

# AARON YAO-SMITH

aaron@aaronyaosmith.com ◊ aaronyaosmith.com ◊ (607) · 288 · 2756

github.com/aaronyaosmith ◊ linkedin.com/in/aaronyaosmith

## EDUCATION

---

**Cornell University, Ithaca, NY**

August 2017 - Present

Intended B.S. in Computer Science, May 2021; GPA: 3.78 of 4.00

**Rel. Coursework:** Data Structures & Functional Programming · Discrete Structures · Intro. to Data Science

## EXPERIENCE

---

**Singer Lab, Dana-Farber Cancer Institute and Harvard Medical School**

May 2018 - Present

*Research Assistant, Biostatistics and Computational Biology*

*Boston, MA*

- Designed and implemented from scratch tool for statistical analysis of RNA sequencing data (Python: pandas)
- Tool uses novel statistical method to find multi-gene sets marking specific cell type. Traditional methods used in current research are limited to single-gene markers; this is less specific and thus less effective in cell type prediction
- Reduced predictive error by >20% over single-gene in most 2-gene cases; further reduction expected for  $\geq 3$ -gene
- Research abstract was accepted and presented at conference (Single Cell Genomics, Broad Institute of MIT and Harvard, Cambridge, MA; October 29-31, 2018. Program at [weizmann.ac.il/conferences/SCG2018/program](http://weizmann.ac.il/conferences/SCG2018/program).)
- Optimized algorithm continually: reducing initial code's >6 hour sample dataset runtime down to 6 minutes

**Steffel Lab, D'Amore-McKim School of Business, Northeastern University** May 2018 - August 2018

*Developer/Research Assistant*

*Boston, MA*

- Designed and implemented web-based psychological research study from research group's qualitative specification
- Deployed study to preliminary experimental group of 300 subjects (Heroku, Django, Python: oTree)
- Collated purely qualitative design specifications from 5 researchers (professors and PhD student) at 3 universities
- Trained two new full-time graduate RAs in order for later deployment to larger audience using Amazon mTurk

**AIDE Design Team, AguaClara Cornell**

September 2018 - Present

*Developer*

*Ithaca, NY*

- Elected as AIDE Design team lead for upcoming Spring 2019 semester
- Parameterize water treatment plant design documents in team of 3 given water needs of individual communities; eliminating design time-cost of already low-cost plants, decreasing barrier of entry for those without clean water
- AguaClara Cornell has partnered with local NGOs to provide clean water to 83,000 in Honduras and Nicaragua

## LEADERSHIP AND PROJECTS

---

**Co-Founder, Team 6731, FIRST Robotics**

2016 - 2017

Co-lead for team of 20. NE District Rookie Inspiration Award, 2017. NE District Rookie All Star Award, 2017. Recruited members, organized meetings and team structure. No town FIRST Robotics team existed prior.

**'COMET: Identifying candidate marker panels.'** [github.com/aaronyaosmith/HG\\_marker\\_detection](https://github.com/aaronyaosmith/HG_marker_detection)

Statistical analysis tool designed for research use. CLI and data visualization. AWS deployment TBA. Python.

**'Incentive compatible disclosure' psychological study** [github.com/aaronyaosmith/ICD\\_survey](https://github.com/aaronyaosmith/ICD_survey)

'Game theory'-style multi-subject study, with interaction between/input by subjects via web interface. Python.

## SKILLS

---

**Languages:** Python (sklearn, pandas, numpy) · JavaScript/HTML/CSS · OCaml · Matlab · Bash scripting

**Software:** Emacs · Vim · PyCharm · VS Code · Git · Linux (incl. UNIX utilities)

## OTHER ACTIVITIES

---

**Coursework:** Waves and Thermal Physics, Honors · Math. Physics (Real and Complex Analysis) · Electronic Materials

**Clubs:** Cornell Taiwanese American Society · Cornell Kendo Club · Cornell Chorale