David Streid

david.streid.dev@gmail.com
linkedin.com/in/davidstreid • github.com/DavidStreid • davidstreid.com

PROFILE

Engineer with strong educational background in computer science, biology, & statistics. Experienced in enterprise software and data visualizations. Interested in the future of genomics/healthcare and eager to apply skillset to related challenges in bioinformatics.

EDUCATION

John Hopkins University, Advanced Academic Programs — Online
Master of Science in Bioinformatics — May 2020

Columbia University, Columbia College — New York, NY
Bachelors of Arts in Computer Science & Biology — May 2016

Favorite Coursework: Ubiquitous Genomics; Natural Language Processing; Machine Learning; Human Evolutionary Genetics

TECHNICAL SUMMARY

Languages - Java, Python, JavaScript/TypeScript, Bash, SQL, MatLab, C#, R; *Frameworks* – Nextflow, Spring, Node, React/Redux, Angular; *DevOps* – AWS, Docker, Gulp, Gradle, Singularity, IBM Spectrum LSF; *Operating Systems* – Linux, macOS

PROFESSIONAL EXPERIENCE

Variantyx — New York, NY
Senior Bioinformatics Engineer — Nov 2021-Present

- Add features to mature pipeline such as analyzing structural variant pathogenicity & variant-calling in hard-to-map regions
- Perform regular updates of databases from Clinvar and HGMD with automated ETL and difference-validation scripts

Memorial Sloan Kettering, Integrated Genomics Operations — New York, NY
Bioinformatics Software Engineer — July 2019-Nov 2021

- Developed bioinformatics pipelines on LSF cluster for high-throughput sequencing processing with workflow languages & docker
- Built React/Node applications for researchers to analyze sequencing project statistics output by bioinformatic tools such as an LDAP login application, a website displaying sequencing analysis, and sequencing project tracker integrating with our LIMS
- Analyze and implement new technologies for scaling output, such optimal compression algorithms & Illumina-DRAGEN platform
- Collaborated with researchers on projects such as gene-expression analysis & mapping improvements with modified references

Goldman Sachs — New York, NY
Technology Analyst — Dec 2017-July 2019

- Built full-stack react/redux/slang application automating trader configuration requests & integrated w/ existing java application. Application saves hours of deployment time per request & maintains traceable state of approvals for our team and compliance
- Member of five-person team decommissioning legacy trading system & onboarding workflow w/ less operational support/setup
- Bimonthly support requiring intimate knowledge of application & interconnected trading systems to resolve production issues
- Run deployments and resolve build issues as a member of deployment team that does monthly, phased deployments

E*Trade — Menlo Park, CA
Staff Software Engineer — July 2016-Nov2017

- Reduced latency of APIs in java to handle thousands of daily REST/SOAP requests by implementing multithreaded design
- Wrote entire test folders for multiple applications and maintain 80% code coverage using Mockito/JUnit to handle all test cases

SELECTED PROJECTS

nf-fastq-plus (nextflow, python, bash) — github.com/mskcc/nf-fastq-plus

- Nextflow pipeline detecting raw sequencing output (BCL) and conditionally processing data through different aligners, reference genomes, and analysis toosl (e.g. RNA vs. WGS) to produce analyzed FASTQ and storing stats later displayed for researchers

fitness-simulation (Typescript, angular) — https://davidstreid.com/projects/fitness

- Simulation of natural selection that allows for modification of trait mutation rate and their effect on fitness

TruCal Variant Recalibrator (wdl, google cloud, python) — github.com/SVAI/MutRackers

- 2017 SVAI AI Genomics Hackathon Submission - Supervised learning pipeline classifying true and false genetic variants in variant callers like GATK to identify variant caller mistakes. Wrote process to analyze 300X coverage BAM files hosted on google cloud

RESEARCH EXPERIENCE

Cornell University, Department of Plant Science — Ithaca, NY
REU Research Assistant, Advisor: Professor Klaas van Wijk — June 2013-Aug 2013

- Researched nucleoid defects in mutants of Clp protease and the relation of ClpS to DNA repair genes in plant chloroplasts
- Analyzed DNA/protein function via immunoblotting & transformation of E. Coli plasmids & Agrobacterium binary vectors.
- Presented at Plant Genome Research Program Symposium - "Exploring the Role of Plastid Clp Protease in DNA Maintenance"