This is in support of an FY19 funding request for a new Soil Interpretations Generator.

The Overview text below was exported from Rational 5/13/2020

HLBR effort was put on hold 6/8/2018 – reviews of overview and HLBRs were in progress

# Overview

## Purpose

The purpose of this document is to define the high-level business requirements for a Soils Interpretations Generator. An Interpretations Generator is a type of decision tool that uses soil and site property data from the soils database to generate soil ratings for land use based on specific interpretation criteria. The criteria used are based on the proposed use of the soil and the result is usually provided as a list of soil limitations or characteristics for the intended use. An example would be a soil used for landfill. If a soil has properties that would result in contamination of groundwater, the soil would have limitations for landfill. Soil properties affecting groundwater contamination would be particle–size (water would flow through sandy soils quickly and contaminate groundwater), or the soil may be on a steep slope resulting in water running off from the landfill and contaminating the surface water.

The current Interpretations Generator operates solely within the NASIS transactional database (as opposed to the public-facing Web Soil Survey Database) by generating interpretive output based on evaluations and ratings of soil properties stored in the database. The evaluations and rules as well as the properties are available only to those who have permissions to use the transactional database. Improvements to the whole process of modeling soil properties could be improved allowing users within the agency, outside users, and cooperators to develop their own interpretations using data from the public-facing database. Internal users would have the option to use either database, as needed. The new Interpretations Generator will be able to assist the public and the agency conservation planners by providing custom interpretive output that they prefer in a timely and more flexible manner.

* Ability to adjust the soil properties from site specific locations and generate on the fly interpretations (from ready-to-use interpretations) for desktop and mobile applications.
* Ability for users to create new interpretations outside of the transactional NASIS database.
* Ability to use geospatial layers from many formats to integrate into the soils data for developing more spatially explicit interpretations.
* Ability to design and build engaging maps for our partners, both within and outside the NRCS, using a variety of technologies including using some open source statistical, GIS, and model building tools This document is further supported by the NRCS Requirements Development and Management Plan. This effort will give the Investment Review Board (IRB) the necessary information for project prioritization and approve the project for funding.

## Background

A soil interpretation is a model that will rate a soil’s suitability for some use, such as the likelihood of finding usable gravel, or the potential for problems when building a manure pit. Interpretations are derived from the soil properties in the National Soil Information System (NASIS) database using a set of algorithms which can be tailored to model an interdisciplinary expert decision making process. There are three components of an interpretation, which are called

• Properties

• Evaluations

• Rules

Each of these is a class of business object in NASIS and has a specialized data entry module in the application.

Interpretations are usually developed and tested by highly trained and experienced users. They can then be used to produce a report of selected interpretations. Interpretation results or ratings are also computed and stored during the process of exporting data from NASIS. Instructions on developing and using interpretations are found in the NASIS documentation.

The current Interpretations Generator operates solely within the NASIS transactional database (as opposed to the public-facing Web Soil Survey Application) by generating interpretive output based on evaluations and ratings of soil properties stored in the database. The evaluations and rules as well as the properties are available only to those who have permissions to use the transactional database. Improvements to the whole process of modeling soil properties could be improved allowing users within the agency, outside users, and cooperators to develop their own interpretations using data from the public-facing database. Internal users would have the option to use either database, as needed. The new Interpretations Generator will be able to assist the public and the agency conservation planners by providing custom interpretive output that they prefer in a timely and more flexible manner.

• Have the ability to adjust the soil properties from site specific locations and generate on the fly interpretations (from ready-to-use interpretations) for desktop and mobile applications.

• Modernize the technology and platforms to provide enhanced ability for users to create new interpretations using soils data.

• Improved ability to use geospatial layers from many formats as input into for developing more spatially explicit interpretations.

• Provide enhanced geospatial technologies to design and build maps as exportable products for our partners.

## Scope

The intent of this document is to provide high-level requirements for an Interpretations Generator that will allow individuals to develop and test soil interpretations using the “official” published soil data as well as related data from external sources and perform an analysis in sufficient detail to support the following:

* IT impact analysis by the Enterprise Architecture Team
* An estimate for the Level of Effort (LOE) and cost for the remaining phases of the project: requirements, analysis, design, development, testing, and implementation.

