

23 Things Physical Samples

An overview of practical, free, online resources and tools that you can begin using today!

Physical samples introduction

These articles provide an overview of the existing resources or challenges that are particularly relevant to physical samples:

1. 'Let's make it easy: A workflow for physical sample metadata rescue'.

Hills 2015: <https://doi.org/10.1016/j.grj.2015.02.007>

2. 'Going Digital: Persistent Identifiers for Research Samples, Resources and Instruments'.

Plomp 2020: <http://doi.org/10.5334/dsj-2020-046>

3. 'A Digital Repository for Physical Samples: Concepts, Solutions and Management'.

Devaraju et al. 2017: https://doi.org/10.1007/978-3-319-67008-9_7

Persistent identifiers

A persistent identifier is a long-lasting reference to a document, file, web page, or other object.

4. A global persistent identifier for physical samples: **IGSN** is an alphanumeric code that is assigned to specimens and related sampling features to ensure their unique identification and unambiguous referencing.

Klump et al. 2021: <http://doi.org/10.5334/dsj-2021-033> / <https://iqsn.github.io/> / <https://doi.org/10.5281/zenodo.31788>

5. The **Research Resource Identifiers (RRIDs)** supports persistent and unique identifiers for referencing a research resource (such as antibodies, organisms) in biomedical research.

<https://scicrunch.org/resources> / <https://doi.org/10.12688/f1000research.6555.2> / <https://doi.org/10.1016/j.neuron.2016.04.030>

6. **PIDINST**: persistent identifiers for analysis instruments

<https://doi.org/10.5334/dsj-2020-018>

7. When do you use a specific identifier?

See Damerow et al. 2021: <http://doi.org/10.5334/dsj-2021-011>

Metadata

Metadata describes physical samples and allows the developments of standards so that data is more easily discovered, understood, and preserved.

8. Find databases and metadata standards using re3data and FAIRsharing:

<https://www.re3data.org/> / <https://fairsharing.org/>

9. **Metadata 'bullseye'**: Working towards a common kernel for sample metadata that are shared among a community of practice. More specific metadata can be added around the kernel.

See Wyborn et al. 2020: <http://doi.org/10.5281/zenodo.4694740>

10. **The Digital Archaeological Record (tDAR)** uses the UK Archaeological Data Services good practices

<https://guides.archaeologydataservice.ac.uk/g2gpwiki/>

11. Set up a **metadata standard**. See for example the process for an ISO standard for Data and Specimen Provenance in Life Sciences:

<https://doi.org/10.5281/zenodo.5093125>

Citing Data

There are existing initiatives that are helping to promote the scholarship of data by encouraging and enabling data citation, assigning identifiers to data, creating links between documents and data, and helping users properly attribute credit to data producers.

12. **State how others should cite the data** to make it easy for people to refer to your resources, see **RRIDs** and **DataCite** for examples:

<https://www.rriids.org/draft-texts> / <http://bit.ly/RDAthing14>

Data Licensing and Ownership

13. In general **Creative Commons Licenses** are used for data derived from samples, with CCo license commonly used for metadata about the data.

<https://creativecommons.org/licenses>

14. Examples of **sample material transfer agreements** have been set up by the Biotechnology Industry Organization and addgene: <https://help.addgene.org/hc/en-us/sections/201196819-Material-Transfer-Agreements-MTAs-> / <https://www.wipo.int/tk/en/databases/contracts/texts/bio.html>

Physical data tools

15. Some **Electronic Lab Notebooks** have sample management modules
RSpace: www.researchspace.com / eLabNext: www.elabnext.com / Labguru: www.labguru.com/inventory / Chemotion: <https://www.chemotion.net/> /
Labfolder: <https://www.labfolder.com>

16. Examples of other **software tools** include the **R-package baRcodeR** to generate barcode labels for physical samples (Wu et al. 2020: <https://doi.org/10.1111/2041-210X.13405>) and **Menoci** which allows tracking of samples (<https://menoci.io/>, <https://doi.org/10.1186/s12859-020-03928-1>).

Repositories

Existing infrastructures for physical samples include:

17. **Interdisciplinary Infrastructure for samples** such as iSamples
<http://isamples.org> / Davies et al. 2021: <https://doi.org/10.1093/gigascience/giabo28>

18. **The System for Earth Sample Registration (SESAR)** is a community platform that helps make Earth, Environmental, and Planetary Science samples more discoverable, accessible, and reusable
www.geosamples.org

19. **Index to Marine and Lacustrine Geological Samples (IMLGS)** is a database maintained on behalf of the Curator community that helps scientists and others discover, learn about, and obtain geologic materials from participating repositories.
<https://www.ngdc.noaa.gov/mgg/curator/>

20. **Biological Research Samples Stock Centre List** by Mary Donaldson
<https://doi.org/10.5281/zenodo.4923489>

21. Or set up a **new repository**, for example, a Biobank: (www.ncbi.nlm.nih.gov/pmc/articles/PMC6777713), or a **directory** such as the BBMRI-ERIC

Directory of Biobanks (<https://directory.bbmri-eric.eu/menu/main/app-molgenis-app-biobank-explorer>). You can follow recommended practices for **Open Archival Information Systems** (<http://public.ccsds.org/publications/archive/650xom2.pdf>) or get inspired by 'Managing Natural Science Collections: A guide to Strategy, Planning and Resourcing' by Huxley et al. 2020: <https://doi.org/10.4324/9780429426599>.

Communities of Practice

Several communities are already involved in conversations about metadata and persistent identifiers for physical samples with the larger community of researchers, technologists, funders, publishers, and others to develop solutions and share best practices.

22

- a. **Earth Science Information Partners (ESIP)** has a **Physical Samples Cluster** and holds an annual conference with sessions that address physical samples in research (https://wiki.esipfed.org/Physical_Sample_Curation).
- b. **Sampling Nature** is a research coordination network organising workshops and working groups (<http://samplingnature.org/>).
- c. **Environmental Systems Science (ESS-DIVE)** stores and publicly distributes data from observational, experimental, and modeling research (<https://ess-dive.lbl.gov>).
- d. **Distributed System of Scientific Collections (DiSSCo)** for **Natural Science Collections** and **digital specimens** (<https://www.dissco.eu/key-messages/dissco-in-a-nutshell> / <https://bit.ly/DigitalSpecimen>).
- e. the **Humanities Commons** have several groups that work on **primary sources** (<https://hcommons.org/groups>).

23. Join the Research Data Alliance!

Belong to an international community who builds social and technical bridges to enable data sharing. It's free to join by visiting the website, then subscribe to the **Physical Samples and Collections in the Research Data Ecosystem Interest Group** <https://www.rd-alliance.org/groups/physical-samples-and-collections-research-data-ecosystem-ig>



*Based on 23 Things: Libraries for Research Data (2016).
Research Data Sharing Without Barriers www.rd-alliance.org RDA 2021*

Contributors: Esther Plomp, Sarah Ramdeen, Jens Klump, Lesley Wyborn, Clint Edrington, Mary Donaldson, Joan Damerow, Sharif Islam

Physical Samples and Collections in the Research Data Ecosystem Interest Group

<https://www.rd-alliance.org/groups/physical-samples-and-collections-research-data-ecosystem-ig>