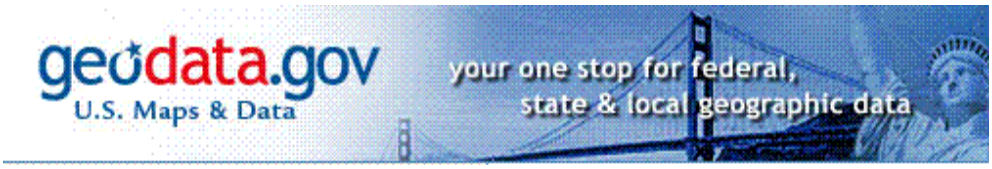


Publishing Metadata to the Geospatial One-Stop Operational Portal – Final

Department of the Interior
Geospatial One-Stop
October 2005

Prepared for:

Department of the Interior
Geospatial One-Stop
12201 Sunrise Valley Drive
Mail Stop 590
Reston, VA 20192

Prepared by:

ESRI
380 New York Street
Redlands, California 92373-8100

Document History

Version	Date	Description	Person Responsible
0.1	09/16/2005	Baseline document	Caroline Wright Charlie Macleod
0.2	09/20/2005	Draft Version Submitted for Acceptance	Marten Hogeweg
0.3	10/05/2005	Review comments incorporated	Caroline Wright Charles Macleod
0.4	10/10/2005	Included additional support for getMap	Caroline Wright

Table of Contents

1.0 Introduction 1

1.1 Purpose of This Document 1

1.2 Intended Audience of This Document 1

1.3 Document Overview 1

1.4 Related Documents 2

2.0 Process Overview 3

3.0 Validation Rules 4

3.1 Element Must Be Present and Be Non-null 4

3.2 Check Resource Type 5

3.3 Check Dates 6

3.4 Check Extent 6

3.5 Data Categories 7

3.6 Resource Types 9

3.7 Dates 10

4.0 Resource Type Rules 12

4.1 Introduction 12

4.2 Resource Types and Online Linkage 12

4.2.1 Content Type – Live Data and Maps (001) 13

4.2.2 Content Type – Downloadable Data (002) 15

4.2.3 Content Type – Other Documents (005) 16

4.2.4 Content Type – Static Map Images (004) 16

4.2.5 Content Type – Geographic Activities (010) 17

4.2.6 Content Type – Other 17

4.3 Processing PublishedDocID and @sourceuri 18

4.4 Standardize Date Formats 18

4.5 Including a Publication Source Attribute 20

Appendix A – Glossary of Terms 21

Appendix B – Data Category Domain Values 22

Appendix C – Resource Types Domain Values 25

1.0 Introduction

This guideline for publishing metadata to the Geospatial One-Stop Operational Portal (GOS2) was prepared by ESRI, Inc. for the Department of Interior Geospatial One-Stop (DOI-GOS) in partial fulfillment of the Geospatial One-Stop Operational Portal contract (GSA GS-35F-5086H).

1.1 Purpose of This Document

The GOS2 publisher community uses a wide range of editors to create metadata. GOS2 supports metadata that is structured according to the FGDC Content for Digital Geospatial Metadata, and the ISO 19115 metadata specifications.

Both metadata specifications provide information on the types of information and the structuring of that information for metadata that is created about a resource. The two standards have different rules as far as mandatory and optional elements are concerned and are open for interpretation with respect to the ways in which specific aspects of a resource can be described. Analysis of the existing metadata, within the geodata.gov repository, has identified many variations in the specifications to create metadata documents.

In an effort to standardize the interpretation of these metadata *variations*, a set of rules have been implemented. The rules are captured within an XSLT transform that parses and interprets the metadata content to categorize the metadata into 1 of 10 possible resource-type categories. These categories, and processing rules, are described within this document. The XSLT is also available as a download from the geodata.gov website and can be run by publishers to determine the categorization that will be applied to their metadata (*prior to publishing* to the geodata.gov portal).

1.2 Intended Audience of This Document

This document is intended for GOS2 publishers who want to create and publish metadata to the GOS2 portal.

1.3 Document Overview

This document is divided into the following sections:

- Introduction: This section introduces the document in the context of the project.
- Process Overview: This section provides an overview of the process of publishing metadata to the GOS portal, the validation of the metadata, and the application of the rules that have been defined.

- Validation Rules: This section describes the rules to which a metadata document must adhere in order to be publishable to GOS. These can be considered Go/No-Go criteria

- Resource Type Rules: These are the processing rules by which the resource type of a metadata document is determined.

- Appendixes.

