Transitioning to the FGDC Draft Geologic Map Database Standard: A Washington State Geologic Survey Pilot Project

by Meredith C. Payne, Anne C. Olson, David K. Norman, David Jeschke

Washington Division of Geology and Earth Resources
MS 47007
Olympia, WA 98504-7007

NSDI Cooperative Agreements Program
FGDC-endorsed Standards Implementation Training and Outreach Project (Category 5: 2012) Final Project Report

Date: March 31, 2014

Agreement Number: G12AC20145
Project title: Transitioning to the FGDC Draft Geologic Map Database Standard: A Washington State Geologic Survey Pilot Project
Final report
Organization: Washington Department of Natural Resources, Division of Geology and Earth Resources, PO Box 47007, 1111 Washington St SE, Olympia, WA, 98504, www.dnr.wa.gov
Principal Investigator: Meredith C. Payne, 360-902-1584, meredith.payne@dnr.wa.gov

Executive Summary
The Division of Geology and Earth Resources (DGER) at the Washington State Department of Natural Resources (WADNR) recognizes the importance of disseminating digital geological map data in a format that insures the highest level of quality, consistency and accuracy for our users. As such, we undertook a pilot project with the assistance of an NSCI CAP grant to implement and explore the process of converting a single digital geological map (a legacy quadrangle) as a test case from our own agency’s schema to a draft-standard schema (NCGMP09), currently under consideration by the FGDC. During this project, we established good working relationships with the NCGMP09 standard authors/developers, as well as with other users who have already adopted this model; their advice has proved invaluable in increasing the efficiency of our efforts. We maintained a steady stream of communication with our collaborators through email, telephone, and a project tracking website that we built for this purpose. We kept continual documentation of our procedures and refined them as we created a number of Python scripts designed to transfer our test data from DGER schema to NCGMP09. Although written for Esri’s ArcGIS 10.1, the scripts and the associated toolbox should be easily updatable to versions 10.2
and beyond. Finally, and perhaps of most value to interested parties outside of our agency, are our notes and observations detailing the advantages and challenges associated with the potential adoption of the NCGMP09 data model. These notes and recorded experiences serve to inform decisions made by our agency’s geologists and data managers as well as the developers and promoters of NCGMP09.

Project Narrative
The Washington State Division of Geology and Earth Resources (DGER) uses geologic maps as primary media for disseminating geologic information to the public, and is committed to sharing their geologic map data in a format that insures the highest level of quality, consistency and accuracy. As such, DGER recognizes the need for a national geologic data standard that will ensure consistent quality, streamline production, and facilitate sharing across various Canadian and U.S. agencies. To address this need, the USGS National Cooperative Geologic Mapping Program (NCGMP) has devised a draft FGDC (Federal Geographic Data Committee) standard for storage of geologic map data and ancillary information in a structured geodatabase model. Currently, this draft standard, known as NCGMP09, is at the advanced development stage. The NCGMP09 model is a geodatabase storage format for geologic feature data and associated information; it is currently available online as part of a USGS Open File Report (Soller, 2011). The NCGMP09 geodatabase model builds on an established FGDC symbol standard “FGDC/USGS Digital Cartographic Standard for Geologic Map Symbolization” (FGDC Document Number FGDC-STD-013-2006, available at http://ngmdb.usgs.gov/fgdc_gds/).

The goal of this project was twofold: (1) to test the feasibility of transitioning from the “in-house” schema DGER currently uses to produce geologic maps and (2) to provide useful feedback to the developers of the NCGMP09 draft geodatabase standard in the form of (a) a cookbook, detailing our migration process, and (b) a record of thoughts, challenges, and solutions pertaining to the specific experience of DGER during the migration process.

