



NICIS

NATIONAL INTEGRATED
CYBERINFRASTRUCTURE SYSTEM

DIRISA

AN INITIATIVE OF:



science, technology
& innovation

Department:
Science, Technology and Innovation
REPUBLIC OF SOUTH AFRICA



Open Science and Research Data Management in South Africa: Origins, Challenges, Benefits, and Impact

Dr. Katlego W. Phoshoko,
Data Services and Applications Manager: DIRISA

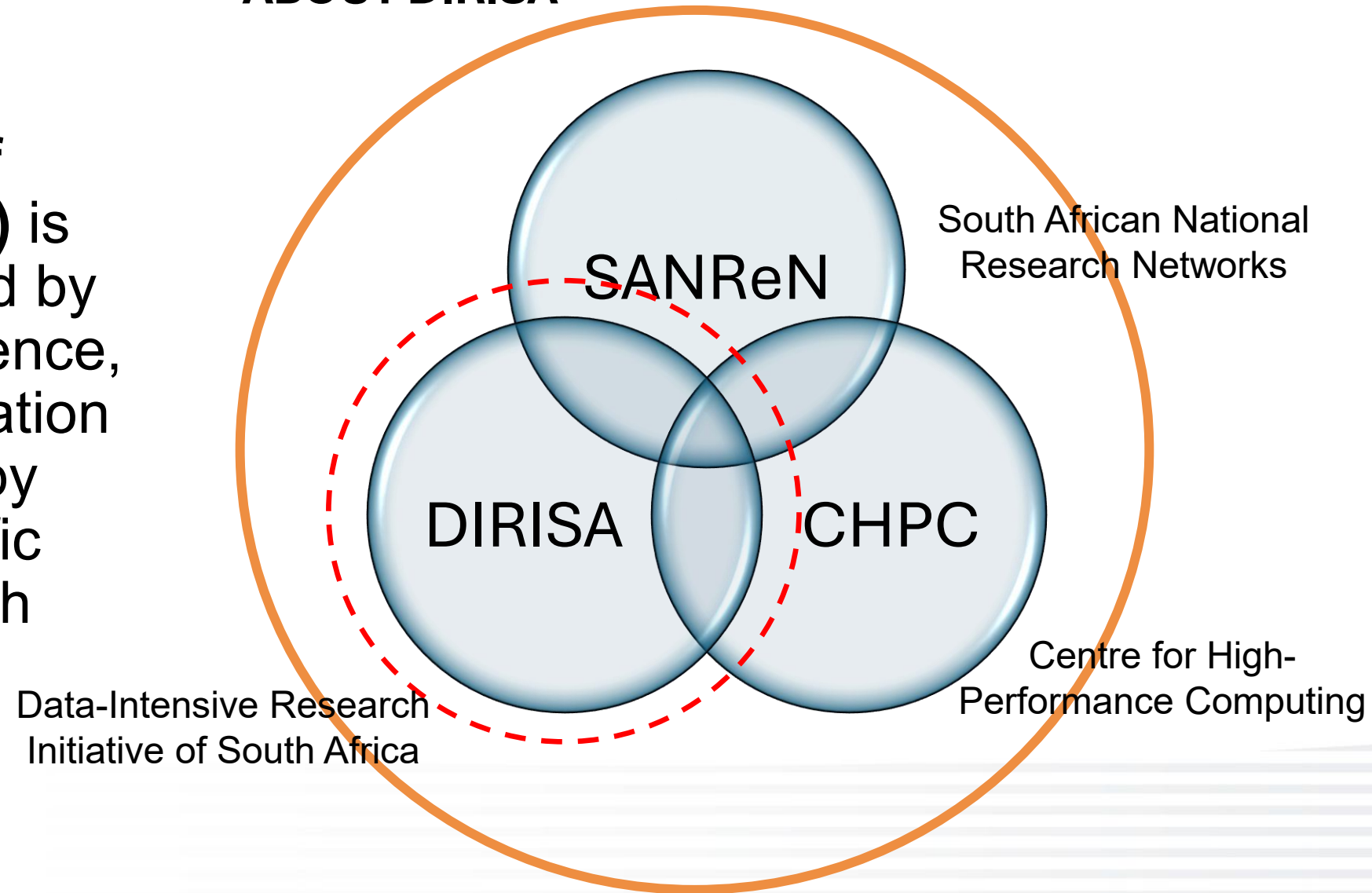
University of Pretoria, International Open Access Week 2025 Series,
Lunch hour session, Theme: "Who Owns Our Knowledge?", 20 October 2025



About NICIS-DIRISA

ABOUT DIRISA

The Data Intensive Research Initiative of South Africa (DIRISA) is an initiative established by the Department of Science, Technology and Innovation (DSTI) and managed by the Council for Scientific and Industrial Research (CSIR).



