

DATA INTEGRATION PLAN FOR NATURAL RESOURCES REVIEW COUNCIL DECISION SUPPORT

Natural Resources Review Council GIS Subcommittee

Version 1.0

Data Integration Plan for Natural Resources Review Council Decision Support

Natural Resources Review Council
GIS Subcommittee

July 16, 2013

M. Lee Allison, Arizona Geological Survey, Chair
Glen Buettner, State Forestry
Janel Day, Arizona Geological Survey
Joyce Francis, Game & Fish
Victor Gass, Environmental Quality
Ryan Johnson, State Land
Gene Trobia, State Cartographer, State Land
Jeffrey Wilkerson, Transportation

Contents

Executive Summary.....	4
Natural Resources Review Council (NRRC) Mission.....	9
Natural Resources Review Council (NRRC) GIS Subcommittee Mission	9
Charge to the Subcommittee.....	9
Executive Order.....	9
Expanded Charge	10
Meeting the Charge	10
GIS Coordination	10
Clearinghouse	10
Planning	12
Mitigation Banking.....	12
NRRC Agency Data Holdings	13
Needs and Requirements.....	13
Existing Data, Services, and Resources	13
Data and Services in Each State Agency	13
Data and Services Needs for Each State Agency.....	14
Software Applications in Use	14
Technical Expertise	15
Existing GIS Costs to NRRC Agencies.....	15
GIS Coordinating Bodies.....	15
Arizona Geographic Information Council (AGIC)	15
National States Geographic Information Council (NSGIC)	16
Existing Documents, Surveys, and Assessments.....	16
Opportunities for Discovery, Access, and Integration of Federal Data, Materials, and Projects	18
Western Regional Partnership	18
Federal Open Data Policy.....	19
System Design and Architecture	20
Data Discovery and Access.....	20
State GIS Data Repository (AZGEO)	21
Arizona Data and Document Repository	22
State-Federal Information Clearinghouse.....	22

Integration of System Elements.....	27
Education and Training	27
Risk assessment	28
Gap analysis of current capabilities	28
Discussion	29
Recommendations	31
Costs and Timeline	31
References	33
Appendix List for NRRRC Data Integration Plan	34
1: Executive Order 2013-02	35
2: NRRRC GIS Subcommittee Scope of Work	37
3: Federal Agencies in the Western Regional Partnership	39
4: AZGEO Frequently Asked Questions	40
5: AGIC Guidelines for Geospatial Data Sharing	42
6: NRRRC GIS State Survey	50
7: U.S. Geoscience Information Network (USGIN) System Elements	53
8: USGIN Community Standards, Profiles, Protocols, and Software	57
9: Acronyms	59

Executive Summary

In January 2013, Governor Jan Brewer released Executive Order 2013-02, “Establishing the Arizona Natural Resources Review Council” (NRRC) to “develop land and natural resource management strategies for Arizona and coordinate with state natural resource agencies and their existing management plans.” In response to the Executive Order the NRRC established four subcommittees: Geographic Information Systems (GIS), Planning, Clearinghouse Development, and Mitigation and Conservation Banking. Each subcommittee established an individual scope of work, which when combined, would help roadmap a set of processes and policies for meeting the goals of Governor Brewer’s Executive Order.

This report comprises the roadmap recommendations of the GIS committee to accomplish the following specifics within the executive order:

Task 5. “The Council shall develop a coordinated and centralized Geographic Information System database model that identifies current and future management priorities for designated land and natural resource areas.”

Task 6. “The Council shall identify and prioritize legal, legislative and incentive-based needs that protect and maintain state interests related to wildlife, land, water and other natural resources.”

The NRRC GIS Subcommittee proposes the establishment of an integrated state-wide data system to act as a GIS clearinghouse, a Single Point of Contact for the State of Arizona (AZSPOC) to receive Federal notifications, and an online catalog of all NRRC Agency documents. The integrated data system seeks to maintain agency independence in software selection and choice for internal data and document management while providing end-users simplified search functionality across NRRC Agencies. In addition, this plan incorporates the work of the Arizona Geographic Information Council (AGIC) through the integration of AZGEO, AGIC’s comprehensive plan for a centralized GIS repository and clearinghouse (AGIC, 2013).

The conceptual design for the state-wide data system is displayed below. End users, including NRRC Agencies, will access the data information resources – both document and GIS – through the NRRC User Interface to the Catalog. The Catalog incorporates metadata from each of the NRRC Agencies, AZGEO, and the proposed AZSPOC simplifying search across land-use related data. Within the AZSPOC, Figure 2, demonstrates the notification and response processes recommended in this approach.

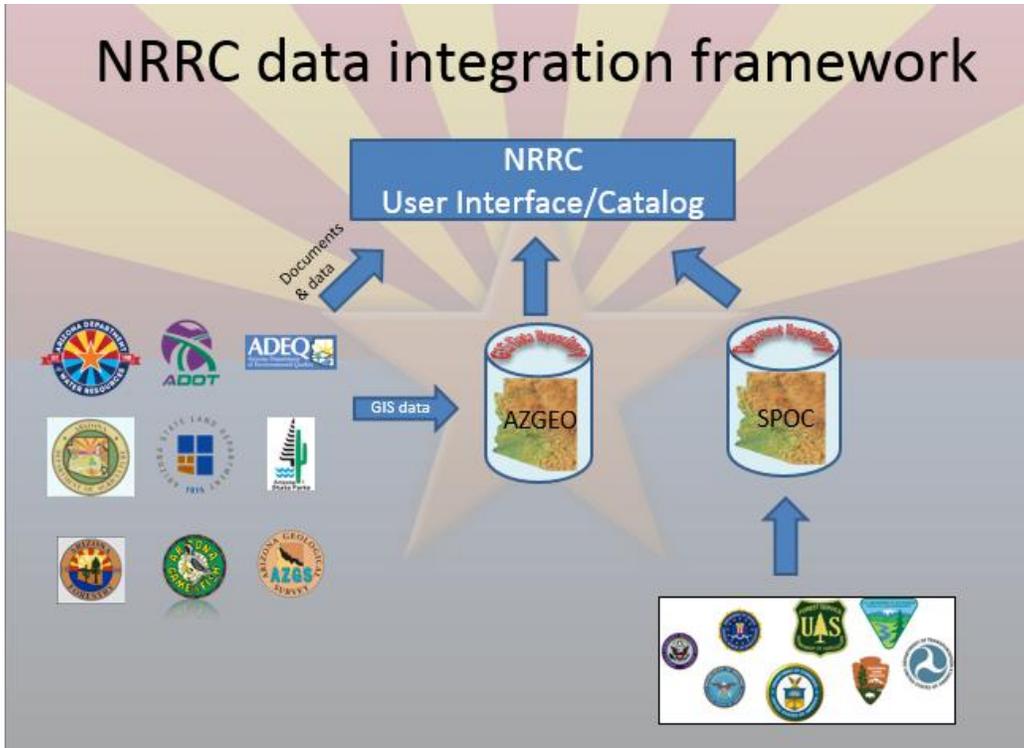


Figure 1: Conceptual Design of AZ Data System

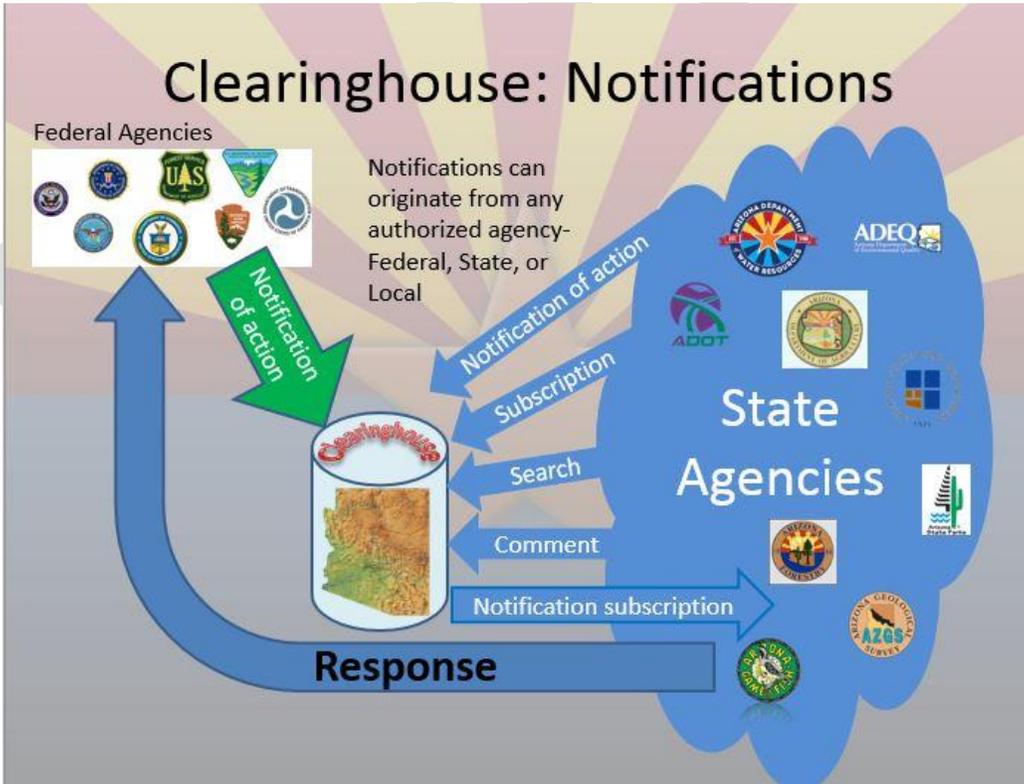


Figure 2: Federal Notifications Clearinghouse

The approach was selected due to the requirements collected during the NRRC committee and subcommittee work. Requirements requested by the NRRC included:

- Clearinghouse of Federal notices, announcements, requests for review, and documents that are digital, georeferenced, and keyworded by providers; and provides automatic notification based on user criteria
- A catalog of State GIS and other relevant agency data and documents that allows NRRC members to discover and access both Federal materials, as well as State documents and other data resources
- A user interface that improves NRRC member's ability to collaboratively discover, review, and analyze statewide GIS and other data resources alongside Federal documents and materials
- Ability to integrate State and Federal GIS and other online data seamlessly in real time
- Ability to maintain independent Agency databases and content management systems while providing support for Agencies not currently using an in-house database or content management system.
- All State agency data is digital, online, georeferenced, keyword tagged, and interoperable

To accomplish the ambitious goals associated with this plane, we propose two work phases: Implementation & Catalog Development and User Interface Development & Data Integration.

Implementation & Catalog Development

The first phase includes:

- Final requirements gathering, including an assessment of all in-house databases and content management systems, metadata collection strategies (if any), and estimate of work-required to complete the NRRC User Interface and Catalog.
- Determine, with the assistance of NRRC member agencies, the appropriate metadata for use in the NRRC Catalog. Provide assistance to NRRC Agencies with processes to transform existing databases and content management metadata for statewide data integration.
- Review existing metadata within AZGEO – the proposed repository for State GIS data – and provide a software converter for metadata (data about the data) of holdings for integration into the State-wide Catalog.
- Implement beta version of AZSPOC using existing Federal partners (e.g. Western Regional Partnership, Data.gov, etc.). This includes the research on the use of free text-analytic software to determine keywords and georeferencing information from these documents automatically as an alternative to requiring federal agencies to keyword and georeference when submitting notifications.
- Use existing off-the-shelf software to link three modules through an integrated metadata catalog, the GIS clearinghouse (AZGEO), an online catalog of agency document repositories, and Federal documents and materials (AZSPOC).

- Provide free and open-source software solutions to agencies that do not already have a document or data management software in place.
- Deploy an open-source node as an aggregator for AZGEO, AZSPOC, and NRRC agency catalogs

Estimated implementation time for the Implementation Phase by the Primary Coordinating Agency: A minimum of 600 hours, plus 40 hours per agency requiring a software metadata wrapper. Estimated implementation time for the other NRRC Agencies: TBD based on data holdings

User Interface Development & Data Integration

The second phase includes the incorporation of additional data holdings that are not currently digital and the refinement of the beta interfaces. The following steps are included:

- Provide for an Enhanced User Interface for collaboration and any additional functionalities identified in the scoping and implementation phase as well as beta use of the system.
- Focus on deploying additional web services for State GIS data not included in the Implementation Phase (this would include new data, newly digitized data, or data that is more unique to each agency, and thus lower importance to the overarching NRRC mission).
- Focus on digitizing legacy and existing hard-copy data, including metadata creation, archival scanning, and keywording/georeferencing.
- Build additional user applications based on the beta implementation phase.

Estimated implementation time for the Adoption Phase varies upon the success of the Implementation Phase.

Challenges for a Statewide Data Integration Platform

AZGEO is being developed under existing federal grants for other uses however, in order to meet the requirements of the NRRC, it can be repurposed to support NRRC in the near term. Long term availability of AZGEO and its costs are unclear, thus, posing a sustainability question of the centralized GIS repository.

The Arizona Geological Survey (AZGS) is developing open-source catalog search and data integration capabilities under separate federal contracts and agreements that may be leveraged at least in part to support NRRC in the near term. Ability to customize the services and long term support under these federal grants are unclear.