1.4 Related Documents

Following is a list of all relevant documents for the project that this project management plan relates to:

- US Government's Statement of Work for Geospatial One-Stop Operational Portal (RFQ No. 1435-04-05-RP-36180), dated October 15, 2004

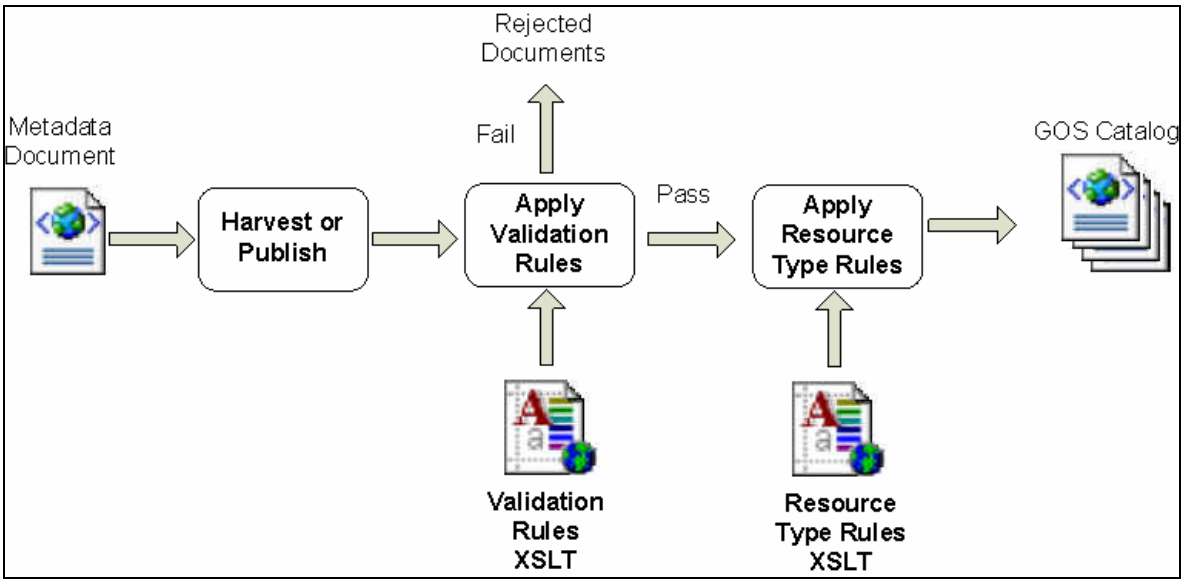
- ISO 19115:2003(E) – Geographic information – Metadata [1].

- Federal Geographic Data Committee. FGDC-STD-001-1998. Content standard for digital geospatial metadata (revised June 1998). Federal Geographic Data Committee. Washington, D.C. [2]

2.0 Process Overview

This section provides an overview of the process of publishing metadata to the GOS2 portal, the validation of the metadata, and the application of the rules that have been defined.

Figure 1 Publication process



There are a number of ways to publish metadata to GOS. Registered publishers can fill out a basic online form, upload individual metadata documents, or register a catalog of metadata for harvesting.

Regardless of the path chosen to publish metadata to GOS, the metadata will be processed as outlined in Figure 1 above.

First, the metadata will be validated to ensure it meets the minimum requirements defined for GOS. Note that these minimum requirements are less strict than the FGDC or ISO minimum mandatory set of metadata elements. If a metadata document does not meet the minimum requirements, it will be rejected and will not be published to GOS.

If the metadata passes validation, it will be categorized by the resource type rules XSLT.

Once a document has been categorized, it is published to the GOS repository (or *catalog*).

3.0 Validation Rules

This section describes the Go/No-Go criteria to which a metadata document must adhere in order to be published to geodata.gov. The latter part of the section illustrates how some of the verified tag content is used within the search capabilities of the godata.gov portal.

The rules are currently broken out into three (3) categories:

- FGDC required elements and/or values,
- ISO required elements and/or values,
- And the default, which is either the presence of an FGDC or ISO, required element and value.

The output of the validation XSLT is available as an xml document (or *xml formatted report*). It contains a listing of each check performed, the status of the check, and, if a check failed, the reason/description for the failure.

The specific checks are listed below

3.1 Element Must Be Present and Be Non-null

Many elements fall into this category and are verified for their presence in the document and must be non-null:

	Both (Default)	
Check	**FGDC**	**ISO**
Origin	/metadata/idinfo/citation/citeinfo/origin	None defined
Title	One of: /metadata/idinfo/citation/citeinfo/title, /metadata/dataqual/lineage/srcinfo/srccite/citeinfo/title, /metadata/idinfo/crossref/citeinfo/title	One of: /metadata/dataIdInfo/idCitation/resTitle, /metadata/dqInfo/dataLineage/dataSource/srcCitation/resTitle
Abstract	One of: /metadata/idinfo/descript/abstract, /metadata/idinfo/descript/purpose, /metadata/idinfo/descript/supplinf	One of: /metadata/dataIdInfo/idAbs, /metadata/dataIdInfo/idPurp, /metadata/dataIdInfo/suppInfo
Subject	/metadata/idinfo/keywords/theme/themekey	One of: /metadata/dataIdInfo/tpCat/TopicCatCd[@value !=], /metadata/dataIdInfo/descKeys[keyTyp/KeyTypCd/@value = 005]/keyword
Organization	One of: /metadata/metainfo/metc/cntinfo/cntperp/cntper, /metadata/metainfo/metc/cntinfo/cntorgp/cntorg	None defined

	Both (Default)	
Check	**FGDC**	**ISO**
Address	/metadata/metainfo/metc/cntinfo/cntaddr /addrtype	None defined
City	/metadata/metainfo/metc/cntinfo/cntaddr /city	None defined
State	/metadata/metainfo/metc/cntinfo/cntaddr /state	None defined
Zip or Postal Code	/metadata/metainfo/metc/cntinfo/cntaddr /postal	None defined

Note: The FGDC themekey and ISO TopicCatCd (evaluated in the 'subject' validation check) are used by the geodata.gov discovery service as searchable keywords or "data categories". Supplemental information on the data categories can be found in section 3.5 below.

