

Curtin Open Knowledge Initiative

Open Knowledge profile of South African higher education and research institutions

Key policy and performance data: Library public access, open access, and open data

Report and data prepared by Katie Wilson, Chloe Brookes-Kenworthy, Richard Hosking, Chun-Kai (Karl) Huang, Lucy Montgomery, Cameron Neylon, Alkim Ozaygen.

Curtin Open Knowledge Initiative,
Curtin University,
Perth, Western Australia, Australia

November, 2019

Introduction

This report presents a profile of progress towards institutional open knowledge (OK) in selected universities and research institutions in South Africa. For the purpose of this report, we analysed data from forty-six educational and research institutions in South Africa. This number includes including twenty-six public universities and universities of technology identified by the South Africa Department of Higher Education and Training (n.d.), and twenty research institutes and councils identified through the GRID identifier website (Digital Science & Research Solutions, 2019) and other sources.

Methodology

Data Sources

Data were collected from several key sources to ascertain the existence of institutional policies, statements, services and facilities supporting open knowledge. We gathered documents from university and research institution public websites relating to library public access, open access, open access publishing funds, institutional repositories, open data repositories, data sharing and research data management.

We employed a Jupyter notebook supported by a small library of Python code to identify and retrieve policy and related documents. Using the Bing search engine API, the notebook tool executes a search against the URL for specific university website recorded in the Global Research Identifier Database - GRID (<https://www.grid.ac/>). The search returns five pages for each website from which we select, view and download relevant documents and web pages.

Manual searching supplemented the automated Jupyter notebook process and tools where necessary to find additional documents and to confirm the presence or absence of policies and repositories. In addition, we used the following directories and collections:

Directory of Open Access Repositories (DOAR) (<http://v2.sherpa.ac.uk/opensoar/>)

Registry of Open Access Repository Mandates and Policies (ROARMAP) (<https://roarmap.eprints.org/>)

African Digital Research Repositories (<https://www.internationalafricaninstitute.org/repositories>)

For the Open Access performance data, the Curtin Open Knowledge Initiative (COKI) group collects and analyses items with DOIs representing institutional research output from bibliographic databases Microsoft Academic (MSA), Scopus and Web of Science. [Crossref](#) metadata collected for each DOI identified provides information on publication dates. DOIs are checked against [Unpaywall](#) to determine the open access status for institutional publications.

Search terms and definitions

The following search terms were used to locate relevant documents and statements on websites:

- library access, library membership, external users, borrowing
- open access policy, open access mandate, open access funds, open access publishing
- institutional repository, data repository, open data repository
- data sharing, sharing data, open data
- data research management, data management plan

We searched or browsed directories by continent, region and/or country, according to the organisation each of the tools used. Terminology differs somewhat between universities and research institutions.

Library public access indicates an institutional library is open to external, unaffiliated or non-academic members of the public, as expressed in a policy document, statement or information page. This can include individual researchers, students from other universities as well as day visitors. Fees for access and membership are often charged by universities for external users access and membership. A small number of research councils provide details of publicly accessible libraries or information centres.

Open access (OA) policy or mandate is a publicly available policy document or statement indicating institutional support for sharing and deposit of research publications in an OA repository. The 2003 Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities of 22 October 2003 is one of the milestones of the Open Access movement.

Institutional repository refers to a repository created or maintained by an institution for the purpose of deposit of research publication output. Such a repository may be shared by more than one institution.

Open Access funds may be made available by an institution to enable academics/researchers to publish research in gold OA journals through funding or subsidising article processing charges (APCs) or book processing charges (BPCs).

Research data management policies or statements outline the institutional position and guidelines on managing, storing, preserving and accessing data produced as part of research and throughout the research lifecycle. This may include a statement on open access to data, and details of data management plans.

A data sharing policy or statement sets out an institutional position on the sharing of, or open access to research data or datasets, and may refer to the deposit of such data in an institutional data repository or external data repositories.

An open data repository is a repository established and maintained by a single institution or multiple institutions, providing open access to data and datasets produced as part of research projects conducted by institutional academics and researchers.

Limitations

The main limitations relate to the availability and non-standardisation of relevant data on websites and the incompleteness of some documents. Dates of policy approval are not always present in policy documents. References made to institutional policies in guides and statements on webpages do not always include links to those documents. The lack of connective links between policies and information regarding services and facilities such as repositories required additional searching. The availability of OA funds to cover the costs of journal APCs and book BPCs is assessed through publicly accessible data from websites, but may not include the provision of funds within faculties, departments and schools not publicly advertised.

