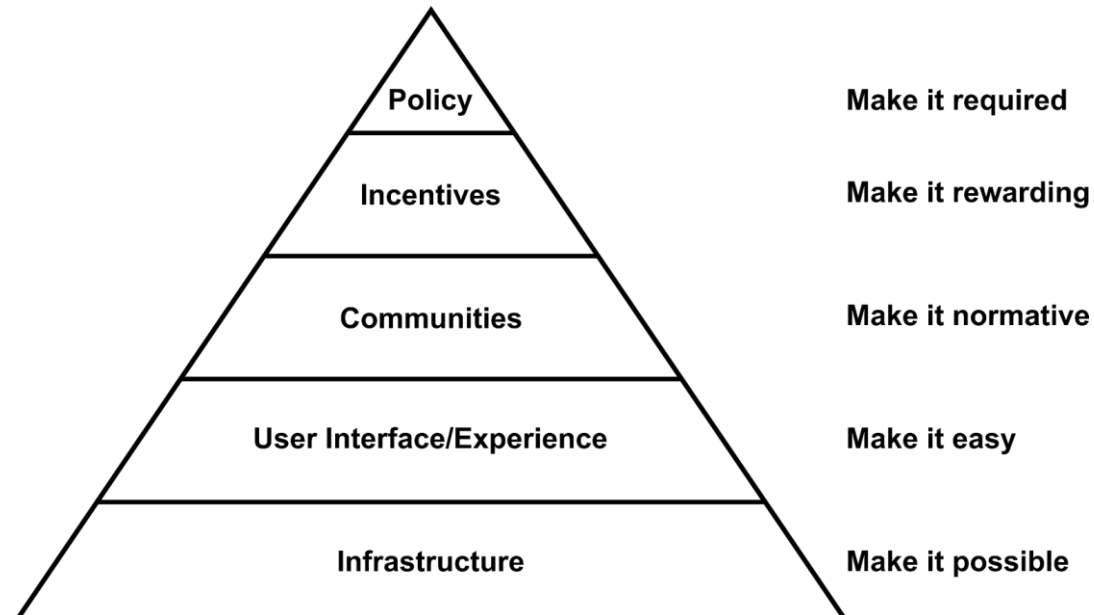


**Software CaRD**  
Curation Dashboard

HERMES > Authors > Michael Meinel				
ORCID [id]	0000-0001-6372-3853	Harvester Location	cif	CITATION cif v01 2024-09-01T17:34:22
Given name [givenName]	Michael	Harvester Location	cif	CITATION cif v01 2024-09-01T17:34:22
Family name [familyName]	Michael	Harvester Location	cif	CITATION cif v01 2024-09-01T17:34:22
E-Mail address [email]	michael.meinel@dlr.de	Harvester Location	cif	CITATION cif v01 2024-09-01T17:34:22
	led02@me.com	Harvester Location	git	CITATION cif v01 2024-09-01T17:34:22
	michael_meinel@web.de	Harvester Location	orcid	CITATION cif v01 2024-09-01T17:34:22
Affiliation [affiliation]	German Aerospace Center (DLR e.V.)			

Curation

- Rollout of HERMES workflows at two German universities
  - Heterogeneous stakeholders
  - Empirical requirements elicitation & usability optimization
  - New deployment targets: DSpace, MyCoRe
  - Development of targeted capacity building & outreach material

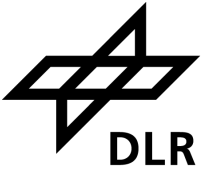


# Automating research *data* publication with rich metadata?



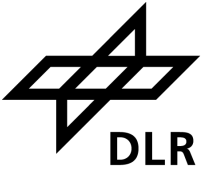
- Data ≠ software
  - Executability
  - Evidence vs. tool
  - Creative work vs. facts/observations
  - Change cycles & dependencies
- Data ≠ software?
- Do we have the same problems in data as in software?
  - Dynamicity
  - Workflows
  - Infrastructure

# Practicalities of using HERMES for data



- Clear focus on research *software*
- HERMES data model: CodeMeta++ (v3.0, Jones et al., 2023)
- Linked Open Data
- Interaction through `HermesContext` API with internal data (nested `dicts`)

# Conclusion



- HERMES automates the publication of software with rich metadata
- HERMES supports FAIR software as FAR as possible
- Can HERMES be used to publish FAIR research datasets?
  - Yes, but it would take some work
- Should HERMES be used to publish FAIR datasets?
  - It depends ...



