

Sean Dai

sean@seand.ai ◊ 646-489-3984 ◊ U.S. Citizen ◊ NYC/ATL, USA ◊ GitHub: LatencyTDH

- WORK EXPERIENCE
- ◊ **Senior Software Engineer, Microsoft** Jun. 2020 – present
Redmond, WA (Remote)
I work on an exabyte-scale data lake platform in Microsoft's commerce organization.
 - Project: Slurp - architected and built a large-scale data ingestion platform used by multiple commerce teams; designed solutions for stream processing, change data capture, schema evolution, data cleansing, & orchestration; set best practices for development, security, compliance, & cost optimization.
 - Saved \$2 million in infrastructure costs by migrating a legacy revenue recognition system to Azure; reduced the latency of a critical API by 1000X.
 - Led a team and migrated 200 business-critical, distributed cron jobs running on a legacy platform to Spark using Azure Data Factory with zero downtime and customer impact.
 - Single-handedly deprecated 300K lines of C# and SQL for a legacy invoicing platform; reduced \$100k in annual hardware, development, & operational costs; optimized daily batch jobs to run in 85% less time.
 - Developed data pipelines & user interfaces with Apache Atlas and Azure Purview to catalog siloed production assets; improved data governance and SOX compliance.
 - Launched a family spending limit feature for the Xbox team leveraging a real-time query service.
 - ◊ **Software Development Engineer, Amazon.com** May 2017 – Jun. 2020
Seattle, WA
 - Brought in \$100+ million of revenue by automating the international expansion process of the Small & Light program; reduced the time to launch by 90%; led the engineering effort to launch the program in Germany, Italy, France & Spain.
 - Saved \$14 million annually in shipping costs for Amazon by launching core free shipping in the U.S. for Small & Light items in one week.
 - Launched a new Java backend microservice to handle ~500,000 RPS for real-time fee calculations.
 - Designed and trained various DeepAR+ forecasting models in AWS Sagemaker to better predict gross merchandise sales of small and light items over time.
 - Designed and built an automated system to scan the Amazon product catalog and identify low-priced, fast-moving, profit-generating items with Apache Spark, EMR, S3, & Redshift.
 - ◊ **Software Engineering Intern, Amazon.com** Summer 2016
Seattle, WA
 - Developed high-throughput, low-latency APIs and serverless computing solutions for third-party sellers.
 - ◊ **Undergraduate Research Assistant in Computer Vision** 2015 – 2017
Computational Perception Lab, Georgia Institute of Technology, Atlanta, GA
 - Automated the collection of fine-grained human activities; open-sourced the Let's Dance dataset on Kaggle.
- HONORS & AWARDS
- ◊ Zell Miller Valedictorian Scholarship (full tuition) at Georgia Tech 2013 – 2017
 - ◊ Amazon FBA Hackathon – Best UX Design 2018
 - ◊ Faculty Honors List at Georgia Tech (5 semesters) 2013 – 2017
- EDUCATION
- ◊ **Georgia Institute of Technology**, Atlanta, GA. 2013 – 2017
B.S. in Computer Science, summa cum laude.
- SKILLS
- ◊ **Domain Experience:** Distributed Systems, Big Data, Cloud Computing, Computer Vision
 - ◊ **Languages:** Python, Java, C#, JavaScript, C/C++, Go, Powershell, Bash
 - ◊ **Databases/Frameworks/Tools:** SQL (MySQL, PostgreSQL, SQL Server, Synapse), MongoDB, DynamoDB, CosmosDB, Redis, Spark, Kafka, Docker, Kubernetes, NumPy, PyTorch, Langchain, ReactJS
 - ◊ **OS/Platforms:** Windows, MacOS, Linux (Ubuntu, RHEL, Kali), AWS, Azure, Google Cloud
- PUBLICATIONS
- ◊ Daniel Castro, Steven Hickson, Patsorn Sangkloy, Bhavishya Mittal, **Sean Dai**, James Hays, Irfan Essa. Let's Dance: Learning From Online Dance Videos. *arXiv preprint arXiv:1801.07388*. 2018.
- SIDE PROJECTS
- ◊ **Co-Founder, Quantitative Trading Firm – Marymoor Capital**
A multi-strategy fund building automated trading systems to arbitrage DeFi protocols & other financial markets.