

Enabling FAIR (Findable, Accessible, Interoperable and Reusable) principles for ice core samples through the application of unique identifiers

Kershaw, C. T. ¹

¹ National Science Foundation-Ice Core Facility, U.S. Geological Survey

Introduction

Physical sample curation has evolved dramatically over the last decade, advanced by the idea that physical samples and their metadata must be Findable, Accessible, Interoperable, and Reusable (FAIR). Our first step in achieving these FAIR goals is through the use of unique and persistent identifiers. The National Science Foundation-Ice Core Facility (NSF-ICF) is developing capabilities to assign unique identifiers to its samples to advance these FAIR principles for ice cores and derivative samples. In the process of achieving this goal, we are compiling controlled vocabularies, creating standard-based metadata profiles, and documenting best practices to facilitate adoption of these practices by other ice core repositories and users. We seek community input on all aspects of this work. Persistent identifiers will be assigned as International Geo Sample Numbers (IGSNs). Initially, registration will be assisted by the System for Earth Sample Registration (SESAR), who will also facilitate community access to the ice core and sample metadata profiles, while the USGS develops these capacities internally. The establishment of profiles and practices will provide the resources needed to expedite the adoption of persistent identifiers (IGSNs) by other ice core repositories (Plomp, 2020).

Community engagement that includes ideas, criticisms, and opinions is essential for widespread adoption of IGSNs for ice core samples. It is through this input that one of the most imperative goals of this project will be accomplished: establishing a set of ice core and sample metadata profiles that are available to the ice core community for use.

Three Primary Methods of Gathering Community Input

1. Online Surveys and Questionnaires (Broadest Reaching, Less Interactive)
2. Online Virtual Meetings (Smaller Audience, Most Interactive)
3. Presentations and Outreach at Conferences (Broad Reaching, Somewhat Interactive, Restricted by COVID-19)

