**Idea Document – National Soil Information System (NASIS) Modernization**

**Key Stakeholders**

* + Luis Tupas, Deputy Chief, SSRA
  + David Lindbo, Director, SPSD – Business Owner
  + Dave Hoover, Director, NSSC – Initiating Change Leader
  + Drew Kinney, National Leader, Soil Business Systems – Sustaining Change Leader
  + Michael Robotham, National Leader, Technical Soil Services/World Soil Resources – Sustaining Change Leader
  + Joel Brown, National Leader, Ecological Sites –Sustaining Change Leader
  + Steve Peaslee – Subject Matter Expert (SME)
  + Charles Ferguson – SME
  + Dylan Beaudette – SME
  + Adolfo Diaz – SME
  + Kevin Godsey – SME
  + George Teachman – SME
  + Jason Nemecek – SME
  + Tammy Cheever – SME
  + Kyle Stephens – SME
  + Gerry Moore – SME

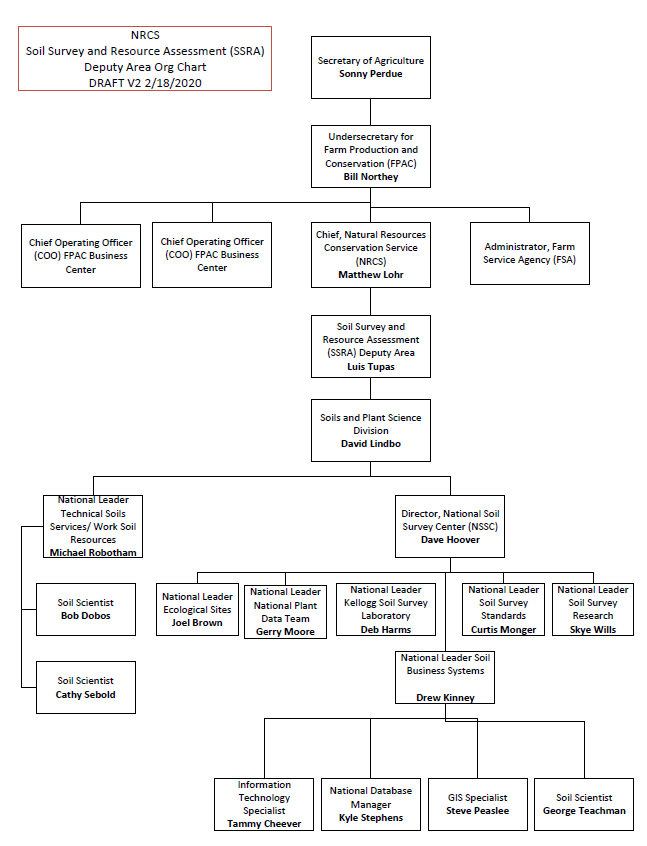
**Background**  
On June 16, 2017, the U.S. Department of Agriculture (USDA) established the Farm Production and Conservation (FPAC) Mission Area. The FPAC Mission Area is USDA’s focal point for the nation’s conservationists, farmers, ranchers, producers, forest landowners, as well as federal, state, 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 2,200 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.

Within NRCS is the Soil Science and Resource Assessment (SSRA) Deputy Area that has operational control of the Soil and Plant Science Division (SPSD), David Lindbo, Director. The SPSD has operational control over the Technical Soil Services/World Soil Resources branch, Michael Robotham, National Leader and over National Soil Survey Center (NSSC), Dave Hoover, Director. Within the NSSC are there are 5 branches and 1 team. They are:

* Soil Business Systems – Drew Kinney, National Leader
* Ecological Sites – Joel Brown, National Leader
* Kellogg Soil Survey Laboratory – Deb Harms, National Leader
* Soil Survey Standards -- Curtis Monger, National Leader
* Soil Survey Research -- Skye Wills, National Leader
* National Plant Data Team, Gerry Moore, National Leader
* The SPSD manages the United States soil, plants, and ecological sites resource inventory and provides soil interpretations. The NSSC branches and team are responsible for providing technical leadership and expertise in support of efforts to expand the availability of soils and plant information and the understanding of soils and plants.

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.