# Thanks!



HERMES:

[software-metadata.pub](https://software-metadata.pub), [docs.software-metadata.pub](https://docs.software-metadata.pub)  
[github.com/softwarepub/hermes](https://github.com/softwarepub/hermes), [github.com/softwarepub/ci-templates](https://github.com/softwarepub/ci-templates)  
[pypi.org/project/hermes/](https://pypi.org/project/hermes/), HMC project ZT-I-PF-3-006

Me:

Stephan Druskat, [stephan.druskat@dlr.de](mailto:stephan.druskat@dlr.de), [fosstodon.org/@sdruskat](https://fosstodon.org/@sdruskat)

References:

Druskat, S., Bertuch, O., & Struck, A. (2023). Towards Research Software-ready Libraries. ABI Technik, 43(3), 168-178. doi: [10.1515/abitech-2023-0031](https://doi.org/10.1515/abitech-2023-0031).

Chue Hong, N P., et al. (2022). FAIR Principles for Research Software (FAIR4RS Principles). RDA. doi: [10.15497/RDA00068](https://doi.org/10.15497/RDA00068).

Druskat, S., & Theis, S., (submitted). Challenges in designing research infrastructure software in multi-stakeholder contexts, in 27th International Conference on Human-Computer Interaction, Gothenburg, Sweden, submitted.

Mejias, G., (2023). DOIs for research software: Increasing Visibility, Connectivity, Citability. Zenodo. doi: [10.5281/zenodo.7985199](https://doi.org/10.5281/zenodo.7985199).

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

Meinel, M., Druskat, S., Kelling, J., Bertuch, O., Knodel, O., Pape, D., & Kernchen, S. (2024). hermes (v0.8.1). Zenodo. doi: [10.5281/zenodo.13311079](https://doi.org/10.5281/zenodo.13311079).

Kernchen, S., Meinel, M., Druskat, S., Fritzsche, M., Pape, D., & Bertuch, O. (2024). Extending and Applying Automated HERMES Software Publication Workflows. arXiv. doi: [10.48550/arXiv.2410.17614](https://doi.org/10.48550/arXiv.2410.17614).

Fritzsche, M., & Meinel, M. (2024). hermes-plugin-python (0.2.0). Zenodo. doi: [10.5281/zenodo.13168127](https://doi.org/10.5281/zenodo.13168127).

Jones et al. (2023). CodeMeta: an exchange schema for software metadata (Version 3.0). [Online]. Available: <https://w3id.org/codemeta/v3.0>

# HERMES: Example

[github.com/sdruskat/hermes-demo](https://github.com/sdruskat/hermes-demo)



hermes-demo Public Pin Unwatch 1 Fork 0 Star 0

main 1 Branch 0 Tags  Add file Code

sdruskat Initial commit f005271 · 7 minutes ago 1 Commit

README.md Initial commit 7 minutes ago

README

## hermes-demo

A repository to demonstrate automated software publishing with hermes

### About

A repository to demonstrate automated software publishing with hermes.

[software-metadata.pub](#)

continuous-integration hermes  
research-software software-publication  
fair4rs

- Readme
- Activity
- 0 stars
- 1 watching
- 0 forks

### Releases

No releases published  
[Create a new release](#)

# HERMES: Example

[github.com/sdruskat/hermes-demo](https://github.com/sdruskat/hermes-demo)



hermes-demo Public

Pin Unwatch 1 Fork 0 Star 0

main 1 Branch 0 Tags

Go to file Add file Code

sdruskat	Format README	92f83b8 · 2 minutes ago	4 Commits
.gitignore	Ignore .hermes/ cache folder	6 minutes ago	
CITATION.cff	Add citation file	2 minutes ago	
README.md	Format README	2 minutes ago	

README

## hermes-demo

A repository to demonstrate automated software publishing with hermes.

**Cite this repository**

If you use this software in your work, please cite it using the following metadata. [Learn more about CITATION files.](#)

APA BibTeX

Druskat, S. (2024). HERMES demo (Version v0.)

