DIGITAL MAPPING TECHNIQUES 2024

The following was presented at DMT '24
May 13 - 16, 2024

The contents of this document are provisional

See Presentations and Proceedings
from the DMT Meetings (1997-2024)
http://ngmdb.usgs.gov/info/dmt/

Tearing Down Data Silos -- An overview of the IGWS's work to consolidate years of borehole data into a centralized database

By Matt Johnson (Assistant Director for Information Services, Indiana Geological and Water Survey)

The Indiana Geological and Water Survey (IGWS) has numerous geospatial data sets related to geological mapping. Many of these data sets were part of individual projects and have not been fully accessible to the entire organization, rendering the data hidden from further research at the IGWS. This inaccessibility and lack of knowledge have caused duplicated efforts and limited the scope of future projects. During the 2020 STATEMAP cycle, a report called "Progress Toward Aggregating Map Compilation Data Sets into a Centralized Database" looked at the data and suggested a plan of action to make it more available. This presentation will discuss the various data sets, the progress that has been made so far, and continued plans to consolidate data into a centralized, accessible location.

Tearing Down Data Silos

Status of a centralized stratigraphic point dataset for Indiana

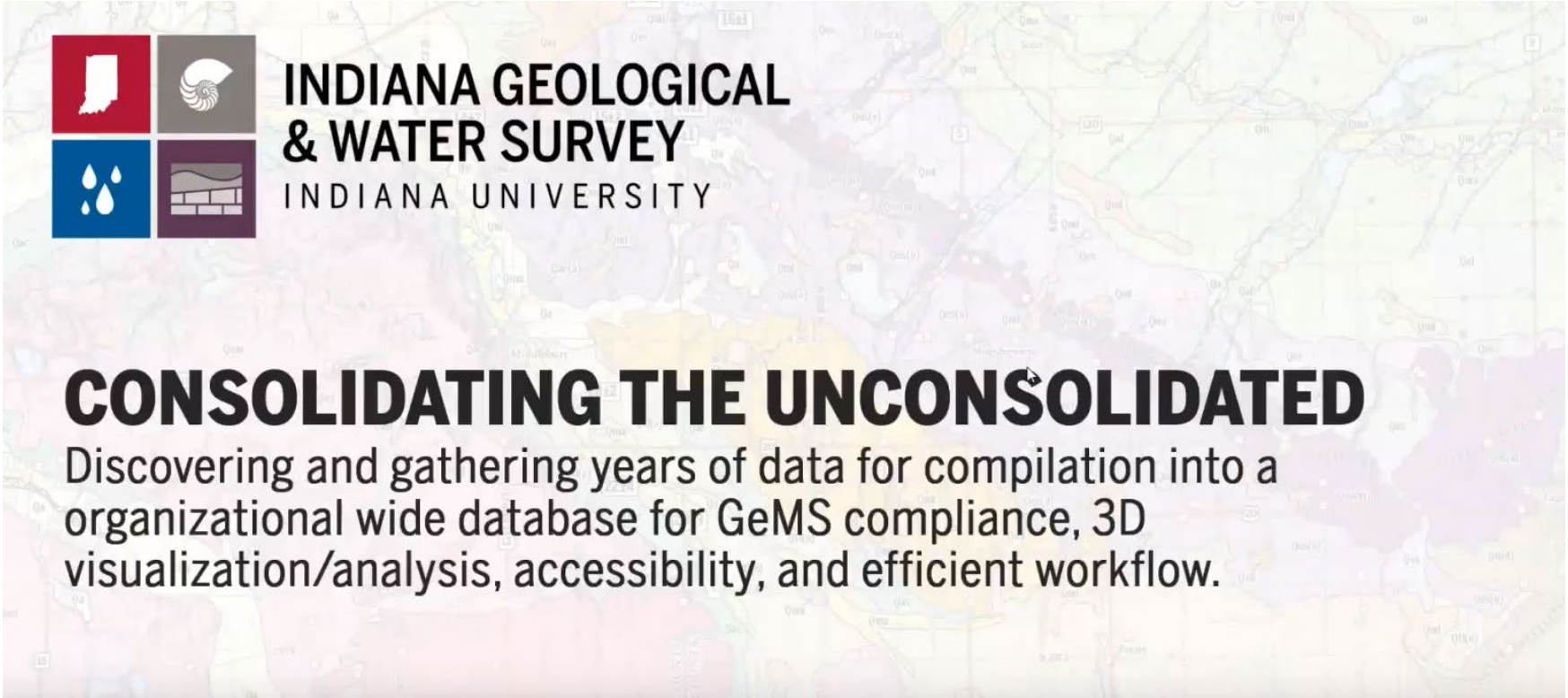
Matt Johnson

Assistant Director for Information Services

INDIANA GEOLOGICAL & WATER SURVEY

INDIANA UNIVERSITY

100 Year Old Silo Demolished YouTube - https://www.youtube.com/watch?v=nksEVPQ6cvs



August 10, 2020, CDEFG:3D presentations by Indiana and Canada
GeMS Wiki - https://dggs.alaska.gov/gemswiki/doku.php?id=start:resources

Project 4 – STATEMAP Supplement Part 1, Indiana Geological and Water Survey:

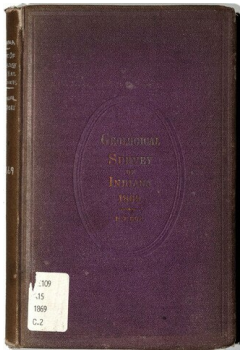
Progress Toward Aggregating Map Compilation Data Sets into a Centralized Database– Year 1 of 5

proposed workflows documented in this report. In addition to data assessment and ingesting, one staff member will attend a one-day training session on Esri branch versioning, depending on future funding and class availability. This training will allow the infrastructure and workflows described above to be built using industry best practices. While the major focus of this report and projects as described above is on the many point data sets at the IGWS, future work will include taking the digital data of published maps and moving it into the GeMS format where appropriate for the IGWS long-term mapping plan.

furthering the organization's mission. Along with inaccessible data sets, many other data sets have been copied multiple times for multiple projects and vary in how current they are. This adds complexity to finding the most up-to-date, reliable data set when new projects start and, in many cases, duplicates efforts to process and obtain the same information routinely. This report summarizes the progress during year one of a planned five-year project to aggregate these data sets into a centralized database that will be fully accessible to researchers at the IGWS and will contain select data sets that will be shared publicly.

IGWS data silos go way back…

- Paper…
- to Lotus 1-2-3…
- to spreadsheets…
- to access databases…
- to local file servers…
- to 5 ¼” & 3 ½” floppies…
- to CDs…
- to DVDs…
- to flash drives…
- to external hard drives…
- to SQL databases…
- to personal, file, and SDE geodatabases…
- to personal computer and cloud storage….WHEW!



1869 Annual Report

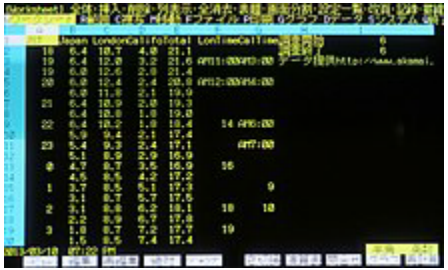


Table 1. Known IGWS data sets

Known point data	Description	Data stored in
BR_Boreholes	Major water-well updates from 2015 bedrock surface work	Shapefile
Coal (NCRDS & ICQD)	These datasets are mostly up-to-date and relevant data ready to ingest	Access, geodatabase, shapefile
Confidential data	Various data used internally for surface positions, etc. that was provided from industry	Varies
Cores Information System	Export out of Petroleum Database Management System (PDMS) with updates	Excel, Geodatabase, Microsoft SQL
Gamma logs	Multiple copies exist and most will need linking with water-well or borehole	Excel, Microsoft SQL, Petra, scans, LAS files
HVSR (Tromino)	Multiple copies exist in various projects	Access, Petra
iLITH (deprecated)	This dataset will probably be removed	Access, Microsoft SQL, shapefile
Illinois Basin data	Mostly stored in Petra databases but are subsets of PDMS with updates	Petra, Others?
Known siloed data	Data stored with individual projects	See Table 2
Other data	Data for projects on staff computers or within printed publications	Varies
PDMS	IGWS Petroleum Database Management System	Microsoft SQL
Seismic	Copies exist within PDMS and other projects	Microsoft SQL, shapefile
UNC_Boreholes	Major water-well updates from 2015 bedrock surface work	Shapefile
Unknown siloed data	Data from past projects on DVDs, staff computers, or misplaced on file servers	Varies
Water wells (IDNR)	Yearly and project-based data provided by IDNR (many copies and many versions)	Access, shapefile, geodatabase

Table 2. Siloed data from past projects

Project name	Data stored in	Year
Bartholomew County	Access	2014
Bedrock Surface	Shapefiles	2016
East South Bend 100k quad	Access, Excel, Petra, shapefiles, geodatabase	2021
Elkhart 100k quad	Access, Excel, Petra, shapefiles, geodatabase	2018
Flatwoods	Access, Excel, shapefiles	2019
Heartland	Access	2004
Illinois Basin	Mostly stored in Petra databases but are subsets of PDMS with updates.	???
Lawrence County	Access, geodatabase	2012
Logansport 100k quad	Excel, geodatabase	2018
Morgan County	Access, Excel, Petra, shapefiles, geodatabase	2016
Patoka Lake	Access, Excel, shapefiles	2020
South Knox 100k quad	Excel, geodatabase	2021
Teays Valley	Access, Excel, shapefiles	2021

Ingested data to date

- Indiana NCRDS Data
  - ~19k locations with ~235k stratigraphic picks
- Illinois Coal Data
  - ~6k locations with ~17k stratigraphic picks
- Lawrence/Monroe County Database
  - ~23k locations with ~62k stratigraphic picks
- Indiana Heartland Database
  - ~84k locations with ~137k stratigraphic picks
- Industry Confidential Records
  - ~2k locations with ~9k stratigraphic picks
- 2021, 2022, 2023 STATEMAP Data
  - ~1k locations with ~5k stratigraphic picks
- 2022-2024 Internal Bedrock Mapping
  - ~800 locations with ~951k stratigraphic picks

Current Enterprise Database

~130k Locations (~136k)

~307k Strat Picks (~460k)

Database and tool demo

1. Webmap showing stratigraphic picks

- With links to resources in Institutional Digital Asset Management System
- Bore hole descriptions, pXRF, core photos, etc.

2. Branch Versioning

- Multiple users editing through portal enterprise feature services with related table
- Can be used with Field Maps and Online Web App

3. New surface/contact creation tool

- Semi-automated approach to create bedrock contacts using enterprise branched database

Questions or comments?

Matt Johnson

mrj21@iu.edu

INDIANA GEOLOGICAL & WATER SURVEY
INDIANA UNIVERSITY