ArcIMS® Metadata Services

An ESRI® White Paper • May 2002

ArcIMS Metadata Services

An ESRI White Paper

Contents	Page

ArcIMS Metadata Services

Introduction to Metadata

Metadata is "data about data"—a summary document providing content, quality, type, creation, and spatial information about a data set. It can be stored in any format such as a text file, HTML page, or database record. Because of its small size and compatible format, metadata has the highly valuable characteristics of being shareable and easy to manage. Metadata can be stored with the data set it details or in a separate location with external linking to its data sets. By creating metadata and sharing it with others, information about existing data becomes readily available to anyone seeking it. This greatly facilitates data discovery and reduces data duplication. These characteristics are what drive the goals behind ESRI® ArcIMS® Metadata Services.

Introduction to ArcIMS Metadata Services

ArcIMS Metadata Services provide for a central online repository of metadata. They are modeled after the Geography Network℠, which is a global online resource for discovering and accessing geographic content. Working over the Internet, Metadata Services are accessible to anyone with access to the Web. Metadata Services are similar to a search engine but geared specifically toward geographic data. For each data set in a search, you can view the complete metadata for the data set, view an interactive map if available, and in some instances, download the data.

The primary goals of Metadata Services are

■ Sharing data

■ Allowing data producers to easily publish their metadata to a central repository

■ Allowing data consumers to easily search for metadata in a central repository

■ Offering "Geography Network in a Box"—an easily customizable application enabling others to create their own geography network

Metadata Server Architecture

Metadata Services run as services on the Metadata Server, which is part of the ArcIMS Spatial Server.

The architecture of the Metadata Server encompasses three existing ESRI products.

1. ArcIMS: Provides the framework and architecture on which the Metadata Server runs.

2. ArcGIS™ ArcCatalog™ application: Serves as the authoring and publishing tool.

3. ArcSDE™: Stores published metadata in records inside a relational database accessed via ArcSDE.

A fourth component of the ArcIMS Metadata Server is the Metadata Explorer, a Web application that allows you to search and browse the contents of a Metadata Server from a Web browser.

The following diagram and description show how all of these components work together:

Figure 1
Metadata Server Component Interaction

1. Metadata is authored and managed in ArcCatalog.

2. Metadata is published to the Metadata Server using drag and drop operations. The published documents are stored in an ArcSDE database.

3. The Metadata Server can be searched via the Web using a variety of clients such as the Metadata Explorer or ArcCatalog.

The Metadata Server is a type of server inside the ArcIMS Spatial Server. In this document, however, Metadata Services are regarded as a combination of the Metadata Server component of the ArcIMS Spatial Server and the ArcSDE database in which published metadata documents are stored. For more detailed information regarding the ArcIMS Spatial Server, refer to the white paper *ArcIMS 4 Architecture and Functionality*.

Creating Metadata Services in ArcIMS

A Metadata Service runs on the ArcIMS Spatial Server. The service uses an ArcSDE database as the repository into which all published metadata documents are stored. Creating a Metadata Service is analogous to creating an ArcIMS Image or Feature Service. A metadata configuration file is used as input to a Metadata Service. The configuration file contains specific information about the database that will be associated with the Metadata Service, defines certain service parameters, and enables the addition of custom administrative features to the database. Services are administered using the ArcIMS Administrator.

J-8909

Figure 2
ArcIMS Services Including a Metadata Service
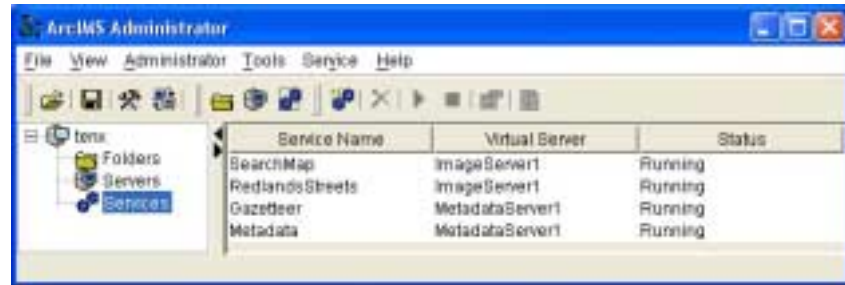
| | **ArcSDE Considerations** | ArcSDE plays a vital role in Metadata Services. It is the interface to the database that ArcIMS relies on to store, search for, and retrieve metadata documents. All requests from ArcIMS are translated to Structured Query Language (SQL) and then sent through ArcSDE to the relational database. The response follows the same path in reverse from the database to the ArcIMS server. Database connection information, such as the ArcSDE server, database, user name, and password, is specified in the metadata configuration file. Several parameters are also available to tune the ArcSDE database to specific needs. |

