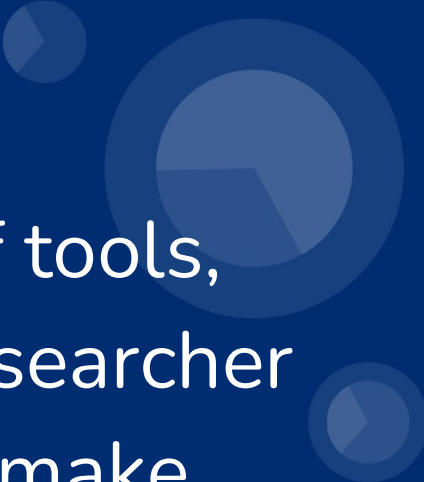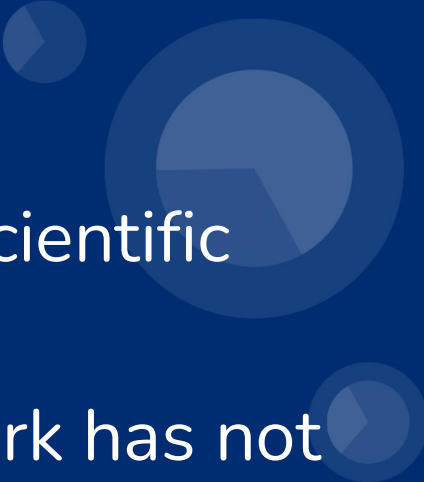# Citing Research Software

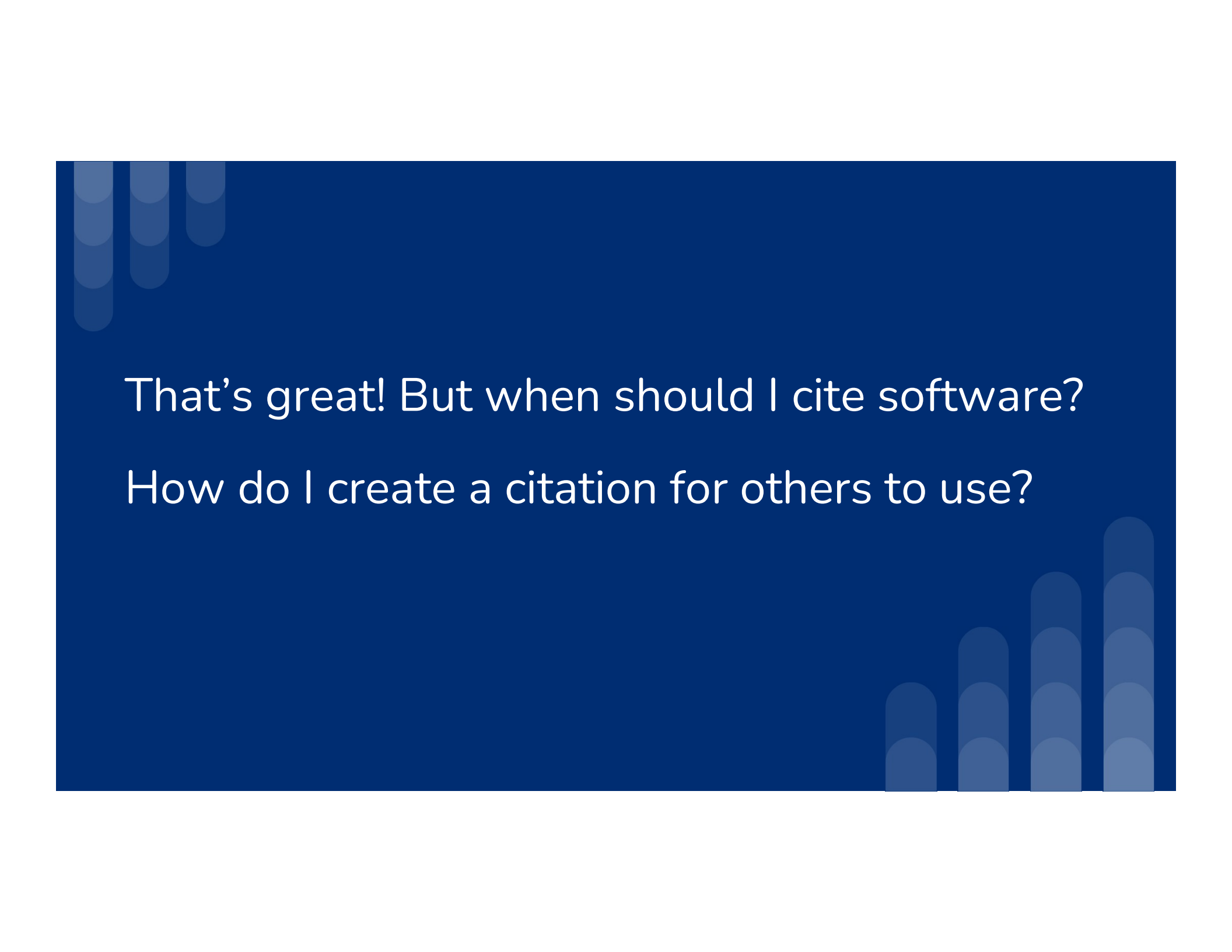## JHU Open Source Programs Office
Explainer

Research software is the collection of tools, code, or code libraries that allow a researcher to generate new data or analyze and make meaning of existing data.

