

A roadmap for implementing PIDs in Africa

This roadmap is prepared by the LIBSENSE Working Group on Infrastructure – open access journals, repositories for publications and data and open discovery services.

LIBSENSE is a collaborative effort to promote open science, facilitate sharing of resources and knowledge, and advance the development of digital services and infrastructures in African higher education and research institutions. It is led by the West and Central African Research and Education Network (WACREN) in collaboration with sister regional African RENs (ASREN and UbuntuNet Alliance). Other participating partners include NRENs, libraries, library associations, universities and research communities in Africa, in conjunction with international partners such as COAR, EIFL, University of Sheffield, National Institute of Informatics (Japan), GEANT, OpenAIRE and LA Referencia.

Contact: libsense@ren.africa



Introduction

Sharing and disseminating African research is vital for global knowledge commons. It facilitates participation and engagement, fostering collaboration within research communities. This active involvement of scholars and other knowledge creators contributes to developing a relevant publishing landscape. Regional research productivity is critical to driving growth and progress within African society. It is essential, however, to understand that this is just one aspect of the broader scholarly communication ecosystem. In addition to research productivity, other components contribute to the effective dissemination and accessibility of scholarly content. Open infrastructure provides resources and tools supporting different research workflow stages, including dissemination and outreach. These initiatives aim to foster collaboration and ensure that scholarly information sharing is multidirectional, acknowledging that scholarship is a collective endeavour.

Persistent identifiers (PIDs) are another crucial element of the scholarly communication ecosystem. PIDs serve as long-lasting references to digital objects, enabling easy and reliable access to research outputs of various types. PIDs can enhance the cohesion and discoverability of African scholarly content, ensuring that research remains accessible and traceable.

For Africa to establish sustainable publishing practices and progress collectively as a continent, embracing these components of the scholarly communication ecosystem is essential. Africa can increase its scholarly outputs and actively contribute to the global

research community by prioritising open infrastructure, encouraging participation and utilising PIDs.

Implementing a PID strategy is crucial for African institutions and researchers. PIDs enable the showcasing of local scholarship, improving its visibility and discoverability. Research outputs can be accurately identified using PIDs, regardless of location or affiliation. Additionally, PIDs shed light on contributions within collaborations, whether they involve fieldwork or authors.

Moreover, PIDs are vital in making data FAIR (Findable, Accessible, Interoperable, and Reusable). They facilitate the discovery, longevity, and transparency of research data, contributing to a more robust research ecosystem. This is particularly important for African research, considering the continent's challenges regarding low research productivity and access to scholarly content. Uneven research infrastructure, limited ICT support, publishing practices, and funding compound these challenges. To address these barriers, it is crucial to address the imbalance created by publishing models driven by the Global North, which often exclude African researchers and their scholarship. By implementing PIDs, research collaboration can be improved, ensuring partnerships are more balanced and unbiased between the Global North and the Global South.

To orchestrate an effective strategy, it may be beneficial for the components of the publishing ecosystem to be provided and supported by regional and National Research and Educational Networks (NRENs). This approach would help level the playing field by addressing the unevenness in resources across the continent and enable an Africa-wide strategy for scholarly communication. Considering the cost implications of implementing a PID strategy is vital, as current financial models for consortia are still prohibitive for many African countries. Sustainable funding models must be explored to ensure PIDs' successful implementation and adoption in African research and scholarly communication.

Recommendations

Governance and Finance: Develop appropriate community governance and sustainability models for long-term community engagement

- Establish national and regional PID consortia to oversee governance and explore opportunities and financing models. NRENs and RRENs can play a crucial role in leading governance, and they could also facilitate interoperability with different PID initiatives. The cost model for a PID strategy may encourage a (new) role for RRENs and NRENs as standard PID providers consortia are currently too expensive for many individual African countries.
- Establish governing councils or committees and advisory groups of major research stakeholders and representatives that could provide oversight and management and a space to plan and share information about opportunities and activities in the PID systems. Enforcing a governance structure that defends against a potential commercial takeover is essential.
- Find a compromise and a balance on the institutional, national, regional and continental levels to reduce the costs.

- Seek seed funding for African PIDs for a while.
- Advocate for flexible and fair (equitable) fee models with PID providers requesting to factor in purchasing power parities conversion rates.

