

Award No. G10AP00113

National Geological and Geophysical Data Preservation Program

**Washington State Metadata Project:
Continuation of Geotechnical Report Metadata Production**

Final Technical Report

Robert Berwick

Isabelle Sarikhan

Washington Department of Natural Resources

Division of Geology and Earth Resources

PO Box 47007

Olympia, WA 98504-7007

(360) 902-1431 / (360) 902-1425

(360) 902-1785 fax

robert.berwick@dnr.wa.gov / isabelle.sarikhan@dnr.wa.gov

Award Term: July 1, 2010 to June 30, 2011

Submitted: September 30, 2011

Abstract

The Washington State Department of Natural Resources, Division of Geology and Earth Resources, houses many geologic and geophysical data collections, 15 of which have been inventoried in the National Digital Catalog.

In 2009 we selected three collections as high priorities for data preservation, based on their usefulness and perishability: (1) a collection of rock cores, (2) a collection of rock cuttings, and (3) a continually growing collection of geotechnical reports. We created metadata records for 47 boreholes that produced rock cores, 309 boreholes that produced rock cuttings, and 56,302 boreholes represented in 10,000 geotechnical reports, for a total of 56,658 total metadata records for the year 2009 delivery.

The 2010 project objective was a continuation of the work began in 2009 on the geotechnical report collection, creating sample-level metadata and housing this information in a database linked to an interactive mapping application to make it more accessible to the public.

For the year 2010 we produced 28,196 metadata records representing 8,053 geotechnical reports. This information was compiled into XML format in accordance with the National Digital Catalog specifications, and submitted for inclusion in the catalog.

We also launched the Washington State Geologic Information Portal with the Subsurface Geology Information System theme in October of 2010. The Subsurface Geology Information System now hosts 84,509 points representing wells and boreholes found in 18,053 reports from the 2009 and 2010 projects. DGER also procured dedicated storage outside of the DNR firewall to house PDFs linked in the National Digital Catalog and to the geospatial data in the Washington State Geologic Information Portal to allow public access to the reports and data through the internet.

Table of Contents

Abstract.....	1
Introduction	2
Project Goals	3
Methodology.....	3
Results.....	4

Introduction

The Washington State Department of Natural Resources, Division of Geology and Earth Resources (DGER), has identified 15 collections of geologic and geophysical data, which have been inventoried in the National Digital Catalog.

In 2008, David K. Norman, Washington State Geologist and Division Manager for DGER, produced the first draft of the *Geological Data Preservation Plan of Washington*, which states that the primary goals for preserving physical collections are “to ensure no subsurface samples are lost or destroyed from unique or critical holes and to ensure that surface samples are properly recorded and added to the National Catalog.”

With these goals in mind, three collections were chosen for year 2009 because they were most at risk of loss or degradation and held high value geologic and geophysical data. These collections were:

- P901-Rock cores
- P902-Rock cuttings
- P961-Geotechnical reports

For year 2009 of the National Geological and Geophysical Data Preservation Program (NGGDPP), we produced xml metadata records for 47 boreholes that produced rock cores, 309 boreholes that produced rock cuttings, and 56,302 boreholes represented in 10,000 geotechnical reports for a total of 56,658 total metadata records for submission to the National Digital Catalog.

For two of those categories, rock cores and rock cuttings, as much information as was available and met the minimum requirements of the NGGDPP was entered into the National Digital Catalog. Next we began processing our large, ever growing collection of geotechnical reports.

Geotechnical reports are an indispensable source of subsurface geologic and geophysical information, and are produced at considerable expense. Unfortunately, geotechnical reports are generally not made available for long-term use; most local governments commonly dispose of these reports at the conclusion of the project for which the borings were drilled, and geotechnical firms typically keep their records of geotechnical investigations private. A countless number of geotechnical reports submitted to local governments have already been lost. Over the past several years, DGER has made an effort to collect geotechnical reports for borings drilled in Washington to preserve their long-term value. DGER has collected these reports over the years as supporting information for mapping and hazards studies, but has

stepped up efforts to collect as many reports as possible for preservation and access by contacting local jurisdictions and engineering firms and requesting copies of any reports they have. Notably, DGER recently acquired the electronic geotechnical report collection of the Pacific Northwest Center for Geologic Mapping Studies (GeoMapNW); this collection was at risk of loss due to the uncertainty of continued funding for the organization, and preserving this collection became a priority for DGER.

