

Erin H. Wilson

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

Data Scientist with 11+ years of computing and analysis experience across industry and academia ▪ Creative analyst who can connect ideas across software, biology, and sustainable technology ▪ Adept communicator (written, data visualization) who can adapt explanations of complex topics to diverse audiences ▪ **Dedicated to a career where I can drive solutions in climate health and sustainability**

Keywords: data science, microbiology, science communication, predictive modeling, interactive visualization, omics data
Tools: Python (PyTorch, pandas, sci-kit learn, numpy, Altair, seaborn), AWS, Git, Jira, SQL

Education

Ph.D. Computer Science, University of Washington, Seattle, WA	2023
▪ NSF GRFP Fellow; Computational Molecular Biology certificate	
M.S. Computer Science, University of Washington, Seattle, WA	2019
B.A. Computational Biology, Carleton College, Northfield, MN	2014
▪ <i>Magna Cum Laude</i> ; awarded Distinction on senior thesis	

Work & Research Experience [Industry || Academia]

Data Scientist, LanzaTech	Feb 2024 - Present
<ul style="list-style-type: none">▪ Technical analysis: deep investigation of diverse data streams (genomics, fermentation variables) to support genome engineering and commercial operations; create interactive visualizations to support data exploration and understanding▪ Project Management: diagram detailed information flow between commercial measurement equipment and databases (data collection, transformation, handoffs); manage in-house computational biology platform to extract, process, and visualize key data (commercial bioreactors, sequencing results, metagenomics); lead internal community for learning and sharing data visualization techniques	
Graduate Researcher, University of Washington <i>/Lidstrom, Beck labs/</i>	2017 - 2023
<ul style="list-style-type: none">▪ Built computational frameworks to accelerate metabolic engineering efforts in methanotrophic bacteria, a promising carbon removal platform; established a suite of new synthetic promoter tools for constitutive strong expression▪ Characterized the effectiveness of deep learning models for discovering gene regulatory elements from RNA-seq experiments in microorganisms with limited data	
Visiting Researcher, DTU Biosustainability Institute, Denmark <i>/Nielsen lab/</i>	Apr-Aug 2022
<ul style="list-style-type: none">▪ Used machine learning methods (ICA) to identify independent gene modules in methane-consuming bacteria	
Data Science Intern, Zymergen	Summer 2018
<ul style="list-style-type: none">▪ Prototyped deep learning models for predicting regulatory DNA features in microbe genomes	
Associate Scientist, Scientific Computing, Amyris	2014 - 2017
<ul style="list-style-type: none">▪ 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)▪ Bioinformatics & Databases: Developed software pipeline and database schema to translate metabolic pathway designs into DNA build instructions for DARPA-funded project to produce sustainable materials ("Automated Scientist"); helped maintain bioinformatics pipelines and supported scientists in interpreting mutation and coverage data.▪ Communication: Facilitated technical communication between biologists and software engineers; held weekly office hours for 1x1 bioinformatics and GSL support; engaged with attendees at Bay Area science nights about Amyris' sustainable technology	
Intern, Scientific Computing, Amyris	December 2013
<ul style="list-style-type: none">▪ Implemented data visualization tool to overlay experimental data on yeast metabolic pathway	
Research Assistant, Computational Biology, University of Minnesota <i>/Myers lab/</i>	Summer 2013
<ul style="list-style-type: none">▪ Analyzed chemical-genetic interaction data and coded pipeline to predict gene targets for drug candidates	
Research Abroad, Coastal Marine Ecology, University of Queensland	"Winter" 2013
<ul style="list-style-type: none">▪ Conducted observational field research in various Australian ecosystems (reef island, rainforest, desert gorge, intertidal zones)	
Research Assistant, Evolutionary Computing, Carleton College <i>/Goings lab/</i>	Summer 2012
<ul style="list-style-type: none">▪ Experimented with populations of mutating digital organisms to examine the evolution of complex logic	
Research Assistant, Genetics, University of California, San Francisco <i>/Ahituv lab/</i>	Summer 2011
<ul style="list-style-type: none">▪ Performed ChIP-seq experiments on mouse limb tissue to find DNA elements involved in limb development	

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	

Science Communication

Scientific Outreach & Tutorials

- “Modeling DNA Sequences with PyTorch.” (2022) [Tutorial](#) in Towards Data Science, a Medium publication.
- “The Light Side of Genetic Engineering.” (2019) [Article](#) in OneZero, a Medium publication.
- “Genetic Constructor and GSL - The Best of Both worlds.” (2016) [Blog post](#) with Autodesk Bionano Research.

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*
- 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. 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. H. Wilson** et al. (2016) “Genotype Specification Language.” *ACS Synthetic Biology*. ([link](#))

Leadership, Volunteering, & Activities

- Mentor for intro Python students, [Paper Airplanes](#), Women in Tech program (2023 - Present)
- Research mentor for an undergraduate student (2020-2023)
- Peer mentor (2018-2022) and Peer Mentorship Program Organizer (2019-2020) for incoming PhD students
- New Grad Orientation organizer (2018)
- 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
- Recreational data visualization: “Mistborn: The Final Eyebrow.” (2021) [Article](#) in Towards Data Science.
- 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
- [MeadoWatch](#) field data collector, UW Biology citizen science project (2019 - present)
 - Collect wildflower blooming data in Mount Rainier National Park
- Carleton Varsity Athletics (Div III)
 - Women’s Soccer (4 seasons), Women’s Tennis (1 season)
 - Received MIAC (Minnesota) Academic All-Conference honors (2011, 2012, 2013)