Receiving AZSPOC materials from federal agencies as "digital, georeferenced, and keyworded" may or may not be feasible, since they are not specifically mandated to do so under the 1982 Executive Order. In that case, we hope to utilize existing text-analytic software which may be capable of providing rudimentary keyword and georeferencing information for Federal Notifications.

While challenges are present, this roadmap will assist NRRC Agencies in providing a state-of-the-art integrated data system for capturing state GIS data, receiving and responding to Federal notifications, and provide a catalog of NRRC Agency documents and data while respecting data ownership and existing Agency systems.

DRAFT

Natural Resources Review Council (NRRC) Mission

Governor Jan Brewer's Executive Order 2013-02, "Establishing the Arizona Natural Resources Review Council" (NRRC) established the NRRC to "develop land and natural resource management strategies for Arizona and coordinate with state natural resource agencies and their existing management plans" (Appendix A). Participants in the NRRC include:

- Arizona Game and Fish Department
- Arizona State Land Department
- Arizona Department of Environmental Quality
- Arizona Department of Water Resources
- Arizona State Forester
- Arizona Geological Survey
- Arizona State Parks Department
- Arizona Department of Agriculture

In response to the Executive Order the NRRC established four subcommittees: Geographic Information System (GIS), Planning, Clearinghouse Development, and Mitigation and Conservation Banking. Each subcommittee was charged with the development of a scope of work to accomplish the goals of the NRRC. This document is in response to the GIS Subcommittee's Scope of Work as presented to the NRRC in April of 2013, and modified in June of 2013 (Appendix B).

Natural Resources Review Council (NRRC) GIS Subcommittee Mission

The Natural Resources Review Council (NRRC) was directed to develop a coordinated GIS capability as part of its mission to "develop land and natural resource management strategies for Arizona and coordinate with state natural resource agencies and their existing management plans." When considered in conjunction with other needs and requirements defined by the Council to carry out its duties, it has become evident that the GIS functions are an integral part of a more comprehensive decision support system, that includes access, discovery and analysis of NRRC and federal agency data and documents, a Clearinghouse for federal materials that will serve as the Arizona Single Point of Contact (AZSPOC), with notification and search capabilities, and a user interface.

The GIS Subcommittee's expanded charge is to develop plans on how to achieve this within the constraints and requirements identified by Council members and the subcommittees.

Charge to the Subcommittee

Executive Order

The Executive Order directed that "The Council [NRRC] shall develop a coordinated and centralized Geographic Information System database model that identifies current and future management priorities for designated land and natural resource areas." In response, the Council established a GIS Subcommittee at its first meeting to develop a plan for carrying out this task.

The NRRC Chair appointed Council member Lee Allison, AZGS, as Chair of the Subcommittee and each NRRC member agency was invited to appoint a member to the Subcommittee. The GIS Subcommittee prepared an initial Scope of Work (Appendix B) that was presented and approved at the next NRRC meeting.

Expanded Charge

The Subcommittee informed the Council at the June 3, 2013 meeting that to meet the needs and requirements implicit in the Executive Order, the GIS system should provide for full data integration across the NRRC agencies and be scalable and transformable for other state agencies and able incorporate data from relevant federal agencies.

Meanwhile, the NRRC Clearinghouse Subcommittee developed their needs and requirements and recognized the technical challenges in creating the clearinghouse and the potential for leveraging the expanded scope of work of the GIS Subcommittee. Subsequently, the chairs of the two subcommittees met with information technology and policy staff in the Governor's office and agreed that the data integration system for NRRC should provide the framework for the NRRC Clearinghouse.

Meeting the Charge

GIS Coordination

All of the NRRC agencies are using Esri-based GIS software, although not all use the same versions of that software. We can more easily talk to one another about how to perform data analysis if we're all using a single software "toolbox" (e.g. ArcGIS)

Clearinghouse

The Scope of Work for the Clearinghouse Subcommittee defined their charge:

"A deliverable that was discussed by the Natural Resources Review Council (NRRC) included the development and implementation of a document review Clearinghouse. In order to better organize within the state and achieve the Governor's expectations, there is a need to establish a mechanism that would ensure all appropriate agencies are getting information and that all agencies are able to respond to the proposed action. A clearinghouse will serve as the coordination point for the federal government to request state review and comment, which will meet their requirements for coordination with the state. An automated system would be desirable for the agencies."

The Clearinghouse Subcommittee developed a preliminary set of requirements for the proposed clearinghouse:

- Ability to accept and share documents and/or internet links to documents
- Unlimited Agency access to documents with searchable capabilities (e.g., keyword searches, geographic search using bounding boxes, etc.)

- Ability to send notifications to agencies of the availability of documents based on identified criteria
- Ability to use a workspace format for sharing agency reviews and comments among the agencies (note: new capabilities with open source Github offer a no-cost alternative to expensive commercial software with steep learning curves)

In discussion with the Council on June 25, it was learned that NRRC agencies are using a variety of document management systems, including Xerox Docushare, Microsoft Sharepoint, and CKAN, and that some agencies have no central system or common approach to document management.

Discussion at the Council meeting concluded that no agency wants to change their current system, or be required to convert their documents over to another format. For agencies without a current system, they expressed concerns over costs to acquire commercial software and to digitize paper documents for entry into a management system. The availability of free, open source software was offered as a one solution to the first issue.

The agencies recognized that the solution has to involve the ability to integrate documents and other digital materials from disparate and incompatible formats not only among state agencies but from a large number of federal agencies that have similar challenges. The group also expressed the desire to be able to search for documents by key words and geographically, using tools such as a bounding box or coordinates to identify materials covering specific areas.

A key requirement for the Clearinghouse is for automatic notifications to NRRC partners of the posting of anything in the Clearinghouse, using a variety of faceted specifications (e.g., one or more designations of location, agency, topic, etc).

The Clearinghouse Subcommittee discovered that a Presidential Executive Order from 1982 is still in effect, requiring agencies to submit their notices, announcements, reports, etc, to a central Single Point of Contact (SPOC) if the state so designates. Arizona maintained a SPOC until early in this century but it lapsed.

Historically, federal agencies would mail announcements of the availability or the materials themselves to the state SPOCs and the appointed state manager would distribute them to the appropriate state agencies.

With email and digital documents now ubiquitous, it is logical that the SPOC become an electronic clearinghouse. This requires two key elements – the state must set up and maintain an online repository for digital materials, and the federal agencies must agree to submit them to AZ SPOC in the required format. Optimally, the documents should be geo-referenced and include catalog information (“metadata”) including keywords, provided by the originating agency, otherwise Arizona will have the task of developing those attributes for a massive number of materials. Alternatively, we may be able to take arbitrary text documents and automatically extract keywords and geo-referencing, although that still puts the burden on NRRC.

Planning

The Planning Subcommittee defined their scope of work regarding their charge as

“The second deliverable that was discussed included the identification of priorities for legal, legislative and incentive-based programs and associated funding. Examples of challenges that were discussed included recent air quality regulations and travel management plans on federal lands. The Council will be putting forward priorities and recommendations for legislation and budget for the Governor’s review. In order to better position the state to be effective in the administrative and judicial areas, the state needs to create an inventory of plans/projects around the state. This process includes proper notification to the state through a clearinghouse. This approach may assist state agencies that may have a plan affected by a federal decision and protect their interests.”

The need for a clearinghouse is specifically identified to include federal documents and materials and a notification process for alerting NRRC agencies of plans and projects. This is consistent with the needs and requirements defined by the Clearinghouse Subcommittee.

Mitigation Banking

The Mitigation Banking Subcommittee plans include the need for access to and use of GIS and other data and documents from state and federal agencies. It appears that the system being proposed here would meet their needs.

NRRC Agency Data Holdings

This section provides an overview of the NRRC determined Needs & Requirements, as well as an assessment of existing resources based on an initial survey of member agencies GIS departments and review of existing documentation.

Needs and Requirements

Discussions within the Subcommittee and among NRRC members elicited a number of policies and constraints in providing services to the Council:

- Foster compatibility among state, federal, and local data for NRRC tasks
- Build on and leverage existing capabilities of and standards used by NRRC agencies
- Incorporate work done by AGIC and other State entities
- Use AGIC as a long-term resource
- Allow each agency to maintain their internal systems, i.e., don't require any agency to convert databases, documents, or to change systems, in order to be able to participate in the NRRC system

Cumulatively, these needs and requirements, coupled with the NRRC mission, effectively define the need for a decision support system for natural resource issues.

Existing Data, Services, and Resources

NRRC GIS subcommittee members have data and services to distribute and data and services they need. This includes GIS and other data, as well as documents and other digitized materials.

Data and Services in Each State Agency

Last year, AGIC conducted a survey of GIS data producing agencies in Arizona. The survey collected information from agencies about the availability of their GIS data. The results of the survey are presented in the Arizona Geospatial Clearinghouse Implementation Plan.

Of the fourteen data producing agencies identified in the survey, seven are NRRC members. The table below summarizes survey results for the seven NRRC agencies. The table is modified from the table presented in the Arizona Geospatial Clearinghouse Implementation Plan

Agency	Data Produced	Is data visually available on map on web?	Is data downloadable from web?	Data service available	Is data current on web?
ADOT	Road Centerline, Mileposts, traffic data, functional classification	yes	No	yes	No
ADEQ	Water, Air, Waste data	Yes	No	yes	Yes
ADWR	Wells, rivers, surface drainage	No	Yes	no	No

AGFD	Wildlife habitat, corridors, designated hunting areas etc.	no	no	no	no
AZGS	Geothermal data and other geologic data	yes	no	no	yes
ASLD/ALRIS/SCO	State land sales and status mapping	no	no	no	no
ASP-SHPO - AZSite	aggregated cultural site mapping	yes	no	no	yes

Data and Services Needs for Each State Agency

NRRC GIS subcommittee members need data that originates from other sources. This data will typically originate from Federal, State and/or local government sources.

AGIC identified 19 data layers its strategic plan that would support GIS data needs for State agencies. These data, their status and source are identified in the table below.

Data Layer	Status	Source
Transportation	Incomplete	Federal, state
Orthoimagery	Complete	Federal, local government
Cadastral (parcels)		County
Land Ownership (not parcels)	Complete	Federal, state, county
Geodetic control	Ongoing	Federal, state, utilities
Land use	Incomplete	Local government
Elevation	Complete to 10m	Federal, local government
District boundaries (tax-related)	Incomplete	Federal, state, county
Master address file	Incomplete	State, county
Administration units (non-tax)	Complete	County, city, tribe
Hydrography	Complete	Federal
Structures (critical infrastructure)	Incomplete	
Geographic names	Complete	Federal
Geology	Ongoing	Federal, state
Land cover	Outdated	Federal
Demographics	Complete	Federal
Non-critical infrastructure	Incomplete	
Environmental/sensitive areas	Complete	Federal, state, environmental
Soils	Complete	Federal

Software Applications in Use

NRRC agencies use a combination of proprietary and open source software for creating, managing and distributing geospatial data. Proprietary GIS and database software includes a variety of Esri GIS software (e.g., ArcGIS, ArcServer, ArcGIS Online), Microsoft Access, Microsoft Excel, Xerox DocuShare, Microsoft Sharepoint, and SQL. Open source software for content management and Web services includes Geoserver, PostgreSQL, Drupal, and Django.

Technical Expertise

Among the NRRC agencies there are the following technical expertise on staff: spatial and non-spatial data editors, field data collectors, map digitizers, cartographers, informatics specialists, software developers.

Existing GIS Costs to NRRC Agencies

Arizona state agencies likely spent over \$1 million in FY13 on Esri GIS software purchases, licenses, and training. A more accurate number is not available without extensive analysis of the state finance system, AFIS, or by surveying the agencies. We worked with ADOA to identify keywords in AFIS but found every agency labels each transaction using their own terminology, with is not only not consistent across agencies but within agencies for different products and services. The actual expenditures could be significantly higher.

This does not include the personnel costs for GIS staff in the agencies that use the software, nor the IT costs for servers and storage.

The purpose of including this estimate is to give an indication of the investment the State currently makes in GIS data alone. The costs for managing other digital data and documents will be harder to quantify.

GIS Coordinating Bodies

The NRRC GIS subcommittee identified three coordinating bodies that provide the framework and guidance for the development of this report. These coordinating bodies are the Arizona Geographic Information Council, the National States Geographic Information Council and Western Regional Partnership.

Arizona Geographic Information Council (AGIC)

The Arizona Geographic Information Council (AGIC) is a primary GIS resource for the NRRC GIS Subcommittee, for GIS data and services, and policy.

AGIC has been coordinating GIS efforts in Arizona for nearly two decades. AGIC was established by Executive Order 89-24. The mission of AGIC is “to coordinate the development and management of geographic information in Arizona. AGIC promotes the use of GIS and related technologies to address problems, develop plans, and manage the natural, economic and infrastructure resources of the state.”

AGIC has been incredibly successful in pursuing this mission. Several geospatial planning documents AGIC has produced are available for download on the AGIC website. Of these documents, NRRC GIS Subcommittee Report incorporates research and guidance from the following documents:

1. [2013 AGIC Geospatial Data Sharing Guidelines](#)
2. [2010 AGIC Strategic Plan](#)
3. [2010 Business Plan](#)

National States Geographic Information Council (NSGIC)

National States Geographic Information Council (NSGIC) is the national equivalent of AGIC. The mission of NSGIC is “to promote statewide geospatial coordination activities in all states and to be an effective advocate for states in national geospatial policy and initiatives, thereby enabling the National Spatial Data Infrastructure (NSDI).”

