Lead Supervisors

Dr Gerardo Serra Gerardo.serra@manchester.ac.uk, University of Manchester
Dr Giulia Grisot giulia.grisot@manchester.ac.uk, University of Manchester

Project: Revisiting Development and Crisis: Conceptual Change in African Economics

REDEV investigates the intellectual history of economics in postcolonial Africa by reconstructing how key concepts such as *development* and *crisis* evolved across economic writings produced between the 1960s and the 1990s in countries including Ghana, Nigeria, Sudan, and Zimbabwe. This internship will support the development of a scalable computational pipeline for analysing a newly digitised corpus of African economic texts, contributing to the creation of a reproducible infrastructure for large-scale textual analysis within the REDEV project. The intern will work on developing and testing computational workflows that enable analysis across the corpus, combining natural language processing (NLP) techniques with historical interpretation.

The project involves the development of a reproducible computational pipeline using Python and other computational resources. Tasks will include:

- preprocessing and cleaning a large multilingual corpus of economic texts (OCR correction, tokenisation, metadata structuring);
- implementing computational methods for concept detection and semantic change analysis;
- experimenting with word embeddings and contextual language models to trace how key economic terms (e.g., *development*, *dependency*, *crisis*) evolve across time and national contexts;
- developing scripts and workflows designed for scalable execution on high-performance computing infrastructure.

The intern will work with the supervisor in DH to ensure the workflow follows best practices in reproducible research software and research data management. This will include version-controlled code, structured data pipelines, and documentation suitable for reuse within the broader project.

Why HPC Resources Are Needed

The REDEV corpus will consist of thousands of pages of OCR-processed economic texts drawn from multiple national contexts and languages. Computational analysis will require training and evaluating large NLP models, running iterative topic modelling and embedding models, and performing repeated corpus-wide similarity calculations.

Expected Outputs

The internship will produce:

- a reproducible HPC-enabled text analysis pipeline for historical economic corpora
- documented Python scripts and workflows suitable for reuse in REDEV
- exploratory analyses of conceptual patterns across the corpus
- improved research data management practices for the project

The work will contribute to the digital infrastructure underpinning REDEV and support the long-term computational analysis of African economic thought.