

Erin H. Wilson

🏠 erinhwilson.com
✉ erinhwilson@gmail.com
☎ 415 - 233 - 0835
🌐 [erinhwilson](https://www.linkedin.com/in/erinhwilson) in [LinkedIn](https://www.linkedin.com/in/erinhwilson)

Data Scientist with 11 years of computational analysis experience across industry and academia ▪ Quantitative data generalist for life sciences with programming skills in machine learning (model design, training, evaluation), genomics, and interactive data visualization ▪ Creative scientist who can connect ideas, navigate relationships, and find solutions across software, science, and commercial teams ▪ Effective communicator who can adapt explanations of technical topics to engage audiences with diverse backgrounds ▪ **Dedicated to a career cultivating a more sustainable world**

Keywords: *data science and visualization, microbiology, predictive modeling, industrial-scale biotech*

Tools: *Python (PyTorch, pandas, sci-kit learn, numpy, Altair, seaborn), AWS, Git, Jira, Tableau, SQL, LucidChart*

EDUCATION

- | | |
|--|-------------|
| Ph.D. & M.S., Computer Science, University of Washington, Seattle, WA | 2023 |
| ▪ NSF GRFP Fellow; Computational Molecular Biology certificate | |
| B.A. Computational Biology, Carleton College, Northfield, MN | 2014 |
| ▪ <i>Magna Cum Laude</i> ; awarded Distinction on senior thesis | |

WORK & RESEARCH EXPERIENCE ([INDUSTRY](#) || [ACADEMIA](#))

- | | |
|--|----------------------|
| LanzaTech, Data Scientist /AI & Computational Biology/ | 2024 - 2025 |
| <ul style="list-style-type: none">▪ Data analysis & visualization: performed root cause investigations of biological diversity and fermentation process data to troubleshoot operation of commercial-scale carbon recycling technology; created custom interactive visualizations to support scientists' data exploration; founded "Data Sketches" community for peer learning and teaching effective data viz techniques▪ Project management: led effort to document detailed flow of diverse data streams across 6 commercial facilities (measurement equipment, quantitative transformations, human handoffs, database storage) and presented improvement plan to leadership; coordinated weekly meetings for computing teams (managed schedule, recruited guest speakers, balanced technical deep dives with high-level updates); managed biological data visualization hub to process and display genomics data (accessed 500 times/month)▪ Biological modeling: monitored microbe communities in commercial-scale bioreactors; analyzed protein binding data (DAP-seq) across genomes to identify functional gene groups and DNA sequence patterns in carbon emissions-consuming bacteria; extracted importance scores from ML models that predict enzyme/molecule binding affinity; proposed mutation changes to swap enzymes' preferred molecular partner | |
| University of Washington, Graduate Researcher /Computer Science/ | 2017 - 2023 |
| <ul style="list-style-type: none">▪ Deep learning for methane mitigation: trained and evaluated neural network models on DNA sequences to predict gene activity in diverse conditions for <i>M. buryatense</i>, a promising microbial engineering platform for methane removal▪ Bioengineering tools: Built computational framework to predict DNA elements that influence gene activity; collaborated with wet lab scientists to design experimental validation of computational predictions | |
| DTU Biosustainability Institute, Visiting Researcher /Quantitative Modeling/ | Spring 2022 |
| <ul style="list-style-type: none">▪ Used machine learning methods (ICA) to identify functional gene modules (iModulons) in <i>M. buryatense</i> | |
| Zymergen, Intern /Data Science/ | Summer 2018 |
| <ul style="list-style-type: none">▪ Prototyped deep learning models for predicting gene regulatory regions in non-model microbe genomes | |
| Amyris, Associate Scientist /Scientific Computing/ | 2014 - 2017 |
| <ul style="list-style-type: none">▪ High throughput strain engineering: Developed software pipeline and database schema to generate and store build instructions for DNA constructs (hundreds per week) designed by the "Automated Scientist" (DARPA project to produce 450 molecules); tracked DNA parts' status through Amyris' high-throughput strain construction pipeline▪ Genotype Specification Language (GSL): Implemented software features and trained biologists/external collaborators to use open-source DNA design language invented at Amyris. (co-authored: article, textbook chapter, blog post, poster)▪ Communication: Facilitated technical meetings between biologists and software engineers; held weekly office hours for 1x1 genomics and GSL support; engaged with attendees at Bay Area science nights about Amyris' sustainable technology | |
| Amyris, Intern /Scientific Computing/ | Winter 2013 |
| <ul style="list-style-type: none">▪ Implemented data visualization tool to overlay experimental data on yeast metabolic pathway | |
| University of Minnesota, Research Assistant /Computational Biology/ | Summer 2013 |
| <ul style="list-style-type: none">▪ Analyzed chemical-genetic interaction data and coded pipeline to predict gene targets for drug candidates | |
| University of Queensland, Research Abroad /Coastal Marine Ecology/ | "Winter" 2013 |
| <ul style="list-style-type: none">▪ Conducted observational field research in various Australian ecosystems (reef island, rainforest, desert gorge, intertidal zones) | |
| Carleton College, Research Assistant /Evolutionary Computing/ | Summer 2012 |
| <ul style="list-style-type: none">▪ Experimented with populations of mutating digital organisms to examine evolutionary dynamics | |
| University of California, San Francisco, Research Assistant /Genetics/ | Summer 2011 |
| <ul style="list-style-type: none">▪ Performed ChIP-seq experiments on embryonic mouse tissue to find DNA elements involved in limb development | |