Community Input

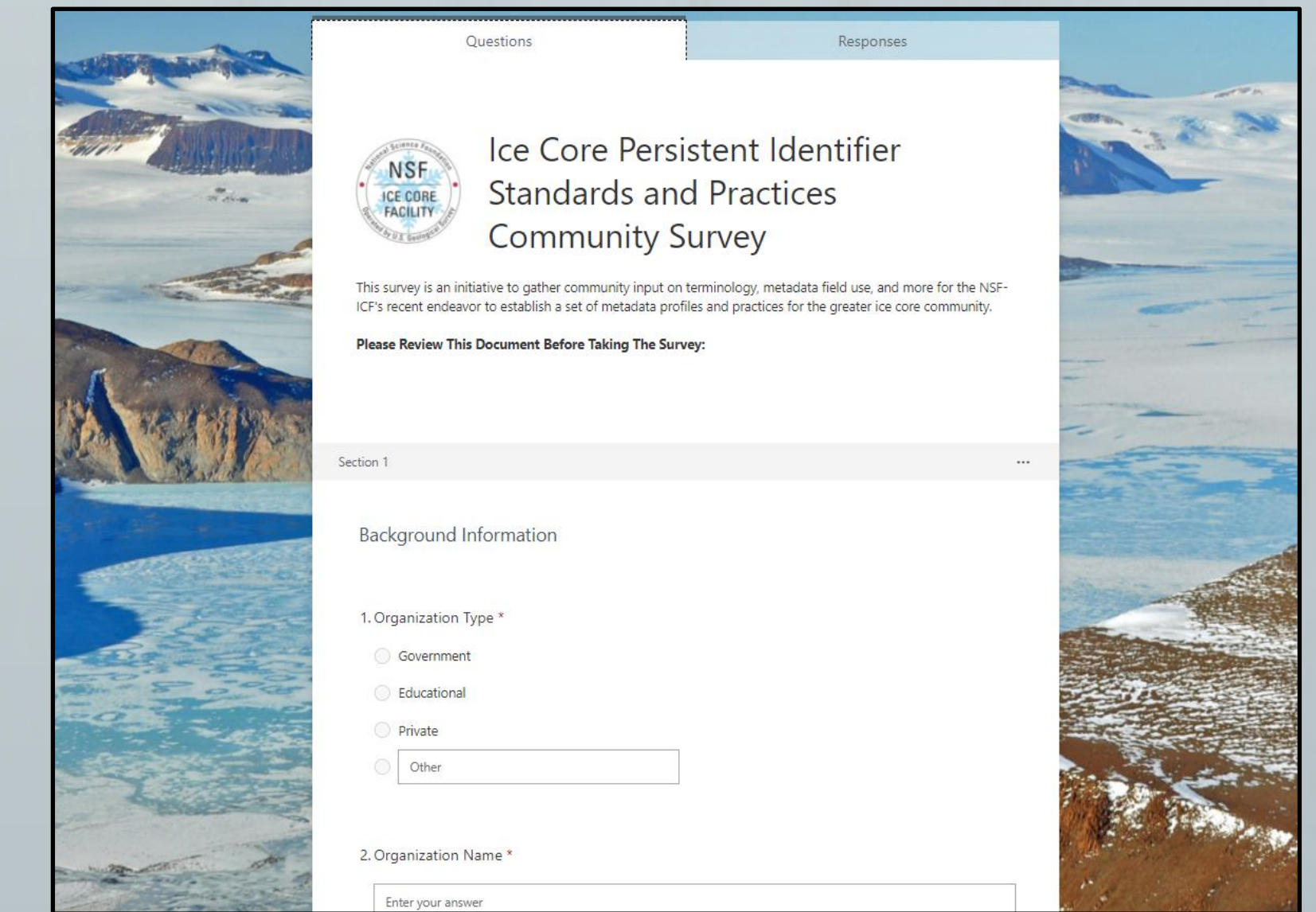
Community Input Plans

Inquiry Emails to Ice Core Community Members

Outreach at Conferences (e.g. AGU & ESIP)

