Software citation in ocean sciences: towards the reproducibility of results

Oceanography | Data analysis | Data citation | Reproducibility | SeaDataCloud | ODIP | Python | Jupyter | GitHub | Zenodo | orcID

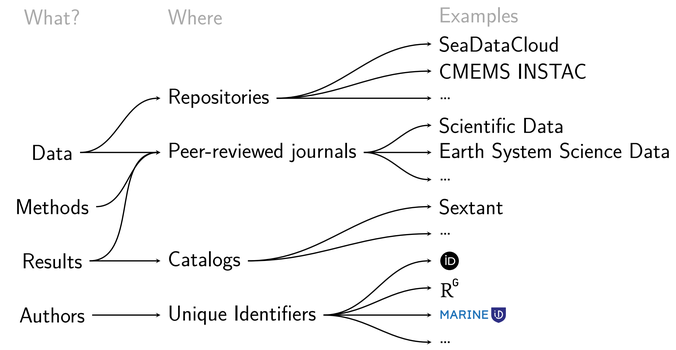
What will we talk about

- Digital Object Idenfitier: unique alphanumeric string assigned to identify content and provide a persistent link to its location on the Internet.
- GitHub: web-based hosting service for version control using `git`
  http://github.com/
- ORCID: persistent digital identifier to distinguishes researchers
  https://orcid.org/

Zenodo : a repository to deposit scientific papers and/or research data
https://zenodo.org/

How can we foster reproducibility

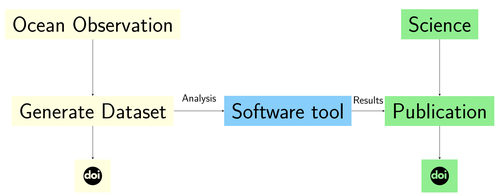
By making { Data, Method, Results } available and citable



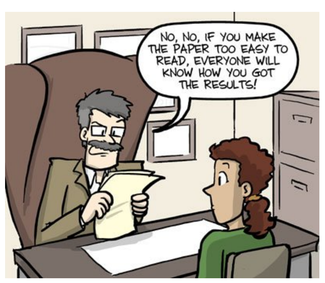
Unique identifiers for each components

– minted on { Datasets, Data products (e.g., climatologies) }

– identifies authors



Is this sufficient to go from data to results

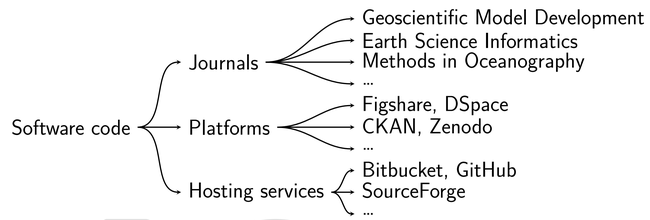


No!

The software tools should also be properly versioned and cited.

Making the code available

If the code of the software tool(s) is open, they are several ways to distribute it:



Research platforms: online infrastructures whose objective is to persistently store and archive digital artifacts relevant to research: articles, data, images, code, model outputs, ...
Comparison ↓.

Comparison

Tool	CKAN	DSpace	Figshare	Zenodo
Open Source	Yes	Yes	No	Yes
	ckan/ckan	DSpace/DSpace	–	zenodo/zenodo
Licence	Affero GNU GPL v3.0	BSD	–	GPL-2.0
1st released	November 2011	November 2002	January 2011	May 2013
Main technology	Python	Java	–	Python
Deployment	Local	Local	Cloud	Cloud
Integration with	No	No	Yes	Yes
Integration with	Yes	Not direct	Yes	Login

Let's have a closer look at Zenodo

✓ ingest all research outputs and any file format
✓ DOIs assigned to have uniquely citable files
✓ integrated into reporting lines for research via OpenAIRE.

