Establishment of a National Index of Borehole Information (NIBI): White Paper on Requirements and Recommendations

by

Mojisola A. KunleDare1 and Michaela R. Johnson2

1Delaware Geological Survey via Intergovernmental Personnel Agreement, Newark, DE; 2U.S. Geological Survey, Denver, CO

Contents

[Abstract 3](#_Toc70684876)

[Introduction 3](#_Toc70684877)

[Objectives and Approach 4](#_Toc70684878)

[Timeline 5](#_Toc70684879)

[Envisioning the Index 5](#_Toc70684880)

[Use Cases 6](#_Toc70684881)

[Requirements 7](#_Toc70684882)

[Schema 9](#_Toc70684883)

[Web Interface 11](#_Toc70684884)

[Summary and Conclusions 12](#_Toc70684885)

[Recommendations and Planning for Implementation 13](#_Toc70684886)

[Acknowledgements 13](#_Toc70684887)

[References Cited 14](#_Toc70684888)

[Appendix 14](#_Toc70684889)

[A. NIBI\_List of Surveyed Organizations 14](#_Toc70684890)

[B. NIBI\_Use Cases Compilation 14](#_Toc70684891)

[C. NIBI\_Attributes and Schema Template 14](#_Toc70684892)

Figures and Tables

[**Figure 1. Project Timeline, July to December 2020** 5](#_Toc70684893)

[**Table 1. Examples of compiled NIBI use cases** 7](#_Toc70684894)

[**Table 2. Minimum attributes for NIBI data schema with inclusion classification (required or recommended).** 9](#_Toc70684895)

[**Figure 2. NIBI schema diagram showing the table relationships (1:1 or 1:Many)** 10](#_Toc70684896)

[**Table 3. NIBI schema elements and definitions.** 11](#_Toc70684897)

[**Figure 3. Example of web interface modeled on the GeoLog Locator** 12](https://doimspp.sharepoint.com/sites/GS-NGGDPPBoreholeInventory/Shared%20Documents/General/NIBI_White_Paper_Drafts/WhitePaper_National_Index_of_Borehole_Information_20210505.docx#_Toc70684898)

# Abstract

The National Index of Borehole Information (NIBI) is a collaboration between the U.S. Geological Survey (USGS) and State Geological Surveys to meet current stakeholder and future users’ needs for accessing subsurface data. Currently, borehole inventories (drill core, oil and gas wells, coal, test holes, water wells, etc.) are compiled and locally served in individual databases and repositories. Querying these resources requires a separate search within each application. While the data reside in unique databases many of the components are contextually similar and could be mapped to a common index for enhanced search and discoverability. NIBI will serve as an index platform for the borehole information within these individual databases, providing the connection back to the holding repositories for full access to the resources. This access will facilitate multidisciplinary research in areas such as geology, climate, carbon sequestration, hydrogeology, and energy and minerals resources and lead to improved understanding of the subsurface.

# Introduction

An urgent need exists for the establishment of a National Index of Borehole Information (NIBI), which compiles information about subsurface datasets, including modern and historical data assets and physical samples from state repositories and USGS collections. NIBI will be a centralized index of information (metadata) about existing resources to facilitate search, discovery, access, and re-use of these highly valuable and distributed geological assets. The data itself will continue to be curated and managed at the repository level. The comprehensive search interface will advance discovery of distributed geoscience assets to inform national-scale geospatial tools, increase discovery and access to local natural resources and facilitate subsurface characterization.

Subsurface characterization was identified as one of five common earth science themes that require modern understanding, which is enabled by borehole characterization (The National Academies of Sciences, Engineering, and Medicine, 2020). Subsurface characterization supports multidisciplinary and disciplinary research, including hydrogeologic, climate, geotechnical (e.g., carbon sequestration, substrate identification and classification), and energy and minerals investigations which advance our understanding of earth processes and result in safeguarding national interests such as water security (quality and quantity), climate studies, critical minerals security, infrastructure security and integrity, energy resources and hazards mitigation.

