

# Abiodun Modupe, Ph.D

+1-832-774-3303 | [abiodunmodupe@gmail.com](mailto:abiodunmodupe@gmail.com) | [Website](#)

[in](#) [abiodun-modupe](#) | [dupsys](#) | [G](#) [Google-Scholar](#) | [dupeabbey](#)

Richmond, Houston, Taxes - 77406, USA

## SUMMARY

Experienced Machine Learning Engineer and Data Scientist with a strong foundation in advanced analytical techniques and a passion for developing innovative solutions to complex problems. Skilled in leveraging deep learning frameworks and big data methodologies to drive impactful insights and optimize system performance.

## EXPERIENCE

### • Dupson LLC

System Support

Nov 2025 - present

Richmond, TX. Part-time

- Building an IT service framework, including a help desk, networking, cybersecurity, and cloud solutions to meet the growing demand for business.
- Installed, configured, and maintained enterprise-level IT infrastructure for clients, including networks, servers, and operating environments.
- Delivered systems integration services and technical support for financial software deployments across private and public sector clients.
- Implemented user-access controls and security protocols to ensure system integrity and compliance with organisational requirements.
- Conducted training and consultation sessions to transfer knowledge on system operation, maintenance, and troubleshooting to client IT teams.
- Carried out quarterly maintenance, support, and installation of software patches to enhance IT productivity, including network management tools, the MS Office suite, user-access controls, security protocols, and enterprise systems

### • University of Pretoria

Lecturer & Big Data Science (BDS) Coordinator

Nov 2019 – present (Remotely)

Pretoria, South Africa. Full-time

- Programme Coordination.
  - \* Oversight of curriculum, student mentorship, research supervision, and module delivery for the Big Data Science programme.
  - \* Collaborating with stakeholders to gather requirements and propose effective strategies for the implementation of the BDS degree curriculum and [capstone project](#) .
- Teaching & Curriculum Development
  - \* Designed and taught several postgraduate courses, such as MIT801 – Intro to Machine & Statistical Learning, MIT808 – Big Data Science Project, COS801 – Deep Learning, COS781 – Data Mining, & COS760 – Natural Language Processing.
  - \* Led the redesign and evaluation of course content and assessments by integrating project-based learning to improve student engagement and learning outcomes.
- Postgraduate Supervision & Mentorship
  - \* Supervised or co-supervised over 15 MIT dissertations, as well as several MSc and PhD students across several research domains, such as NLP, vehicle damage estimation, sign language recognition, & computational science
  - \* Guided students in writing for publication, software prototyping, and real-world deployment.
- Research & Project Leadership
  - \* Led or co-led grants on multilingual NLP, misinformation detection, financial inclusion, and SDG impact analytics.
  - \* Guided students in identifying research frontiers and shaping projects that culminated in their dissertations.
  - \* Published in top journals (IEEE Access, Applied Sciences, etc.) and conferences (SMC, AFRICON, etc.).
- Community Engagement & Conference Involvement
  - \* Served on programme committees (SAICSIT, SACAIR).
  - \* Participated in community-driven initiatives such as Deep Learning Indaba, Masakhane, and AI4D.
  - \* Delivered invited talks both locally and internationally (e.g., UP Colloquium, [IEC: E-Voting South Africa](#), [UMich: All You Need Is Community](#), [Responsible AI and Capacity Building at AI G20 conference](#), & [AI4D Partners Meeting:2025](#)).

- **First City Monument Bank** [🌐]

Mar 2004 - Dec 2008

System Analyst – Back End

Lagos, Nigeria. Full-time

- Contributed to the design, development, and support of FlashMeCash as an electronic banking platform that supports payments, collections, and retail services (e.g., transfers and bill payments).
- Worked closely with operational divisions to ensure the availability, integrity, and performance of the FlashMeCash platform, including daily monitoring and reporting of transaction volumes and system status.
- Supported the implementation and optimization of FlashMeCash API connectivity to the core banking system, ensuring smooth account-to-eWallet transfers (savings, current, and other accounts) and accurate balance updates.
- Managed system troubleshooting, debugging, patch deployments, and system migrations to improve platform performance and ensure SLA compliance.
- Provided system documentation and completed the operational handover to stakeholders.

