DiSSCo related output

This template collects the required metadata to reference the official Deliverables and Milestones of DiSSCo-related projects. More information on the mandatory and conditionally mandatory fields can be found in the supporting document 'Metadata for DiSSCo Knowledge base' that is shared among work package leads, and in Teamwork > Files. A short explanatory text is given for all metadata fields, thus allowing easy entry of the required information. If there are any questions, please contact us at info@dissco.eu.

Title
DiSSCo Prepare Milestone report MS 5.3 "Documentation of PIDs relevant for DiSSCo technical infrastructure"

Author(s)	**Identifier of the author(s)**
Sabine von Mering	https://orcid.org/0000-0003-2982-7792 (SvM)
Julia Pim Reis	https://orcid.org/0000-0002-5357-6148 (JPR)
Falko Glöckler	https://orcid.org/0000-0002-7127-2738 (FG)
Wouter Addink	https://orcid.org/0000-0002-3090-1761 (WA)
Robert Cubey	https://orcid.org/0000-0001-7902-3843 (RC)
Mathias Dillen	https://orcid.org/0000-0002-3973-1252 (MD)
Anton Güntsch	https://orcid.org/0000-0002-4325-4030 (AG)
Elspeth Haston	https://orcid.org/0000-0001-9144-2848 (EH)
Sharif Islam	https://orcid.org/0000-0001-8050-0299 (SI)
Mareike Petersen	https://orcid.org/0000-0001-8666-1931 (MP)

Affiliation
Museum für Naturkunde - Leibniz Institute for Evolution and Biodiversity Science

Contributors
Kessy Abarenkov https://orcid.org/0000-0001-5526-4845
David Fichtmueller https://orcid.org/0000-0002-0829-5849
Alex Hardisty https://orcid.org/0000-0002-0767-4310
Claus Weiland https://orcid.org/0000-0003-0351-6523
Matt Woodburn https://orcid.org/0000-0001-6496-1423

Publisher
DiSSCo Prepare

Identifier of the publisher

Resource ID
https://doi.org/10.34960/1xr6-hr45

Publication year
2021

Related identifiers

Is it the first time you submit this outcome?
Yes

Creation date
01/11/2021

Version
1

Citation
von Mering, S. et al. (2021): DiSSCo Prepare Milestone report MS 5.3 "Documentation of PIDs relevant for

DiSSCo technical infrastructure". DiSSCo Prepare. https://doi.org/10.34960/1xr6-hr45

Abstract

The DiSSCo Prepare Milestone report MS5.3 "Documentation of PIDs relevant for DiSSCo technical infrastructure" compiles information on different persistent identifiers and discusses its relevance to DiSSCo. Different alternative Handle-based PID schemes are discussed. While not all PIDs will be directly used by the FAIR (findable, accessible, interoperable, reusable) data architecture of the DiSSCo research infrastructure, it is important to be aware of other developments for interoperability reasons.

Content keywords
scientific

Project reference
DiSSCo Prepare (GA-871043)

WP number
WP5

Project output
Milestone report

Deliverable/milestone number
MS5.3

Dissemination level
Public

Rights

License
Attribution 4.0 International (CC BY 4.0)

Resource type
Text

Format
PDF

Funding Programme
H2020-INFRADEV-2019-2

Contact email
sabine.vonmering@mfn.berlin

DiSSCo Prepare WP5 – Milestone report

MS5.3 Documentation of PIDs relevant for DiSSCo technical infrastructure

Work package lead: Mareike Petersen (MfN)

Authors: Sabine von Mering (MfN), Julia Pim Reis (MfN), Falko Glöckler (MfN), Wouter Addink (Naturalis), Robert Cubey (RBGE), Mathias Dillen (MeiseBG), Anton Güntsch (BGBM), Elspeth Haston (RBGE), Sharif Islam (Naturalis), Mareike Petersen (MfN)

Contributors: Kessy Abarenkov (U Tartu), David Fichtmueller (BGBM), Alex Hardisty (U Cardiff), Claus Weiland (Senckenberg), Matt Woodburn (NHM)

MUSEUM FÜR
NATURKUNDE
BERLIN

Abstract

The DiSSCo Prepare Milestone report MS5.3 "Documentation of PIDs relevant for DiSSCo technical infrastructure" compiles information on different persistent identifiers and discusses its relevance to DiSSCo. Different alternative Handle-based PID schemes are discussed. While not all PIDs will be directly used by the FAIR (findable, accessible, interoperable, reusable) data architecture of the DiSSCo research infrastructure, it is important to be aware of other developments for interoperability reasons.

Keywords

PID, PID schemes

INDEX

1. Introduction & background

The Distributed System of Scientific Collections ([DiSSCo](#)) is a new Research Infrastructure (RI) of European natural science collections (NSC) currently in its Preparatory phase ([DiSSCo Prepare](#)). The DiSSCo RI aims to create a new business model for one European collection that digitally unifies all European natural science assets under common access, curation, policies and practices that ensure that all the data is easily Findable, Accessible, Interoperable and Reusable (FAIR principles; Wilkinson et al. 2016).

Persistent identifiers (PIDs) are an essential element of global data infrastructures and fundamental for the digital transformation of collections-based science. They facilitate unambiguous citation and tracking of physical samples thus allowing linking of specimens, data and publications; and serve as identifiers but also connectors. Such connections can be recognized by both machines and humans (machine- and human-readable), which reveals and gives access to a wide range of associated information, ensuring that relationships can be understood, knowledge gained and conclusions to be reached.

Acting as a long-lasting reference to a digital entity or resource, PIDs are used to uniquely and unambiguously identify digital representations of natural science objects. PIDs are identifiers that are **globally unique**, **resolvable** (i.e. can be expressed as an URI which can take users or machines to a resource or information about a resource) and they are actively managed so that they remain persistent in the long term. Therefore, PIDs play an important role in digital preservation of data.

Persistent identifiers were originally developed to address challenges arising from the distributed and disorganised nature of the internet and the so-called "link rot" that made it difficult to maintain a persistent record or digital resources including research data (see Klump & Huber 2017 and references within). PIDs have been in use for over 20 years and there is ongoing discussion on which PID schemes to use in a given community.

There are several types, categories and levels to broadly group PIDs (see Meadows et al. 2019):
- identifiers for researchers, organizations, and research objects and outputs;
- open, i.e. fully interoperable vs. proprietary, i.e. for use within a single organization;
- local to an individual organization, national or global.

A highly desirable quality for PIDs is to have FAIR metadata. As well as significantly reducing the risk of reference rot, this enables the discovery of open, interoperable, well-defined (FAIR) metadata containing provenance information in a predictable manner – and the PIDs themselves are also open. DOIs are a good example.

