Carl Vitzthum

Brookline, MA • carlvitzthum.github.io • carl.vitzthum@gmail.com

Motivated developer seeking to expand my technical and leadership skill set. Extensive experience planning and building robust open-source applications on interdisciplinary teams.

Languages:	*Python, JavaScript, R, Bash, MATLAB/Octave*
Technologies:	*AWS, Elasticsearch, Pyramid, ReactJS, Docker, Travis CI, SQLAlchemy, UNIX*
Processes:	*Version Control with Git, RESTful API Design, Unit and Integration Testing, Cloud DevOps, Code Review, Python Packaging, Machine Learning*

HIGHLIGHTED EXPERIENCE

Harvard Medical School • Boston, MA **2016 - Present**

Led back-end engineering and DevOps of the 4D Nucleome Data Coordination and Integration Center (4DN-DCIC), a NIH-funded center supporting 119 genomics labs with our cloud-based data portal. Ongoing responsibilities were developing features, writing tests, and managing the AWS infrastructure for portal and associated projects. Below are some highlights from my 3.5 years at HMS:

Senior Software Developer **2019 - Present**

AWS, Elasticsearch, Pyramid, ReactJS, Docker, Travis CI, PostgreSQL

- Led back-end development and DevOps of the Clinical Genomics Analysis Platform (CGAP), a collaboration between HMS and Brigham and Women's Hospital
- As acting lead developer for CGAP, established back-end roadmap and best practices for the MVP
- Hired and onboarded multiple developers, including the future lead developer for CGAP
- Developed back-end features for *Tibanna,* a cloud-based pipeline runner (published in *Bioinformatics)*

Back-End Developer **2018 - 2019**

- Became lead developer for the 4DN-DCIC and solely responsible for the back-end and DevOps
- Created highly scalable, cloud-based system for indexing 4DN portal metadata into Elasticsearch
- Implemented a cloud-based, Dockerized JupyterHub solution available to 650+ 4DN portal users

Junior Software Developer **2017 - 2018**

- Created *Foursight*, a serverless web application used internally to schedule and display customized Python scripts for data portal management
- Created *hic2cool,* an actively used bioinformatics Python package published on PyPi
- Developed *dcicutils*, a Python API for the portal published on PyPi
- Overhauled Elasticsearch setup for the portal, including index mappings and query building

Junior Scientific Programmer **2016 - 2017**

- Bioinformatics analysis and data management on the HMS compute cluster for the Park Lab
- Back-end and front-end web development for the 4DN data portal

EDUCATION

Harvard Extension School • Cambridge, MA **July 2019**

Data Science Graduate Certificate (4 courses)

Colby College • Waterville, ME **BA, May 2016**

Major: Computational Biology • *Minor:* Chemistry **GPA: 4.00, *summa cum laude***

OTHER EXPERIENCE

Research Fellow • The Jackson Laboratory • Bar Harbor, ME **Summer 2015**

- Developed a pipeline for discovery of enriched genetic variants linked to late-onset Alzheimer's disease (LOAD) from whole exome sequencing data
- Adapted an existing web application for visualization of LOAD variants using HTML, jQuery, and D3.js

Computer Science Tutor and Grader • Colby College • Waterville, ME **2014 - 2016**

Research Assistant • Colby College • Waterville, ME **2012 - 2014**

Research Assistant • Norwich University • Northfield, VT **Summers 2011 - 2013**

MISCELLANIA

- Created *motif-finder*, an open-source Python package for a novel method of conserved motif identification in protein and DNA sequences (2019). This work was based off of an honors thesis done with Prof. Andrea Tilden (2016)
- Volunteered at Roxbury Prep Middle School to run a programming club (2018)
- Avid Ultimate Frisbee player currently captaining a Boston-based team

---

ACADEMIC PUBLICATIONS

- Lee S, Johnson J, et al. **Tibanna: software for scalable execution of portable pipelines on the cloud.** *Bioinformatics.* 2019
- Lodato MA, et al. **Aging and neurodegeneration are associated with increased mutations in single human neurons.** *Science.* 2018
- Sherman MA, Barton AR, Lodato MA, et al. **PaSD-qc: quality control for single cell whole-genome sequencing data using using power spectral density estimation.** *Nucleic Acids Research*. 2018
- Doczi MA, Vitzthum CM, Forehand CJ. **Developmental expression of Kv1 voltage-gated potassium channels in the avian hypothalamus.** *Neurosci Lett*. 2016
- Hinkle KL, Anderson CC, Forkey B, et al. **Exposure to the lampricide 3-trifluoromethyl-4-nitrophenol results in increased expression of carbohydrate transporters in S. cerevisiae.** *Environ Toxicol Chem*. 2015