

Kelly Sovacool

PhD Candidate

✉ sovacool@umich.edu
🌐 sovacool.dev
in [kelly-sovacool](#)
🐙 [kelly-sovacool](#)

Education

- 2018-present **PhD Bioinformatics**, *Dept. of Computational Medicine and Bioinformatics*, University of Michigan.
○ Advisor: Patrick D. Schloss
- 2014-2018 **BS Biology**, *Dept. of Biology*, University of Kentucky.
○ Minor: Computer Science

Skills

- Languages & Tools R, Python, Bash, C++, Snakemake, R Markdown, Quarto, LaTeX, conda, git, GitHub, SLURM.
- Research gut microbiome composition, amplicon sequence analysis, metagenomics, supervised machine learning, data visualization, reproducible manuscripts.
- Software R package maintenance, continuous integration, high performance computing.

Research Experience

- 2019-present **Graduate Student Researcher**, *Schloss Lab, Dept. of Microbiology and Immunology*, University of Michigan.
○ Develop and benchmark bioinformatics pipelines and software for microbial ecology.
○ Analyze 16S rRNA gene amplicon sequence data.
○ Apply machine learning methods to gut microbiome classification problems in colorectal cancer and *C. difficile* infection.
○ Collaborate with other scientists on microbiome projects and mentor junior lab members.
- 2018-2019 **Rotation Student Researcher**, *Program in Biomedical Sciences*, University of Michigan.
○
- 2015-2018 **Undergraduate Lab Assistant**, *Moseley Bioinformatics Lab, Dept. of Molecular and Cellular Biochemistry*, University of Kentucky.
○ Developed a computational tool in Python for identifying sets of orthologous and paralogous gene products in whole genomes to facilitate collinearity analysis and detection of gene duplication events.

- 2016-2018 **BIO395 Independent Research Student**, *Weisrock Lab, Dept. of Biology*, University of Kentucky.
- Developed bash scripts and a SNP calling pipeline in Snakemake.
 - Population structure analysis of the *Ambystoma tigrinum* species complex.
 - Bayesian species delimitation of the *Desmognathus fuscus* species complex.
- 2015-2016 **Undergraduate Lab Assistant**, *Jaromczyk Lab, Dept. of Computer Science*, University of Kentucky.
- Maintained the *Epichloë festucae* genome project database.
 - Analyzed RNA-seq data of *Chenopodium quinoa* and coffee ringspot virus.

Teaching Experience

- Jan-Apr 2023 **Graduate Student Instructor**, *Dept. of Computational Medicine & Bioinformatics*, University of Michigan.
- BIOINF 576: Tool Development for Bioinformatics
- 2019-2022 **Facilitator & Capstone Project Mentor**, *Girls Who Code at U-M DCMB*, University of Michigan.
- Weekly Club during the school year and week-long Summer Experience for high schoolers to learn Python for data science
- 2018-present **Workshop Instructor & Helper**, *U-M Carpentries*, University of Michigan.
- 2-day Software Carpentry workshops teaching computational skills for reproducible research
- Jun 2022 **Instructor**, Virtual.
- Intro to R & RNA-Seq Workshop for ASM Microbe conference attendees
- Apr 2019 **DNA Day Ambassador**, *Michigan DNA Day*, Pioneer High School, Ann Arbor, MI.
- Epigenetics & Scientific Journeys
- Mar 2019 **Workshop helper**, *Graduate Society of Black Engineers and Scientists*, University of Michigan.
- Data Visualization with Python Workshop
- Mar 2019 **Capstone Activity Leader**, *Females Excelling More in Math, Engineering, & the Sciences*, University of Michigan.
- Binary Numbers through Ozobots with GWC at U-M DCMB
- 2012-2018 **Tutor**, *freelance*.
- for high school and college students in Biology, Calculus, Chemistry, Computer Science, and Bioinformatics.

Service

- 2019-present **Executive Committee Member: Open Source Technical Lead, *Girls Who Code at U-M DCMB*.**
- Plan, apply for funding, develop curriculum, and maintain resources to teach introductory Python programming & data science to young women+.
 - Facilitate collaborative development and maintenance of our open source teaching resources.
 - Organize our year-round Club and annual Data Science Summer Experience for high school women+.
- 2019-present **Organizer & Maintainer, *U-M Carpentries*.**
- Co-lead development & maintenance of a curriculum for workshops teaching programming skills for reproducible research.
 - Maintain the website, develop curriculum, and organize workshops.
 - Collaborate with U-M Women in Science and Engineering to organize workshops for women+.
- 2021-present **Mentor, *Schloss Lab*.**
- Mentor an undergraduate student in building reproducible machine learning models to predict `C. difficile` infection severity from gut microbiome composition.
- 2021-present **CoderSpaces co-host, *U-M ISR Data Science Hub*.**
- Hold office hours at a weekly virtual help session for data science practitioners
- 2021-present **Peer reviewer.**
- PLOS ONE (1)
- 2019-2021 **Graduate Student Coordinator, *U-M Data Analysis Networking Group*.**
- Organize monthly meetings & a one-day symposium for researchers to sharpen their data analysis skills.
 - Apply for funding through a Rackham Interdisciplinary Workshop grant.
- 2009-present **Volunteer Sound Engineer, *various churches and non-profit organizations*.**
-

