|  |  |
| --- | --- |
|  | USGIN  U.S. Geoscience  Information Network |

Procedures and conventions for deploying web services for the United States Geoscience Information Network

Title:

Procedures and conventions for deploying web services within the United States Geoscience Information Network

This version: 0.5

Creator:

USGIN Specifications Drafting Team

Editors:

Stephen M. Richard, Celia Coleman, Jordan Matti

Creation date:

2011-12-16

Last revision date:

2012-01-03

Document Status:

Draft in development

Publisher:

Arizona Geological Survey

Description:

This document outlines procedures for deploying interoperable Web Map Services (WMS) and Web Feature Services (WFS) for the United States Geoscience Information Network (USGIN).

Contributor:

See acknowledgements

Document Identifier:

http://repository.usgin.org/uri\_gin/usgin/dlio/456

Notices

Neither the USGIN project, nor any of the participating agencies, take any position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither do they represent that there has been any effort to identify any such rights.

This document and the information contained herein is provided on an "AS IS" basis and USGIN DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Revision History

| **Version** | **Date** | **Comments** | **By** |
| --- | --- | --- | --- |
| 0.5 | 2011-12-16 | Draft document prepared by Celia Coleman and Jordan Matti put into USGIN spec template and registered in repository. Review and edit of text | Stephen Richard |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Acknowledgement

Many individuals and organizations have contributed to or inspired the development of these Guidelines.

The USGIN Specifications Drafting Team include (alphabetically):

Stephen M. Richard – Arizona Geological Survey (AZGS)

Celia Coleman – AZGS

Jordan Matti – AZGS

Funding Provided by (chronologically):

National Science Foundation under EAR-0753154 to the Arizona Geological Survey acting on behalf of the Association of American State Geologists; 2009.

US Department of Energy under award DE-EE0001120 to Boise State University; 2010.

DOE award DE-EE1002850 to compile and publish geothermal data from state geological surveys integrated with the NGDS

Contact Information

If you have questions or comments regarding this document, contact:

Stephen Richard or Celia Coleman

Arizona Geological Survey   
416 W. Congress St., Suite 100.  
Tucson, Arizona 85701-1381   
Phone: 520.770.3500   
Fax: 520.770.3505

E-mail

[metadata@usgin.org](mailto:metadata@usgin.org)

Online

[http://usgin.org](http://usgin.org/)   
[http://lab.usgin.org](http://lab.usgin.org/)

Table of Contents

[1 Introduction 6](#_Toc311963727)

[1.1 System Technical Design principles 6](#_Toc311963728)

[1.2 Data Access 7](#_Toc311963729)

[1.3 USGIN Tools 8](#_Toc311963730)

[1.4 Document Structure 8](#_Toc311963731)

[2 Identifiers 9](#_Toc311963732)

[3 USGIN Service Profiles 10](#_Toc311963733)

[3.1 Metadata Profile 10](#_Toc311963734)

[3.2 Catalog Service 10](#_Toc311963735)

[3.3 Data Services 10](#_Toc311963736)

[3.3.1 Legacy data and data products 10](#_Toc311963737)

[3.3.2 Map services 10](#_Toc311963738)

[3.3.3 Feature services 10](#_Toc311963739)

[3.3.4 Grid services 10](#_Toc311963740)

[4 Implementation and Deployment 11](#_Toc311963741)

[4.1 OGC Services 11](#_Toc311963742)

[4.1.1 Map service 12](#_Toc311963743)

[4.1.2 Feature service 13](#_Toc311963744)

[4.2 XML encoding 13](#_Toc311963745)

[4.2.1 XML Schema Specifications 13](#_Toc311963746)

[4.2.2 Section 5.2: XML Schema Validation 13](#_Toc311963747)

[4.2.3 Section 5.3: Validation Standards 14](#_Toc311963748)

[5 Resource registration 15](#_Toc311963749)

[6 Licensing 16](#_Toc311963750)

[7 Hardware and Hosting 17](#_Toc311963751)

Tables and Figures

# Introduction

The United States Geoscience Information Network (USGIN )is a partnership between the Association of American State Geologists (AASG) and the US Geological Survey (USGS) working to make geoscience information easier to find, distribute, and analyze. Intended users include government agencies (regulatory, land management), commercial users (mineral exploration, engineering, environmental), researchers, and educators. The network consists of three major components:

* A collection of service definitions, interchange formats, and vocabularies
* Catalogs that lead users to resources
* A community of users – data providers and consumers

The USGIN initiative is initiating and promoting community development of these components. We recognize that the cornerstone to adoption of community specifications and practices is the demonstration of their utility, and documentation in sufficient detail and clarity that new adopters can understand the specifications, their implementation requirements. This document is a contribution to documentation of specifications and practices developed or adopted by USGIN and recommended for community use.

