

Reproducibility for Everyone

Tools and Resources

Society for Neuroscience Global Connectome Social

January 13, 2021

@repro4everyone | www.repro4everyone.org | #repro4everyone

Data Management

- Kbroman Lab** <http://kbroman.org/dataorg/> (short primer on data storage and handling)
- Purdue Library** <http://guides.lib.purdue.edu/c.php?g=353013andp=2378292> (short primer on data management and file naming conventions)
- Data One Best Practices** <https://www.dataone.org/best-practices> (detailed resource on how to handle data throughout its life-cycle)
- Mantra** <https://mantra.edina.ac.uk/> (free online course for those who handle digital data)

Electronic Lab Notebooks (ELN)

- Harvard University ELN guide** <https://tinyurl.com/Harvard-ELN> (great summary of current ELNs)
- SciNote** <https://scinote.net/> (open source, free)
- OSF** <https://osf.io> (Open Science Framework, free)

Code

- Github** <https://github.com/> (code repository; free for public repos)
- Jupyter Notebooks** <http://jupyter.org/> (open source web-app for creating and sharing live code, equations, and more)
- Code Ocean** <https://codeocean.com/> (computational reproducibility platform; free to upload, share and publish executable code with DOI; pay for more computing time over freemium limit)
- Conda and BioConda** <https://conda.io/docs/> and <https://bioconda.github.io/> (operating system independent package environment manager for the command line)
- Docker and Biocontainers** <https://docs.docker.com/> and <http://biocontainers.pro> (container ecosystem to package code and data on the command line)
- Binder** <https://mybinder.org/> (tool to make your GitHub repository an online docker image run in the cloud)
- Galaxy** <https://usegalaxy.org/> (web and graphic interface based bioinformatics platform. Needs local set-up for larger data handling)

Reagents

- Addgene** <https://www.addgene.org/> (nonprofit plasmid repository)
- CiteAb** <https://www.citeab.com/> (antibody search engine with results sorted by citations)
- Quartzy** <https://www.quartzy.com/> (manage lab inventory)
- ICLAC** <https://iclac.org/> (registry of false or misidentified cell lines)

Methods

Bio-Protocol <https://bio-protocol.org/> (peer-reviewed protocol journal; free to read and publish)
protocols.io <http://protocols.io/> (open access repository of science methods; free to read and publish)

Data

DataDryad <http://datadryad.org/> (curated digital repository; free to access, \$120 to publish dataset up to 20GB)
Figshare <https://figshare.com/> (free digital repository, 5GB per file limit)
Zenodo <https://zenodo.org/> (free digital repository; 50GB per dataset limit)

Data Visualization

Beyond Bar Graphs <https://tinyurl.com/ecrbeyondbargraph> (free tools and resources for creating more transparent figures for small datasets)
Interactive Dotplot Tool <http://statistika.mfub.bg.ac.rs/interactive-dotplot/> (create dot plots, box plots, violin plots, show subgroups or display clusters of non-independent data)
Interactive Line Graph Tool <http://statistika.mfub.bg.ac.rs/interactive-linegraph/> (examine different summary statistics, focus on groups, time points or conditions of interest, examine lines for any individual in the dataset, view change scores)
Other free tools https://twitter.com/T_Weissgerber/status/953334933019398145

R
Tutorial - Plotting in R https://www.youtube.com/watch?v=sf_li1XV664
Customized interactive visualizations (Shiny)
<https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01782/full>
Ggplot2 <https://ggplot2.tidyverse.org/>
Claus Wilke's blogpost <http://serialmentor.com/blog/2018/1/23/fundamentals-of-data-visualization>

Python

A collection of useful resources https://github.com/schmelling/python_materials
Data Analysis and Visualization in Python [Data Carpentry: An Introduction to Python for Data Analysis and Visualization - Tracy Teal PyCon 2016 Tutorial](#)
PyData Resources <https://numfocus.org/sponsored-projects> (incl. Matplotlib, Numpy, Pandas, and many more important for data analysis and visualization)

Statistical Analysis

Handbook of Biological Statistics <http://www.biostathandbook.com/> and <http://rcompanion.org/rcompanion/> (webpage by John H. McDonald and others from University of Delaware with pdf download links to free book on stats in Biology and its R implementation)
Scipy stats lectures <https://scipy-lectures.org/packages/statistics/index.html> (lecture on stats in python using scipy) see also <https://www.statsmodels.org/stable/index.html> for more stats in python

Nature Statistics for Biologists resources

<https://www.nature.com/collections/qghhqm/content/practical-guides>

Estimation Statistics <http://www.estimationstats.com/#/>

Literature Related to Reproducibility Resources and Tools

SfN's Resources to Enhance Scientific Rigor

<https://neuroonline.sfn.org/collection/promoting-awareness-and-knowledge-to-enhance-scientific-rigor-in-neuroscience>