ArcSDE Considerations

ArcSDE plays a vital role in Metadata Services. It is the interface to the database that ArcIMS relies on to store, search for, and retrieve metadata documents. All requests from ArcIMS are translated to Structured Query Language (SQL) and then sent through ArcSDE to the relational database. The response follows the same path in reverse from the database to the ArcIMS server. Database connection information, such as the ArcSDE server, database, user name, and password, is specified in the metadata configuration file. Several parameters are also available to tune the ArcSDE database to specific needs.

In addition to the tables that are automatically created by ArcIMS for a Metadata Service, ArcSDE can be further customized to give the database additional functionality when receiving requests and sending back responses. The default database consists of a set of tables used to store the published documents and associated information. In addition to these tables, it is possible to customize the database by adding another admin table that can store custom parameters associated with each published metadata document.

Creating, Publishing, and Managing Metadata Using ArcCatalog

Before metadata is published, it must first be created. The default authoring tool is ArcCatalog, the data management tool that is part of ArcGIS. Metadata created in ArcCatalog is stored in XML format. Any XML document, as long as it has valid XML syntax, can be published to the metadata repository.

Two editors are available in ArcCatalog for metadata creation. One follows the Federal Geographic Data Committee standard and the other follows the International Organization for Standardization standard. The output is an XML document.

Figure 3
The FGDC Metadata Editor
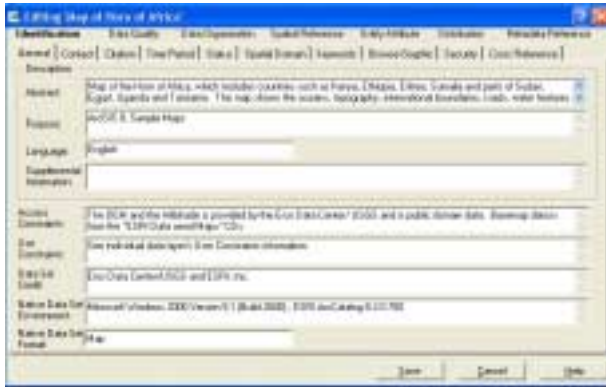
Figure 4
The ISO Metadata Editor

Although an XML document can be any format and can contain any information, ArcIMS has the concept of content validation. When enabled, five pieces of information are required to be present in the metadata before it can be published: title, publisher, spatial extent, data theme, and content type. This information is required for the Metadata Explorer and ArcCatalog to perform searches and display results. If the information is not present and validation is enabled, an error is generated and the document cannot be published. If validation is disabled, however, the Metadata Server can be customized to work with any content or content standard desired.

Once metadata has been authored, it is ready to be published. ArcCatalog can be used as the publishing tool. During the publishing process, the metadata is copied to the repository in the ArcSDE database and indexed for fast searching. The metadata document itself is stored in the database as a binary large object (BLOB).

Once a document has been published, it is possible to assign extra attributes to the document. One such attribute is that of relationships. Published documents can be assigned relationships between one another either within the same folder or across folders. When a data set is returned as part of a result set, the metadata for that data set will also include a hyperlinked list of all related documents, giving the user the ability to find similar data quickly.

Figure 5
Properties window in ArcCatalog lets you assign
related documents to a published data set.