Advocacy and Capacity Building: Set up African communities of practice

- Collaborate with PID organisations on advocacy and capacity building for different stakeholders: funders; government departments; universities and research institutes; publishers; researchers; academic libraries, in collaboration with their universities' research offices, are well positioned to run PID training programmes. PIDs should also be included in the open science training for post-graduate and post-doctoral researchers.
- Identify continental, national and institutional PID champions to inspire wider adoption.
- African communities of practice will serve as venues for raising issues and sharing solutions about implementing PIDs and research information more broadly, coordinating information exchange and the exchange of experience.
- It will provide an avenue for sharing training material (toolkits) on advancing PIDs at the institutional level.

Interoperability, implementation and integration

- Increase the adoption and use of PIDs and improve the scale and depth of PID integration in everyday research workflows and continental open science infrastructures.
- Establish reliable (meta) data exchange, especially with funders, on the national, regional, continental and global levels. Funders and institutions should seek to implement ORCID iDs for researchers in grant application, peer review, and project reporting workflows. Set up data exchange pilots on the national and regional levels between institutions, funders and ministries to obtain information about publications by automatically harvesting them from universities' information systems.

Choosing the right PIDs

- Follow these criteria for selecting the right PID:
 - Openness (open data, open source)
 - Transparency
 - Non-proprietary service
 - Trust
 - Inclusivity
 - Interoperability
 - Community-led and community-governed (cannot be owned or bought)
 - Free of charge for users - a combination of no-charge and for-a-minor fee APIs
 - Sustainability
- Focus on open, non-commercial and community-driven efforts for assigning PIDs to objects — publications, data, software, etc. — such as Archival Resource Keys (ARKs)

and Digital Object Identifiers (DOIs), PIDs for people (e.g. researchers) — ORCID iDs, and PIDs for organisations — Research Organization Registry (ROR), that are critical for building and maintaining reliable and robust links between objects, people, communities and infrastructures, eliminating name ambiguity and pointing persistently on the location of a digital object.¹

PIDs in Africa

There still needs to be more awareness in our institutions on how to optimally use PIDs in publications and scholarly outputs. Many publications are produced without any PIDs or researcher IDs. Institutions that understand the need for PIDs usually face the challenge of financing and locally acquiring them for their publications. Some institutions (particularly private institutions) have directly contacted organisations like DataCite and CrossRef to obtain PIDs, including DOIs, but there are more sustainable approaches. Financial considerations must also be addressed as part of the implementation so that whatever PID strategies are adopted are affordable and sustainably funded.

Multiple PID systems exist for different digital objects, operated by various service providers and used by different research disciplines. ISBN on a book, ISSN on a journal, Google Scholar profiles, and handle services in repositories — allowing for a short URL, which is persistent, for citation and discovery on the web — are some of the most commonly used PIDs with well-established use strategies and workflows.

These PID systems are explained in more detail below.

Education Identity Federations - eduID

Digital identity for research and education (eduID) is increasingly important in Africa. This is because more and more services that rely on trust and identity systems for access are emerging. In addition, the adoption of federated identity access is increasing globally. Several national identity federations in Africa have joined [eduGAIN](#), the global international interfederation service that enables the exchange of identity and authentication information among participating academic and research organisations worldwide. These include eduID.ng (Nigeria), eduIDM.ma (Morocco), FIDERN (Zambia), RAFIKI (Kenya), RIF (Uganda), SAFIRE (South Africa) and the SomaliREN Identity Federation.

However, not all countries in Africa have national identity federations². Institutions from such countries can participate through a proxy federation in [eduID.africa](#), the framework for trust and digital identity collaboration established by the three African regional RENs.

¹ Madden, Frances, van Horik, René, van de Sandt, Stephanie, Lavasa, Artemis, & Cousijn, Helena. (2020). Guides to Choosing Persistent Identifiers - Version 3 (Version 3). Zenodo.

<https://doi.org/10.5281/zenodo.4192174>

² <https://connect.geant.org/2023/03/06/eduid-africa-advances-federated-identity-management-in-the-education-ecosystem-in-africa>

Federated identity is essential for a PID roadmap in Africa. First, it can help to raise awareness of PIDs and their importance. Second, NRENs and RRENs are often the operators of national identity federations or catchall federations. This means they are trusted organisations that institutions can approach or work with to obtain other PIDs, making NRENs potential custodians of a national prefix.