A basic USGIN design philosophy is that software components for the network are to the maximum extent possible based on open-source software, existing standards, and World Wide Web infrastructure and browsers. Existing or emerging standards and technology should be used whenever possible rather than developing new ones. Data should be published using schema and encoding that are well documented in the public domain so that anyone may utilize that data. Applications should be developed with attention to their portability and reusability. This philosophy promotes system architecture based on existing well documented, widely used specifications, decoupled components with well defined interfaces, and attention to documentation of software resources developed for network use. The intention is that anyone can access data published by the USGIN, and data that is published and registered in the catalog system according to USGIN specifications is a network resource.

The goal of USGIN service-based information distribution is interoperability, defined as the capability to communicate, execute programs, or transfer data in a manner that requires the user to have little or no specialized knowledge. In a service context, the operational test for interoperability is whether a software client designed to use a particular service can operate without problems connected to different servers that offer that service. Interoperability requires consistent understanding and implementation of service operation, request parameters, interchange formats for messaging between client and server, and content semantics to the level proscribed by the service specification and profile in use. This document outlines procedures and conventions for deploying interoperable USGIN web services.

We anticipate that as the network evolves, service deployment procedures will change with the introduction of new service protocols and new hardware and operating system environments. This document is constructed in the context of WMS versions up to 1.3.0, WFS up to 1.1.1, CSW up to 2.0.2, mostly targeted to Windows 2008 IIS or Linux Apache (Ubuntu) server environments. Subsequent editions of these guidelines will be necessary as content models, encoding schemes, computational architecture, and use cases evolve.

## System Technical Design principles

USGIN is intended to facilitate locating, obtaining, assessing, and utilizing information resources. The term resource is used in a very general sense to mean any identifiable item of interest. This includes, but is not limited to, documents, datasets, services, functional units, and physical objects. Resources are made available by a community of data providers.

Resources become USGIN resources by creating a metadata record that describes the resource according to a shared content model and inserting that metadata into the catalog system. The metadata description can be indexed for discovery by search engines, provides information about resource provenance and quality so users can evaluate the resource for their application, and provides information describing how to access the resource. The access instructions should be in a format that can be utilized by software clients to automate the access process and minimize the amount of user interaction required to bring the resource to their desktop.

Any search client that implements the USGIN catalog service profile should be able to conduct search against any system catalog that also implements the profile. This means that there can be multiple portals and client applications for accessing system resources; it requires that a single client can search different catalogs in the system without the user having to reconfigure the software.

Structured data are provided through USGIN services that have published protocol and documented interchange formats. A service profile defines a particular base service protocol (e.g. OGC CSW2.0.2, WFS 1.1.0, WMS1.3.0), information interchange format, and other parameter and usage conventions necessary to enable interoperability. Multiple data providers that present information using that service profile should be accessible to any client that implements that service profile with minimum operator intervention.

## Data Access

One of the basic objectives of USGIN is to simplify access to and utilization of geoscience data. For document based resources and datasets that do not have specifications for interchange protocols, data schema, and file format, data access simply involves creating metadata to describe the resource, providing the metadata to a network catalog, and making the resource available in a web-accessible location linked to from the metadata. The data access service is http and standard web file access processes.

Map-based portrayals of spatially referenced information can be made available as documents, or as web map service layers. Metadata must be provided to a network catalog that describes the map portrayal source information, and the options for obtaining the map in various distributions—hard copy, as a standalone file, or using a web map service. Delivery of data in pre-packaged documents or files, or as map images through a map service may be thought of as providing ‘data-products’. These are bundles of information assembled by some authority (the author, compiler, editor, etc.) that has processed, integrated, and possibly interpreted raw data in some way to add value. These data products are primarily designed and intended for human users.

