# AKSHAY AGGARWAL

/Akshayanti   in/akshayaggarwal02   ✉ akshayantimatter@gmail.com   📍 Prague, Czechia

## PROFILE SUMMARY

Machine Learning Engineer with a strong foundation in NLP, GenAI technologies, and scalable software engineering. Passionate about building reliable, production-ready AI systems that integrate with existing infrastructure. Experienced in LLM-powered tools using GPT-3/4 and techniques like Graph RAG and Model Context Protocol (MCP). Skilled in designing data pipelines, automating workflows, and delivering tools like semantic search engines, intelligent chatbots, and contract testing frameworks. Committed to bridging research and deployment through clean code and collaborative practices.

## TECHNICAL SKILLS

- **Programming Languages**: Java, Python, Bash, HCL (Terraform)
- **Machine Learning & AI**: Tensorflow, scikit-learn, NLP, Model Deployment, Deep Learning, LLMs, HuggingFace, Keras, LangChain, LangGraph, MCP, RAG, Graph RAG, LLM Agents
- **Big Data & Engineering**: Kafka, RabbitMQ, Docker, Kubernetes, Data Pipelines
- **Cloud & Deployment**: AWS, FastApi, Jenkins, Buildkite, CI/CD Pipelines

## GENAI & ML PROJECTS

- **GenAI Pipeline for Efficacy Analysis** (Cisco, In Progress)
    - Designed and implemented a GenAI pipeline to assist researchers in triaging weekly security detections, reducing manual analysis time and false positive fatigue
    - Leveraged Graph-based Retrieval-Augmented Generation (Graph RAG) with Model Context Protocol (MCP) to contextualize detection definitions and weekly trigger data
    - Integrated external threat intelligence (MITRE ATT&CK) to enrich the detection context and improve the quality of AI-assisted recommendations.
    - Flagged historically noisy or high-FP detections by analyzing historical firing patterns and false positive rates, enabling faster researcher prioritization
    - Reduced manual detection review workload by ∼80%, while increasing true positive identification consistency across weekly triage cycles.
    - **Technologies:** Python, GPT-4, MCP, Graph-based RAG, RAG, OpenAI, MS Azure
- **Intelligent FAQ Chatbot for Internal Documentation** (Cisco)
    - Designed and deployed an FAQ chatbot leveraging Langchain, GPT-3, and RAG over enterprise knowledge using Model Context Protocol (MCP)
    - Final version boosted response precision by 30% and reduced manual documentation search effort by 60%.
    - **Technologies**: Python, Langchain, OpenAI, FastAPI, MCP, RAG
- **Semantic Ticket Linking Tool for JIRA** (Cisco)
    - Developed a command-line tool to identify and rank relevant historical tickets based on semantic similarity

- Reduced manual effort in identifying duplicates or similar past work by improving triage efficiency by 35% and increasing related-ticket detection by 50% during internal QA cycles

- **Technologies**: Python, FAISS, Sentence Transformers, JIRA Rest API

## SOFTWARE ENGINEERING EXPERIENCE

### Software Engineer (Forensics, Secure Network Analysis) — Cisco, Czechia (August 2023 - Present)

- Designed and implemented data ingestion pipelines to process high-volume telemetry data from multiple sources, improving data accuracy and analysis efficiency

- Developed automated data visualization dashboards for on-premises security products, enabling insights for threat detection and network monitoring

- Led migration from RAML to code-generated OpenAPI specifications, improving documentation accuracy, maintainability, and cross-team interoperability

- Integrated Schemathesis into the CI pipeline for automated API contract testing; served as the team's go-to contact for API test reliability and spec validation, reducing integration time by 50%

- Led high-performing teams in AI-focused hackathons, securing multiple podium finishes and recognition for innovation

- **Technologies**: Java, Scala, Docker, RabbitMQ, Jenkins

### AI Community & Mentorship

- **Facilitate a bi-weekly internal AI book club** centered around GenAI frameworks and applied implementations, with a discussion on topics including, but not limited to, Model Context Protocol (MCP), Retrieval-Augmented Generation (RAG), Graph RAG architectures, Agentic Architectures

- Combine research discussions with hands-on project work, enabling participants to apply concepts in areas like document question-answering, intelligent triage systems, and LLM-driven tooling

- **Mentor teammates on integrating ML and LLM-based techniques** into production workflows, covering use cases such as semantic search, autonomous agents, and pipeline automation

- Fostered a culture of AI-first experimentation, with **12 active participants** and $\sim 60\%$ **adoption of GenAI tools and patterns** across teams

### Software Engineer (Real Time Insights, Flex) — Twilio, Czechia (November 2020 - June 2023)

- Optimized a Kafka Streams pipeline, reducing real-time data processing latency by about 40%, from 45s to 25s

- Led Infrastructure-as-Code (IaC) adoption using Terraform, automating resource provisioning and cutting deployment time by 35%.

- Enhanced testing frameworks with mutation testing, chaos testing, and load testing, improving test coverage by 40%

- Reduced incident response times by improving on-call runbooks and introducing better alerting thresholds during weekly rotations

- **Technologies**: Java, Kafka, Docker, Terraform, Buildkite, Jenkins, Selenium, AWS

## EDUCATION

- **M.S. in Computational Linguistics** — Charles University (Czechia) & University of the Basque Country (Spain)

- **B.Tech. in Information Technology** — Vellore Institute of Technology (India)

## SELECTED PUBLICATIONS

1. A. Aggarwal and C. Alzetta, "**Atypical or Underrepresented? A Pilot Study on Small Treebanks**", CLiC-it 2021, CEUR-WS.org, 2022. [Online]. Access Link

2. A. Aggarwal and D. Zeman, "**Estimating POS Annotation Consistency of Different Treebanks**", Workshop on Treebanks and Linguistic Theories, ACL, 2020. [Online]. Access Link

3. A. Aggarwal, **Consistency of Linguistic Annotation**, M.S. thesis, Charles University, 2020. [Online]. Access Link