

# Open License Guide - Data

Updated as of: 10/13/20

**DISCLAIMER:** This document reflects the policy objectives of the ITS JPO, specifically the ITS JPO Data Access Guidelines, and applies ONLY to research projects funded, either fully or partially, by the ITS JPO. This document may not be suitable for other U.S. DOT or non-U.S. DOT missions. For any questions or concerns regarding the applicability of this document or the ITS JPO Data Access Guidelines, please contact the ITS JPO at [data.itsjpo@dot.gov](mailto:data.itsjpo@dot.gov).

For information concerning the assignment of open licenses to federally-funded custom-developed *source code*, including analytical code used to analyze data, see the ITS JPO Source Code Guidelines and the Open License Guide - Source Code.

## I. Introduction

Assigning the proper open license to data created or collected during a research project is critical to facilitating public access and maximizing the ITS JPO's return on investment. Per the [ITS JPO Data Access Guidelines](#) and wider federal guidance, including the Open Data Policy and OPEN Government Data Act, project teams must assign open licenses to federally-funded data as it is collected or created. The term "open license" is used to refer to any legally binding instrument that grants permission to access, re-use, and redistribute a work with few or no restrictions.<sup>1</sup> The assignment of an open license encourages reuse by placing prospective users on notice as to how they may use the data. What license a project team will use should be agreed upon in the project contract, grant, or other funding agreement.

## II. Creative Commons Zero (CC0 1.0 Universal)

The ITS JPO requires the assignment of [Creative Commons Zero \(CC0 1.0 Universal\)](#)<sup>2</sup> to all ITS JPO-funded data. CC0 waives any copyrights that may apply to the work and places both the data compilation and the underlying data in the public domain. CC0 was selected due to its applicability to data collected or created by both federal government employees and non-federal government employees, its wide and accepted usage (used also by [OpenFDA](#), [18f](#), and [Project Open Data](#)) and lack of requirements and restrictions that would impede reuse.

Other open licenses, such as [CC BY](#), or [CC BY-SA](#) are not the ITS JPO's top choices for licenses as they place restrictions on reuse and are not suitable for use by U.S. government employees since they require the author to grant a copyright license to users. Under Section 105 of the Copyright Act, when data is collected or compiled by federal government employees, the government cannot claim a copyright to the work and by extension grant a copyright license.

Other popular licenses, including GNU General Public License, Apache 2.0 and MIT are designed for software and not ideal for data. Proprietary licenses where the owner reserves and restricts all rights, commonly requiring a fee for usage, should be avoided due to their conditions or restrictions. A project team considering the use of data or software assigned a proprietary license should consult their legal counsel.

These aforementioned licenses (CC BY, CC BY-SA, GNU GPL, Apache 2.0, MIT, and proprietary licenses), all place some restrictions on reuse, including requiring that users provide attribution to the original author or requiring the assignment of the same license on copied or modified works. Creating a dataset derived from hundreds of sources, each with a license requiring attribution, creates a complex web referred to as "attribution stacking." These restrictive open licenses reduce the ability of the public to freely use the work and are not preferred over CC0, even for works created by non-U.S. government employees (e.g. works created by contractors or grantees).

In the event that a project team objects to assigning CC0, the ITS JPO may permit the assignment of [CC BY](#), or [CC BY-SA](#) on a case-by-case basis. These licenses, while not preferred over the CC0 license, include limited restrictions on reuse and are widely used and accepted.

Project teams will retain existing licenses for any pre-existing data integrated into project solutions. In the event pre-existing data is used, the ITS JPO only requires that project teams assign CC0 to new federally-funded data. Any data derived from existing data must respect the existing data's license and any restrictions. If the existing data used is assigned CC0 or a similar license, the derivative data may be assigned CC0. If derivative data is generated through the use of existing software, the project team must respect the software's rights with regards to "use." Project teams must make all licensing relationships with pre-existing data clear in applications and documentation, including Data Management Plans (DMPs).

The below chart provides additional guidance from the ITS JPO with regards to licenses. For specific inquiries or clarification on other licenses, please contact the ITS JPO at [data.itsjpo@dot.gov](mailto:data.itsjpo@dot.gov).

