The Government Information Locator Service: Discovering, Identifying, and Accessing Spatial Data

The Government Information Locator Service (GILS) is a new federal initiative to assist the public in discovering, identifying, and locating government information. GILS may play a special role in helping the spatial data community to search and retrieve information about spatial information resources created, collected, or held by federal agencies. GILS is a decentralized standards-based approach to network discovery and retrieval. The basic components of GILS are structured records (i.e., the GILS Locator Records) with standardized data elements that describe and provide access information to federal information resources; agency-based information servers hosting the Locator Records; client software to initiate information retrieval transactions; and ANSI/NISO Z39.50 as the communications protocol between clients and servers. This paper provides an overview of GILS and discusses ANSI/NISO Z39.50, the American National Standard for information retrieval, and its use in GILS. The paper concludes with a discussion of the implications of GILS for the discovery and use of spatial data.

INTRODUCTION

A major barrier in the effective use of distributed electronic networks and the information available through them is the lack of adequate mechanisms for discovering and retrieving information that is relevant to users' information needs. Nowhere is this more true than in the realm of federal information resources and more specifically geographic and spatial data created, collected, or held by federal agencies.

The Government Information Locator Service (GILS) is one mechanism that will assist users to search for and retrieve information. GILS is a decentralized standards-based approach to network discovery and retrieval and uses ANSI/NISO Z39.50, the American National Standard for information retrieval. GILS is a network-accessible service through which the public will be able to identify and locate federal information resources. GILS will help users identify what resources exist and provide information about those resources to allow the users to determine the utility of the information resource for their needs. Additionally, GILS provides
users the necessary information for accessing or acquiring the resources. While the focus of the initial GILS effort has been on federal information resources, its use is not limited to those resources. Some state governments and other organizations are investigating the use of GILS to assist their user communities in discovering and retrieving information.

This paper provides an overview of GILS and discusses Z39.50 and its use in GILS to facilitate information retrieval in the increasingly distributed electronic environment of federal information. The paper concludes with a discussion of the implications of GILS for the discovery and use of spatial data.

BACKGROUND AND CONTEXT

GILS is a new federal initiative to assist users in discovering, identifying, and accessing or acquiring federal information resources. Through the use of existing technology, GILS provides a framework and specifications for agencies to describe their information resources and make those descriptions available to users directly via the Internet or through any number of intermediaries. Additionally, GILS is intended to assist government agencies in the management of federal information resources. Adequate management of information resources, enhanced access to public information, and dissemination of agency information products are three primary goals of GILS. It is important to see these goals as information policy goals and to understand that GILS is a policy-driven initiative.

A brief overview of the current policy and technology context within the federal government will assist in understanding the emergence of GILS. For many years, the primary policy instrument for federal information has been the Office of Management and Budget's (OMB) Circular A-130, "Management of Federal Information Resources." Originally published in 1985, A-130 directed federal agencies in their information management activities. Beginning in the late 1980s, OMB began the process of revising A-130. This effort culminated in the publication of substantial revisions to the circular in 1993 and in 1994 (Office of Management and Budget, 1994b). While maintaining a focus on federal information resources management, sections of the revised circular also addressed agencies' responsibilities in providing access to, and dissemination of, government information. In part, the revised A-130 set the stage for GILS.

Other activities have also shaped the policy context of the current GILS initiative. For example, two research studies examined problems and design issues related to a federal information locator system (McClure et al., 1990; McClure et al., 1992). Those studies suggested that a user-based approach to designing a locator system, policy analysis and policy advocacy, and an awareness of technology trends and information technology standards can be particularly effective to connect the needs of in-
formation users and providers in achieving broader federal information policy goals.

The development of GILS occurred within the larger policy context reflected, for example, in the National Information Infrastructure (NII) and the National Performance Review (NPR) initiatives. The NII and NPR are two examples of important policy statements that reflect the Clinton administration’s encouragement of the government’s use of electronic networks to increase public awareness of, and access to, government information and services. The NII initiative encourages agency efforts to expand electronic dissemination through diverse channels (Information Infrastructure Task Force, 1993). The GILS is seen as a federal contribution to the NII by providing a “virtual card catalog” to government information. The National Performance Review, released by Vice President Gore in September 1993, encourages agencies to develop electronic dissemination programs that employ the right balance of public and private sector efforts (National Performance Review, 1993). The current incarnation of OMB Circular A-130, however, provides important enabling policy for GILS. The circular directs agencies much more forcefully in their responsibilities in the areas of information access and dissemination. Agencies are, among other things, created to:

- help the public locate government information maintained by or for the agency;
- disseminate information on equitable terms;
- develop aids to locating information;
- establish and maintain inventories of all agency information dissemination products;
- use electronic media and formats to make government information more easily accessible and useful to the public; and
- use information technology standards to create open and interoperable information systems.

GILS is a policy and technology response to the needs of the public to identify, locate, and access or acquire government information. OMB recently formalized the policy on GILS in its Bulletin 95-01, “Establishment of Government Information Locator Service” (Office of Management and Budget, 1994a). The bulletin directs federal agencies to inventory their information resources and begin describing those resources in GILS Locator Records.

Before users can access or acquire information, they must determine whether that information exists, whether there are restrictions on its use, and what methods are available for accessing or acquiring it. A locator serves exactly that purpose. According to McClure, Ryan, and Moen (1992), a locator is a “machine-readable database that identifies different information resources (e.g., databases, libraries, clearinghouses, print
publications, bulletin boards, etc.) and describes the information available in these resources. Usually, the locator does not provide the actual information, but rather points the user to the information sources that do provide the needed information” (p. 2). Locator information is metadata or data about data. A locator is a point of entry for locating government information, regardless of the format and content of that information, and tells the user: (1) what information is available on a particular topic, (2) where that information is located, and (3) how the user would access that information.

AN OVERVIEW OF GILS

GILS is an agency-based, Internet-accessible, locator service. Direct users of GILS will connect to GILS servers via the Internet to find information about a wide range of federal information resources. The public will use GILS either directly or through intermediaries (the intermediaries obtain GILS information as direct users themselves or from other intermediaries [see Christian, 1994, p. 4 for a description of the various classes of GILS users]).