The National Geological and Geophysical Data Preservation Program (NGGDPP) coordinates with USGS Mission Areas and regions to support the preservation of data and geoscience material resulting from USGS work. The Program also administers a Grants Program that provides funds to State Geological Surveys for geological and geophysical data preservation efforts. In collaboration with the NGGDPP, the Earth Mapping Resources Initiative (Earth MRI) supports State Geological Surveys in the preservation of historical assets including borehole data to support compilation of known critical minerals resources on a national scale. NGGDPP also coordinates with the National Cooperative Geologic Mapping Program (NCGMP) and the US GeoFramework Initiative’s focus on seamless multi-dimensional geologic data provision for the United States, for which preserved subsurface data serves as a supporting asset and resource.

The Nation needs improved access to subsurface data and samples (including drill core, oil and gas well logs, minerals, test hole logs, etc.) currently residing in individual databases and repositories. Querying these resources requires a separate search within each application. To address this, the NGGDPP initiated a six-month project, involving federal and state agencies, focused on developing recommendations for the requirements of an index and visioning a geospatial user interface for access. The proposed national product - the National Index of Borehole Information (NIBI) – will be developed as a collaboration between USGS and State Geological Surveys to meet current stakeholder and future users’ needs. It is essential for the State Geological Surveys to be involved in envisioning the role of the index as they are a primary data provider and stakeholder. This white paper describes the recommendations on requirements for the development of the National Index of Borehole Information.

# Objectives and Approach

The objectives of the project were to establish system requirements for the National Index of Borehole Information (NIBI) and provide recommendations for user community interaction (use cases), data compilation and ingestion, and visualization (interactive geospatial web application) compiled through interactions with State Geological Surveys and USGS personnel. A State Geological Survey representative, working with the NGGDPP Associate Program Coordinator, was engaged to define recommendations for establishing the National Index of Borehole Information (NIBI).

Preliminary work was carried out within the Program to investigate examples and best practices for compiling and serving data resources by surveying existing web resources from state, federal, and international agencies who are serving similar data online (Appendix A). Additional resources, such as the [OGC GeoSciML 4.1](http://docs.opengeospatial.org/is/16-008/16-008.html) (2017) standard, were consulted for examples of borehole schema elements and vocabularies. A baseline of common borehole header elements, application use cases, and potential candidates for a web interface were compiled from these efforts.

A National Index of Borehole Information (NIBI) working group was then convened from state and federal agency representatives (see [pp. 13-14](#_Acknowledgements)) to assist with the task of coming up with recommendations for the components and themes necessary to implement the borehole index. Members were chosen to represent a diversity of geography, experience with hosting and serving data, technological and technical expertise, prior and potential participation in the NGGDPP, and experiences with different data types. The working group was charged with helping to define the minimum metadata requirements to describe the resources in the index, providing recommendations on metadata conventions and schema, including identification of controlled vocabularies, refining, and providing additional use cases, and providing suggestions for the web application architecture.

Timeline

The six-month project timeline consisted of tasks that allowed for iterative definition and revision of the elements and recommendations within the Program and in collaboration with the working group as outlined below:



**Figure 1. Project Timeline, July to December 2020**

# Envisioning the Index

NGGDPP leadership, in collaboration with Earth MRI and consultation from state geologists, established the vision for the National Index of Borehole Information (NIBI) as a national resource that will provide an index of borehole locations and the minimum information needed to access samples or data from the borehole from the host repository. Where the term ‘data’ is used, this will include physical materials and samples, and logs associated with the borehole. The index will reside on the USGS National Digital Catalog platform (NDC) and have a flexible, intuitive, web interface with map and text-search capabilities. It is intended that the information about data from the boreholes will continue to be held at the repository level, or on the NDC if they are held there, and NIBI will point to these resources. For our purposes, a borehole was defined as “a narrow hole drilled to establish the nature of, sample, test, or monitor soil, bedrock or contained fluids and gases or for abstraction of water or mineral” (Gaaloul et al., 2018). This allows for the inclusion of all potential sources of subsurface information from shallow to deep. While the index will capture borehole datasets preserved under NGGDPP, with metadata in the NDC, submissions of associated assets not preserved under NGGDPP funding will be encouraged and welcomed. NIBI is intended to be a national resource, therefore other inventories and databases of subsurface resources will be supported for inclusion, for example, coal drill holes, oil and gas wells, water well logs, and geotechnical borings.