For DiSSCo's envisaged FAIR Digital Object (FDO) infrastructure, PIDs for the digital objects should be based on the [Handle system](#) to be compliant with the Digital Object Architecture as described by the DONA foundation and CORDRA as its reference implementation. There are currently eight PID schemes using the Handle system which could potentially be used as a PID scheme for a Digital Specimen, these have been described in Hardisty et al. (2021):
- Digital Object Identifier ([DOI](#))
- International GeoSample Number ([IGSN](#))
- European PID Consortium ([ePIC](#))

- Five-digit prefix ([CNRI](#))
- Second-level prefix
- Three-segment prefix
- Two-digit top level prefix
- National-level services

A PID scheme relates not only to the technical elements but the whole arrangement around PIDs for using and operating them including the ownership, authority, governance and financial elements. The comparison of the different PID schemes resulted in DOI as the preferred option for Digital Specimen (with a tailored metadata schema; Hardisty et al. 2021).

PIDs can uniquely link physical objects to digital artefacts, to records of transactions, the identification of specific vocabulary terms and concepts, etc. Different PID schemes can be used for identifying different things - DOIs for documents and datasets, for example; ORCiD for persons, ROR for organisations. These are described in more detail below.

DiSSCo, which is planned to commence full operations in 2026, will have services for indexing, enriching and assisting reuse of specimen data, and needs PIDs and PID services
- to support the ambition for [Digital Specimens](#), virtual collections, workflows, etc. on the Internet;
- for loans and visits like implemented in [ELViS](#), for annotations, citations, attribution of work and microcredits;
- to pursue aims of common policies and procedures; and to transform work practices.

PIDs for Digital Specimens complement identifiers of the physical specimens themselves and/or their corresponding digital database records in institutional collection management systems (Hardisty et al. 2021). Examples of such identifiers include the CETAF Stable Identifiers (Güntsch et al. 2017), the International Geo Sample Numbers (IGSN; Lehnert et al. 2019), GUIDs (Globally Unique Identifier), Darwin Core Triplets (institutionCode:collectionCode:catalogNumber, https://dwc.tdwg.org/rdf/), or any other combination of institution/collection codes and catalog numbers.

Community involvement was and is crucial to reach a broad consensus related to the future adoption of certain PID schemes. This has been done via online consultations and discussion forums. A consultation on Digital Specimens Persistent Identifiers (PIDs) for the operation of the DiSSCo RI took place in October 2020 (https://www.dissco.eu/dissco-pid-consultation/). Another global and virtual consultation hosted by GBIF under the umbrella of the [Alliance for Biodiversity Knowledge](#) has taken place in 2021. In Topic 7 of this community consultation, [Persistent identifier (PID) schemes](#) have been discussed.

The discussion on technical convergence of DiSSCo's Digital Specimen concept and the similar concept from the Extended Specimen Network strategy of the Biodiversity Collections Network (BCoN) in the USA (BCoN 2019, Lendemer et al. 2019) is expected to reach consensus on the new term 'Digital Extended Specimen' (DES) circumscribing the Digital Specimen and Extended Specimen ideas in one technical concept.

Terms and acronyms related to and relevant for the DiSSCo infrastructure are described in the **DiSSCo Knowledgebase** [Glossary](#).

2. PIDs relevant for DiSSCo technical infrastructure

2.1 Identifiers for metadata

2.1.1 Identifier for people (researchers and other agents)

Name	**Open Researcher and Contributor ID (ORCID iD)**
Focus	ORCID provides a persistent digital identifier (an alphanumeric code called ORCID iD), to uniquely identify scientific and other academic authors and contributors.
Further reading	https://orcid.org/ https://github.com/ORCID
Use Cases	Widely used. There are approximately 1235 ORCID member organizations.
Example ID	https://orcid.org/0000-0002-0767-4310 (Alex Hardisty)

Name	**Wikidata Q number (QID)**
Focus	Wikidata makes use of identifiers for both internal organization of the knowledge base and for its connection to other databases. Wikidata is also a hub/broker for other identifiers.
Further reading	https://www.wikidata.org/wiki/Q43649390 https://www.wikidata.org/wiki/Wikidata:Identifiers
Use Cases	Wikidata provides Q numbers for items on people (all those featured in Wikipedia and many more).
Example ID	https://www.wikidata.org/wiki/Q63764 (Louisa Bolus) https://www.wikidata.org/wiki/Q6694 (Alexander von Humboldt)

Name	**International Standard Name Identifier (ISNI)**
Focus	ISNI is an ISO certified global standard number especially for contributors to creative works and those active in their distribution, including researchers, inventors, writers, artists, visual creators, performers, producers, publishers, aggregators, and more. The focus is to assign to the public name(s) of those persons a persistent unique identifying number in order to resolve the problem of name ambiguity in search and discovery.

	ISNI aims to act as a bridge identifier across multiple domains and is becoming a component in Linked Data and Semantic Web applications.
Further reading	https://isni.org/
Use Cases	ISNI holds public records of over 12.75 million individuals (of which 2.94 million are researchers) and of 1,588,535 organizations.
Example ID	https://isni.org/isni/0000000032197769 (Amalie Dietrich) https://isni.org/isni/0000000121013124 (Aimé Bonpland)

Name	**Virtual International Authority File identifier (VIAF ID)**
Focus	VIAF is an international authority file that combines several authority files in an authority data service. It is a joint project of several national libraries and operated by the Online Computer Library Center (OCLC).
Further reading	http://viaf.org/
Use Cases	VIAF identifiers are widely used in library catalogues but also added to biographical articles on Wikipedia and incorporated in Wikidata.
Example ID	https://viaf.org/viaf/98043389/ (Ernst Mayr)

Name	**International Plant Name Index (IPNI) ID**
Focus	The International Plant Names Index (IPNI) is a nomenclatural index of names of vascular plants from Family down to infraspecific ranks. IPNI IDs are also provided for botanical authors. The older index of authors of plant scientific names is incorporated in IPNI.
Further reading	https://ipni.org/about
Use Cases	International Plant Names Index (IPNI) https://www.ipni.org/ Plants of the World Online (POWO) http://www.plantsoftheworldonline.org/ World Flora Online (WFO) Portal http://www.worldfloraonline.org/

Example ID	https://www.ipni.org/a/12653-1/ (Carl Linneaus)

Name	**Authors of plant names**
Focus	An index of authors of plant scientific names. Includes flowering plants, gymnosperms, pteridophytes, bryophytes, algae, fungi and fossil plants. Full names, dates of birth and death when known, recommended abbreviations and groups in which names have been published, are given for each author. Since 1998, the database of author names and their standard forms has been maintained and updated online as part of The International Plant Names Index (http://www.ipni.org/).
Further reading	http://www.kew.org/data/authors.html https://github.com/tdwg/prior-standards/tree/master/authors-of-plant-names Brummitt & Powell (1992), http://rs.tdwg.org/apn/doc/data/1992
Use Cases	http://www.ipni.org/
Example ID	https://www.ipni.org/a/2432-1