3.2 Check Resource Type

Determining the resource type is the basis for the resource type rules XSLT. The resource type is further described in section 3.6 and extensively covered in section 4. The checks are as follows:

1. Determine the presence and non-null value of:

	Both (Default)	
Check	**FGDC**	**ISO**
Resource Type	One of: /metadata/distinfo/resdesc, /metadata/Esri/resourceType	One of: /metadata/distInfo/distributor/distorTran /onLineSrc/orDesc, /metadata/Esri/resourceType

2. If a non-null element from the relevant node-set (see above) is found, its value is stripped of all white space and it is converted to a numeric value.

   - If the value is numeric, it is checked to see if it is a number from 1 to 10 (matching the resource type). This includes string variants such as 01,001,008,09,010,10, etc.

   - If the value is non-numeric, the string is converted to lowercase, is stripped of all white space and is compared with the following possible string values. The node value must "start with" one of the following strings (remember – white space and case are ignored):

     'livedata', 'downloadabledata', 'offlinedata', 'staticmapimage', 'otherdocument', 'document', 'application', 'geographicservice', 'clearinghouse', 'mapfile', or 'geographicactivit'.

3.3 Check Dates

Both dates and date ranges are checked. There must be at least one non-null, valid date node in the document. Dates must conform to the ISO 8601 standard, specifically a profile of the ISO 8601 identified in a Note to the W3C identified in http://www.w3.org/TR/NOTE-datetime. The date format is described in extensive detail in section 4.4 below.
Note: The keywords 'unknown' and 'present' are also treated as 'valid'.

The specific date elements checked are listed below:

	Both (Default)	
Check	**FGDC**	**ISO**
Date	Any of: /metadata/idinfo/timeperd/timeinfo/mdattim/sngdate/caldate /metadata/idinfo/keywords/temporal/metadata/idinfo/timeperd/timeinfo/sngdate/caldate /metadata/metainfo/metd /metadata/dataqual/lineage/procstep/procdate /metadata/metainfo/metrd /metadata/idinfo/citation/citeinfo/pubdate	Any of:/metadata/mdDateSt /metadata/dataIdInfo/dataExt/tempEle/TempExtent/exTemp/TM_GeometricPrimitive/TM_Instant/tmPosition/TM_CalDate/calDate /metadata/dataIdInfo/dataExt/tempEle/TempExtent/exTemp/TM_GeometricPrimitive/TM_Instant/tmPosition/TM_DateAndTime/calDate /metadata/dataIdInfo/idCitation/resRefDate[refDateType/DateTypCd/@value = 001]/refDate /metadata/dataIdInfo/idCitation/resRefDate[refDateType/DateTypCd/@value = 002]/refDate /metadata/dataIdInfo/idCitation/resRefDate[refDateType/DateTypCd/@value = 003]/refDate
Date Range	Both of: /metadata/idinfo/timeperd/timeinfo/rngdates/begdate /metadata/idinfo/timeperd/timeinfo/rngdates/enddate	Both of: /metadata/dataIdInfo/dataExt/tempEle/TempExtent/exTemp/TM_GeometricPrimitive/TM_Period/begin /metadata/dataIdInfo/dataExt/tempEle/TempExtent/exTemp/TM_GeometricPrimitive/TM_Period/end

3.4 Check Extent

This check verifies the presence of a bounding box and that the coordinates are in decimal degrees and fall within a specific range (-90 to 90 or -180 to 180 for North and South and East and West coordinate values respectively).

Specifically, the check is broken down into 3 components:

- First, it determines if a bounding box is present.
- Second, it determines if the bounding box contains 4 child nodes

- Third, it checks each of the child nodes (North, South, East, West) to determine if they are in decimal degrees and are within a valid value range.

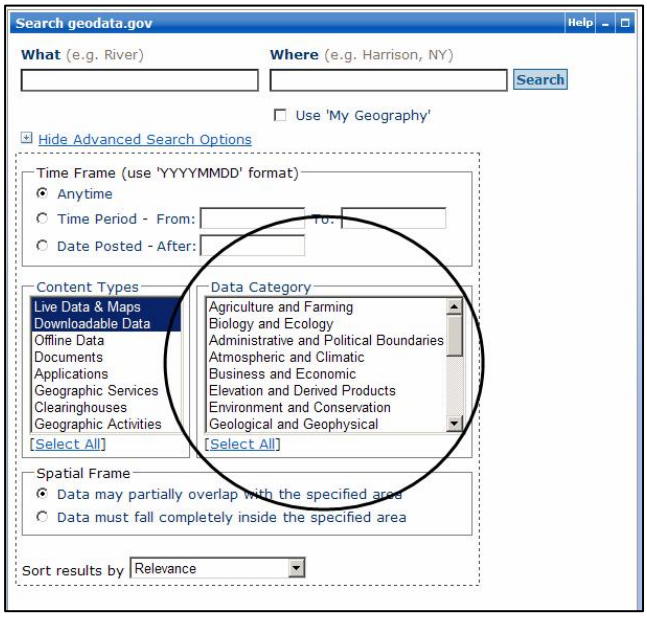
The specific extent elements checked are listed below:

	Both (Default)	
Check	**FGDC**	**ISO**
Box is present and has 4 children	/metadata/idinfo/spdom/bounding	/metadata/dataIdInfo/geoBox
Check coordinates Between 180 and -180	/metadata/idinfo/spdom/bounding/westbc /metadata/idinfo/spdom/bounding/eastbc	/metadata/dataIdInfo/geoBox/westBL /metadata/dataIdInfo/geoBox/eastBL
Check coordinates Between 90 and -90	/metadata/idinfo/spdom/bounding/northbc /metadata/idinfo/spdom/bounding/southbc	/metadata/dataIdInfo/geoBox/northBL /metadata/dataIdInfo/geoBox/southBL

3.5 Data Categories

The geodata.gov search interface supports searching for resources described in metadata based on the thematic data category.

Figure 2 Searching by Data Category



The FGDC metadata standard [2] does not prescribe a specific domain of values that may be used to define the Data Category of a specific resource described in metadata. To allow

searching of metadata according to data categories, GOS publishers must adhere to using the domain of values defined in Appendix B of this document. The domain values are based on the ISO metadata standard [1], which standardizes the supported data categories regardless of the specific metadata format the user wishes to use (ISO or FGDC).

In FGDC metadata, the Data Category for the resource is defined in the XML tag identified by the following XPath expression:

```
/metadata/idinfo/keywords/theme/themekey
```

The FGDC codes must be typed exactly as they are shown in Appendix B. For example, for a shape file containing cell phone antennas, you would use the theme code "utilitiesCommunication" for the first keyword in the FGDC metadata document as in the following example:

```
<metadata>
      …
      <idinfo>
            …
            <keywords>
                  <theme>
                        <themekey>utilitiesCommunication</themekey>
                  …
```

For an ISO metadata document the TopicCategoryCode for the FGDC utilitiesCommunication category is 019. The corresponding fragment would look like this:

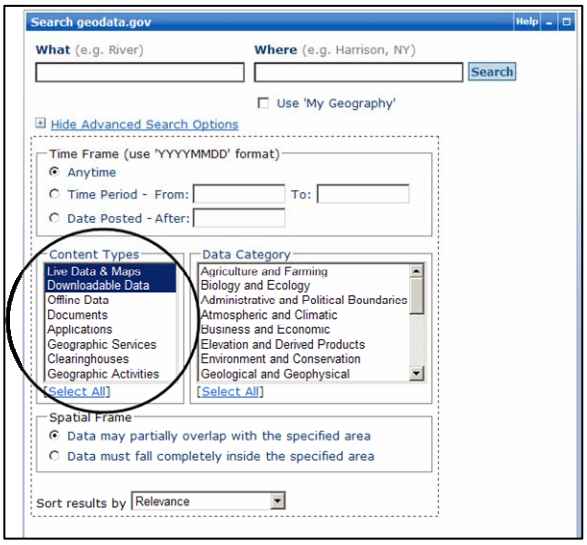
```
<metadata>
      …
      <dataIdInfo>
            …
            <tpCat>
                  <TopicCatCd value="019" />
            </tpCat>
            …
```

Note: At least one themekey must contain a valid value. The order of the themekeys does not matter. However, one of the themekeys must correspond to the FGDC codes, see Appendix B.

3.6 Resource Types

The geodata.gov search interface supports searching for resources based on the type of content that is described in metadata.

Figure 3 Searching by Content Type



In fact, by default geodata.gov will search for metadata that have content types Live Data and Maps or Downloadable Data assigned.

Every metadata record needs to have a valid content type assigned to be accepted by GOS when published or harvested. The allowable values for the content types are defined in Appendix C of this document.

In FGDC metadata, the Content Type for the resource is defined in the XML tag identified by the following XPath expression:

```
/metadata/distinfo/resdesc
```

The content types must be typed exactly as they are shown in Appendix C. For example, to describe the Environmental Information Management System developed by the EPA Office of Research and Development, you would use the content type domain value "Clearinghouses" for the resource description tag in the FGDC metadata document as in the following example:

```
<metadata>
      …
      <distinfo>
            …
            <resdesc>Clearinghouses</resdesc>
            …
```

In FGDC metadata, the Content Type for the resource is defined in the XML tag identified by the following XPath expression:

```
/metadata/distInfo/distributor/distorTran/onLineSrc/orDesc
```

For an ISO metadata document the content type code for clearinghouses is 008. The corresponding XML fragment would look like this:

```
<metadata>
      …
      <distInfo>
            …
            <distributor>
                  …
                  <distorTran>
                        …
                        <onLineSrc>
                              …
                              <orDesc>008</orDesc>
                              …
```

In the search results list of geodata.gov, in addition to a button for adding a Live Data and Map service to a map, a button is displayed that opens a website referenced in the metadata document. The URL for this button is determined if the online linkage value is valid and an additional online linkage metadata element is found in the document.

Section 4.0 contains details on the format of the online linkage information that needs to be included for the different content types.

Note: resedesc takes precedence over orDesc if both are present. If an Esri resourceType tag is present then both resdesc and orDesc take precedence over it.

