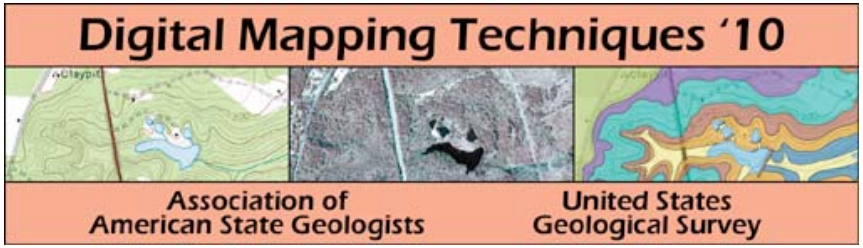


The following was presented at DMT'10
(May 16-19, 2010).

The contents are provisional and will be
superseded by a paper in the
DMT'10 Proceedings.

See also earlier Proceedings (1997-2009)
http://ngmdb.usgs.gov/info/dmt/

THE NATIONAL GEOTHERMAL DATA SYSTEM: THE GEOSCIENCE INFORMATION NETWORK IN ACTION

by Ryan Clark, Stephen Richard and Wolfgang Grunberg - Arizona Geological Survey

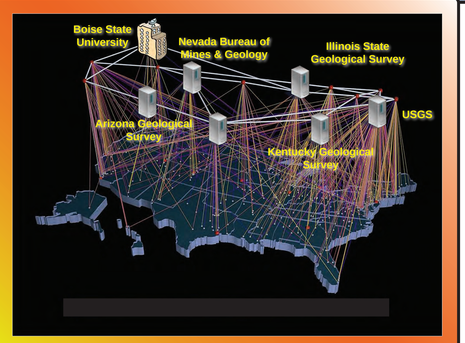
What is the National Geothermal Data System?

The NGDS is a distributed network of databases that are collectively building a system for acquisition, management and maintenance of geothermal and related data.

Users of the system will include federal and state agencies, researchers, decision makers, the general public, educational institutions, the geothermal industry, and financial institutions. The NGDS will be designed by using and adapting existing technology as well as emerging informatics standards and protocols.

Major participants in the NGDS to date are: GeoStrat Digital Information System, Energy & Geosciences Institute, Great Basin Center for Geothermal Energy, Geo-Heat Center, Stanford Geothermal Program, and U.S. Geoscience Information Network. The U.S. Geological Survey, Bureau of Land Management, and Geothermal Energy Association will contribute to building the system as will a number of community task forces that target specific issues.

What are the Goals of the Data System?



Access to these distributed data will be provided through a desktop application (Geothermal Desktop) as well as via the web sites of member databases. Online web-services can also be accessed through a variety of GIS Client applications, such as ArcMap and Google Earth.

A distributed system means all data does not come from one site. It is about data sharing and interoperability among the linked sites so data can be easily discovered and downloaded in compatible formats and easily aggregated by the user to meet their particular needs. The NGDS must make it as easy as possible for the users within the realistic framework in which we will be working.

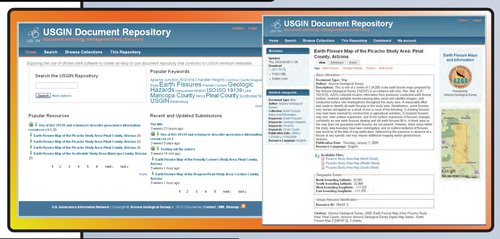
The second phase of this project aims to expand and enhance the National Geothermal Data System (NGDS) by creating a national, sustainable, distributed, interoperable network of state geological survey-based data providers that will develop, collect, serve, and maintain geothermal-relevant data that operates as an integral compliant component of NGDS. We will bring data from the State Geological Surveys (via AASG) into the NGDS, by digitizing at-risk legacy, geothermal-relevant data (paper records, samples, etc), publishing existing digital data using standard NGDS data services, and through limited collection of new data in areas lacking critical information.

Online Document Repository

For new and existing digital documents that are not online and do not already have metadata written, we're developing an online document repository. This open-source, Drupal-based web application allows the documents to be made available online, while providing a simple form for metadata creation. This metadata can be harvested into the NGDS Catalog.

These document repositories will be run by regional hubs, although any interested group may also decide to run their own.

(beta version) **repository.usgin.org**

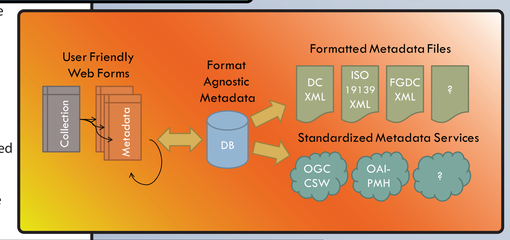


Online Metadata Editor

For existing online datasets and services without useful metadata, an online metadata assistant is being developed.

Another open-source, Drupal-based web application, this system will speed up manual metadata creation through user-friendly web forms. Interoperable, expandable, conceptual metadata records are stored in a database and can be exported in a number of standard formats.

Again, these records can be made available to the NGDS catalog, making your data available to a wide audience.

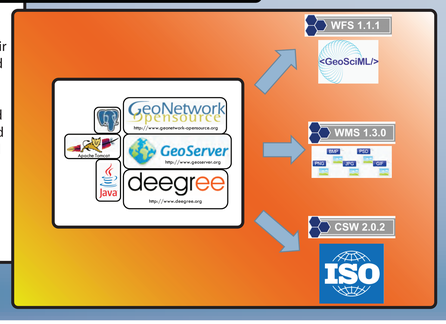


Amazon Virtual Machine Images

Some data providers may not have the hardware that would be required to put their own web-services online. These providers can either work through a regional hub to provide their data to through the hub's servers, or may be interested in hosting their data on an Amazon Virtual Machine.

