Evaluating Geographic Information Systems Technology

Stephen C. Guptill
U. S. Geological Survey, 521 National Center, Reston, VA 22092

ABSTRACT: Computerized geographic information systems (GISs) are emerging as the spatial data handling tools of choice for solving complex geographical problems. However, few guidelines exist for assisting potential users in identifying suitable hardware and software. A process to be followed in evaluating the merits of GIS technology is presented. Related standards and guidelines, software functions, hardware components, and benchmarking are discussed. By making users aware of all aspects of adopting GIS technology, they can decide if GIS is an appropriate tool for their application and, if so, which GIS should be used.

INTRODUCTION

ORGANIZATIONS WITHIN ALL SEGMENTS of our society must quickly respond to complicated problems involving a wide variety of geographically referenced data sets (such as natural resource, socioeconomic, epidemiologic, or facilities management data). Administrative and regulatory responsibilities assigned to government agencies are placing tremendous pressure on existing information delivery systems. Traditional methods of acquiring, storing, and analyzing spatially referenced data are proving to be too costly and inflexible in meeting these growing needs. Computerized geographic information systems (GISs) are emerging as the spatial data handling tools of choice for solving complex geographical problems.

A GIS is a system of computer hardware and software designed to allow users to collect, manage, and analyze large volumes of spatially referenced data and associated attributes. Because GIS technology allows analysts to process and interrelate many more kinds of data than were previously feasible, GIS users have the potential to greatly improve traditional missions, such as data collection, research, assessment, and information delivery. However, potential users of a GIS should understand all aspects of adopting this technology. Is a GIS an appropriate tool for your application, and if so, which GIS should be used? To help answer this question, this paper describes a process for evaluating the technology.

THE UTILITY OF GIS TECHNOLOGY

Spatial data analysis is a multidisciplinary concern. Geographic, medical, sociological, military, urban planning, and Earth science activities, among others, require spatial data analysis. Spatial data sets are frequently heterogeneous, having data elements such as soils, land use, and population statistics, and are often comprised of data sources with differing scales, coordinate systems, accuracies, and areal coverage. The data will originate from source material in multiple formats, such as text, maps, charts, or remotely sensed imagery. In some instances, information is extracted from spatial data sets for incorporation into multidimensional numeric models. The management and analysis of spatial data in a hard-copy environment to support such modeling and other analytical procedures is often tedious and cumbersome, inhibiting the efficient achievement of goals.

Spatial data sets are unique in providing the geographic locations of features, related to known coordinate systems; in specifying attributes of features that may be independent of location, such as color, cost, and size; and in describing the spatial and topological relations among features in the data set. GISs are specifically designed to manage and analyze spatial data sets with such characteristics.

At present, a number of GIS packages have been developed within the public and private sectors. Some packages are application specific while others are more generic in nature. Development within the GIS field is continuing and more GIS packages will be available in the future. Users will often be able to apply existing systems directly to their applications, eliminating the need for costly system development work. This trend is changing the question of GIS implementation from one of developing a system to meet user requirements, to one of selecting the best existing system or combination of systems that meets user needs.

CHARACTERISTICS OF GIS

The term GIS typically has been quite loosely applied (Cowen, 1988; Parker, 1988). There is some agreement, however, that GIS can be defined in terms of general functional characteristics. The functional components of a GIS can be grouped into five broad categories: user interface, system/database management, database creation/data entry, data manipulation and analysis, and display and product generation.

The *user interface* is the method by which the human operator communicates with the various database and GIS application modules. The user interface consists of software capabilities that simplify and organize the interaction between the user and the GIS software (e.g., menus, help screens, and graphic displays).

The *database management component* provides the environment within which the GIS functions and the means by which the data are controlled. The system management environment is furnished by the operating system of the host computer. GIS database functions parallel those of a nonspatial database management system (DBMS), but with extensions beyond the addition, deletion, revision, and Boolean retrieval capabilities of a standard DBMS. The GIS DBMS contains hardware and software facilities for the storage, retrieval, and update of spatial information (in both alphanumeric and digital graphic forms), and incorporates storage structures to minimize data redundancy and to aid spatial searches. In addition, the GIS DBMS (like nonspatial DBMS) must have the file management capabilities to handle a potentially large archive of data files.

Database creation/data entry refers to the process of bringing data into the electronic environment of the GIS. A GIS database is often conceptualized as a series of thematic categories or topics (sometimes termed layers) of information held within the database (Figure 1). These layers may contain information that has been captured from aerial photography, remote sensing satellites, conventional maps, or other sources.