Federal agencies will develop and maintain GILS servers. These decentralized agency-based GILS servers enable ongoing maintenance responsibilities to be carried out by those who understand and manage information resources.

GILS servers are machine-readable databases that contain Locator Records describing federal information resources. A Locator Record consists of a number of data elements (i.e., GILS Core Elements) that identify and describe an information resource (GILS Core Elements are noted in uppercase letters throughout this paper). Several data elements are included in Locator Records to facilitate GILS navigation and network-based access to information. For example:

- Each Locator Record contains a CONTROL IDENTIFIER in the form of a Uniform Resource Identifier (URI). An agency’s server may contain Locator Records with CONTROL IDENTIFIERS that identify Locator Records from other agencies’ servers. This data element allows GILS Locator Records to be replicated on multiple servers for the convenience of GILS users.
- Each Locator Record contains an AVAILABILITY element that informs the user how to procure the described information resource. If the information resource is an electronic information system or electronic document, the AVAILABILITY element includes AVAILABLE LINKAGE information in both human- and machine-readable form. The network linkage information may be used to connect to, and access, the electronic information resource.
Different agencies may create or offer Locator Records describing the same information resource (these may be existing Locator Records that have been replicated and/or modified or entirely new Locator Records). These multiple records can offer different views of a single resource from the particular perspectives of the agencies creating/modifying a Locator Record.

The semantics of the Locator Records coupled with client software that understands these semantics and building upon the ability of the Z39.50 protocol to provide a uniform interface to multiple autonomously managed servers provide the user with the impression of seamless navigation among these distributed servers. The semantics of the Locator Records facilitate elimination of duplicate records further fostering the impression of a single system built out of autonomous distributed servers.

Each GILS server can be represented by a Locator Record in other GILS servers. Some of these servers will include references to all other GILS servers, and these might be regarded as a kind of "directory of directories." However, GILS itself does not assign any hierarchical status to specific servers nor does it specify a "root server." Rather, the structure and content of the GILS Locator Records enable, for example, the aggregation of Locator Records in "directories" that could be offered by one or more federal agencies or other organizations.

A GILS server accessed using Z39.50 in the Internet environment acts primarily as a pointer to information resources. GILS servers will support searching (i.e., accept a search query and return a result set or diagnostic messages) and may support browsing (i.e., accept a well-known search query and return a list of Locator Records in brief display format). Direct users must have prior knowledge of at least one GILS server and its network address and must be able to access it to enter the GILS. Once connected to a GILS server, users supported by appropriate clients that understand the GILS Profile, may navigate through single or multiple GILS servers by following the links provided in the Locator Records. The use of the national standard for network information retrieval, Z39.50, provides for interoperability between clients and multiple servers. GILS, then, is a distributed resource consisting of agency-based servers accessible via Z39.50 that provide users with the potential to discover, identify, and locate federal information resources.

THE GILS PROFILE

The Federal Information Processing Standards (FIPS) Publication 192 lays out the specifications for GILS applications (National Institute of Standards and Technology, 1994). A profile is "a set of one or more base standards, and where applicable, the identification of chosen classes, sub-
sets, options and parameters of those base standards, necessary for accomplishing a particular function" (International Organization for Standardization/International Electrotechnical Commission, 1992, p. 2). Profiles are also referred to as "functional standards," "implementation agreements," or "specifications." Since open systems standards often include choices and options, profiles specify the values and parameters of a standard for an application to increase the likelihood of interoperability and interworking of separate implementations. A profile, then, is a set of implementation agreements that guide implementors in applying one or more standards in a specific and limited context. Separate implementations will have an improved likelihood of interoperability and interworking when they conform to a common application profile.

The development of the GILS Profile was completed in 1994 under the auspices of the research project, "Expanding Research and Development on the ANSI/NISO Z39.50 Search and Retrieval Standard," coordinated by Syracuse University and the United States Geological Survey and funded by the Interagency Working Group on Data Management for Global Change. A project team comprising experts in Z39.50 implementations, system implementations, and information organization, and representatives of federal agencies developed the specifications (i.e., the GILS Profile) for initial GILS implementations (for a complete description of the research project and the development of the GILS Profile, see Moen and McClure [1994]).

The GILS Profile development project built upon a previous study, Identifying and Describing Federal Information Inventory/Locator Systems: Design for Networked-Based Locators (McClure et al., 1992). That study, which was conducted for OMB, the National Archives and Records Administration (NARA), and the General Services Administration (GSA), recommended that each federal agency establish a network-accessible locator that describes its information resources. The study also recommended that agencies use Z39.50 as the appropriate information retrieval protocol to achieve a distributed standards-based Government Information Locator Service.

The Government Information Locator Service (Christian, 1994) provided the project team with high-level requirements for GILS. Based on those requirements, the project team delineated assumptions about the operation and information flows of GILS and developed functional requirements. This process allowed the project team to identify a subset of Z39.50 and other existing and emerging standards that would support these functional requirements.

The GILS Profile fully specifies the use of Z39.50 for the GILS application. The profile also addresses other aspects beyond Z39.50 for information servers that are GILS conformant. The current version of the
profile only addresses servers, and while acknowledging a client’s role in information retrieval, the Profile does not specifically address client functionality. A GILS client, however, will be able to interoperate with any GILS server. While the GILS Profile specifies many aspects of GILS applications, it does not address information system characteristics such as interface requirements, the internal structure of databases that contain GILS Locator Records, or search engine functionality.

Z39.50 provides a key part of the foundation for GILS. This standard enables the interoperability of a variety of systems and hardware platforms in a client/server environment for the purposes of information retrieval.

