

Brian Xi

Email: brianx@berkeley.edu | Phone: (602) 617-8881 | Location: Berkeley, CA

Website: xibrian.github.io | LinkedIn: linkedin.com/brian-xi

EDUCATION

University of California, Berkeley

Graduating May 2024

B.A. in Computer Science, B.A. in Data Science (Urban Science Concentration)

GPA: 3.6

- Coursework: Human Computer Interaction, Convex Optimization, Probability and Random Processes, Algorithms, Data Structures, Data Science, Databases, Econometrics, City Planning, Environmental Justice
- Student Organizations: The Student Advocate's Office, The Berkeley Forum, University Chorus
- Honors: Data Science Honors Program, Out 4 Undergrad Tech (2022, 2021), CAA Leadership Award (2020)

WORK EXPERIENCE

Global Policy Lab

Jan 2022 – Aug 2023

Research Data Scientist

Berkeley, CA

- The Aerial History Project uses machine learning on a unique archive of aerial images taken by 20th century British surveyors to study environmental change and urbanization in developing countries
- Stitched 10,000+ aerial images covering three African countries over three decades using SURF computer vision and Adam optimization techniques to create cohesive mosaics that are compared with modern satellite images
- Validated mosaics by conducting exploratory data analyses in Jupyter Notebooks, using GeoPandas to visualize and inspect mosaics against historical cover maps to achieve stitching errors of around 10 meters
- Improved data pipeline scaling to handle 32+ terabytes of aerial images by using high performance computing, while contributing to supporting, open-source documentation to help future researchers

Amazon Web Services

May 2022 – Aug 2022

Software Development Engineer Intern

Seattle, WA

- Amazon HealthLake is a service that analyzes healthcare data with machine learning to improve patient care
- Led HealthLake feature redesign that uses natural language processing to detect health conditions from medical text and convert it into standardized, FHIR-compliant formats that customers can query to discover trends
- Improved scalability so that 770 FHIR resources are created each second in HealthLake with 99.9% availability using Lambda, SQS and S3, viewable at docs.aws.amazon.com/healthlake/latest/devguide/med-example.html

U.S Department of Transportation Volpe Center

May 2021 – Jan 2022

Pathways Data Scientist Intern

Cambridge, MA

- Developed crowdsourced vehicle blind spot measurement app using Django framework with 400+ vehicle records in order to prevent vulnerable road user fatalities involving large trucks
- Implemented function that saves photos using Cloudinary and Airtable APIs to validate crowdsourced entries
- Built scalable data visualizations and filterable HTML table using D3.js and DataTables to empower users to analyze vehicle data, viewable at <http://blindzonecalculator.herokuapp.com/visualize>

UC Berkeley Office of the Registrar

May 2021 – Aug 2022

Classroom Management Program Coordinator

Berkeley, CA

- Managed after-hours classroom reservations by using Google Scripts to generate daily room reports and led a Chromebook rehaul to digitize reports for team of 30+ student employees, improving efficiency

SKILLS & INTERESTS

- **Languages:** Python, Java, C, SQL, HTML, CSS, JavaScript, R
- **Technologies/Software:** AWS, Git, Heroku, Django, React, QGIS, Jupyter Notebook, GeoPandas, Figma
- **Interests:** Art History, Public Transit, Tennis, Airlines, Traveling, Data Journalism, Wikipedia