

# Mahdi Erfanian

✉ merfan2@uic.edu | 🏠 merfanian.github.io | 🔗 linkedin.com/in/merfanian | 🐙 github.com/merfanian

Chicago, Illinois, USA

Ph.D. Candidate at UIC and Research Intern at Microsoft with a strong publication record in top-tier venues (VLDB, ICML, Data Engineering Bulletin) specializing in LLMs, multimodal data management, generative AI, algorithmics and efficient AI systems. My research combines generative AI and multi-modal databases to solve complex information retrieval challenges, leading to systems that outperform SOTA baselines like OpenAI's CLIP by 200%.

## EDUCATION

---

### • University of Illinois Chicago

December 2022 - Present

Direct Ph.D., Computer Science

Chicago, USA

- Thesis: Generative AI for Multimodal Data Management
- M.S. in Computer Science (awarded en route to Ph.D.), Fall 2025
- GPA: 4.00/4.00

### • Sharif University of Technology

September 2017 - June 2022

Bachelor of Science, Computer Engineering

Tehran, Iran

- GPA: 3.70/4.00
- Ranked 60th nationwide among 150,000 participants in the Iranian University Entrance Exam (Konkour).

## RESEARCH INTERESTS

---

Generative AI, LLMs, Multimodal Data, Vector Databases, Information Retrieval, Algorithmics, Distributed Systems

## SELECTED PUBLICATIONS

C=CONFERENCE, J=JOURNAL S=IN SUBMISSION

- [S.1] **Delulu-FIM: A Verified Multi-Lingual Benchmark for Code Hallucination Detection in Fill-in-the-middle Tasks**  
Erfanian, M., Garg, A., Liu, X., Golnari, P., Fu, S. (2026). Manuscript submitted for publication in *NeurIPS 2026*
- [S.2] **NeedleDB: A Generative-AI Based System for Accurate and Efficient Image Retrieval using Complex NL Queries**  
Erfanian, M., Asudeh, A. (2026). Manuscript submitted for publication in *VLDB 2026*
- [S.3] **Task-aware Data Augmentation using Generative AI for Group-distributional Robustness**  
Erfanian, M., Glavic, B., Asudeh, A. (2025). Manuscript submitted for publication in *The VLDB Journal*
- [S.4] **Needle: A Generative AI-Powered Multi-modal Database for Answering Complex Natural Language Queries**  
Erfanian, M., Dehghankar, M., Asudeh, A. (2025). Manuscript submitted for publication in *KDD 2026*
- [C.1] **An Efficient Matrix Multiplication Algorithm for Accelerating Inference in Binary and Ternary Neural Networks**  
Dehghankar, M., Erfanian, M., Asudeh, A. (2024). In *The 2025 International Conference on Machine Learning*
- [C.2] **Chameleon: Foundation Models for Fairness-Aware Multi-Modal Data Augmentation to Enhance Coverage of Minorities**  
Erfanian, M., Jagadish, H. V., Asudeh, A. (2024). In *Proceedings of the VLDB Endowment*  
DOI: 10.14778/3681954.3682014
- [C.3] **FairEM360: A Suite for Responsible Entity Matching**  
Shahbazi, N., Erfanian, M., Asudeh, A., Nargesian, F., Srivastava, D. (2024). In *Proceedings of the VLDB Endowment*  
DOI: 10.14778/3685800.3685889
- [J.1] **Coverage-based Data-centric Approaches for Responsible and Trustworthy AI**  
Shahbazi, N., Erfanian, M., Asudeh, A. (2024). In *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*
- [C.4] **Graph Pattern Mining Paradigms: Consolidation and Renewed Bearing**  
Dias, V., Ferraz, S., Vadhvani, A., Erfanian, M., Teixeira, C. H., Guedes, D., ... Parthasarathy, S. (2023, December). In *2023 IEEE 30th International Conference on High Performance Computing, Data, and Analytics (HiPC)*  
DOI: 10.1109/HiPC58850.2023.00040

