Title: **Moving GIS into the Ocean Realm: Meeting the Need for Intelligent Data**

Authors: Simon Evans, Jeanne Murday and Richard Lawrence.
Environmental Systems Research Institute Inc.
Redlands, California, USA

Abstract:

Recent advances in marine data sensing technology, such as multibeam echo sounders, raise a series of data management challenges for the mariner. The days of small, localized data collection are vanishing fast and it is now more appropriate to ask how many gigabytes of data will be generated daily during survey operations. With vast amounts of data now a reality, the questions of data storage, maintenance, and access must be addressed. Additionally, marine data collection has become extremely expensive. It would seem logical that the need for the owner to extract maximum added value is paramount; however, in the past, this has been a difficult proposition owing to the dynamic nature of ocean data.

This paper will promote the use of Geographic Information System (GIS) technology in order to provide effective and efficient data management for the mariner. The use of GIS in the marine industry has been restricted in the past due to hardware and software limitations. Environmental Systems Research Institute, Inc. (ESRI), has recently completed research on an implementation of an object-oriented GIS data model. The model lends itself well to database designs and applications associated with the collection, documentation, distribution, and analysis of large amounts of data (i.e., a data warehouse). It also provides the end user with the options of setting states and behaviors for each data object. Using this approach provides the means of capturing/modeling ocean data and enabling the existence of "intelligent data objects."

Introduction

The fact that Oceans cover approximately 71% of the Earth's surface is well documented, as is the continuing exploration and exploitation of its natural resources. The majority of populated land areas of the world are connected by rivers and lakes; indeed whole provinces are maritime in character.

Several distinct working disciplines are presently active in the Ocean arena (e.g., seismic surveying, physical and biological oceanography, hydrocarbon exploration and exploitation). A common element of all of these activities is data, which is captured for a variety of reasons, whether to increase understanding of Ocean processes or provide decision support.

Significant technological advances have been made in providing accurate spatial positioning on a worldwide scale (e.g., Differential GPS). At the same time rapid advances have been made to offer a plethora of specific marine data sensors (e.g., Swathe Echo Sounders, Digital Side Scan Sonar, and Continuous Ocean profilers) that provide the user with a sophisticated level of interoperability.

Data Issues

In terms of data generation the technological advances already stated result in the capture of huge volumes that where incomprehensible just 15 years ago. It is now standard operational procedure to collect data in a digital format and realistically talk in megabytes at the very least, if not gigabytes, depending on duration of the task and sensors used.

It is also standard practice to provide a level of processing to the raw data collected, therefore ensuring that the quality of data captured is acceptable and ultimately of value. Coinciding with the technological advances made in data capture a number of user-specific applications/data models have been developed, thus ensuring adequate use of the data is achieved.

There is one indisputable fact about gathering any data from the ocean; it is an extremely expensive operation. This is due to the vast array of equipment and personnel required to gather the data and the amount of time taken to perform the task in question. It is therefore logical that the maximum use and therefore benefit of any data collected from the ocean must be sought, in effect a sense of "value" added.

To enable any user to maximize the potential of any ocean-related spatial data acquired, it is necessary to have a management regime that supports capture, maintenance, and warehousing.

Traditional GIS

Over the last thirty years there has been a revolution in the processing of geographic information of many kinds. Beginning with researchers who adapted the then current computer technology to the creation of simple maps and driven by the need of governments to deal somehow with the enormous amounts of spatially referenced information they were collecting, a commercial GIS industry arose. That industry, now supported by a very large and diverse group of users (public, private, academic, and nonprofit), has provided the modern GIS technology that is now used by millions of people around the world.

GIS has been identified in the past as an essential technology for the dissemination of oceanographic-related data, cataloging, archiving, display, and mapping (Wright, 1996). However, that actual implementation of such systems has been relativity slow, as GIS technology has generally concentrated on data that is normally stable over time and those variables that relate to relatively solid kinds of objects such as roads, vegetation, topography, and structures. It has been suggested that an object-oriented approach is best suited for dealing with the dynamic and multidimensional nature that is inherent in ocean-related data (Li, 1995).

