

EMILIO GAGLIARDI

emilio.gagliardi@gmail.com • 780.885.9850 • emilio-gagliardi.ca

AI engineer and platform architect who takes retrieval-augmented AI systems from first commit to paying customers. Currently leading technical development at Policy Solver: a multi-tenant RAG platform through which municipal governments query their bylaws, policies, and procedures and receive cited, auditable answers. Sole architect of its data platform — relational, vector, and graph databases working as one retrieval system — and of the Azure ecosystem that runs it. Trains and ships the platform's ML models. Grounded in experimental science: a research master's built a lasting practice of hypothesis testing, calibrated thresholds, and evidence over intuition. A decade of founding and running technical ventures shapes how standards get written, costs get governed, and other developers get direction.

EXPERIENCE

Policy Solver | Lead AI Engineer

Jun 2025 – Present

Contract Jun 2025; full-time Jan 2026. Lead technical development of a multi-tenant, retrieval-augmented AI platform for municipal governance — five production repositories, ~540K lines of code, 2,500+ commits.

- Designed and operate a data platform spanning three database classes: **PostgreSQL/Supabase** with row-level-security tenant isolation and RPC-driven task queues (176 migrations); **Zilliz/Milvus** and **Qdrant** vector stores running hybrid dense + BM25 sparse retrieval fused by reciprocal-rank fusion; and a **Neo4j knowledge graph** synchronized from Postgres through typed column contracts that make schema drift a build failure rather than a production incident.
- Built the document-intelligence pipeline end to end on **Dagster Cloud**: SharePoint ingestion via Microsoft Graph delta sync, Azure Document Intelligence parsing, semantic chunking, summarization, entity extraction, and embedding — sensor-driven, with lazy task fan-out, adaptive concurrency, and automated replica scaling.
- Designed the **graph retrieval framework**: a rule-based query planner over Neo4j with tenant guards, circuit breakers, and a structured error taxonomy, fused with vector retrieval in a three-lane RRF architecture.
- Trained and deployed production ML: a **SetFit/ONNX section classifier** with confidence-thresholded heuristic fallback; a **character-trigram language model** trained on 34M characters that gates heading quality; and spaCy out-of-vocabulary scoring — each calibrated against production corpora, with classification provenance recorded per prediction.
- Built the original **SharePoint implementation** (content-type schemas, metadata governance, Graph API provisioning and tenant onboarding tooling) and the original **Azure application ecosystem** — Container Apps, Blob Storage, Service Bus, Key Vault, Document Intelligence, Entra ID — integrating every service into one platform.
- Authored the engineering system other developers build within: 13 enforced coding standards, a spec-driven architecture process with **212 ratified proposals**, cross-repository error-code taxonomies, and the operational runbooks that make the platform runnable by people other than its author.
- Diagnosed and remediated a **\$9,944 cloud-billing defect** (47,000 runaway pipeline runs), then designed the token-ownership contract and cost caps that prevent recurrence — cost governance is now a documented, load-bearing invariant of the ingest pipeline.
- Instrumented the platform with **Langfuse** tracing, structured logging, and per-stage observability from document ingestion to cited answer.

PortalFuse | Data & AI Consultant

2023 – 2025

Built the automated security-intelligence reporting product for a Windows patch-management SaaS.

- Engineered **Kedro** ingestion-transformation-publication pipelines aggregating CVE and KB-article data from the National Vulnerability Database and the Microsoft Security Response Center across Windows and browser products.
- Designed and automated a weekly, monthly, and quarterly reporting cadence. The quarterly edition analyzes ~250 CVEs per quarter — CVSS v3.1 scoring, CWE categorization, severity and exploitability distributions across 19 visualizations — and closes with strategic guidance addressed to CISOs.
- Delivered dual-perspective analysis — CVE-centric risk and KB-centric remediation — so administrators move from threat to patch inside a single document.

Wild Rose College | Technical Lead

2021 – 2023

- Directed technology across ecommerce, LMS, and marketing systems; rebuilt the site theme spanning WooCommerce and LearnDash; advised on IT budget and vendor selection.

- Designed the ActiveCampaign automation integrating WooCommerce and LearnDash that tracked and communicated with thousands of customers and drove **\$3.5M in first-year sales**.

Earlier career | Founder & Owner

2011 – 2022

Founded and ran **Uplyft** (2015–2022), a web-development consultancy modernizing retail clients' web properties and automation, and **Eco Evolver** (2011–2015), a controlled-environment agriculture design and education company — including a fully automated 25' × 7' living-wall installation for an elementary school.

EDUCATION & CREDENTIALS

MSc, Psycholinguistics — University of Alberta. Experimental design, statistical inference, and computational study of language — the research foundation of the data science practice.

BA (Hons), Psychology — University of Alberta

Microsoft Professional Program in Data Science (2019) • **DataCamp Data Scientist Associate** (2023)

TECHNICAL SKILLS

AI / ML: RAG architecture, hybrid retrieval (dense + sparse, RRF), Haystack, LiteLLM, prompt engineering and Langfuse prompt management, SetFit, ONNX, spaCy, BGE-M3 embeddings, cross-encoder reranking, intent classification, MCP server development

Data platforms: PostgreSQL/Supabase (RLS multi-tenancy, RPC design, migrations), Neo4j/Cypher, Milvus/Zilliz, Qdrant, Redis, Azure Blob Storage

Orchestration: Dagster (Cloud Serverless), Kedro, Airflow, Azure Data Factory

Cloud & DevOps: Azure (Container Apps, Service Bus, Document Intelligence, Key Vault, Entra ID, Log Analytics), Docker, GitHub Actions CI/CD, pytest, mypy, ruff

Languages & APIs: Python, SQL, JavaScript/TypeScript; FastAPI, Microsoft Graph, SharePoint provisioning, REST

Observability: Langfuse tracing and cost tracking, structured logging (structlog), error-code taxonomy design, production runbooks