Name	**ZooBank authors**
Focus	ZooBank, the Official Registry of Zoological Nomenclature, provides UUIDs for currently 83,000+ authors, i.e. anyone who (co-)authored a publication containing nomenclatural acts (e.g. 'original descriptions' of new scientific names for animals, emendations, and lectotypifications) or is contributing to ZooBank content.
Further reading	http://zoobank.org/About
Use Cases	http://zoobank.org/
Example ID	http://zoobank.org/Authors/8C466CBE-3F7D-4DC9-8CBD-26DD3F57E212

2.1.2 Identifier for (research) organizations and their subunits

Name	**Research Organization Registry (ROR) identifier**
Focus	ROR is a community-led project to develop an open, sustainable, usable, and unique identifier for research organizations of the world. ROR's scope is the affiliation use case and connecting organizations to research outputs. ROR is interoperable with other identifiers and ROR IDs map to GRID, ISNI, Crossref Funder ID, and Wikidata. ROR records include basic metadata about an organization to help with discoverability and disambiguation (including alternate organization names, URL, location, and relationships to other organizations). ROR can support multiple languages and character sets. All ROR IDs and metadata are provided under the Creative Commons CC0 1.0 Universal Public Domain Dedication.
Further reading	https://ror.org/ https://github.com/ror-community
Use Cases	Focused specifically on the affiliation use case, and widely applied, e.g. in DataCite, Crossref and ORCID datasets.
Example ID	https://ror.org/039zvsn29 (NHM London) https://ror.org/033xtdz52 (Botanic Garden Rio de Janeiro)

Name	**Wikidata Q number (QID)**
Focus	Wikidata makes use of identifiers for both internal organization of the knowledge base and for its connection to other databases. Wikidata is also a hub/broker for other identifiers.
Further reading	https://www.wikidata.org/wiki/Q43649390 https://www.wikidata.org/wiki/Wikidata:Identifiers
Use Cases	Wikidata also provides QIDs for items on research organizations.
Example ID	https://www.wikidata.org/wiki/Q838691 (MNHN Paris)

Name	**Ringgold Identifier (RIN)**
Focus	Ringgold Identifiers (Ringgold IDs or RINs) are persistent numeric unique identifiers for organizations in the (scholarly) publishing industry supply chain. Ringgold IDs link to ISNI, IPEDS, NCES, Open Funder Registry IDs. Working to link identifiers with institutional identifiers (e.g. ORCID); ORCID also uses Ringgold IDs.
Further reading	https://www.ringgold.com/ringgold-identifier/
Use Cases	Ringgold's Identify Database contains almost 600,000 organizations in all countries and sectors including academia, corporations, hospitals, and government entities. (only accessible with an account)
Example ID	Ringgold ID 1438 (University of California Berkeley) https://www.ringgold.com/identify-sample/

Name	**GBIF Registry of Scientific Collections (GRSciColl)**
Focus	GRSciColl is a comprehensive, community-curated clearinghouse of collections information originally developed by Consortium of the Barcode of Life (CBOL). The collections registry includes data about the world's scientific institutions, collections and associated staff members. It spans all scientific disciplines, including earth and space sciences, anthropology, archaeology, biology and biomedicine, as well as applied fields like agriculture, veterinary medicine and technology.
Further reading	https://www.gbif.org/grscicoll https://github.com/gbif/registry Grosjean et al. (2021), https://doi.org/10.3897/biss.5.74354
Use Cases	GBIF is working with partners to establish a standardized, interoperable flow of information between this registry and other key collections infrastructures. GRSciColl runs a weekly synchronization with Index Herbariorum (IH). Following content migration from iDigBio Collections, GRSciColl shares an integrated registry delivering consistent information with shared editing access to users of both GBIF.org and iDigBio.org.
Example ID	https://www.gbif.org/grscicoll/institution/b9684b7b-2550-47a0-93c5-4bb6fc28cea5

2.1.3 Identifier for grant-giving organizations

Name	**Crossref Funder Registry (Crossref Funder ID)**
Focus	The Crossref Funder Registry is an open and unique registry of persistent identifiers for grant-giving organizations from around the world. The Funder Registry and associated funding metadata allows everyone to have transparency into research funding and its outcomes. It provides grant-giving organization names (funder names), identifiers (funder IDs), and associated grant numbers thus building connections between funders and research outputs. Funding data is used by funders to track the publications resulting from their grants, including use of facilities, equipment, salary awards, etc. The list is available to download as an RDF file, and is freely available under a CC0 license waiver.
Further reading	https://www.crossref.org/services/funder-registry/ https://www.crossref.org/documentation/content-registration/administrative-metadata/funding-information/ https://github.com/CrossRef/open-funder-registry
Use Cases	Used with Open Ukrainian Citation Index (OUCI)
Example ID	https://search.crossref.org/funding?q=501100001659

2.1.4 Identifier for taxa

Name	**Catalogue of Life (COL) Stable identifier**
Focus	CoL has an algorithm to generate stable identifiers for name usage.
Further reading	https://www.catalogueoflife.org/2021/04/14/stable-ids https://www.catalogueoflife.org/ https://data.catalogueoflife.org/
Use Cases	Uses LSIDs as a means of identifying species and higher taxa in these large species catalogue.
Example ID	https://www.catalogueoflife.org/data/taxon/BTB (Lauraceae Juss.)

Name	**International Plant Name Index (IPNI) ID**
Focus	The International Plant Names Index (IPNI) is a nomenclatural index of names of vascular plants from Family down to infraspecific ranks. IPNI IDs are available for taxon names as well as for author names and publications.
Further reading	https://ipni.org/about
Use Cases	International Plant Names Index (IPNI) https://www.ipni.org/ Plants of the World Online (POWO) http://www.plantsoftheworldonline.org/ World Flora Online (WFO) Portal http://www.worldfloraonline.org/
Example ID	https://www.ipni.org/n/30117681-2/ (*Fragaria × ananassa*)

Name	**Index Fungorum Registration Identifier**
Focus	The Index Fungorum is a global fungal nomenclator that contains names of fungi (including yeasts, lichens, chromistan fungal analogues, protozoan fungal analogues and fossil forms) at all ranks. Index Fungorum uses Registration IDs and UUIDs.
Further reading	http://www.indexfungorum.org/
Use Cases	http://www.indexfungorum.org/
Example ID	http://www.indexfungorum.org/names/NamesRecord.asp?RecordID=550550

Name	**World Flora Online (WFO-ID)**
Focus	WFO-ID is a unique, immutable identifier for each name in the Backbone, regardless of its taxonomic and nomenclatural status. For machine readability, Representational State Transfer (REST) services are available. WFO-IDs are resolvable as URIs with the prefix "http://www.worldfloraonline.org/taxon/".
Further reading	http://www.worldfloraonline.org/

	Borsch et al. (2020), https://doi.org/10.1002/tax.12373
Use Cases	World Flora Online (WFO) Portal http://www.worldfloraonline.org/
Example ID	http://www.worldfloraonline.org/taxon/wfo-0000173973 (*Caesalpinia echinata* Lam.)

