

Anshu Kiran Sharma

Graduate Research Assistant — +1(971)977-8316; aks.anshu03@gmail.com; anshukiran.com

RESEARCH INTERESTS

natural language processing, large language models, agentic AI, computational linguistics, artificial intelligence, machine learning, narrative understanding, cognitive science

EDUCATION

- **Doctor of Philosophy (PhD) in Computer Science** January 2021 – Present
Florida International University, Miami, FL *Expected graduation August 2026*
 - **Major Advisor:** Prof. Mark A. Finlayson
- **Master of Science (MS) in Computer Science** 2021 – 2024
Florida International University, Miami, FL
- **Bachelor of Science (BSc) in Computer Science and Information Technology** 2013 – 2017
Tribhuvan University, Kathmandu, Nepal

RESEARCH / ACADEMIC EXPERIENCE

- **Cognac Laboratory, Florida International University** May 2022 – Present
Graduate Research Assistant *Miami, FL*
 - Conducted research on narrative understanding and AI-generated content detection, developing transformer and LLM-based pipelines for narrative similarity, resume classification, and candidate ranking
- **Pacific Northwest National Laboratory (PNNL)** March 2025 – August 2025
Ph.D. Research Intern *Remote*
 - Built a cloud-native multi-agent LLM orchestration system for scientific workflows in the Language Intelligence Team (ACMD, PCSD) within a multi-institution Department of Energy collaboration
- **Florida International University** May 2021 – Apr 2022; Jan 2025 – Apr 2025; Aug 2025 – Dec 2025
Graduate Teaching Assistant *Miami, FL*
 - Graded assignments and projects for *Introduction to Machine Learning* (class size 29); *Human-Computer Interaction* (class size 24); *Introduction to Cryptocurrencies* (class size 120); *Advanced DBMS* (class size 80)

SELECTED PROJECTS

- **Kernel Narrative Similarity** May 2022 – Present
Project Lead, Narrative similarity pipeline
 - Created annotation guidelines and constructed the first 50-narrative dataset for kernel event detection; led a team of 4 annotators to build training corpus
 - Benchmarked transformer architectures (BERT, RoBERTa, T5, Mistral, Llama3) for kernel event detection; LoRA-fine-tuned Llama3 achieved best performance ($F_1 = 0.695$)
 - Developing an end-to-end narrative similarity pipeline incorporating causal event representations
- **Quantum Recruit** August 2025 – Present
Project Lead, AI resume detection and candidate ranking system
 - Built the first labeled corpus of 420 resumes annotated as authentic, AI-assisted, or AI-generated
 - Developed baseline models for detecting AI-generated and AI-enhanced resumes, and benchmarked on the corpus (LR, SVM, Random Forest, XGBoost, BERT, Longformer), achieving $F_1 = 0.953$ using XGBoost
 - Developing a transformer-based semantic ranking system for candidate-job matching
- **Accelerate: Integrated Platform to Predict Degradation of Catalysts** March 2025 – August 2025
Task Lead, Multi-agentic LLM co-pilot development
 - Implemented a multi-agent LLM framework that acted as a scientific co-pilot to aid natural scientists in research process, from hypothesis generation to running simulations to result interpretation using LangGraph and langgraph-supervisor; integrated supervisor-based task routing and tool-calling agents; deployed to AWS
 - Designed and conducted evaluation protocols to assess co-pilot's effectiveness in supporting scientific tasks

PUBLICATIONS

- [Conference Paper] **Sharma, A.**, Castiblanco-Melendez, M., Morales, A., & Finlayson, M. A. (2026). “Once Upon a Kernel: Extracting Important Events from Narratives.” *LREC 2026*
- [Conference Paper] Loizidou, A., **Sharma, A.**, Esquivel, A., Finlayson, M. A., & Ocal, M. (2026). “Corpus and Baselines for Distinguishing Authentic, AI-Generated, and AI-Enhanced Resumes.” *LREC 2026*
- [Workshop Paper] Acharya, A., **Sharma, A.**, Parker, D., Vega, T., Ashraf, R. A., Isenberg, N., Strube, J., & Rallo, R. (2025). “LABMATE: Language Model Based Multi-Agent System to Accelerate Catalysis Experiments.” *SC Workshop on Frontiers in Generative AI for HPC Science and Engineering: Foundations, Challenges, and Opportunities, 2025*
- [Workshop Paper] Islam Erana, T., Islam, A. A., **Sharma, A.**, & Finlayson, M. A. (2026). “COGNAC at SemEval-2026 Task 4: Evaluating Narrative Components with LLMs for Hard Story Similarity Cases.” *SemEval 2026*
- [Preprint] Acharya, A., Vega, T., Ashraf, R. A., **Sharma, A.**, Parker, D., & Rallo, R. (2026). “A Cloud-based Multi-Agent Workflow for Science”

ARTIFACTS

- [Code & Dataset] Corpus for Kernel Event and TimeML Event Detection in Fairytales – 2026 – *FIU Dataverse*
- [Dataset] Corpus for Distinguishing Authentic, AI-Generated, and AI-Enhanced Resumes – 2026 – *FIU Dataverse*

AI / ENGINEERING TECHNICAL SKILLS

Machine Learning & NLP: LLM fine-tuning (LoRA, PEFT), NLP modeling, LLMs (GPT, Llama, Mistral, Claude, Gemini), classical ML; **Frameworks & Libraries:** HuggingFace, Transformers, TensorFlow, PyTorch, LlamaIndex, LangChain, LangGraph; **Agentic AI:** multi-agent systems, agent orchestration, tool-augmented reasoning, LLM-based pipelines; **Cloud:** AWS, GCP, Azure, HPC (UF HiPerGator, PNNL Constance), containerized workflows; **Programming:** Python, SQL

STUDENT MENTORING

- **Alejandro Morales**, Undergraduate Research Intern, October 2023 – May 2025
- **Miguel Castiblanco Melendez**, Undergraduate Research Intern, October 2023 – May 2025
- **Amanda Chacin-Livinalli**, Undergraduate Research Assistant, November 2021 – December 2023
- **Break Through Tech Guild, Miami**, Mentored non-computer science undergraduate students for five days, addressing real-world challenges in a recurring event held every semester break from July 2022 – December 2023

EVENTS AND SERVICES

- **Peer Reviewer** – SemEval 2026; SOCIO-NLP 2025; COLING 2025; LaTeCH-CLfL 2025; FuturED 2024
- **Professional Member** – ACL

INDUSTRY EXPERIENCE

- **Deerwalk Inc.** November 2018 – February 2021
Software Engineer *Kathmandu, Nepal*
 - Developed web applications using Grails, Angular, and AWS (S3, SES, SNS) with test-driven development
 - Built data processing and analytics pipelines using Java Cascading, Hadoop, EMR, Redshift, and MySQL for US healthcare data
 - Authored technical documentation and collaborated with Product Owners to drive project delivery
- **Insight Workshop** August 2017 – November 2018
Junior Software Engineer *Kathmandu, Nepal*
 - Developed backend systems and APIs using Django, handling deployment, migration, and query optimization
 - Managed AWS infrastructure (EC2, RDS, S3, Lambda, CloudWatch) and built Android applications with custom database design

LANGUAGES

Nepali: Native Language; **English:** Native Proficiency;
Hindi: Minimum Professional Proficiency **Urdu:** Spoken only, Minimum Professional Proficiency