Data entry is the process of loading data into a GIS database. Data in a computer-compatible format (such as digital remotely sensed data) can be loaded directly. A database may also be

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 55, No. 11, November 1989, pp. 1583-1587.

0099-1112/89/5511–1583$02.25/0

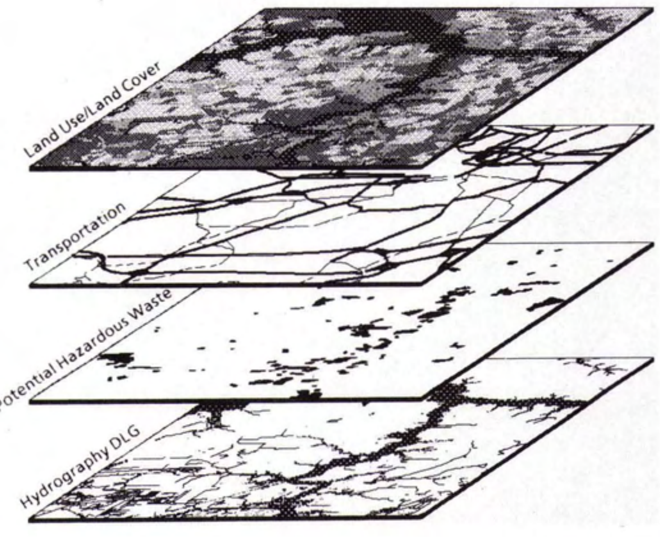

FIG. 1. Concept of separation of categories within a GIS.

created by digitizing, or scanning, maps or by digitizing information on aerial photographs to create a computer-readable data set. This effort is sometimes referred to as data capture (Guptill, 1985; Chrisman, 1987). Two types of data are generally collected: locational or geographic, and accompanying feature attribute data. Locational information is usually digitized from existing graphics, maps, or images. Attributes identify what the features represent in the form of numeric or textual information (e.g., a soil type, feature name, or road class).

Costs to create the data sets to use in a GIS are far in excess of the costs of the hardware and software. In creating data sets for the GIS, data encoding schemes, topological data structuring, attributing of information, and selection of file structures must all be performed in the correct manner to support the user's applications. Failure to capture the correct data, in the correct form, with adequate attributes could result in the GIS being unable to support its intended users.

Emphasis on the sharing of data, or the use of preexisting databases by users of GIS, is increasing. The maintenance of digital data sets of national extent by Federal agencies, such as the U.S. Geological Survey, Bureau of the Census, and the Soil Conservation Service (Federal Interagency Coordinating Committee on Digital Cartography, 1987) is inducing progress in the improvement of data translators and the establishment and use of digital data standards. As a result of these activities, the cost of database creation may gradually decline. The emphasis will then shift to updating and revising the data and to the addition of specialized data themes.

If spatial information capture and management are performed properly, the accuracy and the original meaning of the data will be maintained. The GIS user may then proceed to the primary GIS activities, *data manipulation and analysis*. Spatial analysis tools are used to model, make predictions, and reach conclusions about problems of interest. Such analysis involves combining data from multiple spatial data categories and performing analytical, statistical, measurement, and other operations on the GIS data sets to transform the data into information suitable for a given application. Spatial analysis techniques include compositing areas, and performing proximity searches, topographic analysis, and clustering and aggregation operations. These operations are ideally performed in an interactive mode on the spatial component of the data, as well as on the nonspatial attribute data. Such operations can range in complexity from simple Boolean queries to reclassification and creation of entirely new map displays.

The typical GIS has extensive capabilities for *display and product generation*. Maps, charts, graphs, and tables resulting from the use of its analysis and modeling capabilities can be produced. The displays used will range in complexity from tabular reports and simple monochrome plots to publication quality three-dimensional color graphics.

EFFECTIVE USE OF GIS

Successful GIS implementation and application requires an organization's personnel to be cognizant of the capabilities and limitations of GIS technology and to carefully evaluate the needs of system users and applications. The variety of possible GIS applications and users makes it impractical and inappropriate to provide strict criteria for GIS implementation.

