

# 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	<b>PhD Bioinformatics</b> , <i>University of Michigan</i> , Advisor: Patrick D. Schloss
2014-2018	<b>BS Biology</b> , <i>University of Kentucky</i> , Minor: Computer Science

## Experience

2019-present	<b>Graduate Student Researcher</b> , <a href="#">Schloss Lab</a> , University of Michigan <ul style="list-style-type: none"><li>○ Develop, benchmark, and maintain bioinformatics workflows and software packages in R &amp; Python.</li><li>○ Build machine learning pipelines for human gut microbiome classification and prediction problems in colorectal cancer and <i>C. difficile</i> infection.</li><li>○ Collaborate with other scientists on microbiome projects and mentor junior lab members.</li></ul>
Jan-Apr 2023	<b>Graduate Student Instructor</b> , <i>Dept. of Computational Medicine &amp; Bioinformatics</i> , University of Michigan, BIOINF 576: Tool Development for Bioinformatics <p>Develop curriculum and teach students the principles of software development in R &amp; Python. Topics: software design, implementation, testing, documentation, issue tracking, peer review, and release.</p>
2019-2022	<b>Executive Committee Member: Open Source Technical Lead</b> , <a href="#">Girls Who Code at U-M</a> <p>Facilitated collaborative design, development, and maintenance of our curriculum on Python for data science to teach young women+ in Girls Who Code clubs.</p>
2015-2018	<b>Undergraduate Lab Assistant</b> , <i>Moseley Bioinformatics Lab</i> , University of Kentucky <p>Developed a Python package to identify homologous gene products for comparative genomics.</p>

## Open Source Contributions

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

## Continuing Education

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

## Publications

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