

# Uche J. Okechukwu, M. Eng.

## Data Scientist & Machine Learning Engineer

✉️ uche.okechukwu@gmail.com

📞 +1 825 365 4543

📍 Calgary, Canada

🔗 <https://ucheokechukwu.streamlit.app/>

LinkedIn <https://www.linkedin.com/in/uche-okechukwu>

Github <https://github.com/ucheokechukwu>

Medium <https://medium.com/@uche.okechukwu>

### Profile

---

A Data Science and Machine Learning Engineer with a knack for delivering solutions on a wide variety of DS/ML topics. From developing an LLM-based Ted Talk Recommender Engine to delivering rigorous Data Research Studies for Performance and Marketing, my skills are as versatile as my enthusiasm for tackling new challenges. Experienced with working in diverse and cross-functional teams and eager to bring my knowledge, experience and enthusiasm to your company to solve problems and generate value for clients.

### Skills

---

#### LLM Dataset Generation

Prompt Engineering | Rubrics Benchmarking | Golden Response Generation

#### LLM Output Integrity Assessment

Rubrics Evaluation | Metrics-based Evaluation | Chain-of-thought Fine-tuning | Reinforcement Learning Through Human Feedback (RLHF)

#### Programming

Python | SQL | R | GraphQL | Javascript | HTML | CSS | Bash

#### Natural Language Processing

Hugging Face Transformers | LangChain | Llama | OpenAI

#### Data Visualization

Tableau | PowerBI | Matplotlib | Seaborn | Plotly

#### Data Collection and Wrangling

Requests | Selenium | Beautiful Soup

#### Machine Learning Libraries

Tensorflow Keras | PyTorch | Pandas | Numpy | SciKit Learn | Jupyter

#### UI / UX

Streamlit | Chainlit

#### Cloud Computing

AWS | GCP

#### Engineering

Electrical | Instrumentation | Process Automation and Control | Project Management | Ethics & Transparency

### Recent Projects

---

2024/08

**SQL Chatbot: Using LangChain and OpenAI to query AWS-hosted SQL server** ↗

2024/03

**Yelp Chatbot on Streamlit: LLM RAG chatbot that queries structured & unstructured data on Yelp open datasets. Data hosted by Neo4js Graph DMS. Also integrates routing API** ↗

2023/11

**Fact Checker Study: Building 5 LLM-claims datasets to assess the performance of an NLP-based fact-checker tool for Originality.ai** ↗

2023/11

**Yelp Review Study: Collecting data on Yelp Reviews and conducting experiments for Originality.ai** ↗

2023/11

**Stock Prediction with Deep Learning and Reinforcement Learning** ↗

## **Education**

---

2023 – 2023	<b>Diploma in Data Science</b> <i>Lighthouse Labs</i>
2017 – 2021 Lowell, United States	<b>Electrical Engineering, Ph.D</b> <i>University of Massachusetts</i>
2016 – 2017 Lowell, United States	<b>Electrical Engineering, M.Eng.</b> <i>University of Massachusetts</i>

## **Publications**

---

2019	<b>A Less Convoluted Approach to 3D Pose Estimation</b> ↗ <i>IEEE International Conference on Healthcare Informatics</i>
------	---

## **Professional Experience**

---

2024/01 – Present	<b>AI Engineer</b> <i>Outlier. Mercor. SuperAnnotate</i> <ul style="list-style-type: none"><li>Developed and refined prompts for training and fine-tuning LLMs.</li><li>Created evaluation rubrics to guide the training and assessment of LLMs.</li><li>Evaluated LLM model outputs via rubrics benchmarking and metrics.</li><li>Generated benchmark-quality golden outputs using specialist and generalist knowledge to serve as ground-truths for LLM model training datasets.</li><li>Audited and refined chain-of-thought reasoning by identifying errors and optimizing reasoning steps to ensure logical consistency and correct final outputs.</li><li>Provided subject matter expertise in engineering and coding to support data generation and annotation for LLM training.</li><li>Integrated ethical compliance and social accountability in LLM training datasets.</li></ul>
2023 – 2023 Canada	<b>Data Scientist</b> <i>Originality.ai</i> ↗ <ul style="list-style-type: none"><li>Planned and delivered research studies: data collection, wrangling and analysis, report writing and presentation.</li><li>Developed datasets via web scraping with Requests and Selenium, LLM prompting and SQL querying.</li><li>Designed and executed experiments: hypothesis testing, defining metrics.</li><li>Optimized NLP models via Regex and Hugging-face transformer models.</li><li>Created technical reports: visualizations of research findings with tools like Matplotlib, Seaborn and Plotly.</li></ul>
2017 – 2021 United States	<b>Research and Teaching Assistant</b> <i>University of Massachusetts</i> <ul style="list-style-type: none"><li>Designed and executed own research on Human Pose Estimation, presenting and publishing findings, providing peer review to other team members.</li><li>Directed the undergraduate laboratory – overseeing classes, assessing students, and general administration of the labs and equipment.</li></ul>
2004 – 2014 Nigeria	<b>Senior Instrument / Project Engineer</b> <i>Shell Petroleum</i> <ul style="list-style-type: none"><li>Designed front-end engineering of process control and safety systems for several facilities.</li><li>Developed technical requisitions, technical assurance tasks like fire and gas mapping, and inspected and commissioned plant-wide control system.</li><li>Managed project engineering: contractor management, stakeholder engagement, social impact assessments, scheduling, budgeting, and health and safety requirements.</li></ul>