G10AP00117

Technical Report

AZ Geoscience Information Network (AZGIN), Phase III

Digital conversion of legacy GIS data

By M. Lee Allison and Stephen M. Richard
Principal Investigators

Project Period 07/01/2010 to 12/31/2011

March 10, 2012

TABLE OF CONTENTS

Introduction 3
Scope and Implementation 3
Digital Infrastructure 3
Digital Infrastructure produced (Results) 4
Summary 4
Datasets Converted 4
Availability 8
Acknowledgements 9
Bibliography and Links 9

Introduction

As part of the Energy Policy Act of 2005, the U.S. Geological Survey was tasked with creating a National Geological and Geophysical Data Preservation Program (NGGDPP). This Program is intended to work principally with the nation's geological surveys to: (1) archive geologic, geophysical, and engineering data, maps, well logs, and samples; (2) provide a national catalog of such archival material; and (3) provide technical and financial assistance related to the archival material.

The Program is envisioned as a national network of cooperating geoscience materials and data repositories that are operated independently yet guided by common standards, procedures, and protocols for metadata. The holdings of all collections will be widely accessible through a single, common, and mirrored Internet-based catalog, the National Digital Catalog, thus maximizing the availability of and interconnectedness of all the collections.

Scope and Implementation

In 2010 - 2011, the AZGS was provided with $32,882 for Digital Infrastructure. Work to be completed under the Digital Infrastructure priority, focused on migration of the Arizona Geological Survey's extensive digital geologic data collection from obsolete software formats and data schemas into our current enterprise database and into open source, xml-based, standard formats for improved preservation and access via standard web services.

The AZGS has published 71 paper maps in our Digital Geologic Map (DGM) series since 2000 and 27 digital map data products in our Digital Information (DI) series since 1993. Of the 98 total publications produced by the Arizona Geological Survey, we identified 56 that were in digital file formats (e.g. Arc/Info coverages) or database schemas that were obsolete and no longer in use. The objectives of the Digital Infrastructure priority were to (1) convert these 56 obsolete digital products to current database and open-source standard formats for enhanced preservation and access; (2) make the digital data more easily accessible and universally usable via the AZGS enterprise geodatabase and web services.

Digital Infrastructure

The project plan was to develop a consistent workflow to: (1) first check the quality of the existing data; (2) map vocabulary from older or non-standard vocabulary to our current vocabularies; (3) develop SQL queries to load data from existing tables and feature classes into National Cooperative Geologic Mapping Program (NCGMP) 2009 database scheme (NCGMP09, http://ngmdb.usgs.gov/Info/standards/NCGMP09/) ESRI personal geodatabases, and; (4) verify that all geologic unit description information from the original datasets is captured in the geodatabase.

Digital Infrastructure (Results)

Summary

Following the project plan, we developed and implemented a consistent workflow for the 56 AZGS digital datasets identified as obsolete under the Digital Infrastructure Priority.

The quality of each dataset was first reviewed to ensure data integrity, correctness and completeness. Datasets were classified as being of high quality or low quality. Older or non-standard vocabulary from the high-quality datasets was mapped to our current vocabularies, including the geoscience vocabularies that match the CGI GeoSciML vocabularies. With few exceptions, symbols from the FGDC Digital Cartographic Standard for Geologic Map Symbolization were assigned to each feature. Data from the high-quality datasets were loaded into the appropriate tables and feature classes in the National Cooperative Geologic Mapping Program (NCGMP) 2009 database scheme ESRI file geodatabases. A final review has been conducted to verify that all geologic information from the original datasets is captured in the NCGMP geodatabase.

Datasets converted

Datasets identified under the Digital Infrastructure priority and their current status can be found in Table 1.

Table 1

Series#	Status	Scale	Dataset
	24 K data, digitized from AZGS black-line OFR maps by contractor. Non-standard data format	24K	Oracle Junction (AZGS OFR 00-04), Samaniego Hills (AZGS OFR 99-17), southern Roskruge Mountains (AZGS OFR 00-06), North Butte (AZGS OFR 97-04)
	geometry & tables loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Stone, Paul, 1990, Preliminary geologic map of the Blythe 30' by 60' quadrangle, California and Arizona: U.S. Geological Survey Open-File Report 90-0497, 1 sheet, scale 1:100,000.
DGM-03	geometry & tables loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Surficial geologic maps and geologic hazards of the Green Valley - Sahuarita area, Pima County, Arizona [Sahuarita, Green Valley, Esperanza Mill, and Twin Buttes 7.5 min]
DGM-08	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Combination geologic map of the Galleta Flat West 7.5' Quadrangle, Pima and Cochise Counties, Arizona
DGM-09	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geologic map of the Mescal 7.5' Quadrangle, Pima and Cochise Counties, Arizona
DGM-13	geometry & tables loaded into to NCGMP;	24K	Surficial geology and geologic hazards of the Amado-Tubac area, Santa Cruz and Pima Counties, Arizona

