Dorcas Ojo

Data Scientist | Bioinformatics Research |Plant Genomics | Health Informatics

Biologist / Data scientist with good experience transforming health and genomics data into actionable insights. Passionate about applied AI, research translation, and building ethical, data-driven solutions that improve plant breeding, food security and healthcare outcomes.

Core Competencies & Tech Stack

Leveraging advanced data science, bioinformatics, and machine learning to solve complex biological and healthcare challenges through reproducible, ethical, and domain-driven analytics.

Bioinformatics & Domain Data Science

High-dimensional transcriptomic data analysis using enrichment, clustering, and network modeling to decode biological complexity.

Tools: BLAST, MEGA, DAVID, KEGG, GO Enrichment

Programming & Analytics

Building efficient data pipelines and computational models for biological and health datasets.

Languages: Python, R, SQL

Libraries: Pandas, NumPy, scikit-learn

Machine Learning & AI

Applying supervised and deep learning architectures for predictive analytics and pattern recognition.

Models: CNNs, XGBoost, LightGBM

Visualization & Reporting

Translating complex results into actionable insights with interactive visual dashboards.

Tools: Power BI, Tableau, Matplotlib, Seaborn

Cloud & Deployment

Deploying scalable and accessible data science solutions for collaboration and reproducibility.

Platforms: AWS, Azure, Streamlit, Flask

Compliance & Data Governance

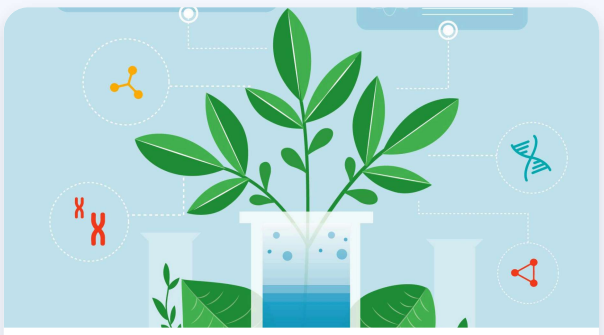
Ensuring ethical research practice and regulatory adherence for sensitive health datasets.

Frameworks: GDPR, NHS Data Governance

Projects

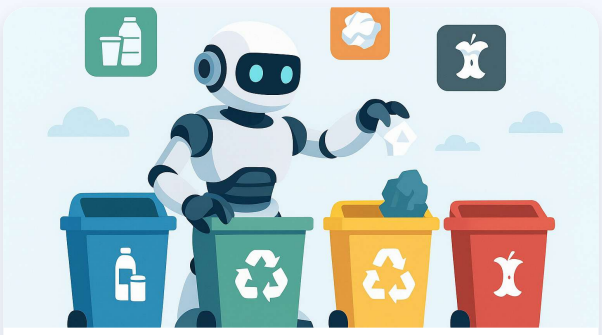
Predictive Health Outcomes

Built ML models using NHS datasets to forecast patient outcomes and enhance clinical resource allocation. Delivered dynamic dashboards in Power BI for operational use.



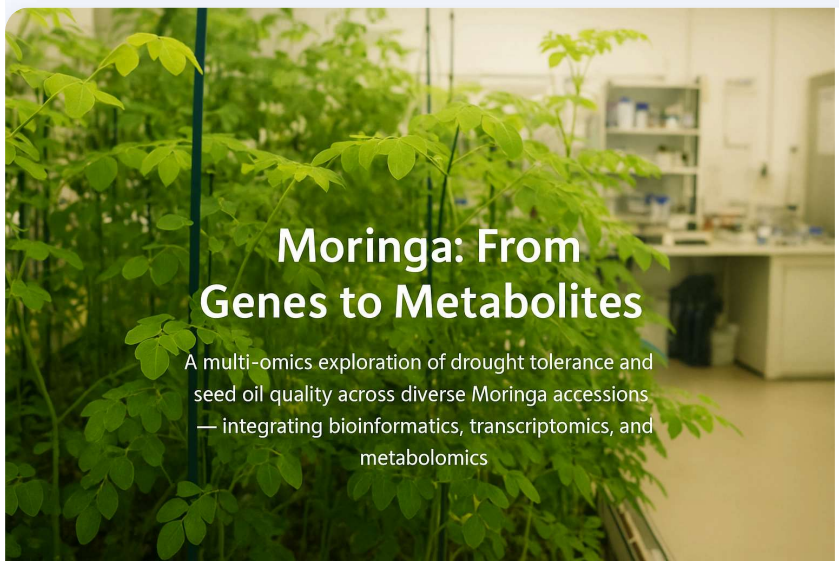
Genomics Classification Model

Created and optimized MobileNetV2 CNN for biological image classification, achieving high accuracy and improving computational biology workflows.



Waste Classification Streamlit App

Designed an AI-driven Streamlit web app for automatic waste sorting and recycling classification using computer vision.



Moringa: From Genes to Metabolites

A multi-omics exploration of drought tolerance and seed oil quality across diverse *Moringa* accessions — integrating bioinformatics, transcriptomics, and metabolomics to uncover genetic and metabolic insights.

- Constructed **de novo transcriptome assemblies** of 49 Moringa accessions using integrated bioinformatics pipelines.
- Performed **RNA-Seq, differential expression, GO/KEGG pathway enrichment**, and co-expression network analysis.
- Discovered high-quality **SNPs** for use in GWAS and metabolic profiling of Moringa seed oil.

Education

PhD, Biology (Plant Genomics)

University of York, UK | 2025

Thesis: "From genes to metabolites: A Multi-Omics exploration of drought tolerance and seed oil quality in diverse Moringa."

MTech, Environmental Biology

LAUTECH, Nigeria | 2012

Thesis: "Determination of Pulp and Paper Making Suitability Indices of some Nigerian Species of Leguminosae: Caesalpinoideae."

PGDip Education

Usmanu Danfodiyo University, Nigeria | 2011

BTech (Hons), Pure & Applied Biology

LAUTECH, Nigeria | 2004

Professional Development & Memberships

- R for Data Science: Analysis and Visualization Certificate (2023)
- Member, WHPC – Women in High Performance Computing
- Member, American Society of Plant Biologists (ASPB)

Publications & Conferences

- *"Manuscript in preparation."*
- Publications on ResearchGate
- Publications on Google Scholar
- Presented findings at BIFOR 2022 (University of Birmingham), EBNet ECR23 (University of Edinburgh), ASPB 2016 (Texas, USA).

Engagement & Community

External Collaborations

- Collaborated with Covenant University/UNESCO on plant stress-tolerance research.
- Worked with Rothamsted on oil metabolic profiling studies.

Community Service

- Soapbox Science Speaker
- Cake-Bake-Sale Fundraiser
- Food Bank Volunteer

Publications & Awards

AI for Healthcare Decision Support

Published in the Journal of Health Informatics (2023). Developed ML-based clinical prediction framework improving hospital readmission prediction accuracy by 15%.

Best Research Award – Data Science Symposium

Recognized for outstanding applied research integrating AI and health data analytics.

Computational Biology Research Grant

Secured competitive funding to develop novel ML approaches for genomic feature selection.

Contact