

Software must be cited

Silvio Peroni

silvio.peroni@unibo.it – <https://orcid.org/0000-0003-0530-4305> – [@essepuntato](#)

[Open Science \(A.Y. 2020/2021\)](#)

[Second Cycle Degree in Digital Humanities and Digital Knowledge](#)

[Alma Mater Studiorum - Università di Bologna](#)



Choosing the license for software

It is important to choose the right license for your code

A description of licenses is available at:

<https://opensource.org/licenses>

<https://www.gnu.org/licenses/license-list.en.html>

Suggestion: choose a license that maximises the software reuse (i.e. with limited constraints)

One possible choice to this direction is [ISC](#)

Action item: you have to specify the license of your code in the GitHub repository – we can even decide to use just one license for all, included in the root of the repository

Get a DOI for your software

You can get a DOI for your software available on GitHub (to cite it within publications), following the guide available here:

<https://guides.github.com/activities/citable-code/>

We need to use GitHub releases for enable that and to link Zenodo with GitHub

Action item: I ask each group to put all (and only) the code in a new repository under your control (I can create it within the *open-sci* organisation, if you agree), and to link it with Zenodo

Please, use [semantic versioning](#) for assign version number to the releases

Are DOIs enough?

For the specific needs of reproducibility, referencing software artifacts via DOI, when it is available, is often sufficient

Please do not reference to software artifacts by just pointing to the project website or the current development repository, since it is necessary to have at hand a precise reference to the software source code used for the experiment, as well as long-term archival of the referenced artifact

Even better if the reference itself allows integrity checking upon artifact retrieval, enabling researchers that are attempting replication to rule out corruption or tampering with the digital objects as a potential cause for non-reproducibility

DOIs do not implement an integrity checking mechanism

Software Heritage identifiers

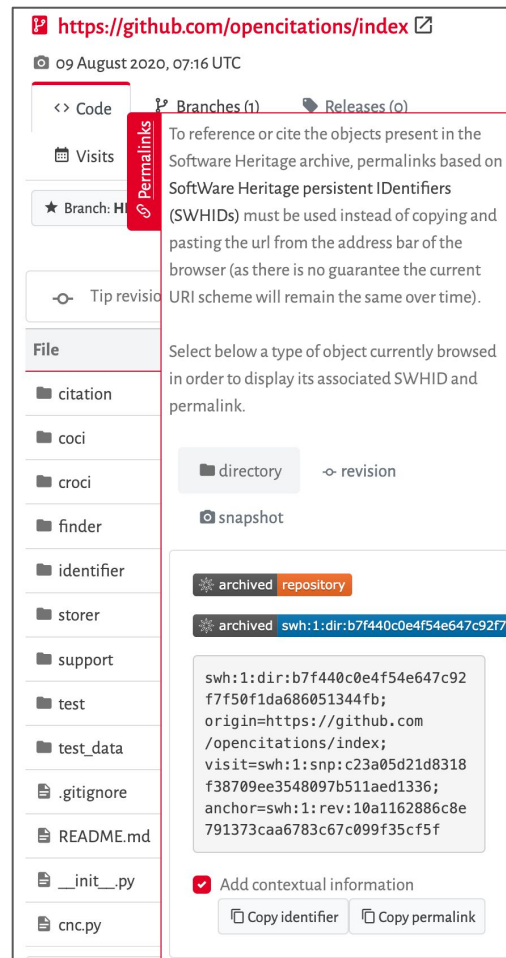
In case you have to cite software that you reuse for any reason within your one software, you can use the [Software Heritage archive](#) to retrieve the identifier of the archived version of the software to cite

Shape of the Software Heritage ID:

`swh:<schema_version>:<object_type>:<object_id>`

E.g.:

`Swh:1:dir:b7f440c0e4f54e647c92f7f50f1da686051344fb`



The screenshot shows the GitHub page for <https://github.com/opencitations/index>. The page title is "Permalinks" and the content explains that to reference or cite objects in the Software Heritage archive, permalinks based on Software Heritage persistent identifiers (SWHIDs) must be used instead of copying and pasting the URL from the browser address bar. The page includes a "File" sidebar with a list of files: citation, coci, croci, finder, identifier, storer, support, test, test_data, .gitignore, README.md, __init__.py, and cnc.py. The main content area shows a "directory" revision selected, with a "snapshot" button. Below this, there are two "archived repository" entries, with the second one selected. The selected entry shows the SWHID: `swh:1:dir:b7f440c0e4f54e647c92f7f50f1da686051344fb` and its origin: `https://github.com/opencitations/index`. The page also includes a "Add contextual information" checkbox and buttons for "Copy identifier" and "Copy permalink".

End

Software must be cited

Silvio Peroni

silvio.peroni@unibo.it – <https://orcid.org/0000-0003-0530-4305> – [@essepuntato](#)

[Open Science \(A.Y. 2020/2021\)](#)

[Second Cycle Degree in Digital Humanities and Digital Knowledge](#)

[Alma Mater Studiorum - Università di Bologna](#)

