



INTERNATIONAL
STANDARD
SERIAL
NUMBER

INTERNATIONAL CENTRE

Identifiers for Open Science: Recommendations for National Policymaking in the Caribbean

Dr. Gaelle BEQUET

Director

Contents

International identifiers and open science

Recommendations from the Knowledge Exchange group

Examples of national identifier policies

Identifying researchers: Options and Issues

International identifiers and Open Science

- A socio-technical artifact
- Central role of the International Organization for Standardization (ISO): standardization process; ISO/TS 22943:2022 on Identification Principles
- An identifier is a "sequence of characters that uniquely designates a referent. Identifiers can be used to specify the referent. In some cases, they can substitute for the referent or be used to retrieve the referent or its metadata."
- Evolution of identifiers : International Standard Book Number (1970), International Standard Serial Number (1975), International Standard Music Number (1993), International Standard Audiovisual Number (2002), Digital Object Identifier (2012) based on Handle technology, International Standard Name Identifier (2012), Research Activity Identifier (2022)

International identifiers and Open Science

- Between 2000 and 2019, the number of researchers in **OECD countries** rose from an average of 6 per 1,000 employees to 9 per 1,000.
- In **India**, the number of researchers rose from 110 to 255 per 1 million between 2000 and 2017.
- The **Universities Worldwide** website listed 6,864 universities in 178 countries in June 2004 and 9,826 in 209 countries in August 2022.
- **Budapest Open Access Initiative** (2001) advocates the creation of Open Access journals and institutional Open Science archives; publicly-funded research must be made available free of charge (and not for a fee in commercial publishers' journals).
- **Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities** (2003)
- **Recommendation on open science adopted at UNESCO's 41st General Conference in 2021**: the success of open science requires the establishment of **documentary infrastructures** including journals, articles, academic repositories, archives and scientific data, combined with bibliometric and scientometric tools, all of which rely in particular on **unique persistent identifiers** but also on the mobilization of a range of social groups (researchers, political decision-makers, entrepreneurs and citizens).

International identifiers and Open Science

- Identifiers enable the discoverability, identifiability and traceability of scientific results, products and activities throughout the research cycle.
- The FAIR principles are the guidelines for implementing open science: making scientific activities and entities findable, accessible, interoperable and reusable requires the use of identifiers.
- Identifiers can be used to interconnect databases and map open research on an international level: for example, [OpenAire](https://graph.openaire.eu/) <https://graph.openaire.eu/>
- The ISSN portal (<https://portal.issn.org>) uses the ISSN as a pivot identifier for aggregating journal data.

Recommendations from the Knowledge Exchange group : Research governance

- A1. Key stakeholders (research funders, national libraries, research organizations and infrastructures supporting interdisciplinary research) to discuss the national strategy for identifiers.
- A2. Define a PID strategy including all priority identifiers and explain their purposes (document identifiers, people, instruments, data sets) and how they interact with information systems.
- A3. Create governance instruments at national level, e.g. a consultative council for identifiers
- A4. Keep abreast of initiatives to create new PIDs and their socio-technical context
- A5. Make comparisons with other national identifier strategies, and participate in international forums dealing with this subject.
- A6. Communicate!

Recommendations from the Knowledge Exchange group : Research funding agencies

- B1. These agencies must be represented on the national working group on national-level identifiers or at least be informed of its work.
- B2. They must be familiar with the PIDs relevant to their activity, including the evaluation of project proposals, the results of funded research and the identification of grants.
- B3. Assign identifiers to the various grants wherever possible
- B4. Use specific identifiers for funded researchers
- B5. Keep abreast of developments concerning identifiers, which should bring added value to the fellowship management process.
- B6. Keep abreast of donor coordination initiatives at national and international level, promoting and associating with them wherever possible.

Recommendations from the Knowledge Exchange group : Identification service providers

C1. Ensure the sustainability of the initiative from a technical and economic point of view; a contingency plan must be put in place in the event of any interruption in service provision.

C2. Supplier business models must be clear and transparent

C3. Data and process documentation must be open and allow for community recovery in the event of failure.

C4/C5/C7. Study the possibility of participating in an international federation for the coordination and governance of identifiers that could be created in the future.