Figure 1. High Level Organizational Chart

**Vision**  
Within the next three years, the Soil and Plant Science Division (SPSD) intends to modernize soil, plants, and ecological sites resource inventory supporting tools to enhance the customer services through adaptable, efficient, scalable technical solutions to streamline processes to free up resources for mission critical work. SPSD intends to eliminate the risks associated with outdated technology to increase efficiency, data integrity, quality, and availability.

Within the next three years, the Soil Business Systems Branch intends to modernize and integrate our information systems to meet our customers' needs, mission requirements, and to support putting conservation on the ground. Within the same time period, the team intends to adapt new technologies and applications to address emerging trends in soil and plant sciences.

Within the next three years, the Ecological Sites Branch intends to have the Ecosystems Dynamic Interpretative Tool (EDIT) program within the family of the soil and plants information systems, integrating it with other conservation programs support tools.

Within the next three years, the National Plant Data Team we will modernize our database and website to include a dynamic content manager; as well, as increasing the processing and delivery of information to our customers. Within the same time period, the team will upload the multi-year backlog of plants data and develop and implement product delivery metrics.

Within the next three years, the Soil Survey Research and Laboratory intends to improve the management of laboratory data systems and leverage improvements of the soil and plants data systems in order to use the combined information in streamlined inventory delivery.

Within the next three years, the Soil Surveys Standards Team intends to evolve training methods to support new technologies, products, and processes in support of natural resources inventory and conservation. To develop and maintain scientific world-wide natural resources inventory standards in order to support conservation that is fact-based, data-driven, science-based and customer focused.

**Business Need**

* + The Soil and Plant Science Division (SPSD) customers must have access to authoritative science-driven up-to-date soil, plant and ecological site data. The data must be easily accessible in multiple formats from a single source through various technologies.

The business need will be addressed by implementing current technological solutions to modernize data production and delivery systems that allow us to adapt to changing customer needs.

**Current State**

The National Soil Survey Center (NSSC) business capabilities were mapped and assessed.

The mapping process included linking NSSC to the NRCS business capabilities; linking NSSC business capabilities to the USDA business capabilities reference model; defining each business capability, listing the activities associated with each business capability, and listing the systems and tools supporting each business capability. Following are the NSSC business capabilities:

* Standards Management
  + Definition. Develop, maintain, implement and enforce standards and provide training on same.
  + Activities.
    - Develop and maintain soil survey standards for mapping, classification; interpretation, investigation, analysis, and for data management;
    - Develop standards for disseminating soil survey information;
    - Develop and implement the National Cooperative Soil Survey training program;
    - Develop and implement the training program for soil, plant and ecological data.
  + Systems/Tools.
    - Adobe Connect
    - Panopto video indexing software
* Cooperative Relationship Management
  + Definition. Negotiate and establish formal work agreements with third parties and manage the execution and fulfilment of those agreements.
  + Activities.
    - Develop and execute cooperative research agreements with Universities;
    - Develop and execute memorandums of understanding (MOU) with other federal agencies;
    - Establish MOU's or Research Agreements for field office soil survey projects specific deliverables;
    - Manage the MOUs through to completion.
  + Systems/Tools.
    - Cooperative Ecosystem Studies Unit (CESU) MS Access database
* Data Collection
  + Definition. Evaluate, design, develop, operate, maintain and improve the data collection network systems and technology that provide reliable soil data and information. Evaluate and develop soil information technology; develop applications; design systems; collect and deliver soil data and information.
  + Activities.
    - Develop, operate and maintain the Laboratory Information Management (LIMS), the National Soil Information System (NASIS), the PLANTS database, and the Ecosystem Dynamic Interpretive Tool (EDIT);
    - Collect soil and ecological data from 130 field offices;
    - Combine data and land attributes with Geospatial info to facilitate mapping.
  + Systems/Tools
    - NASIS
    - LIMS
    - PLANTS database
    - EDIT
    - SQL server
    - SQL Lite
* Data Management
  + Definition. Quality control of collected data. Administer and manage databases.
  + Activities.
    - Work with data models, add data fields, add interfaces with multiple data sources;
    - Internal data quality checks;
    - Monitor system performance.
  + Systems/Tools. No unique systems related to this capability.
* Data Analysis
  + Definition. Measure and examine data and information.
  + Activities.
    - Conduct research on the genesis, landscape distribution, function and behavior of soils;
    - Interpret and predict the impacts of natural and human disturbance on soil and ecosystem function;
    - Conduct research on collected plant and other ecological information;
    - Analyze soil maps to produce the Prime Farmlands inventory.
  + Systems/Tools.
    - SAS Enterprise Business Intelligence
    - JMP Data Analysis Software
    - R for statistical computing and graphics
    - Python
    - ArcSIE
    - ArcGIS
* Data Distribution
  + Definition.
    - Evaluate, design, develop, operate, maintain and improve the data distribution network systems and technology that deliver data and information to the user community.
  + Activities.
    - Develop, operate and maintain systems that provide reliable, real-time, Internet based access to data and information;
    - Make available a database/inventory of soil and ecological resources for land use decisions;
    - Provide user-friendly information that explains how soil characteristics and properties effect everyday life.
  + Systems/Tools.
    - Web Soil Survey (WSS)
    - Soil Data Access (SDA)
    - Soil Data Viewer (SDV)
    - PLANTS database
    - EDIT
    - SQL server
    - Web contribution tool

