1. Potential installation of SpatiaLite DLL(s), Python and Python libraries may require admin or USDA/CEC assstance. – If installation of DLL(s) are required than we will need to consider altering our MVP. Give that spatial will be dealt with in a later phase, we can deal with this down the road but leave the door open to modify the requirements.

Steve in blue - In my experimentation, I have been able to use the standard version of Python (3.7.9) and standard set of libraries that are part of the ArcGIS Pro 2.7 distribution to do all of my work. I am able to create databases and create all attribute tables and constraints using the included sqlite3 library. For the spatial import I have been using arcpy, which works fine with ESRI software, but I plan to replace the spatial import with one of the other non-ESRI Python libraries.

2. Policy and mechanisms for managing hybrid content (e.g. data from multiple sources and/or SSA plus WSS ad-hoc AOIs) is not yet explored or defined. The MVP may already address this point. The MVP states “Database can store multiple soil survey areas”. The source of the SSA shouldn’t matter but the database should be able to store multiple records in the legend table and other tables that store SSA information. This should be true regardless of how the SSURGO data is generated and regardless of how many SSAs are included within a single set of SSURGO text files.

Totally agree. We might want to test the ability to import multiple WSS-AOI’s from a single soil survey area into a single Template database. I’m not sure if the current SSURGO constraints can handle that or if we should even allow that.

3. Use of metadata text files distributed with WSS SSURGO downloads is not considered. – Let me ask this, do we even need to consider these other than ensuring the metadata text files included in the SSURGO download are up to date? Here is another thought, do we even need to retain these files? Other than Steve using them for gSSURGO schema, what purpose do they serve and are they used by customers? If there are select group of people that do need these files (e.g. Steve for gSSURGO), could we simply redirect them to the SQLite database as the source of the metadata and remove these metadata SSURGO text files from the system. That results in one less piece in the puzzle.

When we originally began creating and distributing soils data in the gSSURGO file geodatabase format, I incorporated a static copy of all the ‘mdstat’ tables obtained from the MS Access database (US\_2003 version). After I few years we switched to using a template gSSURGO database with empty ‘mdstat’ tables and imported the tabular data from the textfiles. The reason for this switch was because I discovered that some of the domain values in the mdtatdomdet and mdstatdommas tables were not being kept up to date in the original Access Template database. This resulted in an apparent discrepancy in the some of the maps and reports generated by Soil Data Viewer vs Web Soil Survey. I think maybe it was an issue with the AASHTO domain, but I could be wrong.

4. Definition and delivery of an “empty template” is not finalized. The stories define a path to a manually-created empty template. We do need to refine this. My thinking was that we deliver the same sort of product we deliver for the mdb. That is a database that has the metadata tables and payload tables, along with relationships, indexes, and constraints. The idea was the user downloads the SQLite template and runs some script(s) in some environment to load the SSURGO data into the payload tables. I think this needs further discussion because as I understand the process you proposed, this would not be the case. Instead, the database would be initially constructed or rebuilt by the python script that executes the SQL statements. That may be okay but the “empty template” was us saying we just wanted to do the a similar thing that we did the mdb.

I would agree, at least for the MVP and the first version of the new Template. The look and feel of the Template should remain the same as much as possible.

5. Management of version mismatch (version.txt vs SQLite instance) is undefined. I would like Steve to comment on this one.

I’ve done some preliminary investigation and I sort of understand the concept of the how the database, tabular data and spatial data are versioned. This is a very important piece of the puzzle and deserves a lot more consideration than it has had so far. We haven’t had any changes in SSURGO versions for several years and so we’ve been spoiled. I do remember that the last changes in SSURGO tabular version and database versions and ArMap Addin version did cause a fair amount of support traffic, and those were pretty minor changes. I shudder to think what a major change could trigger if not handled properly and I suspect that even we handle it properly there will still be a lot of support issues. I think we should look at the versioning design and validation very carefully. I could see the potential for expanding upon the versioning design so that we can accommodate version numbers for template database type (SQLite, Postgres, etc), geometry type (ST\_Geometry, WKT, etc), raster type (GeoTIFF, Geopackage Tiled raster, etc). Bring on the Tylenol.

6. Ability to work with “non-ESRI open source programs such as R” is ill-defined. “R” is a language. What other languages beside Python and R are required? Is R in the MVP? We should have said ability to work in other non-ESRI programs. I think ideally the new database could be used in R Studio but I know very little about R and I would defer to Dylan to better define R/R studio requirements. I think this language needs to be modified in the MVP.

I agree that this area needs a lot more work and we need to write some simple stories that would describe our expectations for an MVP in the QGIS and R environments.

7. Interoperability of a populated database is undefined. For example if data are loaded when using Arc GIS Pro, should the same database file be useable in QGIS? I’m not sure how to answer this because I hadn’t envisioned using ArcGIS Pro to load data. The idea was to load the data outside of ESRI.

Same as the answer to number 6. I agree that this area needs a lot more work and we need to write some simple stories that would describe our expectations for an MVP in the QGIS and R environments.

8. Must the SQLite+SpatiaLite or GeoPackage be editable in ArcGIS Pro or QGIS? I think it should be it’s not a requirement. Editing is rare occurrence and is greatly discouraged and should only be left to someone who is comfortable with the database.

Agreed. I think at this time we should look at this new version of the SSURGO database as primarily a format for distributing the official SSURGO data and not as a platform for editing SSURGO data.

9. Should spatial data (shapefiles) be loadable into the SQLite+SpatiaLite or GeoPackagewithout needing to resort to a GIS? I think yes but I would defer to Steve, Adolfo, and Chad.

I think for the MVP I would be OK with using ArcGIS Pro and perhaps QGIS to import SSURGO shapefiles into the new Template database. Ideally though we would want to develop a ‘GIS-software-independent’ method.

10. Business needs to write some stories on what we want the SSURGO Import process to look like. The original tabular import process had fairly limited functionality and was not what I would want, even for the MVP.