The first half of this project’s timeline involved establishing important collaborations and working relationships with key developers and users of the National NCGMP09 draft standard for digital geologic information. Specifically, attendance at the 2012 Digital Mapping Techniques (DMT) meeting in May 2012 allowed the principal investigator to begin dialog with Dave Soller and Ralph Haugerud (USGS, NCGMP09 developers), and Ryan Clark and Janel Day (AZGS, NCGMP09 users and tool creators). Subsequent conversations, and data element and code sharing kept this project on track. Those communications, as well as the NSDI CAP conference calls that were held throughout the project’s timeline, had the added benefit of creating a continuous feedback stream between ourselves and the draft standard developers, whereby some issues were discussed as they arose, allowing the developers to respond and improve their model throughout the project’s duration without having to wait for the results (this report) at the project’s end. As stated above, a primary goal of this pilot project was to provide the NCGMP09 developers with valuable feedback resulting from our experiences with implementing the draft standard, and we are pleased that this goal was realized throughout the length of the project.

Our project began with a research phase into the NCGMP09 geodatabase design, followed by formulation of a migration plan, particularly the determination of how to map feature classes from the DGER schema to the NCGMP09 geodatabase (diagrammatically depicted in Figure 1; a
field assignments spreadsheet is also available on the project website, 
https://sites.google.com/site/wadnrncgmp09/documents in the DGER NCGMP09 Tools directory). This migration was non-trivial, as our in-house schema consists of a file geodatabase containing one polygon, three line, and three point feature classes, whereas the NCGMP09 schema comprises a feature dataset with two polygon, three line and three point feature classes, as well as three geodatabase tables (as seen in Figure 2).

We used a combination of Ralph Haugerud and Evan Thoms’ (USGS) ArcGIS version 10.1 NCGMP09 Toolbox (Haugerud and Thoms, 2013), downloadable from the main NCGMP09 website (USGS) at http://ngmdb.usgs.gov/Info/standards/NCGMP09/, and our own Python scripts to transfer our data into the NCGMP09 geodatabase schema. These scripts underwent continual modification and refinement while this project progressed and various migration procedures were tested. Although we are sure that our scripts are imperfect and far from elegant, we make them available to interested readers on this project’s website (https://sites.google.com/site/wadnrncgmp09/documents), along with directions for implementing them (the DGER NCGMP09 Migration Cookbook)¹. This cookbook is the first of two deliverables promised to the NCGMP09 developers as an outcome of this grant project. Throughout the field mapping and scripting efforts, evaluation by DGER geologists and data managers was recorded. Distilled observations are also available on the project website; this evaluation commentary represents the second deliverable to the NCGMP09 developers. It summarizes DGER’s thoughts and concerns on how well the draft-standard schema performs in meeting our needs for data representation and dissemination.

¹ The Scripts are contained within the DGER_NCGMP09_tools.zip file found in the DGER NCGMP09 Tools directory, while the Migration Cookbook is located in the DGER NCGMP09 Documentation directory.
Figure 1 Chart of NCGMP09 geodatabase structure for the Carnation pilot project. Constructed in MS Visio.
In general, the migration procedure that we established involved writing and running a Python (2.7x) script\(^2\) for each feature class to be migrated between schemas (essentially a field mappings and field calculations script from DGER to NCGMP09). We divided the scripts per NCGMP09 feature class for the reason that some geologic maps will incorporate some feature classes that others do not (for example, not every map would necessarily need a GeochronPoints feature class, as not every geologic map includes age-date data). The execution of the feature class-specific Python script is then followed by running the *Attribute by Key Values* ArcGIS Toolbox tool (Haugerud and Thoms, 2013), which works by building a Python data dictionary from a user-generated list (in text file format) of relationships between attributes of a feature class. The tool uses this information to populate the target feature class attribute table based on those defined relationships.

It is important for the reader to remember that this pilot project was undertaken with testing the NCGMP09 transition for the Washington State Department of Geology and Earth Resources (DGER) in mind. Therefore, the tools built for this project specifically address the “in-house” schema that DGER currently uses. The migration involved a DGER geodatabase containing the digital cartographic information used to produce a STATEMAP geologic map product published in 2010 (Geologic map of the Carnation 7.5-minute quadrangle, King County, Washington by Dragovich and others, 2010). Many archival maps share the Carnation map schema, however, the DGER schema has evolved over time, and thus the Python scripts generated for this cookbook will need to be altered for migration of older DGER legacy datasets. Likewise, we believe that the scripts should be relatively easy to modify to work with other agencies’ data by users with moderate knowledge of Python and the ArcPy geoprocessing module.