SCIENCE COMMUNICATION

Public Outreach & Tutorials

- “How can micro-organisms help us solve macro-problems?” (2023) [Lightning Talk](#) at PNW Climate week (general public)
- “Modeling DNA Sequences with PyTorch.” (2022) [Article](#) in *Towards Data Science* (ML technical tutorial)
- “The Light Side of Genetic Engineering.” (2019) [Article](#) in *OneZero* (general public)
- “Genetic Constructor and GSL - The Best of Both worlds.” (2016) [Article](#) with Autodesk Bionano Research (biotech blog)

Selected Publications & Presentations

- **E. H. Wilson et al** (2024) “Expanding genetic toolkits for acetogens: DNA-affinity purification sequencing (DAP-seq) reveals transcription factor binding trends across 6 Clostridia genomes.” *Poster at JGI annual user meeting*. Walnut Creek, CA
- **E. H. Wilson** (2024) “Limits and opportunities for deep learning approaches in C1 metabolism” *Talk at The Automated Scientist conference*. Hillerød, Denmark
- L. He, J. D. Groom, **E. H. Wilson**, J. Fernandez, M. C. Konopka, D. A. C. Beck, M. E. Lidstrom. (2023) “A methanotrophic bacterium to enable methane removal for climate mitigation.” *PNAS* ([link](#))
- **E. H. Wilson**., M. E. Lidstrom, D. A. C. Beck. (2023) “Probing the limits of deep learning methods for predicting gene expression in non-model microbes.” *Rapid talk and poster at SBFC conference*. Portland, OR
- A. H. Singh et al (2023) “An Automated Scientist to Design and Optimize Microbial Strains for the Industrial Production of Small Molecules.” *bioRxiv*. ([link](#))
- **E. H. Wilson**, M. E. Lidstrom, and D. A. C. Beck. (2021) “A multi-task learning approach to enhance sustainable biomolecule production in engineered microorganisms.” *Tackling Climate Change with Machine Learning, ICML workshop* ([link](#), [recording](#))
- **E. H. Wilson** et al. (2021) “A Computational Framework for Identifying Promoter Sequences in Nonmodel Organisms Using RNA-seq Data Sets.” *ACS Synthetic Biology*. ([link](#), [project page](#))
- **E. H. Wilson** (2020) “Using microorganisms to solve macro problems: untangling the genetic circuitry of methane-eating bacteria.” Invited talks at MIDAS Data Science Symposium, University of Michigan and Virtual Women’s Research Day, University of Washington ([recording](#)).
- **E. H. Wilson**, C. Macklin, and D. Platt. (2018) “Engineering genomes with Genotype Specification Language.” In *Methods in Molecular Biology, Synthetic Biology*. Springer Publishing Company, New York, NY. In Press. ([link](#))
- **E. Wilson**, S. Sagawa, J. Weis, M. Shubert, M. Bissell, B. Hawthorne, C. Reeves, J. Dean, and D. Platt. (2016) “Genotype Specification Language.” *ACS Synthetic Biology*. ([link](#))

AWARDS & FELLOWSHIPS

Scan Design Foundation Fellowship	2022
▪ Support for research and cultural exchange between Danish and American students	
NSF Graduate Research Fellow	2019
▪ Research funding from the National Science Foundation	
Marilyn Fries Fellowship	2017
▪ Awarded first year graduate research funding in Computer Science	
Clare Boothe Luce Scholarship for Women in Physics and Computer Science	2012
▪ Received funding for summer research in Evolutionary Computing	

LEADERSHIP, VOLUNTEERING, & ACTIVITIES

- “Circles & Lines.” (2024) *Minute Motif* (1 slide, 1 minute) talk at PNW Climate Week ([re-recorded](#) for posterity)
- Technical mentor for intro Python students, [Paper Airplanes](#), Women in Tech program (2023 - Present)
- Research mentor for an undergraduate Computer Science student (2020-2023)
- Recreational data visualization:
 - “Do European M&Ms Actually Taste Better than American M&Ms?” (2024) [Article](#) in *Towards Data Science*.
 - “Mistborn: The Final Eyebrow.” (2021) [Article](#) in *Towards Data Science*.
- Youth Outreach at “UW Engineering Discovery Days” and “Introduce a Girl to Coding, Robotics, and Data Science”
 - Developed interactive activity “*Programming Organisms with DNA Puzzles!*” to teach elementary/middle schoolers about metabolic engineering
- [MeadoWatch](#) field data collector, UW Biology citizen science project (2019 - present)
 - Collect wildflower blooming data in Mount Rainier National Park
- Wildlife rescue hospital volunteer, [PAWS](#) (Seattle, WA) and [Wildcare](#) (San Rafael, CA) (2015 - 2019)
 - Treated and cared for injured songbirds in hospital; Co-led youth nature hikes with Education Department
- Carleton Varsity Women’s Soccer (Div III)