

SHINN TANIYA

Pasadena, CA ✉ staniya@caltech.edu ☎ (+1) 310-709-5219

🌐 [linkedin.com/in/shinnosuke-taniya](https://www.linkedin.com/in/shinnosuke-taniya) 🐙 github.com/staniya 🌐 staniya.github.io 🏠 hackerrank.com/staniya1

Education

Harvey Mudd College

Aug 2018 - May 2022

Bachelor of Science in Computer Science with Concentration in Economics (Dean's List)

Claremont, CA

- **Major GPA:** 3.67 (Honors)
- **Computer Science and Mathematics Coursework:** Algorithms, Artificial Intelligence, Computability and Logic, Computer Systems, Computer Vision, Data Structures and Program Development, Differential Equations, Multi-variable Calculus, Natural Language Processing, Probability and Statistics, Software Development

Employment

Lake Day Ventures, LLC

Jan 2023 - Current

Contract Software Engineer, AI & NLP Systems

Remote

- Produced an automated tool for Google Ads by fine-tuning a GPT-2 XL machine learning model on internal advertisement campaign data over the past decade.
- Developing a production-grade Dockerized Flask application with PostgreSQL 15.1, Gunicorn, and Nginx to be managed in AWS ECR and deployed as an AWS Sagemaker endpoint with automatic scaling.
- Creating a Python web scraping program using BeautifulSoup to assist the firm in downloading textual data from client campaigns to enrich internal data and fine-tune ChatGPT2 effectively.

Schmidt Academy for Software Engineering at Caltech

Aug 2022 - Current

Lead Developer For Zwicky Transient Facility(ZTF) Quarter Century Sky Project 🔗

Pasadena, CA

- Designing a PostgreSQL 15.1 processing and archival system with a parallelized, high-throughput processing infrastructure to store photometrically refined light curve data, employing a memory-optimized version of Q3C, a sky indexing scheme programmed in C.
- Equipping the archival system with a photometry data reduction pipeline, API, and a GUI to enable cone searches, data filtering, and periodogram/ other statistical computations.
- Developing a development and production-ready Debian 11 Docker setup for a modular Django application that employs Celery and Redis for task management and a Gunicorn WSGI HTTP server as a reverse proxy and load balancer.

Revorf - Bioinformatics Firm

Oct 2020 - Aug 2022

Lead Developer | R&D Scientist 🔗

Tokyo, Japan

- Led the development efforts to design and implement an AWS data flow ecosystem that improved the company's ability to perform cost-effective and efficient processing of Big Data containing RNA and other functional products.
- Took leadership in the contract with Toshiba Digital Solutions to employ their Simulated Bifurcation Machine+ to obtain an optimal solution to a combinatorial problem involving predicting allosteric regulatory sites for proteins.
- Created a Tkinter GUI framework to enhance the discovery of gene diagnostic agents and defined the product's value proposition to align with the market needs of business partners: RIKEN, top Japanese universities, and significant diagnostic drug development firms.

CYBEX and LongHash (Partnership)

Jan 2018 - Aug 2018

Core Developer For Decentralized Cryptocurrency Exchange

Tokyo, Japan & Shanghai, China

- Developed a full-stack hardware cold wallet in C for EOS and stabilized the CYBEX crypto exchange platform.
- Developed Telegram bots in Python to improve customer experience, crypto transactions, and anti-spam software.

Software Projects

Schmidt Academy and Harvey Mudd College (Partnership)

Aug 2021 - May 2022

Technical Project Manager For Cai Lab Clinic Project 🔗

Claremont, CA

- Oversaw team using Jira to manage Clinic project to visualize spatial genomics for biology and translational medicine (PM tasks: activity & resource planning, team & time & project risk management, progress monitoring & updates).
- Developed a data pipeline to analyze high-dimensional spatial genomic data from seqFISH+ experiments and configured an appropriate file format (Cloud Optimized GeoTIFFs) for storing data in AWS S3.
- Worked with Cai Lab scientists and engineers to provide flexibility to parse genomic data to discover new biological insights, such as identifying phenotypic differences between gene expression targets.
- Developed an AWS Amplify React interface that allows researchers to access, visualize, and analyze the data.

Skills

Software: ASM, Bash, C, C++, HTML5 & CSS, Java, JavaScript, MATLAB, Python, R, Racket, Ruby

Expertise: Agile Methodologies, Application Scaling, Big Data Analytics, Command Lines, Cloud Computing, Database Infrastructure, DevOps Practices, Security Compliance, Software Development Lifecycle Tools, Version Control

Spoken Languages: English and Japanese (Bilingual Proficiency), Chinese (Professional Proficiency — HSK VI certificate)