Geospatial Platform / CKAN Publisher Webinar

25 July 2013

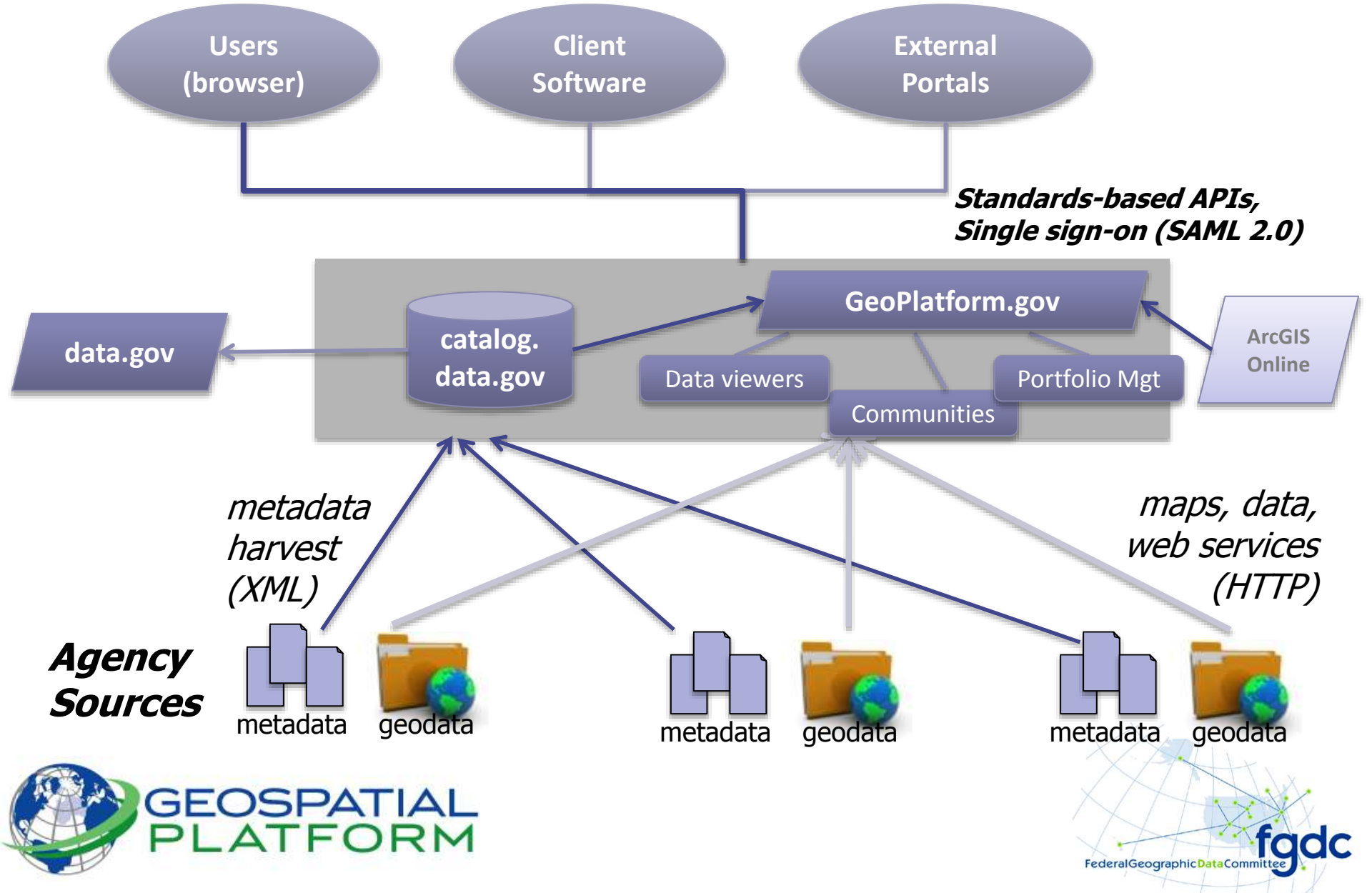
Geoplatform.gov 2.0

- Goals:
  - Facilitate the **use** of geospatial data and tools – links to data, services, and apps
  - Search/Catalog shared with ***data.gov***
  - Focus on NGDAs / A-16 data
  - Focus on non-experts / decision makers
  - Content organized by Communities
    - A-16 Theme Communities
    - Cross-Agency Collaboration Communities
    - Agency “storefront” Communities

GEOSPATIAL PLATFORM

fgdc FederalGeographicDataCommittee

Geospatial Platform architecture



GEOSPATIAL PLATFORM

fgdc Federal Geographic Data Committee

Current approach

- The Geospatial Platform environment is a successor to the GOS and geo.data.gov solutions, harvesting from federal, state, local, academic, and tribal sources
- We seek to provide a customized index and visualization environment for government data in collaboration with data.gov
- Online data, services, and apps are the desired resource types, described in metadata

