



*The Story of SADiLaR*





## The Story of SADiLaR



# T A B L E    O F



# C O N T E N T S



## T A B L E O F C O N T E N T S

### *The Story of SADiLaR*

Acknowledgements	6
<b>CHAPTER 1:</b> SADiLaR - A brief overview	9
<b>CHAPTER 2:</b> Towards the establishment of a national research infrastructure: collaborations across time	21
<b>CHAPTER 3:</b> The future of SADiLaR: Strategy	61
Appendix I: <b>SADiLaR team</b>	75

---

## Acknowledgements

Thank you very much to Professors Langa Khumalo, Attie de Lange, Justus Roux, Tobie van Dyk and Mister Juan Steyn, who all gave of their valuable time and knowledge. Many thanks are also due to the contributing authors of the SADiLaR Newsletters and content availed in the production of the book.

In addition, SADiLaR Management formally conveys its appreciation for the continued support that the Department of Science, Technology and Innovation (DSTI), SADiLaR's Governance Structures and its Nodes provide. Special votes of gratitude go to the late Dr Daniel Adams, Dr Ntombi Ditlopho, Dr Sagren Moodley, Dr Charles Mokonoto, Ms Nokuthula Mthethwa and Mr Imraan Patel, all from the Department of Science, Technology and Innovation, for advice, practical guidance and support. The South African Research Infrastructure Roadmap (SARIR) has contributed much more than financial resources to SADiLaR for it to continue its work in various pockets of excellence. SARIR's support has ensured the creation of a more unified science infrastructure.

Nicholas Meihuizen (commissioned author)



science, technology  
& innovation  
Department:  
Science, Technology and Innovation  
REPUBLIC OF SOUTH AFRICA



*To be taught in one's mother tongue is a basic human right. However, implementation of the necessary framework to achieve this will require significant resources. These include language technologies like grammar editors and spell checkers, language terminologies for academic disciplines and special skills and expertise to support multilingualism. Without these resources and a way for institutions to pool resources and share strategies and expertise, this framework will be a non-starter.*

Professor Langa Khumalo, Chief Director of SADiLaR.<sup>1</sup>



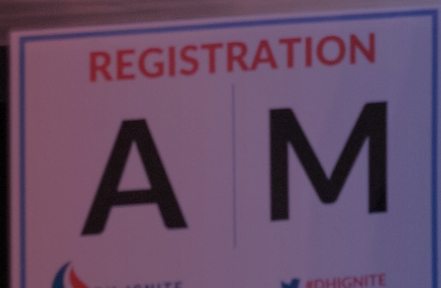
---

<sup>1</sup> SADiLaR. 2021c. <https://sadilar.org/en/sadilar-new-director> Date of access 19 Jul. 2024.





# BRIEF OVERVIEW







# CHAPTER 1

## SADiLaR - *A brief overview*



## **SADiLaR - *A brief overview***

The establishment of the South African Centre for Digital Language Resources (SADiLaR) as a research infrastructure for the promotion of linguistic equity and multilingualism through digital means was premised on the worldwide upsurge of interest and participation in digital humanities, which is the bridge between traditional forms of research practice and digital technology. Locally, it also involves the consolidating of dispersed human language technology initiatives.

The discipline of digital humanities offers scholars new approaches to old problems, including the supporting tools and frameworks necessary for research. It also reveals new problem areas, and provides the means for tackling them. Stemming from this digital upsurge, and underpinning SADiLaR's vision and mission, is an increased availability of digital language resources across disciplines, the growth of Big Data (or extremely large data sets), and access to various types of software (often freely available).



### **SADiLaR's vision**

**To ensure a transformed participative digital future for our official languages for people in all walks of life.**

### **SADiLaR's mission**



1. We enable the use of our official South African languages in research and development activities through active Digital Language Resource aggregation and development towards sustainable multilingualism.
2. We enable the implementation of the new Language Policy Framework for Public Higher Education Institutions through expert tailored support and services towards unlocking epistemological access to knowledge in all official languages.
3. We stimulate and create Digital Humanities communities of practice as an infrastructure.

As a national research infrastructure, SADiLaR is part of the South African Research Infrastructure Roadmap (to be discussed in more detail in Chapter 2). SADiLaR's strategic function is twofold: one, through its digitisation programme it creates, distributes and manages digital language resources and software for all twelve official languages in South Africa (including sign language—recently included among the official languages); and two, through its digital humanities programme it stimulates and supports computational research and development in the humanities and social sciences, and drives human capacity development. The Centre also contributes to activities related to language policy planning and language development and support in

South Africa, especially regarding the implementation of the new Language Policy Framework for Public Higher Education Institutions, to which it is passionately committed and for the success of which it is uniquely situated.

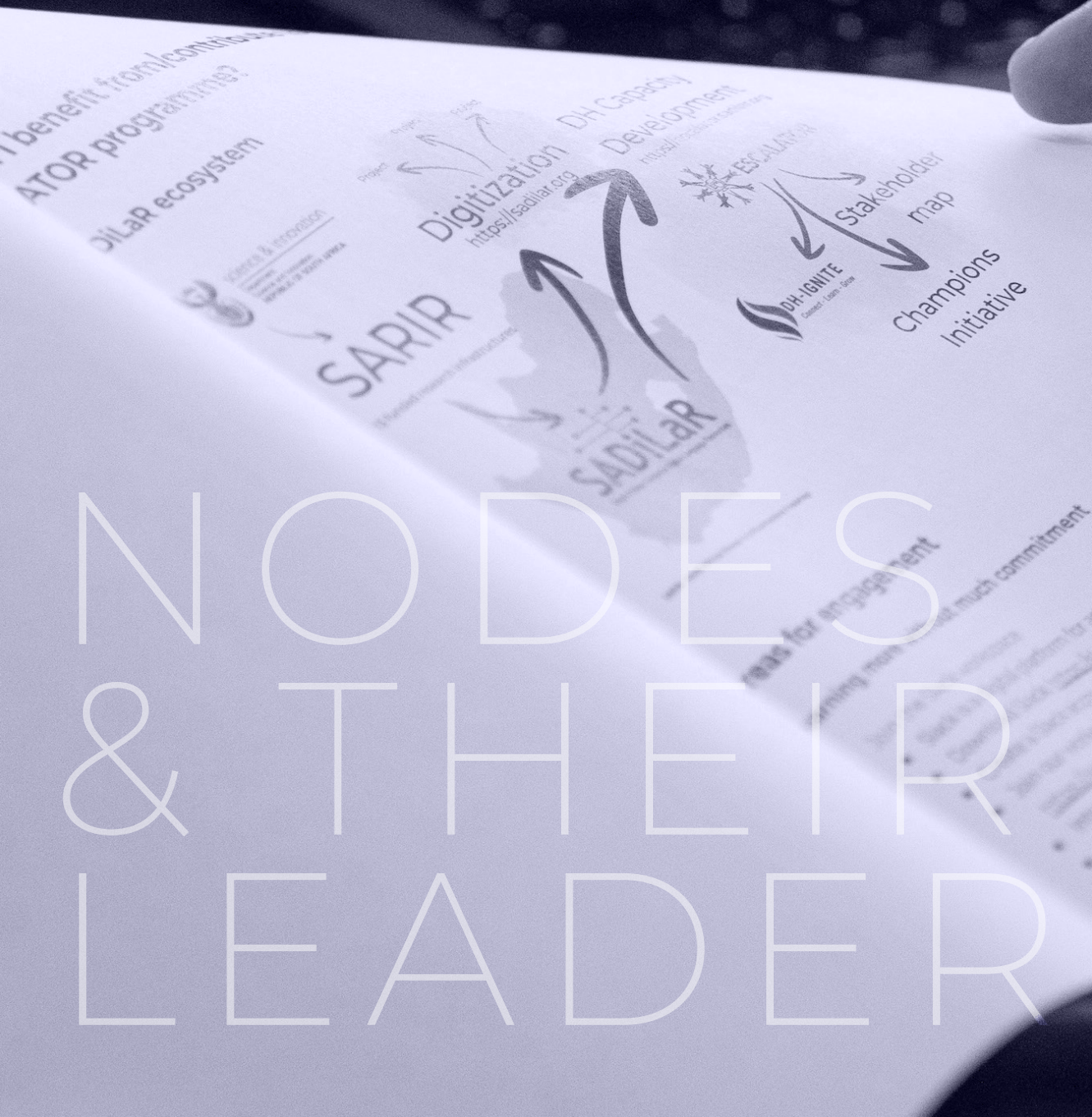
SADiLaR was instituted as a Research Infrastructure during 2016, funded by the Department of Science and Innovation, South Africa. (More details are provided in Chapter 2.) The Research Infrastructure is an independent national Centre, hosted by the North-West University. The 'hub' of the infrastructure is located at this university, but the multiple contributing 'nodes' are to be found at various institutions, following a hub and spoke model.



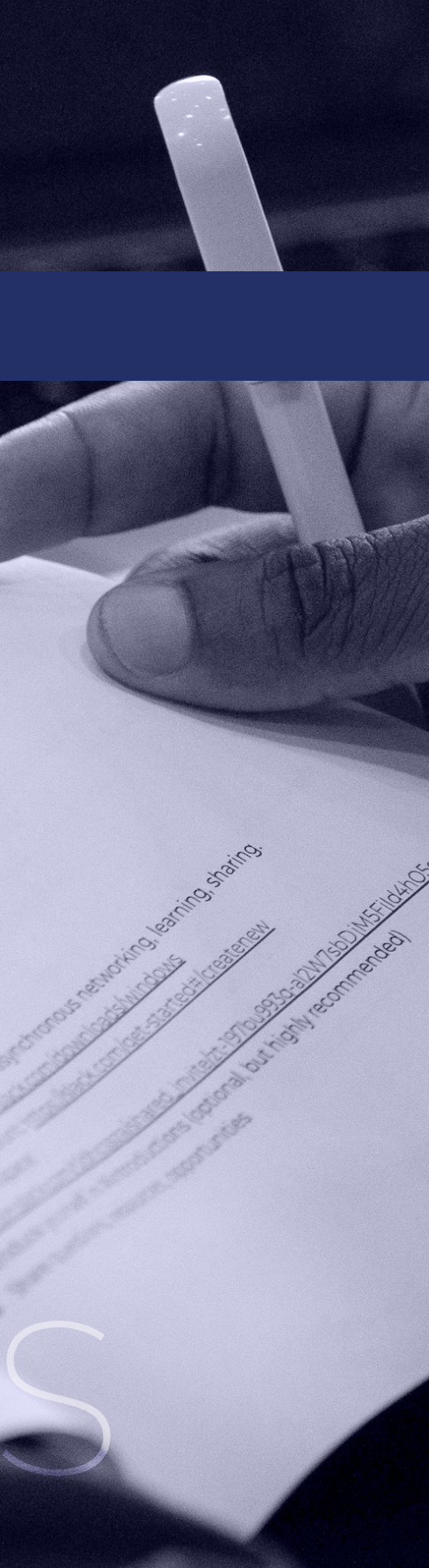
<sup>2</sup> Department of Higher Education and Training (South Africa). 2020. Higher Education Act, 1997 (Act no. 101 of 1997): Language Policy Framework for Public Higher Education Institutions. (Notice 1160). Government Gazette, 43860:3-18, 30 Oct.



