**Idea Document – Dynamic Soil Properties (DSP) and Soil Health (SH)**

**Key Stakeholders**

* + Luis Tupas, Deputy Chief, SSRA
  + Diane Gelburd, Deputy Chief, S & T
  + Amanda Branham, Special Assistant to the Deputy Chief, S&T
  + Dave Hoover, Director, SSRA - NSSC – Initiating Change Leader, Director NSCC
  + Bianca Moebius-Clune, Initiating Change Leader, Science & Technology, Director Soil Health Division
  + Francine Lheritier – Initiating Change Leader - SME
  + Skye Wills – Advisor - SME
  + Kyle Stephens – Initiating Change Leader - SME
  + Brandon Smith - Initiating Change Leader – SME
  + Drew Kinney, National Leader, Soil Business Systems – Sustaining Change Leader

**Background**  
On June 16, 2017, the U.S. Department of Agriculture (USDA) established the Farm Production and Conservation (FPAC) Mission Area. The FPAC Mission Area is USDA’s focal point for the nation’s conservationists, farmers, ranchers, producers, forest landowners, as well as federal, state, local, tribal, and private partners to seek assistance with crop insurance, conservation programs and technical assistance, commodity lending, and disaster programs. FPAC has over 26,000 employees working in 2,200T6 offices located throughout the fifty (50) States, American Samoa, Mariana Islands, Palau, Puerto Rico, and the Virgin Islands.

The FPAC Mission Area realigned three agencies; the Farm Services Agency (FSA), the Natural Resources Conservation Service (NRCS); and the Risk Management Agency (RMA). The realignment also created the FPAC Business Center (FBC), considered the fourth agency within FPAC.

Specific back round on the two divisions that that are the focus of this initiative are:

Soil Science and Resource Assessment Division: SOIL PROPERTIES

Within NRCS is the Soil Science and Resource Assessment (SSRA) Deputy Area that has operational control of the Soil and Plant Science Division, National Soil Survey Center (NSSC), Dave Hoover, Director. Within the NSSC are there are 6 branches. Those impacted for this initiative is the Soil Business systems and Soil Science Research.

The SPSD products are available for internal, other federal agencies, and private or public partners. Their products are critical decision-support tools to conservationists, farmers, ranchers, producers, as well as federal, state, local, tribal, and private partners. Their products are the authoritative source for soil, plant, and ecological information for the United States, its territories, and international cooperators, making it the largest source of natural resource information in the United States.

Science & Technology Division: NRCS Soil Health Division: SOIL HEALTH

Over the last decades the critical importance of applying principles of soil biology and interactions between soil biological, physical, and chemical processes toward system-based soil health management, has gained significant recognition and understanding. Improving soil health on our nation’s agricultural lands will allow farmers and ranchers to simultaneously improve water quality, increase soil water availability, enhance resilience to extreme weather, enhance nutrient cycling, increase carbon sequestration, provide wildlife habitat (including pollinators), enhance rural economic opportunity, and meet the food production needs of a rapidly growing population on a shrinking available land base. Efforts to improve soil health will thus provide significant return on the nation’s conservation investment. In 2014, recognition and demand from stakeholders led to the initiation of the new NRCS Soil Health Division. The purpose of the new Soil Health Division is to incentivize and facilitate producers in implementing science-based, effective, economically viable soil health management systems on the nation’s diverse agricultural lands, in collaboration with partner organizations. Key goals of the Soil Health Division include providing advanced training to diverse audiences, and facilitating soil health assessment, farm/ranch scale soil health management planning adapted to local conditions, and assistance for on-the-ground implementation of soil health management systems

**Business Need**

This initiative is to support the needed data collection for soil health or dynamic soil properties database for storage of dynamic soil properties which include CIG and EQIP funded related soil health field and laboratory metrics, operational outcomes, land use and management information that could be linked to current and past conservation practice standards.

This information will also provide needed data to provide for the compilation and analysis of effective conservation practices for soil health and conservation programs.

The Dynamic Soil Property initiative is to create a Dynamic Soil Properties (DSPs) data hub to include properties that change with land use and management on a human time scale. DSPs are often a measure of soil health, ecosystem change, and conservation evaluation. Successfully documenting DSPs requires novel data collection and analysis. DSP projects combine disciplines and properties as well as collection hierarchies across space and time that do not easily fit in any existing data infrastructure.

