Recommendation 4.0

Recommendation to use IGSN as the standard reference in technical infrastructures to samples where appropriate

Description

Status: Draft 20.5.2025

Motivation for this Recommendation:

The International Generic Sample Number (IGSN) is a globally unique and persistent identifier designed specifically for physical samples and related objects. At Helmholtz, we recommend the use of IGSNs to ensure that samples—and other tangible sources from which data are derived—can be reliably identified, referenced, and linked across research workflows. The motivation to use IGSNs lies in their ability to improve traceability, reproducibility, and data integration across disciplines. By assigning a persistent identifier to a sample, researchers can unambiguously connect it to associated datasets, publications, instruments, and collection metadata, supporting FAIR principles and enabling long-term reuse and verification of research outcomes.

Recommendation

IGSN is used to identify samples in data infrastructures.

For organizations this means:

- a person or unit should be made responsible to maintain the centres IGSNs.

For data curators this means:

- Enable, train and encourage staff to register IGSN when samples are taken.
- Enable, train and encourage staff to record any parent IGSNs with subsamples.

For researchers it means:

- Record an IGSN with any sample taken.
- Record the parent IGSN with any subsample measured

For data infrastructures:

- record a IGSN to identify samples and parent samples and make this data part of the metadata available for harvesting.
- treat IGSN metadata as the primary source of truth and update your own metadata accordingly.

Binding Convention:

	mandatory	conditional	optional
Helmholtz FAIR Principle			

Precondition for Implementation:

Related Recommendations

Parent:

Dependent:

Other: none

Contributors

Names of contributors to this recommendation

Manu

Content

[1] Plankytė, Vaida, Macneil, Rory, & Chen, Xiaoli. (2023). Guiding principles for implementing persistent identification and metadata features on research tools to boost interoperability of research data and support sample management workflows. Zenodo. https://doi.org/10.5281/zenodo.8284206

1. Explanation of the Background and Benefits of the Recommendation

About

The International Generic Sample Number (IGSN) is a persistent, globally unique identifier designed to unambiguously reference physical samples and other material objects in the research lifecycle. It enables reliable citation, tracking, and linking of samples to related data, instruments, people, and publications, making them FAIR—findable, accessible, interoperable, and reusable.

History

Originally developed by the geoscience community in the early 2000s, IGSN emerged from the need

to manage and cite geological samples across laboratories and institutions. It was formalized through the IGSN e.V. foundation in 2011 and has since evolved into a cross-disciplinary identifier supported by the global research infrastructure. Since 2021, IGSNs have been registered through DataCite, aligning their metadata with other research outputs.

Structure

IGSN records consist of a unique identifier (a prefix-suffix structure similar to DOIs) and a metadata record that captures core descriptive information about the sample: sample type, material, collection method, spatial and temporal context, and links to related entities (e.g., datasets, people, institutions). Metadata can be enhanced to fit domain-specific needs while maintaining a consistent structure for interoperability.

Motivation

Using IGSNs improves sample traceability, ensures reproducibility of results, and supports data integration across disciplines. It allows researchers to explicitly reference the physical basis of data analyses, which is critical for verification, reuse, and credit assignment.

Current Use of IGSN

IGSNs are currently used in a range of domains, including geosciences, environmental sciences, archaeology, and biodiversity research. For example, ocean drilling samples from IODP expeditions, sediment cores, rock specimens, water samples, and even archaeological artifacts have been assigned IGSNs. These identifiers help integrate sample-based research into digital infrastructures and link physical materials to datasets and publications, thus enabling transparent and connected science.

2. Possible alternative solutions

1. 1. Internal or Local Identifiers

What: Lab- or institution-specific sample IDs. Pros: Easy to implement, tailored to local needs. Cons: Not globally unique, not resolvable, hard to track across systems or publications.

