Preface to the Proceedings of RAIL 2025

1 Preface

The sixth workshop on Resources for African Indigenous Languages (RAIL) was held on 10 November 2025 at the CSIR International Convention Centre in Pretoria, South Africa. It was co-located with the Digital Humanities Association of Southern Africa (DHASA) 2025 conference, which took place from 11 to 14 November 2025.

The RAIL workshop series aims to bring together researchers who are interested in showcasing their research and thereby boosting the field of African indigenous languages. This provides an overview of the current state-of-the-art and emphasizes the availability of African indigenous language resources, including both data and tools. Additionally, it allows for information sharing among researchers interested in African indigenous languages and also starts discussions on improving the quality and availability of the resources. Many African indigenous languages currently have no or very limited resources available and they are often structurally different from well-resourced languages, requiring the development and use of specialized techniques. By bringing together researchers from different fields such as computational linguistics, sociolinguistics, and language technology, to discuss the development of language resources for African indigenous languages, we hope to boost research in these fields. The RAIL workshop is an interdisciplinary platform for researchers working on resources such as Natural Language Processing tools, Human Language Technologies, data collections, and annotations, specifically targeted towards African indigenous languages. It aims to create the conditions for the emergence of a scientific community of practice that focuses on data, as well as tools, specifically designed for or applied to indigenous languages found in Africa.

At the sixth edition of the RAIL workshop, we received nineteen high-quality submissions, which were all reviewed by three reviewers. The reviewing process was open, which means that the identity of both authors and reviewers were available to both. Finally, seventeen submissions were accepted (89% acceptance rate). As organizers, we would very much like to thank the programme committee for their in-depth reviewing of the submissions and for providing useful recommendations to the authors.

Note that the aim was to get as many people together to share their ideas and research results, while attaining a high-quality event. This led to a workshop, which consisted of a full day of presentations. This publication adheres to DHET's 60% rule and authors in the proceedings come from a wide range of institutions. In total, the audience listened to sixteen presentations. Each presentation consisted of 25 minutes (including time for discussion). After all the presentations, many interesting discussions took place.

2 Program Committee

Akinola Ayodele James, Michigan Technological University, United States of America

Bekker Martin, University of the Witwatersrand, South Africa

Bombardella Pia, North-West University, South Africa

Calteaux Karen, Council for Scientific and Industrial Research, South Africa

du Plessis Gideon, South African Centre for Digital Language Resources, South Africa

Eiselen Roald, North-West University, South Africa

Gaustad Tanja, North-West University, South Africa

Gouws Rufus, Stellenbosch University, South Africa

Griesel Marissa, South African Centre for Digital Language Resources, South Africa

Groenen Gonneke, North-West University, South Africa

John Henry Chukwudi, Universitat Oberta Catalunya, Spain

Augustine Farinola, University of Alberta, Canada

Josef Langerman, University of Johannesburg, South Africa

Lemeko Papi, Central University of Technology, Free State

Louw Henk, North-West University, South Africa

Louw Johannes Abraham, Council for Scientific and Industrial Research, South Africa

Mabuya Rooweither, South African Centre for Digital Language Resources, South Africa

Madlome Steyn Khesani, Great Zimbabwe University, Zimbabwe

Mahloane Malefu, University of the Free State, South Africa

Marais Laurette, Council for Scientific and Industrial Research, South Africa

Marongedze Reggemore, Nelson Mandela University and University of Zimbabwe

Matfunjwa Muzi, South African Centre for Digital Language Resources, South Africa

Matthew Gordon Derrac, North-West University, South Africa

Naidoo Privolin, Council for Scientific and Industrial Research, South Africa

Ogama Chinedu, MYGEECS, Nigeria

Ogungbemi Olarotimi, University of Texas at San Antonio, USA

Onuoha Chialuka, University of Nigeria, Nigeria

Puttkammer Martin, North-West University, South Africa

Roux Justus C., Stellenbosch University, South Africa

Setaka Mmasibidi, South African Centre for Digital Language Resources, South Africa

Sibeko Johannes, Nelson Mandela University, South Africa

Skosana Nomsa, South African Centre for Digital Language Resources, South Africa

Steyn Juan, South African Centre for Digital Language Resources, South Africa

Trollip Benito, South African Centre for Digital Language Resources, South Africa

Van Zaanen Menno, South African Centre for Digital Language Resources, South Africa

Wilken Ilana, Council for Scientific and Industrial Research, South Africa

Wolff Friedel, South African Centre for Digital Language Resources, South Africa

3 RAIL 2025 organizers and proceedings editors

Rooweither Mabuya

Muzi Matfunjwa

Mmasibidi Setaka-Bapela

Menno van Zaanen