Here are some additional details about the importance of federated identity for a PID strategy:

- Federated identity systems enable users to seamlessly access these resources by leveraging their federated credentials across multiple platforms or services. Researchers can efficiently discover and access scholarly content, regardless of the publisher or repository hosting the resource, by combining persistent identifiers with federated identities.
- Federated identity can improve security. Using a single identity to access multiple services makes users less likely to reuse passwords, which can help prevent unauthorised access to accounts. Linking federated identities to PIDs allows access to scholarly resources based on the user's authenticated identity, ensuring appropriate authorisation and security.
- Persistent identifiers are essential for establishing trust and ensuring proper attribution in scholarly communication. They enable accurate citation and referencing of scholarly resources, ensuring credit is given to the original creators or contributors. Federated identity systems can enhance this process by linking the authenticated identities of authors or contributors with their persistent identifiers. This linkage strengthens the integrity and reliability of scholarly information, as the author's identity can be verified and associated with their work.

Overall, federated identity is a valuable tool that can help to improve the efficiency, security, and trust of education services in Africa.

Handle System

The [Handle System](#), established in 1995, is a globally utilised identifier resolution system that provides secure name resolution services. It is managed by the Corporation for National Research Initiatives (CNRI) and serves as a versatile and non-commercial name service. In the Handle System, each identifier, known as a Handle, is composed of a prefix that identifies the "naming authority" and a suffix representing the "local name" of the object. The Handle uniquely identifies an object or resource and facilitates retrieval and management. Notably, the Handle System stores the location of the object and the associated metadata separately. When a Handle is queried through the Handle Service, a server associates the requested Handle with its corresponding metadata and returns that information.

The Handle System provides a robust infrastructure for resolving identifiers and retrieving metadata associated with objects. It serves as a fundamental component of systems such as

the Digital Object Identifier (DOI) system, and repositories can leverage the Handle Service to assign persistent URLs, enabling efficient citation and discovery of scholarly resources online.

Digital Object Identifier (DOI) System

A digital object identifier ([DOI](#)) is a unique alphanumeric string assigned by a registration agency such as the International DOI Foundation to identify content and provide a persistent link to its online location. For example, a publisher assigns a DOI when your article is published and made available electronically.

DOIs utilise the Handle System as the underlying resolving infrastructure. DOIs are assigned by a federation of Registration Agencies worldwide with a mandate to provide processes that ensure that DOI names do not depend on any one agency, guaranteeing their persistence. Each DOI name requires a standard metadata kernel to be defined, and registering DOI names at agencies such as [Crossref](#) and [DataCite](#)³ typically involves a fee. However, the resolution of DOIs is free.

UbuntuNet Alliance joined DataCite as a member in 2021, and its consortium membership is open to all research and education institutions in Eastern and Southern Africa. Registration of DOIs is now part of the UbuntuNet Alliance service portfolio, and there are two member organisations with about 6000 DOIs issued. UbuntuNet Alliance enables research and education institutions with repositories to register and manage DOIs and metadata for their research outputs through this membership. The service is being offered to NRENs and institutions with the expectation that once there are enough members in a country, the respective NREN will be encouraged to set up their consortium.

DataCite launched the Global Access Program (GAP⁴), designed to increase access to PID services and infrastructure and encourage their adoption by communities outside the Global North, supported by a grant from the Chan Zuckerberg Initiative.

The [Grant Identifier initiative](#), led by Crossref and several funders, including Wellcome Trust⁵, JST⁶, European Research Council⁷ and NIH⁸, launched a registration service for grant IDs in 2019. Funders can now register research grants with Crossref. Essentially, DOIs are used in this scheme and assigned to every grant awarded. The African Union, African Capacity Building Foundation, African Development Bank Group, Department of Science and Innovation in South Africa, New Partnership for Africa's Development and the Africa Growth Institute are among African organisations that are part of this initiative.

³ DataCite or Crossref <https://support.datacite.org/docs/datacite-or-crossref>

⁴ <https://blog.datacite.org/datacite-launches-global-access-program-with-support-from-czi/>

⁵ <https://wellcome.ac.uk/>

⁶ <https://www.jst.go.jp/EN/index.html>

⁷ <https://erc.europa.eu/>

⁸ <https://www.nih.gov/>

Open Researcher and Contributor ID — ORCID

Researcher IDs are nomenclatures that uniquely identify a researcher with their scholarly outputs. Common researcher IDs include Google Scholar, Scopus ID, and ORCID. In this roadmap, we focus on ORCID, which stands for Open Researcher and Contributor ID.

[ORCID](#) provides a free persistent digital identifier (an ORCID iD) that a person owns and controls, distinguishing them from every other researcher. Researchers can connect their iDs with their professional information—affiliations, grants, publications, peer reviews, etc. They can use their iDs to share their information with other systems, ensuring they get recognition for all their contributions, saving time and hassle, and reducing the risk of errors.