The DSP data hub would store standalone information about the CIG SH Demo project context (and eventually on aggregation and extrapolation) as well as fetch and link to data in other databases, such as: soil properties (NASIS), ecological sites, states and processes (EDIT), and management information (CR-LMOD). Ingestion of cooperator data into the National Cooperative Soil Survey Repository will be facilitated through storage of soil health methods metadata and associated information.

Intended Audience are the users of a soil health and dynamic soil property database which include CIG Conservation Practice Database (CPD) users, NRCS employees and technical services providers, National Cooperative Soil Survey cooperators (e.g. universities, state and local governments, non-government organizations, and other federal agencies) and other interested individuals. Authentication will likely be required for some portion of intended audience. The published/accessible database will have the same user base, but no level will require authentication.

The purpose of this document is to define the high-level business requirements that will support the collection of the Dynamic Soil Properties and Soil Health data. This will include functionality for soil health (including soil biology), agronomic management, technical soil service activities, and conservation programs.

Functionalities:

1. Interactive relational database storing soil properties, location, field management information, soil properties and metadata over time.
2. Data elements needed
   * 1. Rationale from projects and data collected (why, what does it represent)
     2. Multi-scale hierarchy of field, map unit and sample location information (spatial, temporal, and conceptual)
     3. All other site identifiers available in NASIS
     4. Calculated and derived fields to aid in aggregation and reporting
3. Access for NRCS user data input and data curation
4. Data structure that captures spatial and temporal elements (potentially ArcGIS server and novel databases)
5. Manipulation and reporting capability (output)
   * 1. Soil properties
     2. Management
6. Links to existing corporate databases, tools and applications (NASIS, CR-LMOD etc.)
   1. CIG CPD
   2. NASIS
   3. CR-LMOD – current and historical land use, management and operation
   4. CD/NPAD- conservation plans and conservation practices
   5. EDIT – description of ecological sites, including state and transition models with application to multiple land uses
   6. SCAN or other soil climate and weather data
   7. Outcomes (under development with Soil Health Division)
      1. Function measures
         1. Production/yield
      2. Economic outcomes
         1. Inputs
         2. Profit
         3. Risk assessment?
7. User interface to input data with choice lists that guide users to use the same terminology
8. Links to existing corporate databases – choice lists
9. DSP and Soil Health products available through multiple interfaces, should include:
   1. Spatial and tabular data over time
   2. Soil Survey information about soil series, map units and properties specific to land use and management scenarios
   3. Spatial maps with predicted properties under variable land management scenarios
   4. Soil Health indicator potentials (both tabular and spatial)
10. Accessibility of data:
    1. Query builder- so users do not need to know SQL to extract data
    2. SQL – maintain flexibility to use SQL
    3. Download/package datasets for R analysis
11. Integrated version control for the data
12. Ability to track and curate data with multiple levels of authority
    1. eauthentication will be required for some levels of DSP HUB users

**Vision**

Seamless data entry and delivery of DSP and SH data and supporting information for use in NRCS programs and land managers, by staff, cooperators, and citizens.

**Current State**

Currently, the SPSD supplies the processes for data model standards, data collection planning, data collection, data analysis, data interoperability, data management, data interpretation, research, and product delivery. These processes currently involve manual and automated processes to achieve system integration. The CIG CPD is underway and does not currently have a data store, and needs a ways to connect to other soils data house and produced by NRCS.

Dynamic soil property (DSP) and soil health data are soil properties that change quickly with land use and management. In order to interpret that data and inform conservation programs; metadata about the purpose and arrangement of sampling is needed. Currently data are stored in NASIS, the KSSL LIMS system, so independently developed access databases and a series of spreadsheets kept independently by SPSD, SHD, and NRCS state specialists. Adherence to common definitions and choice lists (or terminology) is uneven and maintained through individual understanding and activities.

The soil survey collects the physical and chemical properties of soils and managing the data in a national database (NASIS). The physical and chemical properties of soils are just two of the major property groups that make up soils. The third and much needed property group, which was not collected during the soil survey and is not in the database is the biological properties group. The inability to capture complex spatial and temporal complexities (including management, weather and initiative/project information) We will never fully understand soils and be able to manage the resource well, until we have collected the biological data and understand the relationships that exist between the physical, chemical, and biological soil properties.

**Future State**

This final delivery product will support the needs of Conservation Innovation and Grants and other systems and related conservation programs in the future.

Intended Users of the dynamic soil property data will include NRCS employees and technical services providers, National Cooperative Soil Survey cooperators (e.g. universities, state and local governments, non-government organizations, and other federal agencies) and other interested individuals. Authentication will likely be required for some portion of intended audience. Other applications may also pull from this data.

