

# Emad Gohari

Senior Machine Learning Engineer · ML Systems, NLP, LLMs — Toronto, ON, Canada

## Contact Info

**Email:** E.Gohari.B@gmail.com

**LinkedIn:** linkedin.com/in/emad-gohari

**GitHub:** github.com/emadg

**Website:** EmadGohari.com

## Summary

Senior machine learning engineer with 8 years of experience building ML, NLP, and LLM systems from research to production. At Thomson Reuters Labs, shipped production GenAI summarization and semantic search, contributed agent configuration and evaluation tooling to a commercially launched agentic AI product, and built model-retraining pipelines and PII-secure data infrastructure for CLEAR entity resolution. Combines hands-on model development with reproducible experimentation and evaluation, working with researchers, domain experts, and product teams.

## Technical Skills

**LLMs & Agents:** LLM Agents & Tool Use, Agent Harness Engineering, Agentic Frameworks (PydanticAI), Multi-Agent Orchestration, Prompt & Context Engineering, LLM Evaluation (LLM-as-Judge, SME Review), Grounding & Source Citation, Model Selection (Quality, Latency, Cost), GPT-4, Claude, OpenAI API

**Machine Learning & NLP:** Machine Learning, Deep Learning, NLP, PyTorch, Hugging Face Transformers, Transformer Fine-Tuning (T5), Model Training & Evaluation (F1 Score, ROC-AUC, MRR, MAP, Recall@k), Text Classification & Text Tagging, Clustering Models, Sentence Embeddings, Semantic Search, Information Retrieval, Knowledge Extraction & Representation (Knowledge Graphs, Ontologies), LSTM, scikit-learn, XGBoost, spaCy, NLTK, pandas, NumPy

**MLOps, Data & Cloud:** MLOps, ML Pipelines, Experiment Tracking (MLflow), Data Pipelines, AWS SageMaker (Studio, Pipelines), Apache Spark (PySpark), Snowflake, AWS (S3, EMR), AWS CloudFormation, PostgreSQL, Elasticsearch, OpenSearch

**Languages & Practices:** Python, SQL, Java, Scala, Bash, Agile Development

## Experience

### Senior Machine Learning Engineer II (Research) at Thomson Reuters

Jul 2021 - Present

Toronto, Canada

Applied ML at TR Labs for CLEAR, Thomson Reuters' due-diligence platform for KYC/AML screening and investigations, and Checkpoint, its tax research platform. Built LLM features and agent workflows, search models, training pipelines, and secure data infrastructure with researchers, subject-matter experts, and product engineering. Used SQL/PostgreSQL for entity-resolution data analysis and name/address candidate blocking before downstream ML matching and ranking.

## Work Projects

### • Entity Resolution Retraining & Snowflake Pipelines: CLEAR — 2026 | Internal Tool

Designed and built an on-demand model-retraining pipeline on SageMaker Pipelines, turning researchers' feature-generation, training, and evaluation code into reusable DAG steps. Enabled parallel experiments tracked in MLflow, cutting experiment turnaround by an estimated ~40%. Built proof-of-concept Snowflake processing and training pipelines and assessed the gaps for migrating AWS research workflows to Snowflake.

**Tech & Skills:** Python, AWS SageMaker Pipelines, MLflow, PySpark, Snowflake, XGBoost, Poetry

- **Agentic Investigation System: CLEAR Investigate** — 2025 - 2026 | Production

Contributed to CLEAR Investigate, Thomson Reuters' commercially launched agentic AI product for investigations. Built a central PydanticAI configuration layer for system prompts, tool descriptions and availability, and agent/subagent wiring, enabling experimenters to reconfigure agents from one place instead of editing scattered definitions. Migrated the team's existing LLM-as-judge evaluation capability into an organized repository for reuse and future CI/CD integration.