Z39.50: A COMPUTER-TO-COMPUTER PROTOCOL FOR INFORMATION RETRIEVAL

Understanding how GILS works requires some familiarity with ANSI/NISO Z39.50-1992, *Information Retrieval Application Service Definition and Protocol Specification for Open Systems Interconnection* (National Information Standards Organizations, 1992). Michael and Hinnebusch (1995) argue that “Z39.50 is the single most important networking standard available today” (p. 15) and state that: “As far as the library community is concerned, it is the most important protocol available today” (p. 21). Z39.50 provides users with the capability to search and retrieve information in the networked environment. Z39.50 recognizes that information retrieval consists of two primary components—selection of information based upon some criteria and retrieval of that information. It provides a common language for both activities. Z39.50 standardizes the manner in which the client and the server communicate and interoperate even when there are differences between computer systems, search engines, and databases.

The National Information Standards Organization (NISO), an American National Standards Institute (ANSI) accredited standards developer that serves the library, information, and publishing communities, approved the original standard in 1988 (referred to as Z39.50-1988 or Version 1). Shortly after the approval of the standard in 1988, a group of Z39.50 implementors began work to enhance and expand the utility of the standard. NISO balloted a revised version of Z39.50 and published the new standard in 1992 (referred to as Z39.50-1992 or Version 2). Continuing development of the standard by the Z39.50 Implementors Group (ZIG) resulted in the third version of the standard (Version 3). A NISO ballot of Version 3 will be completed in the first quarter of 1995.

Oriented initially toward information retrieval of bibliographic records, Z39.50 is not limited to the formats and kinds of data it can handle. New features proposed in Version 3 extend Z39.50 to support full text and images and provide the functionality to accommodate information
retrieval beyond simple bibliographic information. These features emphasize the standard’s utility in complex information infrastructures that process and handle information in various formats.

Z39.50 can be implemented on any platform and enables different computer systems—with different operating systems, hardware, search engines, database management systems—to interoperate and work together seamlessly, and thus supports an open systems environment. Z39.50 supports information retrieval in a distributed, client and server environment where a computer operating as a client submits a search request (i.e., a query) to another computer acting as an information server. Software on the server performs a search on one or more databases and creates a result set of records that meet the criteria of the search request. The server returns records from the result set to the client for processing. In a client/server architecture, software for end-user interaction and display (the client) is separate from the software that manages the information, performs the search, and returns the results (the server). Z39.50 does not address the user interface (e.g., its “look and feel”), but there are protocol specifications and procedures that pertain to the Z39.50 client (referred to in the standard as the Z39.50 “origin”) such as the initiation of an information retrieval query and how it requests specific operations from the server. There are also protocol specifications and procedures that pertain to the Z39.50 server (referred to in the standard as the Z39.50 “target”) such as the manipulation of result sets and the formats in which it returns records to the client. Z39.50 addresses the complex communication among computers for the purposes of information retrieval.

Each database residing on information servers can have unique characteristics. For example, databases may differ in the way they store data and in the access points available for searching. The records in each of the databases may also have different structures and consist of different data elements. The objective of Z39.50 is to support computer-to-computer communication in standard and mutually understandable terms and support the transfer of data between the systems independent of the structure, content, or format of the data in a particular system. However, in individual implementations, servers may be limited to specific formats of data that can be exported and the access points that are supported for searching.

When a database is searched, the client passes a query to the server. The query contains search “terms” (e.g., terms that the user has identified to be matched against access points in the database) and “attributes” of those search terms (e.g., specifying the terms as an “author” or “title,” specifying if the terms are to be “truncated,” etc.). Queries can include different attribute types. For example, if a user wants to search for an author’s name, a “use” attribute specifies the search term as “author.” If the user wants to search for
all books published after a certain date, a "use" attribute specifies the search term is a "date of publication" and a "relation" attribute specifies that the user wants all dates of publication "greater than" a particular date. Z39.50 enumerates these attribute types and their values in registered attribute sets. Standardized and mutually recognized attribute sets allow implementors a common basis for intersystem communication.

After the server executes a search of a database, the server creates a result set consisting of those records that match the criteria of the query. Clients can request that servers return records from a result set, or they can issue additional searches that further qualify a result set or use result sets as arguments in subsequent searches.

When the user wants to display records listed in the result set, ANSI/NISO Z39.50 provides choices about which data elements (i.e., element sets) from the database record the user can request. It also gives choices about the format for transferring the record (i.e., a record syntax) from the server to the client. Z39.50 registers standardized Element Set Names and Record Syntaxes to support client/server communication for this aspect of information retrieval.

Z39.50 is an information retrieval protocol. The standard grew out of an early recognition by people in the library community that users should be able to search remote information systems for information just as they search their own and without the need to learn new search commands and techniques. A Z39.50 implementation enables one interface to search for, and retrieve, information from multiple systems and thus provides end-users with nearly transparent access to other systems.

**GILS AND Z39.50**

There are several basic components of GILS—the data (i.e., machine-readable records) stored in one or more databases; an information server hosting the records; client software used to initiate a query; and Z39.50 as the communications protocol between the client and the server. Z39.50 distinguishes between Z39.50 functions that are the responsibility of the client and Z39.50 functions that are the responsibility of the server; these may be logically associated with the information server and the user's client. The Z39.50 functions pertinent to this discussion are those related to searching, retrieving, and presenting GILS Locator Records.

**GILS Locator Records**

GILS Locator Records consist of a number of data elements. The data elements are used to contain information that identify and describe and provide access information to federal information resources. Figure
### Figure 1. GILS locator record core elements

1 lists the GILS core elements (for the complete semantics of each element, see the GILS specifications in FIPS 192).

Lynch (1992) argues that information semantics in a distributed computing environment must be addressed if the promise of networked information is to be realized. The GILS Profile’s use of a common vocabulary
for data elements and a common information structure of the records is an important step forward. Accuracy, completeness, currency, and consistency of data in the records, however, will be criteria by which the quality of the data can be evaluated. While the technology (i.e., Z39.50 clients and servers) may be able to process locator information, the end-users will be badly served if the data are lacking in quality.