C6. Have mechanisms in place to report technical failures of identifiers

C8. Communicate with users of supplied identifiers (blog and webinars) and report regularly on progress made in implementing a specific identifier.

Recommendations from the Knowledge Exchange group : Universities and research institutes

D1. Be aware of national policy on identifiers

D2. Define corporate policy on identifiers, which should reflect national policy and specify specific features where necessary (local identifiers).

D3. Help researchers understand and use identifiers through training sessions

D4. Implement identifiers in research management information systems (CRIS) and institutional repositories.

D5. Keep abreast of technical identifiers that may emerge from research community initiatives

D6. Get to know the players, e.g. identification service providers and their specific features

Recommendations from the Knowledge Exchange group : Researchers

- E1. Get your login details! Register with ORCID and check your ISNI! Complete the associated metadata in full awareness of how it may be used
- E2. Respect the instructions of funders, publishers and institutions regarding identifiers.
- E3/E4. Monitor emerging identifiers in your field of research
- E5. Apply your institution's PID policy

Recommendations from the Knowledge Exchange group : Publishers

- F1. Ensure long-term availability of PID-identified publications through agreements with long-term archiving agencies and/or national libraries. Maintain the link with publications in the identification service provider's resolver.
- F2. Include entries for the most common identifiers in manuscript submission systems
- F3. Provide researchers/authors with some information on why identifiers are important
- F4. Know the level of maturity of initiatives in the field
- F5. Check that the identifiers in the publications are operational and resolve correctly
- F6. When available, consider including all identifiers for successive versions of an article on the publication's website.
- F7. Publishers of open access publications (Diamond model) should use scientific content identifiers as a minimum and keep abreast of initiatives to make identification more sustainable.

Examples of national identifier policies : JISC, UK

Jisc has selected five priority identifiers:

- Researchers (ORCID iDs)
- Search results (DOI from Crossref and DataCite)
- Grants (GrantId from Crossref)
- Organizations (ROR identifiers, an offshoot of Crossref)
- Projects (RAiD identifiers managed by the Australian Research Data Commons)

Examples of national identifier policies : SURF, NL

Relevant entities for the purpose of registration and reporting of research output

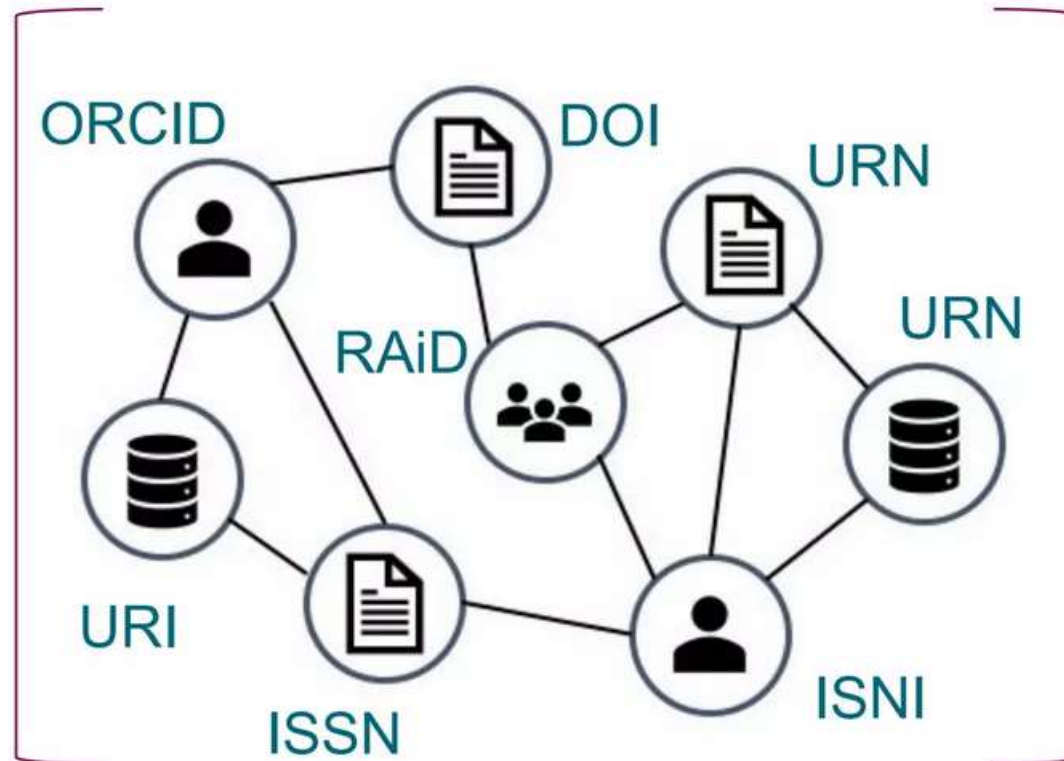
Relevant entities	Motivation (with examples)
Person	Person identifier is needed to connect research outputs with researchers. Integration into an organization's systems requires mature and well-functioning PID infrastructure. Example: ORCID
Organization	Organization identifier is needed to identify to which institution, faculty or department a certain research output belongs to (might be more than one). The system should support the complex and dynamic structures of research institutions. Examples: ROR, Ringgold/ISNI, FunderID (Crossref)
Grant (incl. funding conditions)	Grant identifier is needed to identify how research output is funded. The identification of funder- and journal policies may support the reporting on compliance (e.g. with Open Science policies). The value of these identifiers should be investigated. Examples: DOI GrantID (Crossref), EC Funding ID
Output (Publication, Data, Software, DMP)	Output identifiers are needed to identify, refer and locate research outputs. The outputs should link to the related persons, organizations and grants. Examples: DOI, URN, ISSN, Handle, Etc.
Other (like project)	Other identifiers may be used to support the use case of registration and reporting of research output. In particular: Project identifiers support researchers to document their the outputs during their project. It also helps organizations and funders to track which resources (people and infrastructure) and outputs are related to projects they fund or host. Example: RAiD