Published metadata can be organized on the server in different ways. ArcCatalog is used as the server's document managing tool, providing the ability to create new folders and hierarchies into which the published metadata is placed. Hierarchies can be based on geographic location, subject, or publisher. Structured hierarchies make it easier to browse the contents of a Metadata Server and find data.

Searching Metadata   ArcIMS comes with a Metadata Explorer to search or browse the contents of a Metadata Service. ArcCatalog is another client that can be used for searching and browsing.

The Metadata Explorer is a Web application and runs in both Netscape and Internet Explorer 4.0 and above. Metadata can be searched in two ways: by specifying detailed search criteria or by browsing through the server's contents.

Figure 6
The Metadata Explorer is a Web browser client to the Metadata Server.

The Metadata Explorer offers searching capabilities via geographic extent, content type, data theme, or keyword. Using a combination of search criteria narrows the search results and speeds up the search response times.

The keyword criterion is a simple text box allowing any text to be searched. By default, the full text of metadata is searched for the specified keyword. Criteria for content type and data theme are both chosen from predefined drop-down lists. Geographic extents can be specified by dragging a box around the area of interest on the available map (with the ability to pan and zoom around the map to get a close-up view) or by finding the extent of a place name via a gazetteer and having the gazetteer define a box on the map automatically. The gazetteer is another table in the ArcSDE database listing approximately two million records of geographic place names and their spatial extent.

Once a search has been initiated and a response has been received from the ArcIMS Spatial Server, the search results are presented to the user in a summarized view. The initial display of results shows basic information about each matching data set such as title, content type, publisher, map scale, and a thumbnail of the data if it exists. For each result, it is then possible to view that data set's complete metadata, a listing of all the data sets related to the one being viewed, and the coverage area. If the content type of the data set is specified as being "Live Data and Maps," it may be possible to view an interactive map for this result with the ability to pan and zoom around the map. This gives the user a graphical context to the data that has been found. If the Metadata Explorer has been launched from inside the ArcGIS ArcMap™ application or from ArcExplorer™ and the data is available for download, it may be possible to download the data from the Metadata Explorer into ArcMap or ArcExplorer and work with it—just like any other data set that resides locally on the machine in use.

Figure 7
Display of Browse Results in the Metadata Explorer

The search results, detailed metadata, and interactive maps that are viewable via the Metadata Explorer can also be printed. Printing the summary results will exclude navigational buttons associated with each returned data set.

Searching for data using ArcCatalog presents all the same features that are available in the Metadata Explorer plus some additional features. By using ArcCatalog as the search client, it is possible to search for specific data types such as coverages or computer-aided design (CAD) data. Geographic extent searches are enhanced, date criteria can be included, and multiple keywords can be used. In addition, keywords for only specific metadata fields can be specified.

Once a document has been found, the data that is described in the document can be obtained in a variety of ways. In some instances, the data can be directly downloaded. If the data is a live Image Service, the map can be viewed interactively using the Metadata Explorer's map viewer.

Other ways of obtaining data include

■   Via the contact information present in the metadata.

■   Via online links included in the document. These can be UNC paths to the actual data, links to Zip files, or simple URL addresses to the organization that owns the data.

Searching Other Services
Using the Metadata Explorer or ArcCatalog as the search clients, only one Metadata Service can be searched at a time.

It is possible to search other services using the following methods:

■ Federated searches that propagate a search down the network to all connected metadata servers

■ Harvested searches in which the contents of a Metadata Server are dispersed out to other servers and imported by those servers into their own databases

The Metadata Server uses the harvesting model to achieve the searching of other servers. The reason is that the federated searching model opens the door to unresponsive or slow searches due to network traffic and/or server load of metadata services that are any number of nodes down the line. Not having to rely on these other servers ensures that the response time for a search is always kept within a certain range.

Harvesting works along the principle that metadata documents are harvested from multiple metadata repositories and brought into the harvesting server using import/export tools provided with ArcIMS. Metadata documents are harvested from a server, organized into an export format of files and directories, and made available to any interested parties. The harvesting server can then read this series of documents, importing all the metadata documents into its own database, re-creating all the relationships, and making available the contents of the foreign server's metadata documents as its own. Additional features are available to provide control over which documents are exported or imported based on the date of the data set's last update, the public/private status of the data set, whether the documents are native to the server (as opposed to having been harvested themselves), and the database from which they came.