Project Goals

Our focus for the year 2010 was to continue the work began in 2009 on the geotechnical reports (P961), creating sample-level metadata for entry into the National Digital Catalog, and to make this information readily accessible to the public through hyperlinks to scanned documents and through an interactive mapping application. We estimated the number of individual records contained in the reports to be in the neighborhood of 32,000.

The project continued over a one-year period beginning July 1, 2010, and ending June 30, 2011. It was anticipated that the server space for the imagery would be set up, the new subsurface database structure be in place, and the launch of the Washington State Geologic Information Portal be completed in the first quarter. The remaining time would be focused on (1) scanning documents, (2) processing, reviewing, and normalizing data, (3) creating XML format metadata in accordance with the National Digital Catalog specifications, and (4) integrating the new data with the existing data for use in the Subsurface Geology Information System.

Methodology

Much the same methodology was used as in the previous 2009 project. In order to produce metadata suitable for the National Digital Catalog, we extracted information from various parts of the subsurface database and constructed metadata elements that were not represented in the database. We created active links to the PDF files of each scanned report for each borehole record, and links to the online database.

One way the methodology differed from the previous 2009 project was the creation of a spreadsheet for data entry that utilized drop-down menus and pick lists to normalize data as it was being entered into the spreadsheet. This cut down on the amount of data processing necessary later, during the quality control and quality assurance process.

Most of these tasks, and the exporting of the metadata to an XML file formatted for importing into the catalog, were accomplished using Microsoft Excel and a complex series of computer programming scripts that we had previously custom-built for this purpose.

Results

We scanned documents, collected and processed information, and produced and submitted on June 15, 2011, compliant with NGGDPP standards, for 8,053 geotechnical reports, representing 28,196 boreholes. This number is slightly lower than we originally estimated.

Comparison of the borehole estimate to actual accomplishments for the geotechnical reports collection (P961):

Estimated:	Estimated number of boreholes contained in geotechnical reports	32,000
Results:	Number of boreholes (and associated reports) with complete metadata submitted to the National Digital Catalog	28,196 (8,053 reports)