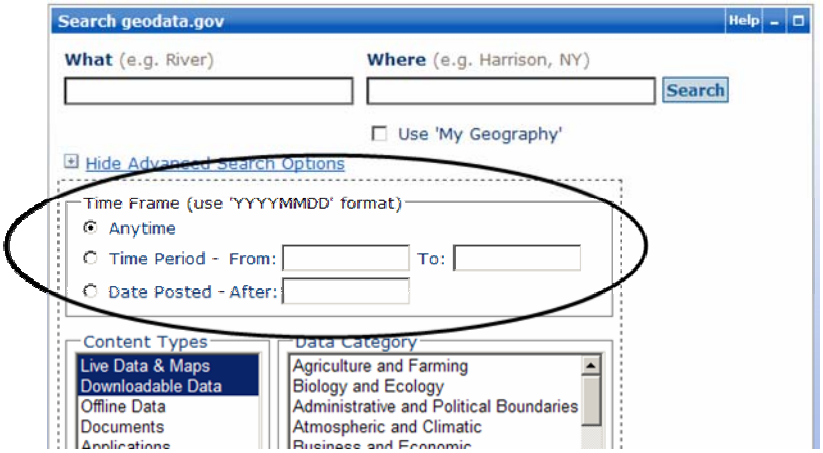
3.7 Dates

The geodata.gov advanced search options allow for searching for metadata based on a time frame. There are two options for searching by time frame:

- Time Period – Searching by time period uses the temporal extent defined in the metadata.

- Date Posted – Searching by date posted uses the time stamp recorded in the database when the metadata document was added to the GOS metadata catalog.

Figure 4 Searching by Time Frame



4.0 Resource Type Rules

This section identifies the rules that are used to derive the metadata resource type.

4.1 Introduction

When creating and publishing GOS metadata the quality of records are improved through the application of a rule set. The rules apply to both the FGDC and ISO standards. At a general level the rules seek to:

- Ensure metadata records with a resource type of Live Data and Maps have a valid map service URL.
- Ensure metadata records with a resource type of Downloadable Data have a valid downloadable data URL.
- Ensure metadata records with a resource type of Static Map Images have a valid map image URL.
- Determine if metadata records with a default resource type can be upgraded to either Live Data and Maps or Downloadable Data.
- Determine if metadata records with a resource type of Geographic Activities are either data requests or planned acquisitions.
- Confirm the existence of a Published DocID or GUID.
- Standardize date formats.

4.2 Resource Types and Online Linkage

For a number of Resource Types, the actual reference to the online resource is defined elsewhere in the metadata document. Online linkage values are detected in the following metadata tags:

Metadata Tag	Standard
/metadata/dataqual/lineage/srcinfo/srccite/citeinfo/onlink	FGDC
/metadata/idinfo/crossref/citeinfo/onlink	FGDC
/metadata/idinfo/citation/citeinfo/onlink	FGDC
/metadata/metainfo/metextns/onlink	FGDC
/metadata/distinfo/stdorder/digform/digtopt/onlinopt/computer/networka/networkr	FGDC
/metadata/distInfo/distributor/distorTran/onLineSrc/linkage	ISO
/metadata/Esri/Server	ESRI
/metadata/Esri/Service	ESRI
/metadata/Esri/ServiceType	ESRI
/metadata/Esri/primaryOnlink	ESRI

4.2.1 Content Type – Live Data and Maps (001)

For the content type Live Data and Maps (ISO code 001) the online linkage metadata tag describes an OGC Web Mapping Service, OGC Web Feature Service, or OGC Web Coverage Service. In order to be usable in geodata.gov, the online linkage value must conform to one of the structures described below.

If a valid value cannot be found for the online linkage of a Live Data and Maps content type, the metadata record is down-graded to Downloadable Data if possible (see below). If that fails as well, the metadata document is downgraded to Content Type 'Other Document' (ISO code 005).

In addition, if more than one online linkage value is provided then the XSLT will use the first one it finds that can be successfully parsed.

4.2.1.1 OGC Map Service URL (WMS, WFS, WCS)

The following conventions are supported:

Either an OGC online linkage must contain a valid ArcIMS OGC connector server URL, which includes an OGC servlet path of the form: http://<Server>/.../com.esri.<Servlet Path>/…

Some examples:

http://mapsite.gov/some-servlet/com.esri.wms.Esrimap

http://151.121.3.218:8100/some-servlet/com.esri.wms.Esrimap

Or, the URL contains an OGC service name embedded within its virtual path as a virtual path directory. This includes the path components /wfs/, /wms/, or /wcs/. The comparison ignores case so /WFS/, /WMS/, and /WCS/ are also legitimate.

Some examples:

http://www.cadcorpdev.co.uk/wfs/SisISAPI.dll

http://webservices.ionicsoft.com/unData/wfs/UN

Or, the URL contains an embedded key-value-pair of service=wfs|wcs|wms. The comparison ignores case so SERVICE=WFS|WCS|WMS are also legitimate.

Some examples:

http://wsdali.spotimage.fr/ionicwcs/coverage/SV5_031122HMX?VERSION=1.0.20&SERVICE=WCS

http://laits.gmu.edu/cgi-bin/NWGISS?VERSION=1.0.0&SERVICE=WCS

http://ot.esri.com/wcs/WCSServlet?VERSION=1.0.0&SERVICE=WCS

Or, the URL contains a "request=getMap" key value pair. In this instance the service is assumed to be 'wms'.