Security Metadata Services use ArcIMS software's existing security model. Permissions to perform various actions on the Metadata Service are defined in an access control list (ACL) file. The level of access to a Metadata Service is defined by four roles. Each role defines the various actions that can be performed on a Metadata Service.

1. metadata_browser: Permission to browse the contents of a metadata server. If a user connecting to a Metadata Service does not specify a user name and password, the metadata_browser role will be assumed by default.

2. metadata_publisher: Permission to publish metadata to the metadata repository as well as create new folders for published metadata organization. Moving or deleting documents is only permissible to the user that published the document, otherwise known as the owner of the document.

3. metadata_service_author: Permission to create metadata for Image or Feature Services residing on the server.

4. metadata_administrator: Permission to administer the database. This administrative role also gives full access to all published metadata documents with the ability to manage the contents of the repository or to reset the database completely.

J-8909

Figure 8
Sample ACL File

```
<?xml version="1.0" ?>
- <AIMSACL>
    <USER name="*" services="Metadata, SearchMap, RedlandsStreets" />
    <USER name="browse" password="browse" services="Metadata" roles="metadata_browser" active="1" />
    <USER name="publish" password="publish" services="Metadata" roles="metadata_publisher" active="1" />
    <USER name="author" password="author" services="Metadata, RedlandsStreets, SearchMap"
        roles="metadata_service_author" active="1" />
    <USER name="administrator" password="administrator" services="Metadata, SearchMap,
        RedlandsStreets" roles="metadata_administrator" active="1" />
</AIMSACL>
```

Client Customization

Metadata Services are customizable on all clients in a variety of ways.

Metadata Explorer

The Metadata Explorer, the default search client, is an application built using JavaServer Pages (JSP). The source code for the Metadata Explorer is included with ArcIMS. The Metadata Explorer is thus an open source and fully available to customization. Several parameters in the metadata server configuration files also give users the ability to customize the interface. The following is a list of features that allow users to customize the Metadata Explorer.

Configuration files

- Requirement of authentication
- Default number of results appearing per page
- Metadata Explorer title
- Service associated with the Metadata Explorer
- Gazetteer availability

Source code modification

- Graphics
- Title
- Help instructions
- Search criteria interface
- Results display (both summarized and detailed view)
- Map display

ArcCatalog

ArcCatalog can be customized in the following ways:

- Search engine and search criteria
- Publishing process
- Metadata editors
- Style sheets for viewing metadata

Custom Clients

In addition to the out-of-the-box clients available for authoring, publishing, and searching metadata, custom clients can be created for accessing Metadata Services. The only

requirement of any such custom application is that the client speaks ArcXML—the language of ArcIMS. Requests sent to the server must be received in ArcXML, and responses that are sent back must understand ArcXML.

Z39.50 Support

Z39.50 is a communications protocol, much like ArcXML, used by servers and clients to send and receive requests and responses. It is the standard protocol used by the National Spatial Data Infrastructure (NSDI) Clearinghouse and its clients for data searches. The NSDI Clearinghouse can be searched using the Metadata Explorer. On the search interface, a check box is present enabling the user to indicate whether a clearinghouse search is to be initiated at the same time as a Metadata Server search. If the check box is turned on, when a search is started, a separate browser window pops up linking to the Geography Network Explorer that initiates a clearinghouse search. The search has the same criteria that were specified in the Metadata Explorer.

Figure 9
Metadata Explorer's Support
of Clearinghouse Searches

Conclusion

ArcIMS Metadata Services constitute a powerful application available with ESRI's release of ArcIMS 4. Metadata Services enable organizations to create their own custom geography networks and give them the ability to manage, organize, and share metadata within their own divisions, across departments, or with other organizations. Metadata Services expand upon ESRI's vision of global data sharing and easier access to existing data.