Online Survey through Microsoft Forms

Virtual Meetings With Ice Core Community Members

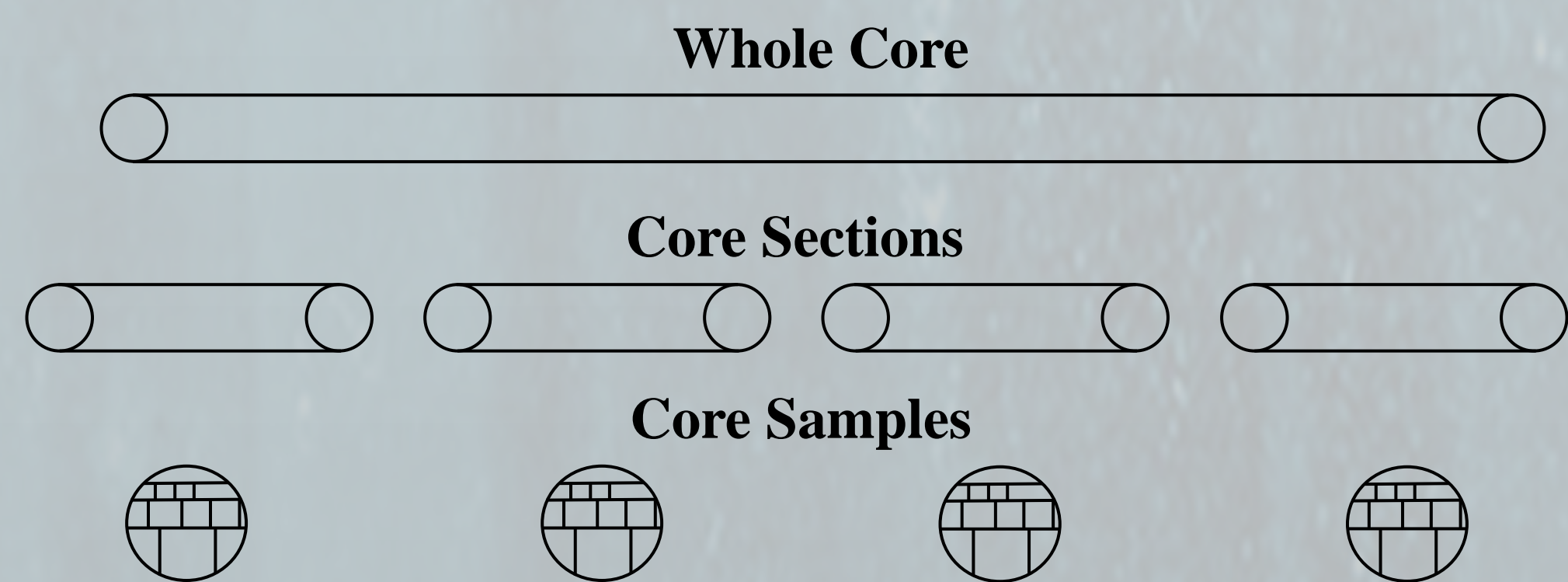


An online survey for community input will be distributed using Microsoft Forms and have four primary sections.

1. A background section to gather information about the participants.
2. A section revolving around metadata.
3. A section concerned primarily with terminology.
4. A small section for additional comments and questions

Mapping ICF Metadata into SESAR Profiles

Metadata included with each IGSN must accurately reflect the core or parts of the core it is derived from.



Metadata Must-haves

- Depth and Size Measurements
- Drill Information/Specifications
- Current Archive
- Location Data (Lat/Long)
- Project/Program Name
- Primary Investigator Information
- Ice Quality Information
- Type of Analysis Being Conducted (Samples Only)

Due to the specialized nature of ice core metadata, translating in-house metadata into SESAR's metadata schema had several roadblocks.

- Metadata that can fit into more than one field
- Deciding whether the metadata provided by our FileMaker database is relevant to include.
- Terminology differences across ice core facilities.
- Variability in metadata field meaning between the three different types of core data (whole core, core sections, and core samples).

Core ID	Section	Top Depth	Bottom Depth	Section Length	Core Diameter (cm)	Section Quality
GISP2D	45	H4.000	45.000	1.000	13.2	
Click diagram to see all of this out for this core						
Available Ice: 2 Pieces						
GISP2D	Top: 44.000	Type: ABC05		50270		Details
TUBE: 45	Bottom: 45.000	Breaks: None				
GISP2D	Top: 44.000	Type: MCA05		50270		Details
TUBE: 45	Bottom: 45.000	Breaks: None				
	Length: 1.000	Fractures: 0				

FileMaker Data

Using the now matched metadata, ice cores will be assigned IGSN's and receive a metadata landing page on SESAR's website.

