Appendix I. Data Management Technical Plan

The national assessment of ecological carbon sequestration is based on a number of national capabilities including remote sensing, expert knowledge consultation, national inventory programs, land-use and land-cover maps, and simulation models. A data-processing, -management, and -serving system will be needed to provide the national assessment with data input, data output, information query and dissemination, and data-archive functionalities. The basic functions of the system should include the following:

- Remote-sensing data access
- Database building of expert knowledge of ecosystem processes and mitigation scenarios
- Database building of inventory and other in situ data (not the proprietary data held by national inventory programs such as the Forest Inventory and Analysis Program (FIA) of the U.S. Forest Service and the National Resources Inventory (NRI) of the Natural Resources Conservation Service)
- Climate, soil, and other biophysical data compilation required for the assessment
- Derived data storage, including geospatial data
- Table, chart, and report production
- Metadata production automation to the extent possible

This data-management system will consist of three subsystems: a data-storage subsystem; a data-conversion and integration subsystem; and a management, mapping, and reporting subsystem (fig. I1).

I.1. Data Storage Subsystem

The data-storage subsystem will be designed to support storage of data types with different spatial and temporal characteristics. Database software and hardware will be selected to provide the capability of handling large data volumes. An existing Sun Fire 4800 server running Oracle's Enterprise Edition Relational Database Management System (RDBMS) and the Environmental Systems Research Institute's (ESRI) Spatial Database Engine (ArcSDE) will be used for the prototype.

Storage of geospatial data and maps will be achieved by using geographic information system (GIS) database techniques. Spatial and nonspatial indexing will be used to enhance the performance on data searching and loading. The data-storage subsystem must be able to hold raster geospatial data, vector geospatial data, tabular data with no geospatial component, and text data.

The data-storage subsystem will be designed to handle the following datasets that will be collected from existing Federal programs:

- Remotely sensed images
- Vegetation, land-cover, land-use, and change maps
- Flux-tower data
- Climate, soil, and biophysical data

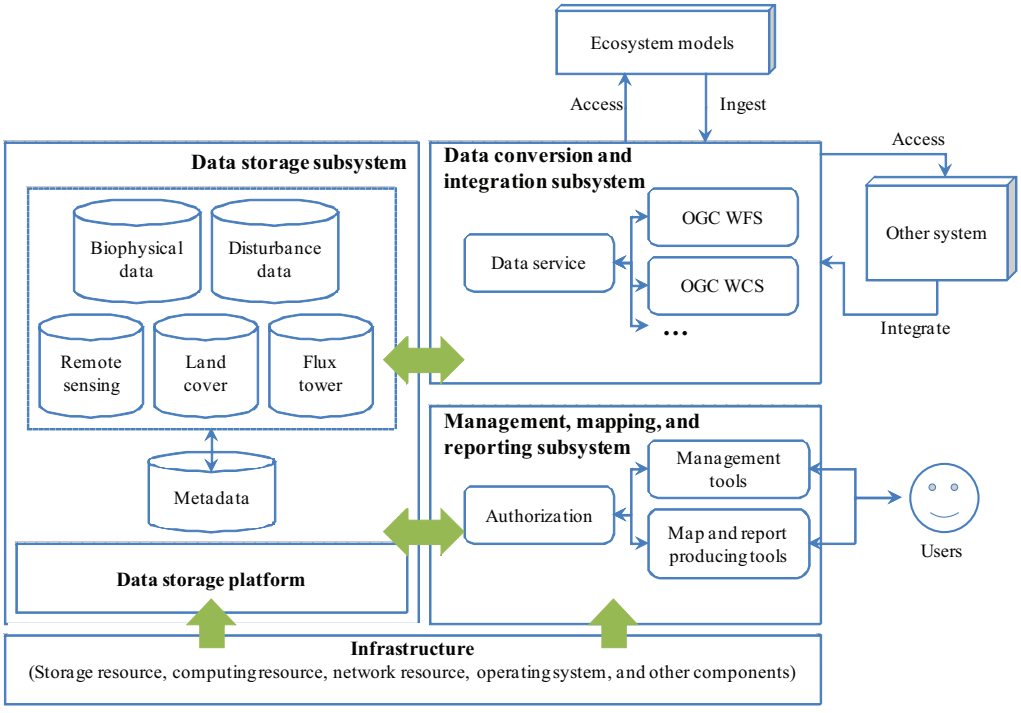

Figure I1. Diagram showing the overview of the proposed data-storage system for the national assessment of ecological carbon sequestration. OGC, Open Geospatial Consortium, Inc.; WCS, Web Coverage Service; WFS, Web Feature Service.

Other datasets will be added to the data-storage subsystem as they are identified as necessary components during the national carbon development assessment. Small to moderate data volumes will be ingested to demonstrate the abilities of this subsystem to handle complex ecosystem data. This subsystem also will provide storage for derived data from the model simulations, as well as maps and reports.