A national initiative of the Department of Science, Technology and Innovation, and Implemented by the CSIR



Outline

- High-level overview
- Foundations of open science
- South African context
- Ownership
- Role of RDM
- Stakeholders and Ecosystem
- Why it matters
- Challenges
- Emerging trends
- Concluding remarks

High-level overview

Imagine a South Africa where research data is freely shared, empowering scientists, farmers, and communities to solve pressing challenges like climate change and health inequities. This is the power of **Open Science** and Research Data Management (RDM).



Given the theme: **"Who Owns Our Knowledge?"**—We explore the critical role of Research Data Management (RDM) in answering this question, particularly within the South African context.



We will explore the South African National Open Science Policy, examining its origins, the challenges it addresses, its potential benefits, and its impact on the national research landscape.



Open Science and RDM democratize knowledge, foster innovation, and align with South Africa's goals of equity and progress.



The Genesis of Openness (Global to Local)

Global Roots

- Open Science emerged from the Budapest Open Access Initiative (2002), advocating free access to research to enhance transparency and collaboration.
- RDM evolved to support open data sharing and reproducibility (Wilkinson et al., 2016).

South Africa's Journey

- Driven by the need to address historical inequities and boost innovation, South Africa embraced Open Science and RDM to make knowledge a public good.

Key Milestones

- Global: UNESCO's 2021 Recommendation on Open Science sets global standards for open data and RDM.
- Regional: The African Open Science Platform (AOSP) promotes data sharing across Africa.
- National: The 2019 White Paper on Science, Technology, and Innovation prioritizes open access and RDM (DST, 2019).

Current Status

- South Africa is building an Open Science ecosystem, with RDM gaining traction but facing uneven adoption.

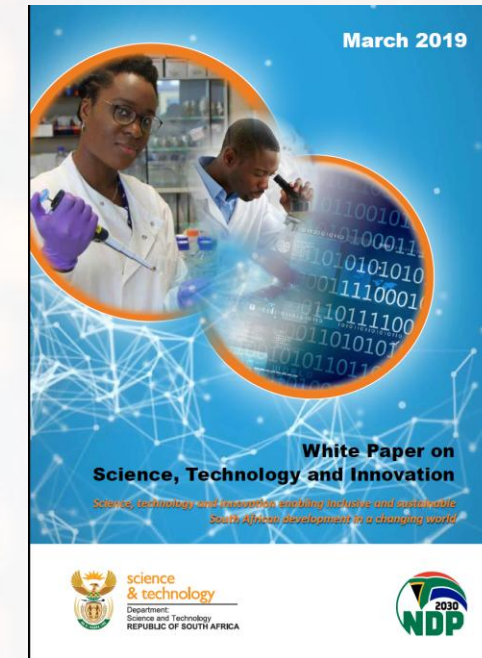
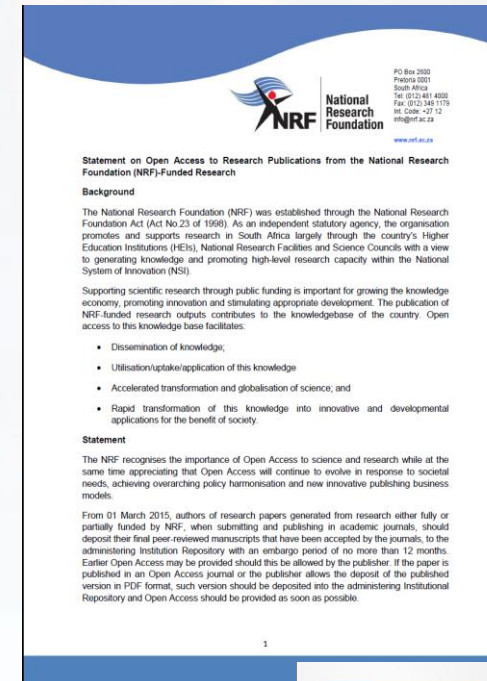


A South African Context



South Africa's Open Science and RDM Framework

- Policy Landscape:
 - The NRF's 2015 Open Access Statement mandates open access for publicly funded research, emphasizing RDM.
 - The 2019 STI White Paper integrates Open Science and RDM into the National System of Innovation.
 - South Africa endorses UNESCO's 2021 Open Science Recommendation, committing to robust RDM practices.
 - SANLIC - Instrumental in negotiating transformative agreements to shift funding from subscriptions to Open Access.
 - Policy Status - The DSTI is actively developing a National Open Science Policy, guided by the principle: "As open as possible, as closed as necessary."
- Infrastructure:
 - DIRISA: Provides RDM services, data repository, and training
 - CHPC: Supports research with high-performance computing facilities
 - Institutional Repositories: Platforms like UCT's ZivaHub enable open data sharing
- Regional Collaboration: AOSP fosters cross-border RDM and data sharing in Southern Africa





These frameworks reshape knowledge ownership, but who holds the power?



