Bill Zeng

Software Engineer · San Francisco Bay Area
415-867-5196 · billzeng2k@gmail.com · linkedin.com/in/billzeng2k

Summary

Product-focused software engineer who turns ambiguous needs into shipped solutions. Owns projects from problem discovery through implementation, rollout, and iteration, with strengths in backend systems and applied ML.

Experience

Persona 2022 – Present
Software Engineer, Identity Foundations (Full-time) · Software Engineer Intern (2021)

Product delivery & applied ML

- Led a **three-engineer team** to build and deploy an in-house OCR engine with accuracy matching the external provider, enabling replacement of a critical third-party dependency and direct control over model improvements.
- Deployed OCR on Kubernetes across **20 production environments**, each with dedicated compute to prevent resource contention across customer groups; right-sized compute using CPU/GPU benchmarks, lowering per-page costs by **approximately 86%** versus third-party OCR.
- Improved document extraction for ID verification used by **94% of customers**, helping win new business by expanding ID coverage to **175+ countries** and raising aggregate field-level accuracy by **15 percentage points**, measured through backtests and live extraction rates.
- Built an LLM extraction pipeline to address verification failures from missing fields, recovering **92% of missing values in backtests**; added model fallbacks and runtime kill switches for controlled rollout.
- Created ground-truth labeling and review tools for OCR and document-alignment models, enabling an approximately **7-percentage-point gain in alignment success** over the previous baseline, measured across **tens of thousands of labeled examples**.
- Fine-tuned OCR with curated real and synthetic data to correct spacing and character errors, improving extraction accuracy by **3 percentage points** over the base model in backtests across **hundreds of thousands of labeled fields**.
- Developed an automated classifier to identify a document's issuing country and type, achieving **99% accuracy** across tens of thousands of backtest examples.

Engineering platforms & reliability

- Proposed and led development of a pre-merge testing platform used **hundreds of times per month** to catch extraction and verification regressions before rollout by testing changes against historical user data in secure, isolated environments.
- Designed and built an A/B testing framework that guided model rollout and rollback decisions across **50+ experiments**, with per-customer impact dashboards and failure inspection tools to pinpoint affected ID card types.
- Cut backtest runtime **from days to hours** by building dedicated infrastructure with autoscaling workers and load-balanced extraction servers, removing processing bottlenecks and isolating backtest traffic from production.
- Re-architected the identity verification pipeline as a directed acyclic graph (DAG) to run processors in parallel, enabling LLM extraction while preventing per-verification latency from rising from **3 to 5 seconds**.
- Built a fuzzy-matching framework adopted by **12 teams**; unblocked customer integrations by supporting comparisons across verification types and languages through shared normalization, transliteration, and customizable matching rules.

Wish Winter 2021
Software Engineer Intern

- Added customs classification (HS) codes to shipping labels to support international deliveries.

Education

University of Waterloo 2018 – 2022
Bachelor of Computer Science *Average: 88%*
Canadian Computing Olympiad – **Silver Medalist**, 2018

Skills

Languages & frameworks: Python, Ruby, TypeScript, JavaScript, SQL, C++, C#, Ruby on Rails, React
Applied ML: Model fine-tuning, evaluation, OCR, LLM integration, ONNX
Infrastructure & deployment: Docker, Kubernetes, Google Cloud (GCP), AWS, Terraform, Argo CD, Datadog