Name	**ZooBank LSID**
Focus	ZooBank is a central, authoritative and comprehensive resource for scientific names in zoology. It provides an authoritative online, open-access, community-generated registry for zoological nomenclature (to register new nomenclatural acts, published works, and authors), as a service to taxonomists, biologists, and the global biodiversity informatics community. Life Science Identifiers (LSIDs) are used as globally unique identifiers for ZooBank registration entries.
Further reading	http://zoobank.org/
Use Cases	Official Register of the International Commission on Zoological Nomenclature (ICZN).
Example ID	http://zoobank.org/NomenclaturalActs/6EA8BB2A-A57B-47C1-953E-042D8CD8E0E2 (*Pseudanthias carlsoni*)

Name	**AlgaeBase PIDs**
Focus	AlgaeBase is a database of information on algae that includes terrestrial, marine and freshwater organisms. For convenience, we have included the sea-grasses, even though they are flowering plants.
Further reading	https://www.algaebase.org/about/
Use Cases	https://www.algaebase.org/
Example ID	https://www.algaebase.org/search/species/detail/?species_id=33389

Name	Integrated Taxonomic Information System (ITIS) TSN
Focus	ITIS provides authoritative taxonomic information on plants, animals, fungi, and microbes of North America and the world. ITIS uses a taxonomic serial number (TSN) system.
Further reading	https://itis.gov/pdf/faq_itis_tsn.pdf
Use Cases	https://itis.gov/
Example ID	https://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=28759#null

2.1.5 Identifier for localities, geographical names and sites

Name	GeoNames
Focus	GeoNames is an open geographical database that contains over 27 million geographical names and consists of over 12 million unique features. Each GeoNames feature is represented as a web resource identified by a stable URI, links to a HTML Wiki page or provides RDF.
Further reading	https://www.geonames.org/about.html
Use Cases	http://www.geonames.org/
Example ID	https://www.geonames.org/2950159/berlin.html (Berlin)

Name	NGA GeoNames
Focus	National Geospatial intelligence Agency (NGA) GeoNames Search is a database that provides geographic names for the guidance of and use by the Federal Government and for the information of the general public. Geographic names have a Unique Feature Identifier (UFI).

Further reading	https://geonames.nga.mil/gns/html/
Use Cases	https://geonames.nga.mil/namesgaz/
Example ID	

Name	**ISO 3166 standard for country codes**
Focus	The "ISO 3166 standard – Codes for the representation of names of countries and their subdivisions" was created and is maintained by The International Organization for Standardization (ISO).
Further reading	https://www.iso.org/iso-3166-country-codes.html https://en.wikipedia.org/wiki/List_of_ISO_3166_country_codes
Use Cases	GBIF https://rs.gbif.org/areas/ IBAN https://www.iban.com/country-codes
Example ID	https://www.iso.org/obp/ui/#iso:code:3166:PT

Name	**Spatial Reference System Identifier (SRID)**
Focus	A SRID is a unique value used to unambiguously identify projected, unprojected, and local spatial coordinate system definitions used by all GIS (geographic information system) applications.
Further reading	https://desktop.arcgis.com/en/arcmap/10.3/manage-data/using-sql-with-gdbs/what-is-an-srid.htm
Use Cases	SRID implementations exist from many different spatial vendors. The EPSG Geodetic Parameter Dataset (or EPSG registry) is one example.
Example ID	

Name	**Wikidata Q identifier (Wikidata QID)**
Focus	Number with a prefix "q" identifying Wikidata entities.

Further reading	https://www.wikidata.org/wiki/Wikidata:Identifiers
Use Cases	Wikidata items for geographical names and entities
Example ID	https://www.wikidata.org/wiki/Q568396 (lake Krumme Lanke in Berlin)

Name	**Getty Thesaurus of Geographic Names® Online (TGN)**
Focus	The TGN is an evolving vocabulary, thousands of TGN place names are added and edited every year. Types of places included in TGN are inhabited places (cities, towns, villages), nations, empires, archaeological sites, named general areas, tribal areas, lost settlements (historically documented, but the precise location is unknown), and physical features. Each record (place concept), name, and much other information in TGN are identified by persistent, unique numeric identifiers.
Further reading	https://www.getty.edu/research/tools/vocabularies/tgn/faq.html
Use Cases	https://www.getty.edu/research/tools/vocabularies/tgn/index.html
Example ID	http://www.getty.edu/vow/TGNFullDisplay?find=Paris&place=&nation=&prev_page=1&english=Y&subjectid=7002980

Name	**Dynamic Ecological Information Management System - Site and dataset registry (DEIMS-SDR)**
Focus	The aim of DEIMS-SDR is to be the globally most comprehensive catalogue of environmental research and monitoring facilities, featuring foremost but not exclusively information about all LTER sites on the globe and providing that information to science, politics and the public in general.
Further reading	https://deims.org/
Use Cases	https://www.re3data.org/
Example ID	https://deims.org/049de4d9-d7db-4b2c-ace5-de8873f5d277

Name	**GADM maps and data**
Focus	GADM provides maps of the administrative areas of all countries, at all levels of sub-division. They provide data at high spatial resolutions that include an extensive set of attributes. They have UIDs - may not be considered as PIDs
Further reading	https://gadm.org/formats.html https://gadm.org/about.html
Use Cases	https://gadm.org/
Example ID	

2.2 Identifier for physical objects (collection items, specimens and samples)

Name	**Natural Science Identifier (NSId)**
Focus	A Natural Science Identifier (NSId) is a universal, unique persistent identifier for digitised natural science specimens (i.e., Digital Specimens) and other associated object types. An NSId will help you unambiguously refer to a specimen you are working with or will help to find a specimen that someone else has told you about by giving you the NSId e.g., as a reference in a journal article. The best DOIs (and other kinds of Handle, including NSId) are opaque ones that carry no information that could potentially become out of date and incorrect.
Further reading	https://dissco.tech/2020/05/28/natural-science-identifiers-cetaf-stable-identifiers/ https://pidforum.org/t/a-global-natural-sciences-identifier-nsid-scheme-for-specimens-and-collections/860
Use Cases	
Example ID	

Name	**CETAF Stable Identifier (CSI)**

Focus	CETAF stable identifiers provide human- and machine-readable access to specimen information.
Further reading	https://cetaf.org/resources/best-practices/cetaf-stable-identifiers-csi-2/ https://cetafidentifiers.biowikifarm.net/ Güntsch et al. (2017), https://doi.org/10.1093/database/bax003
Use Cases	CETAF Botany Pilot https://services.bgbm.org/botanypilot/ CETAF Stable identifiers have been implemented by various CETAF institutions as well as other partners (see https://know.dissco.eu/handle/item/214).
Example IDs	http://herbarium.bgbm.org/object/B100277113 https://data.rbge.org.uk/herb/E00421509 https://www.botanicalcollections.be/specimen/BR0000005516339

