SADiLaR

South African Centre for Digital Language Resources
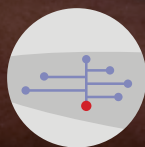
Ensuring a *digital future* for all official languages in South Africa

STRATEGIC PLAN 2024–2029

TABLE OF CONTENTS

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

SADiLaR
South African Centre for Digital Language Resources

1. EXECUTIVE SUMMARY

The South African Centre for Digital Language Resources (SADiLaR)[1] is a national research infrastructure (RI), that is part of the South African Research Infrastructure Roadmap (SARIR). Its strategic function is to create, manage, and distribute digital language resources (DLR)[2] and applicable software for all official languages in South Africa through its digitisation programme, while through its digital humanities (DH) programme, SADiLaR stimulates and supports computational research and development in the humanities and social sciences (HSS).

SADiLaR also contributes as an enabler of activities related to language policy planning, development, and support in the South African context and especially towards the implementation of the new Language Policy Framework for Public Higher Education Institutions.

The infrastructure 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.

As a human-centred RI, SADiLaR understands the importance of the human capacity required in aiming for a transformed digital future in a science-literate and science-aware society. Through its activities and hosted DLR and services, the RI contributes to unlock access to knowledge in all official South African languages.

---

[1] SADiLaR is an RI with a hub-and-spoke structure, with the hub residing as a hosted entity at NWU and six nodes based at different universities and institutions. "SADiLaR" and "RI" both refer to the hub and the nodes.
[2] Digital language resources (DLR) are digital language datasets that can range from plain text corpora to multimodal collections to the digital tools and applications required to parse or to analyse such datasets or new (raw) data. It also includes the building blocks towards end-user products such as developed teaching and learning materials and platforms to facilitate access such material that enables functional multilingualism.

During the next five years, the infrastructure will focus on the following six strategic objectives to enhance the role played by easily accessible language-related technologies, DH for transformational research, and supporting language policy implementation in realising an inclusive and transformed digital future for South Africa.

RESEARCH

Stimulate and advance the scholarship of human language technologies (HLT) and DH in South Africa and across the African continent, and further strengthen Global South knowledge production and dissemination pathways, thereby contributing to global knowledge production.

TECHNOLOGY AND RESOURCES

Enhance the development, deployment, and maintenance of software and technologies in the domains of DLR and DH by strengthening the technical infrastructure.

PROJECTS AND SERVICES

Sustain and enable a broad range of multiphase projects and services in SADiLaR's mandate by increasing internal and external synergies.

GOOD GOVERNANCE

Foster good governance and work towards project sustainability and landmarking of the infrastructure.

COMMUNICATION

Nurture excellent communication, SADiLaR brand management and stakeholder engagement.

TEAM

Promote and nurture staff development and wellbeing through recognition and progression.

2. ABOUT SADILAR

2.1   Background: SADiLaR's new strategy

SADiLaR was established as an RI during 2017 as part of SARIR, a programme funded by the Department of Science and Innovation (DSI).

The RI, hosted by the North-West University (NWU), is made up of multiple nodes with its hub[3] located at the NWU. Through its initial digitisation and DH[4] programmes, various projects were supported within the domains of speech and text processing, digitisation[5], language testing and teaching, child language development and other language development activities.

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.

The RI was setup to fulfill a strategic function which is to create, manage and distribute DLR and applicable software for all official languages in South Africa through its digitisation programme as well as to stimulate and support computational research and development in the HSS through its DH programme.

---

[3] SADiLaR's "hub" refers to the core team that is employed full-time and their programme of flagship activities that are hosted by the research entity at the NWU. This excludes the nodes.
[4] Digital humanities (DH) refers to the paradigm shift in the field of HSS due to advances in the domain of Information Communication Technologies (ICT), and access to digital resources. There are different views of DH, each with its own definition. In general as described by the Department of Digital Humanities at King's College London – "digital humanities form 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". The effect of this is that "over the last decade, humanistic research and teaching has changed. Equipped with digital technology, scholars are rethinking research, publication, and pedagogy". Because language is an important carrier of information and knowledge, it lies at the core of developments in the DH domain, especially when available in a digitised format. In addition alternative media (such as visual and audio) are also investigated in the field of DH.
[5] Digitisation relates to the conversion of materials in analogue or physical format to digital formats. Newly digitised data can be enriched using DLR created and maintained by SADiLaR.