TENET operates a South African ORCID Consortium, allowing beneficiary institutions to realise cost savings in their ORCID membership and access local support ([TENET ORCID Consortium](#)).

ORCID's Global Participation Fund⁹ provided Community Development and Outreach grants to:

- Eko-Konnect Research and Education Initiative ([Eko-Konnect](#))
- Kenya Libraries and Information Services Consortium ([KLISC](#))
- [AfricArXiv](#)
- The Université Virtuelle De Côte D'Ivoire ([UVCI](#))

Research Organization Registry — ROR

[ROR](#) is a global, community-led registry of open persistent identifiers for research organisations. The ROR registry includes unique IDs and metadata for 100,000+ research organisations worldwide. ROR provides an open and community-driven solution to the problem of identifying affiliations in research and connecting them via open scholarly infrastructure to research outputs and researchers. ROR consolidates other identifiers currently in use (e.g., GRID, ISNI, Crossref Funder ID, Wikidata) at a coarser granularity.

Crossref, DataCite, California Digital Library and Digital Science joined forces to create the Research Organisation Registry (ROR).

Archival Resource Key — ARK

An Archival Resource Key ([ARK](#)) is a unique and persistent identifier that references information objects such as research papers, datasets, and software. It is widely employed by libraries, data centres, archives, museums, publishers, and government agencies to ensure reliable access to scholarly, scientific, and cultural resources. ARKs are assigned by the non-profit organisation [ARK Alliance](#), which manages the ARK registration system and provides tools and resources for users. They are well-suited for long-term preservation, offering flexibility, customisation, and granular identification. They excel in managing complex or evolving resources, accommodating versioning and supporting community-driven initiatives.

⁹ <https://info.orcid.org/global-participation-fund-first-round-awardees-announced/>

ARKs are interoperable with DOIs and can be used to link to DOIs and vice versa. This interoperability is made possible by using the Handle System and is beneficial for several reasons. It allows organisations to use the identifier that best meets their needs, regardless of the identified digital object type. It also will enable organisations to link to digital objects in other systems, regardless of the identifier system used by those systems. The interoperability of ARKs and DOIs is important in developing a global infrastructure for managing digital objects. It will make it easier for organisations to find and access digital objects, regardless of location.

One noteworthy advantage of ARKs is that they are free to register and use, making them a cost-effective choice for organisations of all sizes. By utilising ARKs, users can benefit from their unique, persistent, and interoperable nature while avoiding any financial burden associated with their adoption.

Stakeholders and Beneficiaries

Below is an overview of African stakeholders and their roles in the PID landscape:

- **RRENs and NRENs:** can perform the role of custodians of DOIs, be assigned a national DOI prefix or be a regional provider as in the case of the Ubuntunet Alliance. WACREN is also actively exploring the implementation of ARKs.
- **Universities and research centres:** develop institutional policies and strategies that encourage the adoption of ORCID and other PIDs, and define best practices in describing and citing research outputs.
- **Institutional librarians (with the institution's research office):** promote PIDs adoption, provide training, and embed PIDs in institutional services, systems and workflows.
- **Institutional and national repositories:** enable Handle Service and support the adoption of ORCID iDs, DOIs, and other PIDs into metadata records to facilitate the development of a linked metadata framework.
- **Journal editors:** adopt PIDs in journal publishing workflows, e.g., require authors to have an ORCID iD when submitting a journal article.
- **National education and research funding agencies:** embed PIDs in the research funding workflows on the submission and reporting stages, e.g. require researchers to have an ORCID iD when applying for grants; invest in platforms/features that allow funding awards to be pushed to recipients' ORCID records.
- **Researchers' roles:** register for ORCID iDs and complete ORCID profiles by adding research outputs; where possible, deposit datasets and other research outputs with repositories that provide PIDs and use them when citing datasets or research outputs.

Beneficiaries

- **Researchers** benefit from enhanced findability and efficient use for research information—more manageable and unambiguous citation; simplified funding

application; quicker submission of research outputs; easy demonstration of good research management; and simple and early data sharing.

- **Funders** (national education and research funding agencies) benefit from strengthened reproducibility; enhanced data and information provenance tracing and documentation; better and more efficient research data and metadata management and the lifecycle management and dissemination; better visibility and more diverse data and information use across domains; enhanced impact of funded research; easier tracking and impact assessment, and responsible research evaluation.
- **Institutions** benefit through the visibility of their institutional profile due to the discoverability of published materials assigned appropriate PIDs.
- **Conferences and proceedings** producing content can also benefit from using persistent identifiers.