



# RESEARCH SOFTWARE ON WINGS

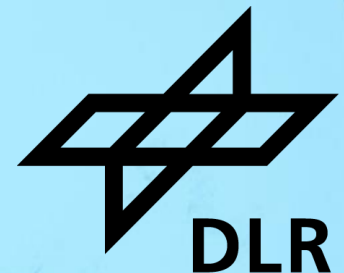
Automating software publication with rich metadata

DOI [10.5281/zenodo.7049910](https://doi.org/10.5281/zenodo.7049910)

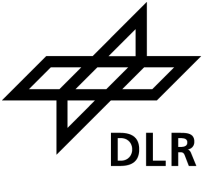
License [CC-BY-4.0 International](https://creativecommons.org/licenses/by/4.0/)

[Stephan Druskat](#)<sup>1</sup>, Oliver Bertuch<sup>2</sup>, Oliver Knodel<sup>3</sup>, Guido Juckeland<sup>3</sup>, Michael Meinel<sup>1</sup>, Tobias Schlauch<sup>1</sup>, Jeffrey Kelling<sup>3</sup>

<sup>1</sup> German Aerospace Center (DLR), Germany. <sup>2</sup> Forschungszentrum Jülich GmbH, Germany. <sup>3</sup> Helmholtz-Zentrum Dresden-Rossendorf (HZDR), Germany  
RSECon UK 2022, Newcastle-upon-Tyne, 2022-09-06



# Overview



- Motivation: Software publication
- HERMES: Automating software publication with rich metadata
- Where are we now?
- Outlook

The background of the slide is a high-resolution photograph of a satellite in orbit above Earth. The satellite is a rectangular platform with two long, parallel solar panel arrays extending outwards. The panels are covered in a grid of small solar cells. The satellite's central body is complex, with various instruments and antennas visible. Below the satellite, the Earth's surface is visible, showing a mix of green landmasses and blue oceans, partially obscured by white clouds. The curvature of the Earth is visible at the bottom of the frame.

# SOFTWARE PUBLICATION

# Motivation: Software publication



## Software publication

enables

enables

enables

enables

Sustainability

Reproducibility

Academic credit



# Software publication state of the art



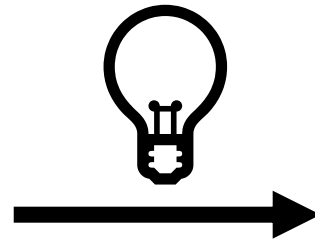
```
model = getattr(spectra, spectrum_dict["type"])

if norm.unit in (u.Unit("erg"), u.Unit("erg cm-3")) and norm_type != "integral":
    raise NameError(
        "Normalisation different than 'integral' available only for 'spectrum_norm' in cm-3"
    )

# check the units of the normalisation
# cm-3 is the only one allowing more than one normalisation type
if norm.unit == u.Unit("cm-3"):
    if norm_type == "differential":
        final_model = model(norm, **spectrum_dict["parameters"])
    elif norm_type == "gamma=1":
        final_model = model.from_norm_at_gamma_1(
            # ...
        )
    else:
        raise NameError(
            "Normalisation different than 'integral' available only for 'spectrum_norm' in cm-3"
        )

# ...

@context: "https://doi.org/10.5063/schema/codemeta-2.0",
@type: "SoftwareSourceCode",
license: "https://spdx.org/licenses/BSD-3-Clause",
codeRepository: "https://github.com/cosimoNigro/agnpy",
contIntegration: "https://github.com/cosimoNigro/agnpy/actions",
dateCreated: "2019-12-17",
datePublished: "2022-01-31",
dateModified: "2021-08-02",
downloadUrl: "https://github.com/cosimoNigro/agnpy/releases/tag/v0.1.6",
issueTracker: "https://github.com/cosimoNigro/agnpy/issues",
name: "agnpy",
version: "0.1.8",
identifier: "10.5281/zenodo.4055175",
description: "agnpy is a python package focusing on the computation of the
applicationCategory: "astrophysics",
funding: "ESCAPE EU H2020 824064",
developmentStatus: "active",
isPartOf: "https://www.astropy.org/affiliated/#affiliated-packages",
```



Software metadata + (software artifacts)

PID + landing page (metadata)

January 31, 2022 Software Open Access

## agnpy

Nigro, Cosimo; Sitarek, Julian; Gliwry, Paweł; Sanchez, David; Craig, Matthew; Vuillaume, Thomas

agnpy is a python package focusing on the computation of the radiative processes of relativistic particles accelerated in the jets of Active Galactic Nuclei (AGN). It includes classes describing the galaxy components responsible for line and thermal emission and calculates the absorption due to gamma-gamma pair production on soft (IR-UV) photon fields.