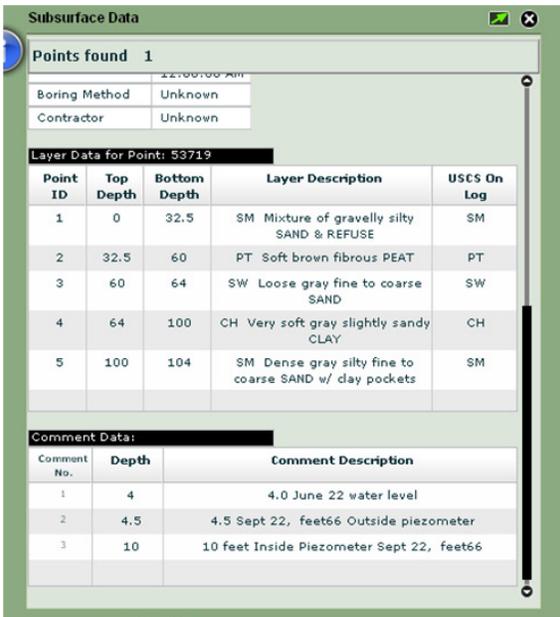
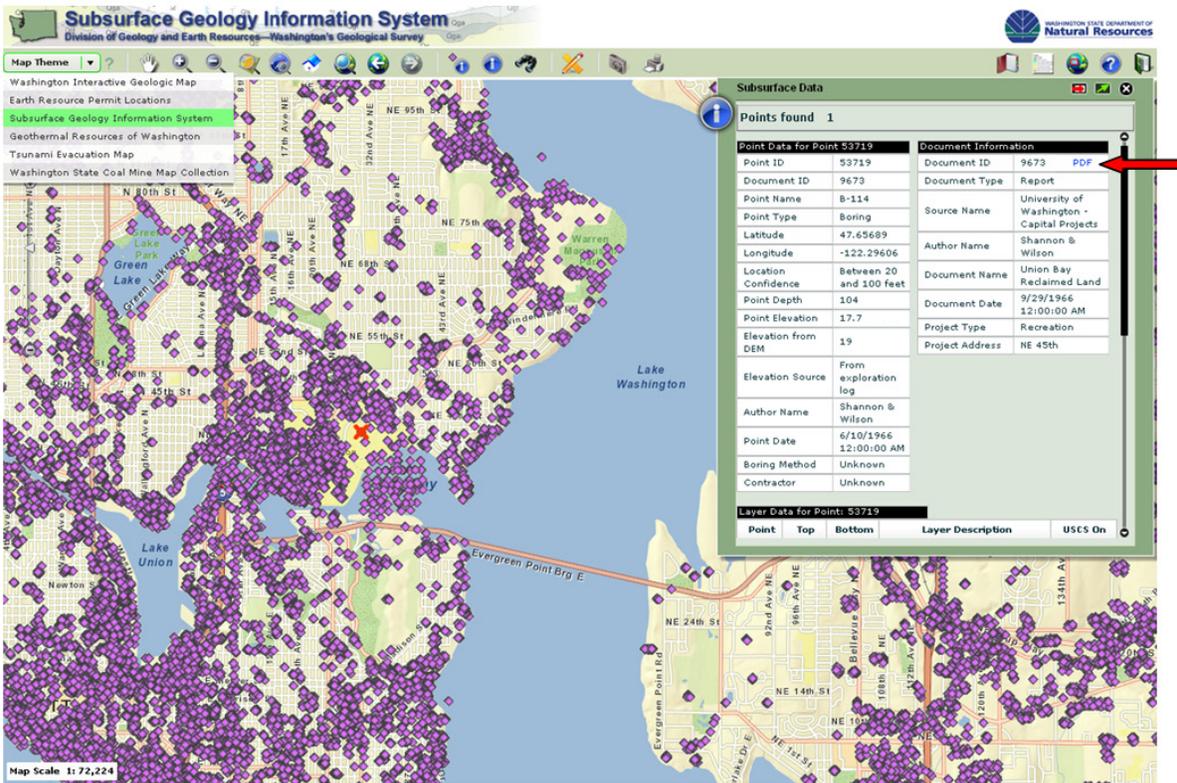
In October of 2010, DGER launched the Washington State Geologic Information Portal, an interactive mapping site designed to address one of the core requirements of the Geologic Data Preservation Plan of Washington, which is to make data available and accessible.

The Washington State Geologic Information Portal utilizes the Arc Server interactive mapping application and is programmed with the Flex API, which is based on the Adobe Flash platform. One of the many strengths of Arc Server with Flex is serving large datasets to the user quickly. With this application the user can navigate through different themes; locate, overlay, and query data; create maps with titles, legends, and labels; create their own points and polygons; and export their maps as PDFs or geo-referenced TIFFs.

The Subsurface Geology Information System is the map theme where the data from the geotechnical reports are now housed. It can be selected from the Map Theme drop-down in the upper left hand corner of the application. When a point is selected with the Point Identify tool, a pop-up gives the user useful information related to that point in four categories;

1. Point Data
2. Layer Data
3. Comment Data
4. Document Information

A hyperlink to a PDF of the scanned document is provided in the Document Information section of the Identify pop-up, indicated by the “PDF” in blue text, as seen in the figure on the following page.



The Subsurface Geology Information System can be selected from the Map Theme drop-down in the upper left hand corner of the application (highlighted in green, as seen in the above map).

A hyperlink to a PDF of the document is provided in the Document Information section (note the blue PDF hyperlink emphasized by the red arrow in the upper right corner of the map above).

Using the scroll bar on to left of the Identify pop-up reveals the layer and comment information as seen in the screenshot to the left.

<https://fortress.wa.gov/dnr/geology/>