

Jonathan Abele

jonathanabele17@gmail.com | (248) 766-9841

EXPERIENCE

FactSet

Software Engineer I

Software Engineer II

Boston, MA

March 2025 - March 2026

April 2026 - Present

- Improved and maintained a high-throughput **NLP data pipeline** processing **hundreds of thousands of articles daily**, supporting company tagging, ESG classification, and sentiment analysis
- Leveraged **RabbitMQ and AWS services (EC2, Lambda, DocumentDB)** to support scalable, event-driven data processing across multiple pipeline stages
- Enhanced an internal web platform used by content curators to review and classify articles, adding dashboards and **data visualizations** to improve visibility into ingestion trends and curator performance
- Conducted large-scale data analysis using **Databricks** to evaluate pipeline performance, identify data quality issues, and uncover inefficiencies in article processing workflows

LLM-based Subtopic Autocuration Project

- Led end-to-end development of an **AI-driven validation system** serving as the final stage in a large-scale production pipeline, improving data quality and system reliability
- Designed a dynamic **prompt-driven classification system** (topic definitions, keywords, examples) with a three-state decision framework (approve, exclude, unsure)
- Deployed to production via API, processing **~20,000 articles/day** and significantly reducing manual curation workload while improving pipeline precision

FactSet Sidecar – Hackathon Project, Selected to demo at FactSet Focus, Austin TX

- A Chrome extension that tags companies in any article and shows live FactSet data in a side panel
- Built the named-entity recognition API integration feeding matched entities to the iframe panel

Data Solutions Intern

May - August 2024

- Built an **AI-powered chatbot** that interprets user queries, retrieves relevant information from large document sets, and generates accurate responses
- Developed a Flask-based backend API and a dynamic frontend using Vue.js
- Implemented a **RAG architecture** integrating OpenAI models with a Databricks vector database for semantic retrieval

Sentinel - AI PR Review System

Personal Project

May 2026 - Present

- Built a GitHub PR review system (Python + FastAPI) with a **multi-step agent pipeline** that posts inline comments
- Designed **parallel reviewer agents** (security, correctness, testing) with a consolidator that ranks findings by confidence
- Developed observability and replay: OpenTelemetry traces, Prometheus metrics, cost tracking, Postgres snapshots

EDUCATION

University of Michigan

B.S.E. Computer Science, Minor in Mathematics

Ann Arbor, MI

December 2024

Summa Cum Laude

GPA: 3.8 / 4.0

Honors: James B. Angell Scholar, William J. Branstrom Freshman Prize

SKILLS

Languages:	Python, C++, C, JavaScript, SQL
Frameworks & Tools:	Flask, Vue.js, FastAPI, Pydantic, Git
AI / Data:	LLM Applications, Model Evaluation, RAG, NLP, Vector Databases
Infrastructure:	AWS, Databricks, Docker, Terraform
Databases:	MongoDB, PostgreSQL
Developer Tools:	Claude, VS Code, macOS, Windows