Ten Simple Rules for Reproducible Computational Research

<http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003285>

Reproducibility in Science

<http://ropensci.github.io/reproducibility-guide/>

Tips for writing good protocols

<https://www.protocols.io/view/how-to-make-your-protocol-more-reproducible-discov-7uahmse>

Managing Laboratory Notebooks

<http://colinpurrington.com/tips/lab-notebooks>

General File and Folder Organization

<https://zapier.com/blog/organize-files-folders/>

File Naming Conventions

<http://www.exadox.com/en/articles/file-naming-convention-ten-rules-best-practice>

Strain Background and Genetic Drift in Mice

<https://www.jax.org/jax-mice-and-services/customer-support/technical-support/strain-background-and-genetic-drift>

Estimation for Better Inference in Neuroscience

<https://doi.org/10.1523/ENEURO.0205-19.2019>

The New Statistics <https://thenewstatistics.com/itns/blog/>

Open Science Tools for Neuroscience

Open Science MOOC <https://opensciencemooc.eu/>

HHMI Janelia's Open Science Resources <https://www.janelia.org/open-science/overview>

Resources shared by the Stanford Center for Reproducible Neuroscience

<http://reproducibility.stanford.edu/resources/>

The International Brain Laboratory (IBL) "An International Laboratory for Systems and Computational Neuroscience" [https://www.cell.com/neuron/fulltext/S0896-6273\(17\)31136-4](https://www.cell.com/neuron/fulltext/S0896-6273(17)31136-4) and their **Resources** <https://www.internationalbrainlab.com/resources>

Kravitz Lab Open Source Neuroscience <https://kravitzlab.com/open-source-neuroscience>

OpenBehavior <http://openbehavior.com/>

OpenEphys <http://www.open-ephys.org/>

DeepLabCut <http://www.mousemotorlab.org/deeplabcut>

Neurodata Without Borders <https://www.nwb.org/about-nwb/>

Literature Related to Reproducibility in Neuroscience

Analytical Transparency and Reproducibility in Human Neuroimaging Studies

doi.org/10.1523/JNEUROSCI.0424-18.2018

Promoting and Supporting Credibility in Neuroscience

<https://doi.org/10.1177/2398212819844167>

Progress Towards Openness, Transparency and Reproducibility in Cognitive Neuroscience

<https://doi.org/10.1111/nyas.13325>

Reproducibility and Rigour in Computational Neuroscience

<https://www.frontiersin.org/research-topics/5964/reproducibility-and-rigour-in-computational-neuroscience#articles>

Scanning the Horizon: Towards Transparent and Reproducible Neuroimaging Research

<https://doi.org/10.1038/nrn.2016.167>

Video Can Make Behavioral Science More Reproducible

<https://doi.org/10.1038/s41562-017-0128>

Double-dipping Revisited

<https://doi.org/10.1038/s41593-019-0398-z>

A Visionary Resource for Instilling Fundamental Principles of Rigorous Neuroscience Research

<https://meetings.ninds.nih.gov/Home/Agenda/20245>

Further Reading

The Future of Graduate and Postdoctoral Training in the Biosciences

<https://doi.org/10.7554/eLife.32715>

Graduate Biomedical Science Education Needs a New Philosophy

<https://mbio.asm.org/content/8/6/e01539-17>

Train PhD Students to Be Thinkers Not Just Specialists

<https://doi.org/10.1038/d41586-018-01853-1>

Rigorous Science: a How-To Guide

<https://doi.org/10.1128/mBio.01902-16>

How Scientists Fool Themselves and How They Can Stop

<https://doi.org/10.1038/526182a>

A Manifesto for Reproducible Science

<https://www.nature.com/articles/s41562-016-0021>

Example Studies

mcSCRB-seq: sensitive and powerful single-cell RNA sequencing

Paper: <https://doi.org/10.1101/188367>

Protocol: [dx.doi.org/10.17504/protocols.io.p9kdr4w](https://doi.org/10.17504/protocols.io.p9kdr4w)

Code: https://github.com/cziegenhain/Bagnoli_2017

TransRate: reference-free quality assessment of de novo transcriptome assemblies

Paper: <https://dx.doi.org/10.1101/2Fgr.196469.115>

Code: <https://github.com/Blahah/transrate>

Tutorial: <http://hibberdlab.com/transrate/>

Experimenting with Reproducibility: a case study of Robustness in Bioinformatics

Paper: <https://doi.org/10.1093/gigascience/giy077>

Code: <https://github.com/sje30/waverepo>

A Bayesian Mixture Modelling Approach For Spatial Proteomics

Paper: <https://doi.org/10.1101/282269>

Code: <https://github.com/lgatto/2018-tagm-paper>