Examples of national identifier policies : SURF, NL

Relevant entities for the purpose of reproducibility and reusability of research output

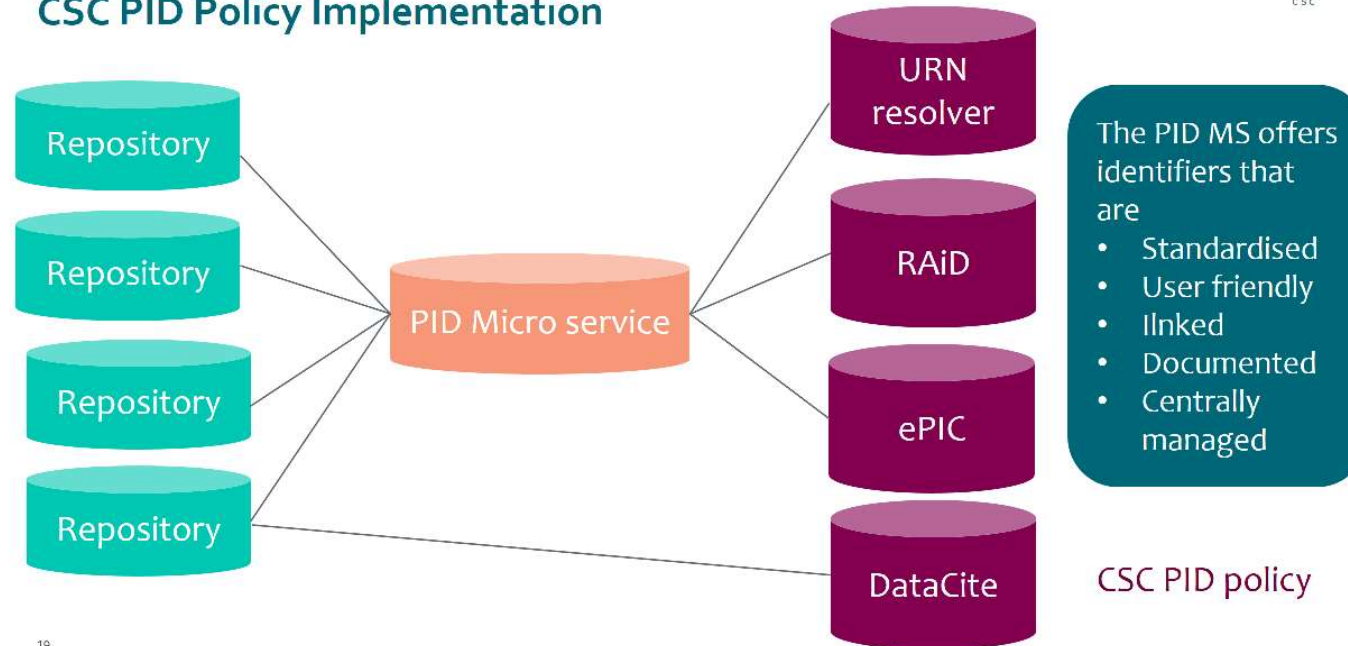
Relevant entities	Motivation (with examples)
Data set	Identifier for data set is needed to locate and identify the source of the reused data. Examples: DOI, Handle
Person	A person identifier is needed to identify to whom the reused data belongs to, and possibly how to access the data (license, embargo, etc.). Examples: ORCID, ISNI, DAI
organization	An organization identifier is needed to identify to which institution the reused data belongs to, and how to access the data (license, embargo, etc.). Examples: ISNI/Ringgold, ROR
Research software	Identifiers for research software (including e.g. scripts, models, analysis pipelines and software packages) are important for reproducing research results, including reference to the specific version of the software. Examples: DOI and SWHID
Instrument, device, sensor, platform, research facility (infrastructure)	Identifying which instruments, devices, etc. are used by the initial researcher(s) is needed to reproduce the research since there might be differences in its application. Internationally the development of such a registry is supported by EUDAT, DataCite and ePIC. Examples: DOI, Handle, RRID, UID
Method	Identifying the research method for transparency and trust to reproduce research. This is (generally) described in the publication. Example: unknown
Publication	Because the research method is (generally) described in the publication it is important to ensure findability of publications. Examples: DOI, Handle, URN:NBN
Sample	Depending on the discipline, samples are collected and digitally represented in research. The identifiers used for samples differ because the type of sample may vary between disciplines. Examples: IGSN, ARK, URN, HTTP URI (CETAF URI), DOI, UUID, RRID, BioSample accession number
Metadata schema	A metadata schema identifier is needed to foster machine readability of the datatypes. There are developments for a metadata schema registry at an European level. Example: unknown
Data type	Identifying the combination of key value units is needed for machine readability purposes. EPIC currently develops a data type registry via Handle.
Data format	Data formats are standardized, but not yet identifiable.