GEOSPATIAL PLATFORM

fgdc
FederalGeographicDataCommittee

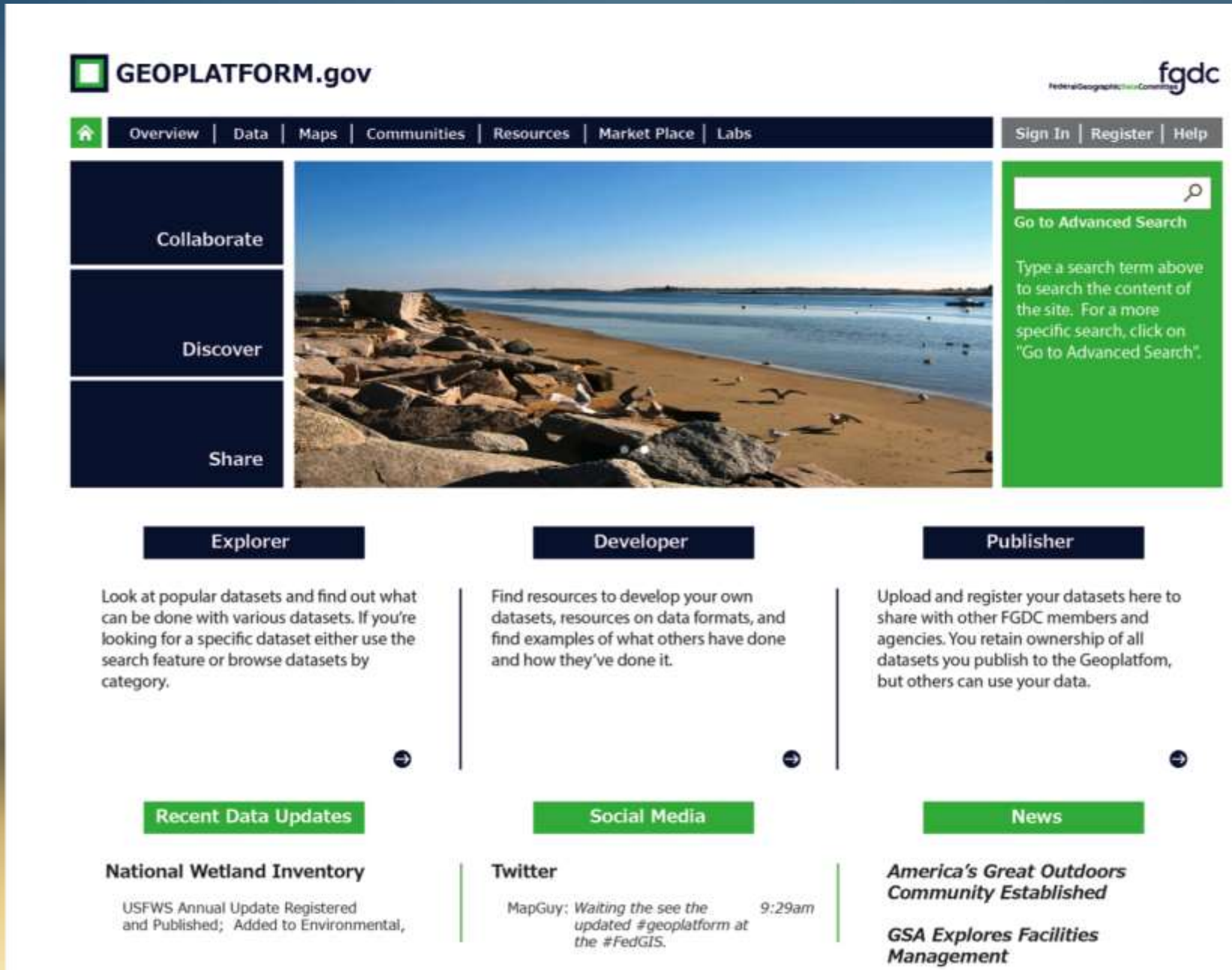
Features of the new environment

- New common catalog.data.gov search engine(CKAN) shared with data.gov
- Geospatial metadata provided by agencies in public Web folders or via OGC CSW protocol
- Format of metadata as FGDC (CSDGM) or ISO 19115/19139 in XML format
- Metadata collections are registered and harvested to refresh the index at catalog.data.gov
- Links in metadata point to download, Web services, or other access URL
- Links to map services are launched in a map viewer; KML and GeoRSS will be supported in the future

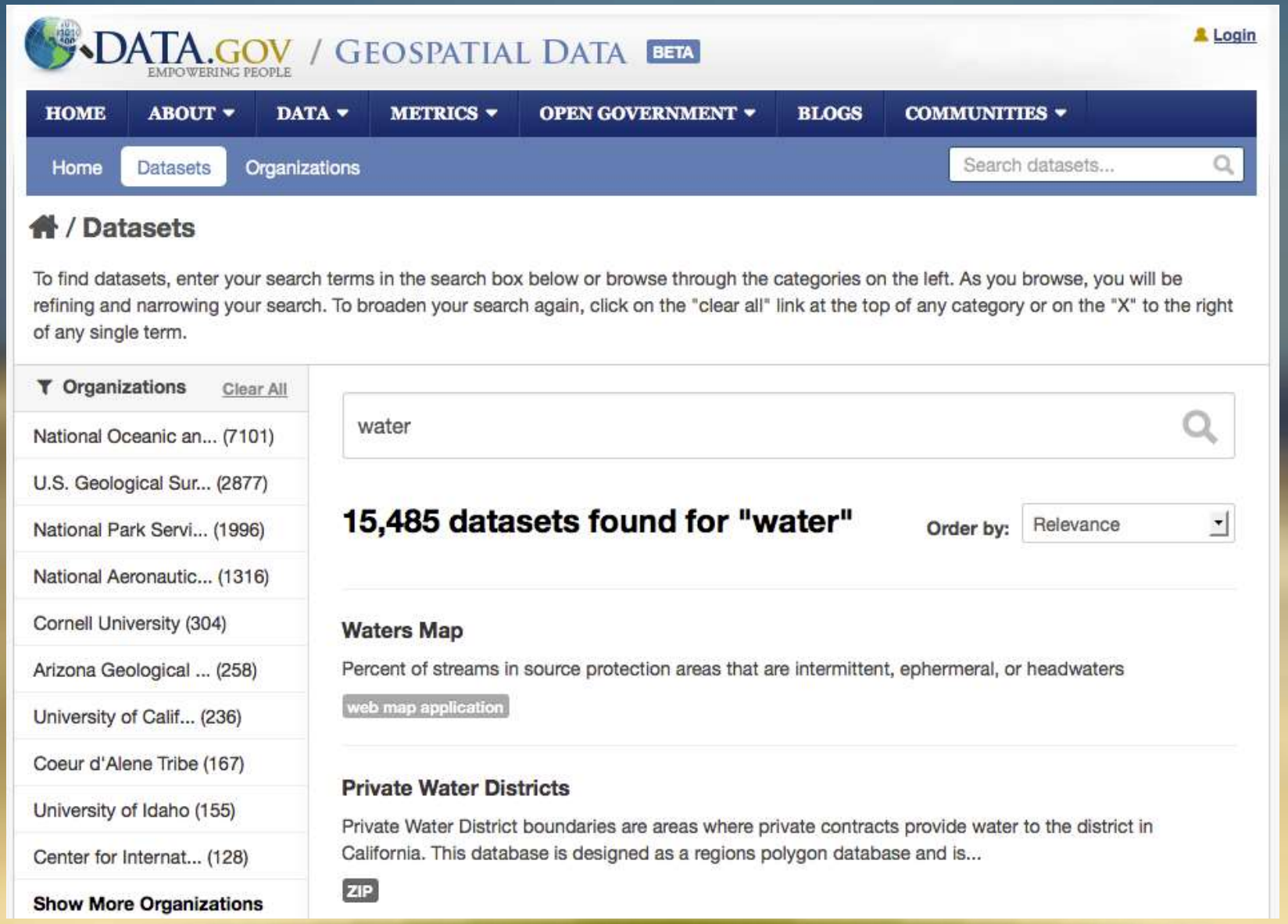
GEOSPATIAL PLATFORM

fgdc
FederalGeographicDataCommittee

Geoplatform.gov (National geoportal)



CKAN open source catalog software



search

Tags Clear All

- United States (122)
- US (122)
- U.S. (122)
- Alaska (122)
- AK (122)
- Aleutians West Cens... (26)
- Lake and Peninsula ... (18)
- Aleutians East Borough (17)
- Kodiak Island Borough (15)
- Skagway-Hoonah-Ango... (14)

Show More Tags

Organizations Clear All

- U.S. Geological Sur... (122)

Show More Organizations

Formats Clear All

122 datasets

Order by: Relevance

None: U.S. Geological Survey Earth Resources Observation and Science (EROS) Center (pointOfContact) × Tags: AK ×

Search facets provide instantaneous combination search filtering on multiple filters as selected by users on the left. This example is of USGS source in Alaska. This is a sort of browse that facilitates data exploration. We plan to populate these search facets based on controlled vocabularies on the processed metadata, such as ISO Topic Category, A-16 theme, agency, resource type, and geographic extent.