For data that are useful to access in a more granular fashion as individual records of observations, features, or data point delivery of data through a web service is warranted. Web services provide access that is tailored for consumption by software processes that can then access data for processing to generate new views, interpretation, mash-ups, or workflows to generate new knowledge and data products. Distribution of data through a service requires design and implementation of an interchange format for delivering that data, and a network service interface that defines the messaging protocol to request data. Metadata describing the dataset and service function must be provided to a network catalog. This approach to data distribution using community interchange formats requires that data providers invest effort to do the data integration work necessary to map their data into network interchange formats. This data integration involves matching field or element names in the schema for source and target data sets, selecting those that contain the information of interest, and then merging content into a single data set with consistent usage of vocabulary and units of measure in a standardized collection of fields or elements.

Data integration may be done by data providers who choose to deliver data in standardized interchange formats, by data consumers who acquire data in heterogeneous formats and schema and figure out how to extract what they need, or data integration may be done by middleware layers that implement transformations between known formats and schema (brokers). The path adopted for the Geoscience Information Network to simplify data access and promote reuse is to develop standard formats and access protocols used to deliver common data sets (e.g. borehole temperature data, heat flow measurements) to consumers. The onus of data maintenance is shifted towards organizations that are tasked with data management and preservation. By documenting data schema, encoding formats and practices for vocabulary usage, data can be put into the ‘data integration’ format when it is made available on the web. Because of its enhanced utility in a standardized format, management and preservation of the data are more strongly motivated. Mapping data into an interchange format is likely to be done more accurately by those who originate the data working in conjunction with data managers who understand the interchange formats. The net effect is a greater likelihood that the federated information system using the documented interchange formats will outlast any particular researcher, data provider, project, or agency. HTML on HTTP, NetCDF, and XML are examples of data integration formats that have achieved wide usage and long term usefulness.

Publication in a data interchange format introduces additional costs into the data delivery process, and this cost dictates that there must be consideration of the benefits obtained. Use of the community interchange formats for data publication requires only one transformation from internal to interchange format to be engineered for each interchange format in use. This may be done at the data origination point, or through broker services. The use of schema and encoding specifically designed for data integration and interchange means data producers and consumers can continue to use internal data formats that are optimized for their business requirements. It is important to remember that the requirements for efficient data storage and maintenance are different from the requirements for client-friendly data delivery. Data providers will need to develop policies to determine what resources should be presented in the data integration format (along with any other formats that the data publisher wants to use), and what data are specialized to a degree that data integration by the providers is not warranted and delivery as packaged data products is sufficient. Criteria for such decisions will likely include how often that kind of data are known or expected to be used, the cost of obtaining or reproducing the data, and the expected useful lifetime of the data.

## USGIN Tools

The US Geoscience information network development effort to data has focused on mechanisms for data registration and discovery, and the development of interchange formats that provide a good balance between expressive richness and complexity. Web-accessible USGIN tools include the following:

The USGIN Portal (<http://usgin.org/>) provides information resources for a broad overview of USGIN, as well as tutorials, a glossary, and links to other USGIN websites.

The USGIN Metadata Wizard (<http://mw.usgin.org>) provides a tool to register online resources by creating metadata records that will be harvested into the USGIN Metadata Catalog. The metadata records are the primary mechanism to support search, discovery, evaluation and access to online resources.

The USGIN Metadata Catalog (<http://catalog.usgin.org/search>) is a web application that supports searching the collection of metadata records to find information network resources.

The USGIN Laboratory (<http://lab.usgin.org/>) is a web site intended to support community development of service profiles, interchange formats, and implementations that are the essence of the USGIN. The site is a portal to detailed information about USGIN specifications and discussion boards for dialogue and problem-solving.

## Document Structure

(this is a working outline—update or remove for final draft)

Introduction (this section) provides background on USGIN

Use of specifications and standards

WMS

Map service interoperability

Styled layer descriptors

Deployment platforms

WFS

Simple vs complex features

Content models

Encoding schemes

Versioning

Services

Content models

Schema

CSW

Gridded data services

Naming conventions

Validation

Service metadata content—getCapabilities documents

# Identifiers

The ability to identify resources and use those identifiers to locate useful representations of the resource is a key functional requirement for binding metadata descriptions to the described resource, locating resources, and for asserting relationships between resources. In the World Wide Web, identifiers are bit streams, and these are most commonly alphanumeric strings that conform to the IETF specification for uniform resource identifier (URI) syntax (IETF RFC3986). For the meaning of the identifier to be discovered, there must be a dereferencing mechanism that yields a description of the identified thing in a form useful to the requesting agent. This requirement has led to the emergence of URI schemes that use the http-protocol for dereferencing, with syntax specified by the more specific scheme defined for use with hyper-text transfer protocol (IETF RFC2616, based on now superseded RFC2396 URI syntax). This is because the world-wide web infrastructure that is already in-place, well tested, and widely deployed can be used to dereference http URIs to electronically accessible representations of resources.