License Name	Status	Pros	Cons
--------------	--------	------	------

Creative Commons Zero (CC0 1.0 Universal)	Permitted: <b>Default</b> for all projects	<ul style="list-style-type: none"> <li>• “Open License”<sup>2</sup></li> <li>• Waives any or all copyright to the fullest extent of the law and places the work in the public domain</li> <li>• Suitable for code authored by federal and non-federal employees</li> <li>• No requirements on users for reuse <ul style="list-style-type: none"> <li>• No attribution requirement</li> <li>• Does not require the same license be applied to derivative works</li> </ul> </li> <li>• Compatible with other open licenses since it does not impose any restrictions on further reuse</li> <li>• Does not preclude <i>recommending</i> attribution <ul style="list-style-type: none"> <li>• <i>Note:</i> The ITS JPO requires the use of issued DOIs and recommends users provide attribution</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Releases the work to the public domain which some project teams may be uncomfortable with <ul style="list-style-type: none"> <li>• <i>Note:</i> The ITS JPO can help clarify any concerns with CC0 at data.itsjpo@dot.gov.</li> </ul> </li> </ul>
<u>CC BY</u>	Permitted in certain cases (exceptions to Default) <ul style="list-style-type: none"> <li>• Contact the ITS JPO at data.itsjpo@dot.gov if an exception is sought</li> </ul>	<ul style="list-style-type: none"> <li>• Open License</li> <li>• Suitable for code authored by non-federal employees</li> </ul>	<ul style="list-style-type: none"> <li>• Can never be used for data compiled by federal employees</li> <li>• Places burden on reuse by requiring: <ul style="list-style-type: none"> <li>• Users give appropriate attribution, provide a link to the license, and indicate if changes were made</li> </ul> </li> </ul>
<u>CC BY-SA</u>	Permitted in certain cases (exceptions to Default) <ul style="list-style-type: none"> <li>• Contact the ITS JPO at data.itsjpo@dot.gov if an exception is sought</li> </ul>	<ul style="list-style-type: none"> <li>• Open License</li> <li>• Suitable for code authored by non-federal employees</li> </ul>	<ul style="list-style-type: none"> <li>• Can never be used for data compiled by federal employees</li> <li>• Places burden on reuse by requiring: <ul style="list-style-type: none"> <li>• Users give appropriate attribution, provide a link to the license, and indicate if changes were made</li> <li>• Users apply the same license as the original</li> </ul> </li> </ul>
CC BY-NC; CC BY-ND; IBM Public License; Sun Microsystems License; DataStock License; and others, with similar or more significant restrictions	Not Permitted: Except on a case by case basis <ul style="list-style-type: none"> <li>• Contact the ITS JPO at data.itsjpo@dot.gov if an exception is sought</li> </ul>		<ul style="list-style-type: none"> <li>• Many place significant burdens on reuse by requiring: <ul style="list-style-type: none"> <li>• NC (Non-Commercial): No reuse in a proprietary work</li> <li>• ND (Non-Derivative): Users cannot transform, alter, or build upon the work</li> </ul> </li> <li>• Non-open licenses place even greater burdens on reuse by forbidding the creation of a derivative work and/or public distribution of a work</li> </ul>
Proprietary licenses, such as the DataStock License; and others, with similar or more significant restrictions	Not Permitted: Except on a case by case basis <ul style="list-style-type: none"> <li>• Contact the ITS JPO at data.itsjpo@dot.gov if an exception is sought</li> </ul>		<ul style="list-style-type: none"> <li>• Proprietary licenses are those in which the owner reserves and restricts all rights, commonly requiring a fee for usage (licensing fee) <ul style="list-style-type: none"> <li>• The copyright holder specifies acceptable uses and privileges for the work</li> <li>• Generally, the work can only be used for a specific purpose and cannot be modified or distributed</li> </ul> </li> </ul>
GNU General Public License; Apache 2.0; MIT	Not Permitted for data		<ul style="list-style-type: none"> <li>• Written specifically to apply to software</li> </ul>
<b>Recommend a License</b>	<ul style="list-style-type: none"> <li>• If there is a license not found on this list that a project team would like to use and appears to be compatible with the provisions found within the ITS JPO Data Access Guidelines, please reach out to the ITS JPO at data.itsjpo@dot.gov</li> </ul>		

### III. Assigning the License

Once a license is chosen, the project team - **NOT the ITS JPO** - is responsible for assigning the license. The open license should be assigned prior to making the data available to the public or simultaneously with collecting or creating the data.<sup>4</sup> For works obtained through a contract, grant,

or other funding agreement, the ITS JPO should obtain rights to the data either via FAR 52.227-14 or 2 C.F.R. 200.315. The government must then hold the project team accountable for following through on this and other terms after award. CC0 should be included in funding language by default, though an alternative license may be proposed and accepted in negotiation and final award.

When project teams place their data in a data storage system, a corresponding DMP must also be included with a link to where the selected license is officially maintained (ex. CC0) per ITS JPO's Data Access Guidelines. Additionally, the assigned license must be explicitly noted in metadata wherever the data is stored. If a project uses pre-existing data with existing licenses, the ITS JPO requires that project teams note this in the DMP and cite this in the data.