Available in **GitHub** and **OpenAIRE**

768 views, 157 downloads

Publication date: January 31, 2022  
DOI: 10.5281/zenodo.593285

Keywords: blazar, agn, jets, radiative processes, jupyter-notebook

Grants: European Commission, ESCAPE - European Science Cluster of Astronomy & Particle physics ESFR1 research infrastructures (824064)

Related identifiers: Supplement to https://github.com/cosimoNigro/agnpy/tree/v0.1.8

Communities: ESCAPE 2020

License (for files): BSD 3-Clause 'New' or 'Revised' License

Versions: Version 0.1.8 (10.5281/zenodo.593285), Version 0.1.7 (10.5281/zenodo.5927787)

# HERMES: Automating software publication



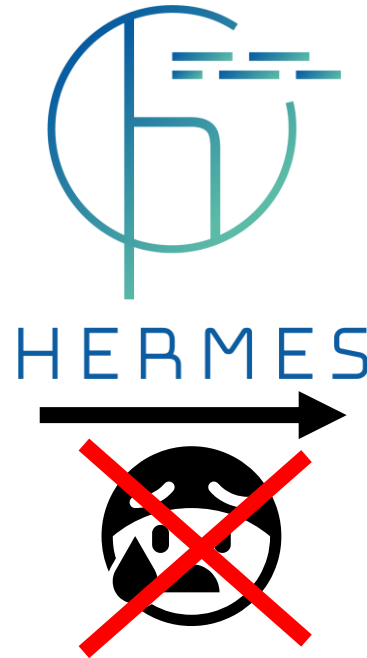
```
model = getattr(spectra, spectrum_dict["type"])

if norm.unit in (u.Unit("erg"), u.Unit("erg cm-3")) and norm_type != "integral":
    raise NameError(
        "Normalisation different than 'integral' available only for 'spectrum_norm' in cm-3"
    )

# check the units of the normalisation
# cm-3 is the only one allowing more than one normalisation type
if norm.unit == u.Unit("cm-3"):
    if norm_type == "differential":
        final_model = model(norm, **spectrum_dict["parameters"])
    elif norm_type == "gamma=1":
        final_model = model.from_norm_at_gamma_1(
            # ...
        )
    else:
        raise ValueError(f"Unknown normalisation type '{norm_type}'")

elif
```

```
{
  "@context": "https://doi.org/10.5063/schema/codemeta-2.0",
  "@type": "SoftwareSourceCode",
  "license": "https://spdx.org/licenses/BSD-3-Clause",
  "codeRepository": "https://github.com/cosimoNigro/agnpy",
  "contIntegration": "https://github.com/cosimoNigro/agnpy/actions",
  "dateCreated": "2019-12-17",
  "datePublished": "2022-01-31",
  "dateModified": "2021-08-02",
  "downloadUrl": "https://github.com/cosimoNigro/agnpy/releases/tag/v0.1.6",
  "issueTracker": "https://github.com/cosimoNigro/agnpy/issues",
  "name": "agnpy",
  "version": "0.1.8",
  "identifier": "10.5281/zenodo.4055175",
  "description": "agnpy is a python package focusing on the computation of the",
  "applicationCategory": "astrophysics",
  "funding": "ESCAPE EU H2020 824064",
  "developmentStatus": "active",
  "isPartOf": "https://www.astropy.org/affiliated/#affiliated-packages",
}
```



January 31, 2022 Software Open Access

## agnpy

Nigro, Cosimo; Sitarek, Julian; Gliwry, Paweł; Sanchez, David; Craig, Matthew; Vuillaume, Thomas

agnpy is a python package focusing on the computation of the radiative processes of relativistic particles accelerated in the jets of Active Galactic Nuclei (AGN). It includes classes describing the galaxy components responsible for line and thermal emission and calculates the absorption due to gamma-gamma pair production on soft (IR-UV) photon fields.

