

ABE FLANSBURG

Greater Houston | work@engabe.com | engabe.com | aflansburg (GitHub)

PROFESSIONAL SUMMARY

Principal Engineer with 12+ years in production software and platform engineering, focused since 2022 on AI systems for value-based care. Builds the clinical-decision and data-platform layers that turn EHR and claims data into reviewable clinical action: agentic care-plan generation, chronic-disease risk stratification, coding and quality-gap surfacing, and the durable runtime and canonical data model beneath them. Co-author of two arXiv papers on preference learning and reinforcement learning for clinical AI in value-based care, both drawn from production deployments. HIPAA/PHI-aware across the stack, shipping agentic systems into live clinical settings where a wrong inference has real consequences for patients.

CORE COMPETENCIES

- **Value-Based Care & Clinical Ops:** agentic care-plan generation, chronic-disease risk stratification, HCC/risk-adjustment and quality-gap surfacing, TCM/CCM touchpoint sequencing, clinician-in-the-loop review and approval gating
- **Data Platform & Modeling:** dbt canonical modeling, BigQuery, Snowflake, EHR/claims integration (athenahealth, HL7 FHIR), Airflow 3 / Cloud Composer orchestration, Terraform IaC, multi-tenant pipelines
- **Agentic Systems & Orchestration:** Temporal durable execution, condition-scoped subagents, clinical skills/tools, MCP interfaces, long-horizon reliability, BAML, FastAPI
- **Retrieval & Grounding:** end-to-end RAG pipelines, vector stores (Redis, Vertex AI), embedding, reranking, semantic search, context engineering
- **LLM Evaluation, Safety & PHI:** agent trajectory analysis, failure-mode and hallucination evaluation, guardrails, Phoenix (Arize) tracing, PHI de-identification (DLP, clinical NER), HIPAA/HITRUST
- **Cloud & ML:** Google Cloud (BigQuery, Vertex AI, Cloud Run, Cloud Composer, GKE, Pub/Sub) and AWS (ECS/Fargate, S3, RDS); Anthropic Claude, Vertex AI/Gemini, model fine-tuning (LoRA/QLoRA PEFT), feature engineering

PROFESSIONAL EXPERIENCE

Principal Engineer & AI Architect

July 2025 – Present

Altitude

- Contributed to **Praxis**, an agentic clinical-recommendation engine that generates guideline-grounded care plans as discrete, reviewable actions (MVAs): condition-scoped subagents for HTN, DM, CKD, and CHF; clinical tools for labs, vitals, medications, ASCVD Pooled-Cohort risk, and drug-interaction checks; cross-condition reconciliation; and clinician accept/reject gating on every recommendation. The platform also surfaces diagnosis and coding-gap suggestions and routes each action to an AI agent or a named clinical role (NP/PCP/care manager/pharmacist), sequenced into TCM (3/7/14-day) and CCM monthly touchpoints. Deployed to pilot health-system tenants.
- Built the **proof-of-concept for the previsit MVA system** and contributed heavily to the initial data platform (pipelines and infrastructure) that Praxis was later built on; spearheaded identifying and integrating the athenahealth API endpoints Praxis relies on for patient context.
- Built the **durable agentic runtime** on Temporal (worker plus agent-runner services), replacing an earlier proxy-based agent runtime with direct SDK integration; OpenTelemetry tracing, connection pooling, and environment-guarded workers running containerized on ECS/Cloud Run.
- Built the **clinical signal-extraction and feature-engineering pipeline** on Vertex AI Pipelines (KFP v2): Gemini batch extraction over unstructured encounter notes across seven chronic conditions (DM, HTN, CKD, obesity, dyslipidemia, CHF, ASCVD), then per-condition control-state classification and patient-level feature matrices for risk stratification. Pub/Sub-triggered, multi-tenant, incremental.
- Led the core data-model migration **off Tuva onto an in-house dbt architecture**: one dbt project per EHR source with config-driven onboarding, matrix Docker builds to Artifact Registry, Cloud Run dbt jobs orchestrated by Airflow into a canonical BigQuery model integrating athenahealth EHR entities.
- Built and cleared for open-source **altitude-athena**, a typed Python client and facade over the athenahealth API (chart, encounters, documents, population health) with a derived-data layer feeding the canonical model.
- Owned the GCP **data-platform infrastructure** as code (BigQuery, GCS, KMS encryption, Pub/Sub, Workload Identity Federation CI/CD) and **Airflow 3 / Cloud Composer** orchestration for multi-tenant BigQuery snapshots and dbt transforms.

- Built a four-stage **PHI de-identification pipeline** for an MIT research data share: column allowlist, deterministic scrub, Google DLP, then a Stanford clinical-NER GPU job (Cloud Run L4) with fail-closed coverage guards and keyed MERGE write-back.
- Drove the team's **LLM observability and evaluation** stack from self-hosted Phoenix (Arize) on AWS Fargate, to Vellum, to Braintrust today, on which the team built an internal abstraction; authored **braintrust-grep** (open source) to catch anomalies and hallucinations by diffing evaluation telemetry against ground-truth data.
- Built internal tooling: a Slack "second brain" on Anthropic managed agents backed by BigQuery and a 300+ page wiki, and an SSO-gated internal visualization service used across the team.