The business capability assessment identified improvement opportunities in the areas of people, processes, information, and technology (PPIT). As shown in figure 2, committing resources to work on these improvement opportunities will also positively impact the NRCS Scientific and Technical Research and Business Management business capabilities.

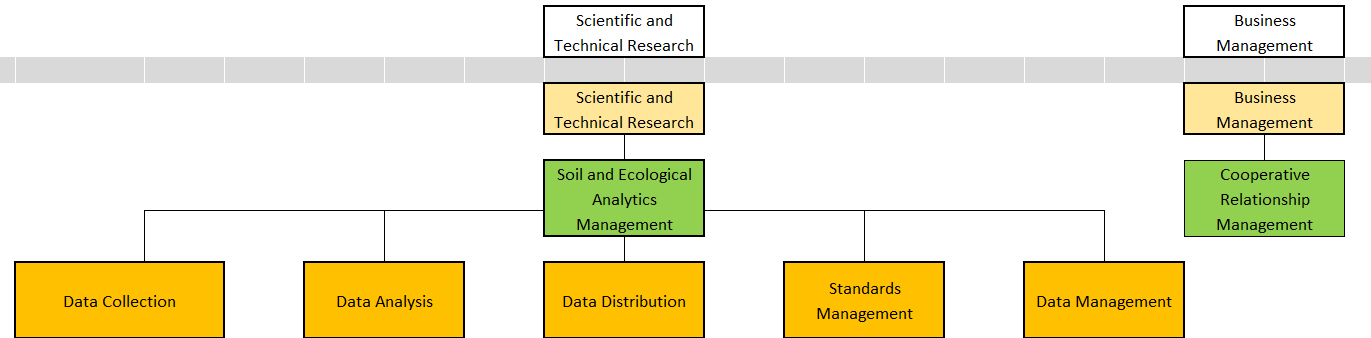


Figure 2. National Soil Survey Center Business Capabilities

Currently, the SPSD relies on outdated technology and 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 time consuming manual and automated processes to achieve system integration.