Preview

agnpy-v0.1.8.zip

- cosimoNigro-agnpy-6abd22
  - github
    - workflows
      - pip-upload.yml 705 Bytes
      - test.yml 1.2 kB
    - .gitignore 403 Bytes
    - .pylintrc 19.5 kB
    - .zenodo.json 1.9 kB
    - LICENSE 1.5 kB
    - MANIFEST.in 322 Bytes
    - README.md 2.4 kB
    - agnpy
      - \_\_init\_\_.py 202 Bytes
      - absorption
        - \_\_init\_\_.py 26 Bytes
        - absorption.py 30.3 kB
      - compton

Files (5.3 MB)

Name	Size	Preview	Download
cosimoNigro/agnpy-v0.1.8.zip	5.3 MB	<a href="#">Preview</a>	<a href="#">Download</a>
md5:87c5a702433f2aaa3d5106d1cafbd024			

Citations (4)

Show only:  Literature (3)  Unknown (1)  Dataset (0)  Software (0)  Citations to this version

- VHE gamma-ray detection of FSRQ QSO B1420+326 and modeling ... Acciari, V. A. et al. (DOI: 10.1051/0004-6361/202039687) 2021 [ADS](#) [ARXIV](#) [DOI](#)
- agnpy: an open-source python package modelling the radiativ... Nigro, C. et al. 2021 [ADS](#) [ARXIV](#)
- Flaremodel: An open-source Python package for one-zone nume... Dall'lar, V. et al. 2021 [ADS](#) [ARXIV](#)
- ADS: 2021arXiv211112926M [ADS](#) [ARXIV](#)

768 views 157 downloads [See more details...](#)

Available in

**GitHub**

Indexed in

**OpenAIRE**

Publication date: January 31, 2022

DOI: [10.5281/zenodo.5932850](https://doi.org/10.5281/zenodo.5932850)

Keyword(s): [blazar](#) [agn](#) [jets](#) [radiative processes](#) [jupyter-notebook](#)

Grants: [European Commission](#)

- ESCAPE - European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures (824064)

Related identifiers: Supplement to <https://github.com/cosimoNigro/agnpy/tree/v0.1.8>

Communities: ESCAPE 2020

License (for files): [BSD 3-Clause 'New' or 'Revised' License](#)

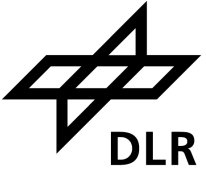
Versions

- Version 0.1.8 10.5281/zenodo.5932850 Jan 31, 2022
- Version 0.1.7 10.5281/zenodo.5927787 Jan 31, 2022

A satellite with two long solar panel arrays is shown in orbit above Earth. The satellite has a central body with various instruments and antennas. The solar panels are composed of many small rectangular cells. The Earth below shows a coastline with green land and blue water, partially obscured by white clouds. The curvature of the Earth and the blackness of space are visible in the background.

# HERMES PROJECT

# HERMES: Project



- 07/2021 – 06/2023
- Aim: Support RSEs in automatically publishing their software with rich metadata



[[arXiv:2201.09015](https://arxiv.org/abs/2201.09015) | [PubPeer](https://pubpeer.com/public?id/220109015)] | [software-metadata.pub](https://software-metadata.pub)