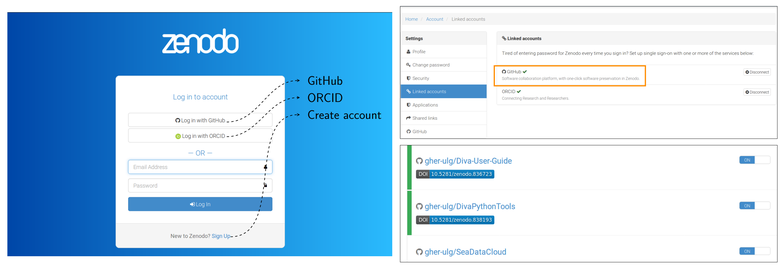


Figure 1: Zenodo log in: using other accounts is great feature. When linked with GitHub, Zenodo creates a DOI everytime you make a release in one of the enabled repositories.

We now have all the pieces to cite the code used in the research:

1. Zenodo log in using or
2. Upload of software code to or to Zenodo
3. Generation of the for a given version of the code

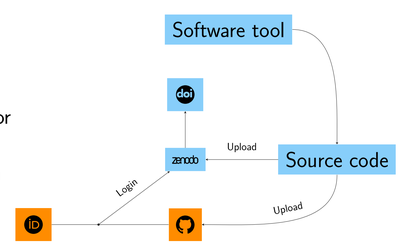


Figure 2: Getting unique identifiers for software codes using Zenodo.

Example: the DIVA interpolation tool

History

1990's: Variational Interpolation Method (Fortran 77) only 2D interpolations
2006 SeaDataNet, code refactory and set of bash scripts
2007 with ODV
2008 code in Subversion , distribution through GHER web page
2009 new modules in Fortran 90 for loops over depth and time
2012 new error calculation technique
2017 new version control system:

1. Switch from SVN to **git**, distribution via , sync with zenodo
2. Enable Diva repository in Zenodo
3. Edit the different *tags* on GitHub to get DOI

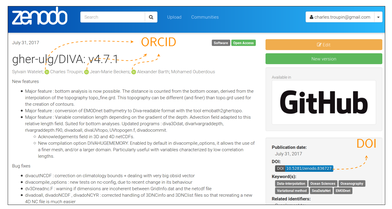


Figure 3: Main page of DIVA software in Zenodo. Note the ORCID logo with the authors and the DOI relative to the code. Share on social media and "cite as" options.

Putting all the pieces together

To ensure reproducibility and traceability, unique identifiers ( ) are attributed to:

- Datasets
- Software tools
- Authors
- Scientific results

Ideally, all the identifiers should be present in the published version of the research paper.

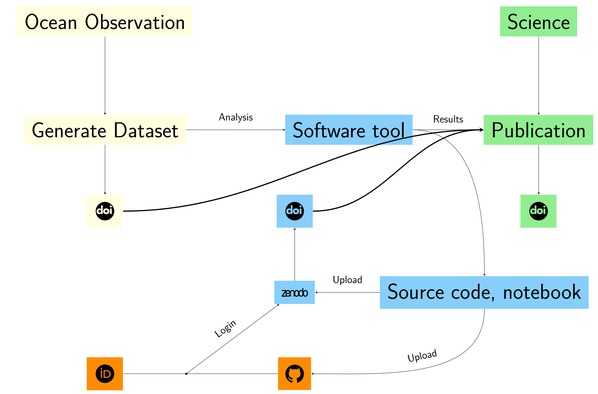


Figure 4: From data to final results: all the components are identified and citable.

Acknowledgements

The work presented in this poster has been developed in the frame of SeaDataCloud–*Further developing the pan-European infrastructure for marine and ocean data management*, Project ID 730960 and ODIP 2–*Extending the Ocean Data Interoperability Platform*, Project ID: 654310.

Source: "*English Communication for Scientists*", Nature, https://www.nature.com/scitable/ebooks/english-communication-for-scientists-14053993/writing-scientific-papers-14239285

Authors: Charles Troupin, Alexander Barth, Sylvain Watelet, and Jean-Marie Beckers

@GHER_ULiege https://github.com/gher-ulg/ http://labos.ulg.ac.be/gher/