As GISs become more widely implemented, their procurement and operation are being more closely monitored for adherence to agency or company information resource management guidelines and procedures. Such scrutiny may be a new experience for GIS program managers who may have first acquired their GIS capability as part of a low-visibility research effort. Extensive guidelines and procedures exist on the topic of information resources management. For example, the General Accounting Office (GAO) has identified the following five basic objectives for the acquisition and operation of information systems (U.S. General Accounting Office, 1986):

- *Ensure System Effectiveness*: System effectiveness is measured by determining whether the system performs the intended functions and whether users get the information needed, in the right form, in a timely fashion.
- *Promote System Economy and Efficiency*: An economical and efficient system uses the minimum number of information resources to achieve the output level the system's users require.
- *Protect Data Integrity*: Data integrity requires that systems have adequate controls over how data are entered, communicated, processed, stored, and reported.
- *Safeguard Information Resources*: Information resources, which include hardware, software, data, and people, need to be protected against waste, loss, unauthorized use, and (or) fraud.
- *Comply with Laws and Regulations*: Compliance with laws, regulations, policies, and procedures that govern the acquisition, development, operation, and maintenance of information systems must be ensured.

Recommendations, such as those contained within the GAO document, should be observed to ensure that a given GIS implementation will comply with an organization's various guidelines and procedures should a review or audit occur.

USER REQUIREMENTS ANALYSIS

GISs are successful when they comprehensively and consistently meet users needs. Development of a successful GIS depends on well-defined user requirements. A user requirements analysis (URA) is a detailed study of the needs of potential system users. The URA should result in a clear statement of end-product characteristics, required production rates, estimated data volumes, and cost/benefit rationale (Stefanovič and Drummond, 1987). Steps in performing the URA include

- Identification of users
- Definition of required products
- Evaluation of work flow
- Estimate of database development efforts
- Inventory of user applications
- Refinement of GIS product characteristics
- Calculation of necessary production rates
- Estimate of data volumes
- Performance of a cost/benefit analysis

The personnel conducting the URA should prepare a report for the organization management. This report should clearly, and in detail, identify the

- Operation, users, and data requirements of the existing system
- Potential uses and users of the GIS
- Digital and hard-copy products required by the users
- Data volumes and production rates the GIS would be required to meet
- Database required to support GIS implementation
- Cost/benefit analysis

The URA report provides managers with a comprehensive description of the data sets, database management capabilities, modeling and analytical needs, and product generation requirements for successful GIS implementation. On the basis of this information the manager must weigh the merits of GIS usage relative to the organization's applications.

The URA cost/benefit analysis is one indicator of the need for GIS technology within the organization. The degree to which intangible benefits were adequately measured and quantified must be considered in a subjective manner when the cost/benefit ratio is evaluated (Gohagan, 1980). When the cost/benefit ratio is marginal, that is close to 1.0, further research may be required, particularly in the intangible benefits, before a decision on GIS implementation can be made on the basis of the cost/benefit ratio.

The manager must also consider the appropriateness of applying GIS technology to the organization's applications regardless of the URA results. The URA documents how an existing organization functions and supports its users and then quantifies and defines an alternative operation on the basis of GIS technology. As outlined, the URA does not address related issues, such as the organization's objectives, goals, or staff capabilities, that are not quantifiable technical issues. Such factors could override conclusions based on solely technical considerations.

SELECTION OF A SYSTEM

If, upon completing a URA and assessing the appropriateness of GIS applications for the organization's missions, it is determined that GIS technology should be incorporated into the operation, evaluation criteria must be devised to serve as the basis for selecting a specific GIS. Evaluation criteria must be clearly specified in order for both the organization and the vendors to have a clear understanding of what is requested and what is required. The evaluation criteria should be incorporated into the specifications used in conducting benchmark testing (Tomlinson and Boyle, 1981).

Hands-on experience with GIS capabilities is valuable in developing reasonable evaluation criteria. A useful method of acquiring such experience is to perform a small-scale pilot project. The pilot project should be designed to test the ability of GIS systems to meet an organization's operating needs. Pilot tests are a source of realistic data on production rates, memory and storage requirements, human interface functionality, and user response to GIS products (Goodchild and Rizzo, 1987).

All or only portions of the GIS designed for the organization may be tested, depending on the organization's familiarity with GIS systems. Testing only the portions of the GIS that are most critical to organizational needs, or that represent the elements with which the organization is least familiar, may be desirable. Pilot test results can be used to refine evaluation criteria that were derived from the URA.