## WORK AND RESEARCH EXPERIENCES

---

- **Ph.D. Research Intern** Jan 2026 - Present  
*Microsoft (Code | AI team) [🌐]* Redmond, WA (Remote)
  - Researching on **mitigating hallucination in LLMs and GenAI as a Code Agent** (Copilot, Codex)
  - Introducing **Delulu-FIM**, the first multi-lingual execution-based verified benchmark for hallucination detection and mitigation in Fill-in-the-middle tasks.
  - Fine-tuning LLMs using curated datasets to reduce the hallucinations.
- **Research Assistant** June 2023 - Present  
*IndexLab, University of Illinois Chicago [🌐]* Chicago, Illinois
  - Researching under supervision of **Dr. Abolfazl Asudeh** on databases, LLMs and responsible data management.
  - Developed **Needle**, an efficient powerful text-to-image retrieval framework that **outperformed OpenAI's CLIP by 200%** in mean average precision on complex natural language queries.
  - Implemented **RSR**, an efficient binary/ternary matrix multiplication method, accelerating model inference time by **24x** compared to the standard **NumPy** baseline and **2.5x** on **Quantized LLMs**
  - Developed **Chameleon**, a fairness-aware data augmentation method that improved model accuracy on under-represented groups by **22%** on the **FERETDB** benchmark on average.
  - Developed **FairEM360**, a framework for auditing and mitigating bias in entity matching.
- **Research Assistant** December 2022 - August 2023  
*Dreese Lab, The Ohio State University* Columbus, Ohio
  - Researched graph and pattern mining under the supervision of **Dr. Srinivasan Parthasarathy**.
  - Developed **SYSML+**, an enhanced stylometry framework that improved author identification accuracy by **3%** over the baseline **SYSML** system.
- **Software Engineer** March 2022 - August 2022  
*Software Engineering Lab, Sharif University of Technology* Tehran, Iran
  - B.Sc. Project: Architected and built a CI/CD pipeline for a containerized, microservice-based web application.
  - Optimized a stock market application's deployment from a monolithic to a distributed architecture, improving scalability and deployment speed.
  - Decreased production Docker image size by **over 95%** compared to traditional build methods, enabling faster deployments.
- **Data Engineer** September 2019 - March 2022  
*Divar Corp. [🌐]* Tehran, Iran
  - *Divar is Iran's largest classifieds platform with over 40 million active users and 200 TB of data.*
  - Engineered and maintained the core data pipeline (Airflow, Spark, S3) processing over **200TB** of data, improving data availability for a team of **40+** data analysts.
  - Deployed a distributed JupyterHub on Kubernetes, enabling on-demand, scalable analysis environments for the data science team and cutting down experiment setup time.
  - Automated ETL deployment workflows using GitLab CI/CD, reducing manual deployment time from hours to minutes and eliminating release errors.
- **Software Engineering Intern** August 2019 - November 2019  
*Rahnema College* Tehran, Iran
  - Designed and developed an E-commerce auction application using a microservice architecture (Spring Boot, ReactJS, Docker).
  - Implemented a push notification service from scratch using Spring Boot and WebSocket.

## SKILLS

---

- **Languages:** Python, Java, Golang, C/C++, SQL, Scala, JavaScript, Bash
- **AI / ML Frameworks:** PyTorch, NumPy, Pandas, Transformers (HuggingFace), NetworkX
- **Data Systems & Databases:** Apache Spark, Apache Airflow, HDFS, PostgreSQL, MySQL, Redis, Milvus (Vector DBs)
- **DevOps & Cloud:** Docker, Kubernetes, Azure AI Services, Azure DevOps, Git, FastAPI, Django, Spring Boot, ReactJS

## PROFESSIONAL ACTIVITIES

---

### • Invited Lectures/Talks

- **CS516. Responsible Data Science and Alg. Fairness:** Guest Lecture on Generative AI and its applications in Fairness, UIC (2026)
- **CS418. Introduction to Data Science:** Guest Lecture on Generative AI and Multimodal Data Management, UIC (2025)

### • Program Committee & Reviewing

- **PC Member:** WWW 2026, CIKM 2026/2025/2024, KDD 2026, NeurIPS 2026
- **Reviewer:** ICLR 2026, NeurIPS 2025 (DynaFront), TKDE 2025/2024, PETRA 2024 (ETHER-AI)

### • Teaching Assistant

- **UIC:** Databases (Spring 24, Fall 23)
- **OSU:** Intro to Java (Summer 23, Spring 23)
- **SUT:** Computer Networks, Architecture, Systems Analysis, Algorithms, Programming (2018–2021)

## NOTABLE COURSES

---

Cloud Computing (A/Grad), Advanced Algorithms (A/Grad), Introduction to Network Science (A/Grad), Algorithms I (A/Grad), Software Engineering (4/4), Modern Information Retrieval (4/4), Functional Programming (4/4), Design of Algorithms (4/4), Data Structures and Algorithms (4/4), Databases Design (4/4)

## EXTRACURRICULAR ACTIVITIES

---

### • Leadership & Volunteer Experience

- **Chicago Persian School:** Volunteer Teacher, developing curricula and instructional sessions (2025)
- **Rayanesh (SUT):** Managing Director of the Computer Science and Engineering Department Publication (2019–2022)