Full Stack Engineer | AI & Platform

April 2024 – July 2025

Remo Health

- Designed and fine-tuned Gemini Flash via Vertex AI (LoRA/QLoRA, 20-epoch PEFT) for a neurology Q&A system; curated a 300-sample labeled evaluation set targeting hallucination, false certainty, and out-of-scope clinical advice as explicit failure modes.
- Built durable AI workflow execution systems processing healthcare data with 99.9% uptime and HIPAA compliance.
- Designed end-to-end healthcare data pipelines integrating EHR systems with AI models for real-time clinical insight, and led MVP/PoC development for clinical decision-support tools in 4–6 week cycles.

Data Solutions Architect (Consultant)

March 2024 – May 2025

Tribunus Health

- Led Google Cloud adoption for enterprise healthcare data workloads, architecting scalable solutions on BigQuery, Dataflow, and Cloud Storage.
- Built data platforms with automated ETL pipelines for batch and streaming healthcare data processing.

Staff Platform Engineer – Data Platform

February 2022 – May 2024

Brightside Health

- Built a RAG-powered practitioner FAQ bot on Vertex AI/Gemini (embedding pipeline, vector store, similarity search for clinical knowledge retrieval) and implemented LLM observability and production tracing with Phoenix (Arize).
- Led platform infrastructure and DevOps, transitioning the team from product platform to data-systems architecture focused on healthcare analytics and ML operations.
- Deployed Kubeflow for ML pipeline orchestration and Camunda for workflow automation on GKE; migrated the enterprise data warehouse from Redshift to BigQuery and the platform from Aptible to GKE.

Staff Engineer

December 2018 – April 2022

Requis

- Architected platform solutions enabling rapid deployment of Ruby on Rails and Node.js applications across multiple environments; implemented infrastructure as code, reducing deployment time by 75%.
- Mentored development teams on system design, code quality, and technical decision-making.
- After handing off the full-stack platform lead role, moved into data engineering and data science work supporting the platform.

eCommerce Software Engineer

October 2017 – December 2018

Rough Country

- Built multi-channel eCommerce tooling (Magento, BigCommerce, Amazon Seller, eBay APIs) and a Node.js order-fulfillment analytics tool for real-time inventory and shipping insight.

Earlier: Technical Services Specialist, Local Government Corporation (2016–2017) · Senior Network Engineer, ERMCO Inc (2016) · Data Engineer, MarkITx (2014–2016), where he built early applied-ML and data packages in Python and Node.js at a VC-backed IT marketplace startup.

PUBLICATIONS

- Singh, P., Gupta, A., Betz, C., **Flansburg, A.**, Ives, B., Lama, S., & Son, J.H. (2026). *Learning from Disagreement: Clinician Overrides as Implicit Preference Signals for Clinical AI in Value-Based Care*. arXiv:2604.28010.
- Singh, P., Gupta, A., Betz, C., **Flansburg, A.**, Ives, B., Lama, S., & Son, J.H. (2026). *Learning to Compress Time-to-Control: A Reinforcement Learning Framework for Chronic Disease Management*. arXiv:2605.09818.

NOTABLE PROJECTS

braintrust-grep — LLM log search for hallucination detection · github.com/aflansburg/braintrust-grep · PyPI · Open Source

- CLI and Python library for regex and structural search over Braintrust evaluation logs, filling a gap in BTQL (no regex, silent misses on structured JSON). Composable predicates for hunting hallucinations, false certainty, and out-of-scope outputs; 84% coverage, fully offline test suite, trusted PyPI publishing via GitHub Actions.

RAG-in-a-Box — visual RAG pipeline debugger · github.com/aflansburg/raginabox · Open Source

- Browser tool that makes every RAG stage inspectable (loading, chunking, embedding, similarity search, context assembly, generation). FastAPI backend, LM Studio local embeddings, in-memory cosine-similarity store.

Open-Source Contributions

- **ScrapeGraphAI** (28k+ stars) — contributor to one of the most-starred LLM-powered extraction frameworks on GitHub.
- **Hashie** (3k+ stars, Ruby) — contributor to the widely used Ruby hash-extension library.
- **incidentbot** (164+ stars) — contributor to the Slack-based incident-management bot.
- **status-trend** — Go terminal dashboard (Bubble Tea, Lip Gloss) that visualizes outage and incident trends across vendor status pages. Additional OSS: bearhugs, amzreviewsscrape.

TECHNICAL SKILLS

- **Languages:** Python, JavaScript/TypeScript, Ruby, Go, C#, Java
- **Data:** BigQuery, Snowflake, Redshift, dbt, Airflow, Dagster, Prefect, athenahealth, Dataflow, Tuva, ETL/ELT
- **Healthcare:** HL7 FHIR, EHR/EMR (Medplum, Oystehr, athenahealth), CDS Hooks, Google Healthcare API, HCC/risk-adjustment coding, HIPAA/HITRUST
- **Backend & Infra:** FastAPI, Ruby on Rails, Node.js/Nest.js, Docker, Kubernetes, Helm, Terraform, Temporal, Camunda, GitHub Actions, CircleCI
- **AI/ML & MLOps:** Anthropic Claude, OpenAI, Vertex AI (Gemini, MedLM), LangChain/LangGraph, BAML, Kubeflow, RAG, vector stores, Phoenix (Arize), LoRA/QLoRA fine-tuning

EDUCATION

Post Graduate Certificate, Data Science & Business Analytics — The University of Texas at Austin, 2020–2021

B.S., Computer Information Systems — DeVry University, 2014–2016

Undergraduate Studies, Psychology — Middle Tennessee State University, 2002–2004