NGGDPP FY 2008 Final Technical Report

Name of the State Geological Survey: Alaska Division of Geological & Geophysical Surveys

Award Number: 08HQGR0132

Project Title: Preparation and submission of catalog metadata for Alaska site-specific geologic data and samples

Principal Investigator(s): Susan Seitz, Analyst/Programmer IV
Alaska Division of Geological & Geophysical Surveys
3354 College Road
Fairbanks, AK 99709
Phone: 907-451-5052
Fax: 907-451-5050
Email: susan.seitz@alaska.gov

Authorized Institutional Representative: Rodney A. Combellick, Deputy Director
Alaska Division of Geological & Geophysical Surveys
3354 College Road
Fairbanks, AK 99709-3707
Phone: 907-451-5007
Fax: 907-451-5223
Email: rod.combellick@alaska.gov

Project Term: September 1, 2008 to August 31, 2009

Technical Report Date: November 25, 2009
Abstract

During the FY 2008 phase of the National Geological and Geophysical Preservation Program (NGGDPP), the Alaska Division of Geological & Geophysical Surveys (DGGS) directly addressed two of the program priorities: improving upon the state’s digital infrastructure and creating site-specific metadata for individual items in geologic data collections. Our database development team created site-specific metadata for three DGGS geologic data collections: (1) data reports on core samples that have been borrowed from the Alaska Geologic Materials Center (GMC) for analysis, (2) geochemical analyses of rock, soil, and stream-sediment samples collected during projects involving DGGS geologists, and (3) geochronology analyses of samples collected during projects involving DGGS geologists. In order to serve the metadata records to the National Digital Catalog, DGGS enhanced its digital infrastructure by implementing a Web Feature Service (WFS) interface. The DGGS WFS allows the site-specific metadata records to be harvested and synchronized automatically by the National Catalog system, thus freeing DGGS staff members of the manual task of uploading data to an additional database on a regular basis. The DGGS database is continually updated via existing business processes with quality control mechanisms programmed to verify the raw data of the sample collections during bulk-uploading processes or within the data entry interfaces to the database.
# Table of Contents

Abstract ........................................................................................................................................... 2  
Table of Contents ............................................................................................................................ 3  
Introduction ..................................................................................................................................... 4  
Prior Work in Recovering, Archiving, and Cataloging Data .......................................................... 4  
FY 2008 Project Scope ................................................................................................................... 5  
Methodology ................................................................................................................................... 5  
Results ............................................................................................................................................. 6  
Conclusion ...................................................................................................................................... 7  
Appendix A: Summary of DGGS Data Preservation Progress ....................................................... 8
Introduction

The Alaska Department of Natural Resources, Division of Geological & Geophysical Surveys (DGGS), has as its statutory mission the responsibility for collecting, archiving, managing, and disseminating geological and geophysical data on the subsurface energy resources, mineral resources, and geologic hazards of the state. DGGS maintains its own databases of this information on servers at its office in Fairbanks, and a collection of nonproprietary rock samples at the Alaska Geologic Materials Center (GMC) in Eagle River representing more than 12 million linear feet of exploration and production drilling. DGGS networks with other state agencies that use these data and that archive related datasets, some of which are proprietary, including the Division of Oil & Gas (DOG), Division of Mining, Land & Water (DMLW), and the Alaska Oil & Gas Conservation Commission (AOGCC). Samples from these collections are examined frequently by private exploration companies and state agencies as a critical step in resource development and management.

