

BRIAN J. ARNOLD

DATA SCIENTIST | COMPUTATIONAL BIOLOGIST | SOFTWARE ENGINEER

CONTACT

651-206-2522

brianjohnarnold@gmail.com

brian-arnold.github.io

SKILLS

Software (Python)

AI/ML

Pipelines

Bioinformatics

Data Analytics

Communication

Strategic thinking

EDUCATION

Harvard University

2009-2015

PhD in Biology

University of Minnesota – Twin
Cities

2004-2008

BS in Plant Biology

PROFILE

Highly adaptive computational biologist with 9+ years of professional data-science experience, specializing in software (Python), machine learning, data pipelines, and detailed analysis of diverse data modalities. Successfully led projects from conception to completion and significantly improved analytical capabilities of research teams by managing large datasets with scalable software. Learns fast, works hard, and navigates uncertainty well.

EXPERIENCE

Senior Data Scientist, Princeton University

2020-Present

- Developed, optimized, and deployed AI model to study gene expression gradients and segment spatial domains in spatial transcriptomics data
- Led image segmentation and machine learning analyses of 3D videos to quantify neural activity from animal behavior experiments
- Managed 2 trainees on large scale sequencing projects involving 700+ samples and 10+ TB of data
- Collaborated with 9+ groups, authored papers in Science, Nature Methods, Nature Reviews Microbiology, and Genome Biology

Senior Bioinformatician, Harvard University

2018-2020

- Created and published workflow to massively parallelize mutation detection in whole-genome sequences
- Designed machine learning classifier to detect horizontal gene transfer in whole-genome alignments

Postdoctoral Researcher, Harvard School of Public Health

2015-2018

- Designed fast, memory-efficient C++ code to simulate models of bacterial evolution and used Bayesian Optimization to fit these models to data