Shuttle Radar Topograp
The Shuttle Radar Topograph
Agency (NGA). Flown aboard

Shuttle Radar Topograp
The Shuttle Radar Topograph
Agency (NGA). Flown aboard

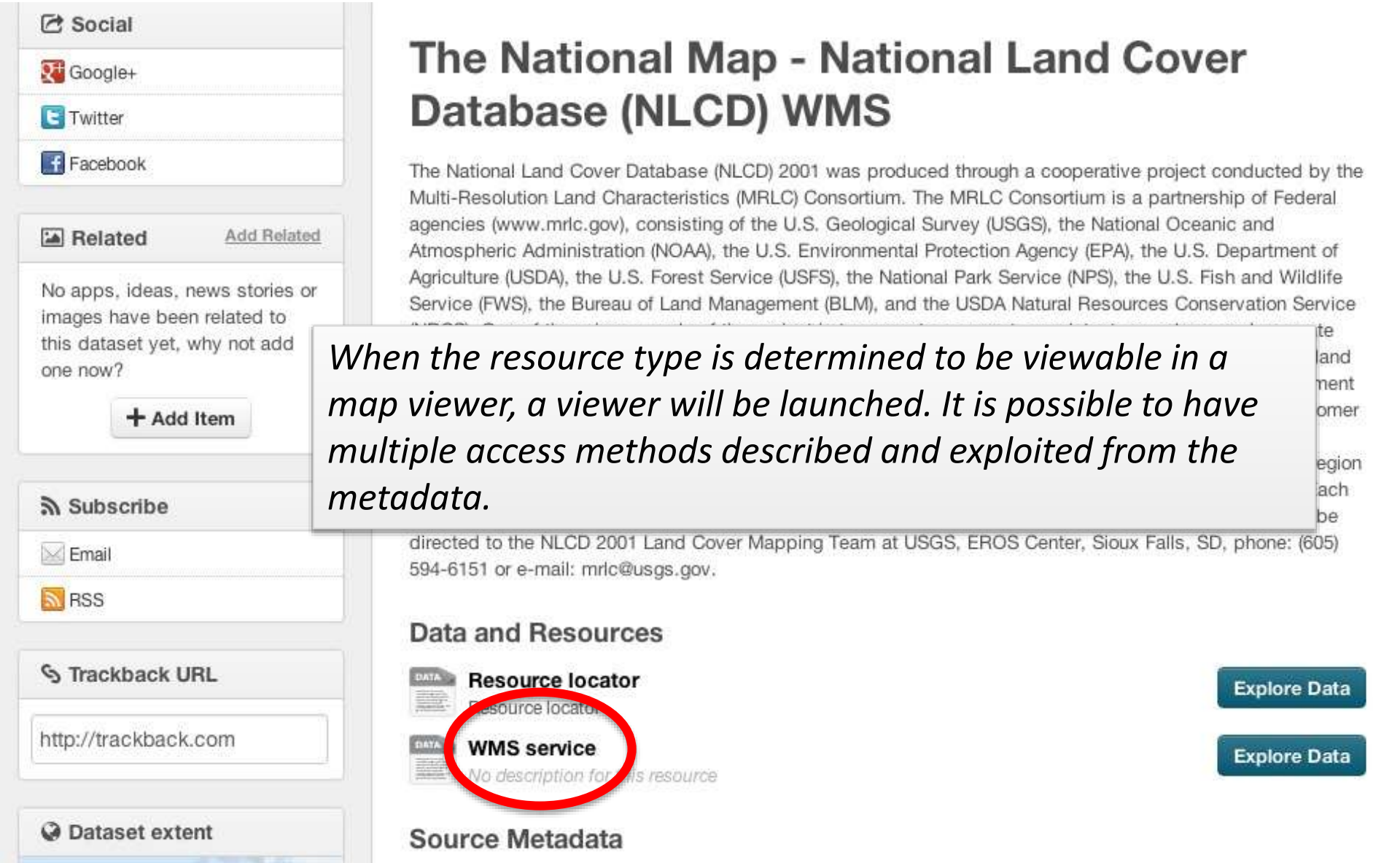
Shuttle Radar Topograp
The Shuttle Radar Topograph
Agency (NGA). Flown aboard

Shuttle Radar Topography Mission: SRTM1N51W177
The Shuttle Radar Topography Mission (SRTM) is a partnership between NASA and the National Geospatial-Intelligence Agency (NGA). Flown aboard the NASA Space Shuttle Endeavour...

Shuttle Radar Topography Mission: SRTM1N51W178
The Shuttle Radar Topography Mission (SRTM) is a partnership between NASA and the National Geospatial-Intelligence Agency (NGA). Flown aboard the NASA Space Shuttle Endeavour...

PLATFORM

fgdc
FederalGeographicDataCommittee


Social
Google+
Twitter
Facebook
Related
Add Related
No apps, ideas, news stories or images have been related to this dataset yet, why not add one now?
+ Add Item
Subscribe
Email
RSS
Trackback URL
http://trackback.com
Dataset extent
The National Map - National Land Cover Database (NLCD) WMS
The National Land Cover Database (NLCD) 2001 was produced through a cooperative project conducted by the Multi-Resolution Land Characteristics (MRLC) Consortium. The MRLC Consortium is a partnership of Federal agencies (www.mrlc.gov), consisting of the U.S. Geological Survey (USGS), the National Oceanic and Atmospheric Administration (NOAA), the U.S. Environmental Protection Agency (EPA), the U.S. Department of Agriculture (USDA), the U.S. Forest Service (USFS), the National Park Service (NPS), the U.S. Fish and Wildlife Service (FWS), the Bureau of Land Management (BLM), and the USDA Natural Resources Conservation Service
When the resource type is determined to be viewable in a map viewer, a viewer will be launched. It is possible to have multiple access methods described and exploited from the metadata.
directed to the NLCD 2001 Land Cover Mapping Team at USGS, EROS Center, Sioux Falls, SD, phone: (605) 594-6151 or e-mail: mrlc@usgs.gov.
Data and Resources
Resource locator
Explore Data
WMS service
No description for this resource
Explore Data
Source Metadata

OpenLayers viewer

https://catalog.data.gov/dataset/usgs-2010-gap-landcover/resource/6ef639f9-5b10-4782-a686-aecd80801fba

eRAS | Quicktime | webforms | ASDI AWS | ASDI Dashboard | IDM UAT | SI SuperInn | WeatherSpark | Leave Form | OBIS Travel

/ Organizations / State of North Dakota /USGS 2010 GAP Landcover /
NDHUB.LANDCLASS_GAP_2010