- **Alpha-Beta Consulting** [🌐]

Jan 1999 - Feb 2004

System Engineer

Lagos, Nigeria. Full-time

- Contributed to the implementation of an enterprise infrastructure called the Electronic Banking System of Revenue Cycle Management (EBS-RCM) to facilitate the digitisation of the direct bank lodgment system (DBLS) revenue collection through participating banks and payers, including payment workflow support, receipting, and real-time monitoring/reporting to enhance IGR collection processes.
- Implemented and design controller, service, and repository layers that underpin the EBS-RCM infrastructure, enabling revenue collection from clients (e.g., banks) via a dedicated API, receipt printing for payees, and real-time reporting of revenue collections to the government.
- Designed and implemented a dynamic querying system that stakeholders can manage, track, reconcile, and report government revenue collections (i.e., internally generated revenue/IGR) collected through banks and other channels
- Wrote documentation and Conducted training on the operation of EBS-RCM.

## EDUCATION

- **PhD in Computer Science**

Graduated Jul 2024

The University of the Witwatersrand

Johannesburg, South Africa

- PhD thesis: *Regularized Deep Neural Network for Post-Authorship Attribution*

- **MSc in Computer Science**

Graduated Aug 2011

Tshwane University of Technology

Pretoria, South Africa

- MSc thesis: *Intelligent Identification of Advanced Fee Fraud Activities on the Internet.*

- **Bachelor of Honour degree in Computer Science with mathematics)**

Graduated Jul 1998

Obafemi Awolowo University

Ile-Ife, Osun-State (Nigeria)

- Final Project Title: *Machine Learning–Based Expert System for Dermatological Diagnosis.*

## PROJECTS

- **Mafoko: Lexicon & Glossary Project for South African Terminology.**

Feb 2025 - August 2025

Tools: Python, LLM, Fine-tuning model, & LRLs

Mafoko [🌐]

- Develop an LLM to address the lack of structured terminological data for South Africa's official languages.
- An open-source dataset (e.g., NOODL<sup>1</sup>) to enhance equitable data governance and local benefit sharing.

- **Leveraging AI to Enhance Risk Understanding of Zoonotic Diseases**

Feb 2025 - Jul 2025

Tools: Python, LLM, Fine-tuning model, & LRLs

Zoonotic-disease [🌐]

- Develop and evaluate an AI-driven approach to deliver reliable rabies information to low-resource languages (LRLs), such as Sepedi.
- Provide a dataset to compare two large language models, GPT-4o and Gemini-1.5 Flash, under both base and fine-tuned conditions.

- **Subword-Information-for-AA-task**

Nov 2023

Tools: Python, LLM, Character embedding, Deep learning, Forensic Science & Linguistics

Subword-AA [🌐]

- Addressed the problem of the ambiguity of hidden words in text documents while preserving the sequential context of multiple candidate authors in given structured and unstructured documents.
- Design a CNN sequential model to perform authorship attribution through writing style pattern recognition for digital forensic pre-trial support.

- **Masakhane WEB: A Machine Translation Web Platform for African Languages**

Feb 2021 - Dec 2021

Tools: Python, Docker, Kubernetes, Java REST

Masakhane-web [🌐]

- Developed translation models based on [Joey NMT](#)
- Provision of LLM as an open-source platform that provides relatively accurate translations for languages across Africa

---

<sup>1</sup>Nwulite Obodo Open Data Licence

[J.1] Thapelo Sindane, **Abiodun Modupe**, Vukosi Marivate. (2025). **Cross-lingual embedding methods and applications: A systematic review for low-resourced scenarios**. *Natural Language Processing Journal*, Vol. 12, pp. 100157. DOI: <https://doi.org/10.1016/j.nlp.2025.100157>

