Tools, tips, & resources compiled by fellows of the MIT Biological Engineering Communication Lab

Getting started

Below is a list of resources compiled for BE Communication Lab fellows relevant to undergraduate, graduate students, and postdocs at MIT. Although there are some MIT-specific resources, this list is also relevant to any STEM student/researcher.

The MIT BE Communication Lab resources are also stored within this repository and are highlighted within the list below. If you have any questions, email us at `mitbecl@gmail.com` .

This list compiles resources to transform data into a clear message through:

- **data analysis and visualization,**
- **figure design,**
- **writing and reference management,**
- **design tools and resources,**
- **professional resources.**

It also includes **resources for reproducibility** and miscellaneous tools for biological engineering research (e.g. plasmid design, flow cytometry analysis, next-generation sequencing).

Legend

Symbol	Meaning
..	no upfront cost
..	open source
..	small cost
..	large cost
..	Computing Package
..	Resource

Computing

- [Programming](#)
- [Computing Clusters at MIT](#)

Programming

Students typical use the follow resources to analyze and plot data for class and research purposes.

- Python – general applicability, open-source; commonly used with Anaconda, a package and environment manager

- R - popular for bioinformatics, genomics, statistics; typically used with RStudio using packages from CRAN

    - RStudio introduction to R - a good place to start for complete beginners.
    - Swirl teaches R within RStudio. A great interactive resource for beginners.

- MATLAB - commercial computing environment offered at MIT for affiliates. See Gnu Octave for an open source alternative.

- Other computing languages/platforms used include Julia and Go, but their user bases are much smaller.

Computing Clusters at MIT

Computing clusters are available at MIT and affiliate institutions for use by students and non-affiliates.

- Athena – computing environment offering remote environments with pre-installed software and file storage

- TIG - CSAIL group offering computing services

- AWS, Google Cloud, Microsoft Azure - commercially available services simple to setup with researcher funds

- C3DDB - Boston-wide resource for life science researchers

- Koch Institute Bioinformatics & Computing Core - offers a variety of cloud computing resources

- McGovern Institute Core - Linux-based cluster offering storage and CPU/GPU performance

- Other institutes (e.g. Broad Institute) and groups offer internal computing resources, inquire directly to gain access

Data Visualization

- Data Visualization Resources
    - Plotting Tools
        - Python Plotting
        - R Plotting

- Other Plotting tools

Data Visualization Resources

- [Trees, Maps, and Theorems: Effective Communication for Rational Minds by Jean-Luc Doumont](#) - The CommLab Bible
- http://serialmentor.com/dataviz/
- https://datavizcatalogue.com/
- http://www.cookbook-r.com/Graphs/
- https://python-graph-gallery.com/
- https://www.data-to-viz.com/
- [Grammar of Graphics](#) - landmark book on foundations in data visualization
- [Plotting One Variable Distributions](#) - BECL-produced resource for plotting bar graphs, boxplots, violin plots, and more. Includes example data set and raw code files.

Plotting Tools

Python Plotting

- [matplotlib](#) – the most popular plotting framework
- Pandas - table management
- bokeh – interactive web-based visualization
- [seaborn](#) – opinionated plotting framework for statistical visualizations
- plotly – interactive web-based visualization
- altair – straightforward visualization framework, biased towards statistical plotting
- [Rpy2](#) - use R code in jupyter notebook
- [BECL notes](#) - BECL-produced resource for python related plotting and getting started.

R Plotting

This is an opinionated summary of key tools for plotting in R, focusing primarily on the [tidyverse](#) group of packages .

- [ggplot2](#) – the most popular plotting framework based on the book, *Grammar of Graphics*
- [plotly](#) – commercially supported interactive web-based visualization tools
- [Shiny](#) – interactive charts and applications on the web, great for displaying public data and generating publication website
- [reticulate](#) - interface with Python via R
- [ggplot2 Cheatsheet sheet](#) - quick overview of ggplot2 plotting functions and aesthetics

- [ggplot2 Tutorial](#) - Harvard tutorial on getting started with ggplot2
- [R Graph Gallery](#) - gallery of plots generated using R
- [R for Data Science](#) - a **comprehensive** resource to become proficient at using R for all data science needs, written by lead instructors at [RStudio](#)

Other Plotting tools

- RAW – fast, easy graphs from Excel or CSV files
- Graphpad Prism – stand-alone plotting program
- Excel – the one and only
- Datawrapper – fast, easy graphs from Excel or CSV files
- Octave – Free version MATLAB
- [WebPlotDigitiazer](#)

Reproducible Analysis

- [General Principles](#)
  - [R workflows](#)
  - [Python workflows](#)

