

# Hossam Zaki

📞 (401) 855 5567 • ✉ [hossam\\_zaki@brown.edu](mailto:hossam_zaki@brown.edu) • 🌐 [hossam.computer](https://hossam.computer)  
<https://github.com/hossam-zaki>

## Education

### Brown University (3.9/4.0)

*Bachelor of Science in Computer Science and Biology,*

**Providence, RI**

*Expected Graduation: May 2022*

- **Related Coursework:** Computational Biology, Computer Vision, Software Engineering, Experimental VR, Practical System Skills, Linear Algebra, Honors Statistical Inference
- **Fall 2020:** Deep Learning, 3D Computer Vision Using Deep Learning, Scientific Computing in C++, Physiology
- **Programming Languages:** Python, Java, Dart, SQL, HTML/CSS, Bash, Javascript, Swift, R, C++
- **Libraries/Frameworks:** Node.js, Angular.js, Android Studio, Flutter, Pandas, Numpy, Tensorflow, Sklearn, PyTorch
- **Scholarships:** Jackie Robinson Foundation Scholar, Merck ACS Scholar

## Work Experience

### Computational Biology (CSCI 1810)

#### *Undergraduate Teaching Assistant*

**Providence, RI**

*September 2020 - Present*

- Iterated upon course content, graded assignments, conducted TA hours, and set up occasional labs
- Worked with the Ethics TA program to develop new CS ethics content for both this course and the department in general

### Wessel Lab at Brown University

#### *Undergraduate Researcher*

**Providence, RI**

*July 2017 - Present*

- Researched the effect of Sea Star Wasting Disease on the embryonic development of Sea Star and Sea Urchin
- Developed pipeline to search and find new genetic motifs in the RNA of germ-line cells using Python and the Knuth-Morris-Pratt algorithm
- Co-authored an abstract presented in Developmental Biology of the Sea Urchin Conference titled: "Bisphenol A exposure differentially affects echinoderm embryogenesis. Developmental Biology of the Sea Urchin"

### Broad Institute of MIT and Harvard

#### *Undergraduate Researcher*

**Boston, MA**

*June 2020 - August 2020*

- Researched effect of mutations in DNA repair genes on whole genome structural variants
- Developed algorithm using Python and R and analyzed data using various statistical tests
- Combined data from several databases including ICGC, PCAWG, Uniprot, Ensembl, and TCGA

### Google

#### *SPS Program Participant*

**San Francisco, CA**

*June 2020 - August 2020*

- Worked with team to develop a real-time voting app using Java and HTML/CSS
- Tested app with several Googlers and gave presentation to several people in the Translate team

### Protein Data Bank of Europe

#### *Software Engineering and Bioinformatics Intern*

**Cambridge, UK**

*June 2019 - August 2019*

- Developed a pipeline to superpose protein structures in a UniprotKB Accession, cluster the proteins based on structure similarity, and show them visually on PyMol. Now deployed on PDBE-KB protein pages.
- Designed Algorithm using Python and SQL and analyzed data using Hierarchical Clustering.
- Compiled the algorithm on over 47,000 Uniprot Accessions which included over 350,000 PDBE entries
- Co-authored and published a paper to Nucleic Acids Research titled "PDBE: Improved findability of Macromolecular structure data in the PDB"

## Software Projects

### Apollo | A Speech-to-Text Note-Taking App for Doctors

<https://github.com/hossam-zaki/apollo> | <http://apollohealth.herokuapp.com/>

- Utilized Google's Speech Recognition Library to capture appointment transcript, and used Knuth-Morris-Pratt to search for relevant details
- Encrypted patient data and appointment transcript using Google Tink
- User-tested with medical students at Brown Medical School, and doctors at Rhode Island Hospital

### Computational Analysis of Sea Urchin Transcriptome

[https://github.com/hossam-zaki/Sea\\_Urchin\\_Transcriptome\\_Analysis](https://github.com/hossam-zaki/Sea_Urchin_Transcriptome_Analysis)

- Developed pipeline to search for known genetic motifs in the transcriptome of the Sea Urchin, as well as discovering new motifs
- Implemented Knuth Morris Pratt Algorithm to linear-time pattern recognition as well as a Suffix Trie to search for the longest most common substring
- Currently developing a machine learning model to search for transcription factors

### X-Ray Classification

<https://github.com/hossam-zaki/X-Ray-Classification>

- Built a neural network to classify X-Ray images in NIH Dataset and added in COVID-19 X-Rays
- Achieved an average of 65% True Positive Rate
- Utilized Tensorflow and Keras frameworks

### Protein Structure Clustering through Superposition

<https://www.ebi.ac.uk/pdbe/pdbe-kb/proteins/P00698>

- Built a tool to cluster proteins based on structural similarity, now used 5,000+ times by labs all over the world