STANDARDS AND GUIDELINES

Standards facilitate integration of GIS technology, not only with other GISs but also with other information systems, through compatibility in data administration, database management, graphics, hardware, and software. Such standardization augments the functionality, flexibility, and productivity of a GIS while extending its availability to a greater audience. Interface standards enable communication between information systems; they include data interchange, database conversion, graphics, software, and hardware standards.

Many of the standards and guidelines for general computer systems and information management are applicable to a GIS. A set of GIS related standards, guidelines, de facto and developing standards, and references is given in Appendix A of Guptill (1988). Special mention should be made of "The Proposed Standard for Digital Cartographic Data." This proposed standard was published as a special edition of *The American Cartographer* (Volume 15, No. 1, January, 1988). Copies of the current version of the document are available from the U.S. Geological Survey (Digital Cartographic Data Standards Task Force, 1988). It consists of four major components: definitions and references, spatial data transfer, digital cartographic data quality, and cartographic features. The standard is an attempt to address the recognized requirement for easy transfer of spatial data from one spatial data handling system to another, with both systems possibly residing on computer hardware and operating system software of different makes.

TECHNICAL ISSUES

Defining a set of processing functions to meet specific application needs is an important step in the design or evaluation of a GIS and is a direct outgrowth of the URA phase. Identifying required functions often begins with a detailed list of required GIS products and their specifications. Further analysis leads to identifying the type of processing functions required to produce each product.

To assist in identifying required functions, a checklist of GIS software functional components is given in Guptill (1988). An example of the entries for Database Creation is given in the appendix. It should be recognized that no single GIS has all of the capabilities listed.

Individual processing functions are often prioritized as either mandatory or desirable capabilities. Mandatory software capabilities, when merged with parameters that quantify specific application needs, such as required response time, accuracy, precision, product generation frequency, and data volumes, lead directly to mandatory hardware capabilities.

Existing GISs are extremely diverse both in functionality and database structure (Smith *et al.*, 1987). Systems use various methods for digitizing, assigning, and storing attribute, coordinate, and topological information. The capability to manipulate, analyze, and display these data varies widely among systems. Capabilities of a given system are often oriented toward providing a specific capability or supporting a specific application area, such as computer-aided design, computer-aided mapping, surveying natural resource management, terrain analysis, and (or) image processing. Users must try to find the best match of system capabilities with user requirements. Benchmark testing can provide information to determine the best match.

Benchmarking is a process in which computer systems such as a GIS are tested for functionality and performance. Benchmarking is accepted as part of the acquisition process by both private industry and the Federal Government. The success of benchmarking as an evaluation technique depends on the extent to which benchmark tests can be constructed that are representative of expected workloads. Ten procedural steps for benchmark construction, as described in FIPS Publication 75, *Guidelines on Constructing Benchmarks for ADP System Acquisition* (U.S. National Bureau of Standards, 1980), are

Step 1. Define benchmarking objectives and complete preliminary activities (such as defining an agency's service, operational, and workload requirements).

Step 2. Quantify the present workload requirements.
Step 3. Survey users to obtain information on present applications and user forecasts of new or changing applications.
Step 4. Forecast future workload requirements.
Step 5. Categorize future workloads. Total workload is partitioned into distinct categories.
Step 6. Determine the relative contribution of each category.
Step 7. Scale each category (weight the running times for each category's set of benchmark problems according to its contribution).
Step 8. Represent workload categories with benchmark problems. Select real or synthetic programs that represent the workload categories identified in step 5.
Step 9. Fine tune each benchmark mix on the present system.
Step 10. Prepare the benchmark package (the documentation of the benchmark mix and the rules for the live test demonstration) and test the benchmark.

Benchmarking should be done following required standards, guidelines, and practices (see also U.S. National Bureau of Standards, 1977).

A key point is that the benchmarking must be done on products and capabilities identified through a URA and must reflect estimated future data volumes. Users must also recognize that the preparation of a comprehensive benchmark may be a major task. The user must weigh factors such as the size of the procurement against the effort required to conduct benchmarking to determine the appropriate level of benchmark testing.

The potential GIS user must have answers to the following technical questions:

- Does the system have the software functions required to perform the necessary applications?
- Will the hardware components perform these functions in an efficient fashion and provide a path for growth with respect to data quantities and analysis work?
- Does the series of tests (benchmarks) provide information to answer the above questions?

If the user can answer yes to these questions, then GIS technology can be successfully implemented in the organization.