THE NODES AND  
THEIR LEADERS



NODES  
& THEIR  
LEADER



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

*The University of Pretoria, Department of African Languages (Digitisation Node), is currently led by Professors Danie Prinsloo and Elsabé Taljard.*

This node **creates language resources for African languages by digitising different kinds of language material built up over many years, including texts, e.g. books, magazines and newspapers, audio and audio-visual material.**

Digitised output is made available to researchers and developers of Human Language Technologies. Included among its valuable materials are historical audio-visual recordings, lexicographic index cards and audio recordings of lectures on linguistics and literature in the official South African languages.







*The University of South Africa, Department of African Languages (Language Resource Development Node), is currently led by Professor Stanley Madonsela and Mrs Marissa Griesel. Former node managers include Professors Sonja Bosch and Mampaka Lydia Mojapelo.*

A principal focus of this node is the **establishment of wordnets for the South African languages, a fundamental resource for the development of core language technologies and their applications.**

A wordnet is a lexical database consisting of words grouped into sets of synonyms. The wordnet indicates various conceptual-semantic and lexical relations, indispensable in automatic language understanding systems. The University of South Africa node also makes multilingual terminology available, especially in the fields of literary studies and linguistics. In addition, the node hosts regular human capital development meetings and training workshops.



*Inter-institutional Centre for Language Development and Assessment—ICELDA (Language Resource Development Node), is led by Professor Tobie van Dyk.*

**This node designs and develops solutions to social issues related to language.**

Different applied linguistic artifacts and services are examples of such solutions: a focus on language testing, language learning and language development for different purposes; a focus on language planning and management matters; building corpora specifically aimed at academic language use; and extending advisory services on best uses of technology regarding multimodal language learning, autonomous learning, blended learning and online delivery. In promoting multilingualism through product design, it also collaborates on and invests in projects with future research and development potential.



*North-West University, Centre for Text Technology—CTexT (Text Node), is led by Professor Martin Puttkammer.*

**This node conducts cutting-edge research in language technology to assist with the development of innovative and relevant technological applications for resource-scarce languages.**

It focuses on the advancement of multilingualism and indigenous languages, both necessary for the establishment of new technologies within Big Data and Artificial Intelligence that are responsive to the unique local context. The node has collected and processed corpora, and developed core technologies such as morphological analysers and part-of-speech taggers for ten South African languages. The node has also produced spelling checkers that are freely available to the public, as well as the open-source Autshumato machine translation systems that provide the public with improved access to information in their mother-tongues, and aid effective public service delivery.



*The Council for Scientific and Industrial Research (Speech Node), is currently led by Dr Laurette Marais. Former node managers include Drs Karen Calteaux and Avashna Govender.*

**This node is involved in localized language technology development, such as in automatic speech recognition, text-to-speech, and natural language processing.**

The text-to-speech application, called Qfrenzy, is the only commercial product of its kind catering for all official local languages. A core focus of the node is applying speech technology in the domain of education, with a view to developing and supporting applications like iSinkwe, designed to help learners overcome reading and learning problems.



*Stellenbosch University, Department of General Linguistics (Child Language Development Node), is led by Professor Heather Brookes.*

**The overall aim of the node is to build a repository of language development data for all South Africa's languages on the SADiLaR platform.**

To do this, a network of researchers across South Africa is working on child language development from birth to adolescence with the aim of making all data available for applied scientists to use to improve language outcomes for South Africa's children. The node's flagship project is an inter-university collaboration focusing on the development of Communicative Development Inventories (CDIs) for all South African official languages.

CDIs are parent report instruments that ask parents/caregivers to report on a child's use of gestures, words and sentences. They can measure language development from 8 to 30 months and are good overall indicators of communicative development. The aim of this project is to get language development norms on which scientists can base the development of tools to accurately assess children's language and identify language delays.





## SADiLaR's domains

SADiLaR supports various projects, which impact on at least the following four domains:



### Language Technology

For the Language Technology domain SADiLaR develops high-level resources and natural language processing tools serving various applications, including: machine translation engines for local languages; automatic speech recognition systems; text-to-speech systems; speech-to-speech translation systems; interactive communication systems; grammar and spelling checkers; and online electronic dictionaries.



### Humanities and Social Sciences

In the Humanities and Social Sciences domain SADiLaR promotes the use of digital data, innovative methods, and software tools that build research capacity and enhance research activities among scholars. The pursuit of previously unanswerable questions and engagement with previously unimagined fields of enquiry become possible.

As an infrastructure SADiLaR is responsive to the imperatives of the Constitution of the Republic of South Africa (1996) and the Use of Official Languages Act (2012) pertaining to the equal development and use of the 12 official languages of the country.





## Socio-economic

The Socio-economic domain is impacted on by reusable digital language resources which are licensed for use in interactive commercial applications in local languages. A living archive documents the nature and use of local African languages along with the cultural heritage practices of different language communities. Also, the more literate a population (implicit in widely available language-enrichment resources), the better the chances for economic improvement.

## Education

The education domain involves a programme geared to multilingual advancement in secondary and higher education, which will unfold over the next decade or so. SADiLaR and its various nodes are in a prime position to inform superior quality higher and secondary education in the fundamental areas of reading, writing and learning. Chapter 3 refers in more detail to this programme.

SADiLaR's commitment to these imperatives can be witnessed throughout the present volume. All these underpin the role of SADiLaR as a non-competing and enabling entity to help ensure a digital future for all official languages in South Africa.



# ESTABLISH- MENT & COLLAB- ORATIONS



## CHAPTER 2

**Towards the establishment of a  
national research infrastructure:  
collaborations across time**



## HISTORICAL BACKGROUND

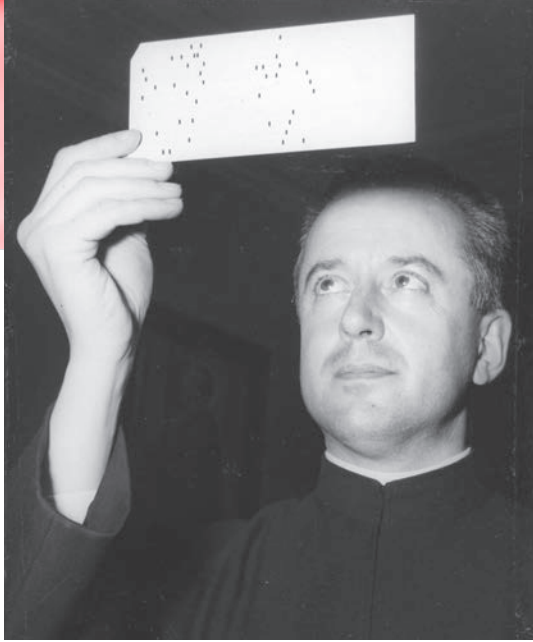
**The story of SADiLaR has its basis in the growth around the world of what has come to be called ‘digital humanities’; locally it built on multiple pockets of excellence and collective efforts in the area of human language technology.**

The growth of digital humanities is largely due to the ever-increasing use of computers and connected devices with internet access and multimodal capturing and analysis capabilities. This is a situation which makes knowledge available on a scale barely conceivable before, and which generates new research methods and questions. As early as the middle of the twentieth century the American inventor Vannevar Bush

imagined a mechanized device he called the ‘memex’, which stored the books, records and communications of an individual, all accessible very quickly; it would be, he wrote, ‘an enlarged intimate supplement to [the user’s] memory’.<sup>3</sup> In 1949 the Jesuit philosopher-priest, Father Roberto Busa, collaborated with IBM to create the Index Thomisticus, a concordance (or a work which situates words and ideas in context) of the eleven million words (involving eleven million punch cards) of Thomas Aquinas.<sup>4</sup> Only the use of computers enabled such a large-scale project to succeed; it was completed in 1980. The text is the largest printed work ever published. The impact of the achievement led to the emergence of the first studies in ‘humanities computing’.

<sup>3</sup> See Bush, V. 1945. As We May Think, in *The Atlantic Monthly*, 176. [www.theatlantic.com/doc/194507/bush](http://www.theatlantic.com/doc/194507/bush).

<sup>4</sup> For a summation of Father Busa’s achievement, see Medium 2024 <https://medium.com/@hamiltonstribley/punch-cards-a-jesuit-priest-chat-bots-and-the-history-of-natural-language-processing-c7a998dd70f> Date of access 30 Jun. 2024. The links to [corpusthomicum.org](http://corpusthomicum.org) give an idea of the influence of this achievement. See Bernot, E. and Enrique Alarcón eds. ‘Index Thomisticus’ web edition. English version. [corpusthomicum.org](http://corpusthomicum.org)



Father Roberto Busa in the 1950s.  
He holds up one of his punch cards.

With the unprecedented rise of the internet in the late 1990s and early 2000s computers could now connect around the world, conveying information quickly and at a low cost. Scholars saw the potential, and several important projects emerged, making library and archives' holdings digitally accessible. From the early 2000s the creation and study of digital texts, images, and sounds (language and music) began to be known as 'digital humanities'. W.G.I. Thomas notes that

the field of digital humanities 'widens the scope of the humanities, opening access to sources, and broadening definitions of scholarly activity'.<sup>5</sup> According to the Kings College London digital humanities site, digital humanities 'forms a bridge between the traditional practices of research and the opportunities afforded by technology, providing scholars with new ways of looking at old problems, and the methods, tools and frameworks to support them in novel avenues of enquiry'.<sup>6</sup> The University of California at Los Angeles site adds that digital humanities 'interprets the cultural and social impact of new media and information technologies', and that it 'creates and applies these technologies to answer cultural, social, historical and

<sup>5</sup> See Thomas, W.G.I. 2016. The Promise of the Digital Humanities and the Contested Nature of Digital, in Schreibman, S. ed, A New Companion to Digital Humanities. Malden: MA and Chichester: Wiley-Blackwell. (Blackwell Companions to Literature and Culture), pp. 523–37.

<sup>6</sup> See Kings College London Digital Humanities. 2024. [kcl.ac.uk/artshums/depts/ddh/about/index.aspx](https://kcl.ac.uk/artshums/depts/ddh/about/index.aspx) Date of access 15 Jul. 2024.





# DIGITAL HUMANITIES

1 10 1001 010101



## THE BEGINNING OF DIGITAL HUMANITIES IN SOUTH AFRICA



**In 1998 attention was drawn to the especial need for reusable digital language resources for research and development in the multilingual South African environment.**

This was achieved by the first personal insight into the fact, a paper by Professor Justus Roux, 'Establishing Speech Resources for the Indigenous Languages of South Africa' (involving in essence HLT), presented at the European Language Resources Association conference in Granada, Spain.<sup>8</sup> Then, in 2001, in close collaboration with electronic engineers and computer scientists at the University of Stellenbosch, Professor Roux led the African Speech

Technology project, sponsored by the Innovation Fund of the Department of Arts, Culture, Science and Technology; the resulting report was titled 'Report on the Development of Human Language Technologies (HLT) in South Africa'. Two further papers stemmed from this report, 'The African Speech Technology Project: Challenges for the Collection of Speech Databases', presented at the 2001 International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques workshop in Aalborg, Denmark. The second paper was titled 'The African Speech Technology Project: An Assessment' and was delivered in 2004 at the international conference on Language Resources and Evaluation, in Lisbon<sup>9</sup>.

<sup>8</sup> Roux, J.C. 1998. Establishing Speech Resources for the Indigenous Languages of South Africa. Proceedings of the First International Conference on Language Resources and Evaluation (LREC) (eds. A. Rubio et al.). Granada, Spain. European Language Resources Association – ELRA. Vol 1, pp.343-350.