# HERMES: Scope

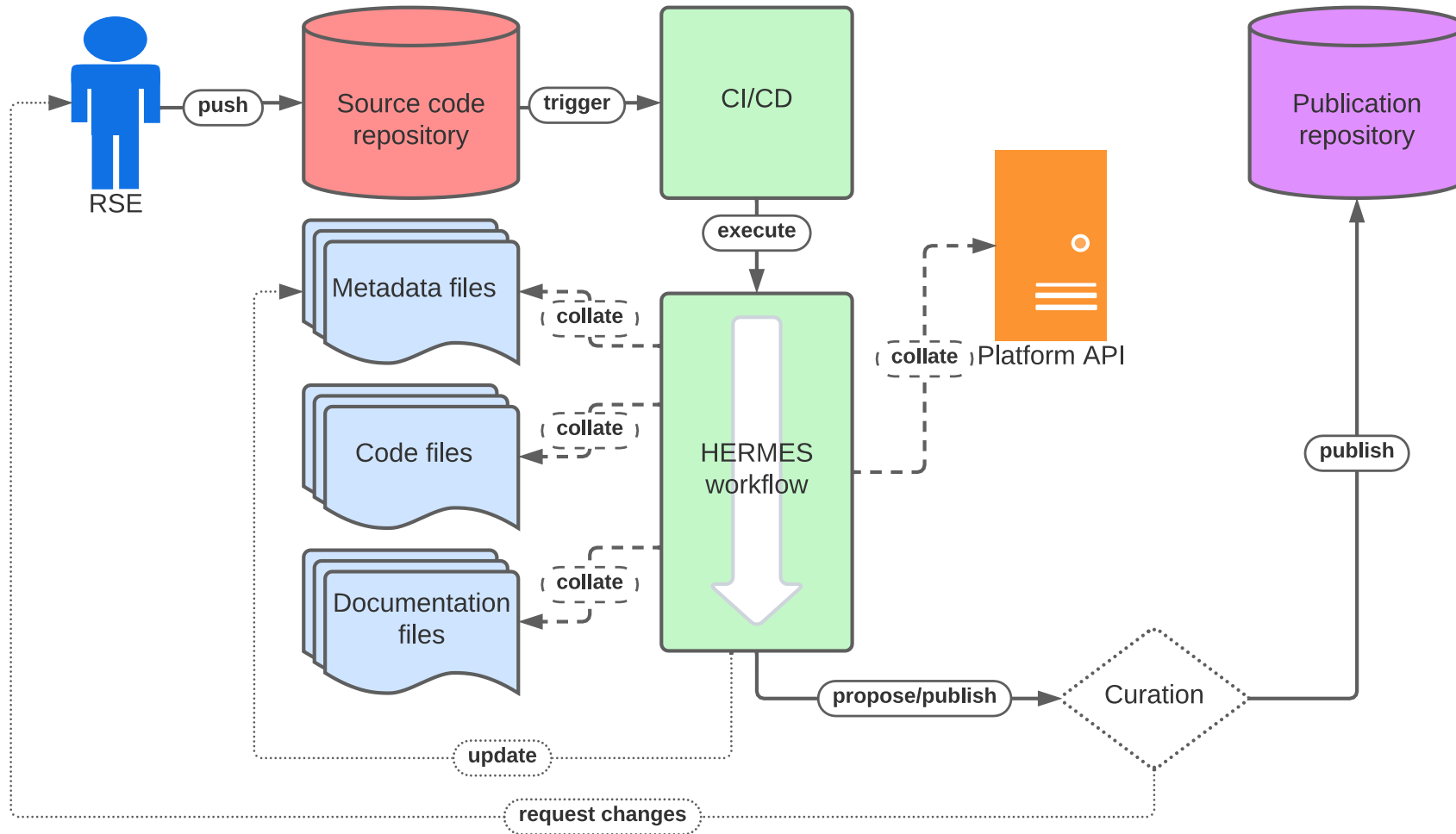
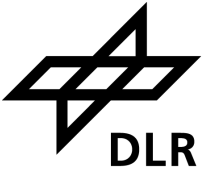


The user receives assistance in depositing software in an automated fashion. This may be used to create publications purely with rich metadata (to be at least FAIR [5], even for closed source software) or with attached artifacts like source code, executables, etc. (to be more easily reusable). To achieve this, HERMES provides