Name	**Digital Object Identifier (DOI)**
Focus	The DOI system provides a technical and social infrastructure for the registration and use of persistent interoperable identifiers, called DOIs, for use on digital networks. A DOI is a persistent identifier or handle used to identify objects uniquely, standardized by the International Organization for Standardization (ISO). DOIs are resolvable and interoperable. DOIs are an implementation of the Handle system.
Further reading	https://www.doi.org/ https://www.doi.org/factsheets/Identifier_Interoper.html
Use Cases	DOIs are widely used mainly to identify academic, professional, and government information, such as journal articles, research reports, data sets, and official publications. However, they have also been used for other types of information resources.
Example ID	

Name	**Archival Resource Key (ARK)**

Focus	ARKs are open, mainstream, non-paywalled, decentralized PIDs that identify anything digital, physical, or abstract. The ARK Alliance is an open global community supporting the ARK infrastructure on behalf of research and scholarship. ARKs are being assigned to a variety of different information resources including museum specimens, digitized documents and objects, historic maps, publisher content, genealogical records, scientific records, datasets, journals, etc.
Further reading	https://arks.org/about/ https://wiki.lyrasis.org/display/ARKs/ARK+Identifiers+FAQ
Use Cases	Since 2001 over 850 organizations across the world have registered and created some 8.2 billion ARKs. The registry includes national and university libraries and archives, art museums, natural history museums, publishers, data centers, government agencies, vendors, and research labs.
Example ID	http://ark.bnf.fr/ark:/12148/btv1b8449691v/f29

Name	**International Geo Sample Number (IGSN)**
Focus	The core purpose of IGSN is to enable transparent and traceable connections between research activities and objects, including samples, collections, instruments, grants, data, publications, people and organizations.
Further reading	https://www.igsn.org/ Lehnert et al. (2019), https://doi.org/10.3897/biss.3.37334 Buys & Lehnert (2021), https://doi.org/10.5438/thhf-kx17
Use Cases	Operates a central registration system for physical samples. Recent partnership with DataCite that intends to support the global adoption, implementation, and use of physical sample identifiers.
Example ID	http://pid.geoscience.gov.au/sample/AU1243

Name	**AAT: Art & Architecture Thesaurus**

Focus	The Art & Architecture Thesaurus (AAT) is a controlled vocabulary used to describe and improve access to information about items of art, architecture, and other material culture. The AAT thesaurus is in compliance with ISO and NISO standards. It is a structured vocabulary of 55,000+ concepts, including terms, descriptions, bibliographic citations, and other information relating to fine art, architecture, decorative arts, archival materials, and material culture. A minimum record in AAT contains a numeric ID, a term, and a position in the hierarchy. AAT could also be used to express that an object in a collection is a "natural" object that occurs in nature and is not made by humans.
Further reading	http://www.getty.edu/research/tools/vocabularies/aat/about.html http://www.getty.edu/vow/AATFullDisplay?find=natural+object&logic=AND&note=&english=N&prev_page=1&subjectid=300404125
Use Cases	http://www.getty.edu/research/tools/vocabularies/aat/
Example ID	http://vocab.getty.edu/page/aat/300404125

2.3 Identifier for literature (scientific articles and other publications)

Name	**Digital Object Identifier (DOI)**
Focus	The DOI system provides a technical and social infrastructure for the registration and use of persistent interoperable identifiers, called DOIs, for use on digital networks. A DOI is a persistent identifier or handle used to identify objects uniquely, standardized by the International Organization for Standardization (ISO).
Further reading	https://www.doi.org/ https://www.doi.org/faq.html https://www.doi.org/factsheets/Identifier_Interoper.html
Use Cases	DOIs are widely used mainly to identify academic, professional, and government information, such as journal articles, research reports, data sets, and official publications. However, they have also been used for other types of information resources.
Example ID	https://doi.org/10.3897/rio.7.e67379 (Hardisty et al. 2021)

The [Biodiversity Heritage Library](https://www.biodiversitylibrary.org) (BHL) is the world's largest open access digital library for biodiversity literature and archives. BHL has been retrospectively minting DOIs (**#RetroPIDs**) for historic publications since 2011, but the focus has primarily been on monographs. BHL's new Persistent Identifier Working Group (PIWG) is dedicated to making the content on BHL persistently discoverable, citable and trackable via DOIs and is (at least initially) focusing on journal articles (Kearney 2021).

The International Standard Book Number (**ISBN**) is a numeric commercial book identifier (comprising 13 digits, earlier 10 digits) that is a permanent and citable reference to the related book. Another identifier, the International Standard Serial Number (**ISSN**), identifies periodical publications such as newspapers, magazines and journals. ISBN and ISSN are not PIDs in the strict sense but important identifiers.

2.4 Other identifiers

2.4.1 Identifiers for images or other media

DOIs and repository specific (stable) URIs are used for images, sound, media, etc.

Name	**Entertainment Identifier Registry (EIDR)**
Focus	EIDR is a global unique identifier system for a broad array of audio visual objects, including motion pictures, television, and radio programs as well as for video service providers, such as broadcast and cable networks. EIDR is an implementation of a digital object identifier (DOI).
Further reading	http://www.eidr.org/
Use Cases	http://www.eidr.org/
Example ID	https://ui.eidr.org/view/content?id=10.5240/EA73-79D7-1B2B-B378-3A73-M (the movie 'Blade Runner')

Name	**IIIF Manifest**
Focus	The International Image Interoperability Framework (IIIF) is a way to standardize the delivery of images and audio/visual files from servers to different environments on the Web where they can then be viewed and interacted with in many ways. It defines several application programming interfaces designed to operate with the storage and presentation of digitized objects via a web-based interface.

A IIIF Manifest is the prime unit in IIIF which lists all the information that makes up a IIIF object. It communicates how to display your digital objects, and what information to display about them, including structure, to varying degrees of complexity as determined by the implementer.

The Manifest is what is shown in a Viewer and is usually the thing that can be imported into viewers and other tools. It usually represents a physical object such as a book, an artwork, a newspaper issue, etc.

The IIIF Manifest is accessible via a URL that points to a document online (in a format called JSON, or JavaScript Object Notation) which a IIIF tool can read and display.

