



# DataCite Commons: A new service for discovering software and software citations

**Martin Fenner**

DataCite Technical Director

<https://orcid.org/0000-0003-1419-2405>



# The FREYA project



Persistent Identifiers (PIDs) for the European Open Science Cloud (EOSC) with a focus on

- New PIDs (e.g. organizations, instruments)
- **PID Graph**: Connections between PIDs
- **PID Forum**: Training and Outreach
- **PID Commons**: Sustainability

Website: [www.project-freya.eu](http://www.project-freya.eu)



FREYA project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777523.



# PID Graph Building Blocks

---



## **PID Graph – Concept**

The connected searchable Graph of publications, datasets, other research outputs, people and organizations.

## **GraphQL – API**

The API and query interface that enables searching the FREYA PID Graph implementation.

## **Jupyter Notebooks**

A collection of openly available notebooks to generate PID Graph reports using the GraphQL API.

## **DataCite Commons**

The web search interface built with React using the GraphQL API to explore the PID Graph.

# Data Sources

The following main data sources are used in DataCite Commons for a total of currently 38,472,046 records:

## DataCite

19,897,461 Works

100% of identifiers and metadata.

## Crossref

8,771,341 Works

7.48% of identifiers and metadata. Import is ongoing.

## ORCID

9,704,912 People

100% of identifiers. Personal and employment metadata.

## ROR

98,332 Organizations

100% of identifiers and metadata.

Additional information comes from these data sources:

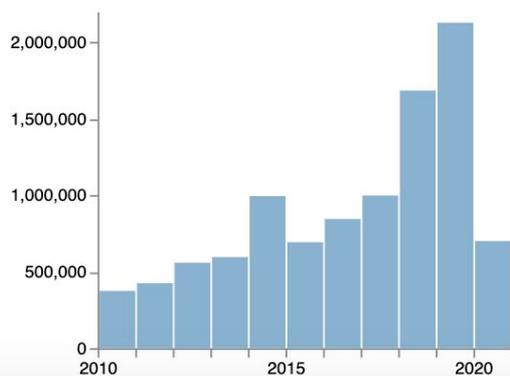
- [Wikidata](#): inception year, geolocation and Twitter account for organizations
- [Unpaywall](#): download link for Open Access content via Crossref

We want to add more!

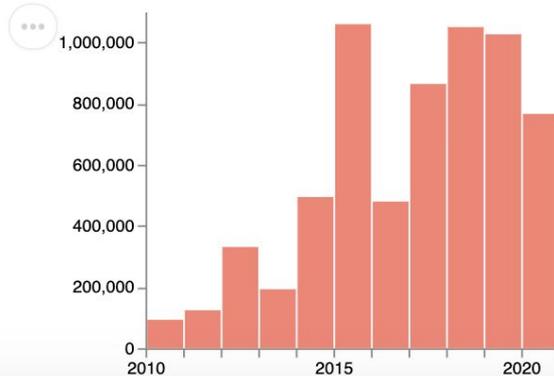
## Works

DataCite Commons currently includes 28,668,802 works, with identifiers and metadata provided by DataCite and Crossref. For the three major work types [publication](#), [dataset](#) and [software](#), the respective numbers by publication year are shown below.

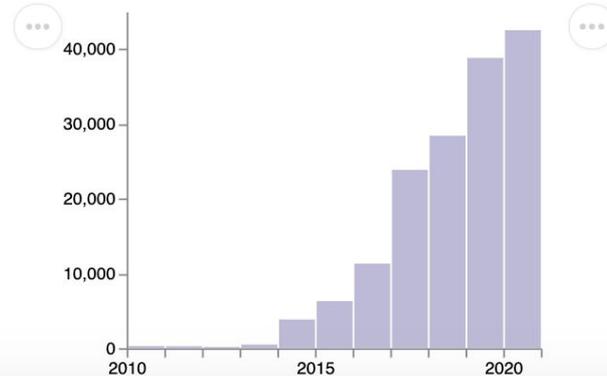
### 14,412,849 Publications



### 7,535,822 Datasets



### 157,363 Software



# Single Search Interface

DataCite Commons

python



About

Support

Works

People

Organizations

## 341 Works

### [datalad/datalad: 0.13.3 \(August 28, 2020\)](#)

[Michael Hanke](#), [Yaroslav O. Halchenko](#), [Benjamin Poldrack](#), Kyle Meyer, [Debanjum Singh Solanky](#), [Gergana Alteva](#), [Jason Gors](#), [Dave MacFarlane](#), [Christian Olaf Häusler](#), [Taylor Olson](#), [Alex Waite](#), [Alejandro De La Vega](#), [Vanessa Sochat](#), [Anisha Keshavan](#), [Feilong Ma](#), [Horea Christian](#), [Jorrit Poelen](#), [Kusti Skytén](#), [Matteo Visconti di Oleggio Castello & Nell Hardcastle](#)

