

ALEXANDER J. DOE

Principal AI Engineer, Staff Software Engineer

San Francisco, CA | open to hybrid/remote

alexander.doe@example.com | (415) 555-0134

linkedin.com/in/alexjdoe | github.com/ajdoe | alexjdoe.dev

Professional Summary

Hands-on AI engineering leader with 12+ years across ML systems, distributed backend platforms, and productized AI. Built and scaled models and infrastructure for fraud detection, credit decisioning, search, recommendation, and document intelligence. Track record of shipping production systems that move revenue, latency, and risk metrics while mentoring teams and instituting MLOps practices.

Career Highlights

- Led delivery of a real-time fraud detection service that cut false positives by 33% at ~10,000 TPS with sub-40 ms P99 latency.
- Shipped a document intelligence pipeline that automated 78% of manual reviews, saving approximately \$4.1M per year in operating costs.
- Owned an end-to-end LLM-augmented retrieval system that increased self-serve resolution by 24% and improved NPS by 8 points.
- Designed a feature store and model registry enabling weekly model releases with rollback in under 5 minutes.
- Scaled the data platform to 4 PB using Delta Lake and Spark, with unit and regression tests for critical pipelines.

Core Skills

Languages: Python, TypeScript/Node.js, Go, SQL (PostgreSQL, Snowflake), Scala (Spark), Bash

ML/AI: XGBoost, LightGBM, CatBoost, PyTorch, TensorFlow/Keras, Hugging Face, ONNX, OpenAI/Anthropic APIs

NLP/GenAI: RAG (vector DBs), prompt engineering, finetuning (LoRA/QLoRA), evaluation harnesses

Data & MLOps: Spark, Airflow, DBT, MLflow, Feast, Great Expectations, Kafka, Delta Lake, Parquet

Cloud & DevOps: AWS (SageMaker, EKS/ECS, Lambda, S3, Step Functions, Glue), GCP (Vertex AI), Docker, Kubernetes, Terraform

Observability: Prometheus, Grafana, OpenTelemetry, Datadog, Sentry

Security/Compliance: PII tokenization, KMS/HSM, IAM, SOC 2, ISO 27001 controls

Experience

Principal AI Engineer, FinPay Technologies (Series C fintech)

2022 - Present | San Francisco, CA

- Architected a real-time fraud detection platform (streaming features + gradient-boosted trees + rules) handling ~10,000 TPS; reduced false positives by 33% and fraud loss by 18% year over year while keeping P99 latency under 40 ms.
- Introduced a feature store (Feast) and model registry (MLflow) with canary and shadow deployments; cut model release cycles from quarterly to weekly with blue-green rollback under 5 minutes.
- Delivered an LLM-augmented RAG workflow for merchant dispute triage; raised first-contact resolution by 24% and reduced average handling time by 31%.
- Led a cross-functional evaluation program: offline AUC/PR-AUC plus online guardrails (KS drift, PSI, safety); instituted bias audits and model cards with SHAP explanations.
- Mentored 7 engineers and 2 data scientists; formalized the interview loop and career ladders.

Staff Software Engineer, ML Platform, ShopHub

2019 - 2022 | Remote

- Owned ML platform roadmap: Kubernetes-based training and inference, feature pipelines on Spark/Delta Lake, and automated retraining via Airflow.
- Implemented a near-real-time recommendation service (approximate nearest neighbors + embeddings) that boosted CTR by 15% and revenue per session by 6.8%.
- Productionized a document-intelligence pipeline (layout-aware OCR + classifier + rules) to automate KYC/KYB verification; automated 78% of reviews and reduced SLA from 24 hours to 30 minutes.
- Built an experimentation service for multi-armed bandits and sequential tests; standardized success metrics and guardrails across product teams.

Senior Software Engineer, CloudScale Inc.

2016 - 2019 | New York, NY

- Designed event-driven microservices with Kafka and gRPC; improved platform reliability to 99.98% and cut mean incident time by 42%.
- Created a data quality framework (Great Expectations) with schema contracts and anomaly alerts, reducing bad-data incidents by 70%.
- Partnered with security on PII tokenization using KMS and structured logging; passed SOC 2 Type II audits.

Software Engineer, BlueOrbit

2013 - 2016 | Boston, MA

- Built high-throughput ETL pipelines on Spark and Airflow, migrating from cron-based jobs and cutting compute cost by 23%.
- Shipped search ranking improvements with learning-to-rank (XGBoost), lifting search conversion by 5.2%.

Selected Projects & Publications

- Open source: 'ledger-etl' (ETL toolkit for financial event streams, 2.1k stars), 'ml-canary' (Kubernetes inference canary controller).
- Patent: 'Adaptive Graph-Boosted Fraud Detection with Streaming Context Windows' (US-P-XXXX).
- Talks: 'Practical RAG for Regulated Industries' at PyData and MLOps World.

Education

M.S., Computer Science (Machine Learning), Northeastern University

B.S., Computer Engineering, University of Massachusetts Amherst

Certifications

AWS Certified Machine Learning - Specialty

Google Professional Machine Learning Engineer

ATS Keywords

Keywords for ATS: AI Engineer, Machine Learning Engineer, Principal Engineer, Staff Engineer, MLOps, LLM, RAG, PyTorch, SageMaker, Spark, Kafka, Kubernetes, Terraform, MLflow, Feast, Airflow, Delta Lake, Vector DB, Retrieval, Prompt Engineering