The Geodatabase Data Model

ArcInfo 8 introduces a new object-oriented data model called the *geodatabase.* The defining purpose of this new data model is to design smarter GISs by endowing them with natural behaviors and to inherently model relationships among features within the data itself. The geodatabase data model brings a physical data model closer to its logical data model. The data objects in a geodatabase are mostly the same objects you would define in a logical data model such as owners, territories, structures, shipping lanes, etc. Further, the geodatabase data model lets you implement a number of behaviors without writing code. Behaviors can be implemented through domains, validation rules, and other functions of the framework provided in ArcInfo 8 technology. Writing software code is only necessary for the more specialized behaviors of features connected to another feature only if certain constraints are met.

Benefits of the Geodatabase Data Model

The common thread throughout these scenarios is that it is very useful to apply object-oriented data modeling to features. Object-oriented data modeling can characterize features more naturally by letting you define your own types of objects; by defining topological, spatial, and general relationships; and by capturing how these objects interact with other objects in a more intuitive approach. Some of the benefits of the geodatabase are

- *A uniform repository of geographic data.* Geographic and oceanographic data can conceivably be stored and centrally managed in one database.

- *Data entry and editing is more accurate.* Fewer mistakes are made because most of them can be prevented by intelligent validation behavior. For users of data sets with few land based features to guide data capture, this alone is a compelling reason to adopt the geodatabase data model.

- *Users work with more intuitive data objects.* Properly designed, a geodatabase contains data objects that correspond to the user's model of data. Instead of generic points, lines, and areas users can work familiar objects such as buoys, ocean currents, water bodies, haloclines, etc.

- *Features have a richer context.* With topological associations, spatial representation, and general relationships, you not only define a feature's qualities, but also its context with other features. This lets you specify what happens to features when a related feature is moved, changed, or deleted. This context also lets you locate and inspect a feature that is related to another. For example, a navigation buoy will consist of several relationships: position, type, and light characteristics, and by deleting the actual buoy the geodatabase will automatically update all inherent relationships.

- *Better maps can be made.* You have more control over how features are drawn and you can add intelligent drawing behavior. Highly specialized drawing methods can be executed by writing software code. For example, this will provide the ability to represent nautical chart features as defined by the International Hydrographic Organization's (IHO) S-57 data standard.

- *Features on a map display are dynamic.* When you work with features in ArcInfo, they can respond to changes in neighboring features. You can also associate custom queries or analytic tools with features. An example of this is scale-dependant viewing, where symbols and features are represented automatically.

- *Shapes of features are better defined.* The geodatabase data model lets you define the shapes of features using straight lines, circular curves, elliptical curves, and Bézier splines (mathematically defined shapes). For example, a thermocline or other continuous feature can now be represented as a continuous surface rather than a discreet entity.

- *Sets of features are continuous.* By their design, geodatabases can accommodate very large sets of features without tiles or other spatial partitions. The potential benefits of this for ocean-related data capture is obvious.

- *Many users can edit geographic data simultaneously.* The geodatabase data model permits work flows where many people can edit features in a local area and then reconcile any conflicts that emerge.

To be sure, you can realize some of these benefits without an object-oriented data model, but you would be at a disadvantage—you would need to write external code loosely coupled to features and prone to complexity and error. A principal advantage of the geodatabase data model is that it includes a framework to make it as easy as possible to create intelligent features that mimic the interactions and behaviors of real-world objects.

ArcInfo 8 is distinguished from antecedent releases as it applies object-oriented methodology to geographic data modeling. A developer interacts with data objects through a framework of object-oriented software classes called the *geodatabase data access object*s. There are three key hallmarks of object orientation: polymorphism, encapsulation, and inheritance.