### In Scope

#### Current System Features

The current NASIS Interpretations Generator has these the following functions to be included in a new solution:

1. Generation of National, State, and Local Interpretation ratings using the transactional Soils Survey database (NASIS). Official data is defined by the National Soil Survey Handbook 644.02a.
2. Ability to deliver the interpretation ratings on the public-facing database with a yearly refresh .
3. The official source of soil information to the public is the Web Soil Survey, a part of the National Soil Information System. This system provides for the collection, storage, manipulation, and dissemination of detailed and general soil survey information. The system includes certified tabular and spatial data at various scales. The new SIG will not replace Web Soil Survey, but will deliver public facing data to the Web Soil Survey.
4. The Interpretations Generator has a graphical user interface.
5. Rule based fuzzy logic model is used for generating interpretations.
6. The system must have the ability to allow the user to have the option to see what value of each soil attribute was fed into the evaluation system so as to be clear what numbers are being used to derive the ratings. For example, “Depth to Bedrock 46 cm”.
7. Allow users to store interpretive model components and allow access to interpretation rules created and managed by other internal and external users in order to share and to use the interpretations effectively.
8. Interpretation model components must be searchable by keywords to allow effective sharing and reduce redundancy across the system.
9. Have functionality to facilitate troubleshooting of interpretive criteria.
10. Reports in text or html format of interpretive output appropriate for export to PDF or other formats.
11. Although the interpretations have to be consistent at any given point in time, the ability to update interpretations based upon new criteria (response curve,  examples?) must be maintained.
12. The interpretations are delivered in several forms in different applications. The Interpretation Generator support sthe following types of data distribution, as described in detail below:
    1. Hard copy reports (Print Capability) for both internal use and external distribution.
    2. Creation of maps using exported data from NASIS top MS Access.
    3. Electronic transfer of interpretation results to other applications such as APEX and Toolkit.
    4. Distribution process using exports to provide data to other applications, such as Web Soil Survey, Soil Data Viewer Microsoft Access databases and downloads of gridded SSURGO (raster based representation of the vector data).
13. The interpretations generator provides the user the ability to parse out the reasons for a given interpretation and show them graphically on a map. As an example, providing the “very limited” rating along with its reason of ‘flooding’, ‘shrink-swell’, ‘depth to bedrock’, and/or ‘slope’.
14. The Interpretation Generator has to allow users to preview results using data from the transactional NASIS.  Notes: not sure what this is asking for.  Is this for current system or is there new functionality being requested?
15. Allow internal users with authorization access the NASIS transactional database for generating interpretations and to test against newly developed data.
16. External users to not have access to the NASIS transactional data. External users will view and utilize only public facing data sources such as WSS and Soil Data Access (SDA).
17. Note:  are there any other features or functionality in the current NASIS system that need to be included in the new system?

.

#### New Features

The New Soil Interpretations Generator must also have these additional features:

1. Generation of National, State, and Local Interpretation ratings using additional sources of data, not just NASIS transactional Soils Survey Data, including public-facing Web Soil Survey (WSS) data (WSS includes SSURGO).
2. Ability to deliver the interpretation ratings on the public-facing database prior to the yearly refresh date. Official data is defined by the National Soil Survey Handbook 644.02a. The official source of soil information is the Web Soil Survey, a part of the National Soil Information System. This system provides for the collection, storage, manipulation, and dissemination of detailed and general soil survey information. The system includes certified tabular and spatial data at various scales.
3. The system must allow integration of the laboratory data into the interpretive process.
4. Use Pedon data to generate a zonal statistics interpretation map.  NASIS aggregates Pedon soil analysis results for a county level.  SIG should allow use of Pedon for a single location rather than aggregated data.
5. The system shall allow use most authoritative data available for attributes from spatial layers  in order to provide spatially explicit data for an interpretation.such as:
   1. PRISM data - The new system must be able to use PRISM spatial and tabular climate data.
   2. LIDAR data - topographic data.
   3. Hydrologic Layer - proximity to other features like wells and surface water, streams and rivers,
6. The new system should be more efficient and be able to process large volumes of data efficiently and quickly.
7. The system must allow forms-based input of data into the interpretive system to allow generation of Interpretation ratings “on the fly” using:
   1. national criteria on locally collected site specific information,
   2. allowing measured and point data to be inserted into the system, such as slope, depth to bedrock, Saturated Hydraulic Conductivity ksat, Cation Exchange Capacity (CEC), and dynamic soil properties from desktop or mobile applications.
8. Create documentation to fully document the rules, evaluations, and properties used to generate a particular interpretation.  Use a content management system so that graphics as well as text can be used to explain the assumptions and document the methods for the child rules and parent rules and allow storage of notes related to each part of the interpretation.
9. During an interpretation process, data  should exists temporarily for multiple soil component iterations which can be queried during the process, but is not retained after the interpretation has completed.
10. The delivery of interpretations to various different applications needs to be improved. The Interpretation Generator should support the following types of data distribution, as described in detail below:
    1. Improved reporting capability for interpretations results resulting in hard copy reports (Print Capability) for both internal use and external distribution.
    2. Mapping reports integrated with the application and allowing direct query and on-screen viewing of maps and tables.
    3. Automated integration to provide data from interpretation results to other applications such as APEX (note: define APEX), Conservation Desktop (CD) (replacement for CST), Soil Data Access application, and Web Soil Survey.
    4. Improved integration and distribution of data to other products, such as Web Soil Survey, Soil Data Viewer Microsoft Access databases and downloads of gSSURGO.  Simply the processes and automate where possible.
    5. The current interpretations generator provides the user the ability to parse out the reasons for a given interpretation and show them graphically on a map. As an example, providing the “very limited” rating along with its reason of ‘flooding’, ‘shrink-swell’, ‘depth to bedrock’, and/or ‘slope’.  Additional features need to provide the capability to map the collective ‘rating and reasons’ for a given interpretation.
11. In addition to fuzzy logic model the system should provide additional options to  process the input data, such as neural networks, and others to be determined.
12. Ability to run multiple versions of interpretations using the same data, but with different modeling criteria and/or different modeling techniques (fuzzy logic, neural networks, etc.).  Ability to compare tabular or mapping display of versions of related interpretive results using same data but different methodology. The current system will allow this, but the relationship between the version is not captured, it has to be defined manually by the user.
13. Ability to use previous criteria and modeling techniques with new set of updated data, create a new version, to compare with an older version.
14. The system utilization of Web Soil Survey data should not impact the performance, responsiveness, or user experience of WSS.
15. Interpretations are for determination of outcome for land use with soil types.  The system needs to allow optimization for a soil type for a land use in a specific interpretation, by calibrating and fine-tuning the model criteria.  Then finding the best fit of the model for the measured responses  (e.g. interpretation weights).  The current system runs the interpretation manually many times, with manual tweaking of criteria.  The new system should allow automated multiple runs with pre-determined inputs for alternative criteria.   for example:  a criteria defining a response curve might need to be tweaked using different scenarios.
16. The current system utilizes aggregated generalized data.  The new system would also allow option to integrate point data (specific location soil samples with characterization test results) for model development.
17. The System should allow external users to export the interpretive results, and/or WSS data for use within other systems or tools.
    1. Larger volumes of data for export will first allow the user to define the data to be exported using filter criteria, such as location, soil type, land use, etc.  The user will then select an export format or thematic map requests.  Format options currently are pipe delimited text or html.  Other formats are needed such as MS Access, SQL, XML, etc.
    2. Data for a single point, soil sample with soil characterization composition test results, could be exported to a mobile platform in a web service input formation such as json, geojson, html, REST, SOAP.  Note: need clarification from Dylan B. (graduate student at Davis working with Professor Bear, who is now and NRCS employee).  NRCS contracts with Davis to do this kind of mobile application.
18. The Interpretations Generator must have an easy to use graphical interface that walks the user through the process of developing interpretations, running them, and obtaining the output.
19. The Interpretations Generator must have the capability to support more complex interpretations by interfacing with external applications that already contain logic to perform part of the interpretation.
20. The Interpretations Generator must have the ability to support both maps generated on the web as well as the ability to supply data to stand alone GIS systems and statistical analysis systems.

### Out of Scope

* The Soils Interpretations Generator does not include delivery of completed models and Interpreted data to the public and end users.
* The Data Delivery to the public and end users is a separate solution.
* Initially the Soils Interpretations may be delivered via existing applications such as Web Soils Survey (WSS), Soil Data Access (SDA) or Soil Data Viewer (SDV) which use the Soil Data Mart (SDM) as the data source.
* As the existing Data Delivery applications are modernized, the Soils Interpretations Generator may need to be enhanced to accommodate new requirements defined for end user.

## Assumptions

* Web based Application will be piloted and or prototyped internally to validate and improve the solution prior to making it available to external users.
* Application will ultimately serve all user types; employees, affiliates, customers, TSPs and general public.
* Authentication is required for all users who have ability to create or modify records using a USDA Enterprise Authentication method.
* Role based authorization sing an enterprise solution will be used to distinguish users with permissions to create, modify or run interpretations.
* Users of the interpretation generator include administrators, soil scientists, data stewards, and external customers.
* The Soils Interpretations Generator will be used to create ‘models’ as rule sets that will be used users to create Soils Interpretations by submitting different data sets. Official Soils Interpretations will be delivered to the public and other users via a public access Data Delivery solution separate from the Interpretations Generator.
* Solution must conform to enterprise technical requirements examples of which are not limited to authentication, authorization, security, PII, 508 compliance, performance, availability, etc. .