Open Source Contributions

Software

mikropml, *User-Friendly R Package for Supervised Machine Learning Pipelines*.

- Co-author and maintainer

schtools, *Schloss Lab tools for reproducible microbiome research (R package)*.

- Co-author and maintainer

mikropml snakemake workflow, *Template for running mikropml with Snakemake*.

- Co-author and maintainer

mothur, *Command-Line Tool for Processing 16S rRNA Gene Sequence Data*.

○ Contributor

mothur snakemake workflow, *Snakemake template for amplicon sequence analysis with mothur*.

○ Co-author

Curricula

U-M Software Carpentry, *Intro to R- the Unix shell- and git for workshops on reproducible research..*

○ Co-author and maintainer

Girls Who Code at U-M DCMB, *Intro to Python for Data Science for Girls Who Code clubs..*

○ Co-author and maintainer

Code Clubs, *Short coding tutorials for lab meetings*.

○ Contributor

U-M DANG!, *repro-packs: Organizing projects for reproducibility and headache prevention..*

○ Author

Intro to R & RNA-seq, *Workshop for 2022 ASM Microbe attendees*.

○ Contributor

Presentations

Talks

Nov 2022 **Bioinformatics Student Research Hour**, *Predicting *C. difficile* infection severity from the taxonomic composition of the gut microbiome*, University of Michigan.

Feb 2022 **Seminar for the KG Jebsen Center for Genetic Epidemiology**, *Intro to git & GitHub*, (Virtual) Norwegian University of Science and Technology.

Mar 2021 **Bioinformatics Student Research Hour**, *OptiFit: a fast method for fitting amplicon sequences to existing OTUs*, (Virtual) University of Michigan.

Apr 2018 **Systems Biology and Omics Integration Seminar**, *Developing a Global Homology Analysis for Comparative Genomics*, University of Kentucky.

Posters

- Jun 2022 **ASM Microbe**, *Predicting the severity of *C. difficile* infections from the taxonomic composition of the gut microbiome*, Washington, DC.
- Jun 2020 **ASM Microbe**, *OptiFit: a fast method for fitting amplicon sequences to existing OTUs*, Virtual.
- Apr 2018 **Showcase for Undergraduate Scholars**, *Developing a Global Homology Analysis for Comparative Genomics*, University of Kentucky.
- Apr 2018 **National Conference on Undergraduate Research**, *Developing a Global Homology Analysis for Comparative Genomics*, University of Central Oklahoma.
- Apr 2016 **Showcase for Undergraduate Scholars**, *Processing RNA-seq Reads of Plants Infected with the Coffee Ringspot Virus*, University of Kentucky.
- Apr 2016 **UT-KBRIN Bioinformatics Summit**, *Processing RNA-seq Reads of Plants Infected with the Coffee Ringspot Virus*, Cadiz, KY.
- Apr 2015 **Showcase for Undergraduate Scholars**, *The Effect of Meditation on Performance*, University of Kentucky.

Awards

Grants and Fellowships

- 2022 **Conference Travel Grant**, \$900, Rackham Graduate School
 University of Michigan.
- 2020-2021 **Rackham Interdisciplinary Workshop Grant**, \$500, Rackham Graduate School
 University of Michigan.
- 2020 **Conference Travel Grant**, \$800, Rackham Graduate School
 University of Michigan.
- 2019-2020 **Rackham Interdisciplinary Workshop Grant**, \$500, Rackham Graduate School
 University of Michigan.
- 2019-2021 **NIH T32 Bioinformatics Training Program Fellow**, Bioinformatics Graduate Program
 University of Michigan.

- 2017 **Oswald Research & Creativity Award**
 2nd place in Biological Sciences, \$200, Office of Undergraduate Research
 University of Kentucky.
- 2017 **Summer Research Grant**, \$2,000, Office of Undergraduate Research
 University of Kentucky.
- 2014-2018 **Presidential Scholarship**, *out-of-state full tuition*, University of Kentucky.
- Honors
- 2018 **Graduated Cum Laude with Departmental Honors in Biology**, University of Kentucky.
- 2018 **Biology Undergraduate Research Award Nominee**, University of Kentucky.
- 2014-2018 **Lewis Honors College**, University of Kentucky.