Version 0.13.3 of Content published 2020 in [Zenodo](#)

Fixes Work around a Python bug that led to our asyncio-based command runner intermittently failing to capture the output of commands that exit very quickly. #4835 push displayed an overestimate of the transfer size when multiple files pointed to the same key. #4821 When download-url calls `git annex addurl`, it catches and reports any failures rather than crashing. A change in v0.12.0 broke this handling in a particular case. #4817 Enhancements and new features The wrapper functions returned by decorators are now given more meaningful names to hopefully make tracebacks easier to digest. #4834

DOI registered June 15, 2017 via DataCite.

[license\\_mit](#)

[Software](#)

<https://doi.org/10.5281/zenodo.808846>

## Publication Year

2020 341

## Work Type

Software 341

## License

MIT 341

CC-BY-NC-4.0 1

<https://commons.datacite.org>

# Filter by Software with Citations

DataCite Commons

citationCount:[1 TO \*]



[Works](#) [People](#) [Organizations](#)

Find all works with at least one citation.

[Support](#)

1,376 Works

## Publication Year

<input type="checkbox"/>	2020	418
<input type="checkbox"/>	2019	243
<input type="checkbox"/>	2018	273
<input type="checkbox"/>	2017	185
<input type="checkbox"/>	2016	98
<input type="checkbox"/>	2015	64
<input type="checkbox"/>	2014	61
<input type="checkbox"/>	2013	19
<input type="checkbox"/>	2012	5
<input type="checkbox"/>	2011	5
<input type="checkbox"/>	2008	1

## A case study for efficient management of high throughput primary lab data

Christian Colmsee, Steffen Flemming, Matthias Klapperstück, Matthias Lange & Uwe Scholz

Source Code published 2011 in [IPK Gatersleben](#)

Source code for the LimsLight primary data management system

DOI registered September 7, 2011 via DataCite.

👤 1 Citations

[Software](#) [English](#)

<https://doi.org/10.5447/ipk/2011/0>

## Illumiprocessor - software for Illumina read quality filtering

Brant C. Faircloth

Content published 2011 in [UC Los Angeles](#)

DOI registered March 18, 2012 via DataCite.

👤 1 Citations

[Software](#)

<https://doi.org/10.6079/j9ill>

## Work Type

<input checked="" type="checkbox"/>	Software	1,376
-------------------------------------	----------	-------

# Proper Navigation of Software Versions on Development Roadmap

DataCite Commons

nbgrader



About Support

Works

People

Organizations

## Publication Year

- 2019 3
- 2017 1
- 2016 1
- 2015 4

## Work Type

- Software 6
- Text 2
- Audiovisual 1

## License

- BSD-3-clause 2

## Language

- English 1

## Registration Agency

- DataCite 7
- Crossref 2

## 9 Works

### Nbgrader V0.1.0

Jessica Hamrick

Content published 2015 in [Zenodo](#)

I'm happy to announce that the first version of nbgrader has (finally) been released! nbgrader is a tool that I've been working on for a little over a year now which provides a suite of tools for creating, releasing, and grading assignments in the Jupyter notebook. So far, nbgrader has been used to grade assignments for the class I ran in the spring, as well as two classes that Brian Granger has taught. If you have any questions, comments, suggestions, etc., please do open an issue on the bugtracker. This is still a very new tool, so I am sure there is a lot that can be improved upon! Thanks so much to all of the people who have contributed to this release by reporting issues and/or submitting PRs: alope107 Carreau ellachao ellis...

DOI registered August 19, 2015 via DataCite.

Software

<https://doi.org/10.5281/zenodo.28549>

### Nbgrader V0.2.0

Jessica Hamrick

Content published 2015 in [Zenodo](#)

Version 0.2.0 of nbgrader primarily adds support for version 4.0 of the Jupyter notebook and associated project after The Big Split. The full list of major changes are: Jupyter notebook 4.0 support Make it possible to run the formgrader inside a Docker container Make course\_id a requirement in the transfer apps (list, release, fetch, submit, collect) Add a new assignment list extension which allows students to list, fetch, validate, and submit assignments from the notebook dashboard interface Auto-resize text boxes when giving feedback in the formgrader Deprecate the BasicConfig and NbGraderConfig classes in favor of a NbGrader class Thanks to the following contributors who submitted PRs or reported issues that were merged/clos...