As part of the five-year evaluation of SARIR infrastructures, the RI successfully completed two rounds of external evaluations where specific recommendations were tabled towards ensuring the sustained long-term impact of SADiLaR. These included: (i) Strengthening SADiLaR's identity and brand regarding digitisation and DH through active human capacity development and engagement with communities to create further diversity and inclusion. (ii) As part of SADiLaR's research activities, Global South knowledge production and dissemination should be strengthened and human capacity development activities should continue in an accessible manner. (iii) For overall sustainability of the RI, matters concerning governance, operations, staff development and financial sustainability must be strengthened further.

Building on these recommendations, SADiLaR's new strategy was developed through broad consultation and consideration of recent developments regarding the RI's enabling role towards e.g. implementing the new Language Policy Framework for Public Higher Education Institutions concerning language policy planning, development and support.

2.2 Strategy

SADiLaR as a human-centred RI understands the importance of the human capacity needed for a transformed digital future in which society is aware of and understands the value of science. Through its activities and hosted DLR and services, the RI contributes towards unlocking access to knowledge in all official South African languages.

SADiLaR, as of 2024, remains the only RI 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[6]. SADiLaR's core mandate[7] is to drive a 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. 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 DH programme which facilitates the building of research capacity by promoting and supporting the use of digital data and innovative methodological approaches in the HSS.

As an emerging programme, SADiLaR, through its established technical infrastructure and DLR management experience, continues to play an important enabling role towards the implementation of the new Language Policy Framework for Public Higher Education Institutions. It is therefore envisioned to develop organically to play an enabling role in matters of language policy planning, development and support.

Our strategy articulates a vision that stimulates and enables the valorisation of DLR development and maintenance, DH research, computational human capacity development in the HSS, organic language-policy-related enablement and the use of the affordances of artificial intelligence in these contexts. SADiLaR's strategy comprises six strategic objectives which actively leverage from the successes achieved since the establishment of the RI. In executing its strategy SADiLaR aims to be locally relevant and effective, playing an influential role regionally and abroad through strategic collaboration and partnerships. The RI will contribute towards a digital future for indigenous and African languages by strengthening and expanding our regional African and international infrastructure networks.

---

[6]Multilingualism can be defined as "the ability of societies, institutions, groups and individuals to engage, on a regular basis, with more than one language in their day-to-day lives" - European Parliament. SADiLaR enables and advances multilingualism through its programmes and activities by creating, maintaining and refining DLR that promote the ability of groups to engage with materials in a multilingual manner. [7]SADiLaR has an enabling function, with a primary focus on the creation, management and long-term distribution of DLR in all official languages of South Africa through its Language Resource Catalogue. It enables research and development in the domains of language technologies and language-related studies in the (digital) HSS.
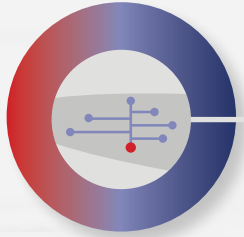
3. VISION

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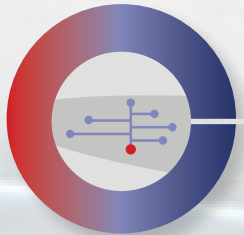
4. MISSION

1. We enable the use of our official South African languages in research and development activities through active DLR 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 DH communities of practice as an infrastructure.
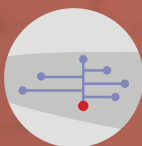
5. VALUES

We espouse our RICH Plus T (RICH^T) principles and foster belonging.

R

SADiLaR undertakes to promote mutual **respect**, courtesy and inclusivity in an environment driven by **research** and productivity.