NSGIC’s goal is to provide “a unified voice on geographic information and technology issues, advocates State interests, and supports its membership in their statewide initiatives. The Council actively promotes prudent geospatial information integration and systems development.”

NSGIC has also been incredibly successful in pursuing its mission. Several geospatial planning documents NSGIC has produced are available for download on the NSGIC website; the Maturity Assessment is discussed below.

Existing Documents, Surveys, and Assessments

The NRRC GIS Subcommittee identified five strategic documents that provide the framework and guidance for the development of this report.

Arizona Geospatial Clearinghouse Implementation Plan Draft

The Arizona Geospatial Clearinghouse Implementation Plan document provides the structure for governance, functionality and technical specifications for the Arizona Geospatial Clearinghouse. The document is currently being drafted, but contains information relevant to the NRRC GIS Subcommittee report.

AGIC Geospatial Data Sharing Guidelines (2013)

The AGIC Geospatial Data Sharing Guidelines document serves as a “best practice guide for Arizona public agencies who engage in the sharing of geospatial data.” The document covers Arizona Revised Statutes as applicable to geospatial data sharing, the benefits of data sharing, data sharing roles, data sensitivity levels, the use of data disclaimers, metadata guidelines and guidelines for agencies to create their own data sharing policy. The document is currently in draft form.

AGIC GIS Strategic Plan (2009)

The AGIC GIS Strategic Plan summarizes the current GIS situation in Arizona; defines AGIC vision and goals; defines financial, organizational and technical requirements for meeting the goals; defines an implementation framework; describes the strategic planning methodology. Strategic Goal #1 is to “facilitate the productive application and sharing of geospatial data and GIS and location based services to address the needs of Arizonans by establishing a Clearinghouse with statewide accessibility.” Strategic Goal #2 is to “achieve greater fiscal responsibility and efficiency through the wise governance of GIS services and geospatial data.”

AGIC Business Plan for the Statewide Geospatial Clearinghouse (2010)

The AGIC Business Plan for the Statewide Geospatial Clearinghouse is focused on Strategic Goal #1 of the AGIC's GIS Strategic Plan, to "facilitate the productive application and sharing of geospatial data and GIS and location-based services to address the needs of Arizonans by establishing a Clearinghouse with statewide accessibility."

NSGIC Geospatial Maturity Assessment (ongoing)

The Geospatial Maturity Assessment (GMA) is a survey conducted by NSGIC. The GMA offers "a common, credible baseline assessment methodology to routinely and continuously monitor and validate statewide geospatial capabilities. Each state was asked to complete eighty three (83) detailed questions that characterize their geospatial programs." The GMA is updated annually.

Survey of States' GIS Clearinghouses

The Arizona Geological Survey conducted a brief survey of the coordinated and centralized GIS data discovery and access systems for states that have data available online. The survey was conducted from the user's perspective to evaluate the similarities and differences between state GIS data distribution systems and to identify states that have successfully implemented coordinated, centralized GIS data systems. Survey methods, results and conclusions are presented in Appendix 6.

The brief survey concludes that most states have similar GIS data distribution system components. The biggest difference lies in the implementation of the system. From the user perspective, this translates to the user interface and user experience.

As expected, the ease of data access and use varied from state to state. Of the forty states that have GIS data online, five states implemented systems that had great strengths that set them apart from other systems.

From the user perspective, some states have well-coordinated, centralized GIS programs while others do not. North Dakota, for example, has a well-coordinated, centralized GIS system. All GIS data for all state agencies is only accessible through one location, the NDGIS Hub Data Portal.

The State of Utah has an easy-to-understand, aesthetically pleasing landing page for their initial data web access point. The State of Utah also distributes data in a variety of formats, enabling users to pick the data distribution method that works best for them.

New Mexico, Missouri, and West Virginia all have a user interface that is simple and intuitive, with no more than 3 clicks to data download. This translates to ease and speed in terms of how easy it is to find and access data

Opportunities for Discovery, Access, and Integration of Federal Data, Materials, and Projects

One of the challenges for NRRC will be to get federal agencies to submit materials to the AZ SPOC and in formats that relieve the burden on the state to digitize them, and create metadata to allow key word and geographic search for relevant items. An existing state-federal partnership and new federal requirements offer intriguing possibilities for Arizona to serve as a testbed and exemplar for state-federal interaction on natural resource and land management issues.

Western Regional Partnership

(WRP - <https://wrpinfo.org/>)

Western Regional Partnership (WRP) works with Federal and State agencies (including Arizona) to develop collaborative, geospatial solutions for the purpose of protecting natural resources and promoting sustainability, homeland security and military readiness. WRP's mission is to provide "a proactive and collaborative framework for senior-policy level Federal, State and Tribal leadership to identify common goals and emerging issues in the states of Arizona, California, Nevada, New Mexico and Utah and to develop solutions that support WRP Partners and protect natural resources, while promoting sustainability, homeland security and military readiness."

WRP membership includes 20 federal agencies, which encompass all or almost all the agencies dealing with NRRC issues in Arizona. The resources, tools, and capabilities they are developing for coordinating land management and land use issues across the five participating western states, should be applicable and transformable to facilitate the NRRC mission.

Preliminary conversations with WRP staff indicate strong interest in this and we plan on proposing it through the WRP GIS Committee and Steering Committee.

Three WRP resources are of initial interest – the GIS repository, Web Mapping Application, and Regional Project Database.

WRP is compiling a repository of an estimated 10,000 GIS layers from participating agencies related to land management and land use across the five state region. These are being made available through the Web Mapping Application, which is built on Esri ArcGIS software.

The first Memorandum of Understanding signed by WRP is with the US Geoscience Information Network (USGIN), which is managed by the Arizona Geological Survey on behalf of a national coalition of state and federal agencies. USGIN is a Web-based, open-sourced, framework for integrating data in a distributed network (as opposed to a central database). WRP has included in its current annual workplan, resources to make the GIS repository in the national distributed data network being developed by AZGS with support from USGS, National Science Foundation, and US Dept. of Energy.

Federal Open Data Policy

Recently, three documents on open data policies at the federal level were released that could. These include an Executive Order, and memos from the Office of Science and Technology Policy (OSTP) and Office of Management and Budget (OMB).

The OSTP Memorandum for Heads of Executive Departments and Agencies, *Increasing Access to the Results of Federally Funded Scientific Research* in February 2013 ensures that federally funded research is “available to and useful for the public, industry, and scientific community” (Holdren, 2013). This requires agencies with research budgets over \$100 million to ensure that “...the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data” (Holdren, 2013). The memo further requires that agencies ensure that publications and metadata produced with federal funds are stored in an archival solution that:

1. Provides for long-term preservation and access to the content without charge
2. Uses standards, widely available and, to the extent possible, nonproprietary archival formats for text and associated content (e.g. images, video, supporting data)

The Executive Order, released May 9, 2013, *Making Open and Machine Readable the New Default for Government Information*, calls for government information to be managed as an asset throughout its life cycle to “promote interoperability and openness, and, whenever possible and legally permissible, to ensure that data are released to the public in ways that make the data easy to find, accessible, and usable” (White House, 2013).

The OMB memo was also released on May 9, 2013 and begins the implementation of the Executive Order by requiring agencies to collect or create information in a way that “supports downstream information processing and dissemination activities” (OMB, 2013). This is specifically requested to assist and enable the development of products and services that benefit the public.

The federal open data initiative promotes distributed data sharing and life cycle maintenance of federally funded datasets. Thus, the NRRC vision of shared and easy data access can assist federal agencies in completing these open-data requirements.

System Design and Architecture

Data Discovery and Access

Based on the needs assessments and requirements described above, the original concept of GIS coordination has evolved into a system for integration of data, both GIS and other, along with documents and other materials from both state and federal agencies. The system requires a clearinghouse that will serve as a Single Point of Contact for federal agencies to submit documents and materials needed by NRRC. The end users want to be able to search among GIS files, data, and documents from state agencies, using keywords and geographic (map-based) search. This requires a catalog of relevant state agency resources that aggregates materials from different software, in different formats, and using different nomenclatures and standards. A user interface (Web portal) nominally would be hosted by NRRC, but it would be preferable for each NRRC agency to be able to provide system access from their own agency portals, in order to facilitate data analysis using applications and software unique or customized for that agencies issues and needs. A general conceptual design incorporating these elements is shown in Figure 1.

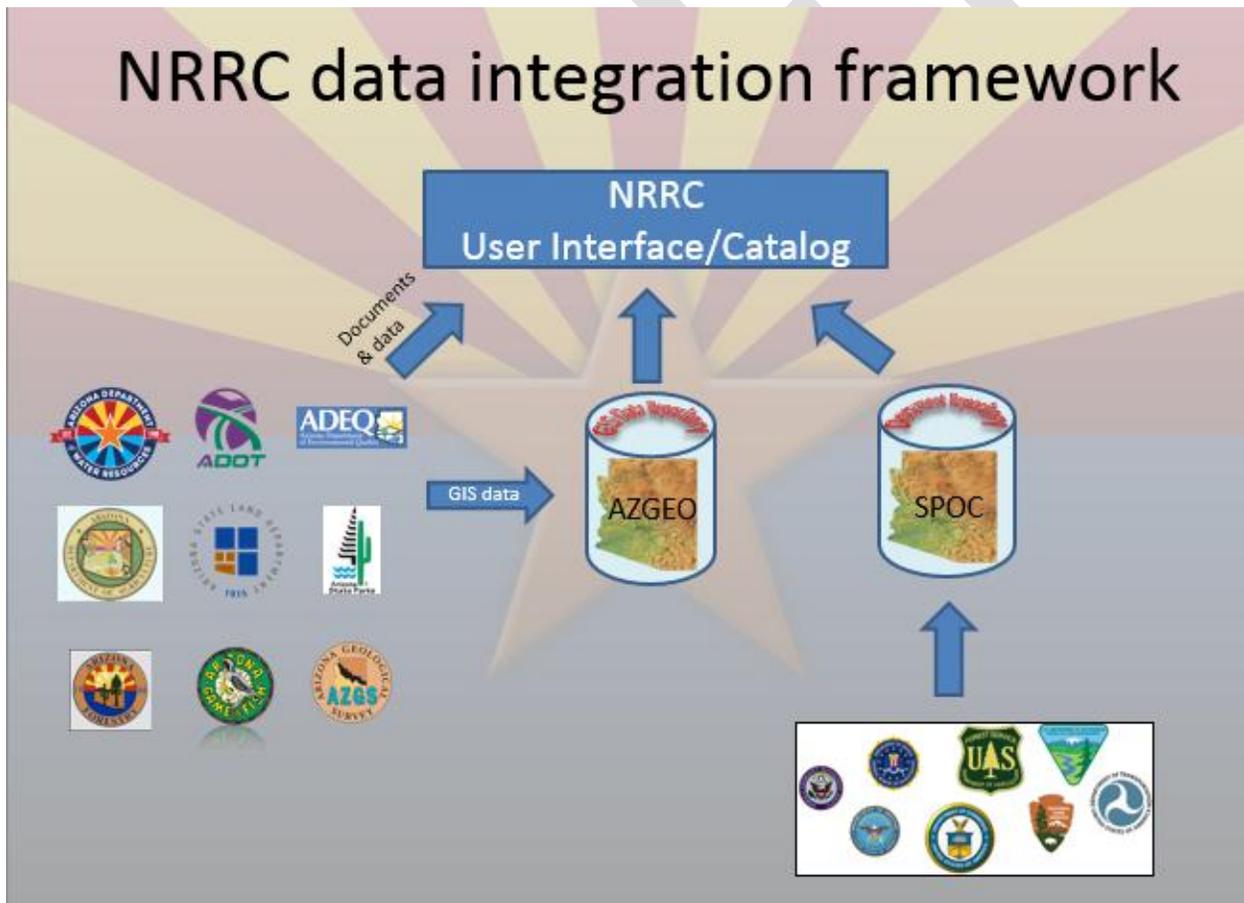


Figure 1. Conceptual design for a data integration framework for NRCC.

A central component of the proposed system is an online catalog to discover all available resources. Characteristics of the catalog include:

- Listing of all state information resources
- Access controls: public vs. state internal
- Search interfaces for people and machines
- Metadata in tiered formats
- Discovery and access to federal data

Tiered metadata describes three levels of interoperability:

Tier 1 = Discover. Whatever you have, make it available online and then write metadata about it. Include the metadata in a catalog, and you're done. One star. Every single dataset that wants to be a part of the system has to accomplish at least this much.

Tier 2 = Access. Make your content accessible using a standard protocol. Many organizations (CKAN, for example) do this by providing a custom data-access API. We would encourage the use of an existing API instead. Tier 2 data are data that are accessible via system-approved data services (e.g. OGC services, etc). You made a WFS? Great. Two stars. [Esri-based GIS products can be made OGC compliant with a click of a button]

Tier 3 = Understand. Make it possible for someone else to understand your data. Either utilize a system-approved content model or else describe your dataset (what do the columns mean??) in your metadata. Way to go, three stars!