SUMMARY

The traditional methods of acquiring, storing, and analyzing spatially referenced data are proving to be too costly and inflexible in meeting information requirements. GISs are emerging as the spatial data handling tools of choice for solving complex geographical problems. However, potential users of a GIS should understand all aspects of adopting this technology. Is a GIS an appropriate tool for your application, and, if so, which GIS should be used? The procedure outlined here will assist managers and technical specialists in answering such questions by providing general guidance for understanding GIS technology in a realistic perspective, evaluating the requirements of possible GIS users and applications, identifying applicable standards for information systems technology, selecting desirable software and hardware characteristics, and conducting benchmark tests to identify optimal hardware and software systems.

ADDENDUM

This paper is based upon a report authored by the Technology Exchange Working Group (TEWG) of the Federal Interagency Coordinating Committee on Digital Cartography entitled *A Process for Evaluating Geographic Information Systems* (Guptill, 1988). The report provides advice and guidelines for those who are involved with the technical issues of evaluating, designing, implementing, or procuring GISs. Standards and guidelines, software functions, hardware components, and benchmarking are discussed. Appendices provide details on standards, definitions of terms used in the report, and working group membership. The TEWG actively promotes the exchange of information and ideas on technology and methods for collecting and using digital cartographic data. The TEWG also monitors developments in GIS technology and documents guidance for utilizing that technology. As editor of the report, I would like to recognize the efforts of the contributing authors Dan Cotter, Bob Gibson, Dick Liston, Dave Pendleton, Elizabeth Porter, Henry Tom, Tim Trainor, and Pete VanWyhe. Copies of the report, *A Process for Evaluating Geographic Information Systems*, are available at no cost from the Federal Interagency Coordinating Committee on Digital Cartography, c/o Debbie Campbell, U.S. Geological Survey, 516 National Center, Reston, VA 22092.

REFERENCES

Broome, F. R., 1986. Mapping from a topologically encoded data base: the U.S. Bureau of the Census example. *Proceedings of AutoCarto, London*, 1986, Vol. 1, pp. 402–411.

Burrough, P. A., 1986. *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford, Clarendon Press, pp. 13–38.

Chrisman, N. R., 1987. Efficient digitizing through the combination of appropriate hardware and software for error detection and editing. *International Journal of Geographical Information Systems*, Vol. 1, No. 3, p. 265–277.

Cowen, D. J., 1988. GIS vs. CAD vs. DBMS: what are the differences? *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 11, pp. 1551–1558.

Digital Cartographic Data Standards Task Force, 1988. *The Proposed Standard for Digital Cartographic Data*. (Available from U.S. Geological Survey, Office of Technical Management, 510 National Center, Reston, VA 22092.)

Federal Interagency Coordinating Committee on Digital Cartography, 1987. *Coordination of Digital Cartographic Activities in the Federal Government*. (Available from the U.S. Geological Survey, Office of Coordination and Requirements, 516 National Center, Reston, VA 22092), pp. 26–32.

Gohagan, J. K., 1980. *Quantitative Analysis for Public Policy*, McGraw-Hill, New York, pp. 181–226.

Goodchild, M. F., and B. R. Rizzo, 1987. Performance evaluation and workload estimation for geographic information systems. *International Journal of Geographical Information Systems*, Vol. 1, No. 1, pp. 67–76.

Guptill, S. C. (ed.), 1985. *Proceedings of the Workshop on Methods and Techniques for Digitizing Data*. U.S. Geological Survey Open-File Report 85-307, 85 p.

———, 1988. *A Process for Evaluating Geographic Information Systems*, Federal Interagency Coordinating Committee on Digital Cartography, Technology Exchange Working Group, Technical Report 1: U.S. Geological Survey Open-File Report 88-105, 144 p.

Parker, H.D., 1988. The unique qualities of a geographic information system: a commentary. *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 11, pp. 1547–1549.

Smith, Terence, Sydhakar Menon, Jeffrey Star, and John Estes, 1987. Requirements and principles for the implementation and construction of large-scale geographic information systems. *International Journal of Geographical Information Systems*, Vol. 1, No. 1, pp. 13–31.

Stefanovič, Pavao, and Jane Drummond, 1987. Selection and evaluation of computer-assisted mapping and geo information systems. *ITC Journal*, 1987-1, pp. 39–44.

Tomlinson, R. F., and A. R. Boyle, 1981. The state of development of systems for handling natural resources inventory data. *Cartographica*, Vol. 18, No. 4, pp. 65–95.

