

Prem Rajkitkul

972-965-7736 • pr3m@icloud.com • prem.md • /in/premrajkitkul • GitHub: premrajkitkul • U.S. Citizen

EXPERIENCE

Amazon, Flex / Speed Offers (Last Mile Delivery)

Austin, TX

Software Development Engineer II

Oct 2023 – Apr 2026

- Built the backend architecture powering driver offers for **all 1-hour/3-hour Amazon.com deliveries**; live in **3000+ U.S. cities, expanding WW in 2026** (<https://www.aboutamazon.com/news/retail/amazon-fast-delivery-orders>)
- Designed the REST to GraphQL re-architecture of an API handling **99% of customer-facing offer traffic (5,000+ TPS)**, adopting a server-driven UI pattern that cut frontend implementation and app release time by **80%**
- Redesigned the schema for a database handling **500M queries/day** and serving as the source of truth for all in-app driver offer actions, reducing an ongoing **5-year migration to <8 weeks** with **0 downtime**
- Identified a scaling bottleneck in a high-traffic code path, cutting database writes per API call from **180+ to 7 (>96% reduction)**, saving **500k writes in 3 months** of rollout) and eliminating repeat high-severity incidents from throttling
- Led the launch of all short-lead delivery experiences in **Japan/Singapore in <6 weeks**, coordinating dependencies across **20+ services** and **15+ teams** spanning Business, Operations, Product, UI/UX, Science, and Engineering
- Ensured availability and health for **multiple Tier-1 services** via a rotational on-call period and scaling review
- Actively mentor junior engineers, contribute to promotion feedback, and guide interns to successful return offers

Software Development Engineer I

Feb 2022 – Sep 2023

- Designed a new API powering “work availability, offer likelihood, and delivery area” maps with **<100ms latency** across **millions of on-demand offers globally** to accelerate driver matching and improve delivery transparency
- Improved new-driver offer acceptance by **~50%** with tenure-based route assignment via a new **event-driven architecture** that backfilled **1M+ records in 2 minutes** and cut daily recomputations from **2M+ drivers to 150K**
- Automated large-scale cancellations (**500K+**) due to station emergencies (weather, injury, fire), improving driver safety, cutting incident response from **3 hours to <30 minutes (-83%)** and saving **2,500+ hours / 4 FTE annually**

Software Engineering Intern

Jun 2021 – Aug 2021

- Developed a self-service tool for driver location transfers (**1k+/ticket**), eliminating a **150-hour/year** manual workflow

Charles Schwab, Cloud Delivery and Services

Austin, TX

Software Engineering Intern

Jun 2020 – Aug 2020

- Created an internal cloud-storage UI that moved engineering teams from direct S3 access to AuthN/AuthZ-backed APIs

The Oden Institute, Computational Materials Lab

Austin, TX

Teaching / Research Assistant

Aug 2018 – May 2019

- Led 10 mentors in teaching algorithm analysis, code optimization, and development practices to 50+ non-majors
- Worked with Dr. Juliana Duncan to evaluate nanoparticle structures for more efficient clean-energy storage systems

EDUCATION

University of Texas at Austin

Austin, TX

BS in Computer Science, BS in Mathematics – GPA: 3.7/4.0

Aug 2017 – Dec 2021

PROJECTS

Agent Relay, Amazon Hackathon — **2nd place (346 participants, 115 ideas, and 81 projects)**

Jan 2026

- Built a background-agent platform for one-click handoff and orchestration of code reviews, engineering tasks, and ticket workflows from Amazon internal tools to coding harnesses (e.g. Claude Code) in users' local/cloud environments

UT Registration Plus, Open-Source Contribution

Jun 2020

- Accelerated and streamlined the university course registration process by developing notification support for class detail changes through a Chrome extension with **20,000+ users, about 40% of the student body**

SKILLS

Languages: Java, Python, Go, C/C++, JavaScript/TypeScript, CSS, HTML, SQL

Technologies: GraphQL, REST APIs, React, React Native

Cloud: AWS (CDK, Bedrock, CloudFront, DynamoDB, ElastiCache, EventBridge, Fargate, Kinesis, Lambda, S3, SNS/SQS)

AI/ML: Model Context Protocol (MCP), Skills, TensorFlow, PyTorch, Scikit-learn, Pandas, NumPy, Matplotlib