<sup>9</sup> Roux, J.C. 2001. The African Speech Technology Project: Challenges for the Collection of Speech Databases. COCOSDA Workshop: Language Resources and Technology Evaluation, Aalborg, Denmark, pp.14-15; Roux, J.C., T. Niesler & P. Louw. 2004. The African Speech Technology Project: An Assessment. Proceedings of the 4th International Conference on Language Resources and Evaluation, Lisbon, pp. 93-93.



Of the Speech Technology project, Professor Roux summarizes: ‘This project highlighted the need for reusable speech resources for the African languages and played a role in the setting up of a Ministerial Advisory Panel, which in turn contributed to several activities which directly and indirectly led to the establishment of SADiLaR’.

Following the appearance of this report local universities interacted with government, as a consequence of which in 2010 Professor Roux, then of North-West University, chaired a task team of the Department of Arts and Culture, set up to create a blueprint for the establishment of a Language Resource Management Agency. In 2012 this resulted in the creation of the language Resource Management Agency, run by the Centre for Text Technology (CTeXT) at the North-West University on behalf of the Department of Arts and Culture. It was a precursor to the establishment of a national resource centre.



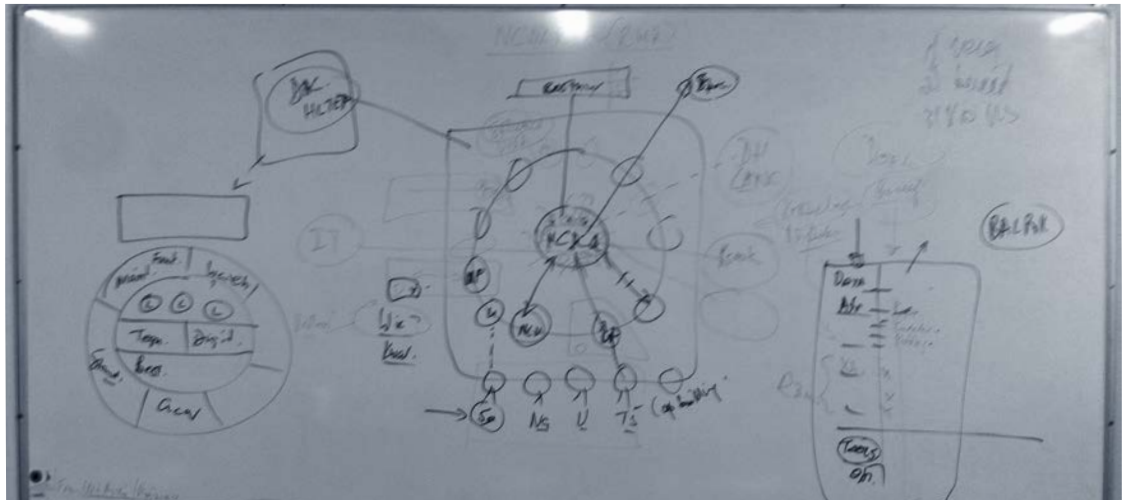
Prof J.C. Roux, Prof J. Swanepoel (Dean), Dr Khalid Choukri (European Language Resources Association President, Paris), Prof Amanda Lourens (Deputy Vice-Chancellor), Dr M. Puttkammer (Director CTeXT), at the inauguration of the Resource Management Agency. Image courtesy of J.C. Roux.

<sup>10</sup> Roux, J.C. 2024. Some personal experiences and activities related to the eventual establishment of SADiLaR [letter]. <sup>26</sup> Aug. 2024.

<sup>11</sup> See DH-IGNITE. 2023. <https://dh-ignite.org/post/2023/02/conversation-justus-roux>. See also, Roux, J.C. 2009. Developing Language Resources in an African Context: The South African Case. Position paper and presentation. FlareNet Forum, Venice, <http://www.flarenets.eu>; Grover, Aditi Sharma, Annamart Nieman, Gerhard B. van Huyssteen and Justus C. Roux. 2012. Aspects of a Legal Framework for Language Resource Management. Proceedings of the 9th International Conference on Language Resources and Evaluation. <http://www.lrec-conf.org/proceedings/lrec2012/index.html>

The Agency (unique in Africa) was intended to run for three years (this was extended to four); it both commissioned projects and mounted its own, and this led to the acquisition, enhancement and distribution of South African language resources and software tools. Its users included not only South African and other African researchers, but international ones as well. **The importance of this enterprise was obvious, as Professor Roux noted,<sup>12</sup> so at this time it was considered strategically imperative to maintain and grow the resources and tools of the Resource Management Agency through the creation of a new centre.** In 2014 the Minister of the Department of Science and

Technology appointed Professor Roux as a 'Champion' (to use the terminology of the Roadmap), to promote the establishment of a National Centre for Digital Language Resources. The first proposal was submitted in November 2014. Presentations were made to three international panels in Cape Town and Pretoria. These were joint committees of the Department of Science and Technology, and the European Commission involved in establishing a new South African Research Infrastructure Roadmap. In 2015 Professor Roux worked on the final proposal for the National Centre in collaboration with various colleagues at North-West University.



Working illustration of the intended 'hub' and 'node' structure of SADiLaR presented in the proposal for the National Centre. Image courtesy of J.C. Roux

<sup>12</sup> DH-IGNITE. 2023. <https://dh-ignite.org/post/2023/02/conversation-justus-roux> Date of access 8 Aug. 2024.



In 2016 the final meta-report was presented to the South African Research Infrastructure Panel of the Department of Science and Technology, Pretoria (now the Department of Science and Innovation). The result was a long-term grant from the Department of Science and Technology in terms of the South African Research Infrastructure Roadmap<sup>13</sup>. Professor Roux notes:

*Usually, these roadmaps only support the natural sciences and not the humanities. We made a presentation to get onto the roadmap so that we could systematically develop digital resources for all official languages, to be used in various language technologies like machine translation and speech recognition.*

Thus, it was Professor Roux who spearheaded the establishment of SADiLaR after the announcement that SADiLaR would form part of SARIR in 2016. He was Director of the Centre until his retirement at the end of 2017<sup>14</sup>. Since retiring, Professor Roux continues to promote digital humanities.

**He was instrumental in helping to establish the Digital Humanities Association for Southern Africa in 2016, for which he served as President for a term.** His expertise in digital humanities and human language technology is recognized globally, and he has contributed to many professional bodies over the years, leading South African delegations to annual International Organisation for Standards Technical Committee

meetings, presenting at international workshops and participating in international scientific committees.



Professor Justus Roux, championed the establishment of SADiLaR, here delivering a keynote at a DH-IGNITE event

<sup>13</sup> A life cycle of 15 years is envisaged with funding of R141 000,000 for the first five years. In Roux, J.C. 2024. Some personal experiences and activities related to the eventual establishment of SADiLaR [letter]. 26 Aug. See more on the Roadmap in the following section.

<sup>14</sup> For a brief biography of Justus Roux and the source of this quotation, see Academy of Science of South Africa Repository: Legends of South African Science: Justus Roux <https://research.assaf.org.za/assafserver/api/core/bitstreams/a1f16261-5138-4b9f-983a-346d4253042c/content>. In a personal note, Professor Roux emphasises the importance of his collaborators in this enterprise.



## THE SOUTH AFRICAN RESEARCH INFRASTRUCTURE ROADMAP

The South African Research Infrastructure Roadmap was introduced by the then Minister of Science and Technology, Dr Naledi Pandor, who noted that ‘access to adequate and relevant research infrastructure is essential to promote the outcomes and quality of research’<sup>15</sup>. She continued:

*Through a joint agreement between South Africa and the European Union, the South African Research Infrastructure Roadmap framework was developed as the basis for prioritising the development of national research infrastructure needs.*

Science and research planners and policy makers had begun to consider the systematisation of research infrastructures since at least 2008, when the G8+5 group of countries established a committee to promote them. Such infrastructures have become widespread in the world of research for obvious practical reasons, centred in the need for a responsible, dependable research environment: **i) they allow funders to budget for specific, expected results; ii) they compel researchers to collaborate for the national good; and iii) they allow researchers to plan future research on the basis of what is likely to be supported**<sup>16</sup>.

The South African Roadmap was conceived as a ‘high-level strategic and systemic intervention to provide research infrastructure across the entire public research system, building on existing capabilities and strengths, and drawing on future needs’<sup>17</sup>.

<sup>15</sup> South African Research Infrastructure Roadmap (first edition). October 2016, p.ii. See [www.dst.gov.za](http://www.dst.gov.za) Date of access 05 Jul. 2024.

<sup>16</sup> Roadmap, p.5.

<sup>17</sup> Roadmap, p.ii.



**Six areas were considered to be in need of research infrastructure development:**

- i) humans and society
- ii) health, including biological and food security
- iii) the earth and the environment
- iv) materials and manufacturing
- v) energy
- vi) the physical sciences and engineering

**Of the 17 research infrastructures originally recommended, 13 were finally selected, following a rigorous reviewing process. What was to become SADiLaR was one of them, under the umbrella of ‘humans and society’, and is listed in the Roadmap as ‘a national centre for digital language resources’.<sup>18</sup>**

As previously noted, SADiLaR was also developed in response to a deep concern expressed in the Constitution

of the Republic of South Africa Section 6, number 108 of 1996, regarding ‘the historically diminished use and status of the indigenous languages of our people’, and the consequent need for the state to take ‘practical and positive measures to elevate the status and advance the use of these languages’. Further, there is a need to ‘promote, and create conditions for the development and use of **i) all official languages; ii) the Khoi, Nama and San languages; and iii) sign language**’.<sup>19</sup>

<sup>18</sup> Roadmap, p.3.

<sup>19</sup> Constitution of the Republic of South Africa 1996.

Of additional significance in the establishment of SADiLaR was the Use of Official Languages Act, 2012, which includes in its brief the need 'to provide for the establishment and functions of a National Language Unit', and to 'ensure that the language unit is provided with human resources, administrative resources and other resources necessary for its effective functioning'<sup>20</sup>. SADiLaR's continuing challenge, then, is to contribute to the development of the country's official languages and to their functional use, through, for example, the deployment of technologies to foster in a practical way effective multi-lingual communication, and thus social unity. Such development entails long term maintenance and enabling access to large corpora of digital data and software tools to interact with the data, as well as the knowledge to acquire these skills, with the ultimate aim of benefitting society at large, academia, industry and government. As the

Roadmap indicates, fundamental to the country's economic advancement is a 'transition from a resource-based to a knowledge-based' environment. This transition is necessary because 'the economy's dependence on raw materials, cheap labour, propriety production technology and privileged access to markets is not sustainable in its current form'.<sup>21</sup>

As part of the later development of the Decadal Plan of the Department of Science and Innovation, SADiLaR directly contributes towards the creation of a scientifically aware and literate society through its strategic focus and specialized projects. It contributes towards the increased access to human-friendly language-related technologies that can be applied to address the country's digital, language and literacy divides; and it contributes to socio-economic development.



<sup>20</sup> Use of Official Languages Act, 2012 (Act no. 12 of 2012): Government Gazette 36392, 26 April.

<sup>21</sup> Roadmap, pp.14-15.