USGIN has developed an http URI scheme with syntax summarized as:

ginURI = "http:" "//" uriHost “/” URIscheme “/” nameAuthority “/” resourcePath [“/” resourceSpecificString] [( “/” / “/” representationPart )]

The reader is directed to *USGIN URI Policies* (<http://repository.usgin.org/uri_gin/usgin/dlio/331>), for an in-depth discussion of requirements and more complete specification of the syntax and intention.

Because URIs are meant to be stable and reliable, whereas the web-locations where representations for identified resources are retrieved are subject to change, most operational identification systems introduce a level of indirection between the http URI and the URL that retrieves a representation of the identified resource. A widely used example is the Handle System (<http://www.handle.net/>). Two handle system implementations that are of particular interest for USGIN are the Digital Object Identifier (DOI®) System (<http://www.doi.org/factsheets/DOIIdentifierSpecs.html>) and International Geo Sample Number (IGSN, <http://www.igsn.org/>). USGIN has deployed a similar URI dereferencing service for identifiers based on the USGIN URI scheme at <http://resources.usgin.org/>.

USGIN recommends that all digital resources (information resources) be assigned Digital Object Identifiers (DOI). Resources that are uploaded to a USGIN repository will be assigned an identifier by that repository, which can be used if a DOI is not assigned. Physical specimens of Earth Material should be assigned IGSN identifiers. The AASG and USGS are developing an IGSN profile for use in the USGIN community, but this is not yet available. This profile will include a definition of the minimum descriptive metadata that must accompany a sample registration.

Any USGIN participant may utilize the USGIN URI dereferencing service. The procedure is as follows (see USGIN URI minting and dereferencing tutorial for more in-depth guidance).

1. Request registration of a naming authority abbreviation. This is a 3- to 6-character string (token) that uniquely identifies the organization minting the URI, and becomes the nameAuthority part of URIs. The role of the name authority token is to identify a subset namespace of the uri-gin namespace that is managed by that authority.
2. URIs are generated in the name authority space by determining a resourcePath URI part for each type of resource to be identified. The conceptual level of typing is analogous to FRBR work or expression (see Tillett, 2004-02).
3. Determined the pattern used to generate resourceSpecificStrings that identify instances of a particular resource type. If resources have hierarchical part-whole internal structure, it may be useful to reflect this structure in a path-like syntax, e.g. book/chapter/section/paragraph. This segment identifies a particular instance of the resource type specified by the resourcePath. If the target resource is versioned, this identifier segment should include versioning information. Versioning at this level should in general be related to the intellectual content of the resource.
4. If particular representations (FRBR manifestations) are to be identified, rules or patterns for generating the representationPart should be defined. Versioning may be specified at this level as well, but in general will have to do with versions of file formats or data structures, not the intellectual or information content.

The key aspect of defining the URI syntax is that it must be possible using regular expressions to define the mapping between the USGIN URI and a URL that will retrieve the desired representation. Different rules for the URI to URL mapping in the dereferencing service may be defined for different URI patterns, as long as regular expressions can be used to assign any incoming URI to the mapping rule to redirect to the correct URL. The USGIN dereferencing service can be configured to accommodate content negotiation with the URI GET request, and to map different representationParts to different content types.

1. Define the regular expressions and mapping rules to map from the naming authority’s URIs to appropriate URLs and content types using the administration interface at resources.usgin.org.

