

# Erin H. Wilson

🏠 [erinhwilson.com](http://erinhwilson.com)  
✉ [erinhwilson@gmail.com](mailto:erinhwilson@gmail.com)  
☎ (415) - 233 - 0835  
🌐 [erinhwilson](#) in [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

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

## Work & Research Experience [[Academia](#) || [Industry](#)]

<b>Graduate Researcher, University of Washington</b> /Supervisors: <i>Dr. David Beck, Dr. Mary Lidstrom</i> /	<b>2017 - 2023</b>
▪ 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	
<b>Visiting Researcher, DTU Biosustainability Institute, Denmark</b> /Supervisor: <i>Dr. Lars Nielsen</i> /	<b>Apr-Aug 2022</b>
▪ Used machine learning methods (ICA) to identify independent gene modules in methane-consuming bacteria	
<b>Data Science Intern, Zymergen</b> /Mentor: <i>Trent Hauck</i> /	<b>Summer 2018</b>
▪ Prototyped deep learning models for predicting the presence of DNA features in microbe genomes	
<b>Associate Scientist, Scientific Computing, Amyris</b> /Mentors: <i>Dr. Amoolya Singh, Dr. Darren Platt</i> /	<b>2014 - 2017</b>
▪ <b>Genotype Specification Language (GSL):</b> Implemented software features and trained biologists/external collaborators to use open-source DNA design language invented at Amyris. (co-authored: <a href="#">article</a> , <a href="#">textbook chapter</a> , <a href="#">blog post</a> , poster)	
▪ <b>Database Development:</b> Developed software pipeline and database schema to translate metabolic pathway designs into DNA build instructions for DARPA-funded project to produce sustainable materials (" <a href="#">Automated Scientist</a> ")	
▪ <b>Bioinformatics:</b> Helped maintain automated whole genome sequencing pipeline and supported scientists in interpreting mutation and coverage data for engineered genomes.	
▪ <b>Communication:</b> 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	
<b>Intern, Scientific Computing, Amyris</b> /Mentor: <i>Dr. Amoolya Singh</i> /	<b>December 2013</b>
▪ Implemented data visualization tool to overlay experimental data on yeast metabolic pathway	
<b>Research Assistant, Computational Biology, University of Minnesota</b> /PI: <i>Dr. Chad Myers</i> /	<b>Summer 2013</b>
▪ Analyzed genetic and chemical-genetic interaction data to predict gene targets for chemical perturbants and coded target prediction pipeline	
<b>Research Abroad, Coastal Marine Ecology, University of Queensland</b> /PI: <i>Dr. Annie Bosacker</i> /	<b>"Winter" 2013</b>
▪ Conducted observational field research in various Australian ecosystems (reef island, rainforest, desert gorge, intertidal zones)	
<b>Research Assistant, Evolutionary Computing, Carleton College</b> /PI: <i>Dr. Sherri Goings</i> /	<b>Summer 2012</b>
▪ 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	
<b>Research Assistant, Genetics, University of California, San Francisco</b> /PI: <i>Dr. Nadav Ahituv</i> /	<b>Summer 2011</b>
▪ Performed chromatin immunoprecipitation sequencing experiments on mouse limb tissue to find DNA elements involved in limb patterning and development	

## Awards & Fellowships

---

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