Writing these scripts highlighted complications with the data migration process, which are exemplified in the *Notes and Observations on the NCGMP09 Migration Process* document (located in the DGER NCGMP09 Documentation directory at

\[^2\] Our Python scripts assume the user is also an ArcGIS user and relies heavily upon the ArcPy package.
Suffice it to say that there is no “one-size-fits-all” migration scripting routine. In addition to needing to alter the scripts for various legacy schemes used in the past, this investigation found that plenty of human intervention is necessary to correctly transfer the data from a single map on an individual feature class-by-feature class basis. This challenge emphasized the need for each script to be written specifically for the migration of each feature class, and then to be run in a logical sequence to build the NCGMP09 geodatabase element-by-element.

Likewise, the tables (DescriptionOfMapUnits, Glossary, and DataSources) in the draft standard geodatabase must be populated by a combination of programmatic and manual methods. As demonstrated by the migration cookbook, the DescriptionOfMapUnits table was the most complex of the three tables to assemble. It involved initial extraction of map unit descriptions from original map information followed by transfer to a style template in Word and ultimately to the geodatabase via an ArcToolbox tool. A manual data-gathering effort was required in order to accomplish the population of some of the table’s fields. These data (RGB fill values and GeneralLithology terms for MapUnitPolys features) were then used as inputs to a script that parsed the values into the proper records via user-assembled data dictionaries. This workflow gives some semblance of the degree of human judgment and manipulation required to migrate a single legacy map to the NCGMP09 schema.

There is no perfect solution to account for the amount of manual intervention that is necessary to successfully migrate the map data from the in-house schema to NCGMP09. The maps are as individual as the mappers that produce them. Therefore, significant time must be allotted, not only for data managers to acquaint themselves with the NCGMP09 model, but also to perform a quality transition of legacy digital map data. Nevertheless, we were able to successfully fulfill our project goal by developing a logical workflow for migrating a legacy map to an NCGMP09 geodatabase that is capable of being used to construct a geological map that accurately conveys the information published in the original map (Figure 3).

In addition to time demands imposed by the basic migration process, further results of this project highlighted several more challenges associated with both the migration to the draft standard and with the NCGMP09 model itself. Primarily, we experienced added complications in the migration process due to the way data are captured in DGER’s geodatabase schema compared to how the equivalent information must be parsed out into the NCGMP09 feature classes. In general, while information detailing feature type, description, and measurement accuracy is usually combined into single (qualitative) fields in DGER’s schema, it is contained within separate fields in NCGMP09 feature classes. Furthermore, the DGER schema does not demand numerical accuracy estimates (for example, LocationConfidenceMeters or OrientationConfidenceDegrees fields for point and line feature classes), as does NCGMP09. As mentioned above, this issue results in limitation of potential migration process automation. Reconstruction of these measurement details from field notes (if possible) would require much human scrutiny and dialog between data managers and mappers. Hence, ample time is needed to review each migration effort that endeavors to recreate numerical accuracy values, which potentially compromises efficiency by demanding staff time to analyze and transfer the data carefully so as to ensure maximum data preservation. This time requirement will undoubtedly concern understaffed agencies faced with NCGMP09 migration. Therefore, it seems obvious that
producing an NCGMP09-compliant digital geologic map database will be less cumbersome to incorporate in future map planning, than to assimilate with past maps.

However, other DGER concerns, such as the heavy reliance of the NCGMP09 model on the FGDC/USGS Digital Cartographic Standard for Geologic Map Symbolization (FGDC-STD-013-2006), which does not meet all of the symbol and definition needs for describing western Washington State surface geology (which is dominated by Pleistocene glacial deposits and Tertiary volcanics) cannot be addressed by time allocation or alteration of DGER mapping or record-keeping practices. Further work must be done to reach a satisfactory solution for addressing this concern.