The National Archives and Record Administration (1995) has developed Guidelines for the Preparation of GILS Core Entries. The OMB stated in its Bulletin 95-01 that NARA should publish guidance for federal agencies on the content of GILS Locator Records to assist those agencies in creating high quality and useful records.

GILS Locator Records are machine-readable records stored in a database on an information server. The use of Z39.50 focuses less on the database record itself than on a standardized representation of the record for the common understanding of the client and server. One aspect of this common understanding is the available access points for searching the record (i.e., the GILS Attribute Set); another aspect is the elements of the record or those parts of the record that the client can request of the server to return to the client for display to the user (i.e., Element Set Names).

Searching GILS: The GILS Attribute Set

Z39.50 logically separates searching for information from retrieving information. Searching is done by formulating a query and passing that query to an information server. Retrieval is done by requesting the server return one or more records, or element of a record, in one or more record syntaxes to the client.

GILS servers may support a variety of search strategies including those:

- to find known items (e.g., where the user knows the exact TITLE of an information resource described in a Locator Record);
- to find resources whose Locator Records contain certain words or phrases;
- to find resources by topic (e.g., using a controlled vocabulary); and
- to find resources whose Locator Records meet other criteria (e.g., specific spatial data coordinates).

Searching in Z39.50 uses the concept of Attributes. "Use" Attributes are access points in a database record that can be searched. The searchable elements of GILS Locator Records correspond to GILS Use Attributes. The GILS Attribute Set is a superset of the Z39.50 Bib-1 Attribute Set and consists of all Bib-1 Attributes and additional specific GILS Use Attributes. Although GILS servers are required to support a minimal set of Use
Attributes, any of the GILS Use Attributes listed in Figure 2 could be used as access points for searching the GILS Locator Record. The exact manner by which the user constructs the query is an interface issue and not specified by the GILS Profile, but users supported by appropriate clients that understand the GILS Profile should be able to specify searches with each of the required Attributes. The extent of available access points or supported Attributes, however, will be implementation specific.

To ensure a minimal level of interoperability among the clients and servers, the GILS Profile requires that servers support a limited number of GILS Attributes (Figure 3 lists these required attributes). If a GILS server receives a query with any combination of these Attributes, it should process the query and never return any of the following diagnostic messages: “Unsupported Use Attribute,” “Unsupported Structure Attribute,” “Unsupported Relation Attribute,” or “Unsupported Attribute Type.” Table 1 displays the combinations of required GILS Attributes.

Retrieval in GILS: The GILS Schema, Element Set Names, and Record Syntaxes

As a GILS server completes a search, it produces a result set and makes that available to a client. The GILS server provides the client with the contents of selected records from the result set using the Z39.50 Present Service. The GILS server must respond to requests that records be presented in any of three Record Syntaxes mandated by the GILS Profile and one of the four Element Set Names specified by the GILS Profile. The exact manner in which records are presented to the user is an interface issue and not within the scope of the GILS Profile. There are three important aspects for retrieving GILS Locator Records: the GILS Schema, the Element Set Names, and Record Syntaxes.

A schema "represents a common understanding shared by the [client] and [server] of the information contained in the records of the database represented by the schema" (National Information Standards Organization, 1995, p. 132). The schema describes and/or defines an abstract record structure for a database record and a tagSet that uses tagTypes and tags to represent the elements in a database record. A schema can represent the hierarchical structure of database records such as the structure of GILS Locator Records.

Z39.50 defines two basic tagSets (tagSet-M and tagSet-G), and these contain elements commonly found in many database records. The GILS Schema uses tags from tagSet-M and tagSet-G, and it also defines a GILS tagSet for elements in the Locator Record that do not correspond to tags already defined in tagSet-M or tagSet-G. There are two general classes of elements in the GILS Schema: (1) Primitive—elements that cannot have locally defined subelements (2) Constructed—elements that can have one or more subelements, any of which may be well-defined or locally defined
Figure 2. GILS use attributes

by the record creator; string tags (i.e., text labels) identify locally defined subelements. Figure 4 presents a selection of the GILS tagSet.

Table 2 is a partial list of the GILS abstract record structure. The GILS tagSet identifies tags from tagSet-M with tagType 1 and tags from tagSet-G with tagType 2. GILS tags are identified with tagType 4. For
The GILS Use Attribute is listed followed by the GILS Use Attribute Number and the corresponding GILS Core Element names.

**Use Attributes:**
- Local Number (12; Local Control Number)
- Author-name Corporate (1005; Originator)
- Date/Time Last Modified (1012; Date of Last Modification)
- Record Source (1019; Record Source)
- Distributor Name (2001; Distributor Name)
- Index Terms—Controlled (2002; Index Terms—Controlled)
- Local Subject Index (29; Local Subject Term)
- Any (1016)

### Structure
- Word (2)
- URx (104)
- Date (5)
- Word List (6)

### Relation
- Greater than (5)
- Equal (3)

**Figure 3. Required GILS Attributes**

example, LOCAL CONTROL NUMBER is defined in tagSet-M as (1,14); TITLE is defined in tagSet-G as (2,1). ACCESS CONSTRAINTS is a specific GILS Locator Record element and is represented in the GILS schema as (4,53). ACCESS CONSTRAINTS is constructed of subelements including SECURITY CLASSIFICATION CONTROL, which is represented in the schema as (4,27). The schema shows the hierarchical structure of ACCESS CONSTRAINTS and SECURITY CLASSIFICATION CONTROL through the Tag Path of (4,53)/(4,27).

Schemas may be used with a particular record syntax, the Generic Record Syntax (GRS) (see below). Z39.50 implementations that use GRS-1 allow the client to request that the server return specific elements of the database record. Since the GILS Locator Record is represented by the abstract record structure in the GILS Schema, all elements are structured and identified for processing.