Use Cases

In order to make the best recommendations on the components of NIBI and the web application, it was important to envision how the application will be used and maintained. User and contributor community needs for the index were investigated, anticipated, and compiled. Various scenarios were documented using the template below, which examined what action or task a particular user may want to perform in the application, why that action is needed and what the application needs to have for these actions to be possible.

The template: As a (**Role**), I want to (**Action**) so that (**Benefit**) and for this I need (**Metadata elements, Features**) allowed for a wide range of possible applications and required elements to be assessed. The major roles (User, Collection Manager and Administrator), and possible use scenarios described for each, reiterate the vision for the NIBI which is to provide an avenue for discovery of the existence of a borehole in a particular location, the availability of data associated with the borehole and means of access to the data. The scenarios also address maintenance of the index and how necessary updates will be achieved. These scenarios informed considerations for the database schema and structure.

The resulting compilation (Appendix B) provides a guide for developing elements of the index and captures functionalities that can be implemented immediately in the initial iteration of the index, as well as several functionalities which may be added in future iterations of the index. The use cases can and will be used to determine the priorities for pilot projects.

**Table 1. Examples of compiled NIBI use cases**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **As a** | **I want to** | **So that** | **For this I need (metadata elements)** | **For this I need (features)** |
| Scientist/Researcher | see where boreholes exist in my area of interest | I can determine viability of a research idea | Location, total depth, samples/core types and availability, logs, previous analyses, purpose | Easy to use web interface (search, display, etc.) |
| Consultant | see where publicly accessible samples exist | I can create derivative value-added products for clients e.g., regional atlas | Location, total depth, samples/core types and availability, repository PID | Easy to use web interface (search, display, etc.), download functionality |
| Collection Manager | see where similar collections to mine are housed | I can determine where to go for information on managing particular types of collections | Location, total depth, samples/core types and availability, logs, previous analyses, purpose, repository PID | Easy to use web interface (search, display, etc.), download functionality |
| USGS Index Administrator | local repositories generate daily materialized view or csv dump that can be scrapped by NGGDPP borehole dba | local repository data can be formatted on a regular basis to update the national database | a standardized schema map between local and national database objects to allow collection of local data from a standardized view or text data file | behind the scenes standardized data communication protocols |
| User (scientist, researcher, consultant) | download information on relevant boreholes | I can use the data on other platforms /applications | Location, total depth, drill date, samples/core types and availability, logs, previous analyses, purpose | Download functionality (ArcGIS, CSV files, etc.) |
| Data Manager/Database Administrator | View usage statistics and user queries | Identify areas of user interest and ways to streamline data delivery |  | Administrator dashboards via database management software or similar to ArcGIS online application statistics |

Requirements

The NIBI will be an index of borehole locations and provide the information necessary to access data or samples from a borehole from its home repository. The data will be held at the repository level, as such, many of the database attributes can continue to be stored at the local level. The minimum attributes required to identify each borehole (Identifiers), provide its location (Location), and provide data availability and access information and other borehole information (Ancillary Information), which would facilitate its inclusion in the index, were compiled.

