Guidance For Using QGIS to Build GeMS-compliant databases   12/11/2023

This is a short collection of best practices for those using QGIS for their geologic map production with the intent of submission to National Geologic Map Database (NGMDB). Note that we, the GeMS Group at the USGS, do not use QGIS frequently; these suggestions come after a review of one submission package that had been assembled in QGIS and after discussions with more experienced users. Please send us (gems@usgs.gov) comments as appropriate.

1. Geopackages are the expected data format for map data as they can be used in both QGIS and ArcGIS Pro. You are welcome to add as much functionality that is native to the format as you like (the most likely being embedding symbology), but keep in mind that not all of that functionality will be available to ArcGIS Pro users. This blog post about what functionality is exposed in ArcGIS Pro is a few years old, but current as far as we know.

2. The NGMDB may decide that other open-source formats can be accepted, but at this time, only Geopackages can be validated with the Validate Database tool from the GeMS Toolbox, and the tool will only run in ArcGIS Pro. If you have no access to ArcGIS Pro, contact us (gems@usgs.gov) for guidance.

3. For symbolizing geologic features in QGIS with FGDC-compliant symbols, see these links:
   https://github.com/moendopi/QGIS-FGDC-Symbols-complete
   https://github.com/BC-Consulting/FGDC-4-QGIS
   https://github.com/rodreras/geologic_icons
   Again, not being frequent users of QGIS, the resources here have not been evaluated.

4. Because the symbology files that QGIS uses (map, layer, and style files) cannot be read by ArcGIS Pro (and may not be readable by other GIS programs), please try to include an FGDC symbol number or fill color values, as appropriate, for all features. These should be placed in all "Symbol" fields in any feature class table and the "AreaFillRGB" field in the DescriptionOfMapUnits table. Add fields as necessary. For example, if the values you use in "Symbol" do not match FGDC symbol names in a style you are using, please add a field such as "FGDCSymbol", find the values you need from the cartographic standard, and put the values in there; use an "AreaFillCMYK" field if you have values from the CMYK gamut but not RGB, etc. You DO NOT have to supply enough symbol information to recreate an exact copy of a published map plate. Just make it possible to roughly show the chosen colors of the unit polygons and some line and point symbolization that distinguish geologic features using (hopefully) commonly understood symbols.

5. The GeMS schema describes using "feature datasets" to separate and collect related feature classes in ArcGIS file geodatabase: notably GeologicMap contains the feature classes ContactsAndFaults and MapUnitPolys. Because Geopackages have no similar container, if you happen to have multiple pairs of line and polygon layers that are topologically related, add a prefix or suffix to the core GeMS table name and the Validate Database tool will attempt to check the topology for those pairs separately. For example, you can add a SurficialContactsAndFaults, SurficialMapUnitPolys pair and a BedrockContactsAndFaults, BedrockMapUnitPolys pair of layers to your Geopackage.

6. See Ben Weinmann's collection of repositories on GitHub for tools to automate some workflows with Geopackages - https://github.com/moendopi?tab=repositories. Please let us know about other resources you use or if you have tools of your own. If you are developer, it may not be too difficult to take the existing GeMS Tools and swap out arcpy methods and objects for those from the GDAL/OGR library.

7. In lieu of ArcGIS Pro project, map, layer, or style files as described in the GeMS documentation, submit any equivalent QGIS files as necessary for quick display. Double-check that you have saved your project with relative paths and that it contains only references to valid layers in the submission package. Include as few layers as necessary to illustrate the geology – that is, you do not need to include basedata layers if all they do is provide geographic context and are available.