

23 Things Physical Samples

RDA Supporting Output



DOI: 10.15497/RDA00067

Authors: Wouter Addink, Joan Damerow, Mary Donaldson, Clint Edrington, Sharif Islam, Jens Klump, Esther Plomp, Sarah Ramdeen, Natasha Simons, Lesley Wyborn, Physical Samples and Collections in the Research Data Ecosystem WG

Published: 8 July 2022

Version: 1.0

Abstract: Physical samples are a basic element for reference, study, and experimentation in research. The 23 Things for Physical Samples aims to provide a reference overview of resources centered on the management and sharing of data on material samples. The output focuses on existing work, recent developments, recommended practices, and community initiatives. The 23 resources are related to the following categories: 1) a general introduction, 2) persistent identifiers, 3) metadata, 4) citing samples, 5) data licensing and ownership, 6) tools, 7) repositories and 8) communities of practice.

Impact: The '23 Things Physical Samples' output will be of value to anyone working with samples and specimens, as well as others interested in this topic. The output provides an overview of resources centered on physical sample management and sharing. Resources included are applicable to any scientific discipline working with sample or specimen data. The output highlights examples and resources from the following research domains: Archaeology, Earth Sciences, Biomedical Sciences, Life Sciences, Biology, Genomics, Environmental Science, and Natural Sciences. The output increases awareness of the various initiatives that focus on physical samples and specimens, ultimately facilitating cross disciplinary data exchange and interoperability.

Language: English

License: [Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

RDA webpage: <https://www.rd-alliance.org/group/physical-samples-and-collections-research-data-ecosystem-ig/outcomes/23-things-physical>

Citation and Download: Plomp, E., Ramdeen, S., Klump, J., Wyborn, L., Edrington, C., Donaldson, M., Damerow, J., & Islam, S. (2021), [Physical Samples and Collections in the Research Data Ecosystem IG](#) (2022). 23 Things Physical Samples. *Research Data Alliance*. [10.15497/RDA00067](https://doi.org/10.15497/RDA00067).

23 Things Physical Samples

A reference resource of practical resources and tools that you can begin using today to manage and share data on material samples!

Physical samples introduction

Physical samples include biological specimens, rock or mineral specimens, soil or sediment cores, plants and seeds, water quality samples, archaeological artefacts, or DNA and human tissue samples. Samples may be used in analyses and can be destroyed in this process, whereas specimens are preserved curated objects that can be continuously studied. The articles below (Things 1-3) provide an overview of existing resources or challenges particularly relevant to physical samples:

1. 'Let's make it easy: A workflow for physical sample metadata rescue' - [Hills 2015](#)
2. 'Going Digital: Persistent Identifiers for Research Samples, Resources and Instruments' - [Plomp 2020](#)
3. 'A Digital Repository for Physical Samples: Concepts, Solutions and Management' - [Devaraju et al. 2017](#)

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](#)). ([video 1](#) / [video 2](#))
5. The **Research Resource Identifiers (RRIDs)** supports persistent and unique identifiers for referencing a research resource (such as antibodies, organisms) in

biomedical research ([Bandrowski et al. 2015](#) / [Bandrowski and Martone 2016](#)). ([rd-alliance.org/PS_RRIDs_August2021_webinar](#))

6. **PIDINST**: persistent identifiers for analysis instruments ([Stocker et al. 2020](#)). As samples are often analysed using analysis instruments, this persistent identifier may provide additional context to the sample and analysis results. ([rd-alliance.org/PID-instruments-May2022_webinar](#))

7. **When do you use a specific identifier?**
See [Damerow et al. 2021](#).

Metadata

Metadata describe physical samples and allow for the developments of standards so that data are more easily discovered, understood, and preserved.

8. Find **databases and metadata standards** using [re3data](#) and [FAIRsharing](#).
9. **Metadata 'bullseye'** is ongoing work towards a common kernel for sample metadata that are shared among communities of practice. More specific metadata can be added around the kernel ([Wyborn et al. 2020](#) / [Klump et al. 2021](#)).
10. **The Digital Archaeological Record (tDAR)** uses the [UK Archaeological Data Services good practices](#).
11. Set up a **metadata standard** like [TDWG](#) and DarwinCore for (biological) specimens or GGBN for genomic samples. See for example the process for an [ISO standard for Data and Specimen Provenance in Life Sciences](#).

Citing Samples

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. Work is ongoing on guidelines specifically for [specimen and sample citation](#).

Data Licensing and Ownership

13. A license is a standardised way to grant permission for reuse of data. In general [Creative Commons Licenses](#) are used for data derived from samples, with the [CC0 license](#) commonly used for metadata about the data.

14. Examples of **sample material transfer agreements** have been set up by the [Biotechnology Industry Organization](#) and [addgene](#).

Physical data tools

15. Some **Electronic Lab Notebooks** have sample management modules ([RSpace](#) / [eLabNext](#) / [Labguru](#) / [Chemotion](#) / [Labfolder](#)) ([rd-alliance.org/physical-sample-webinar1-June2021](#)).

16. Examples of other **software tools** include the **R-package baRcodeR** to generate barcode labels for physical samples ([Wu et al. 2020](#)), [Menoci](#) which allows tracking of samples ([Surh et al. 2020](#)), and [FAIMS](#) for offline data collection in the field.

Repositories

Existing infrastructures for physical samples include:

17. Interdisciplinary Infrastructure for samples such as [iSamples](#) ([Davies et al. 2021](#)).

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 ([geosamples.org](#)).

19. Index to Marine and Lacustrine Geological Samples (IMLGS) is a database that helps scientists and others discover, learn about, and obtain geologic materials from participating repositories.

20. Biological Research Samples Stock Centre List, which lists stock centres that might be suitable for the deposit of biological samples ([Donaldson 2021](#)).

21. Or set up a **new repository**, for example, a [Biobank](#) or a **directory** such as the [BBMRI-ERIC Directory of Biobanks](#). You can follow [recommended practices](#) for **Open Archival Information Systems** or get inspired by 'Managing Natural Science Collections: A guide to Strategy, Planning and Resourcing' ([Huxley et al. 2020](#)).

Communities of Practice

Several communities are already involved in conversations about metadata and persistent identifiers for physical samples, working on solutions and sharing 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.

b. [Sampling Nature](#) is a research coordination network organising workshops and working groups.

c. **Environmental Systems Science (ESS-DIVE)** stores and publicly distributes data from observational, experimental, and modelling research.

d. **Distributed System of Scientific Collections (DiSSCo)** provides research infrastructure and a platform for discussions for **Natural Science Collections** and [digital specimens](#) ([Hardisty et al. 2021](#)).

e. The [Humanities Commons](#) have several groups that work on **primary sources**.

f. **Consortium of European Taxonomic Facilities (CETAF)** is Europe's network of biological and geological collections.

g. The RDA group on [Physical Samples and Collections](#) facilitates cross-domain discussions around key issues related to the digital representation of physical samples and collections ([rd-alliance.org/ps-interdisciplinarysampledata-October-webinar](#)).

h. **Australian Research Data Commons (ARDC) Information Management for Physical Samples Community of Practice** will bring together practitioners in the Australian research sector to foster best practice in the identification, citation and management of physical samples throughout the research lifecycle.

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](#).

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



Citation: Plomp, E., S. Ramdeen, J. Klump, L. Wyborn, C. Edrington, M. Donaldson, J. Damerow, S. Islam, N. Simons, W. Addink, 2022, **23 Things Physical Samples**, Research Data Alliance, [10.15497/RDA00067](#).