For example, a user has submitted a query looking for all Locator Records that describe information resources created by a particular agency (i.e., ORGINATOR). The server creates a result set. The client can then ask to have the result set records returned, and it can specify that the server should return in GRS-1 only the following:

- Control Identifier (4,1)
- Originator (4,52)
- Place Keyword (4,71)/(4,92)/(4,13)
- Place Keyword Thesaurus (4,71)/(4,92)/(4,14)

Being able to specify parts of the record provides the user with additional control over the information retrieval transaction.
MOEN/DISCOVERING, IDENTIFYING, & ACCESSING SPATIAL DATA

TABLE 1
RECOGNIZED AND SUPPORTED COMBINATIONS OF GILS ATTRIBUTES

<table>
<thead>
<tr>
<th>Use</th>
<th>Structure</th>
<th>Relation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Word</td>
<td>URx</td>
</tr>
<tr>
<td>Local Number</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Author-name Corporate</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Date/Time Last Modified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record Source</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Distributor Name</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Index Term—Controlled</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Local Subject Index</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 2
GILS ABSTRACT RECORD STRUCTURE (PARTIAL)

<table>
<thead>
<tr>
<th>Tag Path</th>
<th>Element</th>
<th>Mandatory?</th>
<th>Repeatable?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,10)</td>
<td>rank</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(1,12)</td>
<td>url</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(1,14)</td>
<td>local control number</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>(1,16)</td>
<td>dateOfLastModification</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>(2,1)</td>
<td>title</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>(4,1)</td>
<td>controlIdentifier</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>(2,6)</td>
<td>abstract</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>(4,51)</td>
<td>purpose</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>(4,52)</td>
<td>originator</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>(4,53)</td>
<td>accessConstraints</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>(4,53)/(4,25)</td>
<td>generalAccessConstraints</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>(4,53)/(4,26)</td>
<td>originatorDisseminatorControl</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(4,53)/(4,27)</td>
<td>securityClassificationControl</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(4,71)</td>
<td>spatialDomain</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(4,71)/(4,91)</td>
<td>boundingCoordinates</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(4,71)/(4,91)/(4,9)</td>
<td>westBoundingCoordinate</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(4,71)/(4,91)/(4,10)</td>
<td>eastBoundingCoordinate</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(4,71)/(4,91)/(4,11)</td>
<td>northBoundingCoordinate</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(4,71)/(4,91)/(4,12)</td>
<td>southBoundingCoordinate</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(4,71)/(4,92)</td>
<td>place</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>(4,71)/(4,92)/(4,13)</td>
<td>placeKeyword</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>(4,71)/(4,92)/(4,14)</td>
<td>placeKeywordThesaurus</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

The GILS Profile defines four Element Set Names. When a client submits a request asking the server to return a record using an Element Set Name, the server returns records with a specific set of elements from the database record. The four GILS Profile Element Set Names and the elements contained in each are:

- Element Set Name "B" contains at least Title, Control Identifier, Originator, and Local Control Number
<table>
<thead>
<tr>
<th>Primitive Elements</th>
<th>Recommended Datatype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tag</td>
<td>Element</td>
</tr>
<tr>
<td>1</td>
<td>controlIdentifier</td>
</tr>
<tr>
<td>9</td>
<td>westBoundingCoordinate</td>
</tr>
<tr>
<td>10</td>
<td>eastBoundingCoordinate</td>
</tr>
<tr>
<td>11</td>
<td>northBoundingCoordinate</td>
</tr>
<tr>
<td>12</td>
<td>southBoundingCoordinate</td>
</tr>
<tr>
<td>13</td>
<td>placeKeyword</td>
</tr>
<tr>
<td>14</td>
<td>placeKeywordThesaurus</td>
</tr>
<tr>
<td>25</td>
<td>generalAccessConstraints</td>
</tr>
<tr>
<td>26</td>
<td>originatorDisseminatorControl</td>
</tr>
<tr>
<td>27</td>
<td>securityClassificationControl</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Constructed Elements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tag</td>
<td>Element</td>
</tr>
<tr>
<td>51</td>
<td>purpose</td>
</tr>
<tr>
<td>52</td>
<td>originator</td>
</tr>
<tr>
<td>53</td>
<td>accessConstraints</td>
</tr>
<tr>
<td>71</td>
<td>spatialDomain</td>
</tr>
<tr>
<td>91</td>
<td>boundingCoordinates</td>
</tr>
<tr>
<td>92</td>
<td>place</td>
</tr>
</tbody>
</table>

Figure 4. GILS tagSet (Partial)

- Element Set Name “G” contains all B Element Set elements and Cross Reference
- Element Set Name “W” contains all B Element Set elements and bodyOfDisplay
- Element Set Name “F” contains all elements available in the record.

A GILS server should return to the client all of the elements specified by the Element Set Name if there are data available in the database record. In some cases it may not be possible to encode certain data in the requested record syntax (e.g., some types of locally defined binary data may
SPATIAL DOMAIN (Optional, Not Repeatable): This element is a grouping of subelements that together provide the geographic areal domain of the data set or information resource. Geographic names and coordinates can be used to define the bounds of coverage. Although described here informally, the spatial object constructs should be as defined in FIPS 173, "Spatial Data Transfer Standard."

BOUNDING COORDINATES (Optional, Not Repeatable): This subelement limits the coverage of a data set expressed by latitude and longitude values in the order western-most, eastern-most, northern-most, and southern-most. For data sets that include a complete band of latitude around the earth, the West Bounding Coordinate shall be assigned the value: -180.0, and the East Bounding Coordinate shall be assigned the value: 180.0. The following subelements comprise the Bounding Coordinates:

- WEST BOUNDING COORDINATE: Western-most coordinate of the limit of the coverage expressed in longitude. Domain: -180.0 <= WestBounding Coordinate <= 180.0
- EAST BOUNDING COORDINATE: Eastern-most coordinate of the limit of coverage expressed in longitude. Domain: -180.0 <= East Bounding Coordinate <= 180.0
- NORTH BOUNDING COORDINATE: Northern-most coordinate of the limit of coverage expressed in latitude. Domain: -90.0 <= North Bounding Coordinate <= 90.0; North Bounding Coordinate >= South Bounding Coordinate
- SOUTH BOUNDING COORDINATE: Southern-most coordinate of the limit of coverage expressed in latitude. Domain: -90.0 <= South Bounding Coordinate <= 90.0; South Bounding Coordinate <= North Bounding Coordinate

PLACE (Optional, Repeatable): This subelement identifies geographic locations characterized by the data set or information resource through two associate constructs:

- PLACE KEYWORD: The geographic name of a location covered by a dataset or information resource.
- PLACE KEYWORD THESOURUS: The name of a formally registered thesaurus or similar authoritative source of Place Keywords.