U.S. General Accounting Office, 1986. *Evaluating the Acquisition and Operation of Information Systems*. Technical Paper 2, pp. 11–15.

U.S. Geological Survey, 1987. *Digital Line Graphs from 1:24,000-Scale Maps*. Data Users Guide 1, 109 p.

U.S. National Bureau of Standards, 1977. *Guidelines for Benchmarking ADP Systems in the Competitive Procurement Environment*. Federal Information Processing Standard 42 (FIPS PUB 42-1), 28 p.

———, 1980. *Guidelines on Constructing Benchmarks for ADP System Acquisition*. Federal Information Processing Standards Publication 75 (FIPS-PUB-75), 43 p.

APPENDIX

SOFTWARE FUNCTIONAL COMPONENTS CHECKLIST

A detailed checklist of the kinds of processing functions that are often available in geographic information systems for database creation follows in this section.

Database Creation

Digitizing

Methods

__Manually digitized two-dimensional point and line data.
__Manually digitized two-dimensional full polygon data.
__Manually digitized two-dimensional arc/node polygon data.
__Photogrammetrically digitized three-dimensional point, line, and polygon data.
__Manually encoded cellular data.
__Scanned map data.
__Scanned photographic data.

Tagging

__Assign feature names or codes that may be pointers to feature attributes while digitizing__ or as a separate process__.

Keyboard entry__	Numeric__	Field length__
Menu pad entry__	Text__	String length__
Cursor pad entry__		

__Facility for setting initial default values and duplicating previous entries.

Assigning Topology

__Arc pointers to areas	Automatic__	Manual__.
__Arc pointers to nodes	Automatic__	Manual__.
__Node pointers to arcs	Automatic__	Manual__.
__Node pointers to areas	Automatic__	Manual__.
__Area pointers to arcs	Automatic__	Manual__.
__Area pointers to nodes	Automatic__	Manual__.

__Automatic__manual__polygon assembly from arcs.
__Automatic__manual__identification/linking of complex polygons (for example, polygons with one or more inner rings).
__Automatic snapping of line end points to nodes while digitizing__ or in batch__mode.
__Automatic polygon closure.
__Automatic polygon centroid calculation__ or manual digitizing of centroids __.

Attributes

__Allow for interactive__ or batch__ entry of multiple attributes
__Allow attributes to be associated with features by feature name__ or by digitized coordinate__ (for example, interior polygon coordinate).
__Allow for automatic__ manual__ insertion of calculated area perimeter__ length__ statistics as attributes.

Error Detection and Editing

Raster or Vector Data

__Automatic topologic error checking, graphic display of errors, and facility for interactive correction.
__Format checking__ range checking__ value checking__ on vector__ coordinate data or raster__ pixel data during digitizing__ or in batch__ mode.
__Interactive insertion__ deletion__ changing__ moving__ of vector__ features or raster__ pixels by feature__ or groups__ of features.
__Automatic checking for overshoots or undershoots at line intersections during digitizing__ or in batch mode__ and correction by redigitizing__ or automatic clipping/joining__.

Attributes and Feature Names/Codes

__Interactive insertion__ deletion__ changing__ moving__ of feature names or codes.
__Checking for feature names or codes that are missing.
__Checking for illegal names/codes while digitizing__ or in batch__ mode.
__Entry level__ or batch__ checking for illegal attribute values or combinations of attribute values.
__Query select function for updating groups of graphic__ feature name__ or attribute __records.

Import/Export

__Ability to import the following data-set formats:
MOSS__ MAPS__ AMS__ SAGIS__ GRASS__ ODYSSEY__
USGS DLG (Standard)__ USGS DLG (Optional)__
USGS DEM__ USGS DTM__ GIRAS__ SCS GEF__
USCB DIME__
USCB TIGER/LINE__ USCB TIGER/DATA BASE__ STDS__
FEMA/IEMIS DBMS__ DIGITAL IMAGERY__
OTHER__ OTHER__ OTHER__

__Ability to export the following data-set formats:
MOSS__ MAPS__ AMS__ SAGIS__ GRASS__ ODYSSEY__
USGS DLG (Standard)__ USGS DLG (Optional)__
USGS DEM__ USGS DTM__ GIRAS__ SCS GEF__
USCB DIME__
USCB TIGER/LINE__ USCB TIGER/DATA BASE__ STDS__
FEMA/IEMIS DBMS__ DIGITAL IMAGERY__
OTHER__ OTHER__ OTHER__