[J.1] M. Mathebula, **A. Modupe** and V. Marivate. (2024). **ChatGPT as a Text Annotation Tool to Evaluate Sentiment Analysis on South African Financial Institutions**. *IEEE Access*, Vol. 12, pp. 144017-144043. DOI: <https://doi.org/10.1109/ACCESS.2024.3464374>

[J.1] **Modupe A**, Celik T, Marivate V, Olugbara OO. (2022). **Post-authorship attribution using regularized deep neural network**. *Applied Sciences*, Vol. 12, Issue 7518, pp. 1-24. DOI: <https://doi.org/10.3390/app12157518>

[C.1] Abiodun Modupe, et al. (2023). **Integrating Bidirectional Long Short-Term Memory with Subword Embedding for Authorship Attribution**. In *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 1910-1917. IEEE. 2023, Honolulu, Oahu, HI, USA. DOI: <https://doi.org/10.1109/SMC53992.2023.10393898>.

TECHNICAL EXPERTISE

- **Research areas:** Large Language Models, Information Retrieval, Sentence Alignment, Machine Translation & Multilingual NLP
- **Programming:** Python (Jupyter Notebook, JupyterLab, Google Colab), VS Code, PyCharm, RStudio, Java, & C++
- **Core ML libraries:** PyTorch, TensorFlow, Keras, pandas, NumPy & scikit-Learn
- **Deployment & MLOps:** WS SageMaker, Google Vertex AI, & Azure Machine Learning
- **Datasets & Tools:** PostgreSQL, MySQL, SQL Server, SQLite, BigQuery, Snowflake, Redshift
- **Systems & Infrastructure:** Mac OS X, LINUX (UBUNTU, AND DEBIAN), Windows, Docker, Docker Compose, Kubernetes (basic)
- **Web technology:** HTML, CSS, JavaScript, RESTful APIs, Apache HTTP Server, Apache Tomcat
- **Typography:** L<sup>A</sup>T<sub>E</sub>X, & Endnote

HONORS AND AWARDS

- **Master’s Scholarship – ZAR 30,000** Jun–Jul 2009-2011  
Tshwane University of Technology
- **PhD Bursary Support – ZAR 300,000** Jan 2019 – Dec 2023  
CSIR-DST
- **Masakhane Web ((Co-PI)– USD 5,000.00** 2020 – 2021  
Funded by Mozilla Foundation Machine Translation Project
- **AfriGemma (Co-PI)– USD 200,000.00** Aug 2025  
Funded by Google Dublin AfriGemma Project
- **AI4D DSFSI Lab (Co-PI) – CAD 1,070,500.00** Aug 2024  
Funded by IDRC–FCDO AI4D DSFSI Lab

RESEARCH LEADERSHIP & IMPACT

- **Dataset Creation** Jun 2020 - present  
DSFSI Research group  
◦ Led the development of over 10 first-of-their-kind multilingual datasets used globally
- **LLM Development:** Jun 2020 - present  
DSFSI Research group  
◦ Led the development of a cross-multilingual LLM for domains such as sentiment analysis, information retrieval, language translation, & others.
- **Leadership** Nov 2019 - Dec 2025  
Computer Science Department / University of Pretoria  
◦ Led and co-led ML, NLP, and data science research initiatives with a focus on societal impact, particularly in African and other low-resource languages.  
◦ Foster interdisciplinary collaboration by connecting DSFSI’s research strands with other departments, faculties, and institutions to ensure that research outputs are socially relevant and contextually grounded.

PROFESSIONAL MEMBERSHIPS

- **IEEE**, Membership ID: 95685929 Jun 2020 - Present
- **Data Science Africa** Jul 2022 - Present
- **Stimson/Microsoft Responsible Fellow** Jul 2024 - Present

LANGUAGE

English (Fluent), Yoruba (Fluent), French (Basic) & Interests include: Score, Chess, & listening to the news.

REFERENCES

Available on Request