DOI registered August 19, 2015 via DataCite.

Software

<https://doi.org/10.5281/zenodo.28550>

### Nbgrader: Nbgrader V0.2.1

Jessica B. Hamrick, Matthias Bussonnier, Jonathan Frederic, Brian E. Granger, Alope107, Ellachao, David Tsai, Carol Willing & RedSlug

Content published 2015 in [Zenodo](#)

# Export Metadata in a Variety of Formats, including Codemeta

---

<https://doi.org/10.5281/zenodo.28549>

---

## Nbgrader V0.1.0

Jessica Hamrick

Content published 2015 in [Zenodo](#)

I'm happy to announce that the first version of nbgrader has (finally) been released! nbgrader is a tool that I've been working on for a little over a year now which provides a suite of tools for creating, releasing, and grading assignments in the Jupyter notebook. So far, nbgrader has been used to grade assignments for the class I ran in the spring, as well as two classes that Brian Granger has taught. If you have any questions, comments, suggestions, etc., please do open an issue on the bugtracker. This is still a very new tool, so I am sure there is a lot that can be improved upon! Thanks so much to all of the people who have contributed to this release by reporting issues and/or submitting PRs: alope107 Carreau ellachao ellis...

DOI registered August 19, 2015 via DataCite.

Software

 <https://doi.org/10.5281/zenodo.28549>

TODO: include other identifiers such as SWHID

## Download

---

### Full Metadata

[DataCite XML](#)

[DataCite JSON](#)

[Schema.org JSON-LD](#)

### Citation Metadata

[Citeproc JSON](#)

[BibTeX](#)

[RIS](#)

[Codemeta](#)

# Affiliations, Contributors and Funding

---

## Creators

---

### Langbo Jia

State Key Laboratory of Petroleum Resource & Prospecting, College of Geosciences; China University of Petroleum; Beijing China

### Yong Zhou

State Key Laboratory of Petroleum Resource & Prospecting, College of Geosciences; China University of Petroleum; Beijing China

### Danjie Li

Material Supply Department; Sinopec Shengli Oilfield Company; Dongying China

### Liankai Jia

Shenzhen Branch of China National Offshore Oil Corporation Ltd.; Shenzhen China

### Dakang Zhong

State Key Laboratory of Petroleum Resource & Prospecting, College of Geosciences; China University of Petroleum; Beijing China

### Junlong Liu

Sinopec Exploration & Production Research Institute; Beijing China

### Ruitao Yan

Xi'an Shiyou University; Xi'an China

### Youliang Ji

State Key Laboratory of Petroleum Resource & Prospecting, College of Geosciences; China University of Petroleum; Beijing China

### Lijun Mi

Shenzhen Branch of China National Offshore Oil Corporation Ltd.; Shenzhen China

### Zhen Yi

Shenzhen Branch of China National Offshore Oil Corporation Ltd.; Shenzhen China

## Contributors

---

### S. Tyrrell

Editor

## Funding

---

### National Natural Science Foundation of China

41672098

# Software Citations

APA citation style shows version and label for software.

## Publication Year

2019 1

## Work Type

Text 1

## Registration Agency

Crossref 1

Find out more about the citing publication within DataCite Commons

## Cite

APA ▾

Michael Hirsch, P. D., Mayorov, N., Strandberg, J., Snehil Saluja, VOMAY, Kunas, E., & Estévez, D. (2020). *geospace-code/georinex: add sp3-d format* (Version v1.14.1) [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.595473>

## 1 Citation



## Evaluating Machine Learning & Antenna Placement for Enhanced GNSS Accuracy for CAVs

E. I. Adegoke, Jasmine Zidane, Erik Kampert, Paul A. Jennings, Col R. Ford, Stewart A. Birrell & Matthew D. Higgins

Conference published 2019 via Institute of Electrical and Electronics Engineers (IEEE)

DOI registered February 20, 2020 via Crossref.

Text

<https://doi.org/10.1109/ivs.2019.8813775>

# Software Citations in Publication References

## 2 References

### Publication Year

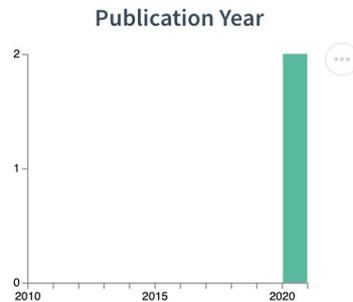
2020 2

### Work Type

Software 2

### Registration Agency

DataCite 2



### [geospace-code/georinex: add sp3-d format](#)

Ph.D. Michael Hirsch, Nikolay Mayorov, Joakim Strandberg, Snehil Saluja, VOMAY, Elias Kunnas & Daniel Estévez

Version v1.14.1 of Content published 2020 in [Zenodo](#)

Thanks to @fmeynadier for adding sp3-d format to Georinex.