## NATIONAL ACTIVITIES AND INTERNATIONAL INTEGRATION AND RECOGNITION

---

### The Digital Humanities Association of Southern Africa

The Field of Digital Humanities is intrinsically multidisciplinary and at its core involves multiple scholars from various backgrounds who come together in a research space and with collaborative projects, which effectively enable and inform novel transformative research in the humanities, consistent with the aspirations of the Decadal Plan.<sup>22</sup>

SADiLaR and many of its funded and current staff members, the NWU host, and other collaborators have played an important role in the growth and adoption of Digital Humanities as a field. Various factors led to the establishment of the Digital Humanities Association of Southern Africa in 2016 at the second National Workshop hosted in Potchefstroom from 4 to 8 April 2016. **These were:** key developments at the North-West

University Research Unit for Literature and Languages in the South African Context; the approval of the creation of a Digital Humanities theme by the then NWU Senate and Institutional Management; and subsequent seed funding, which enabled the hosting of two national workshops and capacity development activities that led directly to the formation of the Digital Humanities Association of Southern Africa.

---

<sup>22</sup> Decadal Plan <https://www.nrf.ac.za/wp-content/uploads/2023/06/STI-Decadal-Plan-2022-23-14Dec22.pdf> Date of access: 14 Oct 2024.



2014

First local awareness campaign (with international guests) + weekly discussion sessions

2015

First DH Workshop SA (23-25 February) with 5 High profile international contributors  
105 participants / plenaries & software training workshops  
Association for DH approved in principle a steering committee was elected to develop a constitution

2016

Second DH Workshop SA (4-8 April) 2 High profile international contributors  
65 participants / plenaries & software training workshops and the founding of Digital Humanities Association of Southern Africa (DHASA) as a Forum the constitution was also accepted

The inaugural conference of the Digital Humanities Association of Southern Africa was held in January 2017 at Stellenbosch University; the Association subsequently hosted successful conferences in 2019, 2021 and 2023. The first conference brought together local and international digital humanities scholars, including Dr Glen Worthey, then Digital Humanities Librarian and Co-leader of the Center for Interdisciplinary Digital Research, Stanford University Libraries, Dr Mark Algee-Hewitt, co-director of the Stanford Literary Laboratory, Professor Paul Arthur, Chair of Digital Humanities and Social Sciences at Edith Cowan University, Australia, and Dr Pip Willcox, then Head of the Centre for Digital Scholarship at the Bodleian Libraries, Oxford University (now Head of the Lambeth Palace Library, UK).

**Topics discussed included: how to create a digital edition; ways to improve user experiences in**

**digital humanities interfaces; automatic text processing tools for South African languages; and orthography-based language modelling for speech recognition.**

The Digital Humanities Association of Southern Africa would, among other matters, promote the understanding and practice of digital approaches to humanities scholarship across the southern African region, and work towards a methodological commons for the digital humanities in this region, providing guidance in the development of standards and expertise, in order to ensure best practice in digital humanities teaching and research. During August 2017 the international Alliance of Digital Humanities Organisations<sup>23</sup> welcomed the Digital Humanities Association of Southern Africa as an observer member, and in 2018 as a full member within ADHO.


<sup>23</sup> The Alliance of Digital Humanities Organisations is a consortium of organisations, which 'promotes and supports digital research and teaching across all arts and humanities disciplines'. See Alliance of Digital Humanities Organisations. 2024. <https://adho.org>.



## NATIONAL ACTIVITIES AND INTERNATIONAL INTEGRATION AND RECOGNITION

---

### CLARIN membership



In November 2018 CLARIN ERIC<sup>24</sup> welcomed SADiLaR as the first observer outside Europe (after a rigorous reviewing process—see Professor Attie de Lange’s comments in the following section).

SADiLaR has since become a fully-fledged member, a decision fully supported by the CLARIN General Assembly at its December 2023 meeting. The CLARIN web page describes the infrastructure as ‘a digital infrastructure which provides easy and sustainable access to a broad range of language data and tools to support research in the humanities and social sciences, and beyond.

CLARIN provides access to multimodal digital language data (text, audio, video) and advanced tools with which to explore, analyse or combine these datasets’. Participating centres are found ‘all over Europe and further afield’, including ‘universities, research centres, libraries and public archives’; SADiLaR is now one of these centres—the only one in Africa. The site notes that ‘all members and observers create and provide access to digital language data collections, digital tools and expertise’<sup>25</sup>.

---

<sup>24</sup> CLARIN stands for Common Language Resources and Technology Infrastructure, and ERIC stands for European Research Infrastructure Consortium. See <https://www.clarin.eu>.

<sup>25</sup> See CLARIN ERIC <https://www.clarin.eu>. Date of access 30 Jun. 2024.



## THE FORMAL LAUNCH OF SADiLaR

**On 7 August 2019 SADiLaR  
was officially launched.**

The announcement from the SADiLaR office noted that ‘the Centre, having been in its incubation and development phase since 2016, has now reached maturity to the level where it was officially launched by the Deputy Minister of Higher Education, Science and Technology, Buti Manamela’.

Speaking at the launch, the Deputy Minister said the global paradigm shift in information and communication technologies ‘has given rise to the ever-expanding interdisciplinary domain of research and development referred to as digital humanities’, SADiLaR being ‘one of government’s responses to the fourth industrial revolution’. Professor Attie de Lange, the then Director of SADiLaR, spoke of the Centre’s enabling function with a focus on all official languages of the country, serving all tertiary institutions.

## He emphasised issues of continuing importance for the Centre:

We regard it as our responsibility to support the research and development in the domains of language technologies and language-related studies.

With our training and workshops in the use of the digital infrastructure and digital tools at national level, together with our partners comprising the Council for Scientific and Industrial Research, the University of South Africa, the University of Pretoria, the Inter-institutional Centre for Language Development and Assessment (ICELDA), and the NWU's Centre for Text Technology (CTexT), SADiLaR aims to unlock the full spectrum of resources to create a broader digitalized footprint of South African indigenous languages.

**We have a very strong focus towards the training of new generation researchers and will also engage in the development of new nodes; that is, collaborative centres<sup>26</sup>.**



Professor Attie and Honourable Buti Manamela at the official launch of SADiLaR

<sup>26</sup> See SADiLaR. 2019. <https://sadilar.org/en/sadilar-launch/>. As indicated in Chapter 1, Stellenbosch University has now also become a SADiLaR partner.



In a personal interview, Professor de Lange spoke of his role in enabling SADiLaR to move from the conception stage to that of a practising entity.<sup>27</sup> The founding process involved establishing the physical infrastructure to deliver the services to be offered. The right people needed to be appointed for each of the official languages and the managerial structure needed to be put in place.

He emphasised that SADiLaR is the first humanities unit in the national research infrastructure system. As such, it was difficult to explain to natural scientists the needs of such a structure according to the requirements of the humanities and social sciences; that is, the qualitative research skills and abilities necessary, compared with the more streamlined, quantitative approaches of the natural sciences. Professor de Lange also had to liaise

with various stakeholders—universities and institutes—to promote SADiLaR and to join this entity as nodes within the context of SADiLaR’s mandate. He also oversaw the legal requirements necessary, the contractual basis of employment at the centre, the rules of a steering committee, and the setting up of project contracts at various institutions. Regarding funding from the Department of Science and Technology, schedules had to be adhered to within a three-year timeline with a pending review after the establishment phase to determine whether or not future funding would take place, a fact which involved intricate timing and overseeing.

A major exercise was to make sure that all the specifications pertaining to SADiLaR met the demanding standards necessary for linking up with CLARIN.



<sup>27</sup> De Lange, A. 2024. Enabling SADiLaR to move from the conception stage to that of practising entity [Personal interview]. 23 Aug., Potchefstroom.

**SADiLaR has been responsive to input and recommendations from partners and stakeholders. Most noteworthy was the recommendations received from two external evaluations that took place. This input as well as new developments in the field informed the creation of SADiLaR's new strategy 2024 - 2029**

This process was started in Professor de Lange's time with SADiLaR, leading to an initial CLARIN membership phase. As noted previously, SADiLaR is now a full member of CLARIN. Professor de Lange also had to communicate 'across the borders of disciplines and groups' in order to promote the Centre. Training sessions for research groups in the use of digital methodologies were a priority. In short, in his two-and-a-half-year tenure with SADiLaR,

Professor de Lange had to manage the issues involving personnel, strategies, and contact with relevant stakeholders such as engagement with publishing houses. He emphasised the extraordinary contributions made by his colleagues, Juan Steyn and Roald Eiselen, without whom SADiLaR could not have become the entity it is today. Professor de Lange also paid a special tribute to Justus Roux, the tireless instigator and founder of SADiLaR.



## A new SADiLaR chief director: Professor Langa Khumalo

In November 2020 the first SADiLaR Newsletter in all the official South African languages was launched. On 1 August in the same year Professor Langa Khumalo was appointed as head of SADiLaR. He holds a PhD in Linguistics from the University of Oslo and an MPhil in Linguistics from Cambridge. Professor Khumalo is proud of the fact that he was trained by the late Professor John Sinclair at the Tuscany Word Centre.

An award-winning author in the field of Intelligent Text Processing and Computational Linguistics, Professor Khumalo became one of the first editors to compile and publish the first corpus-aided monolingual dictionary of isiNdebele; he was also involved in the development of disciplinary terminology for isiZulu, which led to the publication of glossaries such as the illustrated glossary of terms in architecture and the glossary of terms in law. **At his inauguration he noted that ‘language is at the centre of our human existence’.** He continued:

**Language carries and conveys the identity of humanity. It records our human achievements and spurs on our innovation and development. It is notable that our Constitution recognizes eleven official languages. It is a constitutional imperative to develop all of them in order to achieve ‘parity of esteem’ between them<sup>28</sup>.**



Professor Langa Khumalo, Chief Director, SADiLaR

<sup>28</sup> SADiLaR. 2020. <https://sadilar.org/en/sadilar-launch/> Newsletter-nov2020. Date of access 19 Jul. 2024. Since this speech was made, sign language has become the twelfth official language of South Africa, as previously noted.



**In *The Bloomsbury Handbook of Digital Humanities*, Professor Khumalo and Titilola Aiyegbusi note, in a similar vein:**

*In attempting to define language we refrain from using a litany of definitions found in the literature on Language and Linguistics that are drawn from Eurocentric theories. In appreciating what language means, we defer to the explication of mother tongue in isiZulu as simply ulimi lwebele (literally, the language that one sucks from their mother's breast). Mother tongue in this sense means one's first language. In isiNdebele, when a child utters their first word, the elders exclaim: Sekungumuntu! (meaning, It is now a human being). Thus, language is an essential part of that which makes us human; it is at the heart of our human existence.<sup>29</sup>*

<sup>29</sup> Aiyegbusi, Titilola Babalola and Langa Khumalo, Digital Humanities Outlooks beyond the West. In: O'Sullivan, James ed. 2022. *The Bloomsbury Handbook of Digital Humanities*. London: Bloomsbury Academic. p.31.



The pertinent local examples in the above passage speak to the commitment to multilingualism at the heart of SADiLaR's mission. Professor Khumalo's driving perception is the attainment of epistemic freedom, or 'the freedom to think and create knowledge', whether from the 'centre' or the so-called 'periphery' of knowledge generation, which would include the Global South.