Further reading	https://iiif.io/explainers/using_iiif_resources/ https://iiif.io/explainers/using_iiif_resources/#iiif-manifest
Use Cases	https://projectmirador.org/ https://universalviewer.io/
Example ID	Archive of the poet Dioskoros of Aphrodit

Name	**Preston**
Focus	Preston is an open-source software system that captures and catalogs biodiversity datasets. It enables reproducible research: scientists can use Preston to work with a uniquely identifiable, versioned copy of all or parts of GBIF-indexed datasets; dataset registry lookups: institutions can use Preston to check if and when their collections have been indexed and made available through iDigBio; cross-network analysis: biodiversity informatics researchers can use Preston to evaluate dataset overlap between GBIF and iDigBio; and finally, decentralized dataset archival: archivists can distribute Preston-generated biodiversity dataset archives across the world.
Further reading	https://github.com/bio-guoda/preston
Use Cases	https://github.com/bio-guoda/preston Preston tracker
Example ID	

2.4.2 Identifier for nucleotide sequence data and genomic data

The International Nucleotide Sequence Database Collaboration (INSDC; Arita et al. 2021) is the core infrastructure for sharing nucleotide sequence data (NSD) and the corresponding metadata in the public domain. The collaboration is comprised of three partner organizations that keep the identical information through a daily data exchange process that has operated for over 30 years:

- the DNA Data Bank of Japan (DDBJ) at the National Institute of Genetics in Mishima, Japan;
- the European Nucleotide Archive (ENA) at the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) in Hinxton, UK; and
- GenBank at National Center for Biotechnology Information (NCBI), National Library of Medicine, National Institutes of Health in Bethesda, Maryland, USA.

Name	**GenBank (NCBI) accession number**
Focus	Sequence IDs are provided by the sequence database GenBank (NCBI).
Further reading	https://www.ncbi.nlm.nih.gov/genbank/sequenceids/
Use Cases	https://www.ncbi.nlm.nih.gov/genbank/
Example ID	https://www.ncbi.nlm.nih.gov/nuccore/MG215994.1

Name	**BOLD sample IDs**
Focus	The Barcode of Life Data System (commonly known as BOLD or BOLDSystems) is a web platform specifically devoted to DNA barcoding. It consists of four main modules, a data portal, an educational portal, a registry of BINs (putative species), and a data collection and analysis workbench.
Further reading	https://v3.boldsystems.org/
Use Cases	
Example ID	https://www.ncbi.nlm.nih.gov/nuccore/MG215994.1

2.4.3 Identifier for projects

Name	**Research Activity Identifier (RAiD)**

Focus	RAiD is a unique and persistent identifier for research projects. A research project is an activity that takes place over a period of time, has a set scope, is resourced by researchers, research support staff and uses and produces data.
	RAiD acts as a container for research project activities by collecting identifiers for the people, publications, instruments and institutions that are involved.
Further reading	https://www.raid.org.au/ https://ardc.edu.au/services/identifier/ ISO standard under development: https://www.iso.org/standard/75931.html
Use Cases	
Example ID	

2.4.4 Identifier for instruments

Instruments are employed in numerous and diverse scientific disciplines and essential for the creation of research data. An instrument, i.e. a *measuring instrument*, is defined as a "device used for making measurements, alone or in conjunction with one or more supplementary devices" (VIM 2021). Given the importance of instruments and their associated metadata to the assessment of data quality and data reuse, globally unique, persistent and resolvable identification of instruments is crucial. PID assignment for instruments will facilitate the easy implementation of instruments in the metadata for both scientific data publications and journal article publications.

Information about instrumentation used in research has often been missing, but domain specific initiatives have progressed significantly on the implementation of instrument metadata. PID assignment for instruments is now possible and is expected to gain wider adoption in the natural science community in the future.

Name	**Persistent Identifier for instruments (DOI, Handle)**
Focus	The Research Data Alliance Working Group Persistent Identification of Instruments (PIDINST) developed a community-driven solution for persistent identification of instruments. PIDINST developed a universal metadata schema and prototyped schema implementation with DataCite and ePIC as representative persistent identifier infrastructures and with HZB (Helmholtz-Zentrum Berlin für Materialien und Energie) and BODC (British Oceanographic Data Centre) as representative institutional instrument providers.

Further reading	Stocker et al. (2020), http://doi.org/10.5334/dsj-2020-018 Buys et al. (2020), https://doi.org/10.5438/tdk2-2g94 https://www.rd-alliance.org/groups/persistent-identification-instruments-wg
Use Cases	Rolling Deck to Repository (R2R), SENSOR Information System infrastructure
Example ID	

Name	**International Geo Sample Number (IGSN)**
Focus	An International Geo Sample Number (IGSN) provides an unambiguous globally unique persistent identifier for physical samples. IGSN aims to enable transparent and traceable connections between research activities and objects, including samples, collections, instruments, grants, data, publications, people and organizations.
Further reading	https://www.igsn.org/ Lehnert et al. (2019), https://doi.org/10.3897/biss.3.37334
Use Cases	Operates a central registration system for physical samples. Recent partnership with DataCite that intends to support the global adoption, implementation, and use of physical sample identifiers.
Example ID	

2.4.5 Identifiers for collecting events / collection objects

Such identifiers do not exist at present. There could be an identifier for a collecting event such as a moth trap etc. or a collection object which would include all the preparations within it and generally defined by the collector + collector number.

2.4.6 Identifier for trait data

Name	**Identifier for trait datasets (DOI)**

Focus	The Plant trait databaseTRY is using DOIs for datasets comprising trait data.
Further reading	https://www.try-db.org/TryWeb/Submission.php
Use Cases	TRY is working in close collaboration with the Latin American initiative DiverSus (CONICET-Universidad Nacional de Córdoba) and the British Royal Botanical Garden Kew
Example ID	

2.4.7 Identifier for antibodies, model organisms, cell lines and plasmids

Name	**Antibodies Registry ID**
Focus	The Antibody Registry gives researchers a way to universally identify antibodies used in their research. The Antibody Registry assigns unique and persistent identifiers to each antibody so that they can be referenced within publications.
Further reading	https://antibodyregistry.org/about https://scicrunch.org/resources
Use Cases	
Example ID	https://scicrunch.org/resolver/RRID:AB_2178887

Name	**Research Resource Identifiers (RRIDs)**
Focus	The Resource Identification Portal (RRID Portal) was created in support of the Resource Identification Initiative, which aims to promote research resource identification, discovery, and reuse. The portal offers a central location for obtaining and exploring **Research Resource Identifiers (RRIDs)** - persistent and unique identifiers for referencing a research resource such as Antibodies, Model Organisms, Cell Lines, Plasmids, and other Tools (software, databases, services).
Further reading	https://scicrunch.org/resources

Use Cases	https://scicrunch.org/resources
Example ID	https://scicrunch.org/resolver/RRID:AB_2178887

2.4.8 Identifier for software

Name	**Software Heritage**
Focus	It provides unique identifiers, intrinsically bound to the software components. This will ensure that a resilient web of knowledge can be built on top of the Software Heritage archive.
Further reading	https://www.softwareheritage.org/mission/
Use Cases	https://www.softwareheritage.org/
Example ID	SoftWare Heritage persistent IDentifiers (SWHIDs) Example: swh:1:cnt:94a9ed024d3859793618152ea559a168bbcbb5e2