- *Polymorphism* means that the behaviors (or methods) of an object class can adapt to variations of objects. For example, the core behavior of features, such as draw, add, and delete operations, is the same whether the features reside in a geodatabase, or other GIS formats such as coverage or shapefile.

- *Encapsulation* means that an object is accessed only through a well-defined set of software methods, organized into software interfaces. The geodatabase data access objects mask the internal details of data objects and provide a standard programming interface.

- *Inheritance* means that an object class can be defined to include the behavior of another object class and have additional behaviors. You can create custom feature types in ArcInfo and inherit the behavior of standard features.

Guidelines for Geodatabase Design

The structure of a geodatabase—feature data sets, feature classes, topological groupings, relationships, and other elements—lets you design geographic databases that highly resemble their logical data models. For a data modeler, this is the essential reason for the introduction of geodatabases into ArcInfo 8. These are the basic steps in designing a geodatabase:

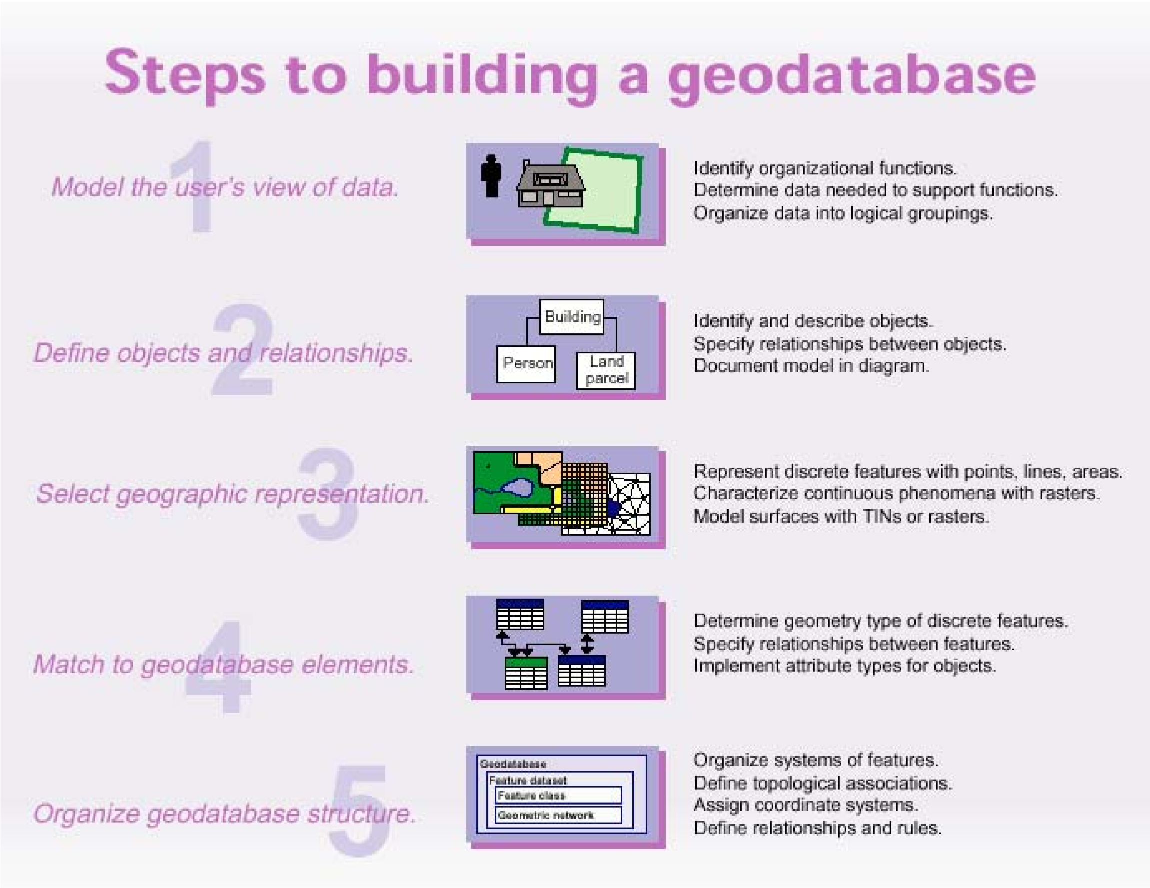
- *Model the user's view of data.* Perform interviews with users, understand an organization's structure, and analyze the business requirements.

