

Philippe Chlenski

pac@cs.columbia.edu | [/in/pchlenski](https://github.com/pchlenski) | [git: pchlenski](https://github.com/pchlenski) | (312) 545-3322 | Brooklyn, NY

EDUCATION

Columbia University

PhD Computer Science
MS Computer Science; GPA 4.0

New York, NY
Expected May 2025
Dec 2021

Yale University

BA Mathematics & Philosophy; GPA 3.9; cum laude with distinction

New Haven, CT
May 2018

Deep Springs College

AA Liberal Arts; GPA 3.8

Deep Springs, CA
Jun 2014

EXPERIENCE

Columbia University

NSF Graduate Research Fellowship Program (GRFP) fellow
Graduate research assistant

New York, NY
May 2021–Present
Sep 2020–Present

- Use geometric deep learning to generate representations of metagenomic count matrices [8, 10].
- Built MiSDEED [1], a synthetic data engine for study design and interpretable testing of causal inference frameworks.
- Worked on coPTR [2], a tool to robustly infer microbial birth rates in sparse, noisy, or otherwise irregular clinical data.
- Planned and supervised 11 students on 16 semester-long research projects in computational biology.

Barnard College

Adjunct Professor, BIOL 3308: Introduction to Microbial Genomics

New York, NY
Jan 2023–May 2023

- Developed lab exercises covering Python, algorithms, and bioinformatics tools for computational biologists.
- Gave weekly lectures, held office hours, and graded all lab assignments and individual research projects.

Uber

PhD Machine Learning Engineer Intern

San Francisco, CA
May 2022–Aug 2022

- Performed exploratory data analysis using SQL and Python to find features predictive of merchant fraud.
- Used Tensorflow to scope and develop a custom deep learning model for real-time anomaly detection.
- Deployed model into production using Apache Spark and Docker, saving up to \$30,000 a week in refund losses.

Argonne National Laboratory

Comparative Genomics Researcher

Chicago, IL
Oct 2018–Sep 2020

- Improved RASTtk genome annotation pipeline with supervised learning models built in Keras and Scikit-learn [3,4,5].
- Developed pipeline for AI-based strain discovery and data analysis to optimize threonine production in *E. coli* [6].
- Maintained and edited a legacy Perl codebase serving 100,000 users per month.

PUBLICATIONS

1. **Chlenski**, Hsu, and Pe'er. MiSDEED: a synthetic multi-omics engine for microbiome power analysis and study design. *Bioinformatics Advances*. 2022. [doi: 10.1093/bioadv/vbac043](https://doi.org/10.1093/bioadv/vbac043).
2. Joseph, **Chlenski**, Litman, Korem, and Pe'er. Accurate and robust inference of microbial growth dynamics from metagenomic sequencing reveals personalized growth rates. *Genome research*. 2022. [doi: 10.1101/gr.275533.121](https://doi.org/10.1101/gr.275533.121).
3. Parrello, Butler, **Chlenski**, Pusch, and Overbeek. Supervised extraction of near-complete genomes from metagenomic samples: a new service in PATRIC. *PLOS ONE*. 2021. [doi: 10.1371/journal.pone.0250092](https://doi.org/10.1371/journal.pone.0250092).
4. Davis *et al.* The PATRIC Bioinformatics Resource Center: Expanding data and analysis capabilities. *Nucleic Acids Research*. 2020. [doi: 10.1093/nar/gkz943](https://doi.org/10.1093/nar/gkz943).
5. Parrello, Butler, **Chlenski**, *et al.* A machine learning-based service for estimating quality of genomes using PATRIC. *BMC Bioinformatics*. 2019. [doi: 10.1186/s12859-019-3068-y](https://doi.org/10.1186/s12859-019-3068-y).
6. Hanke *et al.* Engineering of increased L-Threonine production in bacteria by combinatorial cloning and machine learning. *Metabolic Engineering Communications*. 2023. [doi: 10.1016/j.mec.2023.e00225](https://doi.org/10.1016/j.mec.2023.e00225).
7. **[Submitted]** Moshe *et al.* Efficient algorithms for simulating sequences along a phylogenetic tree.
8. **[Submitted]** **Chlenski**, Turok, Moretti, and Pe'er. Fast hyperboloid decision tree algorithms.
9. **[Submitted]** **Chlenski**, Ricaurte, and Pe'er. Signals of microbial growth learned from single amplicon samples.
10. **[Submitted]** Moretti *et al.* Jet reconstruction using a variational pseudo-marginal framework.
11. **[In preparation]** **Chlenski**, Turok, and Pe'er. Hyperbolic community embeddings for metagenomic samples.