Summary of Findings
NCMP09 is a logical model for dissemination of geologic map information. This investigation found that a significant strength is its preservation of data and observations recorded in the field by geologists. By contrast, the existing schema of DGER is limited to a highly edited synopsis, where much of the information recorded by the mappers in the field does not make it from their notebooks to the final map geodatabase. However, there is considerable benefit to map consumers being able to review interpretations made in the field, and NCGMP09 provides a data structure to do so. That being said, we are concerned that the NCGMP09 model can be overwhelmingly complex and demanding, especially with regards to numerical accuracy assessments for features. In the case of this pilot project, and indeed, in much of our legacy geologic datasets (DGER has geologic maps dating back to the early 20th century), that information is not recoverable. Therefore we ask: what exactly is enough to constitute an NCGMP09 standard-compliant geologic map geodatabase? Can or should the model be altered to allow more legacy datasets to be captured by the NCGMP09 model?

As discussed above, these findings suggest that a different data recording methodology should be assumed by DGER mapping geologists while in the field, should DGER adopt the NCGMP09 schema in the future. Perhaps it would be possible to introduce a “light” version of NCGMP09 suitable for legacy dataset migration. In that way, valuable legacy data could still be presented in an easily exchanged and shared data standard, accessible by public users and other agencies. At the same time, provisions could be made by DGER to collect future mapping data as appropriate for the more robust NCGMP09 model. Regarding the inadequate FGDC/USGS Digital Cartographic Standard for Geologic Map Symbolization (FGDC-STD-013-2006) terminology, DGER would welcome further dialog with the NCGMP09 developers to determine a strategy for representing Washington’s geology when the FGDC-STD-013-2006 symbology and descriptions fall short (for example, should the Notes fields be used for this, or is there likely to be an evolution in FGDC-STD-013-2006?).

Ultimately, DGER recognizes that a plausible next step in the enforcement of the NCGMP09 draft standard would be the requirement that NCGMP09-compliant geodatabases be submitted as deliverables to support U.S. Geological Survey STATEMAP products. Hence, this investigation suggests that the Washington Division of Geology and Earth Resources (DGER) mappers and data managers soon becomes familiar with the standard and plan ahead, so that alterations to field mapping methods and data capture styles in the present will translate to a smoother transition in the future.
Public Access
Interested users can access the project materials used to accomplish this grant work at the project website: https://sites.google.com/site/wadnrncgmp09/. On the Project Documents and Deliverables page, we provide both the archival geologic map geodatabase used for this project, and the resulting NCGMP09 geodatabase generated from it. Furthermore, we include the Python scripts and text files used to complete the migration, as well as the cookbook documentation that explains the data transferal process. We hope that these files might be especially useful in guiding other agencies in making a similar transition. We also include geodatabase, feature class and table metadata in ArcGIS and .xml formats, and an Esri .mxd map document, layer, and style files, so that evaluators can replicate our work. All geodatabases are in Esri’s ArcGIS 10.1 proprietary format, but as the NCGMP09 model advocates inclusiveness, open format shapefiles of the point, line, and polygon features in NCGMP09 format are also offered on our project site.

We encourage investigators to visit our project site, not only to download the products mentioned above, but also to view a list of links to other NCGMP09-relevant websites and to view additional tools, reports, and presentations related to the work carried out for this NSDI CAP grant.

Next Steps
We expect that the documentation generated by this project will be read and evaluated by both the developers of the NCGMP09 model as well as geologists and data managers from other state agencies. We hope that it will be found relevant and informative, and aid in the decision making processes of both the developers, as they finalize the model, and the potential users, who will adopt the model and have need to migrate their legacy data while preparing to collect future data for the new standard schema. DGER’s investigation of NCGMP09 will naturally extend beyond this grant project as in-agency questions continue to arise with respect to the feasibility of formal adoption of the standard. If endorsed, effective implementation of NCGMP09 and a timeline will be discussed by DGER geologists, cartographers, and data managers. Irrespective of the outcome, DGER has joined the pool of agencies with experience in the data conversion process and will be able to assist other agencies making the transition. Participants in this grant also belong to a national steering committee working on NCGMP09, whose work is ongoing.