**Tech & Skills:** Python, PydanticAI, Claude, Agent Harness Engineering, Multi-Agent Orchestration, LLM-as-Judge Evaluation, Prompt Engineering

- **CLEAR Business AI: GenAI Report Summarization** — 2024 | Production

Built the LLM summarization workflow for CLEAR Business entity reports, live in production processing hundreds of reports daily. Designed selective XML extraction for targeted per-section context covering business overview, personnel, adverse media, and legal records. Integrated Bing Search and LLM verification to reduce same-name entity mismatches and linked claims to report sources. Refined extraction and prompts through SME annotation rounds against exemplars for 10-15 companies, and compared GPT-4 and Claude Sonnet on quality, latency, and cost to inform model selection.

**Tech & Skills:** Python, GPT-4, Claude Sonnet, Bing Search API, Prompt & Context Engineering, LLM Evaluation, XML Processing

- **Entity Resolution Data Infrastructure: CLEAR KYC/AML** — 2022 - 2024 | Internal Tool

Built research data infrastructure for CLEAR's entity-resolution system spanning ~800M entities and billions of documents, with a 700 idents/second target. Designed a PII-isolated AWS environment with SageMaker Studio, CloudFormation, and per-researcher profiles; defined a versioned schema reconciling two incompatible annotation sources, with Spark and Python pipelines to merge records and surface label conflicts; and replaced manual runs with scheduled, repeatable SageMaker Pipelines.

**Tech & Skills:** Python, Apache Spark (PySpark), AWS SageMaker, AWS CloudFormation

- **Semantic Search Improvement: Checkpoint Tax Research** — 2021 | Production

Shipped a sentence-embedding query-intent classifier for Checkpoint tax research, distinguishing broad conceptual queries from specific tax lookups to promote higher-level documents. Improved access to conceptual documents by 95% within a ~100ms latency budget. Identified the retrieval gap through SME and customer feedback.

**Tech & Skills:** Python, Sentence Embeddings, Text Classification, Semantic Search

## **NLP & Machine Learning Engineer at INAGO Inc.**

Jul 2018 - Jun 2021

Toronto, Canada

Built NLP and language-understanding systems end to end, from data preparation and model training to evaluation and application integration, including transformer fine-tuning and research collaborations with university and industry partners.

### **Work Projects**

- **Automated Question Generation from Documents** — 2020 | Research

Fine-tuned the open-weights T5 model to generate questions from documents, cutting manual data-curation effort by 40%. Experimented with input representations and evaluated generated text with BLEURT in a research collaboration with university partners.

**Tech & Skills:** Python, PyTorch, Hugging Face Transformers, T5, BLEURT

- **Domain-Specific Language Understanding** — 2019 | Research

Trained domain-specific Word2Vec embeddings and LSTM-based NLU models, with interpretability testing to make model decisions transparent.

**Tech & Skills:** Python, PyTorch, Word2Vec, LSTM

## • **Conversational Question Answering with NLP & LSTM Models** — 2018 | Research

Developed a conversational question-answering system combining syntactic and semantic document analysis with LSTM-based text tagging. Used semi-automatically generated ontologies to interpret user questions and identify intent, and experimented with knowledge graphs for knowledge representation in an industry research collaboration.

**Tech & Skills:** Python, spaCy, LSTM, Text Tagging, Semi-Automatic Ontology Generation, Knowledge Graphs

## **Education**

### **M.Sc. in Computer Science, York University**

2014 - 2018

Toronto, Canada · GPA: 8.17 / 9

Thesis: Interactive Question Answering Using Frame-based Knowledge Representation

### **B.Sc. in Computer Engineering, Amirkabir University of Technology**

2010 - 2014

Tehran, Iran · GPA: 17.18 / 20

## **Publications**

### **Question-worthy Sentence Selection for Question Generation**

Canadian AI 2020, 2020 · co-authored

### **Interactive Question Answering Using Frame-based Knowledge Representation**

York University M.Sc. Thesis, 2018

### **Time Aware Topic-based Recommender System**

Big Data & Information Analytics, 2016 · co-authored

### **A Study on Prediction of User's Tendency Toward Purchases in Websites based on Behavior Models**

Information and Knowledge Technology (IKT), 6th Conference, IEEE, 2014 · co-authored