

Gao Xian Peh

 [gao-xian-peh](https://github.com/gao-xian-peh) |  [gao-xian-peh.github.io](https://github.com/gao-xian-peh) |  pehgaoxian@berkeley.edu |  +1(510)710-4442 |  [pehgaoxian](https://www.linkedin.com/in/pehgaoxian)

ABOUT	I like telling stories using data.	
EDUCATION	Stanford University <i>M.S. Computer Science, Artificial Intelligence Concentration GPA: 4.0/4.0</i>	2018 - Present
	University of California, Berkeley <i>B.A. Economics, Minor in Computer Science GPA: 3.98/4.0</i> Awarded Full-Ride Singapore Government Scholarship (Social Policy)	2015 – 2018
EMPLOYMENT	United Nations ESCAP , Bangkok, Thailand <i>Software Engineering Intern / GIS Consultant</i> (D3.js, jQuery, QGIS, Python, R) <ul style="list-style-type: none">- Created Interactive Data Visualizations with the UN Statistics team.- Developed a method using Gaussian smoothing in QGIS and R to generate population density maps for detailed geographic representations of hazard exposure. The method is now piloted in 6 countries in Asia and the Pacific. Promoted to GIS Consultant.	2017 - 2018
	Ministry of Social and Family Development , Singapore <i>Social Data and Policy Research Intern</i> (R) <ul style="list-style-type: none">- Worked at the Economist Unit to develop R scripts that extracts textual data from casework memos and applies hierarchical clustering to sieve out various characteristics of welfare benefits recipients. Solution enabled better targeting of services and sped up research processes.	2016
RESEARCH	UC Berkeley Electrical Engineering and Computer Sciences Department , CA <i>Tele-monitoring Lab Research Assistant</i> (D3.js, Leaflet, PHP, MySQL, Android) <ul style="list-style-type: none">- Designed and built web and mobile applications to store, query and visualize vital signs and geospatial data under Dr. Daniel Aranki and Prof. Ruzena Bajcsy.	2016 – Present
	UC Berkeley Haas School of Business , CA <i>Urban and Labor Economics Research Assistant</i> (Python, ArcGIS, R Server, STATA) <ul style="list-style-type: none">- Scraped and analyzed ~1.9 million Crunchbase, AngelList and LinkedIn public profiles along with stock prices using Beautiful Soup in Python and R.- Extended the functionality of OSMnx library and utilized Google Maps Elevation and Overpass APIs, OpenStreetMaps and to compute geospatial statistics of road networks.	2015 – 2018
PUBLICATIONS	<p>Daniel Aranki, Uma Balakrishnan, Hannah Sarver, Lucas Serven, Carlos Asuncion, Kaidi Du, Caitlin Gruis, Gao Xian Peh, Yu Xiao and Ruzena Bajcsy. <i>Running Coach – Cadence Training System for Long-Distance Runners</i>. Proceedings of the 11th International Conference on Pervasive Computing Technologies for Healthcare 2017</p> <p>Daniel Aranki, Gao Xian Peh, Gregorij Kurillo and Ruzena Bajcsy <i>The Feasibility and Usability of RunningCoach: A Remote Coaching System for Long-Distance Runners</i>. Sensors Journal Special Issue: Smart Sensing Technologies for Personalized Coaching</p> <p>Zachary A. Pardos, Scott Farrar, John Kolb, Gao Xian Peh, Jong Ha Lee, <i>Distributed Representation of Misconceptions</i>. Proceedings of the 13th International Conference of the Learning Sciences (ICLS 2018)</p>	
PROJECTS	Analyzing Student Misconceptions on Khan Academy (Python, Word2Vec) – Applied Natural Language Processing (NLP) techniques to cluster student misconceptions on Khan Academy	
	Text Analysis of Hilary Clinton’s Emails (Python, Word2Vec) – Exploring email threads using NLP.	
	Interactive Scatterplot Visualizer (D3.js) – A tool built for easy analysis of TSNE output.	
SKILLS	Programming Languages and Frameworks: Python, Java, C, PHP, JavaScript, Android Data Analysis: R, STATA, Matlab, QGIS, ArcGIS Version Control: Git	