The presence or absence of a policy, service, facility or repository recorded in this report was as identified at the time of searching (September - October 2019). If policies, services or facilities were in process, under development or being updated at the time of retrieval this was recorded as a nil presence (No). If we were unable to locate a publicly available policy or repository or if policies or statements were only accessible through a staff intranet this was recorded as No.

The bibliographic data sources used for OA output analysis have substantial biases and limitations with respect to affiliation sources. We believe these are mitigated through our use of three independent data sources. The precise levels of open access should be treated with caution and comparison to other analyses is not straightforward. Comparisons made using other Unpaywall data and within this report may be considered reliable with the caveat that they measure visibility and indexing and not necessarily the up-to-date contents of the repository.

Analysis

In 2015, the National Research Foundation (NRF) of South Africa, a key national funder, issued a statement on open access publications, requiring from 1 March 2015 authors of research papers generated from research either fully or partially funded by NRF to deposit author accepted manuscripts (AAM) or final PDFs into an Institutional Repository, with an embargo period of no more than 12 months (National Research Foundation, 2015). The statement also declared that data supporting publications should be deposited in an accredited Open Access repository. Many institutional websites refer to the NRF Open Access statement.

Following the NRF 2015 mandate, many universities and some research institutions established open institutional repositories for the deposit of research output. In addition, some universities have thesis deposit repositories that may have preceded institutional repositories, and separate repositories for specific digital collections.

Library public access

Twenty-one public universities provide details on their websites indicating access and membership to the library for the public, community or external users who are unaffiliated with the institution, on a day visitor basis and/or for longer term research. Fees are charged by eighteen universities for membership, which includes borrowing of material. Some universities charge for day visitor access. Several universities specify categories of external users and may specifically exclude groups such as school students, or students from other tertiary institutions who are not affiliated through associations and consortia such as CALICO (the Cape Library Cooperative) and CHELSA (Committee of Higher Education Libraries South Africa). Library public access details were unavailable or not current for six university libraries.

Four research institutes indicate public access to their libraries or information centres on their websites and do not charge fees. Those for which details of a library or information centre were not available are recorded as NA (not applicable).

Institution type	Library public access	Fees charged
Universities	21	18
Research institutions	4	0

Table 1. Library public access

Open access

Twelve universities and two research institutions (including the National Research Foundation) display public policies, mandates or statements requiring or recommending OA publication or repository deposit of research output from institutional members. Signatories to the 2003 Berlin Declaration on Open Access (Max Planck Gesellschaft, 2019) include 17 universities and two research institutions. Citing the NRF statement and/or the Berlin signatory on a webpage appears to be a proxy for institutional policies and an indication of institutional commitment to OA in some cases. Four universities make funds available to support researchers publishing in OA. The funds subsidise or cover processing charges (APCs, BPCs) for publishing scholarly articles, conference papers, and books with open access publishers or on open access platforms, and may contribute towards membership fees with OA publishers.

Twenty-nine institutions (24 universities and five research institutes) have open access institutional repositories (IRs) providing public open access to research publications of their staff. The majority of South African

universities host their own or share a repository. Nelson Mandela University, Walter Sisulu University, University of Fort Hare and Rhodes University make up the South East Academic Libraries System (SEALS) and share a digital repository. Some repositories include theses and dissertations, while other institutions have separate repositories for thesis deposit or to store specific archive collections. For example, Stellenbosch University has three repositories: SUNScholar Research Repository (<http://scholar.sun.ac.za/>), SUNETD Doctorates and Masters Submissions (<https://etd.sun.ac.za/>), and SUNDigital Collections (<http://digital.lib.sun.ac.za/>) with digital content from the Library and Information Service's collections. Similarly, the University of Western Cape hosts an institutional repository, (<http://repository.uwc.ac.za/>), the UWC Electronic Theses and Dissertations repository (<http://etd.uwc.ac.za/xmlui/>) and the Africa Higher Education Research Online repository (<http://ahero.uwc.ac.za/>). Several research institutions include lists of publications in place of IRs but these have not been counted as repositories.

