Publishing your data

Øystein Godøy and Markus Fiebig



Outline

- Mandated and long term archives (Øystein)
- Data publications (Øystein)
- PID (Explicit mention DOI) (Øystein)
- Data policies / Licensing (Markus)
- Tracking usage (using DOI) (Markus)
- Repositories: (Markus)
 - NorDataNet (distributed network of data centres)
 - NIRD RDA
 - GAW repositories
 - Repositories for model data
 - Figshare



Mandated and long term archives

- Mandated archive .
 - An archive that has data management as part of the mandate
- Long term archive
 - An archive that has long term funding and mitigation procedures established
- NorDataNet connects existing mandated data archives with long term commitment
- Always choose a mandated archive for your data deposit
- Services offered by the archives vary, but always ٠ select one supporting Persistent Identifiers
 - And preferably one that enables services on top of data allowing them to be integrated in various work flows
- For overview of data centres, check ٠
 - https://www.re3data.org/



Visited 🗧 Getting Started 🜐 N	lorDataNet Arctic 🛛 🕅 XPath a	ind XSLT with I 🖨 Parsing XML and HT 🖨 xslt.md.html 🔞 DuckDuckGo — Priva
data.org		Search Browse - Suggest Resources - Contact 🔶 DataC
Filter	Search	Q Search
Reset all		Toogle short he
bjects 🕀		
ntent Types 🕀	← Previous 1 2	Next → Sort by -
untries 🗆		
stria (2)	Found 27 result(s)	
glum (1) nada (1)	round Er rooun(b)	
opean Union (11)	ACTRIS Data Cent	ire 🚺 🙆 🔘 📼 🗐 §.
and (3)	ACTRIS DC	
many (4)	Subject(s)	Consciences (including Constrants) Atmospheric Science and Oceanography Natural Sciences
ece (1)	Subject(s)	devaciences (including deography) Aunospheric Science and Oceanography Hatural Sciences
(1)	Content type(s)	Raw data Archived data other Structured text Scientific and statistical data formats
rway (27)		Structured graphics Databases
tugal (1) ssian Federation (1)	Country	
aln (1)	Country	Norway European Union
Itzerland (3) ted Kingdom (2)	The ACTRIS DC is desig	ned to assist scientists with discovering and accessing atmospheric data and contains an up-to-date
ted States (3)	catalogue of available da	tasets in a number of databases distributed throughout the world. A site like this can never be complete,
) systems 🕀	but we have aimed at inc	luding datasets from the most relevant databases to the ACTRIS project, also building on the work and
1 	experiences achieved in	the EU FP6 research project Global Earth Observation and Monitoring. The focus of the web portal is
rtificates 🕀	The web site allows your	to search in a local metadata catalogue that contains information on actual datasets that are archived in
a access 🕀	external archives. It is se	t up so that you can search for data by selecting the chemical/physical variable, the data location, the
a access restrictions	database that holds the o	data, the type of data, the data acquisition platform, and the data matrix
abase access 🕀		
abase access restrictions	ENED	
abase licenses 🕀	The European Monitoring and Evaluation Programme	
a licenses 🕀	The European Monitoring	j and Evaluation riogramme
a upload ⊞	Subject(s)	Geosciences (including Geography) Atmospheric Science and Oceanography Geophysics and Geodesy
a upload restrictions 🕀		Geochemistry, Mineralogy and Crystallography Natural Sciences
hanced publication	Content type(s)	Raw data other Databases
titution responsibility type 🗄	comon (polo)	
titution type	Country	Austria Switzerland Finland Norway Russian Federation United Kingdom
and a second sec		

What is a data paper?

A data paper is a peer reviewed document describing a dataset, published in a peer reviewed journal. It takes effort to prepare, curate and describe data. Data papers provide recognition for this effort by means of a scholarly article.



Data papers explained

Unlike a conventional research article, the primary purpose of a data paper is to describe data and the circumstances of their collection, rather than to report hypotheses and conclusions.

. . .

https://www.gbif.org/data-papers



Why publish data papers?

- By publishing a data paper, you will:
 - Receive credit through indexing and citation of the published paper, in the same way as with any conventional scholarly publication, offering benefits to authors in terms of recognition and career building
 - Increase the visibility, usability and credibility of the data resources you publish
 - Track more effectively the usage and citations of the data you publish.

https://www.gbif.org/data-papers



Data publishing and papers

- I. Prepare and publish data through a mandated archive
 - I. Remember to get a DOI!!
- II. Write a data paper
 - I. Cite your datasets using the data centre provided DOI
- III.Write your scientific paper
 - I. Cite your data paper describing the context of the data





Persistent Identifiers

- A persistent identifier is a long-lasting reference to a digital resource.
- An identifier is a label which gives a unique name to an entity: a person, place, or thing.
 - Unlike URLs, which may break, a persistent identifier reliably points to a digital entity.
 - An ORCID iD is an example of a persistent identifier for a person.

https://support.orcid.org/hc/en-us/articles/360006971013-What-are-persistent-identifiers-PIDs-



Different PID types

- There are different PID types for different kinds of resources.
- In the current research environment we most commonly see two varieties:
 - those for objects
 - publications, data, software
 - such as URNs, DOIs, ARKs, Handle
 - those for people
 - researchers, authors, contributors
 - such as ORCIDs, ISNIs

https://www.openaire.eu/what-is-a-persistent-identifier





PID schemes

- Digital Object Identifier (DOI)
 - DOIs are digital identifiers for objects (whether digital, physical or abstract) which can be assigned by organisations in membership of one of the DOI Registration Agencies; the two best known ones are CrossRef, for journal articles and some other scholarly publications, and DataCite for a wide range of data objects. As well as the object identifier, DOI has a system infrastructure to ensure a URL resolves to the correct location for that object.
- Handle
 - Handles are unique and persistent identifiers for Internet resources, with a central registry to resolve URLs to the current location.
 Each Handle identifies a single resource, and the organisation which created or now maintains the resource. The Handle system also underpins the technical infrastructure of DOIs, which are a special type of Handles.
- Persistent Uniform Resource Locator (PURL)
 - PURLs are URLs which redirect to the location of the requested web resource using standard HTTP status codes. A PURL is thus a permanent web address which contains the command to redirect to another page, one which can change over time.
- Universal Resource Name (URN)
 - URNs are persistent, location-independent identifiers, allowing the simple mapping of namespaces into a single URN namespace. The existence of such a Uniform Resource Identifier does not imply availability of the identified resource, but such URIs are required to remain globally unique and persistent, even when the resource ceases to exist or becomes unavailable. The URN term is now deprecated except in the very narrow sense of a formal namespace for expressing a Uniform Resource Identifier.

https://www.dpconline.org/handbook/technical-solutions-and-tools/persistent-identifiers



Choosing a Persistent Identifier Scheme

- Advantages
 - Critically important in helping to establish the authenticity of a resource.
 - Provides access to a resource even if its location changes.
 - Overcomes the problems caused by the impermanent nature of URLs.
 - Allows interoperability between collections.
- Disadvantages
 - There is no single system accepted by all, though DOIs are very well established and widely deployed.

https://www.dpconline.org/handbook/technical-solutions-and-tools/persistent-identifiers



Persistent identifiers: awareness level

- Further reading on persistent identifiers from Australian National Data Service (ANDS)
 - https://www.ands.org.au/ guides/persistent-identifiersawareness







https://www.ands.org.au/guides/persistent-identifiers-awareness