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.
- May 2019 **PyCon Education Summit & Conference**, *The Python Software Foundation*, Cleveland, OH.
- Apr 2019 **MICROBIOL 612.2**, *Riffomonas Minimal R Workshop*, University of Michigan.

Publications

* Indicates co-first author

1. Goodin, M. M., Farman, M., Inocencio, H., Jang, C., Jaromczyk, J. W., Moore, N., & Sovacool, K. L. (2016, August). Processing RNA-Seq data of plants infected with coffee ringspot virus. *Proceedings of the 15th Annual UT-KBRIN Bioinformatics Summit 2016: Cadiz, KY, USA. 8-10 April 2016*. <https://doi.org/10.1186/s12859-016-1154-y>
2. Hagan, A. K., Lesniak, N. A., Balunas, M. J., Bishop, L., Close, W. L., Doherty, M. D., Elmore, A. G., Flynn, K. J., Hannigan, G. D., Koumpouras, C. C., Junior, M. L., Kozik, A. J., McBride, K., Rifkin, S. B., Stough, J. M. A., Sovacool, K. L., Sze, M. A., Tomkovich, S., Topcuoglu, B. D., & Schloss, P. D. (2020). Ten simple rules to increase computational skills among biologists with Code Clubs. *PLoS Comput Biol*, 16(8), e1008119. <https://doi.org/10.1371/journal.pcbi.1008119>
3. Everson, K. M., Gray, L. N., Jones, A. G., Lawrence, N. M., Foley, M. E., Sovacool, K. L., Kratovil, J. D., Hotaling, S., Hime, P. M., Storfer, A., Parra-Olea, G., Percino-Daniel, R., Aguilar-Miguel, X., O'Neill, E. M., Zambrano, L., Shaffer, H. B., & Weisrock, D. W. (2021). Geography is more important than life history in the recent diversification of the tiger salamander complex. *PNAS*, 118(17). <https://doi.org/10.1073/pnas.2014719118>
4. *Topçuolu, B. D., *Lapp, Z., *Sovacool, K. L., Snitkin, E., Wiens, J., & Schloss, P. D. (2021). Mikropml: User-Friendly R Package for Supervised Machine Learning Pipelines. *JOSS*, 6(61), 3073. <https://doi.org/10.21105/joss.03073>
5. *Duda, M., *Sovacool, K. L., Farzaneh, N., Nguyen, V. K., Haynes, S. E., Falk, H., Furman, K. L., Walker, L. A., Diao, R., Oneka, M., Drotos, A. C., Woloshin, A., Dotson, G. A., Kriebel, A., Meng, L., Thiede, S. N., Lapp, Z., & Wolford, B. N. (2021). Teaching Python for Data Science: Collaborative development of a modular & interactive curriculum. *JOSE*, 4(46), 138. <https://doi.org/10.21105/jose.00138>
6. *Lapp, Z., *Sovacool, K. L., Lesniak, N., King, D., Barnier, C., Flickinger, M., Krüger, J., Armour, C. R., Lapp, M. M., Tallant, J., Diao, R., Oneka, M., Tomkovich, S., Anderson, J. M., Lucas, S. K., & Schloss, P. D. (2022). Developing and deploying an integrated workshop curriculum teaching computational skills for reproducible research. *JOSE*. <https://doi.org/10.21105/jose.00144>
7. Sovacool, K. L., Westcott, S. L., Mumphrey, M. B., Dotson, G. A., & Schloss, P. D. (2022). OptiFit: An Improved Method for Fitting Amplicon Sequences to Existing OTUs. *mSphere*. <https://doi.org/10.1128/msphere.00916-21> 1752 N St., N.W., Washington, DC
8. Barron, M. R., Sovacool, K. L., Abernathy-Close, L., Vendrov, K. C., Standke, A. K., Bergin, I. L., Schloss, P. D., & Young, V. B. (2022). Intestinal Inflammation Reversibly Alters the Microbiota to Drive Susceptibility to *Clostridioides difficile* Colonization in a Mouse Model of Colitis. *mBio*, 0(0), e01904-22. <https://doi.org/10.1128/mbio.01904-22>
9. Armour, C. R., Sovacool, K. L., Close, W. L., Topçuolu, B. D., Wiens, J., & Schloss, P. D. (2022). Streamlined implementation of a machine learning model to classify screen relevant neoplasia using reference-based OTU clustering. *bioRxiv*. <https://doi.org/10.1101/2022.09.01.506299>