**IDENTIFIERS:** It was determined that the primary identifier (Unique Borehole Identifier, UBI) used for identifying a borehole in NIBI would be the main means of linking with the host database. NIBI development will examine establishment, implementation, and use of persistent identifiers (PID) for boreholes and repositories. The borehole primary identifier will identify the borehole to NIBI and be related to a unique internal NIBI borehole identifier (e.g., hash of repository PID (ROR) and borehole PID (e.g., International Geo Sample Number (IGSN)). The NIBI repository ID will utilize a PID such as, the Research Organization Registry identifier (ROR), by which entries will be associated with their holding entity (i.e., repository). Contributing repositories will be encouraged to implement the use of RORs. Additional identifiers associated with the borehole within host repositories can be stored in the index alongside the primary to facilitate searches. This attribute will be required for inclusion in NIBI.

**LOCATION:** It was determined that to facilitate interoperability, a standard must be prescribed for location data. For inclusion in the index, the borehole location will be submitted as Latitude and Longitude in decimal degrees (DD) with the World Geodetic System 1984 (WGS84) datum. Location data not already in this standard will need to be converted for inclusion. Since NIBI will link back to local repositories, original location data and information such as accuracy and computation methods will continue to be available to users from those sources. This attribute will be required for inclusion in NIBI.

**ANCILLARY INFORMATION:** It was determined that for efficiency, the minimum information required to find out what data are available on a borehole, and where and how to access these, will be held in the index as links to the local resources or to the repository in cases where data are not available online. Mechanisms will be put in place for regular updates to the index by holding entities as more data are made available online. Minimum complementary information such as Drill Date, Purpose, etc. will be included in the index so that users can quickly assess if the borehole will be of interest and require contacting the relevant repository for further information and access.

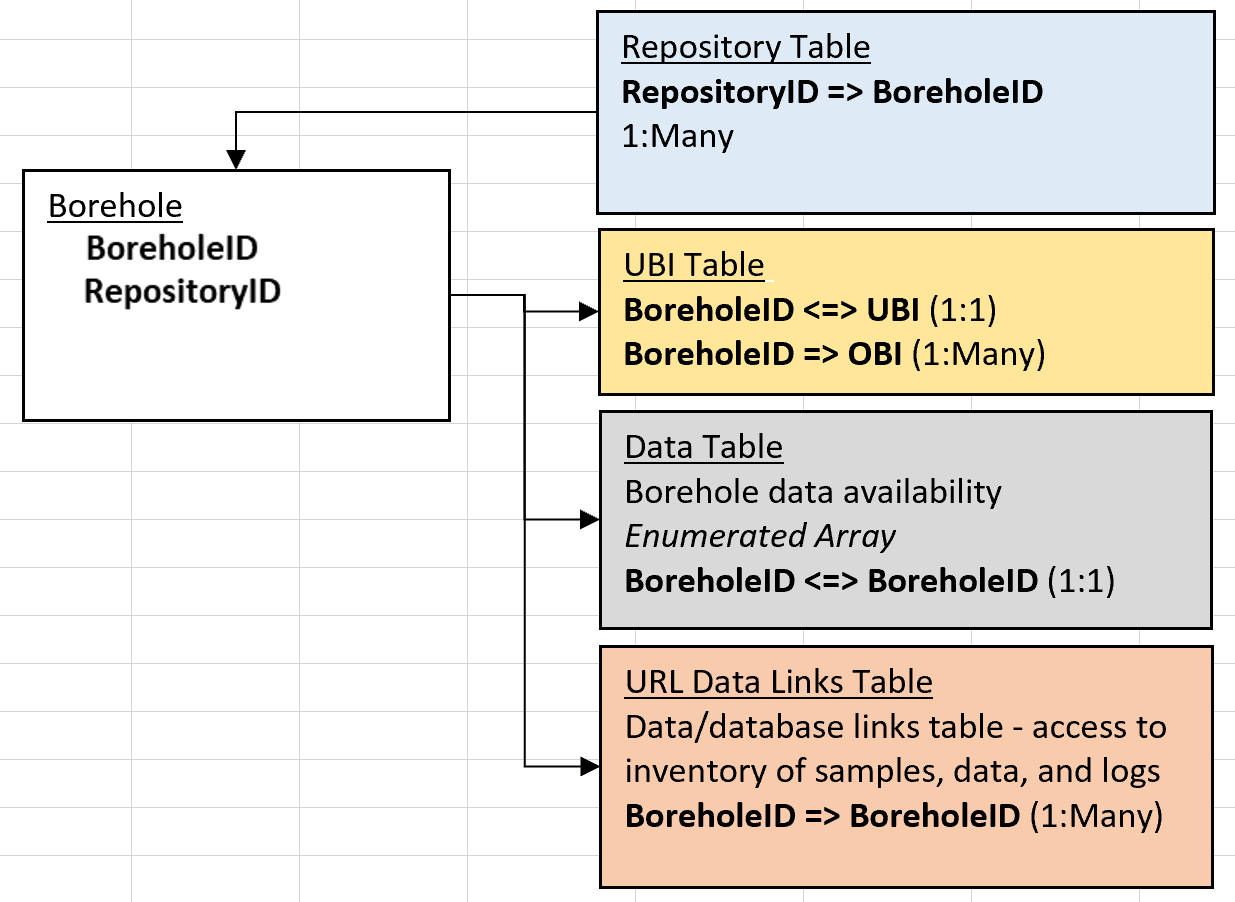
Discussions about, and rankings of, the initial compiled attributes within the working group were used to condense the list into a minimum set of required and recommended attributes with the identifiers and location attributes being required, as discussed above. The working group recognized that in order to capture the diversity of data types, it was essential to not be overly restrictive and as a result there are very few required attributes. The working group also compiled controlled vocabularies to accompany the attributes, where necessary (e.g., Borehole Purpose). The condensed requirements (minimum attributes) and recommended vocabulary are compiled in Appendix C, the main categories are shown in Table 2.