DOI registered May 30, 2017 via DataCite.

🗉 1 Citations

Software

<https://doi.org/10.5281/zenodo.595473>

### [scivision/georinex: increased robustness](#)

Ph.D. Michael Hirsch, Nikolay Mayorov, Joakim Strandberg, Snehil Saluja, VOMAY, Elias Kunnas & Daniel Estévez

Version v1.14.0 of Content published 2020 in [Zenodo](#)

add ability to overwrite .nc files with "overwrite=True" make Rinex3 files handle INRSS add more header parameters make SP3 reading much more robust tolerate blank lines starting rinex file (default up to 10) use "unlzw3" for pure Python .Z files general test / CI / coverage improvements

DOI registered February 10, 2020 via DataCite.

🗉 1 Citations

Software

<https://doi.org/10.5281/zenodo.3660855>

# Researcher Profile

<https://orcid.org/0000-0003-1419-2405>

## Martin Fenner

Martin Hellmut Fenner

Martin Fenner is the DataCite Technical Director since 2015. From 2012 to 2015 he was the technical lead for the PLOS Article-Level Metrics project. Martin has a medical degree from the Free University of Berlin and is a Board-certified medical oncologist.

### Links

Twitter

### Other Identifiers

Scopus Author ID: [7006600825](#)

ISNI: [000000035060549X](#)

GitHub: [mfenner](#)

GitHub username  
from ORCID record

### Other Profiles

ORCID

Impactstory

Europe PMC

Germany

 <https://orcid.org/0000-0003-1419-2405>

## Share

 Email

 Twitter

 Facebook

All with MIT license

## 11 Works

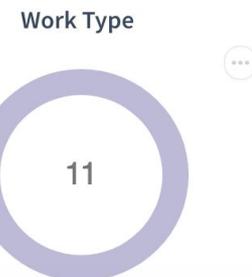
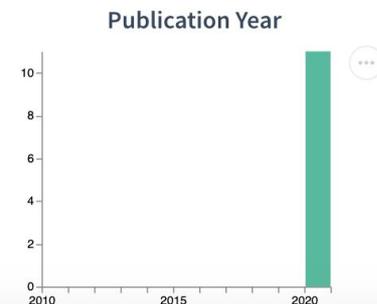
### Publication Year

2020 11

### Work Type

Software 11

### License



# Aggregation by Organization

<https://ror.org/00k4n6c32>

## European Commission

EC

### Links

[Homepage](#)

[Wikipedia](#)

### Other Identifiers

GRID [grid.270680.b](#)

Crossref Funder ID [10.13039/501100000780](#)

Crossref Funder ID [10.13039/501100000893](#)

Crossref Funder ID [10.13039/501100000891](#)

Crossref Funder ID [10.13039/501100000894](#)

Crossref Funder ID [10.13039/501100000887](#)

Wikidata [Q8880](#)

Wikidata [Q20855594](#)

Linked funded works via Crossref Funder ID

[Belgium](#) [Government](#)

<https://ror.org/00k4n6c32>

## Share

[Email](#)

[Twitter](#)

[Facebook](#)

## 683 Works

Filter and search within funded works

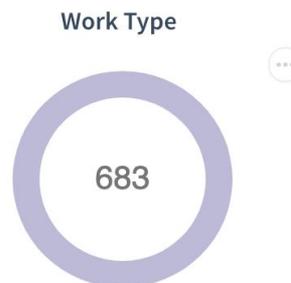
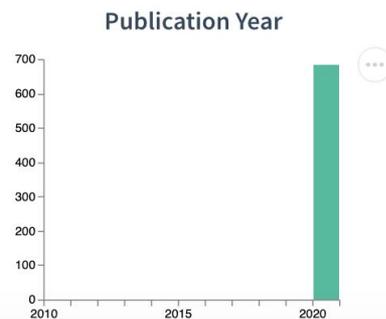
### Publication Year

2020 683

### Work Type

Software 683

### License



# DataCite Commons Official Launch

---



DataCite Commons will officially launch with the FREYA project by the end of October. Until then

- More data, more connections
- Personal accounts for ORCID claiming
- Bug fixes
- Feedback collection

Thanks!

