The Index to Marine and Lacustrine Geological Samples

Community access to common information on sample collections

Kelly J. Stroker\textsuperscript{1}, Jennifer Jencks\textsuperscript{2}

\textsuperscript{1}Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder, Colorado

\textsuperscript{2}NOAA National Centers for Environmental Information (NCEI), Boulder, Colorado

2017 Meeting of the Curators of Marine, Lacustrine, and Geological Samples
Oregon State University Marine and Geology Repository
April 25th – 27th 2017, Corvallis, Oregon
Data discoverability

1. Where do geological repositories fit in Rolling Deck to Repository (R2R)?
   - Bob Arko’s talk

2. Moving sample collections and repositories into the digital age
   - Where we are and where we are going or should go?

3. Is there a need for common data portals: How to improve online discovery?
   - IMLGS is one portal for marine and lacustrine sample information (metadata and in some cases data)
NGDC to NCEI

• In early 2015, NOAA’s data centers merged

• This includes the former National Geophysical Data Center, National Oceanographic Data Center, and National Climatic Data Center

• Enhance integration across science disciplines

• Includes the **World Data Centers for Geophysics, Oceanography, Paleoclimate, and Meteorology** and the International Hydrographic Organization’s Data Center for Digital Bathymetry

• Build upon the full spectrum of weather/climatic, oceanographic, coastal, and geophysical products and services that the National Data Centers previously delivered
National Centers for Environmental Information (NCEI)

• Responsible for hosting and providing access to one of the most significant archives on Earth, with comprehensive oceanic, atmospheric, and geophysical data

• From the depths of the ocean to the surface of the sun and from million-year-old sediment records to near real-time satellite images

Revealing the Past, Interpreting the Present, and Informing the Future
Commitment to Long-term Data Stewardship

We’re in it for the long haul, preserving data for your great-grandchildren

Stewardship in Service of Science
The IMLGS

The Index to Marine and Lacustrine Geological Samples (IMLGS) is, at it’s core, community access to common information on sample collections.

• The IMLGS is a community designed and maintained resource enabling researchers to locate and request sea floor and lakebed geologic samples archived by partner institutions.

• The Index is based on core concepts of community oversight, common vocabularies, consistent metadata and a shared interface.

• The Curators Consortium, now international in scope, meets at partner institutions biennially to share ideas and discuss best practices.

• NCEI serves the group by providing database access and maintenance, a list server, digitizing support (resource dependent) and long-term archival of sample metadata, data and imagery
What is in the Index?

• Over three decades, participating curators have performed the laborious task of creating and contributing metadata for over 209,000 sea floor and lake-bed cores, grabs, and dredges archived in their collections.

• Some partners use the Index for primary web access to their collections while others use it to increase exposure of more in-depth institutional systems.

• The IMLGS has a persistent URL/Digital Object Identifier (DOI), as well as DOIs assigned to partner collections for citation and to provide a persistent link to curator collections.
The Index to Marine and Lacustrine Geological Samples
Community access to common information on sample collections

209,000 Samples
(a core, grab, dredge, or drill hole)

> 6,900 additions since the June 2015 Curators meeting
How Can You Access the IMLGS?

- Curator Home page with access to both text and map based interfaces: https://www.ngdc.noaa.gov/mgg/curator/curator.html
Web Services

• Geospatially-enabled relational database

• ArcGIS Interactive Map:

• ArcGIS REST Services
  http://maps.ngdc.noaa.gov/arcgis/rest/services/
  ✓ Supported Internet Protocols: JSON, SOAP, WMS, WFS
  ✓ Supported interfaces: REST, SOAP, SiteMap, Geo SiteMap
  ✓ Supported operations: Export Map, Identify, Find, Return Updates

• Metadata Catalogs: Geoportal, GCMD, Data.gov

• Linked Data interface maintained by the Rolling Deck to Repository R2R

The IMLGS has a persistent URL/Digital Object Identifier (DOI)
✓ JSP Text/Forms Web Interface
✓ ArcGIS REST Services & Interactive Map
✓ Linked Data Interface maintained by R2R (need to fix)
✓ GRID Arendal annual GoogleEarth KMZ
Since the April 2015 Curators Meeting:

- Carla Moore retired from Federal service after 40 years!
- Minted DOIs for each agency dataset
- A complete inventory of sample data in the IMLGS for each repository was sent out at the end of 2015
- Data management for IMLGS and IODP archive in the hands of Kelly Stroker and Jennifer Jencks
- No sustained funding plan currently in place since Carla’s retirement
- IMLGS is a priority NOAA/NESDIS dataset
Since the April 2015 Curators Meeting
6,918 Cores, Grabs, Dredges & Drill Holes added/in review

<table>
<thead>
<tr>
<th>Repository</th>
<th># Samples/Holes Added</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARF/FSU</td>
<td>308</td>
</tr>
<tr>
<td>BOSCORF</td>
<td>271</td>
</tr>
<tr>
<td>BPCRR</td>
<td>2,776</td>
</tr>
<tr>
<td>ECS</td>
<td>2</td>
</tr>
<tr>
<td>France</td>
<td>2</td>
</tr>
<tr>
<td>IODP</td>
<td>44</td>
</tr>
<tr>
<td>LacCore</td>
<td>419</td>
</tr>
<tr>
<td>LDEO</td>
<td>208</td>
</tr>
<tr>
<td>OSU/OER</td>
<td>68</td>
</tr>
<tr>
<td>OSU</td>
<td>107</td>
</tr>
<tr>
<td>SIO</td>
<td>1</td>
</tr>
<tr>
<td>SOEST</td>
<td>743</td>
</tr>
<tr>
<td>URI</td>
<td>450</td>
</tr>
<tr>
<td>USGS</td>
<td>1,453</td>
</tr>
<tr>
<td>WHOI</td>
<td>62</td>
</tr>
</tbody>
</table>
DOI for each Repository Collection

Archive of Geosample Data and Information from the Columbia University Lamont-Doherty Earth Observatory (LDEO) Lamont-Doherty Core Repository (LDCR)

The Lamont-Doherty Core Repository (LDCR), operated by the Lamont-Doherty Earth Observatory (LDEO) of Columbia University is a partner in the Index to Marine and Lacustrine Geological Samples (IMLGS) database, contributing information to the IMLGS to help researchers discover geological samples curated in their facility. The partner repository also sends related data, documents, and imagery to NCEI for long-term archive, but the originating institution is the definitive source of information related to their sample collection. The LDCR contains one of the world's most unique and important collection of scientific samples from the deep sea. Sediment cores from every major ocean and sea are archived at the Core Repository. Materials available through the NCEI archive (other than core metadata) are core photos and descriptions, both of which were done at the time a core was opened; information/photos of what a core is like now are available upon request. If you would like to request material from LDCR holdings or want to know what other services are offered (grain size, coulometry, MST, etc.), please check the LDCR website for more information.

Show Less
Geosample Deep (Offline) Archive

- 12TB, >1.3M files
- NARA standards/OAIS framework
- ISO standard documentation
- Periodic migration to new archive media
- DOIs: 16 assigned (e.g., IMLGS, DSDP, ODP, IODP archives)

- In progress: API / Public access to all 12TB of geosample data in the offline archive
## API Access

### International Scientific Drilling, 1967-Present (DSDP, ODP, IODP)

<table>
<thead>
<tr>
<th>Request Summary</th>
<th>NOS</th>
<th>DEM</th>
<th>SOD</th>
<th>Multibeam Surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **SOD IDs**: 0127c914-1a6e-45d...
- **Start Year**: Not specified
- **End Year**: Present
- **Ships**: All
- **Legs**: All
- **Holes**: All
- **Subfolders**: core_photos

**Top**: Not Specified  
**Left**: Not Specified  
**Bottom**: Not Specified  
**Right**: Not Specified

### Large Volume Data

- **High-resolution Core Photographs**
- **~20 Data types as text files**
- **ISO metadata (hole/datatype)**
- **For years, the community has wanted API access to files in the archive**
- **We are in the process of setting this up through NEXT (NCEI Extract System)**
- **We will work on all geology after SOD**
Goals and Future Work

1. An up to date IMLGS database

2. Public access to all geosample files in the offline archive via an API

3. More filtering options for the ArcGIS map viewer

4. Adding all IGSNs for more samples in the IMLGS and coordinating with SESAR

5. A completed Submission Agreement between IEDA & NCEI, resulting in improved interoperability between SESAR and the IMLGS

Most important:

1. What do curators want? I would like to hear from you!
Questions/comments?

Kelly.stroker@noaa.gov