I

SADiLaR promises to **inspire innovation** that is effective and contributes meaningfully to the community.

C

SADiLaR commits to **carefully and effectively communicate** with all its strategic partners and stakeholders.

H

SADiLaR undertakes to transform research and teaching in the **HSS** through a **high-impact** DH programme.

T

SADiLaR undertakes to drive its mandate with determination, with commitment to **teamwork and trust** between staff, management, strategic partners and stakeholders.

6. STAKEHOLDERS

For SADiLaR to successfully execute its strategy, the RI must ensure that strategic relationships are strengthened locally, regionally and abroad.

Progress has been made with stakeholder engagement directly and informally as part of conferences, events and collaborations. Synergies are actively sought to maximise impact. SADiLaR will continue to actively liaise with government departments, specific research centres, institutes and language entities. Engagement with relevant public sector entities and civic societies will be further strengthened.

6.1 Stakeholders linked more closely with academic institutions

- African language and other humanities schools, or departments, or centres of excellence

- Computer science schools, or departments

- University publishers

- University executives and councils

- Public and private university language units, or centres

- Research centres or institutes (e.g. Indigenous Knowledge Systems Centres)

- University technology, or innovation units, or centres

- University libraries museums and archives

- Corporate relations, or public relations divisions

- Students, especially mother-tongue enrolment

- SADiLaR's nodes

- SADiLaR governance committee members

- Departments of Basic Education (DBE); Higher Education and Training (DHET); Sports Arts and Culture (DSAC), Pan South African Language Board (PanSALB), and Universities South Africa (USAf)

6.2 Broader stakeholders within Africa and beyond

- Government departments and related national infrastructures

- Regional communities (e.g. African Academy of Languages [ACALAN]; South African Development Community [SADC])

- Civil societies (e.g. Greenpeace, Solomon Projects Language Academy)

- Google, Apple, Facebook, Amazon, Microsoft, the Centre for Digital Humanities University of Lagos (CEDHUL) and other global DH centres

- Financial institutions and other private sector institutions, which could include corporations and organisations with large amounts of text/data (e.g. Saigen / insurance companies)

- Other RIs in SARIR and the European Common Language Resources and Technology Infrastructure (CLARIN) and its member countries. and its member countries

- **The people of South Africa**

- GLAM sector (Galleries, Libraries, Archives, and Museums)

- Language and culture entities not affiliated with a centre or university (e.g. the Zuid-Afrikahuis in Amsterdam) and other relevant international organisations / collections holding African materials

- Researchers in countries similar to South Africa, for instance India, with many languages, many of them under-resourced. Several South American, and also European countries also fit this description

- Academic associations: Southern African Linguistics and Applied Linguistics Society (SALALS), African Language Association of Southern Africa (ALASA), African Association for Lexicography (AFRILEX), South African Association for Language Teaching (SAALT), Network for Digital Humanities in Africa, Digital Humanities Association of Southern Africa (DHASA), Alliance of Digital Humanities Organizations (ADHO)
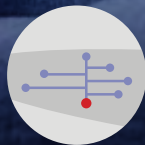
7. FUNDING MATTERS

The RI is fully funded by the DSI and strives to partner with entities, individuals and institutions who share SADiLaR's mission and vision in order to best leverage limited funding. During the next five years, selected external funding applications will be submitted which align with the overall strategy. Commercialisation is foreseen to develop mainly in downstream activities within our stakeholder network.

8. KEY STRATEGIC OBJECTIVES

8.1 STRATEGIC OBJECTIVE 1: RESEARCH

Stimulate and advance the scholarship of HLT and DH in South Africa and across the African continent, and further strengthen Global South knowledge production and dissemination pathways.

KEY EMPHASIS: **Contributing to global knowledge production.**

STRATEGIC RESEARCH AIMS

8.1.1 Cultivate an active DH community of practice

Activities include

DH seminar series, the ESCALATOR[8] programme, Resources for African Indigenous Languages (RAIL) workshops, DHASA, and ADHO conferences.

