

# Tim Farrell

*Data Engineer/ Manager (Biomedical)*

106 Myrtle Street, #2  
Boston, MA 02114  
(856) 904-6017

[tfarrell01@gmail.com](mailto:tfarrell01@gmail.com)  
[tmfarrell.github.io](https://tmfarrell.github.io)

## EXPERIENCE

### Day Zero Diagnostics, Boston MA

#### Manager, Bioinformatics Data Engineering

JUN 2020 - PRESENT

Manage one of [the largest and fastest growing genomics + antimicrobial resistance databases](#) in the world. Increased AI/ ML model dataset size 55+% over 1.5 year period. Led successful migration from legacy to cloud-native systems.

#### Bioinformatics Engineer

MAY 2019 - JUN 2020

Developed/ operated [genomics data workflows/ infrastructure for clinical-grade services](#).

### US Army Reserve, Taunton MA — Operations Officer

MAY 2012 - MAR 2021

Part-time military leader. Activated for DoD COVID-19 response in spring 2020 (3 mo) and deployed overseas 2016 - 2017 (12 mo). Managed operations of ~100-soldier units.

### The Broad Institute, Cambridge MA — Assoc. Computational Biologist

MAY 2017 - JUN 2019

Developed tool for [accurately genotyping polyclonal malaria infections](#) to analyze 60k+ samples from phase III malaria vaccine clinical trial. Developed [scalable, cloud-native genomic pipelines for malaria genomics](#) (parasite and vector).

### Boston University, Dept. of Health, Boston MA — Research Assistant

AUG 2015 - MAY 2016

Built pre-processing pipeline for fMRI connectivity study of Human Connectome data.

### Boston University School of Medicine, Boston MA — Research Assistant

DEC 2014 - DEC 2015

Developed data management and analysis system for 24-primate study on metabolism.

## EDUCATION

### Boston University, Boston MA — MS, Bioinformatics

SEPT 2014 - MAY 2016

Software engineering and data science focus.

### Rutgers University, New Brunswick NJ — BS, Biomedical Engineering

SEPT 2008 - MAY 2012

Bioengineering focus.

## SKILLS

Languages: Python, SQL, Javascript, Java, R

Data engineering: Postgres, Prefect/ Airflow, BigQuery, Metabase, MongoDB, GraphQL, streamlit

DevOps: git, docker, CircleCI, Kubernetes, GCP, APIs

Data science: PyData stack (pandas, seaborn, numpy/ scipy, jupyter), scikit-learn

Data types: genomic, clinical, neuroimaging, imaging (cell microscopy), web data, motion capture

Bioinformatics: Nextflow, WDL/ Cromwell, bwa, GATK, Picard, (sam|bed|vcf)tools, biopython

Genomics: WGS, amplicons, (variant|haplotype)-calling, assembly, annotation, kmers

## PROJECTS

### [covid-geo2geo](#)

Geographic classifier for COVID-19 genomes

[fullstackopen](#) Course on React, Node.js, MongoDB and GraphQL

### [data science bowl 2018](#)

Detecting cell nuclei using MASK-RCNN

### [blood antigenicity prediction](#)

clinical sequencing-based structural variant classifier for Rh antigenicity