The user interface (portal) has to provide data services to not only discover the data (in the catalog) but to access and use the data

- Direct access to maps (GIS services)
- Direct access to structured data records

State GIS Data Repository (AZGEO)

Over the past several years, AGIC has developed a comprehensive plan for a centralized GIS repository and clearinghouse, called AZGEO (AGIC, 2013), with significant input from NRRC member agencies. AZGEO hosts both data and services. Agencies can deposit data in AZGEO that is widely used, or does not change often, while maintaining their own data sets that may be unique to them or change so frequently that it would be burdensome to have to constantly deposit new versions. Also, agencies that do not have the technical support or desire to serve their own data online, may choose to use AZGEO as a repository. Data providers that host their own data can then develop Web services that are available through AZGEO. One current limitation of AZGEO is that it does not allow anonymous access to data layers. That is to say, you could upload your data to AZGEO and get a service, but couldn't consume that service anywhere else (other than AZGEO) without complex authentication requirements.

Eventually, it might be prudent to merge AZGEO with the primary data repository and clearinghouse for the NRRC system by exposing the AZGEO catalog using an agreed upon scheme. AZGEO is currently envisioned as a GIS clearinghouse, but the basic structure and

protocols are such that it could be reconfigured as part of the system to also host and serve non-GIS data and materials.

Arizona Data and Document Repository

Non-GIS data and documents are expected to be maintained by the state agencies that created them or have responsibilities for them. Each agency's file repository would be linked into the system with these parameters:

- Web accessible
- May be managed (DocuShare, Sharepoint)
- Metadata encouraged for files in catalog
- Access control and file management by owning agency

Agencies that cannot or do not wish to host and serve their data and documents would need a repository to provide that service. It's possible that AZGEO could be modified to take on those duties. Another option is that AZGS maintains a repository capable of doing this (repository.usgin.org), albeit with funding from federal grants for other purposes.

Non-GIS data and documents will come in many incompatible formats so interaction among them will be harder to achieve than with state GIS files which can be generally easily moved around.. One of our suggestions is to investigate the use of free text-analytic software to determine keywords and georeferencing information from these documents automatically.

There are also free, open source software solutions that might work for those agencies that don't already have document or data management software in place.

State-Federal Information Clearinghouse

The Clearinghouse is an integral component of the NRRC system because it needs to provide integration of federal materials with state data and documents, for discovery, access, and analysis. The Clearinghouse can be designated as the Arizona Single Point of Contact (AZ SPOC), thus ensuring that it will at least receive the desired materials from federal agencies. What is uncertain is whether we can require that those materials are in digital form or Web services, and whether the federal originators will provide appropriate metadata such as key words and geographic references.

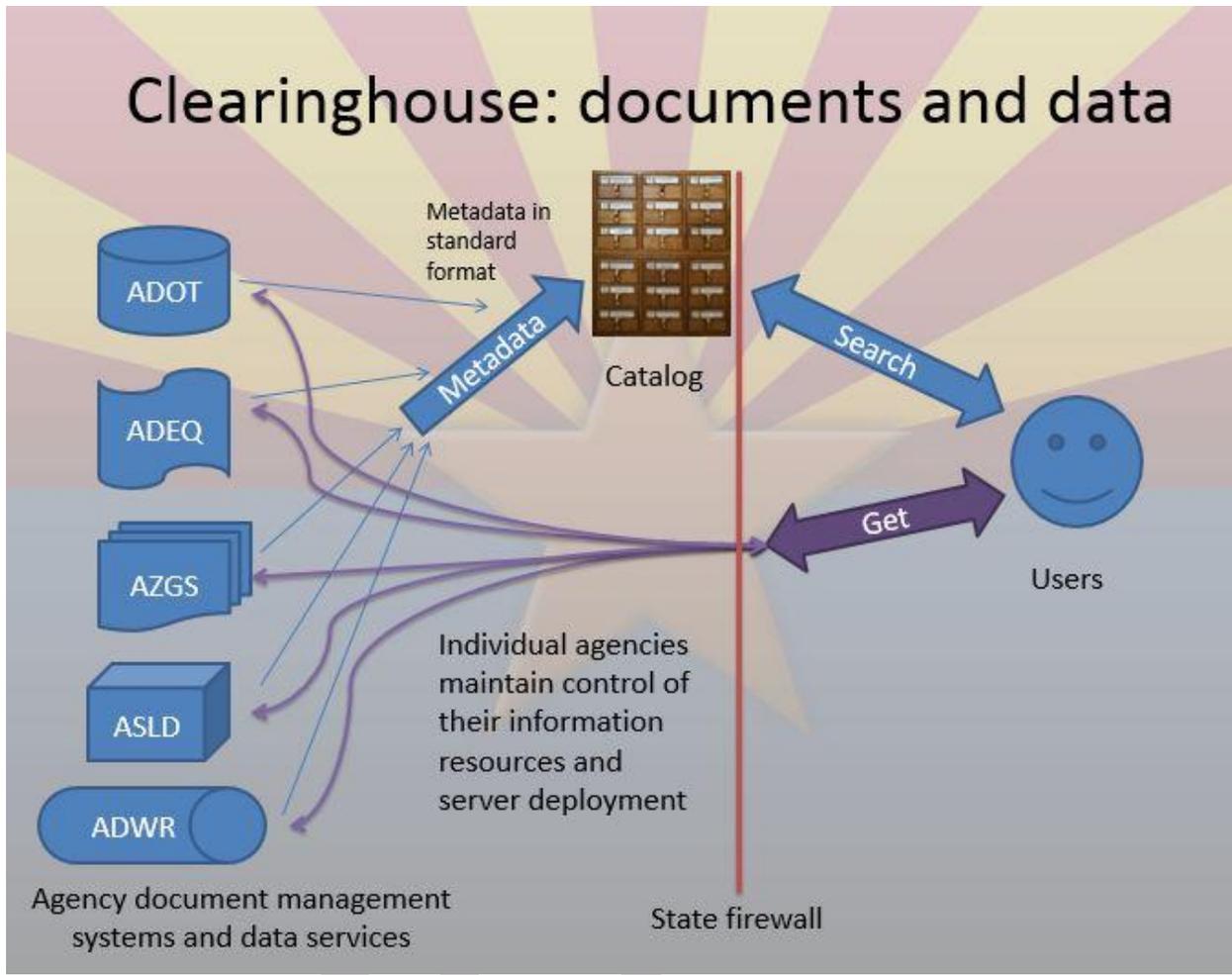


Figure 2. Conceptual design of a clearinghouse for online digital data and documents.

The Clearinghouse must provide a notification system to state agencies including a subscription service to alert appropriate state authorities of new notifications. The subscriptions should be able to be targeted by attributes such as topic, agency, location, etc. This is essentially an application built on top of the NRRC system. Other applications requested are:

- Work space for agencies to review and comment on federal notices
- Collaboration space for agencies to work on a federal notice
- Ability for agencies to maintain portals to the system and use of their apps for specialized analyses

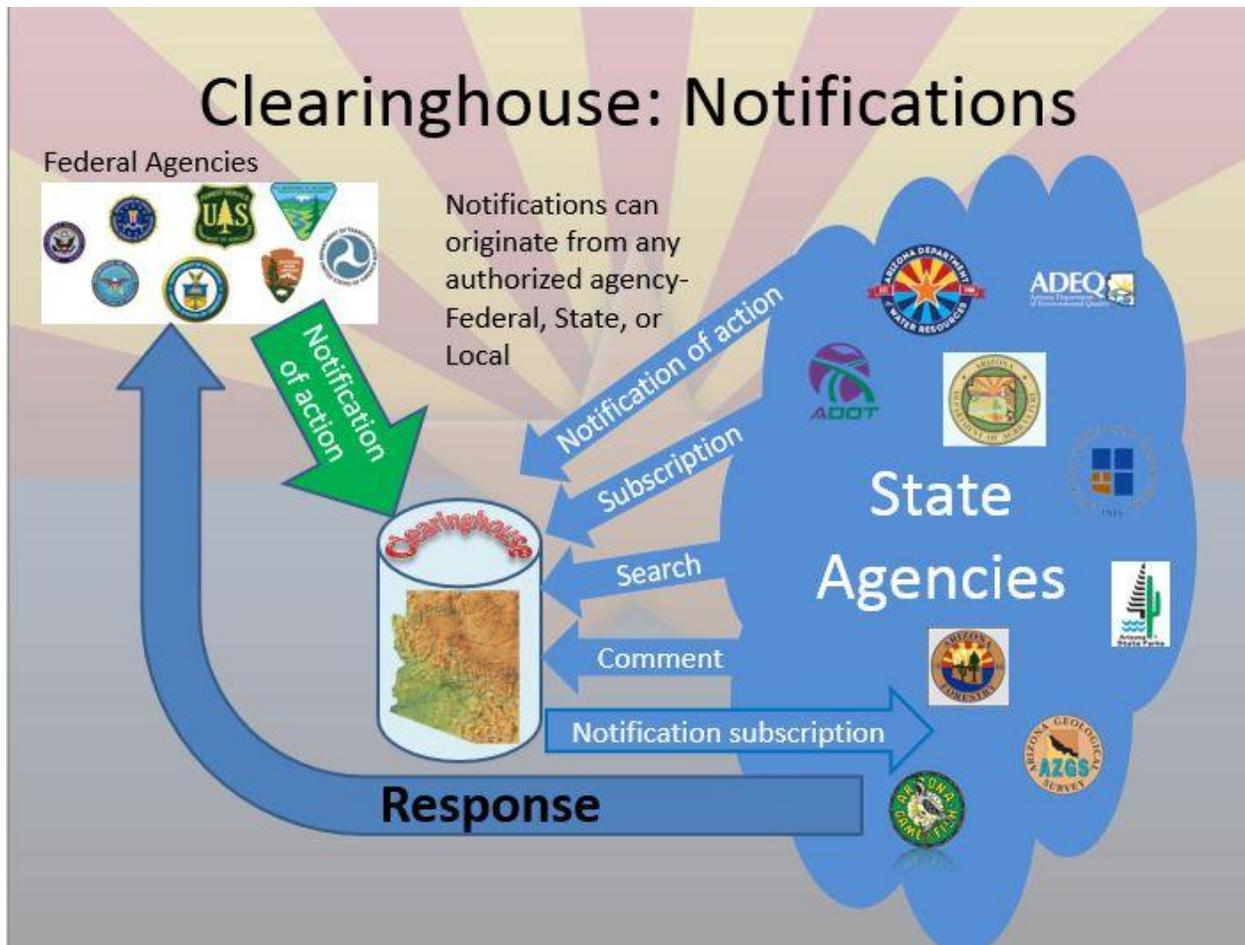


Figure 3. Diagram of notification process for new materials in the AZ SPOC Clearinghouse.

We envision the notification process to be initiated through a Web-based form, accessible to authorized state and federal agency contact points. A management agency would post notifications in a standardized format to a publishing end point. The AZ SPOC Clearinghouse harvests those notifications and also “scrapes” materials from existing sites (e.g. WRP Regional Project Database).

Arizona Single Point of Contact Notifications Clearinghouse

Contact My Account Logout

Manage Collections

- Add to a new Collection

Manage File Attachments

- Please save your record before attaching files

Resource Location

Arizona

Site Information

- About this Repository
- Terms of Use

Create new notification Save Validate

Published / Visible to the public

Title

Description

Publication Date mm/dd/yyyy --:-- --

Resource ID

Keywords

Geographic Extent

North Bound

South Bound

East Bound

West Bound

Figure 4. Mock-up of a possible Web form for submitting new materials to the AZ SPOC, which triggers automatic notification to subscribers based on pre-selected criteria.