An example:

http://msdis-maps.missouri.edu/OGCConnector/servlet/OGCConnector?ServiceName=msdisusgs_wms&request=getMap&SRS=EPSG:4326&format=png&Layers=State

4.2.1.2 ArcIMS Image or Feature Service:

ArcIMS Image Services must have the following format:

http://<server>/image/<service_name>.

ArcIMS Feature Services must have the following format:

http://<server>/feature/<service_name>.

Some examples:

http://mapsite.gov/image/mymap

http://www.example.com/feature/streets

4.2.1.3 Key Value Notation for Live Data and Maps

The third supported format for the online linkage metadata element follows a key=value syntax. Note: this is not referring to key-value-pair arguments embedded within the online linkage value itself. Rather, many metadata documents contain online linkage values where the entire URL itself is specified as a value of a key. Key names are *not* case-sensitive. Supported keys and their usage are:

Key	Value	Description
Server	Server name	Use this key to store the name of the HTTP server and/or the complete servlet URL for an OGC services
service	Service name	Store the name of the ArcIMS service. Not required for OGC services
servicename	Service name	Same as 'service'
servicetype	’image’ or ‘feature’	Store the type of ArcIMS service. This is not required to specify an OGC service and would be ignored.

Key	Value	Description
		'servicetype' can be derived from the 'server' value.

Some examples:

Server=http://www.mysite.gov; service=abc_wms;ServiceType=image

server=http://www.mysite.gov; serviceName=abc_wms;servicetype=feature

sERVer=http://www.mysite.gov/some-dir/wmsconnector/com.esri.wms.Esrimap

Note: This is a significantly different format from where the key-value-pair (in this case 'service=wfs' is embedded as part of the URL argument list. That convention was outlined in section 4.2.1.1

Note: The key-value pairs can occur in any order. Typically, the server value is placed first. A semi-colon must separate each key-value pair.

4.2.1.4 Use of Server, Service and ServiceType Esri element tags.

The Esri metadata element contains three (3) xml elements that can be used to store online linkage information for resource types 001. These elements and their description are provided below.

Key	Value	Description
Server	Server name	Use this key to store the name of the HTTP server and/or the complete servlet URL for an OGC services
Service	Service name	Store the name of the ArcIMS service. **Not required** for OGC services
ServiceType	'Image' or 'feature for ArcIMS or wms, wcs, wfs for OGC	Store the type of ArcIMS or OGC service. Used in conjunction with the server key. Both Server and ServiceType must be present.

For ArcIMS image and feature services **all three (3) tags must be present**. For OGC image and feature services **only the Server and ServiceType tags must be present**. Service is ignored.

4.2.2 Content Type – Downloadable Data (002)

For a content type of Downloadable Data (ISO code 002) the online linkage attribute must conform to one of the following structures:

ftp://<server name>/<file path>/<file name>.<file extension>

http://<server name>/<file path>/<file name>.<file extension>

Currently GOS recognizes the following file types as downloadable files:

.zip,.e00,.gz,.tgz,.dbf,.tar,.shp,.rar,.xls,.txt,.dwg,.dxf,.dgn

Some examples:

ftp://www.co.kootenai.id.us/GIS/Metadata/hydrotxt.zip

http://kgsweb.uky.edu/download/state/KYBASINS.gz

If a valid value cannot be found for the online linkage of a Downloadable Data content type, the metadata document is downgraded to Content Type 'Other Document' (ISO code 005).

4.2.3 Content Type – Other Documents (005)

Any document that is initially classified as 'Other Documents' is evaluated to determine if it can be upgraded to a resource type of Live Data and Maps or Downloadable Data. Typically, this happens when a document is published that has either an invalid or unrecognized resource type element (see section 3.3, Resource Type check above). For a document to be upgraded to a resource type 001, it must either contain a valid URL for a map or feature service or contain the relevant ESRI Server, Service, and ServiceType tags. For a document to be upgraded to a resource type 002, it must contain a valid url that points to a downloadable data source.

4.2.4 Content Type – Static Map Images (004)

For the content type Static Map Images the online linkage metadata element must conform to one of the following structures:

ftp://<server name>/<file name>.<file extension>

http://<server name>/<file name>.<file extension>

Some examples:

ftp://www.co.kootenai.id.us/GIS/Metadata/hydromap.pdf

http://kgsweb.uky.edu/download/state/KYBASINS.MXD

Currently GOS recognizes the following file types as static map images: .gif, .jpg, .jpeg, .bmp, .pdf, .pmf, .tif, .tiff, .cal, .pct, .pict, .eps, .mxd, .av, .mpg, .mpeg, .wmv, .img, and .rm.

If a valid value cannot be found for the online linkage of a Static Map Images content type, the metadata document is downgraded to Content Type 'Other Document' (ISO code 005).

4.2.5 Content Type – Geographic Activities (010)

An important aspect of geodata.gov is the Marketplace. On the Marketplace, two types of information are displayed with the goal of bringing together those who are seeking opportunities for cost sharing with respect to the acquisition of geospatial information.