Who Owns Our Knowledge? Ownership and Power Structures

- Ownership Dynamics:
 - Public Ownership: Publicly funded research (e.g., via NRF) should be openly accessible, supported by RDM to ensure data integrity (NRF, 2015). **The Core Conflict** - Knowledge is a **Public Good** (when publicly funded), but it is captured as a **Private Commodity** (via copyright/paywalls).
 - Private Control: Commercial publishers (e.g., Elsevier) restrict access through paywalls, limiting open data availability (Larivière et al., 2015).
- Power Structures:
 - Government: DSI and NRF drive open access and RDM policies (DST, 2019).
 - Publishers: Control journal ecosystems, creating barriers to open sharing (Larivière et al., 2015).
 - Institutions: Universities promote RDM but face resource constraints. They are the producers and custodians. Their Intellectual Property (IP) policies determine whether research data and code are institutionally owned or belong to the researcher.
 - Researchers: Navigate tensions between open data sharing and career incentives tied to traditional publishing (Tennant et al., 2016).
- South African Perspective: Open Science and RDM shift power toward public ownership, aligning with equity and societal benefit (Fecher & Friesike, 2014).
- The Ethical Challenge: Indigenous Knowledge Systems (IKS) - Western IP laws often fail to protect communal ownership of traditional knowledge. The South African IKS Policy is crucial in addressing this ethical challenge.



Question: How can RDM ensure knowledge serves the public good?



Open Science and Open Data: The Role of RDM

- Open Science Defined: A movement to make research processes, data, and outputs transparent, accessible, and reusable (UNESCO, 2021).
- Open Data and RDM: Open Data involves sharing research data under FAIR (Findable, Accessible, Interoperable, Reusable) and CARE (Collective Benefit, Authority to Control, Responsibility, Ethics) principles, enabled by robust RDM (Wilkinson et al., 2016; Carroll et al., 2020).
- South African Context:
 - Open Data supports national priorities like health equity (e.g., COVID data) and sustainable development (e.g., climate data for agriculture).
- RDM's Role: Ensures ethical data curation, storage, and sharing, aligning with funder mandates and enabling Open Science (NRF, 2015).
- RDM advanced open science and fosters:
 - **Transparency and Reproducibility:** This is essential for scientific integrity and increasing public trust in research outcomes.
 - **Collaboration and Efficiency:** RDM ensures data is organized with clear file-naming conventions, standardized formats, and detailed documentation. This makes it easier for collaborators (both internal and external) to understand and use the data, leading to more efficient research and interdisciplinary discovery.
 - **Maximizing Public Investment:** In line with the FAIR, by preserving and sharing data, RDM ensures that the outputs of publicly funded research are reusable, preventing the waste of resources on generating data that already exists. This accelerates the pace of innovation and knowledge creation.



Who are the key players driving this transformation?



Stakeholders in South Africa's Open Science and RDM Ecosystem

The successful implementation of Open Science relies on dedicated infrastructure and actors:

- Government: DSTI and NRF provide funding and policy direction for Open Science and RDM (DST, 2019; NRF, 2015).
- Research Institutions: Universities (e.g., UP, UCT, Stellenbosch) and councils (e.g., CSIR) generate data and implement RDM practices. Are crucial as the knowledge producers, custodians, and implementers of RDM policies within their communities.
- Infrastructure Providers:
 - DIRISA supports RDM through repositories and training.
 - CHPC enables data-intensive research with computing resources.
- Researchers: Must adopt RDM and open data practices, shifting from traditional norms (Tennant et al., 2016).
- Communities: Citizen science initiatives (e.g., biodiversity monitoring) contribute to data creation.
- Global Partners: UNESCO and AOSP provide standards and collaboration opportunities.



Why it matters



Benefits of Open Science and RDM

Accelerated Discovery: Open Data, supported by RDM, enables data reuse, speeding up research.

Global Collaboration: Facilitates interdisciplinary and international partnerships, enhancing South Africa's research profile.

Increased Visibility and Impact: Open outputs gain wider citations and societal reach. Making research data and publications openly available increases their global impact and visibility, which benefits researchers and institutions.

Social and Economic Gains:

- Knowledge democratization supports SDGs like poverty reduction and health equity, as it enables citizens, policymakers, and civil society to utilize research to address complex national challenges.

- Open Science stimulates innovation in various industries like Agriculture, Biotech, and supports the development of new data-driven industries, fostering a highly skilled workforce and contributing to economic growth.

RDM Advantage: Ensures data quality, reproducibility, and compliance with open access mandates. Making data public allows for scrutiny, validation, and replication, thereby enhancing the rigor and integrity of scientific outputs.