Prior Work in Recovering, Archiving, and Cataloging Data

The State of Alaska received federal funds through congressional appropriations to the U.S. Geological Survey (USGS) and Bureau of Land Management (BLM) between FY 1998 and FY 2004 to recover, archive, catalog, and make publicly accessible minerals-related geologic data and publications that were in danger of being lost or destroyed due to the cancellation of federal minerals programs such as the U.S. Bureau of Mines, and the closing of many private mineral-exploration companies over the past few decades. Through the Minerals Data and Information Rescue in Alaska (MDIRA) program, DGGS has made available to the public, via the internet, many geologic reports, maps, and data that were previously unavailable or difficult to find. The following are just a few of DGGS’s notable accomplishments through MDIRA:

- Completed a full data assessment focusing (but not exclusively) on non-fuels solid-mineral resources data and samples.
- Led a cooperative program to build and maintain a federal/state interagency web-accessible Alaska geoscience bibliography that is geographically indexed and links to online repositories.
- Scanned all previously published DGGS reports and maps.
- Scanned all pre-digital USGS publications and maps with areal subject material relating to Alaska geology.
- Built a public web-accessible catalog for searching and accessing the scanned maps and reports.
- Documented available DGGS digital GIS data with FGDC-compliant metadata.
- Designed and implemented an enterprise data system to catalog and house digital geologic data and information for Alaska.
- Designed and built a database-driven website to distribute DGGS GIS data via the internet.
- Recovered geochemical datasets published in multiple DGGS reports and integrated them into a web-accessible database system.
- Recovered geochronological data from published and unpublished sources and began integrating them into an electronic database that will be web accessible.
- Began cataloging and archiving field project files and unpublished geologic compilations by DGGS and predecessor agencies.
- Began indexing diverse surface and subsurface industry geologic materials housed at the GMC and integrating the index into a web-accessible database.
DGGS requested and received funding under the FY 2007 National Geological and Geophysical Data Preservation Program (NGGDPP) Program Announcement No. 07HQPA0018 to provide collection-level information on its data holdings. DGGS completed evaluations for twelve of its geologic data and sample collections, and entered detailed information about each one into the NGGDPP online inventory system. See the summary table of DGGS collections in Appendix A.

DGGS also received funding under NGGDPP (Order No. 07HQSA0115) to participate in the design and testing of the National Digital Catalog. DGGS helped establish standards for the content, quality, and consistency of metadata describing digital data and physical samples; protocols for searching, transferring, and presenting metadata; clear and consistent ways to explain how users can request access to physical materials; and the extent to which the user interfaces convey branding or meet other organizational needs.

**FY 2008 Project Scope**

In FY 2008, DGGS received funding under NGGDPP (Order No. 08HQPA0015) to implement a Web Feature Service (WFS) interface to deliver site-specific metadata files to the National Digital Catalog for six of its geologic data collections: (1) core samples and drill cuttings stored at the Alaska Geologic Materials Center (GMC), (2) glass slide collection of processed samples at the GMC, (3) data reports on core samples that have been borrowed from the GMC for analysis, (4) geochemical analyses of rock, soil, and stream-sediment samples collected during projects involving DGGS geologists, (5) geochronology analyses of samples collected during projects involving DGGS geologists, and (6) hard-rock surface samples collected by DGGS staff.

**Methodology**

In order to serve the metadata records to the National Digital Catalog, DGGS improved upon the agency’s digital infrastructure by implementing a Web Feature Service (WFS) interface. The database development team installed GeoServer, a Java-based, Open Source server software that allows publishing of geospatial data to the web, on DGGS’s web server to provide the WFS interface. GeoServer utilizes the existing Tomcat application server as its engine and the division-wide Oracle database as its data source.

In the next step, DGGS created its first spatial objects in production mode. We added a geometry column to tables containing point data, created spatial indexes, and transformed stored, sample location coordinate pairs to the proper SDO_GEOMETRY value in North American Datum 1983. The transform functions utilized in these scripts required an upgrade of our Oracle database. A database trigger has been created such that when new records are added to the point data tables, the SDO_GEOMETRY column value is automatically created.

For each data collection, we created database views that query the collection’s source data tables to represent the metadata elements in accordance with the Metadata Profile established by the NGGDPP: CollectionID, Title, AlternateTitle, Abstract, DataType, SupplementalInformation, Coordinates, Alternate Geometry, OnlineResource, BrowseGraphic, Date, DatasetReferenceDate, and VerticalExtent. DGGS also included in each database view the spatial object column for the Web Feature Service. Each database view contains metadata records of an individual collection and is registered as a feature type within GeoServer.

