

# Guotai Shen

Brookline, MA | sgt1796@gmail.com | www.linkedin.com/in/guotaishen | (857) 241-7380 | github.com/sgt1796

## EXPERIENCE

### Stoneye | AI Developer Consultant

08/2024 – Present

- Lead team of 3 to develop a Retrieval-Augmented Generation (RAG) system for TinyTalk app
- Architected and implemented an embedding and search system, using Facebook AI Similarity Search (Faiss) for query acceleration and SQL3 for data storage
- Wrote supportive tools for prompt engineering with auto formatting and reusable prompt functions
- Fine-tuning YOLO computer vision models for GUI object detection

### Dana Farber | Bioinformatics Intern

07/2024 – 09/2024

- Developed WDL scripts to integrate 5 CRISPR analysis tools into DNAnexus platform
- Designed and implemented a WDL workflow to perform comprehensive CRISPR pooled sequencing analysis
- Utilized diverse tech stack including WDL, Bash, R, Python, and Docker for pipeline development

### Exo Therapeutic | Bioinformatics Co-op

07/2022 – 12/2022

- Performed comprehensive analyses including Survival Analysis (TCGA), CRISPR functional co-dependency (Depmap/CCL), and Protein-Domain interactions (Interactome3D/UniProtKB)
- Developed and optimized pipelines for allosteric site prediction in proteins, enhancing database functions
- Identified key protein factors (conservativeness, mutational frequencies) from PDB data to refine allosteric site prediction algorithms

### Boston College | Molecular Genetic Lab

01/2020 – 06/2020

- Conducted research on plasmid design and recombinant DNA techniques, utilizing *Schizosaccharomyces pombe* as a model organism for genetic screening

### BIDMC, Center for Life Science | Sleeping Disorder Lab

05/2016 – 09/2018

- Analyzed EMG and EEG of rats to categorize stages of sleep

## EDUCATION

### Northeastern University | M.S. in Bioinformatics

09/2021 – 12/2023

### Boston College | B.S. in Biology, B.A. in Mathematics

09/2019 – 05/2021

### UMASS Amherst | B.S. in Biology and Pure Mathematics

09/2017 – 09/2019

### Brookline High School | High School Diploma

05/2014 – 06/2017

## PUBLICATION

Gao, Y., Sang, C., **Shen, G.**, & Xu, Y. (2024). *Discovery of genes prominent in ADHD and SUDs: A Meta-analysis and review*. 2nd International Conference on Modern Medicine and Global Health (ICMMGH 2024).

- Performed GWAS analysis and created data visualizations, providing critical technical support for the project

## SKILLS

**Tools:** R, Python, Java, Bash, WDL, Docker, Excel, Minitab, Mathematica, LaTeX, Markdown

### Biology:

- Wet lab: DNA/RNA extraction, PCR, Gel electrophoresis, Cloning, Yeast/bacteria transformation
- Dry lab: Quality control, Genome/Transcriptome Assembly, Differential Expression, Variant Calling, GWAS

## VOLUNTEERING

### Dana Farber Cancer Institute | Ambassador & Resource Center Staff

06/2024 – Present

- Served as patient ambassador and resource center staff, providing frontline support and managing services

**Work Authorization:** Eligible to work in the US without sponsorship.