Contribution ID: **128** Type: **Invited Talk**

® Enhancing Language Translation and Detection for Low-Resource Languages through Sequential and Transformer Models

Thursday, 7 December 2023 11:00 (20 minutes)

The proliferation of digital content has highlighted the disparity in language translation resources, especially for low-resource languages. This research addresses the critical gap in translation and detection technologies for such languages, which is vital for preserving linguistic diversity and ensuring equitable access to information. Our objective is to enhance the accuracy and efficiency of language translation and detection using advanced machine learning models.

We have implemented and compared two different architectures: one serving as a benchmark and the other, transformer models, which boast parallel processing capabilities. These models promise improvements in both translation quality and computational efficiency. Preliminary results indicate that transformer models show significant promise in handling the nuanced structures of low-resource languages.

The implications of this research are profound, offering the potential to democratise information across linguistic barriers and to protect the cultural heritage embedded in language. This study is a step toward bridging the digital divide and fostering inclusivity in the global information ecosystem.

Our presentation at the CHPC National Conference will delve into the methodologies employed, provide a comparative analysis of the models, and discuss the ongoing evaluation of our results. We aim to contribute to the development of more robust and accessible language technologies, particularly for languages that are at risk of digital extinction, including many of South Africa's official languages.

Student or Postdoc?

No. Not a student nor Postdoc.

Primary author: Mrs MBOOI, Mahlatse (CSIR)

Co-authors: Mr NAIDOO, Tyrone (CSIR); RAMBUDA, Thompho (Council for science and industrial research); Mr SEFARA, Tshephisho (CSIR)

Presenter: Mrs MBOOI, Mahlatse (CSIR)

Session Classification: DIRISA

Track Classification: DIRISA