Figure 5. Spatial MetaData in GILS locator records

not be encodable in a USMARC or SUTRS record). In those cases, the data in those elements will not be returned.

A record syntax is a way of representing the retrieval record to return from the server to the client. Record syntaxes allow a server to return GILS Locator Records to a client in a format that the client can process. The GILS Profile requires a GILS server to support three record syntaxes: (1) Generic Record Syntax (GRS-1), (2) Simple Unstructured Text Record Syntax (SUTRS), and (3) USMARC. GRS-1 representation is considered the complete and canonical form since GRS-1 can use the structure provided in the GILS Schema to identify and represent both the well-known and locally identified elements.

SUTRS allows relatively simple clients to accept and display GILS Locator Records to users. SUTRS treats the GILS Locator Record as a block of text with no structure or content designation, and the client should not expect to be able to parse the record to obtain any individual GILS
elements. The client simply presents the record as it is provided by the server. The GILS Profile does suggest a preferred display format for use with SUTRS, but the display format is a concern only of the server and not the client.

USMARC is required since the library community will be a primary user community for GILS. The GILS Profile provides a mapping between the GILS Core Elements and USMARC for use with the USMARC record syntax.

The foregoing has described the basic components of the GILS Profile. Understanding the profile requires a basic understanding of Z39.50, how it works, and the terminology used in the standard.

**GILS AND SPATIAL INFORMATION RESOURCES**

An important set of information resources created, collected, or held by federal agencies relate to spatial data. Like other federal information resources, spatial information resources can be described by GILS Locator Records. GILS will be a useful mechanism for discovering spatial information resources throughout the federal government. Many federal agencies create or collect spatial data. The membership of the Federal Geographic Data Committee (FGDC) indicates the range of agencies involved with spatial data. Each of the Executive Branch agencies that are covered by OMB Bulletin 95-01 are required to create GILS Locator Records for their information resources including those related to spatial data. GILS will enable users to search across these agencies’ Locator Records to discover and locate spatial information resources.

GILS will assist in achieving the goals of OMB Circular A-16 (1990), *Coordination of Surveying, Mapping, and Related Spatial Data Activities*. Those goals include the reduction of duplication and reduction of expense of developing geographic data. For example, a federal agency could use GILS to discover if there are existing spatial information resources collected by other agencies before initiating new projects that might duplicate existing data sets.

Gluck (1994) identifies three primary elements that characterize spatial data selection:

- **Attributes**, which define the contents and characteristics of spatial data
- **Time**, which can provide a time-scale of the coverage of the spatial data
- **User task**, which determines the appropriateness of components of a particular data set or other information resource to solve a particular problem (p. 640).

GILS Locator Records begin to address these selection criteria since metadata about spatial information resources are accommodated by spe-
sific GILS Core Elements. Figure 5 gives the GILS Core Elements that provide spatial metadata along with their semantics as found in the GILS Profile (note that the names and semantics of these elements reflect changes in a revised version of the GILS Profile currently under development as of April 1995).

Appendix A contains two examples of GILS Locator Records that describe spatial information resources. These records, created by the United States Geological Survey, describe:

- National Digital Cartographic Data Base—Large Scale (Record #1)
- Landsat Earth Resources Observations Multispectral Scanners Imagery (Record #2).

The following discussion refers to these records to indicate how a user might search, retrieve, and use these records.

GILS provides the possibility for users to search Locator Records using the spatial metadata elements. The GILS Profile identifies Use Attributes for the GILS Core Elements listed in Figure 5. As one hypothetical example, these Use Attributes could be used to retrieve Record #1 in Appendix A. A user would submit a query to a GILS server with one or more of the Use Attributes West Bounding Coordinate, East Bounding Coordinate, North Bounding Coordinate, South Bounding Coordinate with terms of -179, -66, 72, 24 respectively and a Relation Attribute of “equal” to retrieve the example Record #1. A more likely scenario, however, would be for a user to submit a query with these Use Attributes and a range of bounding coordinates to find GILS Locator Records that describe resources that have a spatial coverage the user is interested in.

It is important to realize, however, that these spatial metadata Use Attributes are not included in the minimum set of Use Attributes required by the GILS Profile. Therefore, this specific searching capability on spatial data fields may or may not be offered by all GILS conformant servers. There is nothing in the GILS Profile, however, that precludes agencies or other organizations from implementing a GILS server that provides robust searching of the spatial metadata elements in GILS Locator Records. An intermediary user of GILS could add value to GILS Locator Records by mounting them on an information system that provides specialized spatial data searching capabilities (i.e., ensures that the SPATIAL DOMAIN elements are access points for the Locator Records).

Record #2 in Appendix A describes a spatial information resource that is accessible via the World Wide Web protocol. This points to another important feature of GILS. If a user retrieves this GILS Locator Record and is interested in accessing the spatial information resource described
in the Locator Record, it is possible to use information in the record to automatically link to the resource. Under the AVAILABILITY element, there are two subelements, LINKAGE and LINKAGE TYPE. Record #2 contains a Uniform Resources Locator (URL) in the LINKAGE element (i.e., http://sunl.cr.usgs.gov:80/glis/glis.html). Appropriate client software would be able to use the LINKAGE information to start up a WWW browser and connect to the spatial information resource pointed to by the Locator Record. Such capabilities imply that the user’s client will need to support multiple protocols in addition to Z39.50 (e.g., http, FTP, telnet, gopher). This scenario of seamless network navigation (i.e., discovery, searching, retrieval, and access) that GILS makes possible is a most exciting prospect.

