

Kai Stewart

SCIENTIST, COMPUTATIONAL BIOLOGY

BioNTech U.S.

617.528.0353 | kai.w.stewart@gmail.com | [kaiwstewart](https://www.linkedin.com/in/kaiwstewart)



Summary

Computational Scientist with 5+ years of experience designing, implementing, and scaling containerized bioinformatics workflows on cloud platforms. Proven ability to build robust data pipelines leveraging ML models, Python, and Nextflow for computational protein design and single cell data analysis. Skilled in bridging computational results with wet lab validation and committed to high software development and engineering standards.

Skills

Programming	Python (Intermediate-Advanced), Git/GitHub, Bash, Linux
Pipeline Development	Nextflow, SLURM
Machine Learning	AlphaFold2, RFDiffusion, RFantibody, ProteinMPNN, ESM-Fold
Cloud Computing	AWS, Docker
Bioinformatics Analysis	10x CellRanger, Scanpy/Muon, Pandas, NumPy, SciPy
Languages	English (Fluent), German (Fluent)

Certifications

AWS AWS Certified Cloud Practitioner 
tecRacer Certified Cloud Engineer - Foundational 
In-progress: AWS Certified Cloud Developer Associate

Work Experience

BioNTech

Cambridge, MA

SCIENTIST — COMPUTATIONAL BIOLOGY

June 2024 - Present

- Develop, deploy, and scale cloud-based protein design ML algorithms to support internal protein engineering efforts.
- Support cross-functional teams, as well as clinical trials, through regular processing and analysis of single-cell data (GEX, CITE, TCR/BCR, spectral flow), aiding hypothesis generation, informing clinical trial manufacturing decisions, and helping to validate drug and vaccine candidates and experimental protocols.
- Patented a modular and scalable peptide-string optimization algorithm as an end-to-end workflow in support of projects ranging from oncology to infectious disease.

ASSOCIATE SCIENTIST — COMPUTATIONAL BIOLOGY

February 2023 - June 2024

- Drove de novo protein design project by developing computational workflows, applying ML tools (ProteinMPNN, AlphaFold, ESM-Fold), implementing analysis frameworks, and scaling analyses in the cloud.
- Engineered fault-tolerant, containerized (Docker) single-cell data processing pipeline using Nextflow on an AWS HPC cluster, automating parallel deployment (10x CellRanger), analysis, and storage, saving the team dozens of hours per month.
- Fostered and implemented coding and collaboration standards, leading a GitHub workshop to introduce colleagues to version control, drive team-wide professional development and improve group-wide code quality.

Getz Lab, Broad Institute of MIT and Harvard

Cambridge, MA

ASSOCIATE COMPUTATIONAL BIOLOGIST

June 2020 - January 2023

- Analyzed bulk RNA and whole-exome sequencing data in collaboration with Dana-Farber Cancer Institute researchers, generating insights into CAR-T cell therapy resistance mechanisms in DLBCL.
- Enhanced DLBCL study power by building a Docker-based data processing tool, enabling efficient handling of terabyte-scale genomic data.
- Partnered with Massachusetts General Hospital researchers on multi-patient single-cell data analysis, uncovering pre-cancerous cellular expression programs in breast tissue harboring BRCA1/2 mutations.

Education

Tufts University

Medford, MA

B.SC. APPLIED PHYSICS - SUMMA CUM LAUDE

- Honors Thesis: A Deep Convolutional Neural Network Approach to Liquid Argon Time Projection Chamber Image Synthesis.