

# Xiaowei (Aaron) Chu

Santa Cruz, CA, 95064 | (831)-346-8352 | [xweichu@hotmail.com](mailto:xweichu@hotmail.com)

<https://www.linkedin.com/in/xweichu>

## Education

**Ph.D. in Computer Science** — University of California, Santa Cruz Sep. 2017 - Jun. 2022

- Research Areas: Distributed Storage System, Big Data
- Courses: Machine Learning, Data Systems, Distributed Systems

**B.E. in Software Engineering** — East China Normal University Sep. 2009 - Jul. 2013

- Courses: Data Structure, Operating System, Database Systems, Software Process, Software Testing & Quality Assurance

## Work Experience

**Technical Support Specialist**— Microsoft Jul. 2013 - Jul. 2017

- Troubleshoot and provided triage and investigation of production environment issues for enterprise customers.
- Debugged source code of Office and assisted product group to fix bugs.
- Provided consulting and training services of Office, O365 and Yammer.

**Intern of Exchange Server Team** — Microsoft Jun. 2012 - Jun. 2013

- Developed case management tools which helped process the routine jobs automatically.
- Familiarized the architecture of Exchange server, Active Directory, and SQL server. And set up labs of for experiments and product features testing.

## Skills

**Programming Languages:** C++, Python, C#, Java

**Tools and Framework:** Git, MySQL, MongoDB, Android SDK, Octave, Matlab, Tensorflow

**Other Keywords:** Distributed storage system, Ceph, Programmable storage system

## Teaching Experience

**Teaching Assistant** of CMPS-111 “Operating System” Jan. 2018 - Jun. 2018

## Projects

**Visualize Big Datasets using Glue** Jun. 2018 - Oct. 2018

Used Ceph cluster to store the datasets and offloaded the computation from Glue (a tool for data visualization) client to the Ceph cluster. Used Dask to do the distributed computation. This enables Glue to work with TB or even PB level datasets efficiently.

**Kinship Verification with an Optimal CNN** Feb. 2018 - Mar. 2018

Our goal was to verify the Kinship given photos of two people. Tried two methods: a 4-layer CNN and “Triplet Probabilistic Embedding Vector” to train a 3-layer network. We also built a web app demo just for fun. Tensorflow, OpenCV, Flask, and Python were used in this project.

**Email Insights** Nov. 2016 - Dec. 2016

Designed a software to retrieve the email data from Exchange Online Sever and then created a dashboard with visualizations of the statistics. Also analyzed the sentiments of the texts in emails to give users more insights. Demo application was deployed on Windows 10 and Cortana was integrated so that voice could be used to interact with the app.