DGGS collection information is generated from the database; the unique identifier of each sample or location is translated into a unique and stable URL for each locality in which DGGS stores data. The OnlineResource metadata element provides this dynamic URL to the National Digital Catalog. The link
provides a summary view of all the information currently cataloged for that locality. As information about the locality is added to the DGGS database through other existing business processes (i.e., condition of the physical samples taken at the location, analysis information, and citations), it also becomes available to users of the National Digital Catalog.

The WFS interface allows the site-specific metadata records to be harvested and synchronized automatically by the National Digital Catalog system, thus freeing DGGS staff members of the manual task of uploading data to an additional database on a regular basis. The DGGS database is continually updated and quality checked by existing business processes. DGGS will utilize the same infrastructure to provide future metadata records to the National Digital Catalog.

**Results**

Three collections are currently available as feature types via the WFS:

- Geochemical analyses dataset of rock, soil, and stream-sediment samples collected during projects that involve DGGS geologists (35,496 records and growing).
- Geochronology analyses (radiometric ages) dataset of samples collected during projects that involve DGGS geologists (3,437 records).
- Data reports on core samples that have been borrowed from Alaska’s GMC facility for various types of analyses. These reports are in PDF format and are indexed by a location-based API (Alaska Petroleum Institute) well number or a hard-rock mineral prospect’s Alaska Resource Data File (ARDF) number (1,020 records and growing).

These collections were originally bulk loaded into DGGS’s central database using data-loading scripts with quality-control mechanisms to verify the raw data of the sample collections during the uploading process. Updates to these collections are made as part of our current business process for existing DGGS applications and public websites.

DGGS also proposed to create metadata records for three additional collections during this phase:

- Glass slide collection of processed samples at the GMC indexed by a location-based API (Alaska Petroleum Institute) well number. The collection includes thin sections, palynological slides, grain mounts, and many other types of processed samples (approximately 160,000 records).
- Hard-rock surface samples collection that consists of hand samples collected by DGGS staff during field projects (projected 100,000 records).
- Core samples and drill cuttings from Alaska’s oil and gas exploration, indexed by API number and housed at Alaska’s GMC (approximately 38,000 records; multiple samples per core).

These three collections are part of the existing GMC inventory catalog, which is currently undergoing integration with the DGGS central database to link sample localities, analysis data, and their physical storage information. DGGS anticipated this data integration would be complete before the end of the FY 2008 NGGDP grant period. However, due to the complexity of the project and staff turnover, the estimated timeframe for migrating GMC data collections into the central repository has been extended into FY 2009. DGGS is continuing this project initiative under State General Funds and will create metadata records for these collections once the data have been transferred. The metadata records will then be served to the National Digital Catalog utilizing the same WFS interface as the initial three DGGS collections.
Conclusion

The FY 2008 phase of the National Geological and Geophysical Data Preservation Program has brought DGGS a step farther in one of the agency’s goals of dynamically disseminating geospatial geologic data to the public. We hope to utilize the Web Feature Service for internal staff and public users to actively connect with DGGS geologic data in Geographic Information Systems (GIS) software, as well as a foundation for future map-interface applications that are being developed under separate funding sources. These accomplishments serve both DGGS data preservation objectives and NGGDPP program priorities. The WFS interface allows any updates to the DGGS’s collection metadata records to be synchronized automatically with the National Digital Catalog, without adding tasks for our staff members with ever-increasing responsibilities. In addition, the WFS methodologies developed during this project will facilitate dissemination of metadata records as DGGS catalogs other legacy data collections.