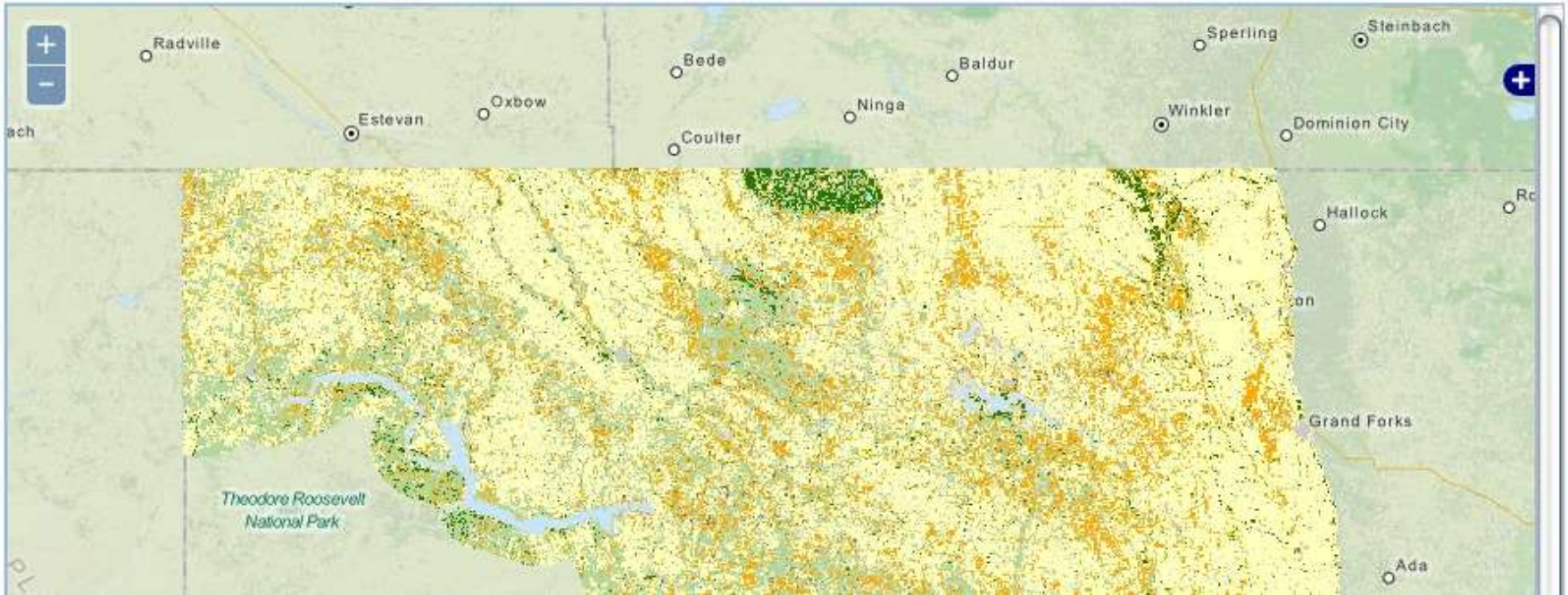
NDHUB.LANDCLASS_GAP_2010

Download

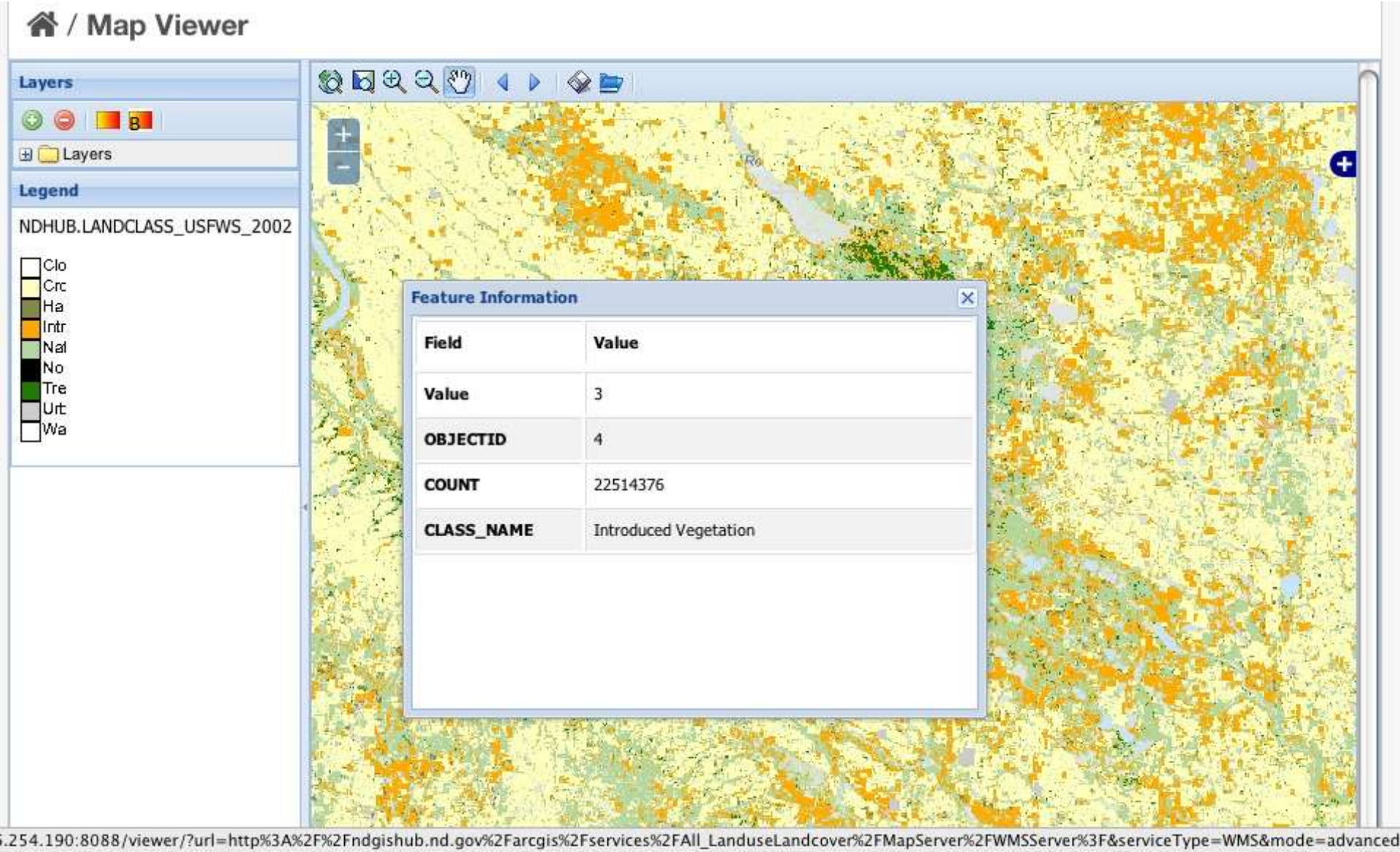
URL: http://ndgishub.nd.gov/arcgis/services/All_LanduseLandcover/MapServer/WMSServer?

