

# 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, gut microbiome composition, amplicon sequence analysis, metagenomics, data visualization, reproducible 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.
- Build machine learning pipelines for human gut microbiome classification and prediction problems in colorectal cancer and *C. difficile* infection.
- 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 Dept. of Computational Medicine & Bioinformatics*

- Developed curriculum and taught Python for data science to young women+.
- Facilitated collaborative development and maintenance of our open source teaching resources.

2015-2018 **Undergraduate Lab Assistant**, *Moseley Bioinformatics Lab*, University of Kentucky

Developed a Python package for identifying homologous gene products for comparative genomics.

## Open Source Contributions

Maintainer	<b>mikropml</b> . User-Friendly R Package for Supervised Machine Learning Pipelines
Maintainer	<b>schttools</b> . Schloss Lab tools for reproducible microbiome research (R package)
Maintainer	<b>mikropml snakemake workflow</b> . Template for running mikropml with Snakemake (Python+R)
Contributor	<b>mothur</b> . Command-line tool for processing amplicon sequence data (C++)
Co-author	<b>mothur snakemake workflow</b> . Snakemake template for microbial amplicon sequence analysis with mothur (Python+R)

## Continuing Education

Jan 2020 **Building Tidy Tools workshop at rstudio::conf**. Rstudio, PBC. San Francisco, CA

Dec 2019 **Winter School in Research Software Engineering**. US Research Software Sustainability Institute. Seattle, WA

Jun 2019 **Software Carpentry Instructor Training**. The Carpentries. University of Michigan