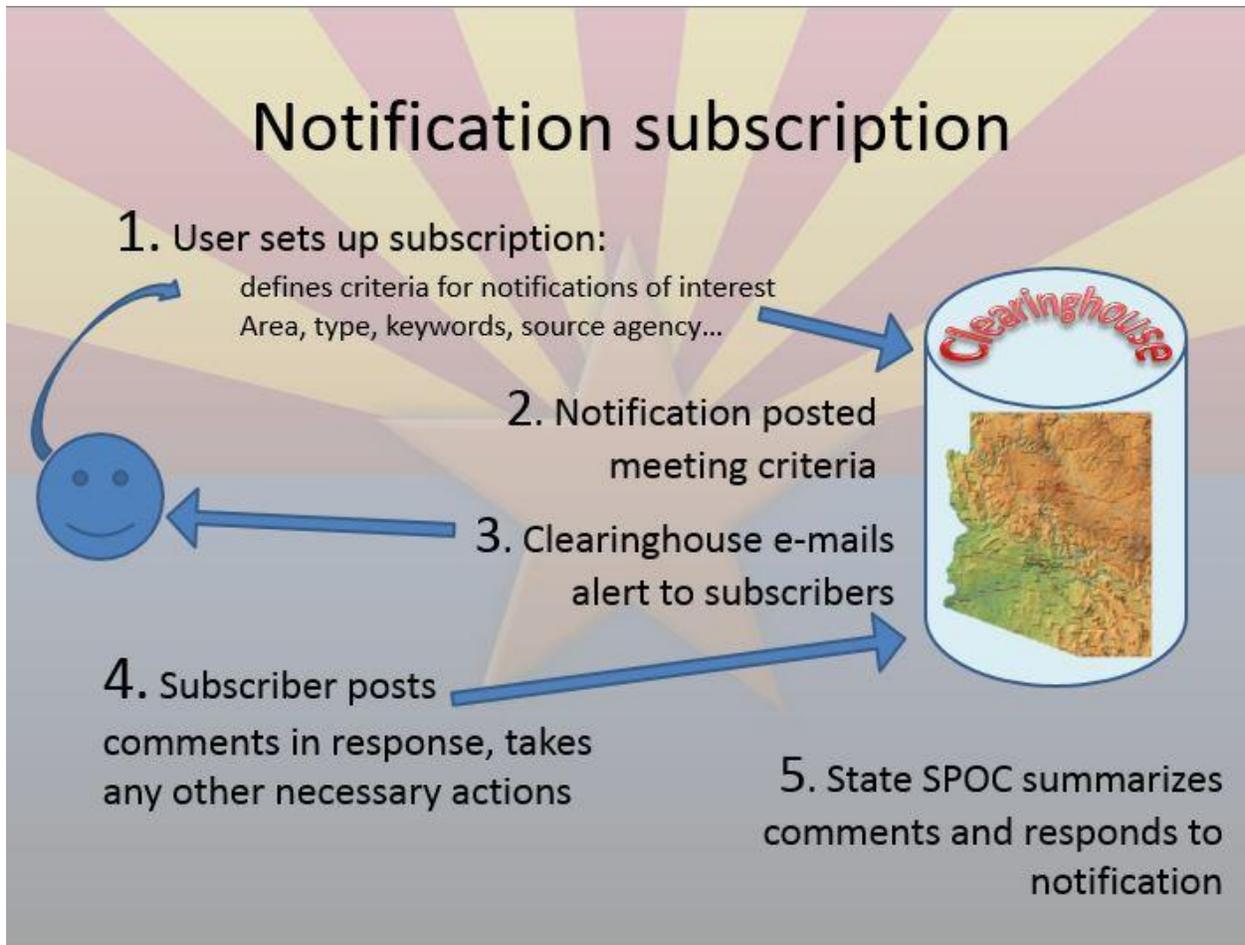


Figure 5. Diagram of clearinghouse notification subscription service with shared work space for comments.

The Utah Resource Development Coordinating Council (RDCC) serves a function very similar to that of the NRRC. RDCC maintains an online Project Management System of federal proposals. RDCC staff add documents to the system, digitizing them when needed, and populate the database with Title, Agency, Status, Comment Deadline information, and a link to view the document. Search is limited to mostly those attributes. There is no geographic search, although the metadata for documents includes some geographic information such as section, township, and range, entered into the database by RDCC staff. While such a database can be easily set up, it requires significant investment by the state to process and interpret federal documents and input the information.

RDCC Project Management System
Public Lands Policy Coordination Office • Resource Development Coordinating Committee

UTAH.GOV SERVICES | AGENCIES | Search all of Utah.gov

Projects
Government Agency Login
Utah.gov
Main PLPCO Site

Project Search
Search by project number, title, or sponsor:
Click [Advanced Search](#) for more options. [click here for search examples](#)
 Include Stream Alt?

Current Projects

Id	Title	View	Status	Agency	Comment Deadline
39203	New Route Designations near Red Canyon and Peters Point Ridge	[view]	Comment Period Open	DOI/Bureau of Land Management	07/29/2013
39202	Foundation Energy MGMT LLC proposes to drill three new gas wells in the ...	[view]	Comment Period Open	DOI/Bureau of Land Management	07/29/2013
39200	Draft EIS--TransWest Express Transmission Project	[view]	Comment Period Open	DOI/Bureau of Land Management	08/23/2013
39182	Gooseberry Mesa Trailhead Expansion	[view]	Comment Period Open	DOI/Bureau of Land Management	07/22/2013
39181	Lower Rocky Pass Riparian Improvement Project for Boreal Toad	[view]	Comment Period Open	DOI/Bureau of Land Management	07/25/2013
39180	Enefit American Oil Utility Corridors EIS	[view]	Comment Period Open	DOI/Bureau of Land Management	07/11/2013
39160	SULA 842 Amend #1 - Abajo Substation Expansion	[view]	Comment Period Open	School and Institutional Trust Lands Administration	07/30/2013
39140	Wheeled Pier at Bear Lake - Easement 400 00266	[view]	Comment Period Open	DNR/Division of Forestry, Fire & State Lands	07/23/2013
39124	Amendment of Easement for Natural Gas Pipeline	[view]	Comment Period Open	DNR/Division of Forestry, Fire & State Lands	07/24/2013
39123	Mexican Mountain Airstrip Maintenance	[view]	Comment Period Open	DOI/Bureau of Land Management	07/17/2013

693 Items Page: 1 of 70 Show: 10 [Download Results](#)

Figure 6. Screenshot of Utah’s Resource Development Coordinating Council Project Management System to track federal natural resource actions.

Integration of System Elements

An NRRC decision support system comprised of the above components can be assembled from existing and emerging elements available from NRRC members. A recent analysis of sustainability for the AZGS-managed US Geoscience Information Network (USGIN) for support of the National Geothermal Data System (Allison, 2013) describes in more detail the specific technical and staff requirements needed for system integration, operations, and maintenance. We have extracted the relevant section of that report and included it as Appendix 5.

Education and Training

One of the key components for maintaining the system is ensuring that end-users have access to appropriate materials and training courses for the system. AGIC holds an annual conference with a variety of GIS training and education elements. AZGS is developing a wide range of education and training materials, tutorials, and guides on open source Web-based data integration methods that will be freely available to NRRC and federal agencies.

Risk assessment

Key findings for Operational Sustainability:

- A distributed system's weakness and strength is that it relies on multiple stakeholders for maintenance and management.
- Virtual servers offer rapid extensibility, less onsite maintenance, and overall lower costs than onsite hardware. In addition, they offer flexibility for rapid replication and disaster recovery. Thus, it is recommended that virtual machines be used for web service deployment and replication when at all possible.
- Long-term developer engagement should be encouraged by maintaining the developers' forums, code-sharing repositories, and system documentation.
- Service Level Agreements (or Recommendations) should be prepared for additional (or potential) nodes.
- Licensing costs must be considered in the long-term maintenance of the system. Extending the use to other user agencies may elicit additional resources.

Gap analysis of current capabilities

The proposed system is focused on working with Tier 1 unstructured data (such as documents) and Tier 2 structured data (such as spreadsheets and databases). The value of Tier 3 data is in interoperability among the data from different providers, but it requires a more sophisticated process including customized interchange formats. Based on preliminary assessments, we do not see much overlap among state agency data sets that would require dealing with Tier 3 structured and standardized data sets.

However, there is greater likelihood of interoperability issues between state and federal counterparts. For instance, Arizona Game and Fish may have data similar to that of the US Fish and Wildlife Service but in different databases, formats, and terminology that are not compatible.

Full interoperability of data across all providers will require significant investment, so we have not included that in the current plan.

Discussion

The solution for the needs and requirements of NRRC collectively define an implicit decision support system to integrate state and federal data (both GIS and other types) and documents in a range of formats.

GIS data from both state and federal agencies appear to be essentially all capable of being exposed as OGC-compliant Web services (ie, WFS, WMS, etc), The AGIC AZGEO repository and clearinghouse capable of serving similar functions for NRRC and AGIC has ancillary policies for data sharing.

Digital documents held by agencies are in multiple formats and stored in disparate document management systems, including commercial and open source software. It

Requirements to implement Phase 1 are almost entirely staff time. At the system level, we estimate:

- AZGEO - ~40 hours to produce metadata in an interchange format, and test harvesting. If the AZGEO portal Web application meets user needs then no further effort is required except maintenance.
- AZ SPOC Clearinghouse - This will require 80-160 hours for a programmer/developer to organize an online repository application similar to that already in use by AZGS or for a programmer to build a customized site with the required function. This includes incorporation of notification and subscription functions. Also budget 40-60 hours for a more complete requirements analysis and planning.
- AZGS has an open source catalog service that is transportable for NRRC. A software converter must be written for each agency's document management system so that catalog listings ("metadata") from all sources can be obtained in a common format for data discovery, access, and integration. Developers familiar with the document management software and interchange format should be able to write such a converter with a 1 to 2 day effort for each agency's DMS.

At the agency level, we will encourage but not require, each data set and document to have key metadata including key words and geographic reference, to facilitate their discovery through the catalog. Agencies would have to determine if they have adequate resources to add or upgrade metadata and over what time frame it could be done.

Agencies that don't have document or data management systems are faced with setting them up and populating them. There are a number of open source software solutions available to help mitigate the costs of doing that, but the major costs will be to digitize the materials and document them adequately (i.e., metadata).

Federal agencies are required by Executive Order to abide by a state's implementation of a SPOC but they are not required to provide materials digitally, in a particular format, or with desired metadata. It will be the responsibility of NRRC to encourage and incentivize federal agencies to submit georeferenced digital documents online with key words. One rationale for this is that those agencies will no longer have to invest the resources to constantly update their contact lists and mail hard copies to a variety of venues. We hope we can convince them to adopt use of an AZ SPOC online submission process and form. One way to facilitate this is to make it a test bed for the new federal Open Data Policy, with agreements at high levels of the state and federal governments.

DRAFT

Recommendations

We recommend that NRRC undertake a phased approach to developing, deploying, and implementing the information resources and tools required for the Council to effectively carry out its mission.

We recommend that a holistic approach to data integration will be most effective in meeting the Council needs, providing enhanced services and functions, minimize the costs of operations, and providing additional benefits to the State beyond those to the Council.

Phase 1 of the plan is to:

- Use off the shelf capabilities to link three modules through an integrated catalog listing materials from: a state GIS clearinghouse (adopting AGIC's AZGEO for this), an online catalog of agency document repositories (using the AZGS document/data repository as a template), and federal documents and materials (harvested from AZ SPOC and WRP Regional Project Database).
- Implement AZ SPOC by engaging Federal partners (e.g., via WRP, Open Data Project, etc) to supply AZ SPOC with digital notifications and documents. Investigate the use of free text-analytic software to determine keywords and georeferencing information from these documents automatically as an alternative to federal keywording and georeferencing.
- Provide free, open source software solutions to agencies that don't already have document or data management software in place.

Phase 2 of the plan is to:

- Deploy enhanced User Interface, with functionality for geographic (e.g., bounding box) and faceted search of the AZ SPOC and agency resources available to NRRC. An NRRC Web portal can be modified from existing portals used for other purposes to provide more robust visualization and search capabilities, functionalities including notifications to relevant parties and collaborative editing of response documents, etc, as well as to accommodate deployment of a range of user applications ("apps").
- Promote the integration of data and documents from different sources and different software through standard OGC data services.
- Accommodate agency-specific analyses of AZ SPOC materials. We need to determine Council and Agency needs more specifically. Many of the analytical tools may be unique or customized to each Agency and should or already reside there. The system will be designed to accommodate and encourage software applications developed by anyone.
- Training and help to state agencies on digital standards, exposing data in compatible formats, and creating Web services to share data and documents.

Costs and Timeline

Implementation of Phase 1 at the system level will require roughly 600 hours of staff time (or 25% of one FTE). Depending on the number of staff involved, the system could be functional in 2 - 3 months.

Implementation of Phase 2 is more difficult to estimate because of uncertainties.

Enhanced User Interface – Time to do this depends on the number and nature of additional functionalities desired.

Data Integration – Focus on setting up OGC services and documenting the structure of each dataset (i.e. “tell me what the columns in your table mean”). There will not be much of a need for developing content models, since data provided by the NRCC agencies are largely unique.

Data Digitization – Work load varies by agency and we estimate it would take 1 to 3 years depending on volume, agency capabilities, and resource availability for those agencies that have not yet digitized their data and documents.

User applications – Development time varies with each one depending on complexity and availability existing components to build on or incorporate.

We assume NRRC principals and staff will contact their cognizant federal counterparts to get them to submit their materials to the AZ SPOC.

References

- Allison, M. Lee, Anderson, Arlene, Chickering, Cathy, Patten, Kim, Richard, Stephen M., Tanikella, Rajanikanth, "Sustaining the National Geothermal Data System: Considerations for a System Wide Approach and Node Maintenance," Geothermal Research Council, 2013 Annual Meeting and Conference, in press.
- Allison, M. Lee, editor, "Sustainability Plan for National Geothermal Data System (NGDS) Operations," submitted to Boise State University, May, 2013, 98p.
- Arizona State Cartographer's Office, ed., "AZGEO Clearinghouse Implementation Plan," July 2013.
- Executive Office of the President, Office of Management and Budget, Memorandum for the Heads of Executive Departments and Agencies, M-13-13, "Open Data Policy – Managing Information as an Asset," May 9, 2013.
- The White House, Office of the Press Secretary, "Executive Order – Making Open and Machine Readable the New Default for Government Information," May 9, 2013, <http://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government>

Appendix List for NRRC Data Integration Plan

1: Executive Order 2013-02	35
2: NRRC GIS Subcommittee Scope of Work	37
3: Federal Agencies in the Western Regional Partnership	39
4: AZGEO Frequently Asked Questions	40
5: AGIC Guidelines for Geospatial Data Sharing	42
6: NRRC GIS State Survey	50
7: U.S. Geoscience Information Network (USGIN) System Elements	53
8: USGIN Community Standards, Profiles, Protocols, and Software	57
9: Acronyms	59

DRAFT

1: Executive Order 2013-02

Executive Order 2013-02

Establishing the Arizona Natural Resources Review Council

WHEREAS, the ability for state natural resource agencies to engage and impact federal land and resource management plans is paramount; and

WHEREAS, cross agency coordination is critical to a sound and timely response to actions on federal lands that will impact state resources and economies; and

WHEREAS, federal land and resource planning processes often do not consider the short and long-term needs of the State of Arizona and local governments; and

WHEREAS, political change at the federal level can impede the ability of state agencies to effectively engage the federal government and result in federal intrusion on state authorities; and

WHEREAS, federal land and resource management decisions are being driven by litigation processes that have resulted in diminished multiple-use and sustained-yield principles; and

WHEREAS, we must protect State interests related to wildlife, land, water and natural resources by actively engaging and countering federal encroachment on State authorities tasked with managing Arizona's natural resources.