The intended final product is to have:

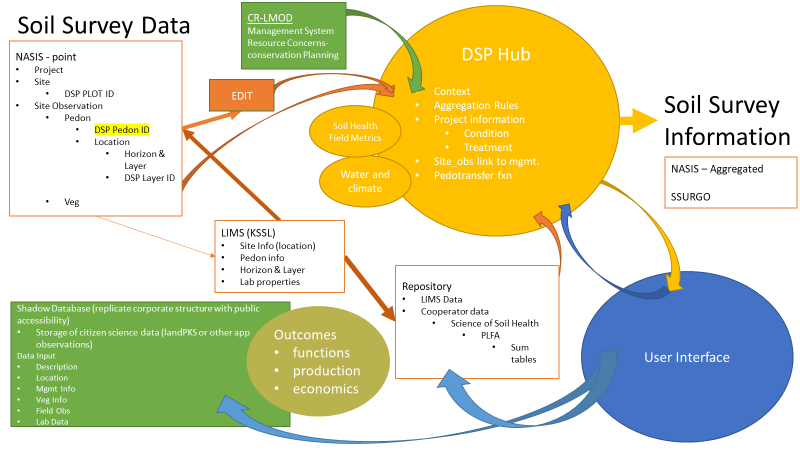
* + - 1. Used to evaluate CIG and “outcomes” of implemented conservation practices
      2. Soil Survey information about soil series, map units, properties specific to land use and management scenarios, and interpretations.
      3. Spatial map will predict properties under multiple scenarios.
      4. Soil health indicator potentials, report card, and reference values will be supplied for all soil series.
      5. Data visualization tools that enable the data to be displayed spatially, tabularly, and temporally
      6. Availability to access raw data or interpreted data

The proposed DSP database would contain four major internal elements:

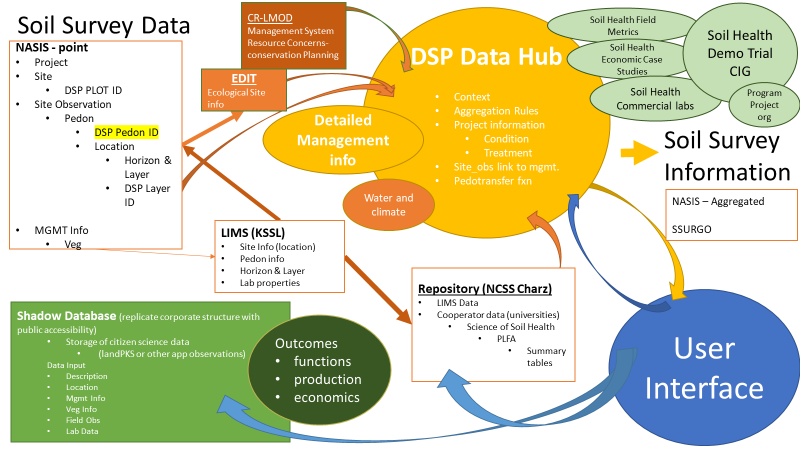
1. Data Source Distinction – the ability to allow multiple entities to input and access data: technical service providers, CIG participants, and internal NRCS personnel.
2. DSP and Soil Health Properties – this includes soil biology and other soil properties not traditionally captured in soil survey databases. This would include both field and laboratory methods such as those that might used for Soil Health Demonstration Trials
3. Management System information – this includes historical and current land use, management system and practice information. This should include all inputs and operations (which can be linked to conservation practices) as appropriate.
4. Hydric soils, hydrology and climate data – this includes hydric soil measures as gathered according to the technical standard, enhanced temporal recordings of water quality and quantity as might be collected through edge-of-field monitoring, and soil water and temperature data in cooperation with the Henry Mount Soil Water and Climate database.

In addition, the system must include:

1. Modular Input Interface – the type and amount of data and information entered would depend of the type of user
2. Integration into Current Corporate Data Structure –
   1. Associations and linkages required to current corporate databases and applicable assessment tools including CIG CPD and soil (e.g. NASIS and EDIT)
3. Multiple Output Streams – both raw data and aggregated information should be available through current and future interfaces.



***Figure x. High Level Schema***



DSP Data Hub:

Transactional (Project Context, Aggregation Rules, Extrapolations Rules)

Links/connections:

CIG CPD

NASIS

CRLMOD

EDIT

Outputs: Reports and API’s

**Business Value**

This will support the CIG CPD project, and serve as the datastore for the CIG soil health demo trial information and other associated CIG information as required. Soil survey databases collect, aggregate and disseminate data and information on a range of physical and chemical soil properties and interpretations. This expansion expands the information provided to customers including soil climate and soil biology. In order to interpret and apply soil information, land use, management and conservation practice information must also be gathered and associated with other soil data elements.