A current initiative will extend the use of Z39.50 for information retrieval of spatial data and will work in concert with GILS. Similar to the GILS Profile, work is now underway to develop a Z39.50 application profile for the content specification of digital geospatial metadata. The GEO Profile development is being coordinated by Douglas Nebert at USGS in cooperation with ASTM Section D18.01.05 on Mapping and GIS for the FGDC (contact Nebert for additional information <ddnebert@usgs.gov>). The GEO Profile supports search and retrieval of geospatial metadata entries and related geospatial data sets accessible on GEO conformant servers. Implementations of the GEO Profile will provide Z39.50 access to existing or new data sets. The implication of this for GILS is that a Z39.50 client that supports both the GILS Profile and the GEO Profile will be able first to search GILS Locator Records and discover particular spatial information resources described by GILS Locator Records and then be able to establish a Z39.50 connection to the actual data sets and use Z39.50 functionality to do sophisticated searching on the data set. The GEO Profile is an important complement to the GILS.

GILS Core Elements provide metadata information for a broad range of information resources including spatial information resources. The spatial metadata elements in GILS Locator Records will assist users of spatial data to discover and select spatial resources that are appropriate to their information problems.

SUMMARY AND CONCLUSION

This paper has provided an introduction and overview to the GILS, the new federal initiative to improve access to, and management of, government information resources. Z39.50 is an important foundation for a decentralized, agency-based, and Internet-accessible GILS. The discussion of Z39.50 and how it is used in GILS demonstrates the utility of using an information retrieval protocol for the discovery of, and access to, federal information resources.
Federal agencies have begun to create GILS Locator Records and are making plans to install Z39.50 access to the information servers and databases that contain the Locator Records. The various user communities that have a stake in accessing federal information resources, especially spatial information resources, will want to encourage agencies to move ahead aggressively in establishing GILS servers and to provide robust searching and retrieval capabilities on those servers. Z39.50 is rich in functionality for searching and retrieving information, and any limits on how users will be able to search GILS Locator Records will likely be a function of the database management and search engines deployed by agencies. Users will want to state their needs to agencies, and press agencies especially for the capabilities of searching the spatial metadata elements in GILS Locator Records.

The OMB Bulletin 95-01 that establishes GILS sets a number of deadlines for agencies. By December 31, 1995, agencies are to have created their initial GILS Core Locator Records. By the same date, these records are to be available online in a form compliant with the GILS Profile. The success of this important policy initiative to provide better public access to government information will depend on the federal agencies complying with the OMB Bulletin. The spatial data user community is an important stakeholder and has the ability to press federal agencies to establish GILS servers so that the community can discover, identify, and locate the wealth of spatial information resources available through the federal government.
APPENDIX A

SAMPLE GILS RECORDS

This appendix contains two GILS Locator Records created by the United States Geological Survey. They describe spatial information resources.

RECORD #1

Title: NATIONAL DIGITAL CARTOGRAPHIC DATA BASE-LARGE SCALE
Acronym: NDCDB/DLG
Originator: USGS/NMD
Local-Subject-Index: AEDD; ALASKA; ALASKA DIRECTORY; ARCTIC; CARTOGRAPHY; DLG; DOIGC; GEODATA; MAP; USGS; ESDD; U.S. Federal GILS

Abstract: Contained are selected US regional coverage of planimetric map features and contours. Also, included are digitized source-map scales varying from 1:24,000 to 1:62,500. The largest amount of coverage exists for the Public Land Survey System (PLSS) and boundary data categories. Data are added on the basis of mapping requirements of the USGS and other Federal agencies. The information is stored in topologically structured DLG-3 format. This portion of the Digital Cartographic Data Base is referred to as US GeoData. The data are distributed through the National Mapping Division. Contact the Earth Science Information Center for documentation, availability of specific coverage, output formats, and current price. The data are not available for general online access.

Format: DIGITAL DATA SETS

Spatial-Domain:
   Geographic-Coverage: UNITED STATES
   Coordinate-System: UTM NORTHTINGS AND EASTINGS; ARBITRARY X, Y COORDINATES
   Coverage-Description: NONE REPORTED
   Bounding-Coordinates:
      West-Bounding-Coordinate: -179
      East-Bounding-Coordinate: -66
      North-Bounding-Coordinate: 72
      South-Bounding-Coordinate: 24

Time-Period:
   Time-Period-Textual: 1979-PRESENT

Availability:
   Distributor:
      Name: USGS/NMD
      Organization: USGS/NMD
      Street-Address: U.S. GEOLOGICAL SURVEY, RESTON-ESIC, 507 NATIONAL CENTER
      City: RESTON
      State: VA
      Zip-Code: 22092
      Country: USA
      Telephone: (703)860-6045
Resource-Description: NATIONAL DIGITAL CARTOGRAPHIC DATA BASE-LARGE SCALE

Order-Process: Contact any below-listed USGS Earth Science Information Center (ESIC) for assistance: Anchorage-ESIC, 4230 University Dr., Rm 101, Anchorage, AK 99508-4664 (907) 786-7011; Anchorage-ESIC, U.S. Courthouse, Rm 113, 222 W. 7th Ave., #53, Anchorage, AK 99513-7546 (907) 271-4307; Denver-ESIC, 169 Federal Bldg., 1961 Stout St., Denver, CO 80225-0046 (303) 844-4169; Lakewood-ESIC, Box 25046, Federal Ctr., MS 504, Denver, CO 80225-0046 (303) 236-5829; Rolla-ESIC, 1400 Independence Rd., MS 231, Rolla, MO 65401 (314) 341-0851; Salt Lake City-ESIC, 8105 Federal Bldg., 125 S. State St., Salt Lake City, UT 84138 (801) 524-5652; San Francisco-ESIC, 504 Custom House, 555 Battery St., San Francisco, CA 94111 (415) 556-5627; Spokane-ESIC, 678 U.S. Courthouse, W. 920 Riverside Ave., Spokane, WA 99201 (509) 353-2524; Stennis Space Ctr-ESIC, Bldg. 3101, Stennis Space Center, MS 39529 (601) 688-3544; Washington, D.C.-ESIC, Dept. of the Interior Bldg., 18th & C Sts., NW, Rm. 2650, Washington, D.C. 20240 (202) 343-8073; Menlo Park-ESIC, Building 3, MS 532, 345 Middlefield Rd., Menlo Park, CA 94025 (415) 329-4309; Reston-ESIC 507 National Center Reston, VA 22092 (703) 860-6045

Technical-Prerequisites:
Data-Set-Type: AUTOMATED
Computer-Type: AMDAHL 5890
Computer-Location: RESTON, VA

Access-Constraints: Access is not restricted unless otherwise noted.