Institution type	OA Policy/mandate	Berlin signatory	OA funds	Institutional Repository
Universities	12	17	4	24
Research institutions	2	2	0	5

Table 2. South Africa Open Access policies or mandates, Berlin declaration signatories, OA publishing funds, institutional repositories

Open access performance

Figure 1 depicts the growth of OA research output in 2000-2018 as a percentage of total research output for the **38** South African institutions included in this report (those with available OA publication data). The **total OA** output reaches 48 % in 2014 and again in 2017. Percentages represent OA research output available through repositories and published journals. In addition to the total **percent OA**, the figures include percentages of **Green OA** (self-archived in an open repository), **Gold OA** (published in an OA journal where all articles are free to read), **Bronze OA** (free to read on a publisher page, but without a license for reuse of reproduction), and **Hybrid OA** (published in a subscription journal with a free license, courtesy of an article processing charge (APC) paid by authors) (Piwowar et al., 2018). Items can be both Green and Gold, Hybrid or Bronze (published). **Green-only OA** refers to OA output archived only in an institutional repository and not published (in the bibliographic sources used).

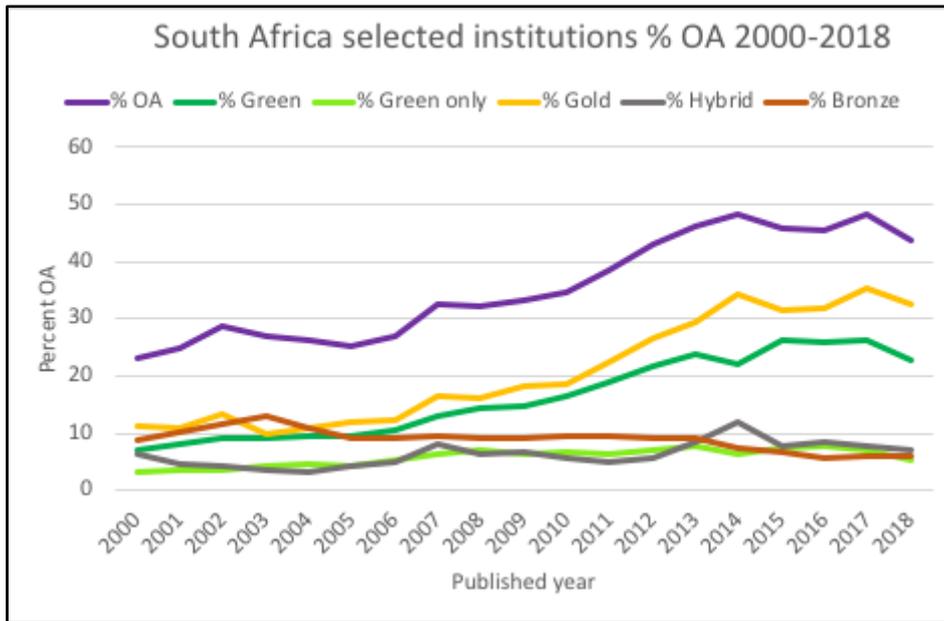


Figure 1: Percentages of OA research output for 38 institutions, 2000-2018. COKI analysis, October 2019.

The number of research output items with DOIs published by the 38 institutions, and the corresponding number of items available as OA is shown in Figure 2.

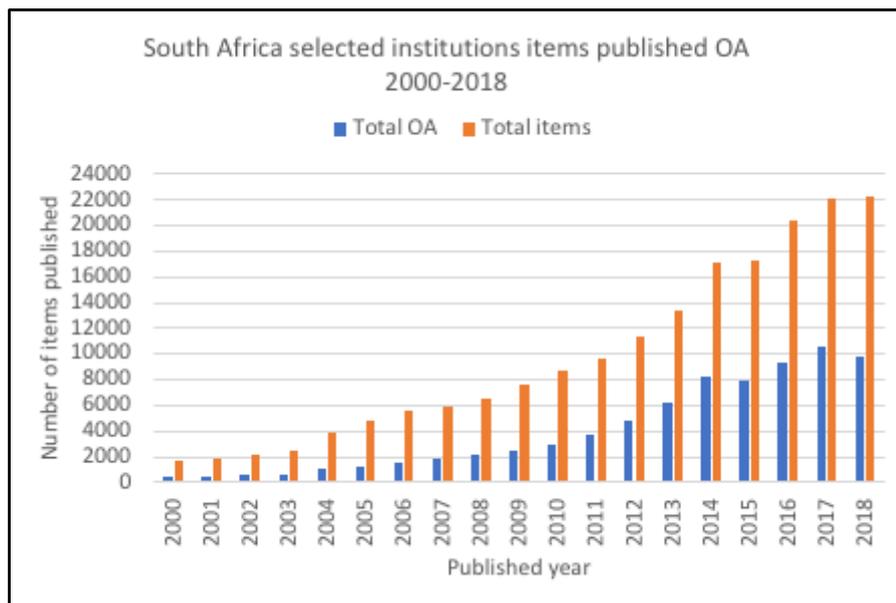


Figure 2: Number of published research items with DOIs and corresponding number of items in OA for 38 South African institutions, 2000-2018. COKI analysis, October 2019

Open data

Nine institutions (five universities and four research institutes) have publicly available research data management (RDM) or data curation policies. Several universities websites indicate an RDM policy is in process or under development, but these are marked as not having a policy (at the time of checking - September/October 2019).