**Table 2. Minimum attributes for NIBI data schema with inclusion classification (required or recommended).**

|  |  |  |  |
| --- | --- | --- | --- |
| **Attribute** | **Classification** | **Attribute** | **Classification** |
| NIBI BOREHOLE ID | Required | CORES | Recommended |
| NIBI REPOSITORY ID | Required | CUTTINGS | Recommended |
| UBI | Required | SAMPLES | Recommended |
| OBI | Recommended | TESTS/ANALYSES | Recommended |
| LATITUDE DD WGS84 | Required | PHOTOS | Recommended |
| LONGITUDE DD WGS84 | Required | THIN SECTIONS | Recommended |
| BOREHOLE NAME | Recommended | REPORTS | Recommended |
| ELEVATION - NAVD88 | Recommended | SEISMIC DATA | Recommended |
| TOTAL DEPTH\_M | Recommended | **Log Category** | |
| DRILL DATE | Recommended | ACOUSTIC | Recommended |
| PURPOSE | Recommended | CALIPER | Recommended |
| BOREHOLE URL | Recommended | COMBINATION | Recommended |
| OFFSHORE | Recommended | COMPOSITE | Recommended |
| REPOSITORY NAME | Recommended | DIRECT PUSH | Recommended |
| REPOSITORY URL | Recommended | ELECTRIC | Recommended |
|  |  | ELECTROMAGNETIC | Recommended |
|  |  | FLUID | Recommended |
|  |  | INTERPRETED | Recommended |
|  | **Identifier** | LITHOLOGY | Recommended |
|  | **Location** | NUCLEAR | Recommended |
|  | **Ancillary Information** | OPTICAL | Recommended |
|  |  | WELL CONSTRUCTION | Recommended |