To be displayed on the Marketplace, metadata needs to be of the content type 'Geographic Activities'. This content type is further specified through three sub-types:

- Acquisitions
- Requests
- None of the above

For FGDC metadata, the element /metadata/idinfo/status/progress contains the sub-type. To support both Planned Acquisitions and Data Requests, the FGDC domain for this metadata element has been extended to include:

Value	Description
request	Data Request
planned	Planned Acquisition

For example:

```
<metadata>
        …
        <idinfo>
                …
                <status>
                        …
                        <progress>planned</progress>
                …
```

The corresponding ISO metadata element is /metadata/dataIdInfo/idStatus/ProgCd. The value of this element is defined in the MD_ProgressCode domain [1] and must be '005' . This identifies the document as an acquisition. There is no ISO equivalent for a 'request' and, as such, this particular resource sub type cannot be determined from the ISO metadata element set.

4.2.6 Content Type – Other

For all other content types, the online linkage information will be used to display a 'Go to Website' button in the search results. This allows for referring to a website of the publishing organization even though the metadata document describes an Offline Dataset.

4.3 Processing PublishedDocID and @sourceuri

If there is a PublishedDocID already within the metadata it is not changed. If there is no PublishedDocId and a GUID is provided as an input parameter (in $p-guid), a PublishedDocId element is added.

In some cases, a sourceuri attribute is also added to the PublishedDocID element. This can be the case in OAI or Z39.50 harvested documents. If a sourceuri attribute is present and is non-null, it is unchanged.

However, a sourceuri attribute can also be added to a PublishedDocId element by passing one in (i.e. a "uri") as an input parameter ($p-sourceuri). If there is no sourceuri attribute (on PublishedDocId) or the sourceuri attribute is null, the input value will be used.

4.4 Standardize Date Formats

The 'recognized' formats are based on a limited sub-set of the 8-digit format identified within the ISO 8601 standard. A complete discussion of ISO 8601 can be found at http://www.cl.cam.ac.uk/~mgk25/iso-time.html.

Specifically, the 'recognized' formats are:

Year:
YYYY (eg 1997)
Year and month:
YYYY-MM (eg 1997-07) or YYYY/MM or no separator, YYYYMM
Complete 8-digit date:
YYYY-MM-DD (eg 1997-07-16) or YYYY/MM/DD or no separator
YYYYMMDD (final, target format)

These recognized input date formats can be converted to the YYYYMMDD format and are based on a profile of the ISO 8601 identified in http://www.w3.org/TR/NOTE-datetime.

Note that all supported input formats have a 4 DIGIT YEAR component. This is in accordance with FGDC Directive 13 (http://www.fgdc.gov/standards/directives/dir13.html), whereby the supported date formats are assumed to be Y2K compliant.

Processing rules for dates:

1) Excluding separator characters, the first 4 digits are assumed to be the year, digits 5 and 6 the month, and digits 7 and 8 the day (as per the 'recognized' formats). Thus YY/MM/DD or MM/DD/YYYY will be interpreted as YYYYMM and YYYYMMDD respectively.
2) A valid year is considered to be in the range 0 A.D. to 2100 A.D. Non-valid years are converted to the year of script creation: '2005'

3) A valid month is assumed to be a value between 1 and 12. If the month is incorrect or missing, a month of '01' is assumed (i.e. January).
4) A valid day is assumed to be a value between 1 and 31. Logic to determine the correct range of days based on the month (e.g. 28, 30, 31) is not implemented. Leap year logic for February is also not implemented. If the day is considered incorrect or missing, a value of '01' is assumed.

Processing rules for time:
A time designation is also supported. The time component will be translated to HHMMSS00 where HH are in the range of 0 to 23 and 00 is a 2 digit right "pad". The recognized formats are:

Complete date plus hours and minutes:
YYYY-MM-DDThh:mm (eg 1997-07-16T19:20)
Complete date plus hours, minutes and seconds:
YYYY-MM-DDThh:mm:ss (eg 1997-07-16T19:20:30)
Complete date plus hours and minutes or hours, minutes and seconds with an am/pm designator:
YYYY-MM-DDThh:mmam, YYYY-MM-DDThh:mmpm,
YYYY-MM-DDThh:mm:ssam, YYYY-MM-DDThh:mm:sspm
(e.g. 1997-07-16T07:20:30pm which will be converted to 192030 plus 2 additional 00's for the pad - 19203000, 2001-06-15T03:35:20am which will be converted to 03352000.)
Special cases:
12 Noon (12pm) is "12". 12 midnight (12am) is 00. All other hours are represented in the range of 01 to 23.

Note: 'T' is a literal that must appear in the string. If the 'T' literal does not appear in the string or the time format is invalid, the time component will be ignored. The Time Zone Designator within http://www.w3.org/TR/NOTE-datetime, if present, will be ignored. All times are assumed to be in the local time of the zone within which this template is being executed.

4.5 Including a Publication Source Attribute

A publication source code can be provided as an input parameter to the resource type rules XSLT. Valid values are:

PubSourceCD Code	Description
0	Harvestor
1	Geography Network
2	GeoCommunicator
3	ArcCatalog
4	ESRI GOS Online Form
5	ESRI GOS XML Upload

The input value is applied to the /metadata/Esri/PubsourceCd element. If a value is not provided, a default value of '0' is used.

