

What is Open Science and Why Should You Care About It?

By: Pamela Shaw, Biosciences & Bioinformatics Librarian, and Matt Carson, Data Scientist

Open science is the practice of science to meet various goals :

- Transparency in experimental methodology and data collection
- Public access to data and publications
- Enhanced reproducibility and replication of results
- Use of web-based tools to facilitate collaboration

Open science as a concept incorporates the [FAIR principles](#) for research data: research data should be Findable, Accessible, Interoperable and Reusable.

There are many definitions of open science. Probably the most-used definition comes from the European funded project [Facilitate Open Science Training for European Research \(FOSTER\)](#). It states:

Open Science is the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.

Why should you care about open science?

Public access to federally-funded data and manuscripts is a requirement.

From the point of view of the public, the issue of taxpayer access to publications and data are probably the most important elements of open science, in addition to “citizen science” projects such as FoldIt and [SETI@home](#).

Much research in the United States is funded by taxpayer dollars in the form of NIH, NSF, and other federal grants. The US government has a long-standing mandate on public access to NIH-funded manuscripts, but is also developing requirements for public access to government-funded research data. For funder mandates on public access to publications or data, please refer to this [SPARC resource on article and data sharing policies](#).

Additionally, a number of journals ([Nature](#), [Science](#), PLOS, and others) require authors to share data and/or methodological detail as a prerequisite for publication in the journal.

Reproducibility and transparency is good science

From a researcher's perspective, the important concepts in open science are those of transparency and availability of methods, data, and source code for data analysis, so that results from one lab can be reproduced by others.

How can you practice open science?

If you want to – or are required to – share data, code, or detailed methodology, there are a number of resources that can help you practice open science.

- **Publish in a “data journal”** – Data journals do not publish traditional research articles. Instead they provide a platform for publication of descriptions of experimental data and methods and links or instructions for accessing the data. Some options are Data in Brief, GigaScience, and Scientific Data. The University of Edinburgh maintains a list of data journals for a variety of disciplines.
- **Find a data repository** – if you don't know of a discipline-specific repository for your data, consult [re3data](#) to find one. You can also deposit data to [figshare](#) or [Data Dryad](#). Galter Library's [DigitalHub](#) accepts some research data from FSM investigators. [Contact our repository team](#) for more information.
- **Share your code and methods** – There are some great options for sharing code for your experimental data analysis. [GitHub](#) is currently the most popular site for storing code, documents, or any types of files associated with open research projects. You can also contribute your code to open source software projects such as [Bioconductor](#) for the R statistical environment and [Biopython](#) for the Python programming language.
- **Use an open source analysis platform** – There are many options available for investigators to run experimental workflows and share them online. For small-to-medium sized genomic data analysis, [Galaxy online](#) is a great choice. You can upload your sequencing data and create custom tracks in the UCSC Genome Browser. If you are a [Cytoscape](#) user, there is a vibrant online community of developers and packages to try out and share.
- **Use the Open Science Framework** – The [Open Science Framework \(OSF\)](#) is a “scholarly commons” developed by the [Center for Open Science \(COS\)](#) to promote and enable open science. On November 6, 2017, Galter is sponsoring a hands-on workshop, hosted by the COS, to introduce the OSF to Feinberg researchers. If you are interested, please register here: <https://goo.gl/EGWx7T>. Space is limited.

Contact [Matt Carson, Data Scientist](#) at Galter Health Sciences Library and Learning Center for questions related to open science or for a consultation.