Schema

The recommendations for a database schema incorporating the compiled requirements, which were converted to required and recommended attributes, and the suggested controlled vocabulary are presented in the Draft Schema (Appendix C). The Draft Schema was finalized with the assistance of a sub-group of the working group with expertise and experience in database schemas. It is based on the minimum attributes needed to adequately represent a borehole’s three main elements in the index. The Identifier and Location fields are required for each borehole while the Ancillary information fields are recommended. The recommended NIBI schema has 1:1 and 1:Many relationships between the Borehole ID table and ancillary data tables. The use of JSON enumerated array vs. individual Boolean fields was adopted as it can handle many levels of hierarchical data and allows for use of picklists based on the suggested controlled vocabulary. The recommended schema structure also allows for pre-determining which field will be searchable in text searches (Figure 2; Table 3)



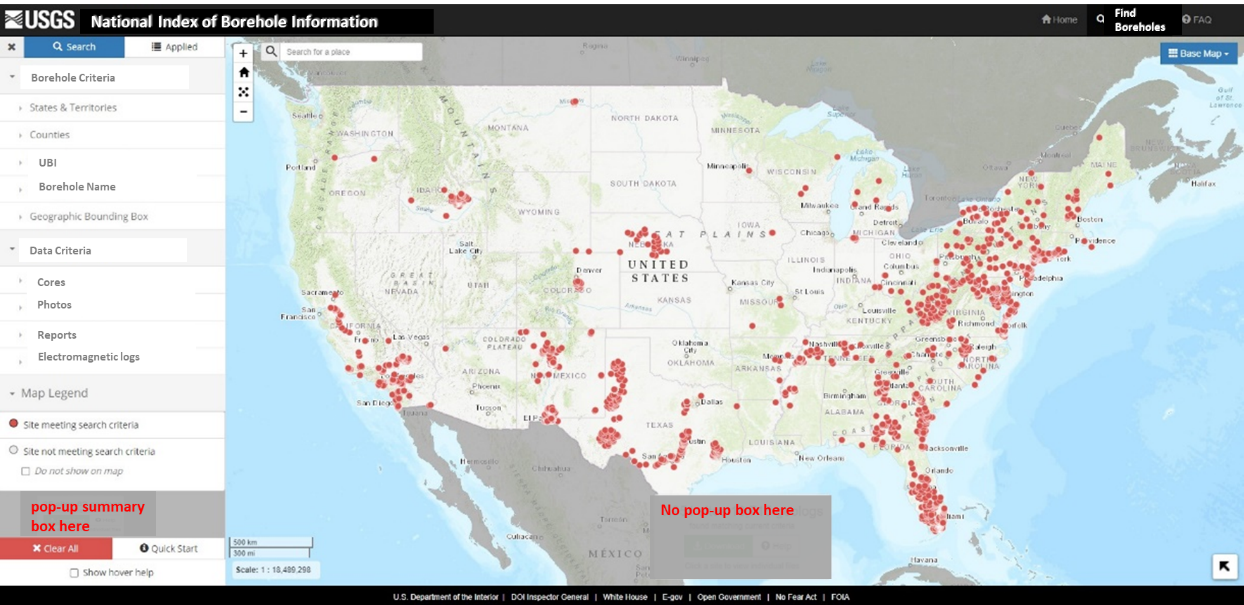
**Figure 2. NIBI schema diagram showing the table relationships (1:1 or 1:Many)**