NOW, THEREFORE, I, Janice K. Brewer, Governor of the State of Arizona, by virtue of the power vested in me by the Constitution and the laws of this State do hereby order and direct as follows:

1. The Arizona Natural Resources Review Council (hereafter Council) is established. The Council shall develop land and natural resource management strategies for Arizona and coordinate with state natural resource agencies and their existing management plans.
2. The Council shall consist of the following natural resource agency directors appointed by the Governor:
 - Arizona Game and Fish Department
 - Arizona State Land Department
 - Arizona Department of Environmental Quality
 - Arizona Department of Water Resources
 - Arizona State Forester
 - Arizona Geological Survey
 - Arizona State Parks Department
 - Arizona Department of Agriculture
3. The Governor shall designate the Chairperson.
4. The Council shall prepare a comprehensive report and plan to the Governor for long-term land and natural resource management. The report shall include and address multiple use and sustained yield approaches, public access issues and sustainable economic development.

5. The Council shall develop a coordinated and centralized Geographic Information System database model that identifies current and future management priorities for designated land and natural resource areas.
6. The Council shall identify and prioritize legal, legislative and incentive-based needs that protect and maintain state interests related to wildlife, land, water and other natural resources.
7. The Council shall provide recommendations to the Governor on a statewide approach to mitigation and conservation banking that includes State government, local governments and the private sector in order to meet long-term natural resource conservation objectives.
8. The Chairperson may form an executive committee or other sub-committees as necessary.
9. The Council shall meet as needed to conduct its affairs.
10. This Executive Order shall take effect immediately upon signature.

IN WITNESS WHEREOF, I have hereunto set my hand and caused to be affixed the Great Seal of the State of Arizona.



Janice K. Brewer
GOVERNOR

DONE at the Capitol in Phoenix, on this 14th day of January in the Year Two Thousand Thirteen and of the Independence of the United States of America the Two Hundred and Thirty-Seventh.

ATTEST:

Ken Blumenthal

SECRETARY OF STATE

2: NRRC GIS Subcommittee Scope of Work

Charge to Subcommittee:

“The Council [NRRC] shall develop a coordinated and centralized Geographic Information System database model that identifies current and future management priorities for designated land and natural resource areas.” - Executive Order 2013-02

Expanded Charge:

The system should provide for full data integration across the NRRC agencies and be scalable and transformable for other state agencies and able incorporate data from relevant federal agencies. June 3, 2012

The system should provide the framework for the NRRC Clearinghouse. June 25, 2013

Proposed scope of work

1. Survey NRRC agencies to identify:
 - Data and services each State agency has
 - Data and services each State agency needs
 - Needs and requirements
2. Survey/Research other states state-wide GIS platforms and policies
3. Identify and locate existing surveys and assessments:
 - NSGIC (National States Geographic Information Council)
Geospatial Maturity Assessment (“a common, credible baseline assessment methodology to routinely and continuously monitor and validate statewide geospatial capabilities. Each state was asked to complete eighty three (83) detailed questions that characterize their geospatial programs.”)
 - AGIC (Arizona Geographic Information Council) data sharing and clearinghouse reports
 - State agencies software applications and technical expertise
4. Query other NRRC Subcommittees on their GIS and data integration needs
5. Identify other GIS and data resources we want to leverage or tap (e.g. Western Regional Partnership, White House Open Data Access Project, Data.gov)
6. Identify best (and failed) practices
7. Compile current total spending on GIS software licenses for Arizona state agencies
8. Conduct gap analysis
9. System Design and Architecture
10. State-Federal Information Clearinghouse
11. Scope of Work for Implementing the NRRC System
12. Prepare budgets for implementation options
13. Compile results into report with recommendations to NRRC

Requirements

Build on and leverage existing capabilities and standards in NRRC agencies
Incorporate work done by AGIC and other State entities

Use AGIC as a long-term resource
Allow each agency to maintain their internal systems
(ie, don't require any agency to convert databases, documents, or to change systems)
Foster compatibility among state, federal, and local data for NRRC tasks

Timeline

Preliminary report to be presented to NRRC at the July meeting
Recommendations made to NRRC at August meeting with intent that agencies can factor them into FY14 budget proposals as warranted

Resources

Rely on volunteer efforts by subcommittee members as time permits

Subcommittee Members

Lee Allison, Arizona Geological Survey, Chair
Janel Day, Arizona Geological Survey
Gene Trobia, State Cartographer, State Land
Ryan Johnson, State Land
Glen Buettner, State Forestry
Joyce Francis, Arizona Game & Fish
Jeffrey Wilkerson, Transportation
Victor Gass, Environmental Quality

Submitted 4-2-13
Revised 6-30-13

3: Federal Agencies in the Western Regional Partnership

Air Force

Army

Army Corps of Engineers

Bureau of Land Management

Bureau of Indian Affairs

Bureau of Reclamation

Customs and Border Protection

Federal Aviation Administration

Federal Emergency Management Agency

Federal Highway Administration

Marine Corps

National Park Service

Natural Resources Conservation Service

Navy

Office of the Secretary of Defense

U.S. Environmental Protection Agency

U.S. Fish & Wildlife Service

U.S. Forest Service

U.S. Coast Guard

U.S. Geological Survey

DRAFT

4: AZGEO Frequently Asked Questions



FREQUENTLY ASKED QUESTIONS

What is AZGEO?

AZGEO (Arizona Geospatial Clearinghouse) provides a data management system that uses a web interface to allow data providers to share either file based data or to publish live data services for consumption by map viewers, while simultaneously determining who may access the data, through security protocols that have been built into the system.

How do I get access to the AZGEO for 9-1-1 Testing?

Please contact Sandra Gilstad: Sandra.Gilstad@azdoa.gov

How do I contribute data?

You will register each dataset through the AZGEO web interface after you have become a registered user. To register a dataset or service, you will need the following information:

- Agency Information
- Access and Use Constraints
- Data Abstract
- Data Purpose
- Keywords (Data Tags)
- Publication Date
- Spatial Extent
- Projection, Datum, Units
- Time period of data
- Project status
- Update frequency
- Next expected update

What data formats can I share with AZGEO?

You may share shapefiles and data services.

How do I share a shapefile?

Please provide a layer file so that your data will display correctly. AZGEO accepts .zip files for upload through the Contribute Data page. You can contribute one shape file per .zip.

How do I share a data service?

The process for sharing a service is the same as for sharing a shapefile, except instead of uploading a file under the "File" tab, you will enter the Service URL under the 'Access' tab.

Who administers AZGEO?

AZGEO is jointly administered by the Arizona State Land Department, GIS Services at Arizona State University, and the Arizona Geographic Information Council (AGIC).

How is data access controlled?

When a dataset is contributed to AZGEO, the data provider will click a box to determine access. The following are examples of choices the data provider will have:

- “Government” Group (all email@something.gov users)
- “Education” Group (all email@something.edu users)
- “E 9-1-1” Group (all E 9-1-1 users)
- “General Public” Group (all AZGEO Clearinghouse users)
- Custom (user selects any combination of existing groups)

Is there a disclaimer?

Everyone who registers to use AZGEO must check a box that indicates they agree to the following:

The Arizona State Land Department (AZGEO Administrator) and Arizona Geographic Information Council (AGIC) AZGEO and Service Contributors provide geospatial data and GIS services to AZGEO in good faith and endeavor to create and maintain accurate data. However, AGIC cannot guarantee that the data provided is always the most current, accurate or complete data for a given location. GIS data is subject to change and, as such, data and services on the portal are also subject to change.

Users of the portal are advised to use AZGEO data and services with caution and independently verify the data’s currency, accuracy and completeness with subject experts and/or local jurisdictions.

Agencies are also able to upload and/or provide links to their specific disclaimers that users must consent to prior to gaining access to the group’s data.

How can I access and use other people’s data through AZGEO?

The AZGEO Clearinghouse features a web-based interactive map viewer that allows users to display data and, if authorized, to ‘clip n ship’ (extract and download data using user defined boundaries). Also, a data catalog allows users to search for a dataset based on data type (file or service), data custodian, data category (based on framework categories), or keywords. Shapefiles can be downloaded as .zip files via the catalog page.

How do I know if anyone is accessing the data I contribute?

AZGEO provides tracking at the server level (who has been accessing the data by domain; for example azland.gov or asu.edu) and at the application level (what dataset has been accessed and how – for example, city boundaries were downloaded as a .zip file 20 times this month). Tracking reports are available to group administrators via the web interface and can be output as .pdf, .xls, and csv.

What is the hardware/software configuration for the AZGEO Clearinghouse?

AZGEO utilizes multiple redundant Web servers, ArcGIS Application Servers, SQL/SDE Database servers and backup servers. Redundancy and the ability to transparently recover from failures is built into the system. The web management system uses a DRUPAL content management system (CMS) with custom authentication and authorization modules.

Guidelines for Geospatial Data Sharing

Intended for use by Arizona Agencies who create, maintain, and share geospatial data

AGIC

ARIZONA GEOGRAPHIC
INFORMATION COUNCIL

Data Committee

Legal and Administrative Committee

April 2013

Letter of Intent

This document serves as a best practice guide for Arizona public agencies who engage in the sharing of geospatial data. AGIC's primary mission is to coordinate the development and management of geographic information in Arizona. The AGIC Vision is to facilitate access to credible, timely, and accurate geographic information to be used by both decision makers and the citizens of Arizona. This type of access or sharing reduces duplication of data, and personnel costs throughout all branches and agencies of government. Agencies may leverage this model of sharing to fulfill public records request.

The guidelines, presented in this document, reference current Arizona State Statutes and are intended for use by all Arizona State agencies who engage in the creation, maintenance, and publication of geospatial data. All other non-State governmental agencies (federal, county, tribal, municipal or other organizational levels (e.g., Councils of Government COGS)), who reside in Arizona, may use and follow these guidelines. If any non-State governmental agency wishes to use these guidelines, the agency must remember that this document does not override or change any national or local laws, rules, policy and regulations in which the non-State agency must comply with. However, it is strongly encouraged that all State agencies who create, maintain, and steward geospatial data adopt a data sharing policy that conforms to current Arizona State Statutes and follows the guidelines recommend in this document.

If you have questions regarding this document or the implementation of a geospatial data sharing policy in your state organization, please contact AGIC or the Arizona State Cartographer's Office.

Arizona Revised Statutes as Applicable to Geospatial Data Sharing

Under Arizona Revised Statutes (ARS) 37-177, the Arizona Geographic Information Council (AGIC) was directed to “facilitate interagency cooperation for the purpose of geospatial data sharing and supporting a geospatial clearinghouse”. It is from this vantage point that AGIC is providing these guidelines for interagency geospatial data sharing. The primary Arizona Statute that provides direction in geospatial data sharing is ARS 37-178.

ARS 37-178. Geospatial data sharing

A. A public agency that shares geospatial data with another public agency may:

1. Share geospatial data without entering into a written agreement with the other public agency.
2. Share geospatial data of which it is the custodian.
3. Retain custodial ownership of any geospatial data provided to other public agencies.
4. Prohibit shared data from being redistributed by recipient public agencies if notification of the prohibition is given.

B. A public agency that shares geospatial data may exempt the data from commercial use fees prescribed in section 39-121.03, subsection A, paragraph 3.

C. A public agency that shares geospatial data of which it is the custodian is not liable for errors, inaccuracies or omissions and shall be held harmless from and against all damage, loss or liability arising from any use of geospatial data that is shared.

D. A public agency that shares geospatial data or receives shared geospatial data may withhold the shared data from public disclosure if the data consist of critical infrastructure information as defined in section 41-1801.

The geospatial data sharing statute impacts policy in the following ways:

- Agencies are not required to charge commercial use fees for geospatial data.
- Agencies may share geospatial data with other agencies without requiring a written agreement.
- Agencies that share their data with other agencies retain their custodianship of the data requiring public data requests to go back to the custodial agencies.

NOTE: Be aware that the intent of ARS 37-178, related to Geospatial data sharing, is not to circumvent ARS 39-121.03 related to public records requests for commercial purposes, but is intended to enable public agencies to more easily share Geospatial data between various government agencies in the interest of the public good. Taxpayer investment in public records should still be considered when private entities request public Geospatial records for commercial purposes. In other words, the intent of ARS 37-178 is not to give away taxpayers investments in public Geospatial records, but to allow agencies to provide Geospatial data without concern for commercial use fees when it is in the best interest of the State of Arizona and the Citizens and Taxpayers of the State.