Ensuring that physical data are properly cataloged and archived should be a critical priority for any organization that provides data to the public. Documentation and ensuring data quality for legacy datasets is crucial to make the datasets meaningful and usable. Once these data collections are converted to digital form, become easily searchable, well documented, and organized by dataset or collection type, users can now focus on merging the data into their own projects, spending more time on analysis, and understanding the implications of their scientific data and observations.
### Appendix A: Summary of DGGS Data Preservation Progress

<table>
<thead>
<tr>
<th>Assess Collection</th>
<th>Catalog Items / Records</th>
<th>Locality Metadata</th>
<th>In NGGDP</th>
<th>Secure Storage and Archive</th>
<th>Digitize Data</th>
<th>Available Online</th>
<th>Direct Internet Link to Sample Level</th>
<th>Establish Process for additions</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGGS geochemical data</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>DGGS geochronology</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>P+</td>
<td>NF</td>
</tr>
<tr>
<td>GMC data reports</td>
<td>C</td>
<td>IP</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>IP</td>
</tr>
<tr>
<td>O&amp;G well samples</td>
<td>C</td>
<td>IP</td>
<td>C</td>
<td>RF</td>
<td>C</td>
<td>P+</td>
<td>P+</td>
<td>IP</td>
</tr>
<tr>
<td>GMC processed slides</td>
<td>C</td>
<td>IP</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>P+</td>
<td>IP</td>
</tr>
<tr>
<td>DGGS surface samples</td>
<td>C</td>
<td>IP</td>
<td>80% IP</td>
<td>IP</td>
<td>P+</td>
<td>70% IP</td>
<td>P+</td>
<td>IP</td>
</tr>
<tr>
<td>Minerals core</td>
<td>C</td>
<td>IP</td>
<td>40% IP</td>
<td>P</td>
<td>P</td>
<td>IP</td>
<td>P+</td>
<td>P+</td>
</tr>
<tr>
<td>U.S. BLM/BOM samples</td>
<td>C</td>
<td>C</td>
<td>0% IP</td>
<td>P</td>
<td>P</td>
<td>C</td>
<td>P+</td>
<td>P+</td>
</tr>
<tr>
<td>DGGS petrographic slides</td>
<td>C</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>AVO rock samples</td>
<td>C</td>
<td>IP</td>
<td>IP</td>
<td>NF</td>
<td>NF</td>
<td>IP</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>AVO geochemical data</td>
<td>C</td>
<td>IP</td>
<td>IP</td>
<td>NF</td>
<td>NF</td>
<td>IP</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>AVO processed samples</td>
<td>C</td>
<td>IP</td>
<td>IP</td>
<td>NF</td>
<td>NF</td>
<td>IP</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Porosity and permeability data</td>
<td>P+</td>
<td>RF</td>
<td>RF</td>
<td>RF</td>
<td>RF</td>
<td>RF</td>
<td>RF</td>
<td>RF</td>
</tr>
<tr>
<td>Organic geochemical data</td>
<td>P+</td>
<td>RF</td>
<td>RF</td>
<td>RF</td>
<td>RF</td>
<td>RF</td>
<td>RF</td>
<td>RF</td>
</tr>
<tr>
<td>Petrographic analyses</td>
<td>P+</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Paleontology reports</td>
<td>P+</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>P-</td>
<td>NF</td>
</tr>
<tr>
<td>Sandstone composition reports</td>
<td>P+</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Seal capacity analyses (MICP)</td>
<td>P+</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Apatite fission track analyses</td>
<td>P+</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Energy-related inorganic geochemistry</td>
<td>P+</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Geothermal data</td>
<td>P</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Grain-size analyses</td>
<td>P</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Measured stratigraphic sections</td>
<td>P</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Gravity data</td>
<td>P</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Permafrost depth and ice content</td>
<td>P</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Geo-engineering radiocarbon ages</td>
<td>P</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Slope stability data</td>
<td>P</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>U-Pb detrital zircon analyses</td>
<td>P</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
<tr>
<td>Paleoseismic data</td>
<td>P</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
</tbody>
</table>

**Status:**
- **C:** Complete
- **IP:** In Progress
- **RF:** Requesting Funds
- **P+:** Planned (funded)
- **NF:** Not funded
- **NA:** Not Applicable