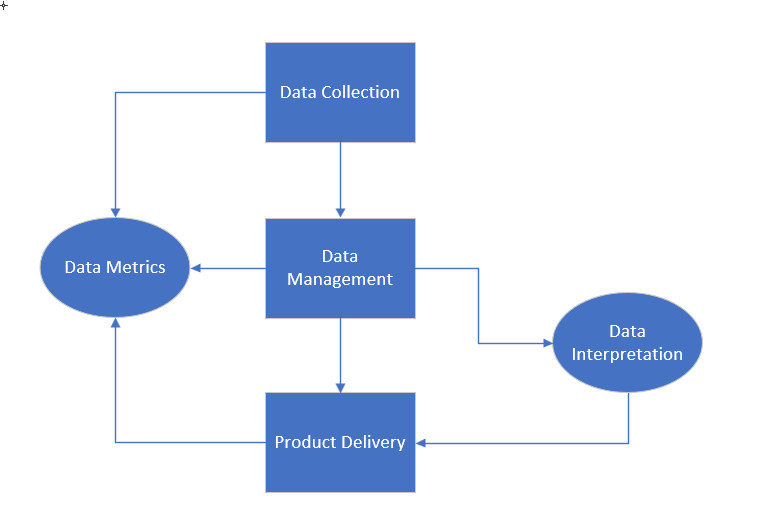


Figure 3. High Level Business Process Workflow

* Data Management, QA, and Analytics
  + Tabular and geospatial data is stored separately
  + Business manually and systematically validates and aggregates tabular and geospatial data
  + Business systematically compares tabular and geospatial data to ensure they are within the same specified date range
  + Tabular data is versioned each time it is changed
  + Geospatial data is versioned each time it is changed
  + Aggregated tabular and geospatial data is versioned each time it is aggregated
  + Soils survey data and plant data is manually extracted, transformed, and loaded into the customer-ready data set
  + Business systematically commits customer-ready data to the soils warehouse
  + Customer-ready data is extracted to create customer-centric products
  + Data retrieval limitations require multiple data extracts in order to rebuild a soil survey area data set for customer use
  + Business creates snapshots of prior aggregated data for ad hoc product creation and trend analysis
  + Data searches and data recalls are limited by security-driven roles and permissions, increase the amount of time to delivery time-sensitive product requests
  + Mandated device scans reduce or abort data management, analysis, and quality assurance procedures
  + Soils survey authoritative data sets are not deployed to production in a timely manner
  + Data is archived and retained permanently
  + Archived data retrieval requires business-created procedures for proper identification
* Product Delivery
  + Each functional area chooses the product that will be delivered
  + Product delivery is accomplished outside the traditional development environments
  + Business partners create products, including a mobile application
  + Cannot adapt to changing demands on technology advances for delivery venues