Benefits:

Sharing geospatial data, resources, and knowledge will cut down on costs associated with developing duplicate data by different agencies. It will also result in productivity improvements

through quicker search results and easier access to relevant data and information. Together, less duplication of effort and more productivity will have a positive impact on the statewide economy and the successful performance of statewide initiatives and programs.

The concept of “collect once, use many times” is a fundamental part of Geographic Information Systems (GIS) data sharing philosophy, with associated economic benefits from leveraging and reusing existing geospatial data and resources. A brief list of benefits, including the ones already mentioned, follows:

- Minimize duplication of effort by collecting geospatial data and resources once and using them many times to improve productivity through quicker and easier access to authorized data
- Reduce the risk of using unauthorized data
- Facilitate self-service and free-up staff time now devoted to filling data orders
- Eliminate bottlenecks and wait times associated with order fulfillment
- Facilitate geospatial data integration (horizontal & vertical) and standardization of geospatial data and resources

The following information provides a common conceptual framework and guidelines to assist agencies in developing geospatial data sharing policies, thereby realizing the benefits mentioned.

Data Sharing Roles:

In sharing geospatial data, it is important to understand the role(s) that your agency may play in order to manage your geospatial data sharing efforts. These are identified as the following:

Data Custodian – These agencies or organizations are responsible for creating the data initially and are responsible for its content and ongoing maintenance, including updates. The Data Custodian is the owner of the data it creates and retains its distribution rights including sales. They are also responsible for creating the metadata describing the data in full. Public records requests must still be directed to the Custodian of the data being requested.

Data Integrator – These agencies or organizations add value to the original dataset by implementing QA/QC procedures and/or adding information to the existing dataset. Data Integrators update the metadata to reflect any changes that are made to the data.

Data Steward – These agencies or organizations are responsible for hosting or disseminating authorized datasets. Data Stewards are responsible for maintaining the authorized data and publishing the most current version of metadata as provided by Data Custodians or Data Integrators.

In some instances one agency may provide all of the above roles. However, in the event that different agencies work together to produce data, the Data Custodian retains the responsibility to process public records requests, regarding the original data, even if they have shared that data with another agency.

Action Item: Determine what data falls under your agencies custodianship and what data requires other agencies to respond to or approve public records requests.

Data Sensitivity Levels:

In establishing a geospatial data sharing guidelines, it is useful to classify data into three distinct levels of sensitivity – or ‘shareability’. The following sensitivity classifications are developed from the Arizona Strategic Enterprise Technology (ASET) data classifications which are necessary for all State agencies to follow, but they also represent a good level of classification for all interagency data sharing:

Level 1 Data: Confidential – Restricted

Level 1 Confidential or Restricted data includes data related to homeland security and public safety, or other non-disclosure purposes. Confidential or restricted data also includes Personally Identifiable Information (PII). This data would not be available for commercial or non-commercial use. It would only be made available if it was required for security related purposes. Data that would fall into this category might include, but is not limited to, critical infrastructure, crime data, emergency response related, event specific information or data that is otherwise considered ‘classified’ information.

Level 2: Confidential – Sensitive (Agency Only)

Level 2 Confidential or Sensitive data includes data that cannot be shared publicly due to socio-political, environmental, or legal constraints. Examples of this type of information might include locations of threatened and endangered species, archeological sites, addresses of individuals, or other data the custodian agency classifies as confidential or too sensitive to share. While this data can be shared with other agencies or outside groups, an official request for the data, with associated approval of the custodial agency, must be completed.

Level 3: Public

Level 3 Public data includes data that can be shared openly with other agencies and the public. Examples of public data include most administrative boundaries, elevation, transportation routes, cadastral data, locations of community anchor institutions and other types of geospatial data considered ‘framework’ information. This data is often provided by, or based upon, Federal data which is public.

As a first step toward developing or updating a data sharing policy, agencies are encouraged to identify the geospatial data, for which they are the custodian, as falling into one of these three categories. More detailed sharing procedures should be based on data sensitivity level.

Action item: For data under your agencies custodianship, categorize datasets as to sensitivity levels 1, 2 or 3 and subsequently:

Determine how different sensitivity level data should be shared:

Display/view only
Data services
Data downloads (web) or media distribution (non-web)

Determine how different sensitivity level data should be accessed

Hard copy
Media-based data distribution (i.e. DVD or USB drive)
File-based distribution on the web
Web mapping application
Data or map service

Disclaimers:

Many agencies and organizations utilize disclaimers when sharing geospatial data in order to ensure the appropriate and applicable data usage, level of access, and/or data accuracy is understood. These Guidelines propose using the AGIC GEOData Clearinghouse (AZGEO) Disclaimer for a standard disclaimer, in order to facilitate consistency among agencies.

Data disclaimers generally address the following:

- Data provided 'as is'
- No warranty concerning accuracy, completeness, correctness
- No liability for damages
- Limit on distribution

To see examples of Arizona agency disclaimers, please see the following:

<http://gisweb.azdeq.gov/arcgis/emaps/?topic=places>

<http://www.azwater.gov/azdwr/gis/>

<http://www.co.apache.az.us/Departments/GIS/GIS.htm>

<http://www.yavapai.us/mis/gis-mapping-applications/>

<http://www.mesaaz.gov/maps/>

Action item: If your agency does not already have a data disclaimer, use these examples to help you create one. Involve your Risk Management Group to insure all agency requirements are met.

Metadata:

Metadata, or the "data about data", is required to share data effectively and judiciously. Metadata should provide sufficient description of the geospatial data to facilitate users finding the resource, determining if the geospatial data is appropriate for their intended use, and information on how to access the geospatial data.

AGIC endorses use of the following 'minimum metadata' content. This content is taken from the Federal Geographic Data Committee's "Content Standard for Digital Geospatial Metadata"

(CSDGM) Version 2 - 1998. (FGDC-STD-001 June 1998). Note: the same content is also included in ISO19115.

- 1) Citation – Basic information about the data set:
 - a) Originator or Steward (who developed the data set; e.g., Arizona Department of Environmental Quality; or “unknown”)
 - b) Title (what the data set is called; shapefile/geodatabase name)
 - c) Published by (name of the person or organization that published the data; often the same as the person or organization in Origin)
 - d) Published at (place where data set was published/produced; e.g., Redlands, California, USA)
 - e) Publication date (date when data set was published, e.g., 2010; unknown; unpublished)
- 2) Description – A characterization of the data set, including its intended use and limitations:
 - a) Abstract (brief summary describing the data set)
 - b) Purpose (why the data set was created)
- 3) Time Period of the Content – Time period information about the data set:
 - a) Date of data set (creation or date range the data covers)
 - b) Status of data (e.g., complete, in progress, planned)
 - c) Data currentness (e.g., “data are accurate as of...” ground condition, publication date, etc.)
 - d) Maintenance & update frequency (e.g., unknown, as needed, annual, monthly, etc.)
- 4) Spatial Reference Information
 - a) Map Datum, Projection and Units
- 5) Bounding Box – Spatial extent of the data set:
 - a) North (north bounding coordinate)
 - b) South (south bounding coordinate)
 - c) East (east bounding coordinate)
 - d) West (west bounding coordinate)
- 6) Keywords – Words or phrases that summarize aspects of the data set:
 - a) Theme (subject covered by the data set – helps to use ISO categories, like agriculture, environment, farming, inland waters, etc.)
 - b) Theme thesaurus (usually “none” is put in here unless you are using a standard controlled vocabulary, like from the Library of Congress or the Getty Museum)
- 7) Data Set Constraints – Statement of use or access restrictions or constraints on the data set:
 - a) Data Sensitivity Level (Level 1: Confidential – Restricted; Level 2: Confidential – Agency Only; or Level 3: Public)
 - b) Access restrictions (e.g., none; or access is restricted to select users)
 - c) Use restrictions (e.g., none; or for use in general planning)
- 8) Contact Information – information relating to the organization (or person) knowledgeable about the data set:

- a) Name of contact person or organization
 - b) Contact person/organization address and phone number
- 9) Distribution Information – information detailing how the data set can be accessed or obtained.
- i) URL for direct online access
 - ii) Ordering instructions for media distribution
 - iii) Service metadata for data service access

For a more complete overview of the FGDC metadata content specification, please see: <http://www.fgdc.gov/csdgmggraphical/index.htm>

Action Item: Define metadata policies

- *What profile will be used*
- *How will metadata be encoded (FGDC XML, ISO XML, other...)*
- *How will metadata be made available*
 - *Bundle with data files*
 - *Web accessible directory*
 - *Insert into a metadata catalog*
- *Metadata maintenance plan*
 - Update workflow*
 - Error correction process*
 - Testing, quality assurance*

Summary:

By first analyzing your agency's geospatial data for content and sensitivity, you can begin to discern the potential for expanding access (or not) to your data. When sharing any data outside of your agency, it would be prudent to first determine the Data Custodian and Data Steward for long term purposes, and possibly some of the Data Integrators. A source of geospatial data is considered complete when it is made available (either through a web site or access point), has any and all disclaimers identified, and is provided with complete metadata documentation describing the data and its accessibility.

It is the intent of these Guidelines to provide agencies with a method of developing data accessibility and sharing procedures, with the desired goal of providing greater and easier access to geospatial data among agencies that maintain it. By providing access to each other's hard earned, and consistently maintained data we can all benefit from a more reliable, accurate, and up-to-date mapping base, ultimately enriching all agency operations and decision making.

6: NRRC GIS State Survey

The Arizona Geological Survey conducted a brief survey on behalf of the NRRC GIS Subcommittee of the coordinated and centralized GIS data discovery and access systems for states that have data available online. The survey was conducted from the user's perspective to evaluate the similarities and differences between state GIS data distribution systems and to identify states that have successfully implemented coordinated, centralized GIS data systems.

Specifically, the survey attempted to answer four basic questions:

1. Where is the data?
2. How do I access the data?
3. What is the data format?
4. Can I use the data?

The answers to these questions indicate how easy or difficult it is to find, access and use state GIS data that is available online.

Research methods

I used google to search for GIS data in each state. When I found a State's GIS website, I searched or browsed for specific GIS data themes. I tried to locate data download and viewing options that each state made available through their GIS data site.

For each site I visited, I collected the following information:

1. The visual aesthetics of the GIS data discovery and access user interface
2. The number of mouse clicks to data download
3. Whether or not the state had a centralized GIS database and/or data distribution system
4. Whether or not data could be downloaded directly from the agency website
5. Data distribution formats
6. Whether or not the website had search and/or browse data capabilities
7. How easy/difficult the search and/or browse capabilities were

I put this information into in an Excel document.

Survey results

Forty states have their GIS data available online. It comes in a variety of formats and can usually be accessed in more than one way. Most data is distributed through a data clearinghouse or a data portal. A data clearinghouse and a data portal are essentially the same thing, they are just called different things depending on the agency's preference.

Where is the data?

Generally speaking, most states have a central, online location where GIS data can be discovered and then accessed through a portal, clearinghouse or some other external site hosted by the data publisher. Other states have multiple online locations where GIS data can be discovered and accessed, usually the state agency responsible for publishing the data.

Of the forty states that have their data online, North Dakota is the only state that appears to have a fully coordinated and centralized GIS. North Dakota has one online location, the North Dakota GIS Hub Data Portal, where all GIS data for the state can only be discovered and accessed. All North Dakota State agencies that have GIS data direct people to this portal from their website.

How do I access the data?

Each state has different data accessibility options. At a basic level, these options include viewing data and downloading data.

Viewing data is facilitated by desktop applications and web applications. An example of a desktop application is ArcMap. An example of a web application is the Arizona Game and Fish Department’s Habimap (<http://habimap.org/habimap/>).

Most states offer direct data download from an online location, often multiple locations. These locations include a website, a portal, a clearinghouse, an FTP site or email access. Depending on the state, it is possible to download a single dataset at a time or bulk download data.

Some state agencies provide data as data services. Depending on the server configuration, data as services can be viewed or downloaded. Data as services can also be added to desktop and web applications. ArcMap is an example of a desktop application and the National Geothermal Data System (NGDS) Feature Search & Map application (<http://data.geothermaldatasystem.org/>) is an example of a web application. The NGDS application allows people to search for a data service, add it to the map, view the tables and geographic distribution of the data in the service and download the dataset.

Utah is the only state that offers a direct database connection.

The table below summarizes data access based on data distribution format. State Totals are the number of states that distribute data in the specified distribution format.

Distribution Format	Access Type	State Totals
Website download	Download	37
FTP	Download	8
Email	Download	3
Database	View, Download	1
ArcGIS Services	View	13
WMS (Services)	View	12
WFS (Services)	View, Download	4
Viewer (Online application)	View	12
Viewer Download (Online application)	View, Download	4

Table: Data Distribution Format, Access Type and State Totals

What is the data format?

Downloadable data formats include geodatabases, shapefiles, spreadsheets, documents, etc.

Can I use the data?

In my search, I specifically wanted data that I could use in ArcGIS. If I could download the data as a shapefile, geodatabase, then I could use it almost immediately. If I could connect to a database or data service, I could also use it almost immediately. Other data formats are possible for use in a GIS, but require a bit more processing.

