Arghya Mukherjee

arghyamukherjee1991@gmail.com | +1 918 330 1803 | arghyam.me | linkedin.com/in/arghyamukherjee | github.com/amukherjee1991

Summary

Engineer who ships AI systems end to end and sits with the people who use them. 8+ years building retrieval, search, and machine-learning products in production: agentic RAG over messy document corpora, real-time inference under 500 ms, and the pipelines, monitoring, and CI/CD that keep them running on GCP and Azure. As comfortable debugging an integration as presenting it to executives, clients, and regulators: briefed the U.S. SEC and a state Attorney General's Office.

Core Skills

AI & Retrieval: Agentic RAG assistants, LLM application design, semantic search and matching, embeddings and NLP (BERT, LLaMA), ontologies and knowledge graphs (JSON-to-RDF), model evaluation; anomaly detection, deep learning (PyTorch, TensorFlow)

Production Engineering: Real-time inference and ML services, API and backend integration, monitoring, drift detection, automated retraining, incident response, CI/CD, Docker, Kubernetes, MLflow

Data & Cloud: Python (pandas, scikit-learn, PySpark), SQL, R, Spark; GCP and Azure; BigQuery, Snowflake, Apache Kafka, Airflow; feature and data pipelines

Customer-Facing: Technical demos and prototypes, executive and regulator briefings, stakeholder workshops, A/B experimentation, self-serve dashboards (Power BI, Tableau, Plotly/Dash)

Experience

State of Oklahoma — Office of Management & Enterprise Services Jan 2025 – Present
Senior Data Scientist (ML / AI Engineer)
- Built and deployed an **agentic RAG assistant** that automates document review and risk flagging over large, unstructured legal and transportation corpora, replacing manual review.
- Own the **full production lifecycle**: real-time inference, monitoring, drift detection, automated retraining, and audit-ready reporting.
- Work directly with non-technical stakeholders to define success metrics and ship **self-serve dashboards**, and mentor analysts on reproducible modeling practice.

Chevron (via Capgemini America) Nov 2025 – Apr 2026
Senior AI / ML Consultant
- **Embedded with an enterprise client's engineering team** to take production ML from prototype to deployment across GCP and Azure, **cutting manual review 73%**.
- Delivered the MLOps and governance their auditors would accept — versioning, monitoring, canary deployments, documentation — plus the feature pipelines and CI/CD behind them.

Anti-Phishing Working Group (APWG) May 2024 – Present
Data Scientist / AI Engineer — Fraud & Threat Detection
- Engineered a **phishing data ontology (JSON-to-RDF)** from behavioral and threat signals that powers retrieval, reasoning, investigation, and case linkage for member organizations.
- Built classification and anomaly-detection models flagging phishing and fraud across **20+ years of data at 97% accuracy**, with per-alert confidence for risk-ranked triage.

Workonward Inc. Jun 2024 – Jan 2025
ML / AI Engineer (Remote)
- Shipped **semantic-matching and classification services in real time**, median latency under 500 ms, with Git-based CI/CD on GCP.
- Ran controlled **A/B experiments** against product metrics, and owned feature engineering, pipelines, and deployment across backend services and cloud infrastructure.

University of Tulsa Jan 2018 – Dec 2024
Research Data Scientist — Financial-Crime Modeling
- Led doctoral research modeling financial crime and market manipulation over large market and transaction datasets.
- **Briefed the U.S. SEC and a state Attorney General's Office** on findings, informing regulatory and investigative action.

Robbie AI — Data Scientist (2023–2024) — real-time computer-vision detection model, accuracy improved 23% at sub-second latency.

PwC — IT Risk Advisory Associate (2015–2016) — client-facing risk assessments and remediation roadmaps for Fortune 500 clients.

Education

Ph.D., Computer Science, University of Tulsa — *Cyber Security & Econometrics* 2024
M.Sc., Computer Science, University of St Andrews | **B.Tech., Computer Science**, WBUT 2014 / 2012

Recognition: Six first-author publications (*Information Processing & Management*, HICSS, ACM CCS DeFi, APWG eCrime).