General Principles

- [Naming files and projects](#), a slide deck compiled by Jenny Bryan (@JennyBryan), software engineer at RStudio

R workflows

- [drake](#) – toolkit to build reproducible workflows that scale
- rapport
- knitr - allows to convert markdown, R, and plots/tables to html or PDF files, similar to Jupyter for python
- workflowr
- here - makes it easy for users to set directories and paths
- ROpenSci

Python workflows

- [Crash course in reproducible research in Python](#)

Writing

- [LaTeX](#)
- [Markdown](#)
- [Citations and Reference Management](#)

- [Comparison of different text editors](#)

LaTeX

- [LaTeX primer](#)
- [LaTeX Thesis Proposal Template](#)
- [LaTeX Thesis Defense Template](#)
- [Rticles](#)
- [Wikibooks](#)

Markdown

- [Pandoc](#) - for switching between .doc/.tex/.md/etc file types
- Microsoft Word

Citations and Reference Management

- [Zotero](#)
- [Mendeley](#)
- EndNote
- Papers
- Readcube
- Jabref

Figures

- [Drawing](#)
- [Image Manipulation](#)

Drawing

- Adobe Illustrator
- [Inkscape](#)

- Microsoft Powerpoint ⠄⠄
- [Affinity Designer](#) ⠄⠄
- [BioRender](#) ⠄⠄

Image Manipulation

- Adobe Photoshop ⠄⠄
- [Affinity Designer](#) ⠄⠄
- [GIMP](#) ⠄⠄ ⠄⠄ The GNU Image Manipulation Program
- ImageJ/Fiji ⠄⠄ ⠄⠄

Design Tools & Resources

- [Color](#)
- [Fonts & Typography](#)
- [Icons](#)
- [Images](#)

Color

- [ColorBrewer](#) - web-based color palette tool with accessibility options ([R package](#))
- [Palettable](#) - similar to ColorBrewer with customizable color schemes ([Python package](#))
- [Adobe Color CC](#) - select color schemes based on color wheel and color harmony
- [Ggsci](#) - color themes inspired by scientific journals, science fiction, and media
- [Viz-Palette](#)
- [GenZ Yellow](#) 🗿
- [Millenial Pink](#) 🗿

Fonts & Typography

- [Butterick's Practical Typography](#) - typography best practices
- [Google Fonts](#) - select from fonts based on characteristics
- [Canva](#) - font combinations based on starter font
- [Font Squirrel](#) - downloadable fonts based on characteristics
- [Neue Haas Grotesk](#)

Icons

- [Noun Project](#) - downloadable icons
- [IcoMoon](#) - more icons

Images

- [Unsplash](#) - downloadable high-quality images

Poster Design

- [Poster Design Tools](#)
- [Poster Templates](#)
- [Poster Galleries](#)

Poster Design Tools

- Adobe Illustrator
- [Inkscape](#)
- Microsoft Powerpoint
- Adobe InDesign

Poster Templates

- [Horizontal and Vertical Microsoft PowerPoint and Adobe Illustrator Poster Templates](#) - by Tyler Toth and Alex Triassi. Takeaway: white-space friendly poster templates to get you started

Poster Galleries

- [null](#)

Scientific Software

- [Chemical Structures](#)
- [Analytical Chemistry](#)
- [Protein Structure Visualization](#)
- [Plasmid Editors](#)
- [Flow Cytometry](#)
- [Microscopy Analysis](#)

Chemical Structures

- [ChemDraw](#)

Analytical Chemistry

- [Mnova](#)

Protein Structure Visualization

- [Chimera](#)
- [Pymol](#)

Plasmid Editors

- [ApE](#)
- [Benchling](#) - a web-based plasmid editor and database
- [Geneious](#)
- [SnapGene](#)

Flow Cytometry

- [Flowjo](#)
- [Cytoflow](#)

Microscopy Analysis

- [CellProfiler](#)
- [ImageJ](#)

Unsorted Weblinks

- BLAST
- TCoffee
- [MPI Bioinformatics Toolkit](#)
- [HMMER](#)
- ClustalOmega
- GSEA
- Weblogo
- Pfam
- Espript

- Primer3
- Ensembl

Professional Resources

- [Github Personal Webpage Boilerplate](#)
- PhD/Post-Doctoral website examples
  - https://jef.works/ (Harvard, Bioinformatics)
  - https://www.nikhitasingh.com/ (MIT Media Lab, AI)
  - https://davidlazar.org/ (MIT CSAIL, Computing)
  - https://www.anishathalye.com/ (MIT CSAIL, Computing)
  - https://slowkow.com/ (Harvard, Immunogenomics)

Miscellaneous & Unsorted

License