1. 2 Accession Numbers in Domain Repositories

What: Identifiers assigned by domain-specific repositories or museums (e.g., GenBank accession numbers, museum catalog numbers).

```
  Pros: Well-integrated in their domains.
  Cons: Often not globally unique, not persistent outside their system, not
interoperable across disciplines.
```

1. 3 Handle System / Custom DOIs

What: Using general-purpose persistent identifiers like DOIs or Handles for samples.

```
  Pros: Technically viable; DOI infrastructure is mature.
  Cons: Lack of community consensus or metadata model for samples unless
built on top of IGSN or similar; harder to ensure consistency and semantic
```

```
clarity.
```

1. 4 ARK (Archival Resource Key)

What: A persistent identifier scheme designed for objects of any type.

```
  Pros: Flexible, openly governed, used by some institutions (e.g., museums,
archives).
  Cons: Less widely adopted in science, lacks built-in metadata requirements
for samples, limited interoperability in research workflows.
```

Why IGSN? While alternatives exist, IGSN is currently the only PID system specifically designed to handle the complexities of referencing physical samples across scientific domains. It combines:

```
  Global uniqueness and persistence
  A structured, interoperable metadata schema
  Community governance
  Integration with DataCite infrastructure
  Support for linking to related PIDs (e.g., ORCID, ROR, dataset DOIs)
```

Therefore, for research workflows that require transparent, machine-readable, and citable links between samples and data, IGSN remains the most suitable and sustainable option.

3. Consideration of the advantages and disadvantages of implementing the recommendation

(quality of content, limitations, interoperability, sustainability: expected future dissemination / technical availability / funding)

4. The Recommendation

IGSN is used to identify samples in data infrastructures.

For organizations this means:

- a person or unit should be made responsible to maintain the centres IGSNs.

For data curators this means:

- Enable, train and encourage staff to register IGSN when samples are taken.
- Enable, train and encourage staff to record any parent IGSNs with subsamples.

For researchers it means:

- Record an IGSN with any sample taken.
- Record the parent IGSN with any subsample measured

For data infrastructures:

- record a IGSN to identify samples and parent samples and make this data part of the metadata available for harvesting.
- treat IGSN metadata as the primary source of truth and update your own metadata accordingly.

Also see [3] Baldewein et al. (2023). FAIR WISH D7 -Standard Operating Procedure for automatic IGSN registration. Zenodo. https://doi.org/10.5281/zenodo.10401380

5. Naming of communities that have already implemented the recommendation

6. Documentation of the test to validate correct implementation

7. Examples of Instances

IGSN is implemented within the Helmholtz association at AWI, GFZ, and Hereon through the FAIRWish Project [4]. See [3] for more information.

Another implementation is documented at the Kiel University (CAU) [6]

8. Further Information

References

[1] Plankytė, Vaida, Macneil, Rory, & Chen, Xiaoli. (2023). Guiding principles for implementing persistent identification and metadata features on research tools to boost interoperability of research data and support sample management workflows. Zenodo. https://doi.org/10.5281/zenodo.8284206

[2] Klump, J., Lehnert, K., Ulbricht, D., Devaraju, A., Elger, K., Fleischer, D., Ramdeen, S., Wyborn, L. (2021): Towards Globally Unique Identification of Physical Samples: Governance and Technical Implementation of the IGSN Global Sample Number. - Data Science Journal, 20, 1, 1-16., DOI: https://doi.org/10.5334/dsj-2021-033

[3] Baldewein, L., Kleeberg, U., Brauser, A., Elger, K., Frenzel, S., Heim, B., & Wieczorek, M. (2023). FAIR WISH D7 - Standard Operating Procedure for automatic IGSN registration. Zenodo. https://doi.org/10.5281/zenodo.10401380

[4] The FAIR Wish Project: https://helmholtz-metadaten.de/de/inf-projects/fair-wish-fair-workflows-to-establish-igsn-for-samples-in-the-helmholtz-association

[5] IGSN Documentation on forschungsdaten.org https://www.forschungsdaten.org/index.php/IGSN

[6] IGSN Service and Documentation at the University Kiel https://igsn.uni-kiel.de/de

Relevant Community Recommendations

9. History of this document