

Matt Coburn

Embedded enterprise delivery

Dallas, TX · mattcoburn.ai

I sit with enterprise customers, learn their workflows, and ship AI systems into their production environments — in regulated, data-heavy domains where correctness, grounding, and auditability are non-negotiable. I've put **agentic AI into production at 2 of the top 3 global automakers, built and deployed** document-intelligence and compliance products to **Fortune 100** customers (MasterCard, Fannie Mae, Freddie Mac), and shipped client-facing analytics to hedge funds including **BlackRock, Two Sigma, and Citadel**.

I'm hands-on end-to-end: **Python, TypeScript, C, Zig**, FastAPI, Postgres, Docker, Kubernetes, AWS — plus the retrieval, evaluation, and governance scaffolding that makes enterprise AI trustworthy. As a founder I went from zero to a **SOC 2-certified** product generating **\$2M** in revenue, which means I understand the whole arc: scoping with a customer, building the integration, deploying into their environment, and owning the outcome.

I do my best work at the boundary between product and customer — turning a hard, messy enterprise workflow into a system that ships and holds up in production. That includes designing RAG architectures and hybrid retrieval systems for grounded LLM applications, then validating them against the workflows customers actually run.

Aristotle — Founding Engineer & Tech Lead

Los Angeles, CA · June 2024 – Present

Founding engineer on an enterprise AI platform **in production at 2 of the top 3 global automakers**. Embedded deeply on deployment, integration, and customer workflows — building agentic systems that automate complex business processes inside regulated customer environments.

Key contributions:

- **Agentic systems in production:** Built tool-using agents as explicit state machines with boundary validation, retry/escalation semantics, and human-in-the-loop controls — designed to survive real customer environments, not just demos
- **Customer-deployed document intelligence:** Shipped a structured extraction pipeline with click-to-source provenance, schema validation, and auditable decision trails, integrated into customer document workflows
- **Grounding & governance guarantees:** No model output reaches a user or changes system state unless it is schema-valid, auditable, and traceable to source — the bar that unblocks enterprise adoption
- **Hybrid retrieval & evaluation:** BM25 + dense retrieval tuned on Precision@K / Recall@K, with regression harnesses wired in as release gates
- **Engineering foundation:** Hired and led 4 engineers; stood up CI/CD, code review standards, release gates, and observability

Tangible Intelligence — Founder & Lead Engineer

Dallas, TX · Jan 2020 – Apr 2023

Founded a document-intelligence company and personally built and deployed it into Fortune 100 customer environments. Generated **~\$2M enterprise revenue** with clients including **MasterCard, Fannie Mae, and Freddie Mac**.

- **Built and shipped the platform:** A no-code, human-in-the-loop extraction system with ontology-driven schemas and click-to-source provenance — configured per customer for lease processing, tenant applications, and screening documents
- **SafeScan — real-time compliance screening:** Built and deployed a sensitive-data detection and anomaly-screening product combining ML, rules, and search; ~\$900K in standalone revenue
- **Deployed into regulated environments:** SOC 2–certified deployments and Fortune 100 enterprise integrations, working directly with customer security and compliance teams
- **Domain depth:** Built AI systems for mortgage-backed-securities document workflows for Fannie Mae and Freddie Mac — extraction, classification, and compliance validation

WorkFusion — VP of Data Science

New York, NY · 2023

Owned the applied-AI roadmap for KYC/AML, document extraction, and screening — shipping compliance AI products into regulated financial customer environments with ~15 data scientists.

- **Compliance AI in production:** Transaction monitoring, risk scoring, and case prioritization for major financial institutions; ~35% improvement in detection outcomes
- **Document intelligence at volume:** LLM-based extraction and auto-labeling for high-volume screening with strict governance and explainability requirements
- **Production reliability:** Security-first MLOps, evaluation, monitoring, and audit practices for regulated deployments

M Science (Jefferies subsidiary) — Principal Data Scientist, Founding Lead

New York, NY · May 2018 – Dec 2019

Built and shipped a client-facing analytics platform to production for hedge-fund clients including **BlackRock, Two Sigma, and Citadel**.

- **Client-facing product:** Built an application layer that let institutional clients apply shared analytics to their **private datasets** with isolation and fine-grained access control
- **Ontology-as-a-Service:** Modeled companies, securities, events, and relationships across public and proprietary data
- **Reproducibility first:** Deterministic identity resolution and relationship semantics built for auditability

Expedia Group (Hotels.com) — Data Scientist

Dallas, TX · Oct 2016 – May 2018

Built large-scale revenue-management systems influencing **\$200M+ in annual advertising spend**.

- **Revenue optimization at scale:** Models estimating expected value of clicks across millions of keywords, driving automated bid optimization across publisher inventory
- **Measurable impact:** Reallocated spend against systemic underperformance for ~240% ROI improvement on targeted segments
- **Production pipelines:** Revenue-critical systems with monitoring, safeguards, and alerting

Earlier Career

Mortgage Industry · Dallas, TX · 2005 – 2008

Originated residential mortgages at Wells Fargo and an independent brokerage — deep, first-hand understanding of underwriting, compliance, and the document-heavy processes I'd later automate with AI.

Technology Stack

AI & platform: Claude, LangChain, LlamaIndex, PydanticAI, pgvector, Pinecone, OpenRouter, Microsoft Azure, Google Cloud Platform (GCP), Databricks, React

Core: Python, SQL, TypeScript, C, Zig

Backend & Infrastructure: FastAPI, Pydantic, PostgreSQL, Redis, Docker, Kubernetes, AWS

Data & Analytics: Pandas, NumPy, Spark, ETL/streaming pipelines, vector databases

AI/ML: PyTorch, HuggingFace, OpenAI/Anthropic APIs, agentic frameworks

Specialties: production AI deployment, compliance screening, structured extraction, AI governance, responsible AI, evaluation frameworks

Education

University of Texas at Dallas

- B.S. - Electrical Engineering, 2013
- Masters Coursework - Computer Science, 2016

What I'm Looking For

A forward-deployed / applied-AI engineering role where I embed with customers and ship AI into their production — agentic systems, document intelligence, and retrieval in regulated, high-stakes environments. I want to own outcomes end-to-end, from the first customer conversation to the system running in their stack.