Appendix A – Glossary of Terms

AXL	ArcIMS XML Language
DOI	Department of Interior
ESRI	Environmental Systems Research Institute, Inc.
FGDC	Federal Geographic Data Committee
GOS	Geospatial One-Stop Operational Portal
OGC	Open Geospatial Consortium
OMB	Office of Management and Budget
USGS	United States Geological Survey
WAF	Web Accessible Folder
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Mapping Service
XML	Extensible Markup Language

Appendix B – Data Category Domain Values

To search by data category, the metadata must contain one of the data category values that GOS understands.

The data categories used in GOS are listed below. Each row in the table contains a descriptive name of a data category, the corresponding ISO Topic Category as defined in the ISO metadata standard [1], the FGDC Theme Keyword defined for that data category, and a description of the data category that is intended to support assigning a data category to a resource.

GOS Category	ISO Topic Category Code	FGDC Theme Keyword	Theme Description
Administrative and political boundaries	003	boundaries	Legal land descriptions
Agriculture and farming	001	farming	The rearing of animals or cultivation of plants. For example, resources describing irrigation, aquaculture, herding, and pests and diseases affecting crops and livestock.
Atmosphere, climatology, and meteorology	004	climatologyMeteorologyAtmosphere	Atmospheric processes and phenomena. For example, resources describing cloud cover, weather, atmospheric conditions, climate change, and precipitation.
Biologic and ecologic	002	biota	Naturally occurring flora and fauna. For example, resources describing wildlife, biological sciences, ecology, wilderness, sea life, wetlands, and habitats.
Business and economic	005	economy	Economic activities or employment. For example, resources describing labor; revenue; commerce; industry; tourism and ecotourism; forestry; fisheries; commercial or subsistence hunting; and exploration and exploitation of resources, such as minerals, oil, and gas.
Cadastral and land planning	015	planningCadastre	Land use. For example, resources describing zoning maps, cadastral surveys, and land ownership.
Cultural, society, and demographic	016	society	The characteristics of societies and cultures. For example, resources describing natural settlements, anthropology, archaeology, education, traditional beliefs, manners and customs, demographic data, crime and justice, recreational areas and activities, social impact assessments, and census information.

GOS Category	ISO Topic Category Code	FGDC Theme Keyword	Theme Description
Elevation and derived products	006	elevation	Height above or below sea level. For example, resources describing altitude, bathymetry, digital elevation models, slope, and products derived from this information.
Environment and conservation	007	environment	Environmental resources, protection, and conservation. For example, resources describing pollution, waste storage and treatment, environmental impact assessment, environmental risk, and nature reserves.
Facilities and structures	017	structure	Man-made construction. For example, resources describing buildings, museums, churches, schools, hospitals, factories, housing, monuments, and towers.
Geological and geophysical	008	geoscientificInformation	The earth sciences. For example, resources describing geophysical features and processes; minerals; the composition, structure, and origin of the earth's rocks; earthquakes; volcanic activity; landslides; gravity information; soils; permafrost; hydrogeology; and erosion.
Human health and disease	009	health	Health services, human ecology, and safety. For example, resources describing human disease and illness, factors affecting health, hygiene, mental and physical health, substance abuse, and health services.
Imagery, basemaps, and land cover	010	imageryBaseMapsEarthCover	Basemaps. For example, resources describing land cover, topographic maps, and classified and unclassified images.
Inland water resources	012	inlandWaters	Inland water features, drainage systems, and their characteristics. For example, resources describing rivers and glaciers, salt lakes, water use plans, dams, currents, floods, water quality, and hydrographic charts.
Locations and geodetic networks	013	location	Positional information and services. For example, resources describing addresses, geodetic networks, postal zones and services, control points, and place names.
Military and intelligence	011	intelligenceMilitary	Military bases, structures, and activities. For example, resources describing barracks, training grounds, military transportation, and information collection.
Oceans and estuaries	014	oceans	The features and characteristics of salt water bodies excluding inland waters. For example, resources describing tides, tidal waves, coastal information, and reefs.

GOS Category	ISO Topic Category Code	FGDC Theme Keyword	Theme Description
Transportation networks	018	transportation	The means and aids for conveying people and goods. For example, resources describing roads, airports and airstrips, shipping routes, tunnels, nautical charts, vehicle or vessel location, aeronautical charts, and railways.
Utility and communication networks	019	utilitiesCommunication	Energy, water and waste systems, and communications infrastructure and services. For example, resources describing hydroelectricity; geothermal, solar, and nuclear sources of energy; water purification and distribution; sewage collection and disposal; electricity and gas distribution; data communication; telecommunication; radio; and communication networks.

Appendix C – Resource Types Domain Values

The resource type codes supported by GOS are listed below. Only one resource type can be defined for a resource.

Content Type Code	FGDC and ISO value
001	Live Data and Maps
002	Downloadable Data
003	Offline Data
004	Static Map Images
005	Document, Other Documents
006	Applications
007	Geographic Services
008	Clearinghouses
009	Map Files
010	Geographic Activities, Geographic Activity

Note: The domain is actually broader than this and is based on empirical analysis of metadata resource type descriptions within geodata.gov. Section 3.2 states that the resource type description string is lowercased, and stripped of all white space prior to performing any comparison. In addition, the resource type description is compared to see it is *starts with* a particular value and not if it *equals* a particular value. Thus, live data, lIVe data and Maps, live data and maps ARCIMS image service, etc. would all be valid resource type descriptions.