The enabling technologies will support soil and plants sciences activities by improving SPSD’ s ability to collect, manage, analyze, and interpret soil survey data applying the latest scientific methodologies. Soil survey products will be delivered leveraging trending technologies in a timely fashion and enhance all other NSCC operations.

**Consequences**  
Without modernization of the SPSD data production and delivery systems, USDA – NRCS conservation programs will be adversely impacted due to inability to maintain and update natural resource information. Additionally, external partners will steadily shift from using USDA-NRCS natural resource information to non-authoritative sources. Another consequence would be that the CIG CPD would not function properly and the CIG soil health demo participants would have not data storage for their projects.

**Business Change**  
SPSD will leverage new enabling technologies to ensure that USDA manages an adaptable soil survey area data store. This will result in a reduction on resources spent managing multiple soil survey area data stores and reduce the latency in data set refreshes.

Frequent data model changes that are deployed on a timely and on an as needed basis will improve conservation planning.

Ability to deliver new products in new formats such as raster.

**High-level Features and Functional enhancements**

A data hub that includes both stand-alone data and information about CIG SH Demo project context (and eventually on aggregation and extrapolation) as well as connections to existing databases in order to fetch and link to data in other databases, such as: soil properties (NASIS), ecological sites, states and processes (EDIT), and management information (CR-LMOD). Ingestion of cooperator data into the National Cooperative Soil Survey Repository will be facilitated through storage of soil health methods metadata and associated information.

* **All users of soil health and dynamic soil property data and information should have access to the database**.
  + NRCS employees and technical services providers
  + Levels of access and authentication will be variable for the transactional portion of the database (may require curation from a database manager)
  + Export or published data, information, and reports will be publicly available without authentication.
* **Development of web services and data delivery systems, or enhancement of existing interfaces, web services, and delivery systems** (e.g. WSS, Soil Data Access).
* **Associations and linkages required to current corporate databases and applicable assessment tools** including CIG CPD, soil (e.g. NASIS, and/or ESIS)

**Requirements**

|  |  |  |  |
| --- | --- | --- | --- |
| **2.1** |  | **Required?** | **Priority** |
| 2.1.1 | Project information  Include purpose of effort, design of sampling or observation  Multi –scale hierarchy of field, map unit and sample location information (both spatially and conceptually)  Include linkages to corporate databases including for example CIG CPD, NASIS, EDIT. | YES |  |
| 2.1.2 | Management information  Include current and historical management operations using CIG CPD and current corporate database (such as CLMOD)  Include links for regional and local climate and weather information | YES |  |
| 2.1.3 | Individual Sampling Units (Point Data)  Include location information including GPS coordinates and depth information based both of depth intervals and genetic horizons (could potentially link to genetic horizon in NASIS)  Include time and date information including local conditions, seasonal climate differences (conditional outside of “typical”, “normal”, “historical” ranges), moisture status, etc.  Include option for element selection (properties, methods, observations and techniques will have multiple, modular inputs and require multi-level relational databases with point measurements). Types of data that will be accommodated:   * Method metadata (include choice lists and options for new method entry) * Monitoring data (temporal replication from yearly to hourly intervals) | YES |  |
| 2.1.4 | Calculated/derived fields  Include capacity to use inputs of multiple data elements to produce results in relevant units. | YES |  |
| 2.1.5 | Aggregated Data and Information  Include capacity to aggregate and report data elements according to spatial and project design elements.  Aggregation should have both an automated (preferred) technique and be customizable through user input (2.1.6) | YES |  |
| 2.1.6 | Output/export through multiple interfaces  Include dynamic soil property elements through existing data sharing portals (e.g. Soil Data Access).  Include options that allow the user to adjust calculation and aggregation scheme. |  |  |
| 2.1.7 | Integrate with existing databases  Previous elements require a robust, modular metadata structure and data dictionary. Elements in or associated with the dynamic soil properties database need to be compatible with existing databases. (e.g. NASIS, EDIT, etc.) and the new CIG CPD. Allowing for this type of functionality may require an adhoc or similar strategy for the associated domain(s) or data dictionary.  It maybe reasonable to consider a transactional GUI that will interface with and populate multiple databases simultaneously | YES |  |

**Business Proposed Execution Roadmap**

Planned Roadmap – **TBD** based on priorities with the planned initiatives