USGS 2010 GAP Landcover

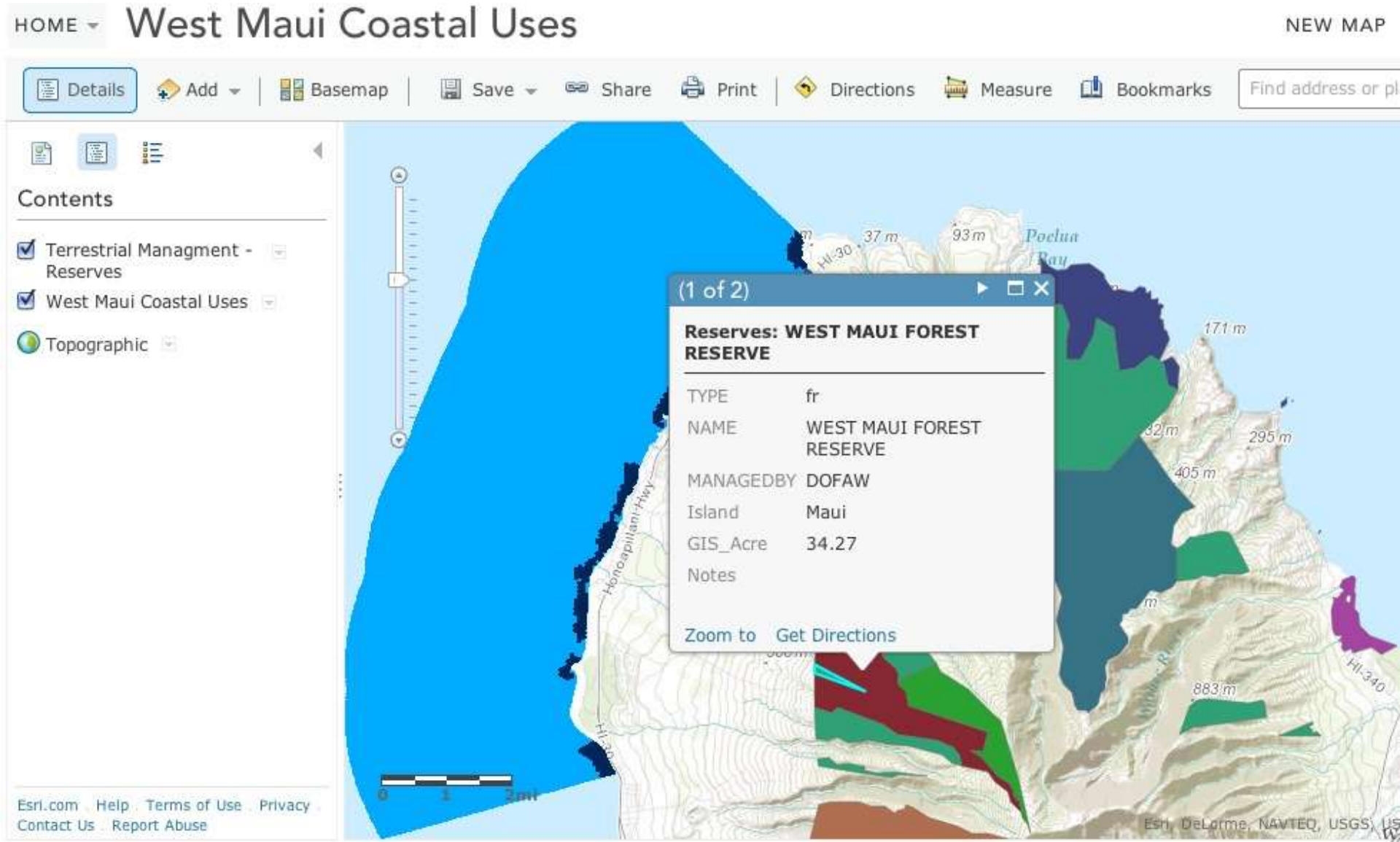
View in Advanced Viewer



Advanced OpenLayers viewer



ArcGIS javascript viewer



Publication process

- FGDC or ISO metadata are prepared by agencies and hosted in a browseable directory (WAF) or in a CSW catalog
- Publisher registers for Admin access
- Access request is reviewed and granted
- Publisher creates “harvest source” descriptions at catalog.data.gov
- Metadata are periodically harvested

GEOSPATIAL PLATFORM

fgdc
Federal Geographic Data Committee

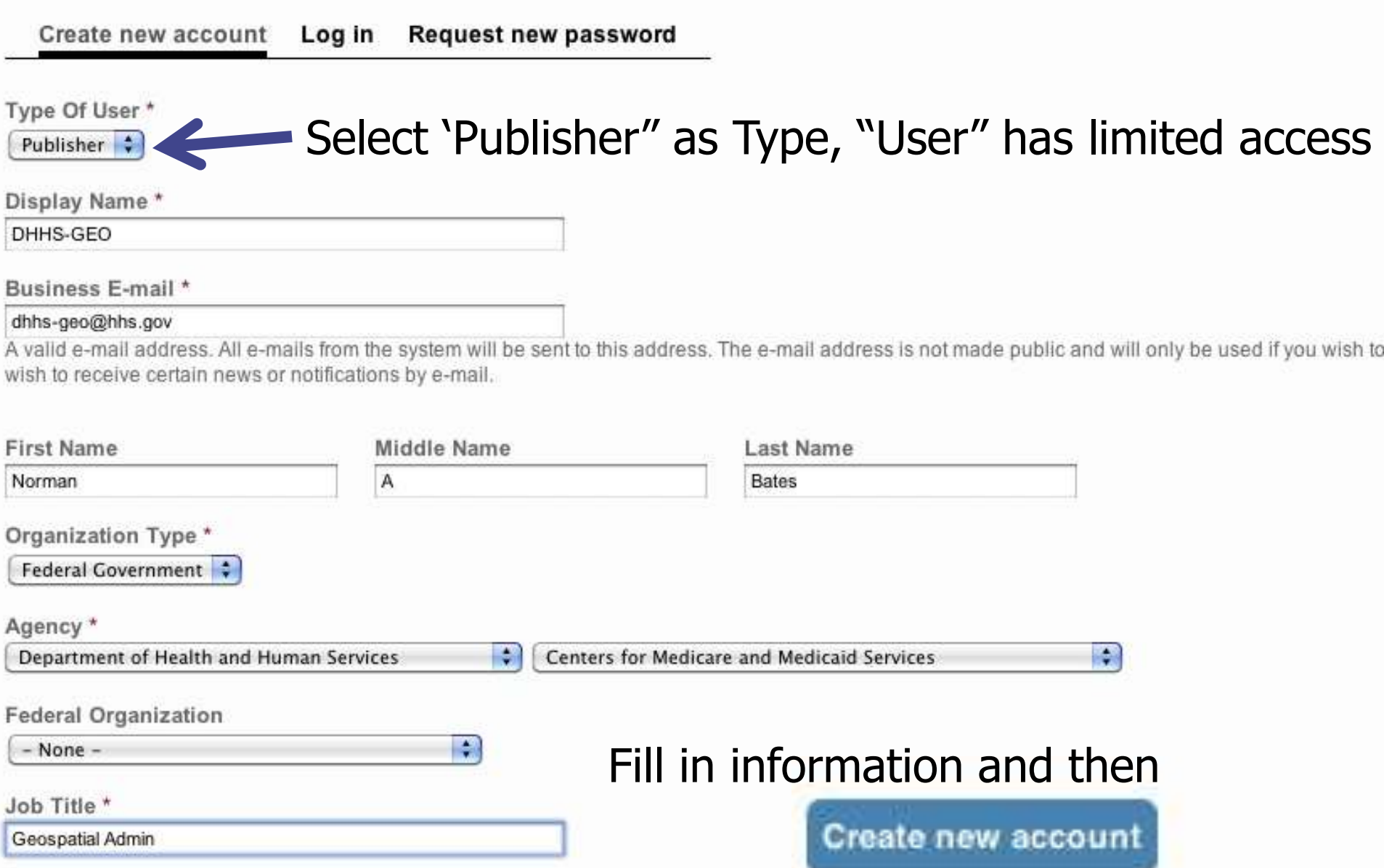
Creating a new Publisher

- If you have received an email invitation from GSA to reset your password, follow the instructions to activate your account - *former geo.data.gov members*
- To create a new account and request publisher role go to the URL:
  - https://idm.data.gov/user/register