Field Matching Excel

IGSN: SIO0000Q4	
IGSN:	SIO0000Q4
Sample Name:	SHOW03HO-0180
Other Name(s):	Core
Sample Type:	Core
Parent IGSN:	Not Provided
Description	
Material:	Sediment
Classification:	Not Provided
Field Name:	Not Provided
Description:	EXHAUST_CODE : x
Age (min):	Not Provided
Age (max):	Not Provided
Collection Method:	Not Provided
Collection Method Description:	Not Provided
Size:	CORED_LENGTH : 0;CORED_DIAM : 4
Geological Age:	AGE : unknown
Geological Unit:	Not Provided
Comment:	STORAGE_METH : refrigerated
Purpose:	Not Provided
Geographic	
Latitude (WGS84):	22.87
Longitude (WGS84):	Not Provided
Northing (m) (UTM NAD83):	Not Provided
Easting (m) (UTM NAD83):	Not Provided
Zone:	Not Provided
Vertical Datum:	Not Provided
Elevation Start:	Not Provided
Elevation End:	Not Provided
Nav Type:	Not Provided
Physiographic Feature:	Not Provided
Name Of Physiographic Feature:	Not Provided
Location Description:	Not Provided
Locality:	Not Provided
Locality Description:	Not Provided
Country:	Not Provided
State/Province:	Not Provided
County:	Not Provided
City:	Not Provided
Collection	
Field Program/Cruise:	SHOW-H
Platform Type:	Not Provided
Platform Name:	Not Provided
Platform Description:	Not Provided
Launch Type:	Not Provided
Launch Platform Name:	Not Provided
Launch ID:	Not Provided
Collector/Chief Scientist:	Whitney, W.
Collector/Chief Scientist Detail:	Not Provided
Collection Start Date:	1966-07-06T00:00:00Z
Collection End Date:	Not Provided
Curation	
Current Archive:	Scripts Institution of Oceanography (SIO)
Current Archive Contact Details:	Not Provided
Original Archive:	Not Provided
Original Archive Contact Details:	Not Provided
Related Samples	
Parents:	No Parents
Siblings:	No Siblings
Children:	No Children
Contact Sample Owner	

SESAR IGSN Metadata Profile

NSF-ICF metadata was exported into an excel file and mapped to SESAR's metadata fields.

Next Steps and Conclusions Moving Forward

The process of registering IGSNs for the NSF-ICF has been very promising.

- The matching of ice core metadata derived from our in-house FileMaker server with SESAR's metadata schema has been relatively straightforward.
- Extensive community input is currently being gathered for consensus on vocabulary designation, terminology, and metadata alignment.
- Once complete, we hope that the ice core community will have the opportunity to utilize the established profiles and vocabularies.

Acknowledgements

- This work was funded by the National Science Foundation through an Interagency Agreement (IAA 2041950) with the U.S. Geological Survey.
- Author's salary was partially funded by the National Association of Geology Teachers internship program with the USGS.
- Special thanks to Richard Nunn and Geoffrey Hargreaves of the NSF-ICF, and to Lindsay Powers and Michaela Johnson of the National Geological and Geophysical Data Preservation Program (NGGDP)
- Special thanks to SESAR for their assistance in helping this project come to life.

References

Plomp, E. 2020. Going Digital: Persistent Identifiers for Research Samples, Resources and Instruments. *Data Science Journal*, 19:46, pp. 1–8. DOI: <https://doi.org/10.5334/dsj-2020-046>

Damerow, JE, Varadarajan, C, Boye, K, Brodie, EL, Burrus, M, Chadwick, KD, CrystalOrnelas, R, Elbashandy, H, Alves, RJE, Ely, KS, Goldman, AE, Haberman, T, Hendrix, V, Kakalia, Z, Kemner, KM, Kersting, AB, Merino, N, O'Brien, F, Perzan, Z, Robles, E, Sorensen, P, Stegen, JC, Walls, RL, Weisenhorn, P, Zavarin, M and Agarwal, D. 2021. Sample Identifiers and Metadata to Support Data Management and Reuse in Multidisciplinary Ecosystem Sciences. *Data Science Journal*, 20: 11, pp. 1–19. DOI: <https://doi.org/10.5334/dsj-2021-011>

Medina-Smith, A, Becker, C, Plante, R, Bartolo, L, Dima, A, Warren, J, and Hanisch, R. 2021. A Controlled Vocabulary and Metadata Schema for Materials Science Data Discovery. *Data Science Journal*, 20(1), p.18. DOI: <http://doi.org/10.5334/dsj-2021-018>

Conze, R, Lorenz, H, Ulbricht, D, Elger, K, and Gorgas, T. 2017. Utilizing the International Geo Sample Number Concept in Continental Scientific Drilling During ICDP Expedition COSC-1. *Data Science Journal*, 16: 2, pp. 1–8, DOI: <https://doi.org/10.5334/dsj-2017-002>