



Contribution ID: 174

Type: **Presentation**

The implications of CODATA's priorities for countries such as South Africa

Monday 13 October 2025 14:52 (11 minutes)

Founded in 1966, CODATA is the Committee on Data of the International Science Council (ISC). CODATA's vision "is of a world in which science is empowered to address universal challenges through the transparent, trustworthy and equitable use of data and information"[CODATA 2025]. CODATA's mission "is to connect data and people to advance science and improve our world"[CODATA 2025]. CODATA has three long-standing priorities:

- Data Policy, aimed at meeting current and urgent challenges at both the international and national levels.
- Data Science and Stewardship for the professions, through practical initiatives such as developing terminology and fundamental constants.
- Data skills capacity building activities, particularly for early career researchers, repository professionals and those researchers interested in the FAIR principles for interoperability: data that are findable, accessible, interoperable and reusable [CODATA 2025].

CODATA's Strategic Plan for 2023-2027 has four thematic priorities [CODATA 2023]:

1. Making Data Work for the Cross-Domain Grand Challenges, to support the plans of the ISC. The ISC identifies these grand challenges when necessary, such as climate change, sustainable development and reducing the risks of disasters. CODATA is doing this through the WorldFAIR+ initiative.
2. Improving Data Policy: this focuses on the principles that data should be FAIR, that is, findable, accessible, interoperable and reusable (see below).
3. Advancing the Science of Data and Data Stewardship: this is to promote evidence-based research and policies and the required systems, standards and infrastructure.
4. Enhancing Data Skills: these are needed to ensure that the data stewardship and science are trustworthy, equitable and transparent [CODATA 2023].

The FAIR principles have become influential and are widely cited. They were developed to facilitate reusing existing data holdings easily and correctly, not just by humans but also automatically by computers [Wilkinson et al 2016]. They are:

- Findable: such as using unambiguous, persistent identifiers and providing metadata that allow the data to be discovered.
- Accessible: such as explicit access conditions and well-described technical access protocols.
- Interoperable: such as standard machine-encoded definitions of the key concepts, variables, etc.
- Reusable: such as clear licensing and details of fair use, and suitable metadata, including provenance and quality [Wilkinson et al 2016].

The FAIR principles have led to related initiatives, such as:

- Machine-actionable FAIR Implementation Profiles (FIPs) to help different disciplines implement the FAIR principles.
- The Cross-Domain Interoperability Framework (CDIF), which provides a framework of standards, particularly for the interoperable and reusable FAIR principles. See Gregory et al [2024].
- The Leiden Declaration on FAIR Digital Objects (FDOs), for individuals and organisations to commit to FAIR data, open standards and increased reliability and trustworthiness [FDO Forum 2022].

WorldFAIR was a successful project aimed at collaboration to implement the FAIR principles, done through 11 case studies. CODATA then launched the WorldFAIR+ initiative to focus on practical guidance and technical recommendations to increase the availability of FAIR data. WordlFAIR+ includes projects around the world, including Data Science Without Borders project, with several African countries participating.

For centuries, the San peoples have been studied by academics, but with concern over being objectified, doubt over usefulness and even perceptions of actual harm, the San leaders initiated the San Code of Research Ethics

[South African San Institute 2017]. With similar initiatives by other Indigenous Peoples around the world, this led to the CARE Principles for Indigenous Data Governance, to balance protecting Indigenous rights and interests with open data, etc [Carroll et al 2020]. The CARE principles are:

- Collective Benefit: Indigenous data must help Indigenous Peoples achieve inclusive development and innovation and realise equitable outcomes.
- Authority to Control: in line with the United Nations Declaration on the Rights of Indigenous Peoples [UNGA 2007], Indigenous Peoples need to be able to govern their own data and have sovereignty to facilitate greater Indigenous self-determination [Hudson et al 2023].
- Responsibility: researchers need to nurture respectful relationships with Indigenous Peoples, including developing their capacities, and embedding the data within the languages and cultures of the Indigenous Peoples.
- Ethics: the research must protect the rights and wellbeing of the Indigenous Peoples throughout the data lifecycles to minimise harm, maximise benefits, promote justice and allow for future use –but the Indigenous Peoples must determine this [Carroll et al 2020].

The CARE principles work together with the FAIR principles, rather than contradicting or competing with them. For South Africa, the CARE principles are perhaps more important than the FAIR principles.

This paper will explore the implications of the priorities listed above for countries such as South Africa, southern Africa or Africa as a whole. Before a data set can become FAIR, it needs to exist, and there are concerns that much data have been lost in South Africa. The Promotion of Access to Information Act [South Africa 2000] is similar to freedom of information legislation in other countries. The Act requires public bodies to compile and make readily available a manual on what records can be accessed and how. However, compliance is poor, even though there are official penalties such as fines or imprisonment for non-compliance [Fourie 2023]. Unsurprisingly, this does not bode well for the availability of data sets, never mind those that comply with the FAIR and CARE principles.

CODATA is looking for more partners to provide case studies for WorldFAIR+, but the obvious limitation for South African organisations is funding. Besides the global problems, the South African economy has been doing poorly for some years.

Primary author: COOPER, Antony (CSIR)

Co-author: Dr VAHED, Anwar (Retired)

Presenter: COOPER, Antony (CSIR)

Session Classification: Presentations Session 1: CAREful Indigenous Data Governance

Track Classification: SciDataCon2025 Specific Themes: CAREful Indigenous Data Governance