Metadata will be collected and stored in this subsystem. To ensure compatibility, the metadata structure will be designed following the standards of the International Organization for Standardization (ISO) and the Federal Geographic Data Committee (FGDC). This subsystem will be designed to synchronize metadata with associated data when new data are added or existing data are updated.

The Oracle database engine, coupled with ESRI's ArcSDE, will be able to store effectively all of the datasets and data types specified in this document. This data-storage solution is expandable to petabytes to accommodate the assessment requirements; a data-flow diagram is shown in figure I2.

I.2. Data Conversion and Integration Subsystem

The data-conversion and processing subsystem will be designed to provide key functions necessary for seamless integration of various datasets, including data and file-format conversion, reprojection, resampling, and other necessary geospatial transformations. These functions also will enable effective dissemination of results and data products derived through the national assessment. This subsystem will provide these functions via Web services for data access, exchange, and processing using standards-based interfaces. Open Geospatial Consortium, Inc. (OGC) and ISO data-service standards, such as Web Feature Service (WFS) and Web Coverage Service (WCS), will be adopted in the interface design. The interactive interfaces will allow other systems or clients to access the data through the Internet. In addition, this subsystem will provide capabilities for ingesting data from other remote data services and processing services. An overview of the processing subsystem is shown in figure I3.

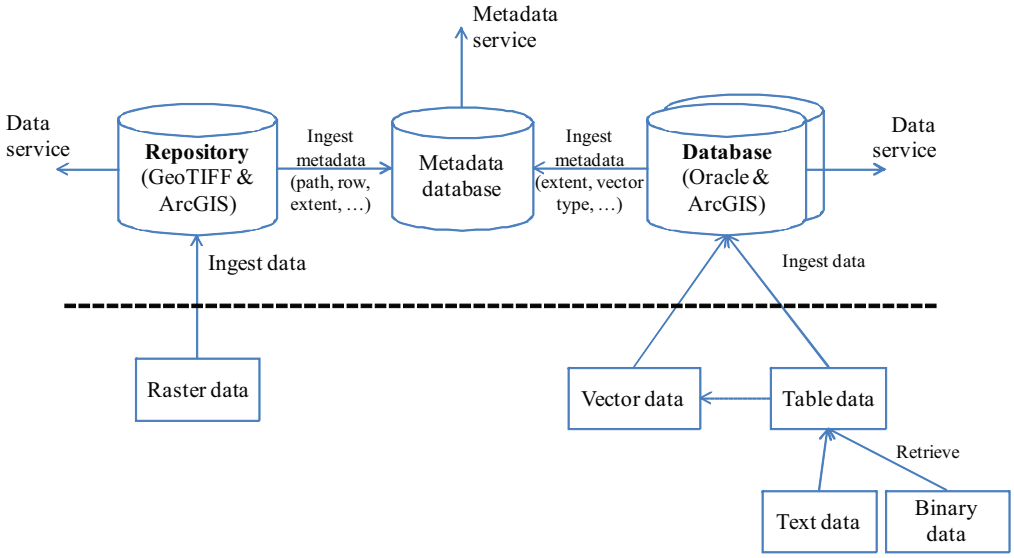

Figure I2. Data-flow diagram for the data-storage subsystem to be used for the national assessment of ecological carbon sequestration.

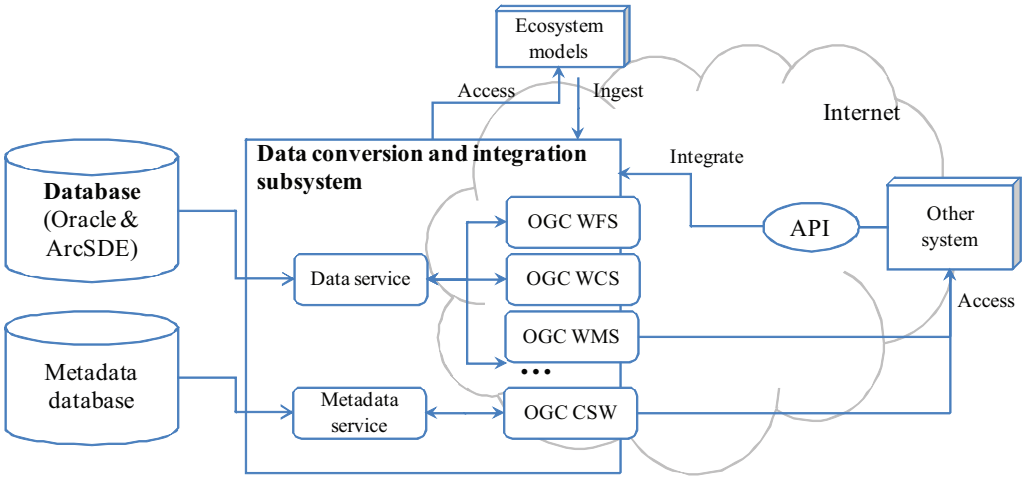

Figure I3. Diagram showing the data conversion and integration subsystem to be used for the national assessment of ecological carbon sequestration. ArcSDE, ESRI's Spatial Database Engine; OGC, Open Geospatial Consortium, Inc.; WFS, Web Feature Service; WCS, Web Coverage Service; WMS, Web Map Service; CSW, Catalogue Service for the Web; API, application programming interface.