Unfortunately, due to time constraints, an NCGMP09-generated correlation of map units and cross section diagrams were not constructed as part of this project. However, we acknowledge that the NCGMP09 model makes provisions for these additions, and we feel that eventual completion of these features will generate additional useful commentary for the draft standard developers, and be an informative exercise for DGER as they continue their evaluation of the standard. Therefore, we hold this as a worthwhile future goal.

Maintaining and updating this project’s website (https://sites.google.com/site/wadnrncgmp09/) will allow the materials to be accessed by interested parties for the near future. It is also possible that this work will be incorporated into links from the official NCGMP09 site page (http://ngmdb.usgs.gov/Info/standards/NCGMP09/), giving it wider circulation in the future.

Lastly, although beyond the original scope of this project, we have worked on generating a publically available online version of the digital geologic map (Figure 3) that would be served
through our agency’s ArcGIS Online for Organizations account. Working with ArcGIS Online as a vehicle to disseminate our data has the implications of reaching a wider audience and increasing the visibility of our agency and our products. If this product is eventually realized, it will be advertised on this project’s website (https://sites.google.com/site/wadnrnegmp09/).

Feedback on Cooperative Agreements Program

What are the CAP Program strengths and weaknesses?
The CAP Program made it possible to assemble a group of subject matter experts over a number of agencies in various locations by providing an official forum for discourse and securing the assistance of outside parties in the project development and execution. We were very disappointed to hear that the 2013-2014 CAP announcement was cancelled, as we wished to write a proposal for continuation of this work.

Where did it make a difference?
Simply put, this project would not have been possible without the assistance of the CAP. The grant provided an FTE, allowing the principal investigator to thoroughly research the problem and to complete the work necessary to achieve the goal of providing the desired feedback, both to our agency (DGER) and to the model developers (USGS). This opportunity also allowed us as grant recipients to connect and remain in dialog with members of the geologic map standard common-interest community. We strongly believe that the CAP funding of our work has increased the amount of relevant information available to decision makers involved with national digital geologic map standard development, as well as to other agencies interested in adopting it.

Was the assistance you received sufficient or effective?
Throughout this project, we remained in regular contact via email and telephone with the NSDI CAP support staff (Brigitta Urban-Mathieux and Julie Binder Maitra). They were instrumental in keeping the project on track by organizing conference calls, and addressing questions in a timely manner. Moreover, the principal investigator wishes to express gratitude for their support and understanding when the project timeline was delayed due to personal circumstances.

What would you recommend that the FGDC do differently?
As we did not have any negative experiences with the CAP organizers or project support team, we do not have recommendations.

Are there factors that are missing or are there additional needs that should be considered?
Not in the case of our project.

Are there program management concerns that need to be addressed, such as the time frame?
As mentioned above, we are grateful for the program’s flexibility, which allowed for the successful completion of the project during a modified time frame.

If you were to do the project again, what would you do differently?
Although it may have seemed burdensome to them at the time, it would have been beneficial for the PI to delegate much more of the project work to her coworkers, so that even more tasks could have been accomplished during the timeline, as well as distributing understanding and expertise of the NCGMP09 model more evenly throughout DGER staff.
ACKNOWLEDGMENTS

The authors wish to thank Ralph Haugerud (USGS) and Dave Soller (USGS), without whose guidance, assistance, and thoughtful commentary this project would not have been possible. The authors also wish to thank Gita Urban-Mathieux (USGS/FGDC) and Julie Binder Maitra (FGDC) for project support. Furthermore, the authors thank J. Eric Schuster and Ian Hubert (WA DNR-DGER) for their valuable insights. Lastly, the authors thank Janel Day and Ryan Clark (AZGS), Kent Brown (UTGS) and Tom Carlson (USGS) for their technical expertise and direction.

REFERENCES CITED


Figure 3. Next Page. NCGMP09 rendition of the Carnation Quadrangle 7.5-minute, 1:24,000-scale geologic map. Generated directly from the NCGMP09 geodatabase in ArcMap.