

# Ethan Bailie

+1(289) 221-0488 | [eabailie@uwaterloo.ca](mailto:eabailie@uwaterloo.ca) | [www.linkedin.com/in/ethan-bailie](https://www.linkedin.com/in/ethan-bailie) | [ethanbailie.github.io](https://ethanbailie.github.io)

## Technical Skills

---

**Programming Languages:** Python, R, SQL, LookML

**Data:** Snowflake, dbt, Looker, PostgreSQL, Mage, Prefect, SQL Server, Spark, Requests, Pandas, Polars

**AI and ML:** LangChain, OpenAI API, Cohere API, Pinecone, JAX, NumPy, PyTorch, TensorFlow, scikit-learn

## Experience

---

### Data Engineer

June 2023 – Present

*Stamped.io*

*Vancouver, British Columbia, Canada (Remote)*

- Engineered an **LLM-driven RAG system** using **OpenAI, Cohere, LangChain, and PineCone**, creating value out of immense amounts of text data
- Improved and maintained **dynamic data pipelines** utilizing **Python, Snowflake, dbt, and SQL**, resulting in a **30% reduction in errors** in critical data pipelines
- Expertly crafted comprehensive **reporting dashboards** with **Looker**, swiftly identifying trends in **critical KPIs** like churn rate and MRR
- Developed **reverse ETL pipelines** with **Segment**, providing enriched data to tools utilized by **key stakeholders**
- Implemented the seamless integration of **Mage**, resulting in a **20% increase in pipeline efficiency** and spearheading a comprehensive data infrastructure overhaul

### Geospatial Data Scientist

January, 2022 – December 2022

*TD Bank*

*Toronto, Ontario, Canada*

- Engineered **data models** using robust **big data ETL pipelines**, combining **SQL, Python, and Alteryx** to identify and prioritize **growth opportunities** in key markets, **generating \$1 million in new leads**
- Crafted interactive **web applications**, analyzing demographic and competitor location data to **identify untapped market potential** and prospective customers, **improving acquisition efficiency by 15%**
- Supported various projects by **accessing and querying data** from data marts and warehouses using **SQL**, collaborating effectively with **cross-functional teams**
- Presented complex **analytical insights** through visually appealing **dashboards**, catering to non-technical executives

### Geospatial Data Analyst

May, 2021 – August, 2021

*Department of National Defence*

*Ottawa, Ontario, Canada*

- Optimized satellite imagery database** by automating the rebuilding process with **Python scripts**, enhancing its utilization across the Department
- Conducted thorough **analysis of satellite imagery and elevation data**, creating realistic simulations for military helicopter pilot training
- Implemented and documented **data standards** and best practices to ensure the integrity and maintenance of databases

## Projects

---

### Paper Recommender

2024

*Personal AI Project*

- Stores abstracts from arXiv in **Postgres**, embeds them into **Pinecone**, and returns the most 'interesting' articles based on my own preferences (sorted using **content-based filtering and Cohere's Rerank**)

### Simple Document Chat

2023

*Personal AI Project*

- Simple application that allows you to chat with uploaded documents to retrieve relevant information. Built using **OpenAI's chat and embedding, Pinecone, and Streamlit** for a user interface

## Education

---

### University of Waterloo

*Bachelor's of Environmental Science, Computer Science Minor*

*Waterloo, Ontario, Canada*