Examples of national identifier policies: CSC-IT, FI



Examples of national identifier policies: CSC-IT, FI

CSC PID Policy Implementation



Examples of national identifier policies : CAICYT-CONICET, AR

67% of Argentine journals are managed and published under OJS (Open Journal System)

Strong trend towards free and open systems

Only 5% of Argentinian journals charge an article processing fee (APC).

Journal prices are modest: between 1,000 and 4,000 Argentine pesos.

Only 37.5% of journals offer permanent identifiers for their articles, in particular the DOI (payment in dollars).

Implementation of an ARK resolver, free identifier, in 2018 (<http://id.caicyt.gov.ar/index.php>):

- Strategy for coexistence with other identifiers: coexistence and/or substitution and/or solution
- Consensus version of ARK including ISSN for unambiguous identification of continuous resources
- ARK-CAICYT considers the ISSN as the authoritative identifier, i.e. NAAN ;

[http://id.conicet.gov.ar/ark:/**ISSN**/Name\[Qualifier\]](http://id.conicet.gov.ar/ark:/**ISSN**/Name[Qualifier])

Examples of national identifier policies : CAICYT-CONICET, AR

EJEMPLO ARK-CAICYT

Bergero, P., Pedrosa, J., & Petrucci, D. (2021). Epidemias, salud y ciencia: interés y percepción de jóvenes de escuelas públicas del conurbano bonaerense antes del COVID-19. Revista De Educación En Biología, 24(2), 66–79.
<http://id.caicyt.gov.ar/ark:/s23449225/r7ewizkmi>

- NAA: <http://id.caicyt.gov.ar>
- NAAN: <ark:/s23449225/>
- Name: r7ewizkmi

Identifying researchers: Options and issues

A brief socio-technical history of ORCID

ORCID and ISNI: similar but different

Towards the "platformization" of research?

