Erin H. Wilson

homes.cs.washington.edu/~ewilson6
erinhwilson@gmail.com
(415) - 233 - 0835
erinhwilson LinkedIn

Recent CS PhD graduate with 10+ years of computational analysis experience (industry, academia) ▪ Creative scientist who can connect ideas across computing, life sciences, and sustainable technology ▪ Programming skills in machine learning, software development, and data visualization ▪ Effective communicator who can adapt explanations of complex topics to diverse audiences ▪ **Seeking a data scientist position where I can be part of solutions for 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. Computer Science, University of Washington, Seattle, WA **2023**
- NSF GRFP Fellow

M.S. Computer Science, University of Washington, Seattle, WA **2019**

B.A. Computational Biology, Carleton College, Northfield, MN **2014**
- *Magna Cum Laude*; awarded Distinction on senior thesis

Work & Research Experience [Academia || Industry]

Graduate Researcher, University of Washington *|Supervisors: Dr. David Beck, Dr. Mary Lidstrom|* **2017 - 2023**
- Built computational frameworks to accelerate genetic engineering efforts in methane-consuming bacteria, a promising carbon removal platform; established a suite of new methanotroph promoter tools
- Characterized the effectiveness of machine learning models for discovering influential genetic patterns from RNA-seq experiments in microorganisms with limited data

Visiting Researcher, DTU Biosustainability Institute, Denmark *|Supervisor: Dr. Lars Nielsen|* **Apr-Aug 2022**
- Used machine learning methods (ICA) to identify independent gene modules in methane-consuming bacteria

Data Science Intern, Zymergen *|Mentor: Trent Hauck|* **Summer 2018**
- Prototyped deep learning models for predicting the presence of DNA features in microbe genomes

Associate Scientist, Scientific Computing, Amyris *|Mentors: Dr. Amoolya Singh, Dr. Darren Platt|* **2014 - 2017**
- **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)
- **Database Development:** 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")
- **Bioinformatics:** Helped maintain automated whole genome sequencing pipeline and supported scientists in interpreting mutation and coverage data for engineered genomes.
- **Communication:** 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 *|Mentor: Dr. Amoolya Singh|* **December 2013**
- Implemented data visualization tool to overlay experimental data on yeast metabolic pathway

Research Assistant, Computational Biology, University of Minnesota *|PI: Dr. Chad Myers|* **Summer 2013**
- Analyzed genetic and chemical-genetic interaction data to predict gene targets for chemical perturbants and coded target prediction pipeline

Research Abroad, Coastal Marine Ecology, University of Queensland *|PI: Dr. Annie Bosacker|* **"Winter" 2013**
- Conducted observational field research in various Australian ecosystems (reef island, rainforest, desert gorge, intertidal zones)

Research Assistant, Evolutionary Computing, Carleton College *|PI: Dr. Sherri Goings|* **Summer 2012**
- 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 *|PI: Dr. Nadav Ahituv|* **Summer 2011**
- 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 **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

- 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, 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 (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)*