8.1.2 Stimulate human capacity development

Activities include

Training regarding DLR and DH practices and methodologies, promotion and creation of Open Science initiatives in collaboration with the South African Humanities Deans Association (SAHUDA), National Institute for the Humanities and Social Sciences (NIHSS), Wikimedia, universities, and the African Open Science Platform (AOSP).

8.1.3 Explore the establishment of an African RI consortium

Activities include

International integration through exploring the establishment of a southern African and an African RI consortium.

8.1.4 Create additional ways to collaborate and utilise the RI

Activities include

Establishing a hub-based virtual laboratory specialisation programme and hub-driven specialised projects that advance development of the different official languages.

8.1.5 Provide strategic DH support and expertise

Activities include

Establishment of DH centres, programmes, and curriculum renewal across public higher education institutions.

---

[8]SADiLaR established the ESCALATOR programme which aims to grow an inclusive and active South African community of practice in DH and computational social sciences by establishing new and connecting existing relationships through mutually beneficial partnerships. ESCALATOR consists of three main activities: 1) DH-IGNITE – a regional event organised within the ESCALATOR programme aiming to support the development of an active community of practice in DH and computational social sciences in South Africa. 2) The ESCALATOR Digital Champions Initiative – a multi-track mentorship and networking programme that complements our other activities. 3) Stakeholder map project – aims to collect and share data on DH and computational social sciences researchers and related activities and initiatives in South Africa.

8.2 STRATEGIC OBJECTIVE 2: **TECHNOLOGY AND RESOURCES**

Enhance the development, deployment, and maintenance of software and technologies in the domains of DLR and DH.
KEY EMPHASIS: Strengthening the technical infrastructure.

STRATEGIC TECHNOLOGY AND RESOURCES AIMS

8.2.1 Promote a culture of quality, sustainability and maintainability

Activities include

Creating a framework to guide high quality development and methods to approach sustainability and longer-term maintenance planning, maintaining existing software packages, systems, and core technologies supporting HLT and computational methodology uptake in the HSS. Continue technology development colloquium series.

8.2.2 Responsive technology development

Activities include

Response to development needs such as a centralised system for terminology management and dissemination, platforms for resource development, dissemination, stimulating the development of tools to support multilingualism, and piloting subtitling services to support multilingual language policies across academia.

8.2.3 Grow African HLT expertise

Activities include

Retaining expertise in the domain of HLT and expanding it across the country to deepen national integration.

8.2.4 Proactive resource maintenance and aggregation

Activities include

Actively supporting the maintenance and aggregation of language resources and providing primarily higher education advisory services where practical.

8.3 STRATEGIC OBJECTIVE 3: PROJECTS AND SERVICES

Enable and sustain a broad range of multiphase projects under SADiLaR's mandate.

KEY EMPHASIS: **Increase internal and external synergies. Re-imagining the hub-and-spoke model.**

STRATEGIC PROJECT AND SERVICES AIMS

8.3.1   Run specialisation projects

Activities include

Specialisation projects run by SADiLaR nodes as approved by formal SADiLaR processes, and development of a more prominent SADiLaR brand that unifies the identity of the hub and its nodes and that raises awareness of the RI.

8.3.2   Strengthen South African Sign Language (SASL) resource development

Activities include

A hub-driven specialisation project to support SASL resource development and the establishment of a SASL community of practice.

8.3.3   Provide consultation and guidance

Activities include

Providing expertise in the development of language policies and plans that further multilingual, multimodal teaching and learning frameworks at public universities. Provide support and advice to expand multilingualism at public universities and across government departments as far as our mandate allows.

8.3.4   Create engagement opportunities for broader community

Activities include

Enhancing the visibility of SADiLaR and broadening its community engagement locally and internationally, creating human capacity development project opportunities to include historically disadvantaged institutions, and fostering open science principles.

8.4 STRATEGIC OBJECTIVE 4: **GOOD GOVERNANCE**

Foster good governance and work towards project sustainability.