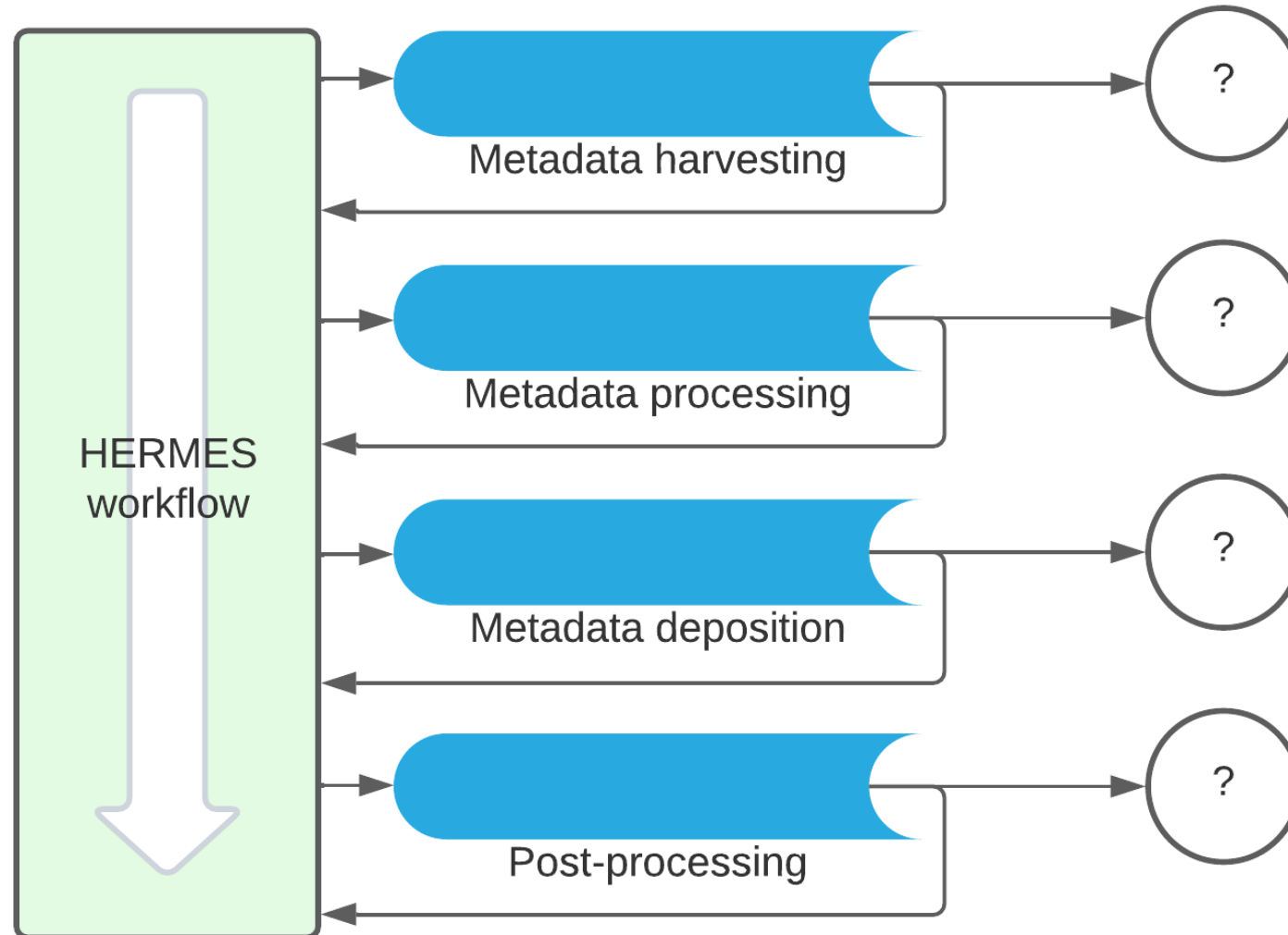
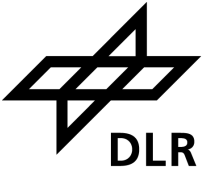
- an extensible, configurable and automatable toolchain with capability to be executed for<sup>15</sup>
  - N software publications in
  - M target publication repositories
  - from the same origin
  - as configured by the user,
- initially harvesting and collating [statically available metadata](#) from formerly described [metadata sources](#) and
- initially targeting
  - [InvenioRDM](#) and
  - [Dataverse project](#)
- for deposits of metadata and artifacts according to curator-defined requirements
- and output of the respective metadata in a structured format (e.g., [CodeMeta files](#)) for further reuse.