Example: Open climate data, managed via RDM, can help farmers adapt to changing conditions, boosting food security.



What challenges stand in the way of these benefits?



Challenges to Open Science and RDM in South Africa



Cultural Resistance: Fear of data misuse or loss of competitive edge hinders open sharing.



Infrastructure Gaps: Limited internet access and outdated systems, especially in rural areas, restrict RDM implementation.



Skills Shortage: Lack of expertise in RDM and data stewardship slows progress.



Ethical Concerns: Balancing openness with privacy, particularly for indigenous and health data, is critical.



Funding Barriers: High costs of open access publishing and RDM infrastructure strain resources.



How is South Africa
addressing these challenges,
and what's on the horizon?



Emerging Trends in Open Science and RDM

- The DSTI is actively developing a National Open Science Policy, guided by the principle: "As open as possible, as closed as necessary."
- UNESCO Recommendation (2021): Promotes inclusive, equitable science; South Africa aligns through RDM policies.
- FAIR Principles: Ensure data is Findable, Accessible, Interoperable, and Reusable. The core technical standard for RDM, focused on maximizing machine-readability and reusability of data.
- CARE Principles: People- and purpose-oriented principles that complement FAIR; ensures data sharing respects governance rights and prioritize ethical governance for indigenous data, aligning with South Africa's equity goals (Carroll et al., 2020).
- Citizen Science: Engages communities in research (e.g., Southern African Bird Atlas Project, SABAP2). The rise of projects involving the public in research, fostering engagement and decentralized data collection.
- AI and Machine Learning: Enhance RDM and data analysis in Open Science projects. These technologies are being increasingly used to manage, analyze, and extract insights from vast, openly available datasets within Open Science initiatives.
- Open Educational Resources (OER): A growing movement towards making teaching and learning materials openly available to address educational equity and supports open knowledge.



Conclusion

- South Africa's Open Science and RDM journey, rooted in global and local imperatives, redefines knowledge ownership, drives innovation, and promotes equity. Challenges like infrastructure and skills gaps persist, but the impact is transformative.
- Open Science and RDM empower South Africa to lead in equitable, impactful research, addressing global and local challenges. Open Science is the vision; RDM is the essential tool and practice that achieves the Open Data component.
- Call to Action:
 - Researchers: Adopt FAIR/CARE principles and prioritize RDM.
 - Institutions: Invest in RDM infrastructure and training.
 - Policymakers: Strengthen funding and open access policies.
 - Communities: Engage in citizen science to co-create knowledge.
- By embracing Open Science and RDM, South Africa can answer "Who Owns Our Knowledge?" with a vision of shared, equitable progress.



References

- Motshegwa, T., Development Of The African Open Science Platform (AOSP) Towards a Continental Open Science Vision. <https://www.rd-alliance.org/wp-content/uploads/2024/05/AOSP-Briefing.pdf>
- Budapest Open Access Initiative (BOAI). (2002). Budapest Open Access Initiative declaration. <https://www.budapestopenaccessinitiative.org/read/>
- Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., et al. (2020). The CARE principles for indigenous data governance. *Data Science Journal*, 19(1), 43. <https://doi.org/10.5334/dsj-2020-043>
- Centre for High Performance Computing (CHPC). <https://www.chpc.ac.za>
- Data Intensive Research Initiative of South Africa (DIRISA). <https://www.dirisa.ac.za>
- Department of Science and Technology (DST). (2019). White Paper on Science, Technology and Innovation. https://www.gov.za/sites/default/files/gcis_document/201912/white-paper-science-technology-and-innovation.pdf
- Fecher, B., & Friesike, S. (2014). Open science: One term, five schools of thought. In S. Bartling & S. Friesike (Eds.), *Opening Science* (pp. 17-47). Springer. https://doi.org/10.1007/978-3-319-00026-8_2
- Larivière, V., Haustein, S., & Mongeon, P. (2015). The oligopoly of academic publishers in the digital era. *PLoS ONE*, 10(6), e0127502. <https://doi.org/10.1371/journal.pone.0127502>
- National Research Foundation (NRF). (2015). NRF statement on open access to research publications. https://chelsea.ac.za/wp-content/uploads/2018/12/NRF_open_access_statement_19Jan2015_V6.pdf
- Tennant, J. P., Waldner, F., Jacques, D. C., et al. (2016). The academic, economic and societal impacts of open access: An evidence-based review. *F1000Research*, 5, 632. <https://doi.org/10.12688/f1000research.8460.3>
- UNESCO. (2021). UNESCO Recommendation on Open Science. Paris: UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000379949>
- University of Cape Town (UCT). ZivaHub: Open data repository. <https://zivahub.uct.ac.za>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

Thank you!