Eleven institutions (five universities and six research institutions) have publicly available data sharing policies or guidelines, sometimes embedded within RDM policies and planning documents, library guides or annual reports.

Eleven institutions have open data repositories: six universities and five research institutes. Several university repositories are hosted through Figshare, and some include datasets in their IR. Where a policy exists, but no institutional data repository, a policy may suggest the use of external discipline-based or national data repositories. The Datafirst Open Data Portal at the University of Capetown (<https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/central/about>) hosts research datasets from individual departments, research units and projects from the University of Cape Town, Rhodes University, University of KwaZulu-Natal, the University of Johannesburg and the University of the Witwatersrand. However, the repository collections are not always cross-referenced or linked at the home institution websites. A survey by Bangani and Moyo (2019) found some concerns about sharing data, such as privacy, ethical issues and data misuse, particularly among emerging researchers, the primary respondents, indicating that data sharing is not yet widely adopted. The research institutes with open data repositories have public statements indicating RDM and data sharing policies.

Institution type	RDM Policy/Plan	Data sharing policy	Open data repository
Universities	5	5	6
Research institutions	4	6	5

Table 3. Research data management policies, data sharing policies, open data repositories

Discussion

Public access to libraries is offered by the majority of South African universities investigated, usually with fees charged for membership with additional privileges, indicating openness to sharing access to knowledge in these institutions. Four research institute libraries are open to the public, the Council for Scientific and Industrial Research, MINTEK, the National Institute for Occupational Health, part of the National Health Laboratory Service and the South African National Biodiversity Institute and do not appear to charge fees.

The presence of institutional repositories (IRs) for the deposit of research publications, theses and dissertation is well established among South African universities. Some show small numbers of documents deposited although it is possible that research output is deposited in other South African and external repositories. An OA policy, mandate or statement was identified in half (thirteen) of the universities. This suggests the adoption and practice of depositing research output in institutional OA repositories may be less widespread among members, but this may also be a reflection of gold OA publication through ScieELO. IRs are reasonably well advertised on institutional websites but more coordination with related information such as policies and guidelines would be good practice (see University of Pretoria). There is less public evidence of OA policies and institutional repositories among the research institutions. The OA performance figures indicate research output available as gold OA has grown since the establishment of the open access ScieELO South Africa in 2013.

The National Research Institute 2015 statement recommends associated research datasets be deposited in an open repository. We found institutional open data repositories in a small number of institutions, but it is possible that datasets may also be deposited in external repositories. Reference was found on one university website to a draft NRF Data Management Plan but no further evidence was located. The Africa Open Science Platform (AOSP) encourages the open sharing of research data through providing tools and systems for managing data, and a three year pilot project for the AOSP ended in October 2019 (Participants of African Open Science Platform Stakeholder Workshop, 2018).

South African institutions are making good progress on the path to openness. Greater focus on sharing data will enhance this, and the continued promotion and clarification of institutional Open Knowledge policies and practices via public websites and internally will expand the national position. Appendix 1 summarises the OKI features for the institutions included in this report.

Recommendations:

- Clarify institutional positions on public access to libraries where relevant.
- Increase institutional promotion of OA commitment through policies and practices.
- Coordinate and link institutional OA policies, practices and explanations with institutional repository entry pages.

Appendix 1: List of institutions and OKI features

Institution	Library public access	OA policy or statement	Berlin declaration signatory	OA funds	IR	RDM policy	Data sharing policy	Open data repository	OA output
Academy of Science of South Africa	NA	Y	Y	N	Y	N	N	N	Y
Africa Health Institute	NA	N	N	N	N	Y	Y	Y	NA
Agricultural Research Council of South Africa	NA	N	N	N	N	N	N	N	Y
Cape Peninsula University of Technology	Y	Y	Y	N	Y	Y	N	Y	Y
Central University of Technology, Free State	Y	N	N	N	Y	N	N	N	Y
Centre for the AIDS Programme of Research in South Africa	NA	N	N	N	N	N	N	N	Y
Council for	NA	N	N	N	N	N	N	N	Y