KEY EMPHASIS: Strengthening governance towards landmarking.

STRATEGIC GOOD GOVERNANCE AIMS

8.4.1 Promote strong and transparent governance

Activities include

Reviewing and updating SADiLaR governance charter and aligning with best practices.

8.4.2 Active succession planning

Activities include

Developing of a formal succession plan, articulating growth opportunities for all staff, and initiating an exchange programme with local and international universities and/or relevant entities to harness and share complementary expertise, and growing communities of practice where members of the community can take more responsibility as part of SADiLaR's ESCALATOR programme.

8.4.3 Financial sustainability

Activities include

Developing a sustainability plan to be less reliant on current sole government funding, identifying synergies with stakeholders and opportunities to leverage funding, and considering co-funding where possible, seeking active partners from the private and public sectors locally and abroad where complementary.

8.4.4 High quality outputs

Activities include

Enhancing holistic quality and quantity of research and technology outputs.

8.4.5 Strengthen relationship with higher education stakeholders

Activities include

Engaging and collaborating with entities such as SAHUDA, NIHSS, USAf, and DHET.

8.5 STRATEGIC OBJECTIVE 5: COMMUNICATION

Foster excellent communication, SADiLaR brand management, and stakeholder engagement.

KEY EMPHASIS: **Strengthening communication and engagement.**

STRATEGIC COMMUNICATIONS AIMS

8.5.1 Grow the SADiLaR brand

Activities include

Making the DSI mandate clear internally and externally, managing expectations with care, ensuring that communication adheres to corporate identity. Further activities are packaging and communicating a distinct message that defines SADiLaR as a locally relevant and effective research infrastructure that is internationally competitive and expansive through its programmes.

8.5.2 Engage key stakeholders on relevant matters

Activities include

Reviewing current and potential new key stakeholders, developing an engagement plan to achieve meaningful and mutually beneficial engagement, and leveraging DSI investment. The following are additional activities: alignment with DSI demographic transformation areas (e.g. of the researcher base and higher levels of research management and the innovation community); institutional transformation (e.g. strengthening previously marginalised institutions based on review findings and changing needs in the National System of Innovation (NSI); an increased footprint of STI activities in non-traditional spaces and targeting non-traditional actors; ensuring greater social and economic benefit to designated priority beneficiary groups; transforming the public's awareness of the value of STI.

8.5.3 Communicate clearly

Activities include

Defining clear communication lines for internal and external stakeholders, choosing appropriate communication platforms, communicating key messages for engaged and informed team and stakeholders, and ensuring articulation and alignment between messages communicated and the strategy/improvement plan activities.

8.5.4 Valorise ethic of belonging

Activities include

Promoting and valorising new ethic of belonging, which embodies SADiLaR's commitment to diversity, inclusivity, and social cohesion for team and stakeholders.

8.6 STRATEGIC OBJECTIVE 6: TEAM

Promote and nurture staff development and wellbeing

KEY EMPHASIS: **Staff recognition and progression.**

STRATEGIC TEAM AIMS

8.6.1 Implement growth orientated performance planning

Activities include

Quarterly task planning linked to the host's performance agreement and appraisal cycles, running cycles with sensitivity and mutual respect, active succession planning and deputising, and active SADiLaR related project involvement.

8.6.2 Prioritise wellbeing in the workplace

Activities include

Creating social platforms to nurture emotional, physical, social and workplace wellbeing, and fostering the ethic of belonging.

8.6.3 Develop recognition systems

Activities include

Introducing an annual research staff and support staff recognition system, celebrating growth and taking on responsibility, and engaging NWU on creation of a SADiLaR organogram to allow for career pathways.

SADiLaR

South African Centre for Digital Language Resources

📞 +27 18 285 2750

✉️ info@sadilar.org

🌐 https://sadilar.org

f www.facebook.com/SouthAfrican
CentreForDigitalLanguageResources

𝕏 SADiLaR_ZA

in South African Centre for
Digital Language Resources