# USGIN Service Profiles

For datasets published via web services, USGIN is currently advocating use of the OpenGeospatial Consortium suite of services, including the Web Map Service (WMS), Web Feature Service (WFS), and for gridded (raster) datasets, the NCAR OpenDAP protocol. Adoption of these particular service protocols is based on the availability of client and server applications, demonstrated utility based on existing services. We recognize that other date encoding schemes and services are constantly being developed and tested, and that a successful USGIN will evolve to utilize new technology as it becomes proven and available. This is meant to be a production system, true cyberinfrastructure, not an experimental prototype. This chapter specifies the core metadata content model, basic service protocols adopted for initial USGIN services. Discussion of implementation of service instances and XML encoding schemes is included in the following chapter.

## Metadata Profile

The USGIN catalog is a collection of metadata records describing resources hosted by data providers. The collection is unified by a shared content model, protocols for searching, and conventions for encoding metadata content. This collection of metadata records is a cloud resource; it is not necessarily housed in any single location. Because the metadata interchange format and content are standardized, metadata can be harvested between any repositories in the network. We anticipate a variety of aggregating servers that may specialize in particular subdomains, offer community annotations or other social networking functions, or offer value added searching for example semantic searches. This section documents the metadata content specification and the service protocol for searching metadata. Encoding of metadata in interchange document is documented in section 4.3 Metadata encoding.

The distinction between data and metadata is tricky. For the purposes of USGIN, we consider metadata first and foremost a finding aid to locate and evaluate resources for use. Metadata records should be created and provided to a network catalog for any resource that is meant to be accessible individually. Individual documents require one metadata record per document (FRBR ‘work’ or ‘expression’, Tillman 2004). 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. These documents might be scans of well logs, scanned reports or publications, or data in a spreadsheet, such as an Excel file.

Datasets include internal record level source information, documenting details of observation or measurement procedure and other information specific to a particular data type. This includes information such as location, data and time of observations, and the source of the data. These metadata are delivered with the data, and only summarized in the dataset metadata that are published to the USGIN catalog.

The required metadata content is documented in ‘

Providing quality information to evaluate system resources requires criteria that can be used to filter data and categorize them according to established and user-defined quality levels. These quality filters will vary depending on the type of data and their targeted use.

Standard measures of "quality" should be available. E.G. variability, bias, systematic error, imprecision, accuracy, precision, reproducibility, etc.

## Catalog Service

CSW

OpenSearch

OAIPMH

## Data Services

The data can be hosted with WMS or WFS capabilities

Aspects of service deployment include naming conventions for feature types and services, validation of feature type XML schema, feature type schema versioning, content of OGC capabilities documents, and use of OGC Styled Layer Descriptor (SLD) documents.

All data will credit the original intellectual source and host server of record for that data.

Information in system must be supported by metadata to document authority and to provide people and projects that compile data the appropriate level of recognition and support

### Legacy data and data products

Existing file and document based datasets or other information resources. Make web accessible, and get metadata in catalog.

For data that are not provided using documented interchange formats, detailed metadata describing the schema and encoding of the data will be necessary to enable reuse.

### Map services

Web services must conform to OGC Specifications. Standards can be found at:

<http://www.opengeospatial.org/standards>

### Feature services

#### Content models

### Grid services

# Implementation and Deployment

This document provides an outline of the specifications for setting up a WMS/WFS in accordance with USGIN specifications. This document therefore assumes that the reader possesses the intent to serve their data as a WMS or WFS in accordance with USGIN specifications:

Furthermore, this document assumes that the reader will serve their data in a web-accessible directory on an appropriately configured server.

## OGC Services

Data originators and providers must be credited in the following locations:

Map Document Properties

Data Frame Properties

Web services must be named in accordance with USGIN conventions

USGIN naming conventions forthcoming

URIs to identify service Type (table)

Data Frame and Layer Name fields[[1]](#footnote-2) must be web-safe and contain no spaces. Web-safe characters include:

A-Z

a-z

0-9

\_ (underscore)

- (dash)

The Layer Name field should indicate the type of resource[[2]](#footnote-3) described by the data being hosted. The following is a list of USGIN-approved layer names (indented lists indicate sub-categories):

ActiveFault

aqSpring/aqWellChemistry

WaterQuality

CommonAnalytes

BaseMetals

MajorDissolvedConstituents

MinorDissolvedConstituents

Nitrogen

WaterDissolvedGas

FreeGas

WaterIsotopes

GasIsotopes

IsotopesDissolved

BoreholeTemperature

DrillStemTest

ThermalConductivity

ThermalSpring

VolcanicVent

Wellheader

The Data Frame field is a concatenation of two strings[[3]](#footnote-4):

The contents of the Layer Name field, but plural

The official abbreviation of the state of origin of the resources described by the data hosted as a web service

Examples of entries in the Data Frame field are as follows:

“AZBoreholeTemperatures”

“CAActiveFaults

“MAaqSpringChemistry”

…etc.