Identifying researchers: Options and issues

- ORCID stands for Open Researcher and Contributor ID.
- **December 2009:** announcement in the journal Nature of the creation of an "author ID system backed by 23 organizations, including Thomson Reuters, Nature Publishing Group, Elsevier, ProQuest, Springer, CrossRef, the British Library and the Wellcome Trust."
- **2010:** ORCID Inc. founded in the USA with the aim of creating a central registry of unique identifiers for active researchers in order to foster academic communication and research; Thomson Reuters gives a copy of the code of its identifier system, namely Researcher ID used by the Web of Science database.
- **2012:** creation of the ORCID Registry
- **2013:** ORCID Inc. becomes a non-profit organization
- ORCID = Hybrid system of self-affirmed identity and identity confirmed by a research organization; data provided by individuals and research organizations.

Identifying researchers: Options and issues

- **ISNI** is a precursor to ORCID: standardization work began at ISO in **October 2006**.
- The ISO working group brings together experts mandated by the International Confederation of Societies of Authors and Composers (CISAC), the Bibliothèque nationale de France, the Federation of Italian Publishers (AEI), Bowker/ProQuest, OCLC, Nielsen Book Data, the British Library, the Conference of European National Libraries (CENL), International DOI Foundation and ISSN International Centre, ISAN International Agency, Société civile pour l'administration des droits des artistes et musiciens interprètes, ISBN International Agency, Swiss Society for the Rights of Authors of Musical Works (SUISA), and the International Federation of Reproduction Rights Organisations (IFFRO).
- **2012**: creation of the ISNI registration authority; OCLC (Leiden) is the operator of the identifier allocation system.
- The ORCID identifier is a URI starting with https with a 16-digit number compatible with the ISO standard (ISO 27729), also known as the International Standard Name Identifier (ISNI), for example <https://orcid.org/0000-0001-2345-6789>.
- **2014**: an agreement is signed between ISNI and ORCID; ORCID positions itself as an open, interdisciplinary identifier created by the research community; ISNI is presented as an identifier with an institutional character because it is created by ISO and supported by national libraries and international rights management organizations; part of the identifiers standardized by ISO is reserved for ORCID.

Identifying researchers: Options and issues

- **ORCID** received support from the **European Commission** through the THOR (2015-2017) and FREYA (2017-2020) projects, which led to its inclusion in the construction of the European Open Science Cloud (EOSC).
- **ORCID** aims to become a pivotal identifier linking researchers to their open access output (thanks in particular to interoperability with CrossRef and DataCite DOIs), whether published by commercial publishers (gold open access) or made available on academic repositories (green open access).
- ORCID has enlisted in its network organizations such as **commercial publishers** with the ability to compel researchers to apply for an ORCID: is this an impediment to the freedom of research defined in the Universal Declaration of Human Rights (article 27)?
- ORCID is associated with **personal metadata** that is subject in the European Union to specific regulations, namely the **General Data Protection Regulation** (GDPR). In the event of a data dispute, the legal action would take place in the USA. Could this be damaging to the interests of researchers and research organizations?
- What about agreements between a university and its researchers to share personal information with ORCID? **What about the consent to share personal data and the right to be forgotten?**

Identifying researchers: Options and issues

- Open science = subversion of traditional science (scientific results questioned, low productivity of researchers, retraction of articles, low involvement of citizens, peer review questioned)
- Open science = online content and service platforms, combining preprint publication with post-publication peer review and the use of new bibliometric indicators (altmetrics).
- The production of value no longer resides solely in the publication of research results, but in the ability to control this production from end to end, and to create data analysis tools that enable research policies to be adjusted: see for example, the PURE research management system marketed by Elsevier.
- An identifier like ORCID can be used for big data analysis to monitor scientific research worldwide.

Conclusion

- Using identifiers in the search process - yes, but :
- Preserve the diversity of identifiers by combining national and international identifiers to maintain autonomy in the event of difficulties with the identifier supplier.
- Choose PID suppliers with genuine expertise in data curation
- Beware of internationally shared metadata, especially those concerning individuals!
- Beware of the costs!
- National research organisations need to get involved in the governance of international identifiers



INTERNATIONAL
STANDARD
SERIAL
NUMBER

INTERNATIONAL CENTRE

**Thank you for your attention.
Any questions? Any comments?**

gaelle.bequet@issn.org