**Table 3. NIBI schema elements and definitions.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ElementName** | **Table** | **Definition** | **Attribute** | **DataType** | **IsNullable** |
| **NIBI BOREHOLE ID** | borehole | Unique borehole ID internal to NIBI (e.g., hash of repository PID and and UBI) | BoreholeID | integer/text? | NO |
| **NIBI REPOSITORY ID** | borehole | Repository PID for organization or geological survey that compiled this record (Source of borehole information) | RepositoryID | integer/text? | NO |
| UBI | borehole | Unique Borehole Identifier (borehole ID - primary ID for identifying borehole in host databases and means of link with NIBI) | UBI | text | NO |
| OBI | borehole | Other Borehole Identifiers - additional identifiers for enhanced search, for example IGSN, API, UWI, USGS station identifier, Local ID, etc. | OBI | text | YES |
| BOREHOLE NAME | borehole | Name of borehole, for example, "Coles", 'Coles-1", "Coles #1" etc. | BoreholeName | text | YES |
| LATITUDE DD WGS84 | borehole | Latitude, in decimal degrees with WGS84 horizontal datum | LatWGS84 | numeric | NO |
| LONGITUDE DD WGS84 | borehole | Longitude, in decimal degrees with WGS84 horizontal datum | LonWGS84 | numeric | NO |
| ELEVATION - NAVD88 | borehole | Elevation of borehole collar or origin, in meters, referenced to NAVD88 | ElevNAVD88 | numeric | YES |
| TOTAL DEPTH\_M | borehole | Total drilled (measured) depth of the hole, in meters | TotalDepth | numeric | YES |
| DRILL DATE | borehole | Date drilling operations started (YYYY, YYYYMM, YYYYMMDD) | DrillDate | date | YES |
| PURPOSE | borehole | Purpose of borehole | Purpose | domain - picklist | YES |
| BOREHOLE URL | borehole | Uniform Resource Identifier, direct link to URL, application, repository URL, etc. hosting borehole data | BoreholeURL | text | YES |
| OFFSHORE | borehole | Flag for offshore location for borehole | Offshore | yes/no | YES |
| BOREHOLE DATA | borehole | Searchable list of data and samples available for the borehole | BoreholeData | enumerated array | YES |

Web Interface

The NIBI will reside on the National Digital Catalog platform, while the data accessed by the index will be held at the state and/or USGS Science Center level. The index is envisioned to have a flexible, intuitive, web interface with map and text-search capabilities. It is currently modelled to have the look and feel of the [GeoLog Locator](https://webapps.usgs.gov/GeoLogLocator/#!/search), with features specific to NIBI.



**Figure 3. Example of web interface modeled on the GeoLog Locator**

Solutions for interaction between the index and the local repository databases were explored and recommendations and suggestions for different aspects are as follows:

* RESTful APIs
* A transformation tool or a link to one, to be provided on the platform for location and elevation data to be converted to the prescribed datum standards
* Incorporation of existing USGS capabilities and services that can spatially select point data based on underlying geographical divisions (e.g., PLSS, Quads, County, State, etc.) without having to build that into the index.

# Summary and Conclusions

The USGS National Geological and Geophysical Data Preservation Program (NGGDPP) and Earth Mapping Resources Initiative (Earth MRI) interests in the establishment of a national borehole index led to a six-month project to develop recommendations for a National Index of Borehole Information (NIBI). The index will host data preserved under NGGDPP on the National Digital Catalog (NDC), although contributions of associated assets not under NGGDPP will be encouraged and welcomed. With the assistance of a working group consisting of representatives of federal and state agencies who are primary data providers and stakeholders, recommendations on the contents and structure of the index were established. This is a necessary first step that confirms the need for the index and the viability of the proposed index as envisioned.

Recommendations and Planning for Implementation

The next steps toward implementation of the National Index of Borehole Information based on recommendations from the current working group as determined through the evaluation of user community needs are as follows:

* The minimum requirements for inclusion in the index are a Unique Borehole Identifier, Location (Latitude/Longitude, decimal degrees, WGS84) and the details of the holding repository (name and/or URL),
* Determine data ingestion method(s),
* Determine mechanism and timing of updates,
* Collaboration with larger groups within USGS with similar platforms to align objectives and features
* Coordination with other national and global groups with similar platforms
* Start communications with prospective developers
* Produce a candidate application for test cases and pilot studies

The conceptual drivers for NIBI were the among the founding principles for the NGGDP Program, to provide a national catalog of geological and geophysical resources. We recognize that there are many national and global efforts with similar and/or complementary objectives, such as Open Geospatial Consortium ([OGC](https://www.ogc.org/)) groups (GeoScience Data Working Group (GeoScienceDWG) and Commission for the Management and Application of Geoscience Information (CGI)-OGC GeoSciML) , the International Union of Geological Sciences Deep-Time Digital Earth ([DDE](https://www.iugs.org/dde)) and the National Science Foundation EarthCube [GeoCODES](https://www.earthcube.org/geocodes). The next steps will include engagement with these groups to better understand existing capabilities, lessons learned, and avoid duplication of efforts, which may help shorten the time frame to implementation.