### Map service

WMS Capabilities Document

Layer Properties

Point data should be displayed using appropriate symbology. USGIN-approved symbology can be found at the following location:

<http://repository.usgin.org/uri_gin/usgin/dlio/206>

The Service Description field should contain a description of the both the service and all layers contained therein.

WMS Capabilities configuration should include an online resource

### Feature service

WFS Capabilities Document

WFS Capabilities configuration should contain the “aasg” prefix[[4]](#footnote-5); prefix should indicate a namespace defined in the following format:

<http://host/uri-gin/authority/xmlschema/resource_specific/version_number>

An example is as follows:

<http://stategeothermaldata.org/uri-gin/aasg/xmlschema/wellheader/1.5>

For more information, see the USGIN URI tutorial (http://usgin.org/content/usgin-uri-tutorial)

## XML encoding

The USGIN initiative uses Extensible Markup Language (XML) as a software client-neutral data interchange format. For basic information about XML, see the USGIN XML tutorial:

<http://usgin.org/content/xml-tutorial>

This section provides specifications for the XML documents employed by USGIN and the NGDS.

### XML Schema Specifications

For practical purposes, a schema is a content model for data entry in a database: schemas indicate where and how data should be entered into a database.

Schema validation is the process by which data that has been entered into a database in accordance with a specific schema is compared with the rules for that schema. Schemas are valid when data is entered in accordance with the rules of a given schema.

USGIN schemas can be found at the following site:

<http://schemas.usgin.org/schemas/>

### Section 5.2: XML Schema Validation

To validate the schema of a web service, it is necessary to find the schema according to which web service data has been entered.

The schema for a given web service can be found by making a DescribeFeatureType request for the web service. To make a DescribeFeatureType request, enter the URL for your web service into a web browser, and append a DescribeFeatureType request onto the URL. Below is a sample web service URL; the DescribeFeatureType request is highlighted in red:

<http://services.azgs.az.gov/arcgis/services/aasggeothermal/AZWellHeaders/MapServer/WFSServer?request=DescribeFeatureType%26version=1.1.0%26typename=Wellheader>

To check the data in the database against a given schema, do a GetFeature request for the web service. This will prompt the service to return database entries, which can be checked against a schema. Below is a sample web service URL; the GetFeature request is highlighted in red.

[http://services.azgs.az.gov/arcgis/services/aasggeothermal/CAWellHeaders/MapServer/WFSServer?service=WFS&request=GetFeature&typeName=Wellheader](http://services.azgs.az.gov/arcgis/services/aasggeothermal/CAWellHeaders/MapServer/WFSServer?service=WFS&request=GetFeature&typeName=Wellheader&maxFeatures=1)

Checking the validity of a web service schema against the service capabilities document can be done manually or with software such as XMLspy

### Validation Standards

The service schema and the NGDS schema should be an exact match for the purposes of interoperability.

## Metadata encoding

Cross walk ISO-csw:record-atom

Although full USGIN metadata must be made available in ISO19139 XML encoding (see USGIN ISO metadata profile), the OGC CSW spec mandates that any CSW server offer a set of ‘core queryable’ and ‘core deliverable’ elements in a csw:record XML schema. The ESRI Geoportal and GeoNetwork OpenSource catalog service implementations both also offer Atom feed, and the ESIP discovery cluster and OGC web service context SWF are also pursuing metadata schemes utilizing Atom. Table xx provides a mapping between the recommended content for USGIN metadata (USGIN metadata content recommendations), ISO19139, csw:record, and an atom:entry xml encodings.

# Resource registration

### Publication of metadata records

Place ISO19139 XML documents for each record in Web accessible folder that is a harvest target

Exposure in a network catalog that may be harvested with ISO19139 XML records. A variety of harvest protocols are in use. CSW—typically a get records request, Harvest operation does not seem to get implemented. OAI-PMH, have to offer USGIN ISO19139 format. Geonetwork has harvest protocol.