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*, [abs/2201.09015](https://arxiv.org/abs/2201.09015).

# HERMES: Concept



# HERMES: Workflow pipelines



- **Metadata**

- Differences in generation, scope, mode, aspects
- Generic software metadata vs. software-specific metadata

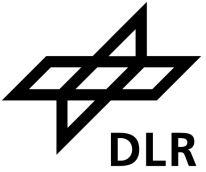
- **Metadata formats**

- Metadata files, snippets, third-party systems, API responses
- Structured vs. unstructured

- **Sources**

- Collectable structured metadata
- (Metadata from minable structured data)
- (Metadata from minable unstructured data)

# HERMES: Outputs

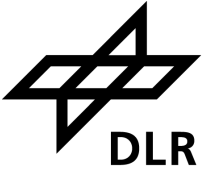


- **Software**
  - Software for software publication workflow automation (workflow runner + modular pipelines)
- **CI templates**
  - GitLab CI, GitHub Actions, Jenkins, Travis CI
- **Improved research software-readiness in publication repositories**
  - Position paper “research software-ready repositories”
  - Respective contributions to Dataverse + InvenioRDM (data models, UI)
- **Training materials**
  - Adaption of open Helmholtz training materials (HIFIS) to include workflow usage
- **Project website**
  - One-stop shop for information and documentation
- **Policy proposals**
  - Proposals for updates to policies/guidelines at Helmholtz and cross-institutional

The background of the slide is a high-resolution photograph of a satellite in orbit. The satellite is a rectangular platform with two long, thin solar panel arrays extending outwards. It is positioned in the center of the frame, with the Earth's surface visible below. The Earth shows a mix of green landmasses, blue oceans, and white cloud cover. The curvature of the planet is visible on the right side, where the atmosphere transitions into the blackness of space.

# PROJECT PROGRESS

# HERMES: Roadmap



## Concept

- Concept paper  /
- Requirements analysis  /

## Implementation

- HERMES tools
- Dataverse / InvenioRDM data model + UI
- Project website  /

## Distribution

- see outputs

# HERMES: Where are we now?



## ▪ **Community outreach and consultation**

- Stakeholders: repository projects, related metadata tools, infrastructure providers
- Users: RSEs, IT departments, computing centres ← **you**

Test-drive & have your say at our **workshop** „Cooking FAIR research software with HERMES“:  
**Wednesday, 7 Sep, 09:00-12:30**, room 2.16

## ▪ **Workflow implementation**

- Harvesting: Citation File Format, CodeMeta, git metadata
- Processing: # *TODO* (unified data model)
- Curation: # *TODO* (user feedback via logs)
- Deposition, post-processing: # *TODO* (CodeMeta files)

## ▪ **Research software-ready repositories**

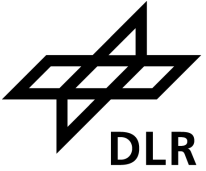
- Dataverse, InvenioRDM



A satellite with two large solar panel arrays is shown in orbit above Earth. The satellite is oriented vertically, with its central body and instruments visible. The solar panels are extended horizontally. The Earth's surface below shows a mix of green land, blue water, and white clouds. The curvature of the planet is visible on the right side of the frame.

# OUTLOOK

# HERMES: Outlook



- **Project ends:**

- Automated research software publication with rich metadata

- **Potential future work:**

- New target repositories:
  - SURESOFT (TU Braunschweig)!
  - Helmholtz research software directory?
  - NFDI publication repositories?
- New metadata types via extended metadata mining
- Curation UI
- Support for research software KPIs

# Thanks!



## Get in touch ...

<https://software-metadata.pub>  
[team@software-metadata.pub](mailto:team@software-metadata.pub)

(... with me)

[stephan.druskat@dlr.de](mailto:stephan.druskat@dlr.de)

Twitter: [@stdruskat](https://twitter.com/stdruskat)



## HERMES at RSECon



Stephan  
Druskat



Michael  
Meinel



Oliver  
Bertuch



Jeffrey  
Kelling



Oliver  
Knodel