- *Define objects and relationships.* Build the logical data model with the set of objects, knowing how they are related to one another.

- *Select geographic representation.* Determine whether vector, raster, surface, or location representation is best for the data of interest.

- *Match to geodatabase element*s. Fit the objects in the logical data model into the elements of a geodatabase.

- *Organize geodatabase structur*e. Build the structure of a geodatabase with consideration of thematic groupings, topological associations, and department responsibility of data.



Technology Trends

A GIS is at its core a database management system enhanced to store, index, and display geographic data. ArcInfo 8 is a significant release of new GIS technology that exploits several important technology trends just as they have become ready for commercial implementation. These trends collectively realize the vision of GIS as a geographically enabled database. The timing of ArcInfo 8 is fortuitous as it occurs during the convergence of several critical developments in software and database technology. The following are the principal trends that shape the technological framework of ArcInfo 8.

Spatial data and databases

When the coverage data model was first implemented, practical considerations led to the spatial component of geographic data being contained in binary files with unique identifiers to rows in relational database tables that stored feature attributes. With performance and functional advances in database technology, it is now possible and advantageous to store all spatial data directly within the same database tables as attribute data. The gain from storing spatial data directly within commercial databases is improved data administration, the utilization of data access and management services, and closer integration with the other databases that an organization manages. Moreover, ArcInfo users can select from any of the industry leading relational databases to host their geographic databases.

User interface

Applications developed for Microsoft Windows® have set a new standard for ease of use and consistency. Users have become accustomed to expected behaviors for mouse interaction, menus, dialog boxes, and the like. These user interface standards have made powerful applications accessible and usable by people who are not computer experts. ArcInfo 8 thoroughly implements the Windows standards for user interface and stands as a new milestone in making GIS software easier to use.

Software component architecture

ArcInfo 8 is built on the Microsoft COM architecture because it is the most robust and reliable component framework for desktop applications. Modern software is built on software component architectures, examples of which are Microsoft Component Object Model (COM), the Common Object Request Broker Architecture (CORBA), and Java Remote Method Invocation (RMI). The idea behind components is to divide software functionality into discrete, independent pieces that can be developed, tested, and combined into programs. By their design, components can be used to build any number of applications without modification. This provides a highly efficient level of software reuse. The benefit of software component architectures is better software quality, better performance, and the ability to update software versions without affecting other installed software.

Programming environment

Mature visual programming environments, such as Visual Basic and Visual C++, have become the norm for application development. The benefits of using these languages are the large pool of experienced programmers and the richness of these environments. It is no longer necessary or desirable to use proprietary macro languages. ArcInfo 8 uses Visual Basic for Applications (VBA) as its embedded macro language for customizing its applications, ArcMap and ArcCatalog. Other COM-compliant languages, such as Visual C++, can be used to extend the geodatabase data model.

Conclusion

The common themes of these technology trends are open standards and interoperability. The benefit of implementing these trends is to take advantage of technology from other industry segments, which lets ESRI concentrate its research and development on core GIS functionality. It is believed that the ArcInfo 8 technology will go some way toward meeting the challenges set by ocean-related data and provide a basis for a true GIS-oriented marine data management regime.

References

Li, R, Qian, L and Blais, (1995) "*A Hypergraph based conceptual model for bathymetric and related data management.*" Marine Geodesy, vol. 18, pp. 173–182

Wright, D. J. (1996) "*ArcView supports sea floor exploration.*" ESRI *ArcUser*, Spring Edition