Toluwani Samuel Aremu

AI Researcher | Applied Scientist

Google Scholar • Website • Medium • Email • LinkedIn • GitHub • Certifications

KEY COMPETENCIES

- **Skills**: Research, Mathematics, Statistics, Machine Learning, Deep Learning, Data Science, Project Management, Programming, Writing.
- **Tools**: Python, VB.Net, PyTorch, Lightning, TensorFlow, Keras, Jax, Scikit-learn, NumPy, Matplotlib, Visual Studio, Visual Studio Code, PyCharm, Jupyter, Latex, Office 365, Google [Docs, Sheets, Slides].
- **Research Interests**: AI Safety, Trustworthy AI, Responsible AI.

EDUCATION

- **Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI), UAE** AUG 2023 – PRESENT
  - Doctor of Philosophy (PhD) in Machine Learning.
  - Research Areas: AI Safety, Self & Collaborative Alignment, Watermarking.
- **Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI), UAE** JAN 2021 – DEC 2022
  - Master of Science (MSc) in Machine Learning.
  - Research Area: Privacy-Preserving ML.
- **University of Ibadan (UI), Nigeria** MAY 2018 – MAY 2020
  - Master of Science (MSc) in Computer Science.
  - Research Area: Cryptography.
- **Adeleke University, Nigeria** OCT 2012 – JUL 2016
  - Bachelor of Science (BSc) in Computer Science.
  - Minor: Philosophy.

RESEARCH EXPERIENCE

- **PhD Student, MBZUAI, UAE** AUG 2023 – PRESENT
  - Research Areas: Safe and Trustworthy Generative AI.
- **Research Assistant, MBZUAI, UAE** FEB 2023 – AUG 2023
  - Research Areas: Safety in Smart Cities, Generative AI, Evaluation.
- **Applied Science Intern, M42, UAE** FEB 2023 – APR 2023
  - Built an end-to-end pipeline to efficiently download and preprocess NHANES dataset.
  - Ensured that the pipeline generates a comprehensive TRIPOD report after use.
- **MSc Student/Graduate Research Assistant, MBZUAI, UAE** JAN 2021 – DEC 2022
  - Research Areas: Privacy-Preserving ML, Responsible AI, Bias Mitigation, etc.

TEACHING EXPERIENCE

- Teaching Assistant, Mathematical Foundations of AI (MTH701) [MBZUAI] 2024
- Teaching Assistant, Object Oriented Programming (OOP) [UI, Nigeria] 2019

VOLUNTEERING EXPERIENCE

- Graduate School Application Mentorship (2021 – PRESENT)
- Associate Editor, MBZUAI Blog (2024 – PRESENT)

- Reviewer: IEEE Access (‘23) | OIS (‘23) | NeurIPS (‘24) | ICCTIS (‘24) | DLI (‘24) | ICLR (‘24) | HRAIM (‘24) | AISTATS (‘24) | SafeGenAI (‘24) | ICML (‘25)

PUBLICATIONS (* denotes equal contribution)

- **T. Aremu** et al., “On the reliability of Large Language Models to misinformed and demographically informed prompts,” AI Magazine (AAAI), vol. 46, no. 1, 2025.
- M. Nwadike, Z. Iklassov, **T. Aremu** et al., “RECALL: Library-Like Behavior In Language Models is Enhanced by Self-Referencing Causal Cycles,” arXiv, 2025.
- A. Diaa, **T. Aremu**, & N. Lukas, “Optimizing Adaptive Attacks against Content Watermarks for Language Models,” arXiv, 2024.
- S. Fares*, K. Ziu*, **T. Aremu*** *et al.*, “MirrorCheck: Efficient Adversarial Defense for Vision-Language Models,” arXiv, 2024.
- N. Tastan, S. Fares, **T. Aremu**, S. Horvath, and K. Nandakumar, “Redefining Contributions: Shapley-Driven Federated Learning,” 33rd International Joint Conference on Artificial Intelligence (IJCAI), Jeju, Korea, 2024.
- **T. Aremu** and K. Nandakumar, "PolyKervNets: Activation-free Neural Networks For Efficient Private Inference," 2023 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML), Raleigh, NC, USA, 2023.

TALKS & PRESENTATIONS

- Ethical Perspectives of Artificial Intelligence, Department of Material Sciences, University of Denver, Virtual. (JUL 2022).
- PolyKervNets: Activation-free Neural Networks For Efficient Private Inference, IEEE SaTML, North Carolina, USA. (FEB 2023).
- SSIVD-Net: A Novel Salient Super Image Classification and Detection Technique for Weaponized Violence, Computing Conference, London, UK. (JUL 2024).

HONORS & AWARDS

● MBZUAI MSc & PhD Fully Funded Fellowship	2021-2027
● UAE Golden Visa for Talented Persons/Specialists in Science	2022
● MBZUAI Award of Appreciation for Iconic Representation and Student Hospitality	2022
● ProjectSet Innovation Challenge for Entrepreneurship (ICE-22)	2022
● Top 100, DeepLearning.AI Data-Centric AI Competition	2021
● NYSC-FRSC Award for the Most Creative Corp Member	2017
● 2015 AUE-NACOSS Award for the Best Programmer	2015

REFERENCES

- Dr. Nils Lukas - Assistant Professor, MBZUAI - nils.lukas@mbzuai.ac.ae.
- Dr. Karthik Nandakumar - Associate Professor, MBZUAI - karthik.nandakumar@mbzuai.ac.ae.
- Prof. Kun Zhang - Professor, MBZUAI & Carnegie Mellon University - kun.zhang@mbzuai.ac.ae.
- Prof. Abdulmotaleb El Saddik - Distinguished Professor, UOttawa - elsaddik@ottawa.ca.