For the prototype, OGC services will be provided by using ESRI's ArcGIS Server. The ESRI software for OGC services integrates seamlessly with Oracle and ArcSDE, and allows for minimal development efforts.

I.3. Management, Mapping, and Reporting Subsystem

The managing, mapping, and reporting subsystem will provide Web-based tools for users to view, search, and update their database remotely. User authentication and access control will be implemented to enable secure data access and preserve data integrity. This subsystem also will provide functionalities for producing maps and reports using datasets collected by other Federal partners or produced through the national carbon assessment, or for extracting subsets of these data and saving them into a database or other desired formats. Web-based techniques will be used in developing this subsystem. The ESRI ArcGIS Server software product allows for Web-based delivery and custom tools to be developed.

I.4. System Technology Components

The planned system architecture is built on technology already in use at the Earth Resources Observation and Science (EROS) Center of the U.S. Geological Survey (USGS). The system architecture will use Oracle's Real Application Clusters (RAC) as its database, running ESRI's ArcSDE management software to add capacity for geospatial data. By leveraging Oracle RAC, the database subsystem can be spread across multiple systems and will provide increased availability and performance while using inexpensive commodity hardware. The diagram in figure I4 shows the anticipated system architecture.

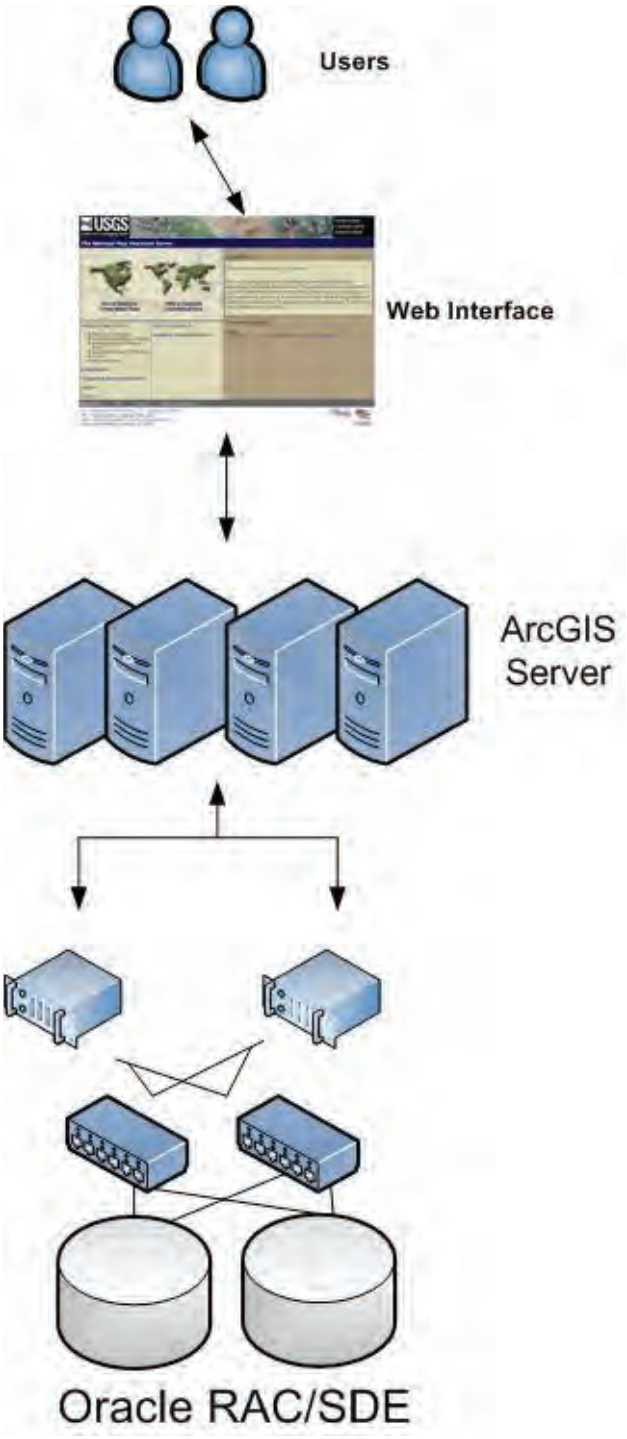

Figure I4. Diagram showing the system architecture to be used for the national assessment of ecological carbon sequestration. RAC, Real Application Clusters of Oracle; SDE, spatial database engine of ESRI.