View citation file

# HERMES: Example

[github.com/sdruskat/hermes-demo](https://github.com/sdruskat/hermes-demo)

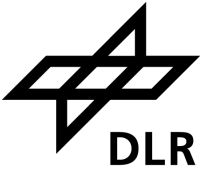


The screenshot shows a GitHub repository page for 'sdruskat / hermes-demo'. The navigation bar includes 'Code', 'Issues', 'Pull requests', 'Actions', 'Projects', and 'Wiki'. The current view is the 'Code' tab, showing a commit by 'sdruskat' titled 'Add HERMES publication configuration'. The file 'hermes.toml' is selected, and the diff shows 10 lines of code (7 LOC) and 150 bytes. The code content is as follows:

```
1  [harvest]
2  sources = [ "cff" ]
3
4  [deposit]
5  target = "invenio_rdm"
6
7  [deposit.invenio_rdm]
8  site_url = "https://sandbox.zenodo.org"
9  access_right = "open"
```

# HERMES: Example

[github.com/sdruskat/hermes-demo](https://github.com/sdruskat/hermes-demo)



ci-templates Public

Edit Pins Unwatch 11 Fork 1 Star 1

main 3 Branches 0 Tags

Go to file Add file Code

poikilotherm Update hermes-ci.yml 71d11c0 · 8 months ago 26 Commits

.github/workflows	added github action to run pre-commit (closes #5)	last year
LICENSES	Add license files	last year
gitlab	Update hermes-ci.yml	8 months ago
.pre-commit-config.yaml	added github action to run pre-commit (closes #5)	last year
README.md	fixed coding style in README.md	last year
TEMPLATE_hermes_github_to_zenodo.yml	Update TEMPLATE_hermes_github_to_zenodo.yml	8 months ago
TEMPLATE_hermes_gitlab_to_zenodo.yml	Fix workflow syntax	last year

README

## HERMES continuous integration templates

This repository contains templates for continuous integration systems that can be used to publish software automatically.

About

Continuous integration templates for automatic software publication with HERMES

[docs.software-metadata.pub](https://docs.software-metadata.pub)

Readme Activity Custom properties 1 star 11 watching 1 fork Report repository

Contributors 4

- sdruskat Stephan Druskat
- led02 Michael Meinel
- daniel-mohr
- poikilotherm Oliver Bertuch

# HERMES: Example

[github.com/sdruskat/hermes-demo](https://github.com/sdruskat/hermes-demo)



```
hermes-demo$ git status
On branch main
Your branch is up to date with 'origin/main'.

nothing to commit, working tree clean
hermes-demo$ git tag -a v0.1.0 -m "v0.1.0"
hermes-demo$ git push origin tag v0.1.0
Enumerating objects: 1, done.
Counting objects: 100% (1/1), done.
Writing objects: 100% (1/1), 162 bytes | 162.00 KiB/s, done.
Total 1 (delta 0), reused 0 (delta 0), pack-reused 0
To github.com:sdruskat/hermes-demo.git
 * [new tag]          v0.1.0 -> v0.1.0
hermes-demo$
```

# HERMES: Example

[github.com/sdruskat/hermes-demo](https://github.com/sdruskat/hermes-demo)



← Software Publication

✓ Copy template and configure to publish on tags #3

🏠 Summary

Jobs

✓ Prepare Metadata for Curation

🕒 Publish Software with Curated Me...

🕒 Cleanup aborted curation branches

Run details

🔄 Usage

📄 Workflow file

## Prepare Metadata for Curation

succeeded now in 18s

- > ✓ Set up job
- > ✓ Run actions/checkout@v4
- > ✓ Run actions/setup-python@v5
- > ✓ Run pip install hermes
- > ✓ Run hermes harvest
- > ✓ Run hermes process
- > ✓ Run hermes curate
- > ✓ Run # Cache current branch for PR close job
- > ✓ Run peter-evans/create-pull-request@v5
- > ✓ Post Run actions/setup-python@v5
- > ✓ Post Run actions/checkout@v4
- > ✓ Complete job

# HERMES: Example

[github.com/sdruskat/hermes-demo](https://github.com/sdruskat/hermes-demo)



Metadata Curation for Commit 8532d954 #1

Open

github-actions wants to merge 1 commit into hermes/curate-8532d954 from hermes/curate-result-8532d954

Conversation 0 Commits 1 Checks 0 Files changed 3

github-actions bot commented 3 minutes ago

Please carefully review the attached metadata.  
If you are satisfied with the result, you may merge this PR, which will trigger publication.  
(Any temporary branches will be cleaned up.)