While CC0 does not require users of the data to provide attribution, it does not preclude authors from requesting attribution in copied, modified, or derivative works. The ITS JPO recommends that project teams ask users for attribution in citations of the data to track reuse. This includes requesting a link to the original dataset and reference to the original dataset's Digital Object Identifier is included in citations to the data.

#### IV. Reaping the Benefits of Open Licenses

Assigning an open license to federally-funded data is required by the ITS JPO. Beyond that, project teams benefit greatly from releasing their data to the public. Common benefits include, but are not limited to:

- Work not being 100% dependent on the author or organization that originally developed it. Users can provide continuous feedback on a project, improving final research results.
- Higher return on investment for federally-funded work by contributing to additional ITS research projects.
- Increased transparency.
- Facilitation of enhanced scientific collaboration and reproducibility.

#### V. Additional Resources

Want to learn more about open licenses? See the links to resources below or reach out to the ITS JPO at [data.itsjpo@dot.gov](mailto:data.itsjpo@dot.gov).

- Timothy Vollmer, **Comment from Creative Commons #149 on Source Code Policy GitHub Page**. GitHub, April 11, 2016. <https://github.com/WhiteHouse/source-code-policy/issues/149>.
- **Common License Types for Datasets**. Data.World. <https://help.data.world/hc/en-us/articles/115006114287-Common-license-types-for-datasets>.
- **Comparison of Free and Open Source Software Licenses**. Wikipedia. [https://en.wikipedia.org/wiki/Comparison\\_of\\_free\\_and\\_open-source\\_software\\_licenses/](https://en.wikipedia.org/wiki/Comparison_of_free_and_open-source_software_licenses/).
- **The Copyright Act, 17 U.S.C. § 105**.
- **Introduction to Intellectual Property Rights in Data Management**. Cornell University. <https://data.research.cornell.edu/content/intellectual-property>.
- Some content adapted from **The Legal Side of Open Source**. GitHub [github.com/github/opensource.guide](https://github.com/github/opensource.guide), used under the CC-BY-4.0 license.
- Vivian Deparday & Robert Soden, **Leveraging Open Source as a Public Institution**. World Bank Blogs, June 13, 2017. <https://blogs.worldbank.org/opendata/leveraging-open-source-public-institution-new-analysis-reveals-significant-returns-investment-open>.
- **OPEN Government Data Act, Pub.L. 115–435**.
- **Open Data Policy, M-13-13**.
- **Open Data**. Creative Commons. <https://creativecommons.org/about/program-areas/open-data/>.
- **Open Licenses**. Resources.Data.Gov. <https://resources.data.gov/open-licenses/>.
- **The Open Source Definition**. Open Source Initiative, March 22, 2007. <https://opensource.org/osd>.
- Dr. David A. Wheeler, **Publicly Releasing Open Source Software Developed for the U.S. Government**. Cyber Security & Information Systems Information Analysis Center, March 11, 2016. <https://www.csiac.org/journal-article/publicly-releasing-open-source-software-developed-for-the-u-s-government/>. *Originally published in Software Tech News, Volume: 14, Number 1*.
- **What are Open Standards?** OpenSource.com. <https://opensource.com/resources/what-are-open-standards>.
- **What is Free Software?** GNU Operating System. <https://www.gnu.org/philosophy/free-sw.html>.

#### Footnotes

<sup>1</sup> Generally, when data is arranged in a creative fashion original to the compiler, that compilation is covered by an exclusive copyright held by the compiler. Copyright gives the compiler the exclusive right to make copies of a work and enact restrictions. Note that this does not apply to works created by the U.S. government, which cannot copyright a work. Due to this exclusive copyright dictated by statute, a non-U.S. government entity making data publicly available does not make it open. Therefore, a suitable copyright waiver or license such as CC0 must be assigned to a work in order to give others the permission to use, disclose, reproduce, prepare derivative works, etc. Deciding what data should be included in a compilation, how to organize the data, or how to relate different data elements qualifies as creative decisions eligible for copyright protection. Data collected during research that has not been arranged, organized, or related in a creative manner has no copyright protection under U.S. law. ITS JPO-funded projects should always err on the side of caution and apply an open license, even if they are unsure if the data is copyrightable.

<sup>2</sup> While technically not a “license,” Creative Commons Zero is a worldwide public domain dedication which satisfies the definition of “open license” and serves an identical purpose.

<sup>3</sup> Organizations that sell data, such as DataStock, employ a proprietary license to limit modifications and reuse.

<sup>4</sup> Waiting to assign a license until after the data is made available to the public (even if the project is in its infancy) prevents collaboration, raises legal issues, and is contrary to the Data Access Guidelines.