Caltech Workshop Reproducibility Handout (from CodeOcean, Addgene, protocols.io) Version 2

Code Ocean,Addgene,protocols.io

Abstract

This is a handout that was created by Code Ocean, Addgene, and protocols.io for the Caltech workshop on reproducibility. For more information and slides from the workshop, please see: https://codeocean.com/workshop/caltech.

Please feel free to clone and modify it. If you do, would be wonderful to see you share the new resource in this group. Also, please suggest other useful resources.

Citation: Code Ocean,Addgene,protocols.io Caltech Workshop Reproducibility Handout (from CodeOcean, Addgene, protocols.io). **protocols.io**
dx.doi.org/10.17504/protocols.io.nibdcan
Published: 27 Feb 2018

Guidelines

Practical tips for reproducibility

1. Plan for reproducibility before you start

- **Write a study plan or protocol** and track new versions.
- **Set-up a reproducible project** using an electronic lab notebook to organize and track your work. Avoid saving proprietary file formats.

2. Keep track of things

- **Preregister** important study design and analysis information. Free tools to help you make your first registration include AsPredicted, Open Science Framework, and Registered Reports. Clinical trials use Clinicaltrials.gov.
- **Track changes** to your files using version control.
- **Document** everything done by hand in a README file and data dictionary. **Karl Broman's**

Data Organization module: http://kbroman.org/dataorg/pages/dictionary.html

3. Report your research transparently

- **Share your protocols and interventions** explicitly and transparently.
- **Write a transparent report**. Guidelines from the Equator Network or processes like Registered Reports can help.

4. Archive + share your materials

i. Data

- Avoid supplementary files, licence, and share your data using a repository. **How to License Research Data**: http://www.dcc.ac.uk/resources/how-guides/license-research-data.

ii. Materials & reagents

- Licence your published materials so they can be reused. **Creative Commons License Picker**: https://creativecommons.org/choose/
- Deposit reagents with repositories like Addgene, The Bloomington Drosophila Stock Center, and ATCC to make them easily accessible to other researchers.

iii. Software

- Licence your code using Code Ocean or Github. **Open Source Initiative: About Open Source Licences**: https://opensource.org/licenses.

5. Further reading:

- **Ten Simple Rules for Reproducible Computational Research**: http://journals.plos.org/ploscompbiol/article?
- **Reproducibility in Science**: http://ropensci.github.io/reproducibility-guide/
- **Open Science MOOC**: https://opensciencemooc.eu/
- **Tools and Resources for Reproducibility Series at protocols.io:** https://www.protocols.io/groups/tools-and-resources-for-reproducibility

Protocol

Reagents

Step 1.

Addgene https://www.addgene.org/ (nonprofit plasmid repository)

CiteAb https://www.citeab.com/ (antibody search engine with results sorted by citations)

ICLAC http://iclac.org/ (registry of false or misidentified cell lines)

Quartzy https://www.quartzy.com/ (manage lab inventory)

Electronic Lab Notebooks

Step 2.

Benchling https://benchling.com/ (free)

Evernote https://evernote.com/ (most popular with biologists but not designed as an ELN)

Labguru https://www.labguru.com/ ($)

sciNote https://scinote.net/ (open source, free)

Open Science Framework https://osf.io/ (free)

Methods

Step 3.

Bio-Protocol https://bio-protocol.org/ (A peer-reviewed protocol journal; free to read & publish)

protocols.io http://protocols.io/ (an open access repository of science methods; free to read & publish)

Code

Step 4.

Github https://github.com/ (code repository; free for public repos)

Jupyter Notebooks http://jupyter.org/ (open source web-app for creating & sharing live code, equations, and more)

Code Ocean https://codeocean.com/ (computational reproducibility platform; free to upload, share & publish executable code with DOI; pay for more computing time over freemium limit)

Data

Step 5.

DataDryad http://datadryad.org/ (curated digital repository; free to access, $120 to publish dataset up to 20GB)

Figshare http://figshare.com/ (free digital repository, 5GB per file limit)

Zenodo https://zenodo.org/ (free digital repository; 50GB per dataset limit)

ANNOTATIONS

Thomas Morrell 05 Feb 2018

CaltechDATA https://data.caltech.edu is another option for anyone at Caltech (free digital repository, no fixed storage limits)