GUANG YANG

http://www.github.com/guang
https://www.linkedin.com/in/gyang8/
San Mateo, CA
garry.yangguang@gmail.com

EXPERIENCE

KOMODO HEALTH, San Francisco, CA
Data Engineer Mar 2015 – Present

As employee #3, I have had the opportunity to add value in different parts of the engineering stack, from scratch:

- Design and implement distributed (Spark, S3) and automated (Airflow) data pipelines that process multiple external data sources to deliver mission-critical insights to end-users
- Lead implementation of internal services (containerized Flask APIs deployed on Kubernetes) that improve transparency (domain-knowledge driven data metrics tracking) and automation (config and metadata tracking)
- Deploy and maintain the data infrastructure (AWS, Spark, Airflow, internal services) that support the data team (13+ developers), leveraging container technologies (Kubernetes, Docker)
- Work with architect and product managers to define requirements and roadmap for scalability of data deployment, data ingestion and internal tooling.

INSIGHT DATA SCIENCE, Palo Alto, CA
Data Engineering Fellow Jan 2015 – Mar 2015

- Built a data pipeline for map analytics in Starcraft II© using replay files.
- Gained experience in large scale distributed data architectures through training, mentorship and projects using Kafka, Spark, Cassandra, and Hadoop ecosystem tools.

BERKELEY COMPUTATIONAL OPTIMIZATION LAB, Berkeley, CA
Graduate Student Researcher Aug 2013 – May 2014

- Formulated and implemented an algorithm to analyze exact reachability for skew-line needle planning in automated brachytherapy
- Published results in an article for the IEEE CASE 2014 Conference

NATIONAL INSTITUTE FOR MATHEMATICAL AND BIOLOGICAL SYNTHESIS, Knoxville, TN
NSF, Research Experience for Undergraduates Jun – Aug 2010

- Built a model in R to simulate the dynamics of Johne's Disease in a U.S. dairy herd. Performed cost analysis comparing existing control strategies and a newly developed testing method
- Published results in a paper for Journal of Biological Systems

TECHNICAL

- *Languages*: Python, SQL, shell
- *Frameworks*: Spark, Airflow, Postgres, Flask, etcd
- *Infrastructure*: AWS (architecture, networking), Kubernetes, Docker

EDUCATION

UNIVERSITY OF CALIFORNIA, BERKELEY, Berkeley, CA
M.S. in Industrial Engineering & Operations Research Dec 2014

- Awarded UC Berkeley Graduate Fellowship

RICE UNIVERSITY, Houston, TX
B.A. in Computational and Applied Mathematics (CAAM) May 2013

- Awarded CAAM-Chevron Undergraduate Research Prize 2013