Transformation from other formats—xslt from FGDC XML, GetCapabilities to ISO19139. CSV generation format used by Metadata template for AASG geothermal data; other similar approaches from flat file formats possible but would have to be engineered by metadata creater.

We sites allowing creation of records and ‘publish’ to catalog. (may use any of above approaches to get created record into catalog)

# Licensing

Contributors can require user consent to license conditions on data (e.g. noncommercial use only)

Open source and open accessibility is preferred to encourage third parties to independently develop software applications that can use the content and services provided by the system

Implement access controls and security to limit access to datasets at discretion of provider

Data owner retains control of access to all data regardless of where it is stored.

# Hardware and Hosting

Server and Hosting specifications and configuration

There are no specifications for server hardware and hosting at this time.

Windows

Linux

Physical servers

Cloud options

# References

Booth, David, 2003-01-28, Four uses of a URL: Name, Concept, Web Location and Document instance: accessed at <http://www.w3.org/2002/11/dbooth-names/dbooth-names_clean.htm> (2010-02-15).

Coyle, Karen, 2010, Understanding the Semantic Web: Bibliographic Data and Metadata: Library Technology Reports, v.46, American Library Association, 31 p., ISSN 0024-2586

Duerst, M., and Suignard, M., January 2005, Internationalized Resource Identifiers (IRIs): IETF RFC-3987, accessed at <http://www.ietf.org/rfc/rfc3987.txt> .

Fielding, R., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and Berners-Lee, T., 1999-06, Hypertext Transfer Protocol -- HTTP/1.1: RFC2616, accessed at <http://tools.ietf.org/html/rfc2616> 2010-02-19.

Gregorio, J.C., and deHora, B., editors, 2007, The Atom Publishing Protocol: IETF RFC 5023.

ISO19115 WG, in prep, ISO 19115-1 Geographic information — Metadata – Part 1: Fundamentals: ISO TC211 N3071

ISO19119, 2005-02-15, Geographic Information—Services.

ISO19119:2005/Amd.1, 2008-05-01, Geographic information — Services, AMENDMENT 1: Extensions of the service metadata model.

Le Hors, A., Raggett, D., and I. Jacobs, December 1999, HTML 4.01 Specification, W3C Recommendation REC-html401-19991224, <http://www.w3.org/TR/1999/REC-html401-19991224>. Latest version available at <http://www.w3.org/TR/html401>.

Mendelsohn, Noah, and Williams, Stuart, 2007-01-02, The use of Metadata in URIs: W3C TAG Finding, accessed at <http://www.w3.org/2001/tag/doc/metaDataInURI-31-20070102.html>.

Nottingham, M., IETF, 2010, Web Linking: IETF RFC 5988, ISSN:2070-1721.

Nottingham, M., and Sayre, R., eds., 2005, The Atom Syndication Format: IETF Network Working Group Request for Comments RFC-4287, accessed at <http://www.ietf.org/rfc/rfc4287.txt> .

Starr, Joan (head of working group) and DataCite Metadata Working Group, 2011-07, DataCite Metadata Schema for the Publication and Citation of Research Data, version 2.2: DOI:10.5438/0005, accessed at <http://schema.datacite.org/meta/kernel-2.2/doc/DataCite-MetadataKernel_v2.2.pdf> 2011-09-30.

Tillett, Barbara, 2004-02, What is FRBR? A Conceptual Model for the Bibliographic Universe: Library of Congress, Cataloging Distribution Service, 8 pages, accessed at: <http://www.loc.gov/cds/FRBR.html> 2011-10-09.

1. Here, **field** is used within the context of a **database**, as in a **database field**. In this case, the **Data Frame** and **Layer Name** fields are contained within the web service **capabilities document**. A web service created in accordance with USGIN specifications will generate an XML document displaying the contents of the associated capabilities document. [↑](#footnote-ref-2)
2. Here, **resource** means, “an identifiable thing that fulfills a requirement.” Though almost anything can be a resource, in the context of USGIN, resources are usually geophysical: faults, wells, volcanic vents, etc. [↑](#footnote-ref-3)
3. A string is simply a series of letters and numbers; a *concatenation* of strings is two strings put together [↑](#footnote-ref-4)
4. In XML, a prefix is often used to represent a namespace; for more information, see the USGIN XML Tutorial at: http://usgin.org/content/xml-tutorial [↑](#footnote-ref-5)