Geoscience									
Council for Scientific and Industrial Research	Y	N	N	N	Y	N	N	N	Y
Durban University of Technology	Y	Y	Y	N	Y	Y	N	N	Y
Human Sciences Research Council	NA	N	N	N	Y	Y	Y	Y	Y
Institute for Maritime Technology	NA	N	N	N	N	N	N	N	NA
Mangosuthu University of Technology	N	N	N	N	N	N	N	N	Y
MINTEK	Y	N	N	N	N	N	N	N	Y
National Advisory Council on Innovation	N	N	N	N	N	N	N	N	NA
National Health Laboratory Service	Y	N	N	N	N	N	N	N	Y
National Research Foundation of South Africa	NA	Y	Y	N	Y		Y	N	Y
Nelson Mandela University	Y	N	N	N	Y	N	N	N	Y
North West University	Y	Y	Y	N	Y	N	N	Y	Y
Nuclear Energy Corporation of South Africa	NA	N	N	N	N	N	N	N	Y

Rhodes University	Y	Y	Y	N	Y	N	N	N	Y
Sefako Makgatho Health Science University	N	N	N	N	Y	N	N	N	Y
Sol Plaatje University, Northern Cape	Y	N	N	N	Y	M	N	N	NA
South African Bureau of Standards	NA	N	N	N	N	N	N	N	NA
South African Medical Research Council	NA	N	N	N	N	N	N	N	Y
South African National Biodiversity Institute	Y	N	N	N	N	Y	Y	Y	Y
South African National Space Agency	NA	N	N	N	N	N	N	N	Y
South African Weather Service	NA	N	N	N	N	N	N	N	Y
Stellenbosch University	Y	Y	Y	Y	Y	N	Y	Y	Y
Technology Innovation Agency	NA	N	N	N	N	N	N	N	NA
Tshwane University of Technology	Y	N	Y	N	Y	N	N	N	Y
University of Cape Town	Y	Y	Y	Y	Y	Y	Y	Y	Y
University of Fort Hare	Y	N	Y	N	Y	N	N	N	Y

University of Johannesburg	Y	Y	Y	N	Y	N	N	N	Y
University of KwaZulu-Natal	Y	N	Y	N	Y	N	N	N	Y
University of Limpopo	Y	N	Y	N	Y	N	N	N	Y
University of Mpumalanga	N	N	Y	N	N	N	N	N	NA
University of Pretoria	Y	Y	Y	Y	Y	Y	N	N	Y
University of South Africa	Y	Y	Y	N	Y	N	Y	Y	Y
University of the Free State	Y	Y	Y	N	Y	N	N	N	Y
University of the Western Cape	N	Y	Y	N	Y	Y	Y	Y	Y
University of the Witwatersrand	Y	Y	Y	Y	Y	N	Y	N	Y
University of Venda	Y	N	N	N	Y	N	N	N	Y
University of Zululand	Y	N	N	N	Y	N	N	N	Y
Vaal University of Technology	N	N	N	N	Y	N	N	N	Y
Walter Sisulu University	Y	N	N	N	Y	N	N	N	Y
Water Research Commission	NA	N	N	N	Y	N	N	N	N

References

Bangani, S., & Moyo, M. (2019). Data sharing practices among researchers at South African universities. *Data Science Journal*, 18(1), Article no. 28. <https://doi.org/10.5334/dsj-2019-028>

Department of Higher Education and Training (n. d.). Universities in South Africa. Retrieved September 16, 2019 from <http://www.dhet.gov.za/SiteAssets/New%20site%20Documents/Universities%20in%20South%20Africa1.pdf>

Digital Science & Research Solutions. (2019). GRID: Global Research Identifier Database. <https://www.grid.ac/>

Max Planck Gesellschaft (2019). Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities <https://openaccess.mpg.de/319790/Signatories>

National Research Foundation (2015). Statement on Open Access to research publications from the National Research Foundation (NRF)-funded research. Retrieved September 16, 2019 https://www.nrf.ac.za/sites/default/files/documents/oastatement_2015.pdf

Participants of African Open Science Platform Stakeholder Workshop, September 2018, Participants of African Open Science Platform Strategy Workshop, March 2018, Advisory Council, African Open Science Platform Project, Technical Advisory Board, African Open Science Platform, Boulton, Geoffrey, Hodson, Simon, ... Wafula, Joseph. (2018, December 12). The Future of Science and Science of the Future: Vision and Strategy for the African Open Science Platform (v02). Zenodo. <http://doi.org/10.5281/zenodo.2222418>

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., ... Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, 6, e4375. <https://doi.org/10.7717/peerj.4375>