GEOSPATIAL PLATFORM

fgdc FederalGeographicDataCommittee

Login screen, create new account



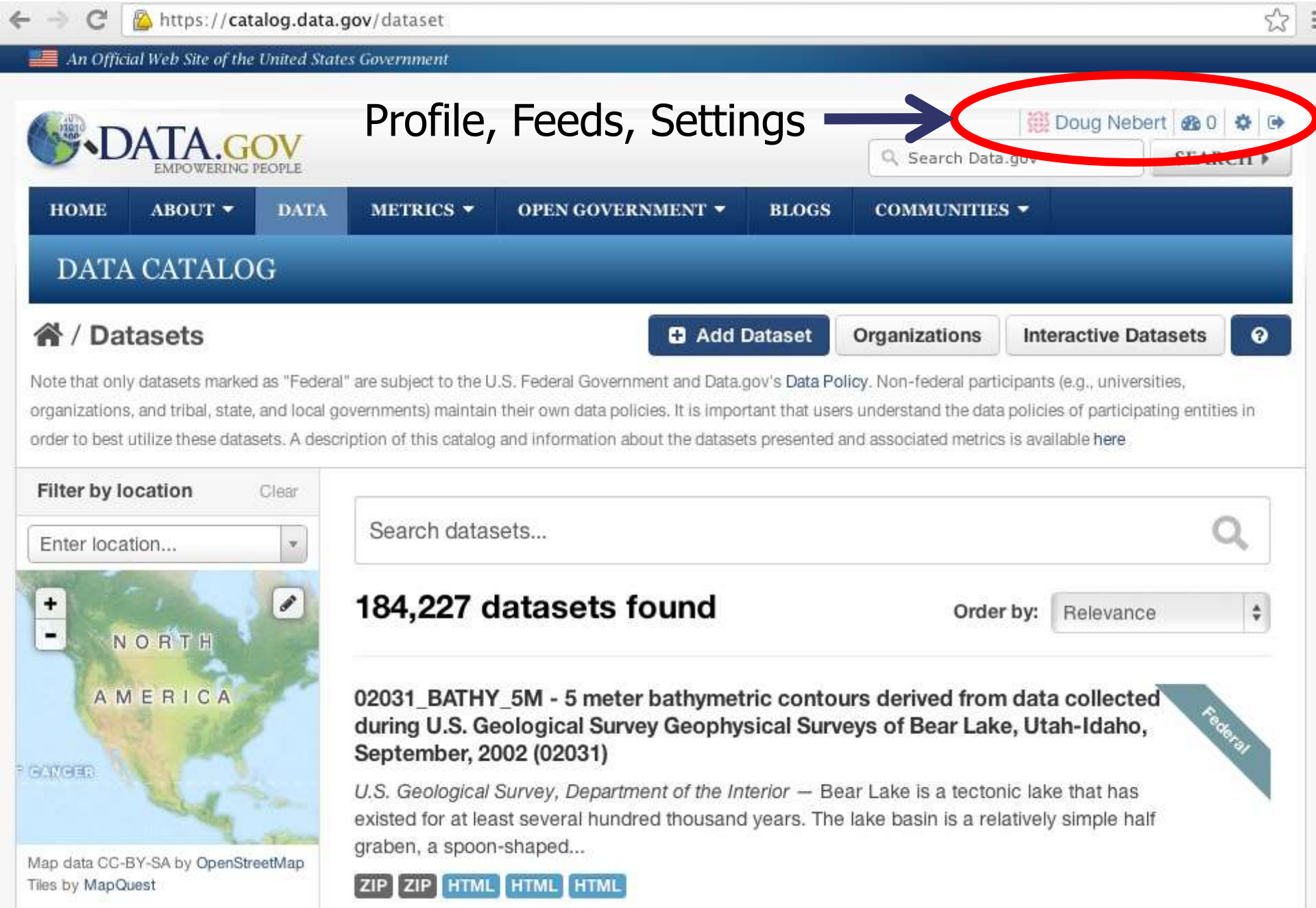
Review and approval

- Usually within a business day, requests for User or Publisher are reviewed and approved and associated with an Organization (domain)
- You will get a confirmation email that the account has been approved. This allows access to data.gov, CKAN, GeoPlatform.gov, and our AGOL backend

GEOSPATIAL PLATFORM

fgdc
FederalGeographicDataCommittee

https://catalog.data.gov/dataset
An Official Web Site of the United States Government
Login with your credentials
Login
DATA.GOV
EMPOWERING PEOPLE
Search Data.gov
SEARCH
HOME
ABOUT
DATA
METRICS
OPEN GOVERNMENT
BLOGS
COMMUNITIES
DATA CATALOG
/ Datasets
Note that only datasets marked as "Federal" are subject to the U.S. Feder
organizations, and tribal, state, and local governments) maintain their ow
order to best utilize these datasets. A description of this catalog and infor
Filter by location
Clear
Enter location...
NORTH
AMERICA
CANCER
Map data CC-BY-SA by OpenStreetMap
Tiles by MapQuest
Search datasets...
184,445 datas
02031_BATHY_5M -
during U.S. Geologic
September, 2002 (02
U.S. Geological Survey,
existed for at least sever
graben, a spoon-shapec
ZIP
ZIP
HTML
HTML
https://idp.data.gov/saml/module.php/core/loginuserpass.php?Au
An Official Web Site of the United States Government
DATA.GOV
EMPOWERING PEOPLE
USER ACCOUNT
Create new account
Log in
Request new password
E-mail *
ddnebert@usgs.gov
Enter your e-mail address.
Password *
Enter the password that accompanies your e-mail.
Login


https://catalog.data.gov/dataset
An Official Web Site of the United States Government
Profile, Feeds, Settings
Doug Nebert
0
DATA.GOV
EMPOWERING PEOPLE
Search Data.gov
HOME
ABOUT
DATA
METRICS
OPEN GOVERNMENT
BLOGS
COMMUNITIES
DATA CATALOG
/ Datasets
Add Dataset
Organizations
Interactive Datasets
Note that only datasets marked as "Federal" are subject to the U.S. Federal Government and Data.gov's Data Policy. Non-federal participants (e.g., universities, organizations, and tribal, state, and local governments) maintain their own data policies. It is important that users understand the data policies of participating entities in order to best utilize these datasets. A description of this catalog and information about the datasets presented and associated metrics is available here
Filter by location
Clear
Enter location...
NORTH
AMERICA
CANCER
Map data CC-BY-SA by OpenStreetMap
Tiles by MapQuest
Search datasets...
184,227 datasets found
Order by:
Relevance
02031_BATHY_5M - 5 meter bathymetric contours derived from data collected during U.S. Geological Survey Geophysical Surveys of Bear Lake, Utah-Idaho, September, 2002 (02031)
Federal
U.S. Geological Survey, Department of the Interior — Bear Lake is a tectonic lake that has existed for at least several hundred thousand years. The lake basin is a relatively simple half graben, a spoon-shaped...
ZIP
ZIP
HTML
HTML
HTML

Set up a harvest

- Go to **Organizations** and search for your organization, or use/save pathname based on your domain name, https://catalog.data.gov/organization/usgs-gov
- Click **Admin** (see/set org details)
- Click **Harvest Sources**
- Click **Add Harvest Source** or edit an existing one

GEOSPATIAL PLATFORM

fgdc
FederalGeographicDataCommittee

https://catalog.data.gov/organization/edit/usgs-gov

An Official Web Site of the United States Government

DATA.GOV
EMPOWERING PEOPLE

Doug Nebert | 0

Search Data.gov | SEARCH

HOME | ABOUT | DATA | METRICS | OPEN GOVERNMENT | BLOGS | COMMUNITIES

DATA CATALOG

/ Organizations / U.S. Geological Survey, Department ... / **Admin**

USGS
science for a changing world

U.S. Geological Survey, Department of the Interior

http://www.usgs.gov/ read more

Edit | Datasets | Members | Harvest Sources | View organization

Title: U.S. Geological Survey, Department of the Interior

URL: catalog.data.gov/organization/usgs-gov Edit

Description: http://www.usgs.gov/

You can use Markdown formatting here

Image URL: http://pubs.usgs.gov/sir/2004/5296/04-5296_confluence_park/usgs_logo.gif

Custom Field: Key: organization_type Value: Federal Government Remove

https://catalog.data.gov/organization/harvest/usgs-gov

An Official Web Site of the United States Government

DATA.GOV
EMPOWERING PEOPLE

Doug Nebert | 0

Search Data.gov | SEARCH

HOME | ABOUT | DATA | METRICS | OPEN GOVERNMENT | BLOGS | COMMUNITIES

DATA CATALOG

/ Organizations / U.S. Geological Survey, Department ... / **Admin**

USGS
science for a changing world

U.S. Geological Survey, Department of the Interior

http://www.usgs.gov/ read more

Edit | Datasets | Members | Harvest Sources | View organization

16 harvest sources

Add Harvest Source

Search sources...

