

Policies across the SADC region are largely shaped by external funding priorities and international policy influences rather than locally developed solutions and frameworks.

Africa-centric perspectives on climate change: using locally encoded LLMs to identify historical asymmetries in policy making in SADC region.

INTRO

This project aims to investigate climate change discourse and policy perspectives across the SADC region using a combination of **policy documents**, **academic research articles**, and community or stakeholder insights from interview **video transcripts**. Different language models are used to analyse this data in order to identify the structural biases such as **Colonial framing**, **External dependency framing**, **Universalist assumptions**, **Local knowledge omission**, and **Differential impact recognition in the system**.

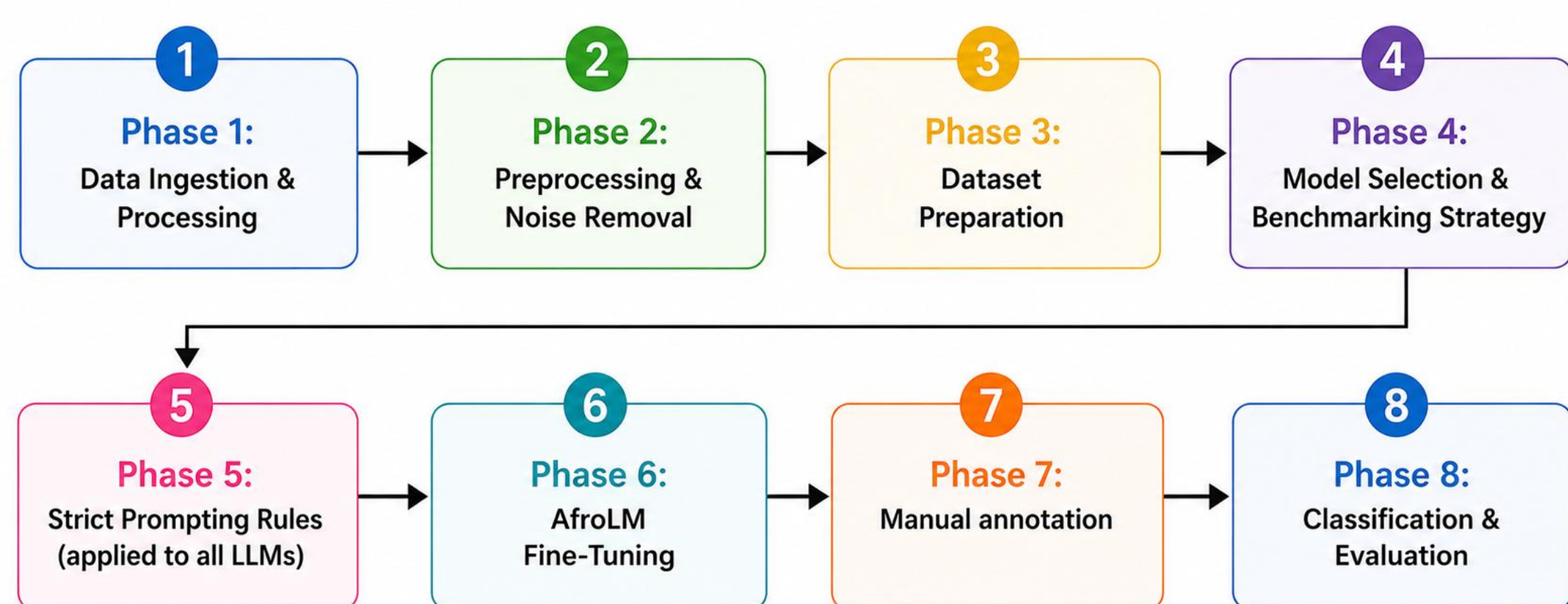
METHODS

1. Generate **JSON** metadata for each document type (**Policy, Video and Articles**)
2. Chunk text by sections;
3. Randomise Split data:

- Generic LLMs (ChatGPT, DeepSeek): **70% train / 30% test**
- AfroLM: **85% train / 15% validation** (additional split)
- Manual Annotation: **30%** original text chunks

4. Output required: filename, source_type, section_title, chunk_index, chunk_text, category, score, justification, evidence, confidence.
5. AfroLM is not a text-generating language model; the first try resulted in 'unclassified' because the model could not understand what was expected, so fine-tuning had to be implemented.