Documentation: Standards for Digital Line Graphs (6 parts), Br. OF Technical Mgmt., MS 510, USGS, Reston, VA. Digital Line Graphs from 1:24,000-scale Maps, Data Users Guide 1, 1986, USGS.

Use-Constraints: These data and information have been approved for release by the Director of the USGS on condition that neither the USGS nor the United States Government may be held liable for any damages resulting from its authorized or unauthorized use.

Status: OPERATIONAL

Point-of-Contact:
Organization: USGS/NMD
Street-Address: U.S. GEOLOGICAL SURVEY, RESTON-ESIC, 507 NATIONAL CENTER
City: RESTON
State: VA
Zip-Code: 22092
Country: USA
Telephone: (703) 860-6045

Purpose: These data and information resources contribute to the USGS mission of Earth science in the public service and the USGS role as the principal source of earth-science research and information for the Nation.

Control-Identifier: USGS0024
Record-Source: USGS/NMD
Date-of-Last-Modification: 9008
**RECORD #2**

**Title:** LANDSAT EARTH RESOURCES OBSERVATIONS MULTISPECTRAL SCANNERS IMAGERY  
**Acronym:** LANDSAT DATA (MSS)  
**Originator:** USGS/NMD

**Abstract:** A Multispectral Scanner (MSS) has flown on board five Landsat satellites to date. Landsat 1, 2 and 3 operated in a circular, sun-synchronous, near-polar orbit at an altitude of approximately 913 km (567 miles), with a nominal 9:30 AM crossing of the Equator during the descending mode. They circled the Earth every 103 minutes, completing 14 orbits per day and viewing the entire Earth every 18 days. The Landsat orbits are selected and trimmed so that each satellite ground track repeats its Earth coverage at the same local time every day. Landsat 4 and 5 also operate in circular, sun-synchronous, near-polar orbit at an altitude of 705 km (438 miles), with a nominal 9:45 AM crossing at the Equator during the descending mode. Each orbit takes nearly 99 minutes, and the spacecrafts complete just over 14 orbits each day, covering the entire Earth (poles excluded) every 16 days. An international Landsat database exists that contains information on the available MSS data acquired by independent international ground receiving stations. These data are directly downlinked to ground receiving stations when in transmission range of the satellite(s). The framing of each scene is based on the World Wide Reference System (WRS), a network of intersecting paths and rows whose junctions define the nominal scene center of each Landsat scene. The WRS path represents the nominal satellite track, a maximum of 251 paths exist for Landsat 1-3 and 253 paths exist for Landsat 4 and 5 data. The WRS row indicator represents scene centers that are chosen at 23.92 second increments along the orbital track, a total of 248 row numbers exist. MSS imagery is available on microimage fiche, otherwise in late 1994, MSS browse imagery is scheduled to be available on GILS. This MSS browse imagery may be viewed via the EDC World Wide Web home page, which allows direct access to GLIS. To date, sixteen unique ground stations have acquired and archived MSS data on a wide variety of media. EROS Data Center periodically receives international database tape catalogs of foreign stations’ data holdings which are added into an on-line Landsat global database (i.e., International Landsat Database).

**Format:** MICROFILM; PHOTOGRAPHS; DIGITAL DATA SETS AND BROWSE

**Spatial-Domain:**
- Geographic-Coverage: GLOBAL
- Coverage-Description: NONE REPORTED

**Time-Period:**
- Time-Period-Textual: 1972-1993

**Availability:**

**Distributor:**
- Name: USGS/NMD
- Organization: USGS/NMD
- Street-Address: EROS DATA CENTER, U.S. GEOLOGICAL SURVEY
- City: SIOUX FALLS
- State: SD
- Zip-Code: 57198
MOEN/Discovering, Identifying, & Accessing Spatial Data

Country: USA  
Telephone: (605)594-6151  
Fax: (605)594-6589

Resource-Description: LANDSAT EARTH RESOURCES OBSERVATIONS MULTISPECTRAL SCANNERS IMAGERY

Order-Process: To place orders, obtain additional information, technical details, ancillary products, and pricing schedules regarding products and services, or international data holdings contact the EROS Data Center, Customer Services Section.

Technical-Prerequisites:
  - Data-Set-Type: AUTOMATED
  - Access-Method: INTERACTIVE AND BATCH
  - Number-of-Records: 650,000
  - Bytes-Per-Record: 154
  - Computer-Type: IBM
  - Computer-Location: SIOUX FALLS, SD


Access-Constraints: Access is not restricted unless otherwise noted.


Use-Constraints: These data and information have been approved for release by the Director of the USGS on condition that neither the USGS nor the United States Government may be held liable for any damages resulting from its authorized or unauthorized use.

Status: OPERATIONAL

Point-of-Contact:
  - Name: CUSTOMER SERVICES
  - Organization: USGS/NMD
  - Street-Address: EROS DATA CENTER, U.S. GEOLOGICAL SURVEY
  - City: SIOUX FALLS
  - State: SD
  - Zip-Code: 57198
  - Country: USA
  - Telephone: (605)594-6151
  - Fax: (605)594-6589

Purpose: These data and information resources contribute to the USGS mission of earth science in the public service and the USGS role as the principal source of earth-science research and information for the Nation.

Control-Identifier: USGS2020

Record-Source: USGS/NMD

Date-of-Last-Modification: 9410
REFERENCES