His views impact on the current local drive to decolonize curricula, adding a valuable nuance: 'The decolonization and transformation of the curricula in Africa and Asia ... would require the incorporation of epistemic perspectives, knowledge, and thinking from the Global South into the growing global knowledge

economy from which everyone draws. Decolonisation is, thus, a call for plurality in sources of knowledge, and not the current negation of "other" worldviews. By embracing rather than rejecting, a unified digital humanities may realize a noble purpose of enriching and harmonizing knowledge and learning and supporting equity in development'<sup>30</sup>.

**Under Prof Khumalo's leadership, with the support of SADiLaR's governance structures, a first strategic plan was created along with the establishment of SADiLaR's governance charter towards a fully operational national research infrastructure.**



<sup>30</sup> Aiyegbusi, Titilola Babalola and Langa Khumalo, Digital Humanities Outlooks beyond the West. In: O'Sullivan, James ed. 2022. The Bloomsbury Handbook of Digital Humanities. London: Bloomsbury Academic. p.37.

## SADiLaR projects, programmes and activities

Since its inception SADiLaR has been proactive in building new networks and being responsive to developments in the Humanities and Social Sciences.

To date SADiLaR has funded multiple multiyear specialisation projects run by its nodes and projects run by the broader HSS community in South Africa. The specialisation projects run to date link to the various specialisation areas of SADiLaR's nodes, which cover text and speech processing, language development (and in particular child language development), language testing, digitisation and the creation of language resources such as wordnets, used in natural language processing.

These node projects aim to systematically address the under-resourced state of our African Languages. All node projects are also reviewed by our Scientific Advisory Committees and undergo formalised review processes. **South African languages can only have a digital future through the continued development of the technological building blocks required to construct representative human language technologies.**

## OPEN CALL PROJECTS

Follow the QR code to SADiLaR's website for a more detailed account of all projects and activities to date.



**SADiLaR announced funding opportunities for open call projects during 2018. The aim of these projects was to stimulate broader participation by the HSS sector in research and resource building related to SADiLaR's Digitisation and Digital Humanities programme.**

Of the open call projects most noteworthy was SADiLaR's first additional node, the Child Language Development node, which was established in 2020. Open calls ranged from transcription projects, child language data collection, word embeddings and language documentation work, to name a few areas. More specialised open call projects were also run under SADiLaR's Digital Humanities Open Educational Resources Champions project.

The DH-OER project, run as part of SADiLaR's ECALATOR programme, promoted the creation and use of Open Educational Resources for digital humanities at universities throughout South Africa. 26 projects were accepted into this programme, to help foster awareness of how important it is to create local context-specific, open-educational resources that will stimulate digital humanities-related research and teaching and learning in South Africa.

## COLLABORATIONS, CONFERENCES AND CREATING NEW CAPACITY IN THE HSS

SADiLaR during its early years made a concerted effort to raise awareness not just about the role of the Infrastructure but also about the broader need to address the limited resources available for African Languages as well as to improve the standard of computational skills in the HSS.

Many collaborative initiatives took place with international partners and entities as part of broader international strategies, such as the UNESCO Year of Indigenous Languages, the SADiLaR-Wikipedia-PanSALB (SWiP) Project and the Technologies for Enhanced Documentation of African Languages: Creating Synergies project, to name a few. To date the infrastructure has run 180 events, which ranged from general awareness to hands-on capacity building.

SADiLaR's long-standing collaboration with The Carpentries played an important role during the Infrastructure's establishment years to further the development of foundational computational skills in South Africa and the broader African continent.

Providing access to resources and technologies is only part of SADiLaR's resource-creation mandate. SADiLaR as a national research infrastructure also played an important role in stimulating the number of platforms that are available for the dissemination of Digital Humanities-related activities; SADiLaR's own digitisation programme is also significant.

**SADiLaR has been instrumental in ensuring the successful hosting of multiple conference of the Digital Humanities Association of Southern Africa to provide avenues for digital humanities-related research to be shared.**

This emphasis on sharing research led to the establishment of the Journal of the Digital Humanities Association of Southern Africa<sup>31</sup>. The SADiLaR Digital Humanities colloquia series was established during the COVID-19 pandemic, to provide an informal virtual forum for dissemination of digital humanities-related activities in South Africa and beyond. By October 2024 more than 40 DH colloquia had been hosted, covering a very broad and comprehensive list of topics.



<sup>31</sup> Digital Humanities Journal of South Africa <https://upjournals.up.ac.za/index.php/dhasa>





Follow the QR code to DHASA's journal and conferences summary

SADiLaR's DH Professor, Prof Menno van Zaanen, played an instrumental role in establishing the annual Resources for African Indigenous Languages Workshops. RAIL provides an interdisciplinary platform for researchers working on resources (data collections, tools, etc.) specifically targeted at African indigenous languages; it aims to foster the conditions for the emergence of a scientific community of practice that focuses on data, as well as computational linguistic tools specifically designed for or applied to African indigenous languages.



Follow the QR code to summary of all RAIL workshops to date

The Research Infrastructure also plays an important role in actively stimulating the growth of traditional research publications and conference presentations. An up to date list of publications and conference presentations to date is available online with a QR to page.



Follow the QR code to summary of all research outputs.



## The value of a research infrastructure

**Though SADIaR's mandate mainly pertains to the official South African languages, the value of an infrastructure providing open access to technological building blocks can be seen in relation to other projects that can leverage the already created resources and platforms in a way that would not have been possible using traditional means.**

One such project that shows in practical terms the importance of building on an existing infrastructure was the creation of a Digital Dictionary of the N|uu Branch of the !Ui-Taa Language. This branch of the language of the ǀKhomani people, traditional hunter-gatherers from the southern Kalahari, is severely endangered. According to researchers Kerry Jones and Bonny Sands, as far as linguists know, only one fluent mother tongue speaker of N|uu is still alive, ǀXuu |Eeki Katrina Esau, known as 'Ouma'. With the aim of consolidating linguistic resources in N|uu collected over approximately 20 years, the N|uu

Language Authority, African Tongue and SADIaR worked together on the project, 'Digital Dictionary Resources for N|uu'. Historical and contemporary research combined to produce a dictionary in three formats, a limited-edition physical dictionary, a digital one and a smartphone application. The project was awarded the 2023 Afrikaans Language and Culture Association Wordwise Prize in recognition of an outstanding contribution to the Afrikaans language.

This achievement aligns well with the original aim of the South African Research Infrastructure Roadmap, which notes that 'access to adequate and relevant research infrastructure is essential to promote the outcomes and quality of research. In order to maximise the return from investment in research, scientists and researchers must have access to modern and appropriate infrastructure'.<sup>32</sup> This access is facilitated primarily via SADIaR's Language Resources Repository.



<sup>32</sup> Roadmap, p.ii.

The Repository is an extension of the Resource Management Agency established by North-West University's Centre for Text Technology (CTeXT), previously mentioned. **The Repository is concerned with the curation, distribution and maintenance of reusable digital resources (text and speech) for the official languages of the nation.**

This data plays a significant role in the development of facilitatory tools for communication among the different language groups, and it also plays a role in placing local languages within the realm of contemporary communication systems. The repository contains over 400 records of items in different languages, some from outside South Africa, all available to interested parties through open access.



Dr Friedel Wolff, SADiLaR's technical manager, involved with the Language Resource Repository

Dr Friedel Wolff, SADiLaR's technical manager, says that 'some of the items themselves describe a resource that is itself multilingual or, for example, software that supports several languages', and which 'might just be what some researcher or software engineer needs to build something exciting' for a local language. Resources include electronic text and speech data, such as domain-specific text collections, wordlists, dictionaries, translation memories and aligned multilingual corpora. Also to be found are multimodal resources and tools, applications and platforms, all of which support the processing of data along with the development of new technologies.

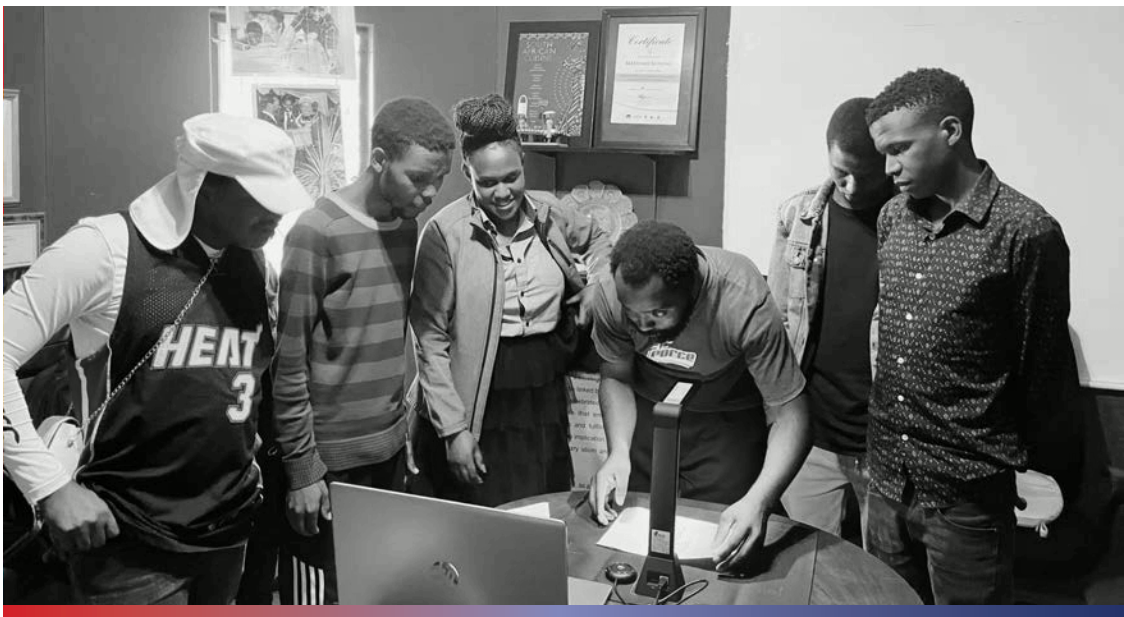
Dr Wolff goes on to point out the data was time-consuming to create, and expensive, and some of it 'required expert knowledge or computing power' available to few. Of great importance is the fact that the resources are given the degree of permanence afforded by '**a centralised access point**'. Dr Benito Trollip adds that the repository should help engender 'a mindset of sharing data to enable and empower the community at large'.

Dr Laurette Marais, manager of SADiLaR's speech node, at the Council for Scientific and Industrial Research, notes that the repository 'has become an integral tool in the planning and execution of the research agenda' of the CSIR Voice Computing research group. This group contributes, for instance, speech data 'aimed at training automatic speech recognition systems'.

She feels that 'any student or researcher in language technology in South Africa should be familiar with the repository and what it has to offer, especially given the resource-scarce nature of our languages'.<sup>33</sup>



Dr Laurette Marais, manager of SADiLaR's speech node at the Council for Scientific and Industrial Research



<sup>33</sup> SADiLaR. 2023b. <https://sadilar.org/en/sadilar-Newsletter-september-2023> Date of access 20 Jul. 2024.





## OPPORTUNITIES FOR GROWTH

### Building computational human capacity: The ESCALATOR programme

Stemming from a need to integrate all SADiLaR training and awareness events the ESCALATOR programme was established in 2020. This is a national programme to grow an inclusive community of practice in digital humanities and computational social sciences in South Africa. ESCALATOR was developed by SADiLaR. Special thanks to Mrs Anelda van der Walt, who through Talarify<sup>34</sup> as partner facilitated and developed the programme as first programme manager.