METHODOLOGY WORKFLOW



RESULTS

- This graph below with the heatmap provides comparison of **average bias scores (0-5)** assigned by different models against each bias category. The graph gives more insights into the comparison model behaviour, intensity of the bias classification and sensitivity of the framing.



Interpretation:

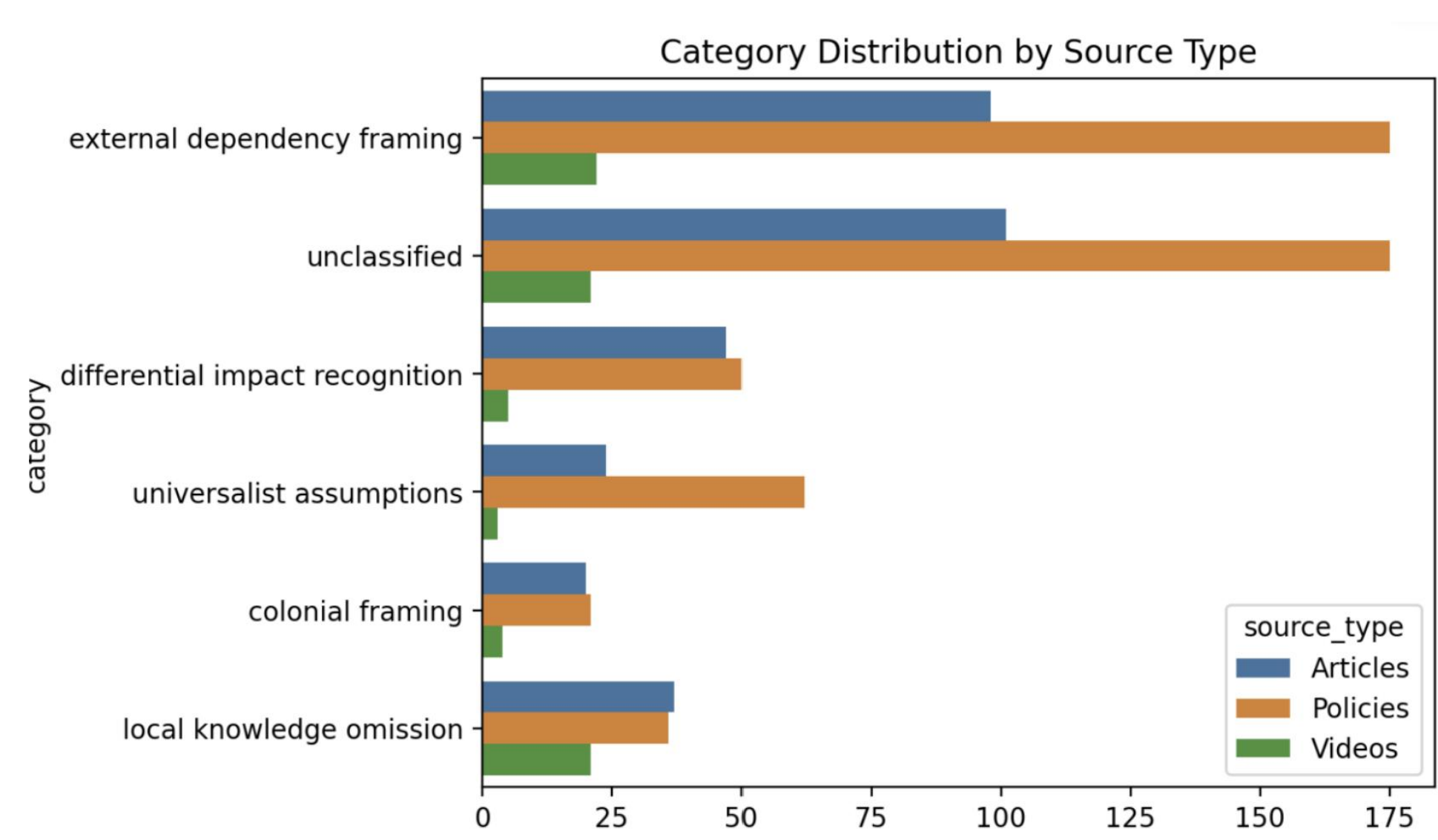
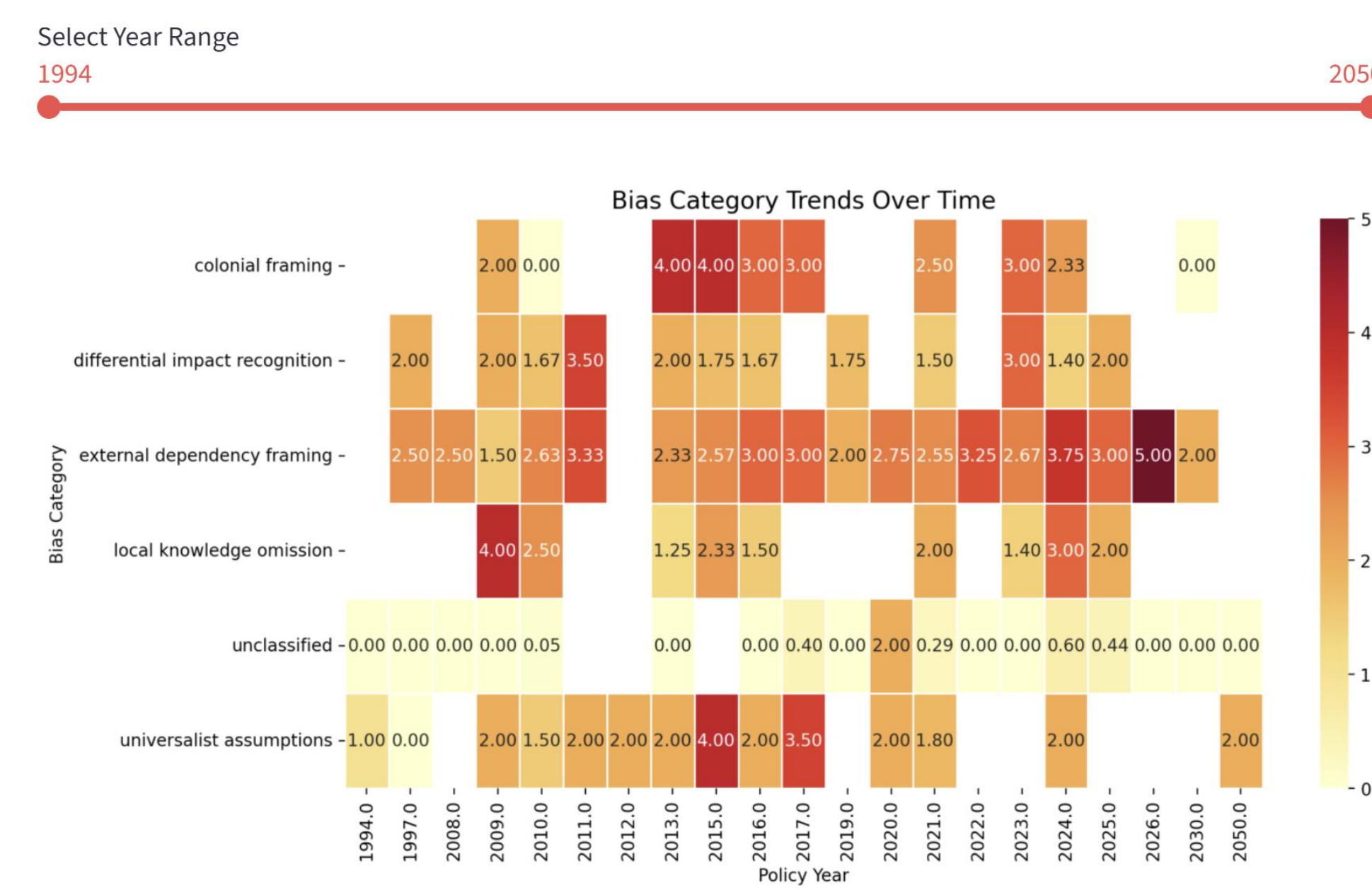
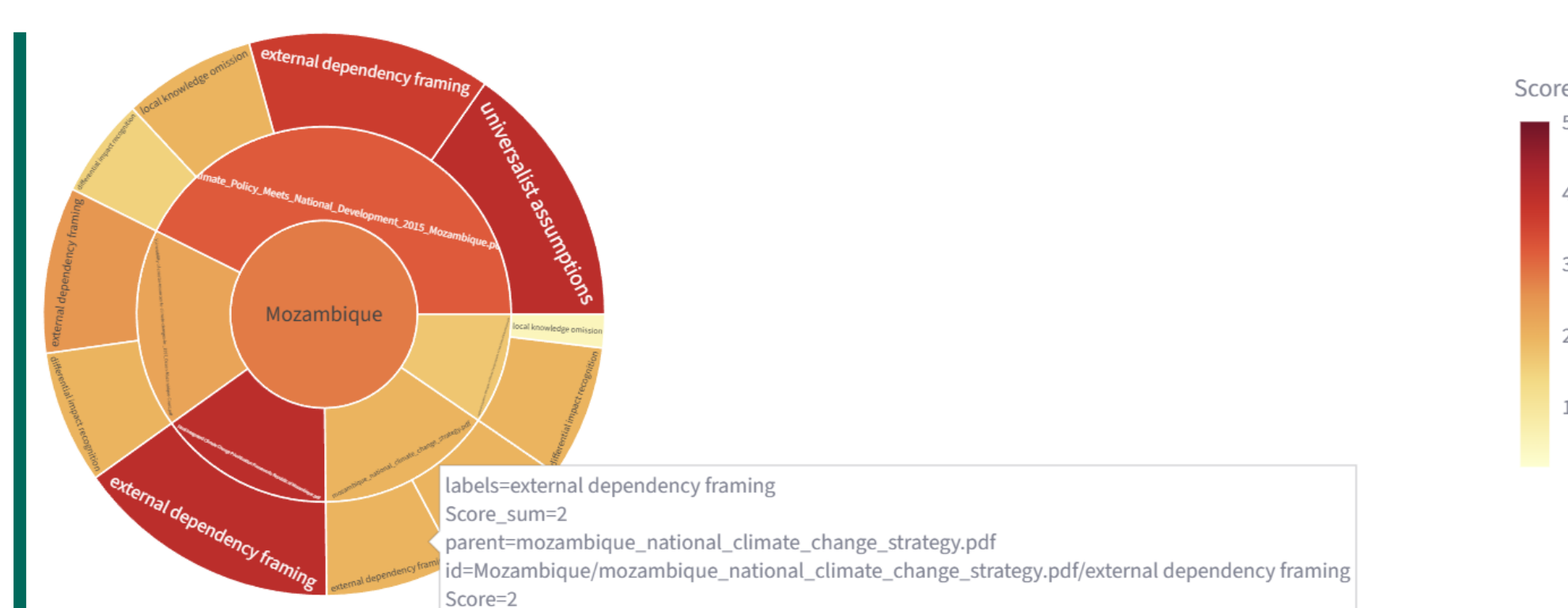
ChatGPT assigned most of its score high value for both local knowledge omission and colonial framing. This suggests higher sensitivity to structural and involvement of policy discourse. Whereas DeepSeek has similar average scores as manual annotation differential impact recognition and for external dependency framing differential, showing it somehow agrees with human interpretations. AfroLM gave similar scoring across all categories.

DISCUSSION

The broader research problem addresses climate policy biases, governance framing and the influence of external institutions Southern African climate discourse. Previously policy analysis methods were manual, complex to scale and quite limited in comparing their capability. Currently, the deployed system addresses these gaps by providing automated classifications, enabling multiple policy comparison while providing better interpretation through visualization.

Next steps:

- Real time policy monitoring by fully automating the system(e.g usage of API-based uploads)
- Usage of other African language models for capturing of indigenous language and adaptation in policy making.
- Real-world deployment for government policy monitoring and regional governance analysis..



Group Name: Byte Squad
Phakisho Kupa, Keabaka Mooketsi

Department of Computer Science

Faculty of Engineering,
Built Environment and
Information Technology

Fakulteit Ingenieurswese, Bou-omgewing en
Inligtingtegnologie / Lefapha la Boetšenere,
Tikologo ya Kago le Theknolotši ya Tshedimošo

Capstone Project - MIT 808

Course Coordinators:
Dr. Vukosi Marivate (vukosi.marivate@cs.up.ac.za)
Abiodun Modupe (abiodun.modupe@cs.up.ac.za)

Scan me

