

# Ying Wang

[me@yingw787.com](mailto:me@yingw787.com) | [linkedin.com/in/yingw787](https://www.linkedin.com/in/yingw787) | [github.com/yingw787](https://github.com/yingw787)

## Experience

---

### Caribou (formerly MotoRefi)

Senior Software Engineer, Data

Arlington, VA

July 2020 - Present

- Maintained continuity of finances for a \$1.1B company. Implemented 6-month project in 1 month by redefining company general ledger in native SQL from scratch, including fully reconciling ledger over 8 iterations with finance. Query pulls 100 fields with <1s latency, impacting 3k funded loans per month.
- Saved company relationship with Credit Karma (80% of top-of-funnel leads) by migrating inherited Apache Beam pipelines to Apache Airflow + GCP BigQuery SQL during internal Sev1 outage in 1 week. Increased IC capacity + removed future Sev1s by cutting back core data tooling from 3 to 2, and improved analyst experience by reducing data warehouse refresh latency from 6 hours to 10 minutes.
- Pushed organizational data culture early by founding and scaling the data engineering team from myself to 10 people (7 FTEs including EM, PM, 3 mids, 1 junior + 3 contractors) as company headcount grew 7x in one year, including writing and publishing job descriptions, evaluating technical screening tools, developing and validating multiple bespoke data practicums, interviewing >25 job candidates for {PM, EM, IC} roles, onboarding new team hires, mentoring/training junior ICs (promoted entry-level => mid), drafting a \$250k+ contract expansion engagement, etc.
- As tech lead, reviewed IC design and implementation of single-touch marketing attribution, offline ads conversion platform hydration (enabling 3x marketing spend w/ only 10% increase in CAC), data system health metrics (99.99% general pipeline uptime) for CTO/VP reporting, Airflow 1.x to 2.x migration, CI for Airflow using k8s + ArgoCD, data quality testing/alerting with bespoke logic + Monte Carlo, in-house Kafka-based streaming ETL framework, etc.
- Improved bottom-of-funnel candidate interview experience for all engineering by writing, testing, and publishing a bespoke Python coding practicum (popularity tied at #1 w.r.t. other practicums).
- Crafted industry-differentiating corporate culture via Culture Pod, developing “Humans of Caribou”, and co-leading wellness employee resource groups.

### Sabbatical (Tooling / Infrastructure Exploration)

Arlington, VA

October 2019 - July 2020

- Created personal dotfiles to formalize development processes: <https://github.com/yingw787/dotfiles>
- Achieved beginner-level proficiency with Haskell: <https://github.com/yingw787/thehaskellbook>
- Built TinyDevCRM, a hybrid OLTP/OLAP streaming BI framework composed of PostgreSQL, Apache Parquet, and HTTP/2 Server-Sent Events: <https://docs.tinydevcrm.com>
- Started Paint Creek Software (software contracting), >\$3k revenue in 2020 (built out a logo comparison PoC using React.js + AWS Mechanical Turk)

Bytes by Ying (<https://bytes.yingw787.com>)

Arlington, VA

- 70,000 page views and 40,000 visitors from ~150 countries with ~10% return rate

**Kinetica DB****Arlington, VA**

Software Engineer | Developer Tools, Data Engineering

February 2018 - October 2019

- Expanded Kinetica I/O (graph-based ETL tool for highly-structured / semi-structured data) functions:
  - Integrated Apache Spark as a data acceleration layer to boost data ingestion speeds by >2000x (1K records in 4 seconds → 20M records in 35 seconds)
  - Implemented engineering processes to shorten sustainable feature development lifecycle from 3 weeks to 1 day, enabling multiple \$1M - \$10M+ enterprise B2B commitments.
  - Implemented support for ESRI Shapefile ingestion supporting GIS use cases within Kinetica, part of the core product offering's market differentiation and value add.
  - Designed and implemented data-driven integration test harness in support of rapid test-driven development with support for test auto-generation, improving mean-time-to-repair by >3x.

**Xometry****Bethesda, MD**

Software Engineer | Full-Stack

December 2016 - February 2018

- Refactored machine learning pipeline for cost and win rate calculations, brought testing coverage to ~80%, and integrated individual testing coverage and linting reports with gated commits.
- Led a team of three SWEs to refactor cost-plus pricing algorithms, contributing to a ~80% speedup.
- Prototyped front-end expediting shipping feature, contributing to primary expedite shipping functionality generating \$10K in first week of production.
- Migrated parts of internal ERP software from Jinja / Angular / jQuery to React.js, boosting developer productivity 5x and improving page performance 3x.
- Migrated parts of landing page from WordPress to bespoke solution by developing Craft CMS templates, allowing sales/marketing to add site content without developer input while scaling to 4000+ requests per day.

**Education**

---

**Duke University** (Durham, NC)

2012 - 2016

B.S.E, Electrical and Computer Engineering (3.23/4.00)