2.4.9 Identifier for patents

Name	**The Lens - Free & Open Patent and Scholarly Search**
Focus	Lens is an online patent and scholarly literature search facility that provides global patent and scholarly knowledge as a public good to inform science and technology. It collaborates with the open access scholarly and open data initiatives, the services are embedded in global public resources like PubMed.
Further reading	https://www.lens.org/
Use Cases	https://www.lens.org/
Example ID	https://www.lens.org/lens/patent/004-676-988-641-380/fulltext?l=en

2.5 Overview of PIDs used for different use cases

	DOI	Wikidata QID	URI	ORCID	ROR	UUID (GUID)	LSID	ARK
People		x		x		(x)	(x)	
Organizations		(x)	(x)		x			
Taxa		(x)	x			x	x	
Localities / sites		x	(x)					
Physical Specimens	(x)		x					x
Digital Specimens	x		(x)			(x)		x
Literature	x		x					x
Images or other media	x							x
Instruments	(x)							
Transaction events						x		

3. Discussion

Community involvement has long been important for the development of PIDs in biocollections (Guralnick et al. 2015) and ongoing discussions in the NSC community related to PIDs are happening during consultations (see above) and on platforms such as the PID Forum. The PID Forum is a global information and discussion platform for persistent identifiers (PIDs), which was initiated by the FREYA project and is now hosted by NISO, the National Information Standards Organization. The PIDforum brings together different research communities working with PIDs and provides a virtual place for sharing best practices (Knowledge Hub), announcing events, asking questions, and having discussions about PIDs (Braukmann et al. 2020).

DiSSCo will not be directly using all of the above-mentioned PIDs or other identifiers. However, it is important to be aware of other developments due to interoperability reasons and to participate in initiatives that support interoperability.
The matrix presented above (Section 2.5) gives an overview of PID systems used for different use cases. While some PID systems are applied for various use cases (e.g. DOIs or Wikidata QIDs), others PIDs are very specific and have been developed for single use cases. In addition, the matrix shows which use cases are using several identifier systems which could limit interoperability.

During the global consultation, the different concepts of sample, specimen, material sample, and subsample had to be clarified. While a material sample is the result of a sampling event, a (catalogued) specimen is the result of a curation process applied to a material sample. Nevertheless, there are some categories of curated objects, i.e. specimens that are not material samples, such as sound recordings or drawings. Specimens can yield subsamples that can themselves be curated (and thus become specimens with identifiers).
Standardizing metadata for heterogeneous sample collections is a present challenge. Beyond samples/specimens, people and organizations are involved with collections, in the curation and management of samples. These persons (agents) need PIDs, such as ORCID IDs and Wikidata QIDs, and ROR IDs for organizations.

Some identifiers gained general importance during the last years. ROR is not only included in DataCite, Crossref and ORCID datasets, it has also been adopted more widely in the NSC community. ROR has reached a level of community adoption that will ensure its long-term place in the persistent identifier and data infrastructure environment. ROR is interoperable with other identifiers and ROR IDs map to GRID (Global Research Identifier Database), ISNI, Crossref Funder ID, and Wikidata.
Since ROR's focus is on the "affiliation use case" identifying organizations and related research outputs, it will probably become more widely used in describing affiliations (formal relationships) and linking between researchers and organizations associated with those researchers, including but not limited to their employer, educator, funder, or scholarly society.

There are numerous different people identifiers, some are more widely used then others, and partly also link to other identifiers. ORCID IDs are becoming widely used in the scientific community. However, there are users that do not want to get an ORCID ID due to e.g. its compulsory use by some publishers, GDPR issues, etc. For deceased persons, the Wikidata QIDs might become a means of choice since it serves also as a hub for many other identifiers. For DiSSCo, current plans involve the use

of ORCID IDs and Wikidata QIDs for people but referring to other identifiers in the metadata should also be possible.

DOIs are the most widely adopted persistent identifier in research data repository systems (Klump & Huber 2017). While the underlying DOI system has a strong commercial backing, other PID systems such as URN and ARK have the backing of national libraries. Sustainability is an essential aspect of PID systems and as Klump & Huber (2017) point out "they do not come for free". To ensure the persistence, resolvability, and discoverability for long periods (e.g., 100+ years) entails a cost. This cost, however, is not so high if compared to the value of research and economic opportunities lost because objects are not properly identified, research datasets lost, etc.

Handle System mechanisms are proposed as the underlying technical and operational infrastructure making the PIDs needed by the natural sciences community persistent and resolvable. DiSSCo will adopt a 'driven-by DOI' persistent identifier (PID) scheme customised to the needs of the natural sciences community (Hardisty et al. 2021). This proposal of adopting DOI as the PID for Digital Specimens is based on a substantial evaluative comparison of 22 Handle System variants (Hardisty et al. 2021).

The cost of operating an appropriate PID scheme based on the Handle System is estimated to be around €1m/$1.2m annually for the 30 billion PIDs needed for Digital extended Specimens in natural history domains. The cost could be shared globally among institutions and/or various research infrastructures but there should be no costs for individual researchers to make use of PIDs.

Other initiatives are also facing the decision to choose the best PID system (e.g. Heritage PIDs) or are working on a sustainable business model to scale to growing demands (IGSN, Global Sample Number, a popular PID currently mostly applied to physical earth samples). The experience of these initiatives can be helpful throughout the process of implementing PIDs in DiSSCo.

DiSSCo has become a member of the International DOI Foundation (IDF) and is working to develop the governance, operations, financing, and service portfolio models, potentially for a new Registration Agency (RA) operating on behalf of the global community. Establishing such a new RA is a practical way forward to support the FAIR (findable, accessible, interoperable, reusable; Wilkinson et al. 2016) data architecture of DiSSCo research infrastructure. This approach is compatible with the policies of the European Open Science Cloud (EOSC) and is aligned to existing practices across the global community of natural science collections.