[create-pull-request] automated change 20b457b

github-actions bot force-pushed the hermes/curate-result-8532d954 branch from ce4ed5a to 20b457b 3 minutes ago Compare

Require approval from specific reviewers before merging  
Rulesets ensure specific people approve pull requests before they're merged. Add rule X

**This branch has no conflicts with the base branch**  
Merging can be performed automatically.

Merge pull request

You can also [open this in GitHub Desktop](#) or view [command line instructions](#).

Add a comment

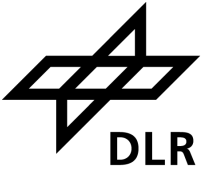
0 / 3 files viewed Review in codespace Review changes

```
1 + {
2 +   "@context": [
3 +     "https://doi.org/10.5063/schema/codemeta-2.0",
4 +     {
5 +       "hermes": "https://software-metadata.pub/ns/hermes/"
6 +     }
7 +   ],
8 +   "@type": "SoftwareSourceCode",
9 +   "author": [
10 +    {
11 +      "@id": "https://orcid.org/0000-0003-4925-7248",
12 +      "@type": "Person",
13 +      "affiliation": {
14 +        "@type": "Organization",
15 +        "legalName": "German Aerospace Center (DLR)"
16 +      },
17 +      "familyName": "Druskat",
18 +      "givenName": "Stephan",
19 +      "email": "stephan.druskat@dlr.de"
20 +    }
21 +   ],
22 +   "codeRepository": "https://github.com/sdruskat/hermes-demo",
23 +   "datePublished": "2024-11-14",
24 +   "description": "This is a project to demonstrate automated software
25 +     publication with HERMES CI workflows.",
26 +   "keywords": [
27 +     "HERMES",
28 +     "software publication",
29 +     "demo",
30 +     "Zenodo Sandbox",
31 +     "GitHub",
32 +     "continuous integration"
33 +   ],
34 +   "license": "https://spdx.org/licenses/CC0-1.0",
35 +   "name": "HERMES demo",
36 +   "url": "https://github.com/sdruskat/hermes-demo",
37 +   "version": "v0.1.0"
38 + }
```



# HERMES: Example

[github.com/sdruskat/hermes-demo](https://github.com/sdruskat/hermes-demo)



← Software Publication

## ✔ Metadata Curation for Commit 8532d954 #4

🏠 Summary

Jobs

- 🔄 Prepare Metadata for Curation
- ✔ Publish Software with Curated M...
- 🔄 Cleanup aborted curation branches

Run details

🕒 Usage

📄 Workflow file

### Publish Software with Curated Metadata

succeeded now in 24s

- > ✔ Set up job
- > ✔ Run actions/checkout@v3
- > ✔ Run actions/setup-python@v4
- > ✔ Run pip install hermes
- > ✔ Run hermes deposit --initial -O invenio\_rdm.auth\_token \*\*\* --file README.md
- > ✔ Run for BRANCH in \$(git ls-remote origin 'refs/heads/hermes/curate-\*' | cut -f2 | cut -d'/' -f3-); do
- > ✔ Post Run actions/setup-python@v4
- > ✔ Post Run actions/checkout@v3
- > ✔ Complete job

# HERMES: Example


[github.com/sdruskat/hermes-demo](https://github.com/sdruskat/hermes-demo)



zenodo Search records... Communities My dashboard

Published November 14, 2024 Version v0.1.0 Software Open

## HERMES demo

Druskat, Stepha  [Hide affiliations](#)

1. German Aerospace Center (DLR)

HERMES demo

### Files

Name	Size	
<a href="#">README.md</a>		<a href="#">Download all</a>
Files (86 Bytes)		
<a href="#">README.md</a>	86 Bytes	<a href="#">Preview</a> <a href="#">Download</a>

md5:ff39cc4401bed90231e1c8796e9b2c2