EarthExplorer

There is no description for this harvest source

Datasets: 0

National Atlas of the United States

There is no description for this harvest source

Datasets: 115

National Geologic Map Database

Frequency Clear All

MANUAL (12)

Show More Frequency

Type Clear All

waf (11)

z3950 (3)

waf-collection (2)

/ Harvest Sources / **Create Harvest Source**

Harvest sources

Harvest sources allow importing remote metadata into this catalog. Remote sources can be other catalogs such as other CKAN instances, CSW servers or Web Accessible Folders (WAF) (depending on the actual harvesters enabled for this instance).

URL: http://bison.usgs.ornl.gov/waf

This should include the http:// part of the URL

Title: BISON

URL: catalog.data.gov/harvest/bison Edit

Description: Biodiversity Information Serving Our Nation (BISON) is a product of the U.S. Geological Survey's Core Science Analytics and Synthesis Program. BISON is an information system that allows users to access, explore, and download U.S. species occurrence data from participating data providers.

You can use Markdown formatting here

Source type:

- ○ CSW Server
- ○ Web Accessible Folder (WAF)
- ○ Single spatial metadata document
- ○ Geoportal Server
- ◉ Web Accessible Folder (WAF) Homogeneous Collection
- ○ Z39.50
- ○ ArcGIS REST API

Harvest types

- OGC Catalog Service (CSW 2.0.2) carrying full FGDC or ISO metadata
- Web Accessible Folder – browse enabled public web folder (looks like ftp)
- WAF for homogeneous collections
- Single metadata file as XML (upload)
- Esri Geoportal Server instance
- Z39.50 server
- ArcGIS Online or Portal for ArcGIS

GEOSPATIAL PLATFORM

fgdc
FederalGeographicDataCommittee

WAF for homogeneous collections

- A new feature in CKAN is to allow for large homogeneous collections to not clutter initial results (e.g. TIGER, Landsat, bathy surveys)
- A parent metadata record is identified
- Resource Description URL contains the location of the WAF in which all child records exist
- Parent record is returned with initial results, with option for user to continue search on the collection

GEOSPATIAL PLATFORM

fgdc FederalGeographicDataCommittee

Update frequency: Weekly

Collection Metadata Url: eg. http://example.com/collection.xml

Validation:

- ○ Autodetect
- ○ ISO 19115 Metadata (ISO 19139 XSD)
- ○ FGDC Minimal Validation
- ○ FGDC CSDGM Version 2.0, 1998 (FGDC-STD-001-1998)
- ◉ FGDC CSDGM Biological Data Profile (FGDC-STD-001.1-1999)
- ○ FGDC CSDGM Metadata Profile for Shoreline Data (FGDC-STD-001.2-2001)
- ○ FGDC Extensions for Remote Sensing (FGDC-STD-012-2002)

Dataset visibility: Public

Organization: usgs-gov

Delete Save

Validation

- Metadata are parsed and validated as either CSDGM or ISO. If it passes validation it is imported into the index
- Options:
  - Auto-detect
  - ISO 19115/19139 from NOAA NGDC
  - CSDGM Minimal validation (Xpath tests)
  - CSDGM Version 2 (1998)
  - CSDGM Biological Data Profile, Shoreline Data Profile, Extensions for Remote Sensing Data

GEOSPATIAL PLATFORM

fgdc FederalGeographicDataCommittee

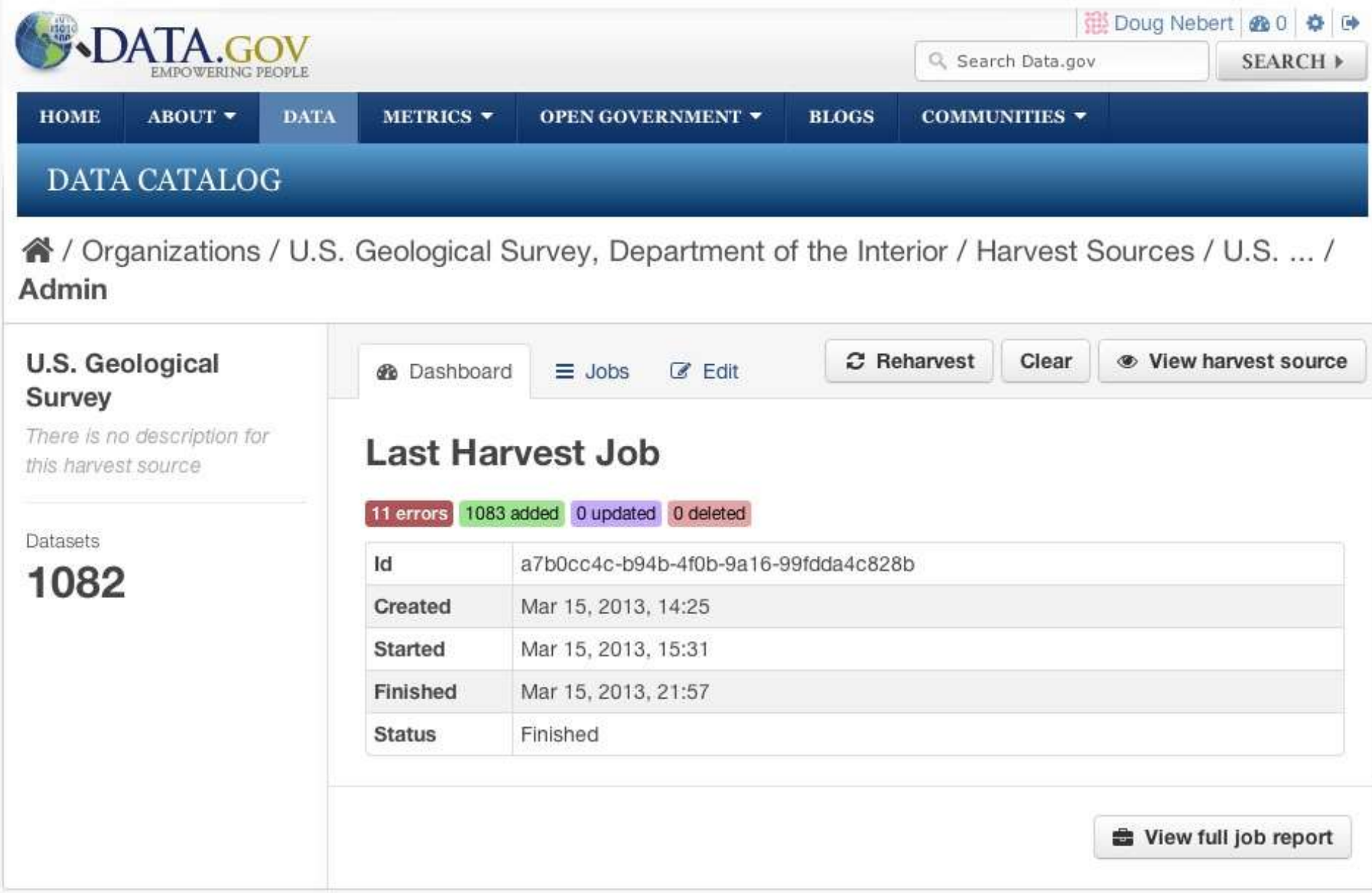
Harvesting

- Harvests are run on a schedule that you set
- Only new and updated records are detected in a WAF – if you need to delete then hit **Clear** to purge and fully reload
- Harvests will be scheduled and should complete within one business day
- An error report is available