Series#	Status	Scale	Dataset
	vocabularies updated; FGDC symbology		
DGM-18	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geologic map of the Fortified Peak 7.5' Quadrangle, southeastern Pinal County, Arizona
DGM-19	missing or incomplete data (Statemap '02)	24K	Geologic map of the Durham Hills 7.5' Quadrangle, Pinal County, Arizona
DGM-20	missing or incomplete data (Statemap '03 Tortolita-BKUP)	24K	Geologic map of the Desert Peak 7.5' Quadrangle, southeastern Pinal County, Arizona
DGM-21	missing or incomplete data (Statemap '02)	24K	Geologic map of the Oro Valley 7.5' Quadrangle and the Pusch Peak area, northeastern Pima County, Arizona
DGM-22	missing or incomplete data (Statemap '02)	24K	Geologic map of the Chief Butte 7.5' Quadrangle, southeastern Pinal County, Arizona
DGM-23	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geologic map of the North of Oracle 7 1/2' Quadrangle, southeastern Pinal County, Arizona
DGM-24	missing or incomplete data	100K	Geologic map of the Sierra Ancha, central Arizona
DGM-26	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geologic map of the Tortolita Mountains, Pinal and Pima Counties, Arizona
DGM-29	missing or incomplete data	24K	Bedrock geologic map of Sentinel Peak (A-Mountain) and Tumamoc Hill, Pima County, Arizona
DGM-30	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology;	24K	Geologic map of the Samaniego Peak 7.5' Quadrangle, Pima County, Arizona
DGM-31	Needs more thorough review	24K	Geologic map of the Twin Buttes 7.5' Quadrangle, Pima County, Arizona
DGM-32	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geologic Map of the Batamote Hills 7.5' Quadrangle, Pima County, Arizona,
DGM-33	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geologic map of the Esperanza Mill 7.5' Quadrangle, Pima County, Arizona
DGM-41	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology;	24K	Geologic map of the Vulture Mine 7.5 Quadrangle, Maricopa County, Arizona
DGM-42	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology;	24K	Geologic map of the Star Well 7.5' Quadrangle, Maricopa County, Arizona
DGM-43	geometry loaded into to NCGMP; vocabularies updated; FGDC sym-	24K	Geologic map of the Tucson SE 7.5' Quadrangle, Pima County, Arizona

Series#	Status	Scale	Dataset
	bology;		
DGM-44	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology;	24K	Geologic map of the Rincon Valley area, Pima County, Arizona
DGM-45	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology;	24K	Geologic Map of the Davis Dam SE 7.5' Quadrangle, Mohave County, Arizona and Clark County, Nevada
DGM-46	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology;	24K	Geologic Map of the Flatiron Mountain 7.5' Quadrangle, Maricopa, Arizona
DGM-47	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology;	24K	Geologic map of the Wintersburg 7.5' Quadrangle, Maricopa County, Arizona
DGM-52	shapefile, simple data structure	1000 K	Estimated Depth to Bedrock in Arizona, Richard et al, 2007
DGM-53	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology;	24K	Geologic map of the Needles NE Quadrangle, Mohave County, AZ
DGM-54	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology;	24K	Geologic map of the Boundary Cone, Mohave County, AZ
DGM-55	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology;	24K	Fortuna quadrangle
DI-04 v2	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Digital Representation of Geologic Map of the Phoenix North 30x60 Quadrangle, by S.J. Reynolds and M.J. Grubensky
DI-05 v2	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Digital Representation of the Geological Map of the Phoenix South 30' x 60' Quadrangle, Central Arizona, by S.J. Reynolds & S.J. Skotnicki,
DI-06	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Digital Representation of the Geologic Map of the Salome 30' x 60' Quadrangle, West-Central Arizona, Richard et al., compilers
DI-07	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Digital Representation of the Geologic Map of the Little Horn 30' x 60' Quadrangle, Southwestern Arizona, Spencer, J. E., compiler
DI-10	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	1000 K	Physiographic Areas in Arizona Used by the Arizona Geological Survey, by R.A. Trapp and S.J. Reynolds