Other data use considerations that were not evaluated in this survey include spatial and temporal resolution, spatial extent and whether or not dataset has appropriate metadata.

Conclusion

The Arizona Geological Survey's brief survey concludes that most states have similar GIS data distribution system components. The biggest difference lies in the implementation of the system. From the user perspective, this translates to the user interface and user experience.

As expected, the ease of data access and use varied from state to state. Of the forty states that have GIS data online, five states implemented systems that had great strengths that set them apart from other systems.

From the user perspective, some states have well-coordinated, centralized GIS programs while others do not. North Dakota, for example, has a well-coordinated, centralized GIS system. All GIS data for all state agencies is only accessible through one location, the NDGIS Hub Data Portal.

The State of Utah has an easy-to-understand, aesthetically pleasing landing page for their initial data web access point. The State of Utah also distributes data in a variety of formats, enabling users to pick the data distribution method that works best for them.

New Mexico, Missouri, and West Virginia all have a user interface that is simple and intuitive, with no more than 3 clicks to data download. This translates to ease and speed in terms of how easy it is to find and access data.

7: U.S. Geoscience Information Network (USGIN) System Elements

The system is designed based on a service-oriented approach using open standards to support data access by a wide variety of software applications, promote novel approaches to data analysis, and foster the development of tools by third parties. It is based on the U.S. Geoscience Information Network (USGIN) which enables users to efficiently find, access, and share geoscience data, reducing the time and effort spent locating and integrating useful information and document new data by providing for information registration (by providers) and discovery (by users) based on standardized catalog services and metadata. The system accommodates resources in various forms, from unstructured text and images to documented, community Web services and interchange formats. To simplify, USGIN is a collection of Web-accessible resources that are registered in online catalogs and conform to data-sharing practices. Because USGIN is a network, it is largely intangible and rather defined by the protocols, interchange formats, and conventions that enable its operation. Thus the documents that define these protocols, interchange formats, and conventions are the artifacts that represent the network. These documents are archived on the USGIN repository (<http://repository.usgin.org>), indexed through the USGIN Catalog, and accessed using standard Web resource retrieval. Links to the major specifications documents are also provided in the references section of this document.

The core components of the network are information exchange specifications and the catalog function. Information exchange specifications are community agreements on the conventions necessary for the interoperable exchange of some particular information. The exchange specification specifies a scope, the content model (descriptive components) for the data items of interest, interchange formats for encoding and transmitting information electronically, and the protocols used to request information. By using an interchange exchange specification a data provider can publish data that will be available to any client that implements the exchange, and a client application can access data from any provider publishing data according to the exchange conventions. This allows data in a variety of formats, organization, and structure to be integrated without having to manually transform it.

The catalog function comprises a collection of metadata records that describe resources accessible through the network (in this case, USGIN), and a special information exchange that defines metadata content, how the metadata collection is searched, and how metadata are encoded in search responses.

As a system, USGIN is a loosely coupled system of independent data providers, client applications, and infrastructure. The infrastructure includes 1) tools for registration of new resources, searching metadata catalogs, authentication, and resource validation; 2) registries for vocabularies, agents, specifications, and interchange schema; and 3) documentation and educational resources. Because network operation is based on information exchange specifications that are independent of any particular hardware or software implementation, all of the operational components can evolve as technology evolves. Use of standard protocols enables data access using off-the-shelf software, both commercial and open-source. Open

Geospatial Consortium (OGC) Web Map Service (WMS) and Web Feature Service (WFS), and OpenDAP NetCDF services are specified in current USGIN information exchanges. These services were specifically chosen due to the open-source nature and existing developer community. Details on the use of these services, and interchange formats for data exchange, are can be found in the References & Links section of this document.

The distributed nature of the system means that stewardship of resources is determined by the resource owner. Participation in the network only requires that a resource provider create metadata that conforms to the USGIN profile, and make the metadata and the described resource available. The network is open; anyone can deploy new nodes and components that implement one or more USGIN specifications, without requiring approval. New specifications can be introduced for service protocols, interchange formats, or vocabularies. Keeping resources under the stewardship of the parties responsible for the information promotes system sustainability because the stewards have a direct connection with the quality of the product rather than submitting the data to a centralized database and related data manager.

Elements of the Network

- **Catalog:** the catalog is a collection of metadata records describing resources that are intended to be considered part of USGIN. A catalog should be thought of as a cloud of records, there may be no individual data-store on the network that contains the entire collection of metadata as catalogs can harvest from one another (improving functionality and search capability) and expose the data to commercial search engines (e.g. through OGC's Catalog Service for the Web, CSW). The use of a standardized encoding scheme for metadata interchange and a standard service protocol for accessing catalog content simplifies interoperability and enables the use of off-the-shelf software to implement the catalog system. The scope of the catalog includes data products or datasets; these are units of information that have been authored, edited or compiled under some stewardship, with some purpose and procedure, and having a common collection of access processes.
- **Metadata:** the data about the data. Individual documents require one metadata record per document. These documents might be scans of well logs, scanned reports, maps, or publications, multimedia files, or data compilation spreadsheets (e.g. Excel or Access file). Some document types may consist of a bundle of files, e.g. Esri shape file. In general, these should be bundled into a single file like a zip archive or UNIX tar file. The metadata must include the URL at which the document can be accessed, if it is not accessible online, the metadata should provide instructions on how to access the document. Full metadata documentation can be found in the linked resources section of this document under USGIN Standards and Protocols Drafting Team, 2010. In order to maximize usability and interoperability, we have adopted ISO metadata schemas for geographic information, ISO 19115 and 19116 as well as ISO 19139 for XML encoding.
- **Community Specifications:** One of the operating principles of USGIN is to not "reinvent the wheel," i.e. to use and extend existing software whenever possible instead of developing new software. Active, free, open-source software communities offer a path to

long term viability of key system components whether or not individual projects continue. USGIN specifications, in general, define conventions and practices for the use of existing components and standards to simplify interoperability.

Scope of Work for implementation

A sustainability study recently completed by AZGS for the National Geothermal Data System, using the US Geoscience Information Network as the data integration framework, identified what components are likely to be necessary to sustain operations.

Servers and Software

This list assumes that network infrastructure (internet connection, switches, firewalls, DNS, etc.) are in place within NRRRC agencies that

- Catalog and repository server: Linux/Tomcat/PostgreSQL/ GeoNetwork or Geoportal (could run on Windows stack as well). Only one is essential, any number is possible.
- To support a catalog system: a server for registries for identifiers and vocabularies, in addition to repositories for system specification documents and resources, like XML schemas that must be web accessible to support service operations.
- Data server: Windows/ArcGIS Server (dbms optional if shapefiles are used as data source), or Linux/Tomcat/PostgreSQL/Geoserver or Mapserver. One server with modest capabilities could serve perhaps 300 Mb of data, depending on load.

Personnel

- Technical IT personnel: Need capabilities to deploy server software, load data, configure services, debug http traffic if there are problems. Data preparation (if standard interchange formats are being used) requires understanding of ETL using SQL queries and other techniques. Some understanding of XML and XML schema is required occasionally; GeoServer requires mapping from XML to database fields in an XML configuration. Someone with understanding of metadata content models and encoding is likely to be essential to get a catalog system working well.
- Outreach and marketing: personnel dedicated to the production and maintenance of documentation and educational materials, as well as face-to-face and online training programs.

Maintenance

- Individuals dedicated to system management, including arrangement of meetings, maintenance of hardware, user help lines, and network operation.
- Software development and testing:

Components for service deployment may be necessary if off the shelf solutions do not meet all requirements; applications for service and conformance testing, development of metadata, and network monitoring.

Transition from current stovepipe data discovery and delivery systems to a loosely coupled, service based architecture can probably be done by large organizations like the USGS with the personnel and hardware they currently utilize, following whatever IT/hardware refresh cycles they currently use. Additions to hardware capacity in the form of servers, bandwidth, and online storage will be necessary to bring new data online, but this would be true no matter what approach to data delivery is adopted. The most significant additional investment will likely be in education of personnel who develop and deploy data and metadata services, in addition to human effort for data integration, documentation, and migration of existing data to new formats and delivery protocols. These investments will provide long-term return in staff capabilities if, as anticipated, these become widely accepted standard operating procedures. An additional return is the expected increase in utilization of the information resources and greater visibility for the agency providing data.

Maintain System Providers/Nodes

Even though each node will make their own choices regarding computer manufacture, operating system, network providers, etc., (another strength of a distributed network) and while there is currently no provision for ongoing federal support of individual nodes on the NGDS, it would be worth considering implementing a Service Level Agreement (SLA). SLAs are a common vehicle in the IT world for specifying expectations including system quality attributes such as interoperability, reliability, availability, recoverability, performance, integrity, confidentiality, etc. Suggested standards could potentially include some or all of the following (Allison, et. al., 2013):

- Minimum storage capacity and provisions for added growth of disk space
- Minimum network connectivity bandwidth
- Regular backup and maintenance procedures according to a schedule, with off-site storage of backups
- Codified disaster recovery plan with alternate facilities (For those organizations without an alternate backup facility, they could agree to serve as a backup facility with another node in exchange for the same)
- Physical system security precautions
- Software system security precautions
- User administration support (for data contributors to that specific node and/or to block inappropriate use from within their network domain)
- Commitment to participate in the system using the standards and protocols established by the lead team

8: USGIN Community Standards, Profiles, Protocols, and Software

The following table represents a list of community standards, profiles, protocols, and software employed by USGIN to address the challenges associated with sharing large amounts of diverse data stored in geographically disparate locations. In general, these are not the only solutions available, but indicate resources that the developers have found useful. USGIN is proposed as the framework for data integration for NRRC.

CATEGORY: APPLICATIONS		
Specification	Description	Uses
Amazon Web Services	Computing infrastructure resources available for rent. More...	Several servers that host USGIN Websites or provide USGIN services are provided by Amazon Web Services.
Apache Tomcat	Free-and-open-source servlet engine. Many key applications for providing Web services can be hosted using Tomcat. More...	Used to host software applications, such as GeoServer and GeoPortal, which provide Open Geospatial Consortium (OGC) -compliant Web services
Django	Open-source Web application framework written in the Python programming language. More...	Used to develop a number of USGIN Web applications, including the USGIN URI redirection engine. More...
Drupal	Modular, extensible, actively maintained free-and-open-source Web site content management framework; capable of running on any PHP-capable Web server. More...	Websites related to the USGIN Initiative built using the Drupal include: The AASG Geothermal Data Website The NDGS Portal , USGIN Lab site The USGIN Website. More...
Esri ArcGIS	Commercial geographic information system software and Web service-compatible server software. Supports both proprietary service protocols ('Geoservices API'), and OGC services.	Commonly used for deploying geospatial data as Web services .
GeoServer	Free-and-open-source Web service-compatible server software. More...	Southern Methodist University, a contributor to the National Geothermal Data System , uses GeoServer to host Web services; the landing page is here .
PostGIS	Extension for PostgreSQL relational	PostGIS databases are used to persist

CATEGORY: [APPLICATIONS](#)

Specification	Description	Uses
	database to support geographic objects and operations. More...	spatial data for ArcGIS, GeoServer, Django, and many other applications.
PostgreSQL	Free-and-open-source object-relational database system with over 15 years of active development. More...	PostgreSQL is used by GeoPortal, GeoNetwork, CKAN. [CKAN is used by data.gov]
Python	A programming language. More...	Python is used for USGIN Django sites and converting Excel spreadsheets into ISO metadata (the application responsible for which can be found at http://github.com/usgin/csvtometadata).

DRAFT

9: Acronyms

ADEQ – Arizona Department of Environmental Quality
ADOA – Arizona Department of Administration
ADWR – Arizona Department of Water Resources
AFIS – State of Arizona Financial System
AGFD – Arizona Game and Fish Department
AGIC – Arizona Geographic Information Council
ALRIS – Arizona Land Resource Information System (part of Arizona State Land Department)
ASLD – Arizona State Land Department
ASP – Arizona State Parks
ASP-SHPO – Arizona State Parks, State Historic Preservation Office
AZDA – Arizona Department of Agriculture
AZGEO – Arizona Geodata Portal (managed by the Arizona Geographic Information Council)
AZGS – Arizona Geological Survey
AZSF – Arizona State Forester
AZSPOC – Arizona Single Point of Contact
DOE – U.S. Department of Energy
FGDC – Federal Geographic Data Committee
FTE – Full Time Employee
GIS – Geographic Information Systems
GMA – National States Geographic Information Council Geospatial Maturity Assessment
ISO – International Organization for Standardization
IT – Information Technology
NGDS – National Geothermal Data System
NRRC – Natural Resources Review Council
NSDI – National Spatial Data Infrastructure
NSF – U.S. National Science Foundation
NSGIC – National States Geographic Information Council
OGC – Open Geospatial Consortium
OMB – U.S. Office of Management & Budget
OSTP – U.S. Office of Science and Technology Policy
RDCC – Utah Resource Development Coordinating Council
REST – Representational State Transfer
SCO – State Cartographer's Office (part of Arizona State Land Department)
SOW – Scope of Work
USGIN – U.S. Geoscience Information Network
WFS – Web Feature Service
WMS – Web Map Service
WRP – Western Regional Partnership
XML – Extensible Markup Language