**By October 2024 SADiLaR had been involved in more than 180 events**



One of its major functions is the Digital Champions initiative, which offers mentoring to scholars and students in digital research by experts and peers from the public research centres and universities. The first intake of mentors and mentees took place in the middle of 2021.<sup>35</sup> ESCALATOR wishes to guarantee the sustainability of the community of practice through this mentoring. It is hoped that the 'digital champions' and other community members will develop the necessary skills and confidence through their own work. The programme is closely aligned with Centre initiatives, and institutional, regional, national and international digital capacity and community development.

<sup>34</sup> See Talarify <https://www.talarify.co.za>

<sup>35</sup> SADiLaR. 2021a. <https://sadilar.org/en/Newsletter-jun-2021> Date of access 18 Jul. 2024.



## DH-IGNITEs

---



**A recent initiative of the SADiLaR ESCALATOR programme is DH-IGNITE, mounted to develop a locally inclusive, working community in digital humanities and computational social sciences.**

The first in a series of regional events planned across South Africa took place in October 2022 at Umhlanga in KwaZulu-Natal, the COVID-19 restrictions having been lifted. The goal of DH-IGNITE is not to teach coding skills, but to introduce researchers to the ranges of options for learning these skills and to show their potential to enhance work. Programme manager, Anelda van der Walt, says DH-IGNITE is ‘a five-to-ten-year project’ to enable computational social sciences and humanities to ‘upskill for the fourth industrial revolution’.

At a DH-IGNITE event in Somerset West in March 2023, Professor Menno van Zaanen, Professor in Digital Humanities at SADiLaR, spoke of the importance of bringing people together in a mutually enriching collaborative environment that is inherently interdisciplinary, as it lies within the fields of humanities, social sciences, and computing, ‘to share and learn about digital humanities’<sup>36</sup>.



---

<sup>36</sup> SADiLaR. 2023c. <https://sadilar.org/en/dh-ignite-western-cape> Date of access 4 Aug. 2024.



## **SADiLaR's contribution towards policy support and implementation**

As language continues to be a barrier for many students at higher education institutions, in 2017 the Department of Higher Education and Training set up a revised Language Policy Framework for Public Higher Education Institutions<sup>37</sup>.

**Its goal is to develop multilingualism in the country's public higher education institutions to remove the language barriers. In September 2021 the Vice Chancellors' Colloquium on the New Language Policy Framework of Higher Education decided that SADiLaR should conduct an audit of higher education institutions to determine what resources institutions have available to them to successfully implement the new framework<sup>38</sup>.**



<sup>37</sup> Language Policy for Higher Education Act 103 of 1997 as amended. The policy framework was put forward in 2020, to be followed by the 2021 VC colloquium. The framework took effect in January 2022.

<sup>38</sup> SADiLaR. 2023a. <https://sadilar.org/en/Newsletter-april-2023> Date of access 20 Jul. 2024.



## Supporting the implementation of the revised Language Policy Framework for Public Higher Education Institutions and the SADiLaR language resources audit, 2022

The core language audit team comprised Professor Langa Khumalo (Executive Director: SADiLaR, and member of CoPAL<sup>39</sup>), Professor Tobie van Dyk (Professor of Applied Linguistics, North-West University, School of Languages), Ms Ilana Wilken (Council for Scientific and Industrial Research), Dr Natasha Ravysse (North-West University, Faculty of Law), Mr Juan Steyn (Operations Director: SADiLaR), Ms Lebogang Boemo (Project Manager: SADiLaR), and Mr Deon du Plessis (Linguist and Researcher: SADiLaR). Additional support and guidance were offered by experts in the field of language planning and management, in particular Professor Bassey Antia (University of the Western Cape), Professor Rosemary Cromarty (North-West University), Professor

Theo du Plessis (University of the Free State) and Mr Johan Blaauw (North-West University). SADiLaR began the audit with two pilot sessions in October 2022 at North-West University and the University of KwaZulu-Natal, to test the waters. The first official audit began at the University of Mpumalanga on Friday 4 November. On 22 and 28 November 2022 audits were carried out at the Vaal University of Technology and the University of South Africa. Responses from the institutions were positive, with Professor Calvin Gwandure, Head of the School of Social Sciences at the University of Mpumalanga, saying that 'this audit afforded us an opportunity to contribute to the success of the Language Policy by stating the resources we would need to ensure implementation'<sup>40</sup>.

<sup>39</sup> CoPAL is the Community of Practice for the Teaching and Learning of African Languages, formed in 2015. CoPAL 2024. <https://usaf.ac.za/communities-of-practice/community-of-practice-for-the-teaching-and-learning-of-african-languages/> Date of access 20 Jul. 2024.

<sup>40</sup> SADiLaR. 2023a. <https://sadir.org/en/Newsletter-april-2023> Date of access 20 Jul. 2024.



## The national language resources audit continues, 2023

At the beginning of January 2023 Professor Tobie van Dyk was seconded to SADiLaR, one of his tasks being to provide primary support with the National Language Resources Audit. The language resources audit was administered at 25 of the 26 public higher education institutions in South Africa.

For the audit, two questionnaires were designed, one directed at staff, and the other directed at students. Questionnaires focussed on such domains in higher education as: language services; teaching and learning (including scholarship); the language of communication and administration; and student life / co-curricular activities. The section following is taken from a summary of the findings regarding language services compiled by Professor van Dyk:



Professor Tobie van Dyk

**In terms of the availability of language services at higher education institutions, it transpired that there are indeed services available. These, however, are seemingly few and far between, and staff are not necessarily aware of their existence. Moreover, language services are typically for the exclusive use of university management, and funding is either not available or insufficient for the use of such services, whether internally or externally. A need was also expressed among student participants for accessible and effective language support services and resources, especially for non-English speakers, to address learning barriers and encourage inclusivity.**

The audit recommended 'that resources should be integrated into teaching and learning activities, rather than presented as stand-alone tools'. Staff members generally highlighted a lack of awareness regarding technical resources. The report proposed that SADiLaR, as 'a non-competing independent national Centre with an enabling mandate', might 'play a key role in providing and modelling both the expertise and technical infrastructure backbone required towards the implementation of the Language Policy Framework'. The report also notes that 'SADiLaR, with its extensive local and international networks and internal expert base, can provide guidance and support on language policy, planning and management, as well as resource development and implementation'.

As part of the South African Research Infrastructure Roadmap, SADiLaR 'is in a unique position to enable and stimulate work pertaining to the advancement of the indigenous African languages in a scientific way'. The report stresses, however, that 'successful implementation of the Policy Framework does not

only depend on a single entity, like SADiLaR. Sector-wide support and collaboration is key'. SADiLaR, though, 'can take responsibility for the long-term preservation and maintenance support of existing and future resources to ensure that investment and generated knowledge does not get lost over time'.<sup>41</sup> (Chapter 3, which follows, provides further information on the audit, along with proposed measures to deal with the challenges registered.)

Van Dyk will also work to 'strengthen existing relationships and networks and establish new ones'. Juan Steyn says that Professor van Dyk's secondment will offer valuable insights regarding the broader process of secondment to SADiLaR's research infrastructure, which will enable the Centre to 'draw on research expertise in a way that is more hands-on and with higher impact than the previous visiting scholar model'. Professor Langa Khumalo adds that Van Dyk's appointment will contribute to ensuring the 'broad national impact' of SADiLaR's work over the next five years, through 'bringing on all the skills and expertise available'.

<sup>41</sup> Van Dyk, T.J.; Wilken, I.; Ravyse, N.; Du Plessis, D.; Khumalo, L. & Steyn, J. 2023. National Language Resources Audit. Potchefstroom: SADiLaR, pp.41, 82-3.



## The SWiP initiative

**SWiP is the acronym for the SADiLaR-Wikipedia-PanSALB<sup>42</sup> collaboration, a project geared to bring together communities of language practice in a practical way to advance the use of local languages; it is hoped that communities will participate in contributing to Wikipedia, the free online encyclopaedia. Launched on 20 September 2023 at UNISA, the project aims at promoting all local languages online.**

Language communities will be brought together and provided with the skills to create and review content submitted to Wikipedia, thereby boosting the digital footprints of the various languages. The project ran from February 2024 to July 2024, offering two-day workshops on in-depth Wikipedia authorship training at ten universities: the University of Venda; the University of Mpumalanga (Siyabuswa Campus); the Vaal University of Technology; the Tshwane University of Technology; the University of South Africa; the Cape Peninsula University of Technology; the University of Zululand; the Central University of Technology; Sol Plaatje University; and Walter Sisulu University. The 20 participants per university will then be in a position to run their own Wikipedia training events.<sup>43</sup>

<sup>42</sup> PanSALB is the Pan South African Language Board, established to foster the development and use of official local languages, and to ensure respect for all languages commonly used in South Africa, including Arabic, German, Greek, Gujarati, Hebrew, Hindi, Portuguese, Sanskrit, Tamil, Telegu and Urdu. See <https://www.pansalb.org>

<sup>43</sup> SADiLaR. 2024a. <https://escalator.sadilar.org/champions/engager/July 2024> Date of access 01 Aug. 2024.

As a consequence of the SWiP initiative, isiNdebele has already featured prominently on Wikipedia, with 128 articles appearing in under a year. Professor Langa Khumalo is confident that ‘the next phase of the project will yield even better results.’ He added that ‘SWiP provides our communities with a platform to grow our languages in the digital space’.

Apart from this major achievement in isiNdebele, at the time of writing, participants in different languages have written 638 new articles and edited 1080 existing ones. This has resulted in the addition of more than 290,000 words, 1820 references, and 122 photos appearing in Wikipedia. According to the SWiP Wikimedia dashboard report, there have been 22.4 million article views, reflecting the substantial global impact the project is making regarding the appearance of African languages and knowledge.



Images taken at the official launch of the SWiP project in September 2023

The second phase of the project begins in November 2024, and will end in October 2027. It will include institutional workshops to enhance knowledge of how Wikipedia can benefit education, the training of facilitators across the country, and ‘Edit-a-thons’ designed to encourage language community contributions to Wikipedia and to encourage more collaboration.



Julius Dantile, Executive Head of Languages at PanSALB, feels that the successful completion of the first phase of the project makes one confident ‘that the data will show linguistic diversity in the digital world’. He sees the collaboration as providing ‘a significant step towards instilling a culture of writing in indigenous African language on digital platforms’. Further, ‘the outcome of the project may increase corpora which will assist the use of artificial intelligence through indigenous African languages’<sup>44</sup>. At the recent SWiP principals meeting the President of Wikimedia South Africa,

Mr Bobby Shabangu announced that ‘It is with great pleasure that Wikimedia South Africa received notice from Wikipedia International that isiNdebele through the work of the SWiP project is going to be moved out of incubation and is now a formal language in Wikipedia’. There will thus be a formal launch of isiNdebele, led by PanSALB, as the latest language to achieve digital ascent. This is a testament to the importance of facilitating epistemological access to new forms of knowledge and skills. This achievement would not have been possible without many collaborators.



Bobby Shabangu  
Photo credits: Neo Ntsoma, Wikimedia  
Foundation contractor CC BY-SA 3.0



**The SWiP project attests to collaborations across time with first collaborations taking place at the 2019 DHASA conference then again the 2021 conference and later DH-IGNITE events which led to further collaboration ideas and finally the SWiP project.**

<sup>44</sup> SADiLaR. 2024b. <https://sadilar.org/en/isindebele-enjoys-major-boost-on-wikipedia-thanks-to-swip/> Date of access 1 Aug. 2024.

