“Getting started with GeMS, part 2” (or “Moving Maps to GeMS, part 2”)
Presentation given as part of the DMT 2020 digital meeting on June 9, 2020.

by Caroline Rose (cmrose@wisc.edu) and Punwath Prum

Introduction

[slide 1] My name is Caroline Rose and I am a GIS Specialist and cartographer for the Wisconsin Geological and Natural History Survey. We have been working for 2 years now on moving previously-published maps into the GeMS schema. For this past year, I have been joined by student collaborator, Punwath Prum. He will show some of his most recent work at the end of this presentation.

Last year I presented in Montana and gave an overview of our not-quite-finished workflow. I call this presentation “Part 2” because it follows up on last year’s. But I intend for this to be understandable and useful even if you didn’t see last year’s talk.

This year’s presentation will briefly review what I offered last year, tie up some loose ends of that workflow, introduce some new documentation that we have recently shared, and discuss where we’re at and what comes next.

I hope this will be useful for any survey staff with upcoming grant funding through the Data Preservation Project or anyone else getting started with GeMS.

Side note

[slide 2] As a side note, I want to take a few seconds to mention this recent book called Data Feminism. It is not just for women, it’s for everyone. This book takes feminst concepts and applies them to data analysis and visualization, which is what we do as cartographers.

So, back to GeMS.

Documentation

[slide 3: github page] Throughout this project, we have been developing documentation, which we are sharing on this github page, including last year’s

slides and presentation script. If you scroll down to the readme section on the lower half of the page, you should see a summary of each document there.
Exactly where I would start with GeMS

[slide 4: flowchart] is with the “workflow overview” PDF. I know these slides have a lot of text on them; I don’t expect anyone to read all of this from the slides; I just want to tell you what’s in these documents. This gives a broad summary in this flowchart, followed by a more detailed table reviewing the steps

[slide 5] This steps through the process of evaluating the initial vector data files, converting file formats,

[slide 6: master glossary] collecting line and point features from each map to build a master glossary.

[slide 7: DMU] drafting a Description of Map Units table that we could ask a geologist to complete in Microsoft Excel

[slide 8] and doing some basic topology checks to establish that contacts and faults align with polygon edges, with no gaps or overlaps

[slide 9: quick-reference sheet] When we start to work on the GeMS-required attribute tables, this is where another piece of our documentation comes in handy.

[slide 10] Here’s a detail view of it. These Quick-reference sheets are really just extracted from the GeMS document, putting all the information for one layer on one sheet of paper. This helped us focus on the essential reference material, one layer or one table at a time.

[slide 11:decisions] populating all of the GeMS required attributes is not always straightforward.
I’ve said this before, but it bears repeating: GeMS is flexible, and there are a lot of decisions to make about how to capture the original map information. As of last year, many of these remained unresolved,

[slide 12: with arrows] but this year we’ve had to answer each of them:

- We have developed our list of standard confidence values,
- We have a master glossary, where all definitions must be listed

- And so on.

[slide 13: done] so as we approach the end of this process, with all of the required layers, and some optional layers, and all of the required tables, I want to give you a summary of how we go about confirming that it is complete: when can we call it "done"?

Pleistocene & other difficulties

[slide 14: Pleistocene] The numerous misspellings of Pleistocene that I've encountered in this project really encapsulate for me the endless variations on details that needed fixing at the end of this process: the many little ways to be almost-but-not-quite done with a database. And there are some documents that help us through this

[slide 15] So I want to discuss these documents and tools to help you verify that a GeMS database is complete. The three items listed on the left here are PDFs provided on our github page, which are meant to be used in combination with some of the USGS scripts.

[slide 16] First, we have this "GeMS Fields Checklist". This is meant to be used in tandem with the Validate Database tool. It's a checklist to manually go through each layer and each field or attribute. It doesn't tell you exactly what to look for; you need to be familiar already with what should be in those fields.

This relates to something I want to reiterate that was said by Tanner from South Carolina at the meeting two years ago: it's worth getting eyes on each map database. Although you can automate many steps, Each map has some unique aspect to the conversion process. This checklist helped me find a variety of errors. It's not just minor typos; I've noticed where field values were truncated accidentally, missing pattern descriptions, among other corrections. Some of this is redundant with the validation script but I've found some of it is not. Try it and decide for yourself whether it helps you find things to fix.

[slide 17] After layers, tables, and topology are settled, writing the metadata was also a process we had to learn. This is a spreadsheet we developed to guide that process, showing where to enter each item in ArcCatalog, and where it will show up in the FGDC xml export. Following this, you would run the metadata tool.

[slide 18] The metadata tool pulls content throughout the database into the metadata records, and this “Metadata Summary for GeMS Fields” clarifies where those items come from, and where they end up in the metadata as well as how each attribute is described in the final metadata. The script is pulling domain values from seven different fields into the metadata, it’s pulling the Fullname attribute from the Description of Map Units, it’s pulling Source from DataSources. So, if I have to make a correction *after* creating metadata, do I need to alter that in the metadata as well? If I found a missing pattern description, no. If I made a mistake in a citation, perhaps.

[slide 19: corrections] So when your coworker inevitably points out something you need to correct, you can trace back from the metadata to the geodatabase -- or -- you can find where it has been copied into the metadata.

A shift in focus

[slide 20] So that was just some of our documentation that we have made available. Going forward, we are beginning to change our focus. While our priority a year ago was developing the workflow to get maps into GeMS, our emerging goal now is to begin to **use** this collection of GeMS-formatted data.

We plan to use the GeMS data to support our statewide 1:500K surficial mapping that is just starting now.
The various geologists collaborating on that project will be able to reference these datasets while working remotely.
I will turn this over to my student collaborator, Punwath, to show some of his latest work.

[Punwath talk about ArcGIS online map]

[slide 21] Thank you. I’ve worked on making cartographic representations for the map. It is a tool in Arcmap for managing feature symbology and it is stored in Geodatabase. We set the rules which are used to draw features on the map. So, whenever we load the data to ArcMap, we will always have the same map view, though on different computers. Also, it improves the data display because there are a variety of rules that can be created for producing the symbols.

[slide 22] For map sharing, ArcGIS Online is a helpful tool for collaborators working remotely and map visualization. This tool is well supported in ArcGIS Pro and ArcMap, and the online map can be published directly from them. Especially when

we have done the cartographic representation, we can publish a map of many counties to ArcGIS online at once and share with the GIS or Geologist teams.

[slide 23] For example, here is the online map of 5 counties showing geomaterial. We can see the beautiful big picture of the geomaterial in the region.

[slide 24] As we wrap up this year's grant, I'm also considering whether to combine the many project overview tables, including master glossary, a full list of the maps, master data sources, and so on. I would be interested to hear if anyone has done this.

Concluding Thoughts

[slide 25] I hope this will be useful to people who are about to start on this project that I began two years ago. Please check out the github repo and / or get in touch. With that, we can take questions.