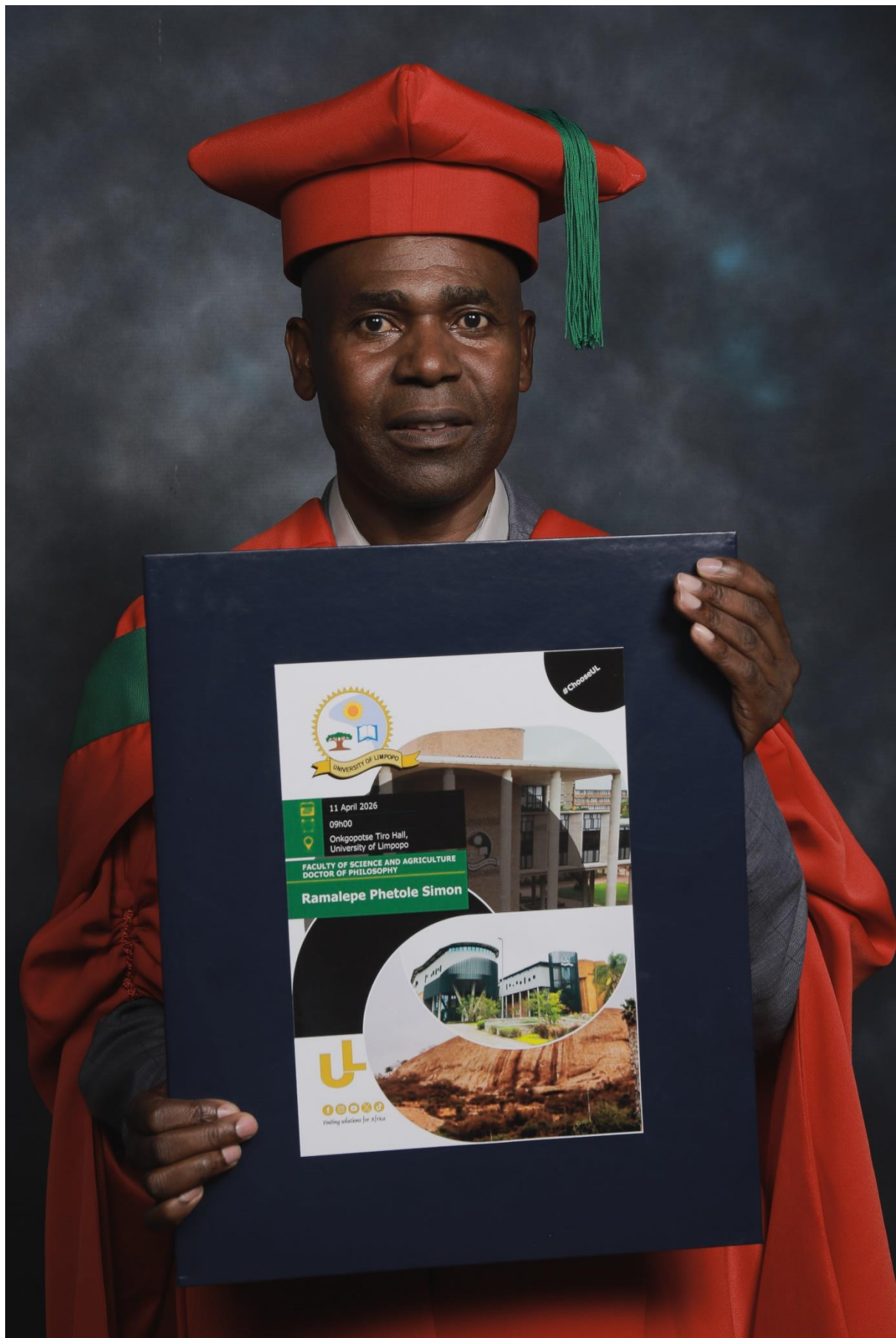




CITATION

Ramalepe Phetole Simon investigated the use of a transformer-based architecture to develop Sepedi-English code-mixed text generation model. A Sepedi radio news corpus was developed to enhance the contextual diversity of a pre-trained baseline model and to adapt it to the news domain. Furthermore, an entity-aware code-mixing framework was designed to generate instruction–response pairs in Sepedi, English, and code mixed forms, resulting in the Sepedi news instruction dataset tailored for an entity-grounded, code-mixed question-answering task. This study presents a methodological advancement in the adaptation of transformer models within low-resource settings. The work has resulted in one high-impact publication and presentations at multiple conferences.