"Research Software Programs." Accessed March 14, 2024. https://researchsupport.harvard.edu/research-software-programs

Software is foundationally important to scientific and social progress; however, traditional acknowledgment of the use of others' work has not adapted in step with the rapid development and use of software in research.

Bouquin, Daina, Ana Trisovic, Oliver Bertuch, and Elena Colón-Marrero. "Advancing Software Citation Implementation (Software Citation Workshop 2022)." arXiv, February 15, 2023. http://arxiv.org/abs/2302.07500.

One way to improve acknowledgement of software is to create and use software citations.

That's great! But when should I cite software?

How do I create a citation for others to use?

When should I cite software?

"You should cite software that has a significant impact on the research outcome presented in your work, or on the way the research has been conducted. If the research you are presenting is not repeatable without a piece of software, then you should cite the software. Note that the license or copyright of the software has no bearing on whether you should cite it."

Chue Hong, Neil P., Alice Allen, Alejandra Gonzalez-Beltran, Anita de Waard, Arfon M. Smith, Carly Robinson, Catherine Jones, et al. "Software Citation Checklist for Authors." Zenodo, October 15, 2019. https://doi.org/10.5281/zenodo.3479199.

In general, you should cite:

- Software (including scripts) you have written yourself to conduct the research presented.
- A software framework / platform upon which the software you wrote relies.
- Software packages, plugins, modules and libraries you used to conduct your research and that perform a critical role in your results.
- Software you have used to simulate or model phenomena/systems.
- Specialist software used to prepare, manage, analyse or visualise data.
- Software being evaluated or compared as part of the research presented.
- Software that has produced analytic results or other outputs, especially if used through an interface.

Chue Hong, Neil P., Alice Allen, Alejandra Gonzalez-Beltran, Anita de Waard, Arfon M. Smith, Carly Robinson, Catherine Jones, et al. "Software Citation Checklist for Authors." Zenodo, October 15, 2019. https://doi.org/10.5281/zenodo.3479199.

In general, you do not need to cite:

- Software packages or libraries that are not fundamental to your work and that are a normal part of the computational and scientific environment used. These dependencies do not need to be cited outright but should be documented as part of the computational workflow for complete reproducibility.
- Software that was used during the course of the research but had no impact on research results, e.g. word processing software, backup software.

Chue Hong, Neil P., Alice Allen, Alejandra Gonzalez-Beltran, Anita de Waard, Arfon M. Smith, Carly Robinson, Catherine Jones, et al. "Software Citation Checklist for Authors." Zenodo, October 15, 2019. https://doi.org/10.5281/zenodo.3479199.

How do I create a citation for others to use?

The best way to create a citation for your work that others can find is to add a .cff file to your code repository.

CFFINIT is a Citation File Format webapp that can generate or update a validated citation file for you based on a few inputs: https://citation-file-format.github.io/cff-initializer-javascript/#/

If you have deposited your code into the Johns Hopkins Research Data Repository, a standard citation will be created based on your deposit's metadata.

Citation files in GitHub

If you add a CITATION.cff file to the default branch of a **GitHub** repository, a link will be added to the right sidebar of your repository landing page with the label "Cite this repository."

Repository visitors can click on the link in the sidebar and copy an APA or BibTex-formatted citation.

Visit https://github.com/JH-OSPO/citation-examples to see it in action!

GitHub Docs. "About CITATION Files." Accessed August 1, 2024.
https://docs.github.com/en/repositories/managing-your-repositorys-settings-and-features/customizing-your-repository/about-citation-files.

Required citation elements

- authors - the authors of the software
  - Note - you can include an ORCID ID with each author name
- cff-version - The Citation File Format schema version that the CITATION.cff file adheres to for providing the citation metadata
- message - a message to the human reader of the CITATION.cff file to let them know what to do with the citation metadata
- title - the name of the software or dataset

Recommended citation elements

- abstract - a description of the software

- date-released - the date the software has been released. Format is 4-digit year, 2-digit month, 2-digit day of month, separated by dashes

- identifiers - the identifiers of the software

- keywords - keywords that describe the work

- license - the SPDX license identifier(s) for the license(s) under which the work is made available. When there are multiple licenses, it is assumed their relationship is OR, not AND

- repository-code - the URL of the work in a source code repository

- version - the version of the software

But wait, there's more!

Extended definitions with examples for all the .cff elements can be found via the Citation File Format Project's GitHub:

https://github.com/citation-file-format/citation-file-format/blob/main/schema-guide.md

Maintaining and Updating Citation Files

It's good practice to regularly review your citation file(s). The CFFINIT tool can be used to create new .cff files or edit existing ones. Events that might kick off a review or update:

- A new release with significant changes or improvements
- New contributors joining the project
- Change or modification to the software's license
- Change in official website or repository location

Questions? Ask the
JHU Open Source Programs Office

ospo@jhu.edu

https://ospo.library.jhu.edu