Series#	Status	Scale	Dataset
DI-11	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Geologic Map of the Mesa 30' x 60' Quadrangle, Arizona, compiled by J.E. Spencer et al., compilers
DI-12	poly geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Geologic Map of portions of the Theodore Roosevelt Lake 30' x 60' Quadrangle, Arizona, Spencer, et al., compilers
DI-13	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Geologic spatial data (1:100,000-scale) for portions of the Globe 30' x 60' Quadrangle, Arizona, Richard, S.M., compiler
DI-14	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geologic Map of the Fountain Hills / Mount McDowell Area, Maricopa County, Arizona, by S.J. Skotnicki
DI-15	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Digital Maps and Geographic Database of the Geology, Geomorphology, and Soils of the San Bernardino Valley, Southeastern Arizona, by T.H. Biggs, R.S. Leighty, S.J. Skotnicki, P.A. Pearthree, and T.R. Orr
DI-16	1997 data format coverage; in New Mexico; include as data preservation activity.	24K	Digital Geomorphic Surface Map and Geographic Database of the Southern Animas Creek Valley, Hidalgo County, New Mexico, by K.R. Vincent and P.R. Krider
DI-17	1997 data format coverage, projection problems in original data (NAD 27-83)	24K	Digital Surficial Geologic Map and Geographic Database of the Northern Tucson Basin and Tucson Mountains, Pima County, Arizona, by P.A. Pearthree et al.
DI-18	2001 data format coverage	24K	Digital Geologic Map and Geographic Database of the Morenci Area, C.F. Ferguson
DI-19	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geologic spatial data for the Roskruge and Waterman Mountains and Western Avra Valley Area, Pima County, Arizona, by C.A. Ferguson, W.G. Gilbert, T.R. Orr, P.A. Pearthree, S.M. Richard, S.J. Skotnicki, J.E. Spencer and Ann Youberg
DI-20	merge with map index, azgeobib; mostly in Mapextents already	1000 K	Digital geographic index of geologic literature for the east half of the Phoenix North 30' x 60' quadrangle, Arizona, by T.R. Orr
DI-23	Stand-alone database; 2001 type data structure; set up as WMS/WFS	1000 K	Database for Mineral Districts in the State of Arizona, by S.M. Richard, editor
DI-24	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geological Data (1:24,000) for the Southeast Phoenix Metropolitan Area, Maricopa and Pinal Counties, Arizona, v. 1.0, by S.M. Richard et al, compilers
DI-26	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Geologic spatial data (1:100,000-scale) for Casa Grande 30' x 60' Quadrangle, Arizona, Richard, S.M., compiler

Series#	Status	Scale	Dataset
DI-27	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Geologic spatial data (1:100,000-scale) for the west half of the Mammoth 30' x 60' Quadrangle, Arizona, Richard, S.M., compiler
DI-28	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Geologic spatial data (1:100,000-scale) for the Silver Bell 30' x 60' Quadrangle, Arizona, Richard, S.M., compiler
DI-30	geometry and tables loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geological Spatial Data(1:24,000) for the northwest Phoenix Metropolitan Area, Maricopa County, Arizona, v. 1.0, by S.M Richard, E. M. Moore, T.R Orr and R.C Harris (compilers),
DI-31	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geologic spatial data (1:24,000-scale) for the southern San Pedro River area, Cochise County, Arizona, v. 0.5, by S.M. Richard and E.M. Moore, compilers
DI-34	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Geologic spatial data (1:100,000-scale) for the eastern part of the Sells 30' x 60' quadrangle, Pima and Santa Cruz Counties, Arizona, v. 1.0, Richard et al., compilers
DI-35 in prep	AZGS08 prototype; geometry imported to root; needs QA	100K	Geologic spatial data (1:100,000-scale) for Ft Huachuca 30' x 60' quadrangle, Pima, Cochise, and Santa Cruz Counties, Richard et al., compilers
DI-36	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	100K	Geologic spatial data (1:100,000-scale) for the Gila Bend and eastern part of the Dateland 30' x 60' quadrangle, Pima and Maricopa Counties, Arizona, Richard et al., compilers
DI-37	geometry loaded into to NCGMP; vocabularies updated; FGDC symbology	24K	Geological spatial data (1:24,000) for the Southwest Phoenix Metropolitan Area, Maricopa County, Arizona, v. 1.0, Richard et al., compilers

Availability

The datasets will be loaded into the AZGS resource repository (http://repository.azgs.az.gov), but this has required a change in agency policy away from sales of digital datasets to open access. Review of the policy changes is under way and the datasets should be accessible by Summer 2012. Datasets in the repository will have metadata pushed to the AzGS/USGIN online catalog for harvest by the National Digital Catalog.

Acknowledgments

Work on this project was completed by Stephen M. Richard, Lee Allison, Janel Day and Mitchell Levine.

Bibliography and links

CGI GeoSciML Vocabularies,

http://resource.geosciml.org/

FGDC Digital Cartographic Standard for Geologic Map Symbolization,

http://ngmdb.usgs.gov/fgdc_gds/geolsymstd.php

National Cooperative Geologic Mapping Program (NCGMP) 2009 database scheme,

http://ngmdb.usgs.gov/Info/standards/NCGMP09/

National Geological and Geophysical Data Preservation Program,
http://datapreservation.usgs.gov