

# Kelly Sovacool

*Bioinformatician seeking to build open source software for reproducible data science*

## Skills

Languages & Tools	R, Python, C++, Bash, Snakemake, git, GitHub, R Markdown, Jupyter, Quarto, LaTeX, conda/mamba, Docker, Singularity, SLURM, Unix CLI
Research	supervised machine learning pipelines, data visualization, reproducible reports & manuscripts
Software	package maintenance, test-driven development, continuous integration, documentation, collaboration & peer review, high performance computing

## Education

2018 - 2023 **PhD Bioinformatics**, *University of Michigan*, Advisor: Patrick D. Schloss

2014 - 2018 **BS Biology**, *University of Kentucky*, Minor: Computer Science

## Experience

- 2019 - present **Graduate Student Researcher**, [Schloss Lab](#), University of Michigan
- Develop, benchmark, and maintain bioinformatics workflows and software packages in R & Python.
  - Build machine learning pipelines for human gut microbiome classification and prediction problems in colorectal cancer and *C. difficile* infection. Publish findings as reproducible papers.
  - Collaborate with other scientists on microbiome projects and mentor junior lab members.
- Jan - Apr 2023 **Graduate Student Instructor**, *Dept. of Computational Medicine & Bioinformatics*, University of Michigan, BIOINF 576: Tool Development for Bioinformatics
- Develop curriculum and teach students the principles of software development in R & Python. Topics: software design, implementation, testing, documentation, issue tracking, peer review, and release.
- 2019 - 2022 **Executive Committee Member: Open Source Technical Lead**, [Girls Who Code at U-M](#)
- Facilitated collaborative design, development, and maintenance of our curriculum on Python for data science to teach young women+ in Girls Who Code clubs.
- 2015 - 2018 **Undergraduate Lab Assistant**, *Moseley Bioinformatics Lab*, University of Kentucky
- Developed a Python package to identify homologous gene products for comparative genomics.

## Open Source Projects

- Maintainer [mikropml](#). User-Friendly R Package for Supervised Machine Learning Pipelines
- Maintainer [mikropml snakemake workflow](#). Snakemake template for building reusable and scalable machine learning pipelines with mikropml (Python + R)
- Maintainer [schtools](#). Schloss Lab tools for reproducible microbiome research (R)
- Contributor [mothur](#). Command-line tool for processing microbial amplicon sequence data (C++)
- Co-author [mothur snakemake workflow](#). Template for reproducible mothur workflows (Python + R)
- Co-author [Python for Data Science curriculum](#) via live-coding for Girls Who Code clubs
- Co-author [Workshop curriculum](#) on reproducible research integrating R, the Unix shell, & git/GitHub

## Continuing Education

- Jan 2020 **Building Tidy Tools workshop at rstudio::conf**. RStudio, PBC. San Francisco, CA
- Dec 2019 **Winter School in Research Software Engineering**. URSSI. Seattle, WA
- Jun 2019 **Software Carpentry Instructor Training**. The Carpentries. University of Michigan

## Publications

 see list here: [sovacool.dev/pubs](https://sovacool.dev/pubs)