## Conclusion to Chapter 2

---

SADiLaR continues to play an enabling role to transform research and teaching in the Humanities and Social Sciences through a high-impact digitisation and digital humanities programme. It organically supports and drives policy implementation and creation. SADiLaR remains, in 2024, the only entity of its kind in Africa.

As a multi-partner entity, the infrastructure will continue to enable the development, research and support of all the official languages of South Africa in pursuit of sustainable multilingualism. Through its digitisation programme, which entails the systematic creation and maintenance of relevant digital text, speech and multi-modal resources related to all official languages of South Africa, SADiLaR plays a direct role in helping to ensure a digital future for South African Languages. The development of appropriate natural language processing software tools for research and development purposes is included as part of the digitisation programme, in

addition to a digital humanities programme which facilitates the building of research capacity by promoting and supporting the use of digital data and innovative methodological approaches in the Humanities and Social Sciences. As a generator of emerging programmes and functionalities, SADiLaR, through its established technical infrastructure and digital language resource management experience, continues to play an important enabling role towards the implementation of the New Language Policy Framework for Higher Education institutions, a matter that will be more fully considered in the following, final chapter.

**SADiLaR as an enabler continues to support initiatives and aggregates African related language technologies and digital language resources that form the technological building blocks that continue to enable collaborations and innovations across time.**



# THE *FUTURE*





## CHAPTER 3

*The future of SADiLaR: Strategy*



## **SIX KEY AREAS**

*Six Key Areas: Research; Technology and Resources; Projects and Services; Good Governance; Communication; and Team*



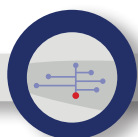
**The Centre's strategic objectives over the next five years are keyed to promote easily accessible, human friendly language-related technologies, digital humanities for transformational research, and language policy implementation in the realm of higher education. The following six strategic objects will be served:**



## 1. RESEARCH

To stimulate and advance the scholarship of human language technologies and digital humanities in South Africa and across the African continent, and thus further strengthen the Global South knowledge production and dissemination pathways, thereby contributing to global knowledge production. This would involve seminar series, workshops and conferences to cultivate an active digital humanities environment. Training in relevant practices and methodologies and the promotion and creation of collaborative

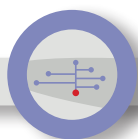
Open Science initiatives will stimulate human capacity development. National and international integration might be achieved through the establishment of an African-wide research infrastructure consortium. Additional means of collaboration might be advanced through hub-based and hub-driven projects to promote local languages. SADiLaR will also continue to provide strategic digital humanities support and expertise.



## 2. TECHNOLOGY AND RESOURCES

To enhance the development, deployment and maintenance of software and technologies in the domains of digital language resources and digital humanities by strengthening the technical infrastructure. A culture of quality, sustainability and maintainability will be promoted through the creation of a framework to guide the necessary processes relevant to all three. This would involve long-term maintenance planning, the maintenance of existing software packages, systems and core supportive technologies regarding human language technologies and computational methodology uptake in the human and social sciences, along with relevant colloquia.

Also essential are responsive technology development for terminology management and dissemination, and resource and tool development, the growing of African human language technology expertise, and actively supporting the maintenance and collection of language resources.



### 3. PROJECTS AND SERVICES

To sustain and enable a broad range of multiphase projects in SADiLaR's mandate by increasing internal and external productive interaction. This will be achieved by, for instance: running specialisation projects with a more prominent SADiLaR profile that unifies the hub and its nodes; strengthening South African sign language resource development; continuing to provide consultation

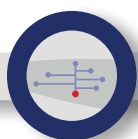
and guidance regarding the development of language policies and plans for furthering multilingual and multimodal teaching and learning frameworks; enhancing the visibility of SADiLaR by broadening its community engagement locally and internationally; and creating human capacity development project opportunities that include historically disadvantaged institutions.



### 4. GOOD GOVERNANCE

To foster good governance and work towards project sustainability and landmarking of the infrastructure. Good governance will be maintained by the promotion of strong and transparent governance, involving reviewing the governance charter and best practices; active succession planning and the articulation of growth opportunities for all staff; ensuring financial sustainability; ensuring high quality outputs; and strengthening relationships with higher education stakeholders.

Of importance is the initiation of exchange programmes locally and abroad. Also important is the development of communities of practice, where community members become responsible members of the ESCALATOR programme. Sustainability plans should consider the need to be less reliant on sole government funding. Partners might be found in the private and public sectors locally and abroad.

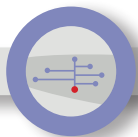


## 5. COMMUNICATION

To foster excellent communication, SADiLaR brand management and stakeholder engagement. Growing the SADiLaR brand involves making the Department of Science and Innovation mandate clear internally and externally, curating expectations and communication in the light of corporate identity, fostering activities and packaging that clearly identifies SADiLaR as both locally relevant and effective, and internationally competitive.

Stakeholder engagement should be mutually beneficial, turning on, for example the following: alignment with Department of Science and Innovation demographic transformation areas, such as the researcher base and higher levels of research management; the institutional transformation of previously marginalised ones based on changing needs in the National System of Innovation; and increasing the footprint of Science, Technology and Innovation activities to ensure greater social and economic benefits to priority beneficiary groups. Clear communication involves the definition of clear lines of communication between internal and external stakeholders to relay key messages, using appropriate communication platforms. An ethics of belonging is also important for SADiLaR, to embody its commitment to diversity, inclusivity and social cohesion, both in the case of the team and stakeholders.





## 6. TEAM

Team—to promote and nurture staff development and well-being through recognition and progression. Growth orientated performance planning is important. This involves quarterly task planning linked to performance agreement and appraisal cycles that are run sensitively and with mutual respect, and evident succession planning, deputising and project involvement.

Wellbeing in the workplace will be fostered through social platforms catering for emotional, physical and social wellbeing, tied to the ethic of belonging. Perhaps most important are recognition systems, with annual acknowledgment of achievements, and the generation of a sense of growth and acceptance of responsibility linked to career pathways<sup>45</sup>.

As part of the five-year evaluation of the South African Research Infrastructure Roadmap, the Research Infrastructure completed two rounds of external evaluations, where specific recommendations were tabled towards ensuring the sustainable long-term impact of SADiLaR. These included strengthening its identity and brand regarding digitisation and digital humanities through active human capacity development in communities to create further enriching diversity and inclusion.

For the overall sustainability of the Research Infrastructure, the areas in need of attention were deemed to be governance, operations, staff development and finance. Building on these recommendations, SADiLaR's new strategy was developed through broad consultation and a careful consideration of recent developments pertaining to its enabling role in implementing the New Language Policy Framework for higher education regarding language policy planning, development and support.

<sup>45</sup> SADiLaR. 2024c. Strategic Plan 2024-2029. Potchefstroom: SADiLaR. Information provided courtesy of Professor Tobie van Dyk.



# NATIONAL EDUCATION

*support  
programme*

National  
education:  
SADiLaR  
sector  
support  
programme



Very important for SADiLaR is its role in national education. SADiLaR has created a higher education support programme designed to support the higher education sector with capacity development, project execution, and resource and infrastructure development. This initiative is key in building and sustaining multilingual teaching and learning capacity across universities.

**Eight key outcomes from the language resources audit have been identified as informative bases to help deal with the many challenges and opportunities in the higher education sector:**





SADiLaR, through its higher education support programme can, in varying degrees, provide support to universities in all of these areas. In doing so, it will contribute to reshaping the linguistic landscape of higher education by promoting the importance of effective management of linguistic diversity, equitable access to educational resources, and by helping to enhance the overall learning experience for students from diverse linguistic backgrounds.

**In the light of the above eight outcomes, then, SADiLaR proposes five projects which will impact on national higher education:**

## **PROJECT 1**

Project 1 is titled ***‘Expertise and support in language planning and management’***, and will, through a series of three Workshops, Workshops 1, 2 and 3, offered between 2024 and 2026, provide a comprehensive introduction to the principles and practises involved in designing, bringing into being and evaluating language policies and implementation plans.

Practical strategies for integrating language policy into institutional frameworks, capacity development, identifying existing resources and fostering institutional collaboration will also be dealt with, as will assessment literacy and means of enhancing translation practises of assessment tools.





## PROJECT 2

Project 2: ***'Expertise and support in language policy and plan monitoring and reporting'***. A fourth workshop will be introduced at regional level in 2025 and 2026, covering language policy and plan monitoring and reporting, dealing with key methods for tracking implementation, evaluating effectiveness and ensuring alignment with designated goals.

## PROJECT 3

***'Expertise and support in multilingual pedagogies training'***. Workshop 5 in the series will be presented regionally between 2025 and 2026, and will offer strategies and tools for teaching in multilingual classrooms, in order to promote inclusive environments. Best practices will be emphasised, along with key theories and models in multilingual pedagogy and means of valuing cultural and linguistic diversity. It will, ultimately, offer action plans for the responsible introduction of multilingual pedagogy in class.



## PROJECT 4

### ***'Academic language development'.***

Multilingual teaching and learning materials have been developed by SADiLaR, which could either be used along with standard academic language development courses, or as standalone materials for self-study and use. Academic writing, academic reading and the link between them will be covered, as well as materials geared towards imparting better academic writing and reading skills. This project will also design relevant pedagogical frameworks for adaptation by institutions according to their needs and approaches.

## PROJECT 5

***'Terminology development'.*** This project is geared towards imparting multilingual terminology development for effective communication across diverse linguistic contexts. SADiLaR already has available data on terminologies, glossaries, wordlists and dictionaries; the Centre aggregates data and has developed a sustainable open source platform for use by institutions according to their requirements. Further data collection and platform development will take place between 2024 and 2026.<sup>46</sup>

<sup>46</sup> Van Dyk, T.J. 2024. Post Language Audit Matters: Higher Education Support Programme. SADiLaR, Potchefstroom. (Unpublished proposal).



## INTERNATIONAL COLLABORATION

While in executing its six strategic objectives SADiLaR aims to be locally relevant and effective, playing an influential role regionally, it would also extend important collaborations and partnerships abroad. Stakeholders have been engaged with, formally and informally, through conferences and other events.

Further collaborations and engagement are foreseen with government departments (such as Departments of Basic Education and Higher Education and Training), research centres (such as Indigenous Knowledge Systems), and institutes and language entities (such as the Pan South African Language Board); this process will continue. Broader stakeholders would include civil societies such as Greenpeace and the Solomon Projects Language Academy, Google, Apple, Facebook, Amazon, Microsoft, the Centre for Digital Humanities at the University of Lagos, the GLAM sector (that is, the sector involving galleries, libraries, archives and museums), financial institutions with large amounts of data, relevant international organisations and collections holding African materials, and researchers in countries similar to South Africa as far as multilingualism goes, such as India and various South American countries.

## *General conclusion*

SADiLaR's enabling function, in improving academic skills and creating new knowledge, contributes indirectly to the country's economy by helping to produce people with high level job skills across various economic sectors. Digital language resources in African languages in particular have a commercial potential regarding (for example) the development of language-based applications for health care, education, job training, e-governance, and entertainment.

More fundamentally, as languages are so closely linked to linguistic communities with their own expressions and means of acquiring knowledge, individuals from these communities interacting with researchers will contribute to a store of unique data, helping to provide a holistic perspective onto the various languages of South Africa, while at the same time helping to preserve what is unique to certain groups. Together we can ensure a digital future for our languages.





