

Erin Wilson

🏠 homes.cs.washington.edu/~ewilson6
✉ erinhwilson@gmail.com
🌐 [erinhwilson](#)
in [Linkedin](#)

Recent PhD graduate with 10 years of computational analysis experience across academia and 2 biotech companies ▪ Creative scientist who can connect ideas across computing, biology, and carbon emissions mitigation ▪ Programming skills in machine learning, software development, and data visualization ▪ Effective science communicator who can adapt explanations of complex topics to diverse audiences ▪ **Seeking a data scientist position where I can be part of solutions for challenges in sustainability and climate health**

Keywords: data science, machine learning, microbiology, science communication, data visualization, predictive modeling
Research Tools: Python (PyTorch, pandas, sci-kit learn, numpy, Altair, seaborn, matplotlib), Tableau, AWS, Git, SQL

Education

Ph.D, M.S., Computer Science, University of Washington, Seattle, WA

2017 - 2023

Advisors: Dr. Mary Lidstrom, Dr. David Beck

- **NSF GRFP Fellow**
- **Research focus:** Building computational and machine learning models to predict genetic responses in methane-consuming bacteria

Visiting Ph.D student, Center for Biosustainability, Technical University of Denmark

2022

Supervisor: Dr. Lars Nielsen

B.A. Computational Biology, Carleton College, Northfield, MN

2010 - 2014

Advisors: Dr. David Liben-Nowell, Dr. Jennifer Wolff

- **Graduation Honors:** *Magna Cum Laude*; awarded Distinction on senior thesis.
- **Studies Abroad:** Coastal marine ecology with University of Queensland, Australia

Work & Research Experience

Intern, Data Science, Zymergen

Summer 2018

Mentor: Trent Hauck

- Prototyped deep learning models for predicting the presence of DNA features in microbe genomes

Associate Scientist, Scientific Computing, Amyris

2014 - 2017

Mentors: Dr. Amoolya Singh, Dr. Darren Platt

- Implemented software features for and trained biologists/external collaborators to use **Genotype Specification Language (GSL)**: an open-source DNA design language invented at Amyris. (co-authored: [article](#), [textbook chapter](#), [blog post](#), poster)
- Developed software pipeline and database schema to translate metabolic pathway designs into DNA build instructions for DARPA-funded project to produce 450 molecules ("[Automated Scientist](#)")
- Helped maintain automated whole genome sequencing pipeline and supported scientists in interpreting mutation and coverage data for engineered genomes.
- Facilitated technical communication between biologists and software engineers; held weekly office hours for 1x1 sequencing and GSL support; gave technical presentations to Amyris R&D (~60 people), Strain Engineering (~30 people), and Automation & Computing groups (~30 people); engaged with attendees at Bay Area science nights about Amyris' sustainable technology

Intern, Scientific Computing, Amyris

December 2013

Mentor: Dr. Amoolya Singh

- Implemented data visualization tool to overlay experimental data on yeast metabolic pathway

Research Assistant, Computational Biology, University of Minnesota

Summer 2013

Principal Investigator: Dr. Chad Myers

- Analyzed genetic and chemical-genetic interaction data to predict gene targets for chemical perturbants and coded target prediction pipeline

Research Assistant, Evolutionary Computing, Carleton College

Summer 2012

Principal Investigator: Dr. Sherri Goings

- Executed experiments with populations of mutating digital organisms to examine the effects of limited CPU resources on the populations' ability to evolve complex Boolean logic functions

Research Assistant, Genetics, University of California, San Francisco

Summer 2011

Principal Investigator: Dr. Nadav Ahituv

- Performed chromatin immunoprecipitation sequencing experiments on mouse limb tissue to find DNA elements involved in limb patterning and 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 for graduate students in Computer Science & Engineering	2017
▪ Awarded first year graduate research funding	
Clare Boothe Luce Scholarship for Women in Physics and Computer Science	2012
▪ Received funding for summer research in Evolutionary Computing	
Technos International Fellow	2012
▪ Selected for cultural exchange program in Tokyo, Japan	

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 Academic Publications & Presentations

- **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 the Symposium for Biomaterials, Fuels, and Chemicals (SBFC). Portland, OR
- A. H. Singh, B. B. Kaufmann-Malaga, J. A. Lerman, D. P. Dougherty, Y. Zhang, A. L. Kilbo, **E. H. Wilson**, C. Y. Ng, O. Erbilgin, K. A. Curran, C. D. Reeves, J. E. Hung, S. Mantovani, Z. A. King, M. J. Ayson, J. R. Denery, C. Lu, P. Norton, C. Tran, D. M. Platt, J. R. Cherry, S. S. Chandran, A. L. Meadows. (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

- Research mentor for an undergraduate student (2020-2023)
- Pre-application Review Service mentor (2021-2022)
 - Provided early application feedback to support prospective PhD students from diverse backgrounds
- Peer mentor for groups of incoming PhD students (2018-2022)
 - Peer Mentorship Program Organizer (2019-2020)
- 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
- 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)
 - Named to MIAC (Minnesota) Academic All-Conference (2011, 2012, 2013)