GEOSPATIAL PLATFORM

fgdc
FederalGeographicDataCommittee

Harvest report



Error reporting

Document Errors

Count	Message
10	Transformation to ISO failed
9	The transformation service returned an error for object {0}: [409] net.sf.saxon.trans.XPathException: A sequence of more than one item is not allowed as the first argument of contains() ("USGS Pacific Science Center, S...", "U.S. Geological Survey, Coasta...")
1	The transformation service returned an error for object {0}: [409] net.sf.saxon.trans.XPathException: Cannot convert string to double: "4,700"
1	Point extent defined instead of polygon for object b0b9fe1e-671d-4275-8b42-a5d43346e75c

Error Report

Document Errors 11 documents with errors

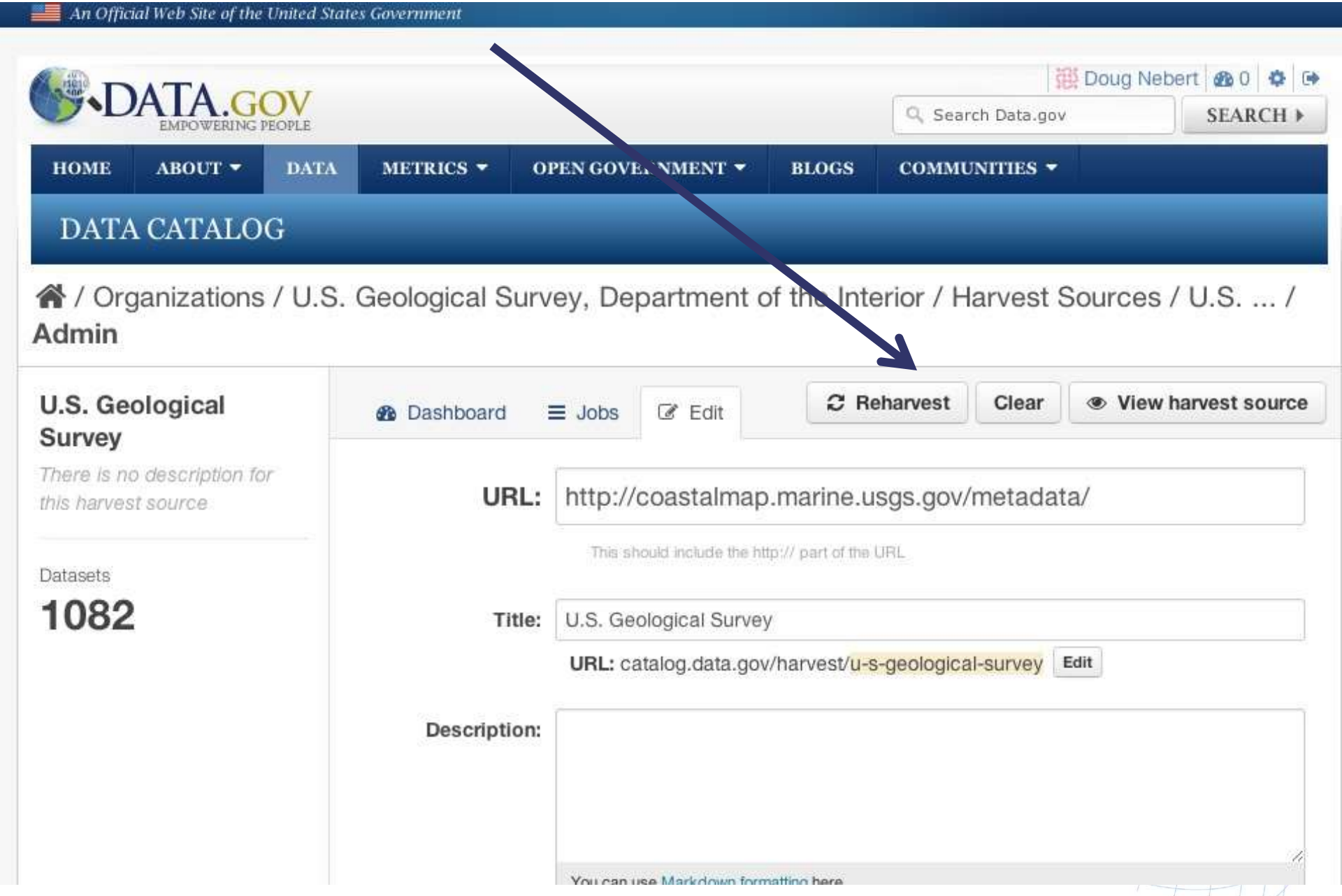
114c4840ba70123873b83405f4a7ca2c63ae76d3 Remote content | Local content

Transformation to ISO failed
The transformation service returned an error for object {0}: [409] net.sf.saxon.trans.XPathException: A sequence of more than one item is not allowed as the first argument of contains() ("USGS Pacific Science Center, S...", "U.S. Geological Survey, Coasta...")

1b372346bbdf9fc3a5b2db7d79aa9e810af4978f Remote content | Local content

Transformation to ISO failed
The transformation service returned an error for object {0}: [409] net.sf.saxon.trans.XPathException: A sequence of more than one item is not allowed as the first argument of contains() ("USGS Pacific Science Center, S...", "U.S. Geological Survey, Coasta...")

Reharvest and Clear



Reharvest and Clear

- If your collection is set to “Manual” or you wish to force a re-harvest of content, then click on **Reharvest**. It will process within a few hours.
- If you wish to flush the contents of the index or have deletions from a prior run, then click on **Clear** first

GEOSPATIAL PLATFORM

fgdc FederalGeographicDataCommittee

Metadata recommendations

- Based on review of the results and in the interest of having metadata be optimally processed by the indexing process, some metadata recommendations were prepared
- These recommendations are summarized on the following pages and will be linked from Resources page of the new geoplatform.gov

GEOSPATIAL PLATFORM

fgdc
Federal Geographic Data Committee

Recommendations

- Provide direct URL to online resources in metadata
- Use the appropriate metadata structure to describe the download or access URL
- Include terms found in published taxonomies in your metadata

GEOSPATIAL PLATFORM

fgdc
FederalGeographicDataCommittee

Recommendations, continued

- Reference browse graphics or thumbnails by URL
- Publish only your original metadata/data
- Publish your metadata only once in data.gov
- Write descriptive titles and abstracts

GEOSPATIAL PLATFORM

fgdc
FederalGeographicDataCommittee

Recommendations, continued

- Collections of homogeneous data need collection-level and member-level metadata records
- Specify content dates
- Use recommended format and protocol codes

GEOSPATIAL PLATFORM

FederalGeographicDataCommittee fgdc

Enhancement requests

- Add a “Clear on Re-harvest” option to the registration/edit page for a given harvest source.
- Add the ability for logged in users of CKAN to list the publishers for their own organization via the UI.
- Support CSW endpoint on catalog.data.gov for external search and syndication.
- Document existing CKAN search API options on the Developers page of Geoplatform.gov.
- Post the minimum ISO and FGDC metadata requirements for a compliant record on geoplatform.gov.
- Allow “Group” affiliation and rights to transfer between Geoplatform and AGOL environments to enable conditional access to records.

GEOSPATIAL PLATFORM

fgdc
FederalGeographicDataCommittee

Questions?

Doug Nebert
FGDC Secretariat
ddnebert@fgdc.gov

GEOSPATIAL PLATFORM

FederalGeographicDataCommittee fgdc