**Future State**

***Will we keep this process workflow in the future?***

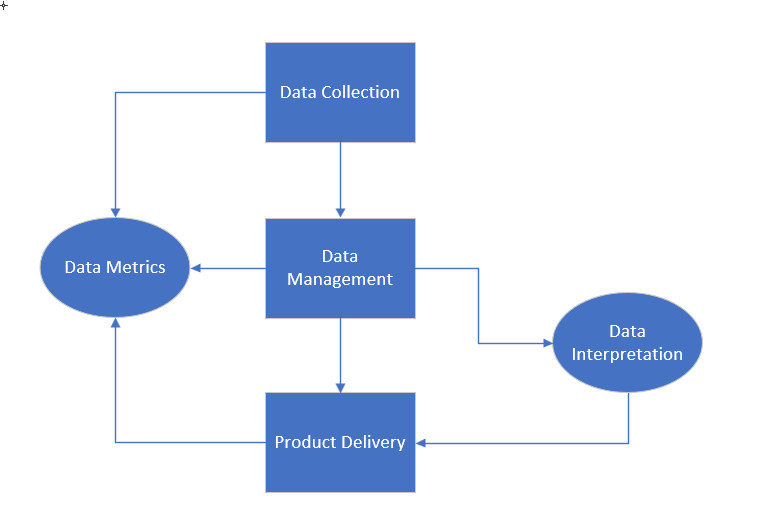


Figure 4. High Level Business Process Workflow

* **Definitions and future requirements of boxes in New Version of Figure 3.**
  + **Data Collection:**
    - *Definition – Collection of point data in the field and production of soil polygons and raster layers for inclusion in database. This includes the part of the process from field work to uploading information into the transactional database in preparation of data management.*
    - Future Needs:
      * One solution for uploading point data to the soils database which will includes quality control processes.
      * Creating a method for uploading information from a point shapefile or feature class directly into the database.
  + **Data Management:**
    - *Definition – Main concept is the conducting of analysis, aggregation, quality assurance, and quality control on data in the transactional database (spatial and tabular) in preparation for publication. This includes the part of the process where field data that has been entered into the transactional database and is managed in preparation for versioning and delivery. Additionally, includes connection to reference data in external databases, managing technical soil service activities, managing soil survey projects, managing database models, and managing all components of OSD.*
    - Future Needs:
      * Create a single database that serves as both the production and publication database for both spatial and tabular data. Will contain versions in order to replace Soil Data Warehouse
  + **Data Interpretation:**
    - *Definition –The generation of ratings that predict soil behavior under specified use and management. Ratings are based on soil properties and site-specific properties such as slope, elevation, temperature, and precipitation. This includes the part of the process where the component interpretation table is generated and published based on a set of rules, evaluations, and properties stored in the NASIS db.*
    - Future Needs:
      * Need an interpretation generator that allows for onsite data to be fed into the interpretation instead of only relying on the data stored in the NASIS aggregated data. **The existing Interpretation Generator in NASIS will need to do the same things as the stand-alone Interpretation Generator, using just authoritative data.**
  + **Product Delivery:**
    - *Definition – Certified aggregate data (soil mapunits) are versioned and the final database model is created and delivered to the public. This includes the part of the process that occurs after the aggregate data is certified up until the end products reach the customers.*
    - Future Needs: NSSC group will work on this and Danielle will incorporate
  + **Data Metrics:**
    - *Definition - Monitoring of database and application usage, with a major focus on who is using our published data, how often it is being accessed, and what are the most common data requests.*
    - Future Needs: *Determine* *who is using our published data, how often it is being accessed, and what are the most common data requests by customers.*
* Data Management, QA, and Analytics
  + Tabular and spatial data are managed from a single point of access. The single point of access must handle multiple data models. The single point of access must handle both vector and raster spatial data.
  + Scalable soil survey area data sets
  + Automated versioning on the aggregated tabular and geospatial data set
  + Soil survey data and plant data is automatically aggregated
  + A State’s soil survey area data set will be retrieved with a single procedure for customer use
  + A proof of concept environment is available with the latest soil survey data are set for ad hoc product creation and trend analysis
  + Group roles and permission are used to enhance the data searches and data recalls from within the DISC environment
  + Mandated device scans are reconfigured, so the current business partner devices are used to their full capacity
  + Soils survey authoritative data sets are deployed to production within 24 hours
  + Data is archived and retained permanently
  + Archived data is catalogued to facilitate efficient and timely retrieval
* Product Delivery
  + Formalize business developed applications
  + Implement product enhancements
  + Reengineer Web Soil Survey, Soil Data Access, and other deliverable methods with a single access dashboard
  + Business intelligence and analytics capability for internal users
  + Minimally operational product should be adaptable to emerging technologies
  + Reduce reliance on proprietary technologies
  + Leverage platform agnostic technologies
  + Deliver a quantitative and qualitative soil survey data set
  + Products should be easily maintainable

**Business Value**  
Conservationists, farmers, ranchers, producers, as well as federal, state, local, tribal, and private partners will have integrated soil, plants, and ecological sites resource inventory. Partners will have up-to-date data to allow them to continue to implement conservation practices that are fact-based, data-driven, science-based, and customer focused.

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 center operations.

**Consequences**  
Without modernization of the SPSD data production and delivery systems, USDA – NRCS farm bill 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.

**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 30% 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**

* As a user in any role I want to eliminate workarounds so I can create efficiencies in delivering data
* As a user in any role I want to reduce the reliance on locally developed applications so that I can reduce the risks associated with locally developed applications
* As a user in any role I want an adaptable and scalable soil survey area data store so that we can deliver customer defined areas of interest
* As a user in any role I want a mean time to repair of one hour so that I can mitigate the currently loss of data as a result of extended outages
* As a user in any role I want a proof of concept environment so that I can create products before they are permanently implemented
* As a user in any role I want the NASIS to support 300 concurrent users so that we can mitigate performance issues
* As a user in any role, I want a single point of entry that provides customer access to all public facing SPSD data that is scalable to a user-defined area of interest and available in different formats.
* As a user in any role I want the SPSD web site coverage failures to be repaired within two hours so that I can prevent stoppage of information delivery to our customers
* As a user in any role I want the SPSD website to present the most recent content so that our customers have actionable decision-support information
* As a user in any role I want the ability to review historical soil survey data for specified years and sites of interest so that I can perform trend analysis
* As a user in any role I want the ability to run ad hoc reports so that I can meet non-cyclic reporting demands

**Business Proposed Execution Roadmap**