We're developing Amazon Machine Images that would allow anyone to spin up a virtual server pre-configured with the software and documentation that you'll need to provide online OGC Services at a much lower cost-of-entry than buying a physical server.

lab.usgin.org/applications/amazon-web-services
aws.amazon.com/ec2



What Are We Doing to Make This Easier? The AZGS At Your Service...

Demo Services

We've been working with our partners at the Great Basin Center for Geothermal Energy, the Geo-Heat Center and Southern Methodist University to spin up some sample services, such as the nationwide heatflow dataset shown here.

These sample services help users to understand the benefits of providing data through web-services, and also gives developers something to work with while creating applications to consume these services.

Look for services at these URLs:
services.usgin.org/gsvr
services.azgs.az.gov/arcgis/rest/services

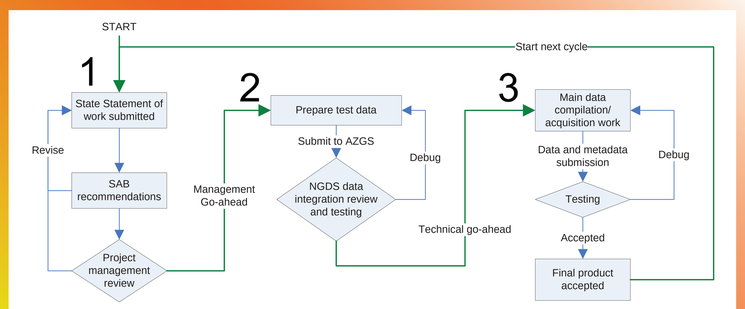
Documentation / Help Desk

The AZGS has the task of assuring that data compiled by the project subcontractors are integrated into the National Geothermal Data System.

This means that we are required to help you generate a effective workflow to get your data online, ideally as web-services.

For those unfamiliar with the setup and maintenance of software capable of providing these services, we're developing lots of online documentation at **lab.usgin.org**. We also will work closely with data providers and regional hubs to make sure that we all know how to keep the system running.

The Process of Contributing Information to the Data System



Participants have two options on how to make their data available:

- Register files in an NGDS-compliant document repository.
- Implement a web service, either at your agency, a NGDS data center (Boise State, regional project hubs at state geological surveys--Kentucky, Illinois, or Nevada), or by arrangement with another agency.

Data will be considered part of the NGDS when it is locatable using the NGDS core catalog, and accessible via the web according to procedures described in the metadata record obtained from the NGDS core catalog.

Metadata should be created and submitted for any resource that is meant to be accessible individually via the web. If metadata already exists, it will need to be checked to ensure that it meets the minimum requirements defined for the system. Tools are being developed by the AZGS to generate acceptable metadata records quickly and easily.

We are planning that each state data compiler will be operating on an approximately one year data development cycle.

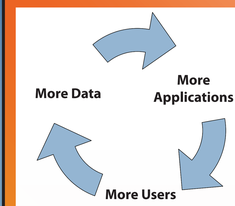
After submission, review, and acceptance of the statement of work for the cycle, the state will prepare a small test data set to prototype the data delivery content and format for each deliverable product. The NGDS development team at AZGS, in collaboration with the NGDS core development team at Boise State, will review this prototype and test integration of the data with the NGDS system. Debug iterations will be made between the NGDS system and state data producer until the prototype is demonstrated to work and provide the necessary content.

At this point the project management team will give the state provider a go-ahead for development of the full data product for that cycle. The prototype dataset should at that point be made 'live' (accessible online) in the system, but flagged as a development data set. It is strongly recommended that as data are developed, they are incrementally added to the prototype delivery dataset so that content is available as early as possible, and problems with scaling will become apparent if they emerge.

When the state producer has completed the data development, they will make a final data submission for technical review and approval. Ideally the final submission will simply be a declaration that the live dataset evolved from the prototype is complete and ready for final review.

Upon approval by the project management, the data producer may submit their Statement of Work for the next data development cycle.

What Are the Long Term Objectives of the Geoscience Information Network?



The process of building the Geoscience Information Network can be likened to the development of the world-wide web:

At first only a few datasets were provided in HTML format and transferred via HTTP. However as people realized the potential, newer and better applications for generating and consuming HTML were built, making it easier for more people to participate, which made development of even better tools more lucrative...

Converting data from poorly documented, cryptic shapefiles to slick, interoperable web-feature services with useful metadata may seem as impossible as turning lead into gold, but not only is it **remarkably easy**, it is becoming **increasingly** necessary if you want people to find and use your data.

This effort hinges on development of a community of practice using a shared collection of protocols for publishing, finding, and delivering digital information online.

We want to convince you that:

1. Web-services are better ways to provide your data than shapefiles. They are literally the future of data sharing.
2. Web-services are becoming an increasingly important part of data generation, portrayal, distribution and acquisition.
3. It is **desperately** important to provide useful metadata for any data you ever want anyone else to be able to use.

After we've convinced you of that, we will see the expansion of this community of practice that will lead to the development of better tools for finding and viewing data, which convince more users to participate, leading to more data, and more tools , and more users, and more data...

FOR MORE INFORMATION:
www.geothermaldata.org
www.usgin.org
lab.usgin.org
ryan.clark@azgs.az.gov

National Science Foundation
WHERE DISCOVERIES BEGIN
U.S. DEPARTMENT OF ENERGY
AASG Association of American State Geologists
USGS science for a changing world
Arizona Geological Survey
416 W. Congress St. Suite 100
Tucson, AZ 85701
AZGS