4. References

Addink W. & Hardisty A. (2020): 'openDS' – Progress on the New Standard for Digital Specimens. Biodiversity Information Science and Standards, 4, e59338. https://doi.org/10.3897/biss.4.59338

Arita M., Karsch-Mizrachi I. & Cochrane G., on behalf of the International Nucleotide Sequence Database Collaboration (2021): The international nucleotide sequence database collaboration. Nucleic Acids Research 49(D1): D121–D124. https://doi.org/10.1093/nar/gkaa967

BCoN (2019): Biodiversity Collections Network. Extending U.S. Biodiversity Collections to Promote Research and Education. American Institute of Biological Sciences. URL: https://bcon.aibs.org/wp-content/uploads/2019/04/BCoN_March2019_FINAL.pdf

Borsch T., Berendsohn W., Dalcin E., Delmas M., Demissew S., Elliott A., Fritsch P., Fuchs A., Geltman D., Güner A., Haevermans T., Knapp S., le Roux M. M., Loizeau P.-A., Miller C., Miller J., Miller J. T., Palese R., Paton A., Parnell J., Pendry C., Qin H.-N., Sosa V., Sosef M., von Raab-Straube E., Ranwashe F., Raz L., Salimov R., Smets E., Thiers B., Thomas W., Tulig M., Ulate W., Ung V., Watson M., Jackson P. W. & Zamora N. (2020): World Flora Online: Placing taxonomists at the heart of a definitive and comprehensive global resource on the world's plants. Taxon 69: 1311-1341. https://doi.org/10.1002/tax.12373

Braukmann R., Cousijn H., Lammey R., Madden F. & Meadows A. (2020, January 30). PIDforum.org - A global discussion platform about PIDs. Zenodo. https://doi.org/10.5281/zenodo.3731058

Brummitt, R. K. & Powell C. E. (1992): Authors of Plant Names. Kew Publishing. ISBN: 1842460854 http://rs.tdwg.org/apn/doc/data/1992

Buys M., Dasler R. & Fenner M. (2020): PIDs for instruments: a way forward. DataCite Blog. https://doi.org/10.5438/tdk2-2g94

Buys M. & Lehnert K. (2021): Bringing together communities: IGSN and DataCite. DataCite Blog. https://doi.org/10.5438/thhf-kx17

Dasler R., Ferguson C., Dohna T., Schindler U., Madden F., Bernal Llinares M., Bunakov V. & Lambert S. (2020): FREYA Deliverable 3.3 Prototypes of New PID Resources. Zenodo. https://doi.org/10.5281/zenodo.4302331

DataCite, Crossref, ORCID & ROR (2021): Making the world a PIDder place: it's up to all of us! - Webinar co-hosted by DataCite, Crossref, ORCID & ROR on September 22, 2021 [further information].

GBIF (2021): Global consultation "Converging Digital Specimens and Extended Specimens - Towards a global specification for data integration" under the umbrella of the alliance for biodiversity knowledge Phase 2, Topic 7. Persistent identifier (PID) schemes. https://discourse.gbif.org/t/7-persistent-identifier-pid-schemes/2664

Grosjean M., Høfft M., Gonzalez M. L., Robertson T. & Hahn A. (2021): GRSciColl: Registry of Scientific Collections maintained by the community for the community. Biodiversity Information Science and Standards 5: e74354. https://doi.org/10.3897/biss.5.74354

Güntsch A., Hyam R., Hagedorn G., Chagnoux S., Röpert D., Casino A., Droege G., Glöckler F., Gödderz K., Groom Q., Hoffmann J., Holleman A., Kempa M., Koivula H., Marhold K., Nicolson N., Smith V. S. & Triebel D. (2017): Actionable, long-term stable and semantic web compatible identifiers for access to biological collection objects. Database 2017: bax003. https://doi.org/10.1093/database/bax003

Guralnick R. P., Cellinese N., Deck J., Pyle R. L., Kunze J., Penev L., Walls R., Hagedorn G., Agosti D., Wieczorek J., Catapano T. & Page E. D. M. (2015): Community Next Steps for Making Globally Unique Identifiers Work for Biocollections Data. ZooKeys 494: 133-154. https://doi.org/10.3897/zookeys.494.9352

Hardisty A. R. (2020): Natural Science Identifiers & CETAF Stable Identifiers. DiSSCo Tech blog post https://dissco.tech/2020/05/28/natural-science-identifiers-cetaf-stable-identifiers/.

Hardisty A., Ma K., Nelson G. & Fortes J. (2019): 'openDS' – A New Standard for Digital Specimens and Other Natural Science Digital Object Types. Biodiversity Information Science and Standards 3: e37033. https://doi.org/10.3897/biss.3.37033

Hardisty A. R., Addink W., Glöckler F., Güntsch A., Islam S. & Weiland C. (2021): A choice of persistent identifier schemes for the Distributed System of Scientific Collections (DiSSCo). Research Ideas and Outcomes 7: e67379. https://doi.org/10.3897/rio.7.e67379

Kearney N. (2021): What Is BHL's New Persistent Identifier Working Group DOI'ng? BHL Blog. https://blog.biodiversitylibrary.org/2021/05/persistent-identifier-working-group.html

Klump J. & Huber R. (2017): 20 Years of Persistent Identifiers – Which Systems are Here to Stay? Data Science Journal 16: 9. http://doi.org/10.5334/dsj-2017-009

Lehnert K., Klump J., Wyborn L. & Ramdeen S. (2019): Persistent, Global, Unique: The three key requirements for a trusted identifier system for physical samples. Biodiversity Information Science and Standards 3: e37334. https://doi.org/10.3897/biss.3.37334

Lendemer J., Thiers B., Monfils A. K., Zaspel J., Ellwood E. R., Bentley A., LeVan K., Bates J., Jennings D., Contreras D., Lagomarsino L., Mabee P., Ford L. S., Guralnick R., Gropp R. E., Revelez M., Cobb N., Seltmann K. & Aime M. C. (2019): The Extended Specimen Network: A Strategy to Enhance US Biodiversity Collections, Promote Research and Education. BioScience 70(1): 23-30. https://doi.org/10.1093/biosci/biz140

Madden, F. & Kotarski R. (2020): Persistent Identifiers at the British Library. https://doi.org/10.23636/1242

Madden F. & Mitchell L. (2021): Persistent Identifiers at Royal Botanic Garden Edinburgh. https://doi.org/10.23636/pvfs-n308

Madden F. & Woodburn M. (2021): Persistent Identifiers at the Natural History Museum. https://doi.org/10.22020/k99s-we61

McMurry J. A., Juty N., Blomberg N., Burdett T., Conlin T. et al. (2017): Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximize utility and impact of life science data. PLOS Biology 15(6): e2001414. https://doi.org/10.1371/journal.pbio.2001414

Meadows A., Haak L. L. & Brown J. (2019): Persistent Identifiers: The Building Blocks of the Research Information Infrastructure. Insights 32 (1): 9. http://doi.org/10.1629/uksg.457

Meadows A., Cousijn H., Gould M., Hendricks G., Petro J. & Simons N. (2021, January 27). PIDs 101: A Beginners' Guide to Persistent Identifiers. Zenodo. https://doi.org/10.5281/zenodo.4574566

Stocker M., Darroch L., Krahl R., Habermann T., Devaraju A., Schwardmann U., D'Onofrio C. & Häggström I. (2020): Persistent Identification of Instruments. Data Science Journal 19(1): 18. http://doi.org/10.5334/dsj-2020-018

VIM (2012): International Vocabulary of Metrology. Basic and General Concepts and Associated Terms (VIM 3rd edition). https://www.bipm.org/en/publications/guides/#vim

Wilkinson M., Dumontier M., Aalbersberg I. et al. (2016): The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3: 160018. https://doi.org/10.1038/sdata.2016.18