SADiLaR's  
TEAM

---



## MANAGEMENT

### **Prof Langa Khumalo** *Chief Director*

Langa Khumalo is a full professor and executive director of the South African Centre for Digital Language Resources (SADiLaR) at North-West University. He is the former director of Language Planning and Development at the University of KwaZulu-Natal (UKZN). He has been an invited speaker and keynote speaker at various conferences and colloquia. He is also a board member of the African Literary Metadata project at Uppsala University (Sweden) and is a member of the Advisory Panel for Language Policy Implementation appointed by the Department of Higher Education and Training (DHET). He is an award-winning author in Intelligent Text Processing and Computational Linguistics and is a recipient of the Staff Excellence Award (2018) for Excellence in Teaching at UKZN. He has received a recognition award from the University of the Free State for his global contribution to multilingualism. He is a fellow of Zimbabwe Academy of Sciences, and an adjunct professor at UKZN.

### **Mr Juan Steyn** *Director: Operations*

Juan Steyn is the director of operations at SADiLaR. He has been involved in multiple digital humanities and educational technology-related projects. He also has a special interest in training and capacity-building through his involvement with the software and data carpentry community and the Digital Humanities Association of Southern Africa.

## MANAGEMENT

### **Prof Menno van Zaanen** *Professor in Digital Humanities*

Professor Menno van Zaanen is a professor in digital humanities and the research manager at SADiLaR. His background is in computer science, computational linguistics, and digital humanities. Menno is particularly interested in incorporating the use of computational techniques into the field of humanities. His research mostly focuses on identifying patterns in symbolic sequences, such as natural language, music, or other modalities related to human cognition, but he is also interested in computational approaches to situations where only limited resources are available, such as low-resource languages.

### **Dr Friedel Wolff** *Technical Manager*

Dr Friedel Wolff is a computer scientist specialising in language technology. As the technical manager at SADiLaR, he is responsible for technical infrastructure and technical standards. Having worked in industry, the NGO sector and academia, his focus has always been on the proper application of information technology, and the sustainability of teams and projects. As a researcher, his core interests are in language resources, translation support, and techniques for addressing resource scarceness.

### **Ms Lihle Sosibo** *Communication Manager*

Lihle has served on several communication portfolios in the public, research, and higher education sectors. She has gained vast competencies in developing and implementing corporate communication and public relations strategies. She is passionate about organisational culture, brand management, events management, content management, media relations, and stakeholder engagement.





## SECONDMENT APPOINTMENTS

### **Prof Tobie van Dyk** *Higher Education Support Programme*

Professor Tobie van Dyk is professor of applied linguistics in the School of Languages at North-West University. He holds a C2 NRF rating and is acknowledged locally and internationally for his work on language for academic purposes. He serves on several professional bodies, both in South Africa and abroad. He is currently tasked with the responsibility of creating the required infrastructure for an emerging programme in SADiLaR, namely the higher education sector support programme.

This is after successfully completing in 2023 the national language resources audit at public universities in South Africa, under the auspices of USAf and the DHET. Formerly, Prof Van Dyk was head of the Centre for Academic and Professional Language Practice and director of the School of Languages, both at NWU.

### **Prof Justus Roux** *Research and Development Projects*

Professor Justus Roux is the founder member and first director of SADiLaR. He is South Africa's official delegate to the CLARIN General Assembly, where SADiLaR has been accepted as the only non-European representative to the EU consortium. His work within SADiLaR involves assessing SADiLaR and CLARIN network research and development projects, potentially leading to wider cooperation with South African academics working in the same fields. He is an emeritus professor in the Department of African Languages at Stellenbosch University.



## RESEARCHERS

### **Ms Andiswa Bukula**

***Digital Humanities Researcher:  
IsiXhosa***

Andiswa Bukula is a digital humanities researcher specialising in the preservation and development of South African indigenous languages, with a particular focus on IsiXhosa. Andiswa contributes significantly to the academic discourse surrounding the intersection of language and technology. Her current research interests are focused on digitisation techniques for the purposes of language preservation.

### **Ms Nomsa Skosana**

***Digital Humanities Researcher:  
IsiNdebele***

Nomsa's research interests include translation studies, terminology development, and the preservation of South African indigenous languages, with a focus on isiNdebele. Given that isiNdebele is an under-resourced language, she is dedicated to advancing its development through research and resource creation. Nomsa is especially interested in developing corpora and digital tools to support and enhance isiNdebele.

### **Ms Rootweither Mabuya**

***Digital Humanities Researcher:  
IsiZulu***

Rooweither Mabuya is a language researcher at SADiLaR. Her research focuses on systematically creating digital text, speech, and multimodal resources to support the development of isiZulu. She is also dedicated to promoting the adoption of digital humanities methods and tools within the isiZulu research community. Her goal is to contribute to the advancement of isiZulu as a language used for administration, education, research, and innovation.

### **Mr Deon du Plessis**

***Digital Humanities Researcher:  
English***

Deon serves as the SADiLaR researcher for South African English. His expertise is in sociophonetics, world Englishes, and Afrikaans phonetics. His research explores how South Africans express their identities through South African English, the language's evolution in a postcolonial context, and its role in building interpersonal connections. Deon aims to enhance access to shared knowledge through English as a common language, to advocate for it as one thread in our multilingual tapestry, and to advance South African research capabilities with new technologies.

**Mrs Mmasibidi Setaka-Bapela**  
***Digital Humanities Researcher:***  
***Sesotho***

Mmasibidi Setaka-Bapela is a digital humanities researcher specialising in Sesotho with expertise in lexicography, translation, and editing. She holds a BA in Media and African Languages (Wits), a BA(Hons) in Applied Language Studies (UP), and an MA in African Languages (UP). She serves on various committees, including the Digital Humanities Association of Southern Africa, the Sesotho National Language Body of the Pan South African Language Board, and the University of Free State's Academic Advisory Board (Department of African Languages). She is a published author and has organised numerous conferences and workshops locally and internationally.

**Dr Respect Mlambo**  
***Digital Humanities Researcher:***  
***Xitsonga***

Respect Mlambo is a Xitsonga language researcher at SADiLaR. His research interests include translation, terminology development and lexicography, as well as sociolinguistics, morphology, semantics, and syntax. He is also engaged in research and initiatives aimed at utilising language technologies and resources to promote indigenous languages.

**Dr Benito Trollip**  
***Digital Humanities Researcher:***  
***Afrikaans***

Dr Benito Trollip is a curious and enthusiastic person with a passion for people. In the research space, he is especially interested in the ways people construct meaning through language. He is also interested in other aspects of research, including intellectual property rights and the ownership, management, and distribution of data. As a SADiLaR researcher, Benito is always discovering new ways to bring language and the digital age together. In his PhD, which he completed in 2022, his focus was on the ways in which Afrikaans speakers use different means to build judgement into their language.

**Dr Muzi Matfunjwa**  
***Senior Lecturer***

Muzi Matfunjwa is a senior lecturer at SADiLaR. His areas of interest are Siswati linguistics and digital humanities. He is interested in using digital language resources to facilitate and stimulate research in low-resource languages in South Africa, especially Siswati. He also focuses on language use in society and the sociocultural nuances that language manifests. He contributes to the digital preservation of Siswati and employs digital tools to conduct research in this language.

## SUPPORT STAFF

### **Ms Jessica Mabaso** ***Chief Administrative Officer:*** ***Operations***

Jessica Mabaso brings over 15 years of experience in business support and administration to her role at SADiLaR. With a strong background in office management, finance, and business development, Jessica excels at developing and maintaining comprehensive administrative processes that improve operational efficiency. Prior to her current role, Jessica served as the chief administrative officer and financial administrator at SADiLaR, where she managed stakeholder relationships, executive support, and financial operations. Her career also includes experience as a senior administrative officer at North-West University and in various administrative and analytical roles at ArcelorMittal South Africa. Jessica holds a diploma in Human Resources Management, a certificate in Public Relations, and is a PRINCE2 certified project management practitioner. She is currently pursuing a postgraduate diploma in Business Management. Known for her excellent communication skills, attention to detail, and ability to build strong relationships across cultural boundaries, Jessica is dedicated to supporting SADiLaR's mission through effective programme management and community engagement.

### **Mrs Lebogang Boemo** ***Project Manager***

Lebogang Boemo is the project manager at SADiLaR, with a background in working with the National Skills Fund (NSF), which supports skills development, education, and training initiatives, including bursaries, scholarships, learnerships, and work-based learning programmes. Projects she has managed include the DHET Language Resource Audit at all 26 higher education institutions. She is currently managing the SADiLaR-Wikipedia-PanSALB project, open call projects and node projects.

### **Mr Marius Coqui** ***Section Head: Project Finances***

Marius Coqui is the section head of Project Finances at SADiLaR. He holds a BA (Hons) in Applied Linguistics and has experience in research project finance. Previously, he was project coordinator at the AIDS and Society Research Unit, and operations manager at the Centre for Social Science Research, at the University of Cape Town.

**Mr Makgobalo Cascious  
Mofokeng**  
*Senior Project Administrator*

Cascious Mofokeng, a PRINCE2 project management practitioner at SADiLaR, excels in planning and problem-solving, inspired by the rich history of African languages. He co-led the Department of Higher Education Language Resource Audit, involving all 26 public universities in South Africa, and contributed to the COVID-19 multilingual terminology list, verified by PanSALB. He also played a significant role in celebrating South Africa's 11 official languages during UNESCO's International Year of Indigenous Languages in 2019, organising collaborative events to promote language awareness and development across universities.

**Ms Phelo Masala**  
*Administrative Officer: Finance*

Ms Masala is a strategically minded financial administrative assistant with six years of experience in bookkeeping, financial reporting and auditing.

**Ms Tshego Rapoo**  
*Communication Officer*

Ms Tshego Rapoo-Maloyi is a media and communications all-rounder with a diverse background in print and broadcast media, social media management, public relations, and corporate communications in the public sector. She is committed to contributing to the success of SADiLaR by providing a comprehensive communication service aligned with its strategies and objectives, and to facilitate two-way communication between the organisation and its stakeholders.

**Ms Nomkhuleko Ngubane**  
*Chief Administrative Officer:  
Executive Office*

Ms Nomkhuleko Ngubane is the chief admin officer in the office of the executive director at SADiLaR. She has experience and interest in strategic administration and project management of language units in South Africa. Her approach is not only to provide strategic administration support but also to tailor it specifically for language units. She has an interest in language and language development within the broad context of higher education and national development. She is currently pursuing her MA in Linguistics with a focus on the language policies of the public higher education Institutions in South Africa.



**Mr Jan-Rudolph Bührmann**  
*Senior Software Developer*

Mr Jan-Rudolph Bührmann is the senior software developer at SADiLaR. His interest in software development grew from an assignment during his BSc (Hons) in Exploration Geophysics. Three years at the Council for Geosciences were followed by a move into software development. Career highlights include implementing deployment pipelines under Agile at Microsoft's UK offices long before the acceptance of such practices, working remotely for Belgian companies, being the onsite representative for an offsite team, being scrum master and having articles accepted onto the front page of DZone. His interest in languages started when he gained understanding of how Afrikaans derived from Dutch while working in the Netherlands.





# SOUTH AFRICAN DIGITAL LEARNING RESOURCES



**SADILAR**  
South African Centre for Digital Learning