# Acknowledgements

The National Index of Borehole Information (NIBI) working group was convened from state and federal agencies. The members are gratefully acknowledged for the time, energy, passion, and ideas that they committed to the task. Their expertise and perspectives shared in a collegial, interactive environment made the project successful.

Many thanks to the working group members (\* denotes Schema sub-group members):

1. Denise Hills (Alabama)
2. \*Kurt Johnson (Alaska)
3. \*Michelle Ladle (Florida)
4. \*Rick Langel (Iowa)
5. Harvey Thorleifson (Minnesota)
6. Heather Arends (Minnesota)
7. Cheryl Seeger (Missouri)
8. Beverly DeJarnett (Texas)
9. Clint Edrington (NOAA - NCEI)
10. \*Eli Mlawsky (Nevada)
11. Chrissi Wood-Smith (Virginia)
12. William Lasseter (Virginia)
13. Daniel Eungard (Washington)
14. Kristine Zellman (USGS -GECSC)
15. \*Mojisola KunleDare (Delaware/USGS Detail)
16. \*Michaela Johnson (USGS - NGGDPP)

The following contributed to discussions on the Schema:

1. Marcella McIntyre (Alabama)
2. Christopher Ramey (Alaska)
3. Aaron Averett (Texas)

The following provided guidance and insights from USGS programs:

1. Donald Sweetkind (USGS - GECSC)
2. Candice Hopkins (USGS – NGWMN)
3. Warren Day (USGS – Earth MRI)
4. Lindsay Powers (USGS- NGGDPP)

Dr. William B. Harrison, III (Michigan), Beverly DeJarnett (Texas), and the USGS Core Research Center are also gratefully acknowledged for their compiled repository lists providing the foundation for the survey of repository databases and web applications.

# References Cited

1. National Academies of Sciences, Engineering, and Medicine, 2020. A Vision for NSF Earth Sciences 2020-2030: Earth in Time. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25761>.
2. Gaaloul N., Eslamian S., Ostad-Ali-Askari K., 2018 Boreholes. In: Bobrowsky P.T., Marker B. (eds) Encyclopedia of Engineering Geology. Encyclopedia of Earth Sciences Series. Springer, Cham. <https://doi.org/10.1007/978-3-319-73568-9_31>
3. OGC Geoscience Markup Language 4.1 (GeoSciML), 2017, Open Geospatial Consortium, <http://docs.opengeospatial.org/is/16-008/16-008.html>
4. U.S. Geological Survey, 2020, GeoLog Locator, <https://doi.org/10.5066/F7X63KT0>.

# Appendix

1. [NIBI\_List of Surveyed Organizations](https://doimspp.sharepoint.com/:x:/s/GS-NGGDPPBoreholeInventory/EcPfYCWxGZpPtTElRckxb8oBebbCsB1a7ej3BcO5nyweVg?e=d2bOlF)
2. [NIBI\_Use Cases Compilation](https://doimspp.sharepoint.com/:x:/s/GS-NGGDPPBoreholeInventory/EZsjvhMdPipBmw-IQIx18bUBVZAtKKE20FNLC7LgMA4Yxw?e=IHNS1o)
3. [NIBI\_Attributes and Schema Template](https://doimspp.sharepoint.com/:x:/s/GS-NGGDPPBoreholeInventory/ERBER8nTkpZCia_ATA81qPABA1_kaLnxlEVM3_nOsuukYg?e=TB9FFm)