SHARED CONTENT AND SERVICES FOR THE GEOWEB

Geographic Data Models

It is now technologically possible to build advanced internet geographic-information servers, introducing a new era of information sharing that some are calling the â€˜GeoWebâ€™. This will demand collaboration between producers and consumers of GI services to ensure their relevance and usefulness. Development of common data models is important for building GeoWeb content. The author envisages directions for the near future and invites the GIS community to start thinking about template design.

Over the past seven years ESRI has facilitated the development of more than thirty community-driven GIS data models. While these efforts have been successful to the extent that many GIS project teams use the models as a starting point for design, new technology and data needs are arriving with evolution of the GeoWeb. This is an interconnected, interdependent set of loosely coupled computer systems that will form the foundation of a new set of capabilities for business, science, government and citizen. Geographic Information Systems are currently providing benefits for many organisations. Desktop, mobile and Web-based applications have been built to help them make better decisions, to be more efficient and to provide better customer services. While these systems are important, recent technological advances could increase the reach and value of geographic information.

The GeoWeb
This â€˜system of systemsâ€™ will require new partnerships and co-ordination. To date organisations have primarily focused on their own mission and information needs but the GeoWeb will require better collaboration across traditional organisational and geographic borders. One key necessity will be consistent datasets, so that applications, Web services and maps can work in different geographical areas. Such datasets and Web services are currently being developed by the commercial sector; ArcWeb Services provide a good example of consistent datasets/products that can work at local, regional, national and global scale. These services are pushing the envelope for advanced capabilities and hosting solutions, but an opportunity exists to significantly expand available datasets and services over coming years. Significant investments by large, mainstream companies such as Google and Microsoft are also being made. While much of this focus will be on simple mapping and location-based services, additional content will drive more applications and uses. Today large volumes of imagery and commercial street datasets exist and consumer applications are limited to the information available in these. Obviously many more applications could be built, but the richness of the data is currently a limiting factor in broadening the suite of web services.

Co-ordination
What really makes up â€˜contentâ€™ for the GeoWeb? It is an interesting question because it stretches our current thinking about data models and the need for consist-ent content. In practice, GeoWeb content is typically pre-cached for multiple display scales. We have all seen the â€˜+â€™ and â€˜- â€™ buttons on a variety of commercial mapping tools and these pre-set scales have a strategy for caching tiles of information on servers to improve response time, reduce network bandwidth and improve scalability for mapping. They also have an approach for Web-service contracts for geo-coding and other tools with well-defined contracts for programmers. Behind the scenes there are also datasets/databases, what most industry practitioners would call the â€˜data modelâ€™, but these new types of service are leading us to consider all aspects of GeoWeb information models as a co-ordinated â€˜data modelâ€™ for the GeoWeb.

Industry-oriented
Our collective challenge is now to expand thinking and design methods to better suit this new style of data model. There are also significant discussion points relating to geo-standards and Web standards, but here the dialogue has shifted with new Web-based standards such as Keyhole Markup Language (KML) from Google. While there are many differences between organisations and the way they do business, the data they tend to manage is often quite similar. In some ways this is not too surprising; there are thousands of payroll systems deployed in many different organisations, most containing the same kind of information as required by local laws and accounting practices. In the same way, many geo-databases are also similar in different organisations. At ESRI we took the approach of building many industry-oriented data models. When new data types or new functionality are required to meet the needs of a particular community the strategy has been to add this to our core platform rather than build separate applications for each community.

Experience Gained
When we first began this approach we had an idea that we could standardise the entire â€˜objectâ€™ model (data/properties and behaviour/methods) and build an object-oriented template. What we discovered is that while there are many similarities in data/properties for objects, there is wide variation in approaches to business logic and object behaviour/methods. This is particularly true for data management systems - the majority of users involved in data-model projects. In order to build consensus we had to retreat to a simpler data-model approach, but there have also been many tools and data types that have been gradually added to support community needs. For example, DHI, the major water-resource institute in Denmark, had built extensive time-series tools for graphing, map display and charting, but these tools were not tightly integrated into ArcMap. A number of other organisations in the water-resource community expressed a desire for these tools and they are now integrated into the platform. Related atmospheric and marine-community research led

our software-development teams to realise that support for scientific-community multidimensional datasets such as NetCDF would be valuable in terms of integrating geographic and scientific information. Support for real-time visualisation is also important for many applications in areas such as emergency management and law enforcement. Finally, many organisations expressed a need for history and auditing capabilities in a Geodatabase; the land-records community is primarily interested in such archiving capabilities. It is through collaboration within multiple user communities that a more complete picture of GIS support for managing time has evolved.

Lessons Learnt

These efforts have been successful to the extent that many GIS project teams use the data models as a starting point for design, and ESRI has been able to target the development of software capabilities to the needs of different communities. The GeoWeb is offering GIS professionals a new set of challenges and opportunities. The context of their data models and best practices is reasonably well understood within each community but the method for arriving at best practices has not been straightforward. Without fuller context for purpose and application the data models are still relatively hard to understand for the uninitiated. Unfortunately, this knowledge is relatively difficult to acquire and communicate without many yearsâ€™ experience in a particular industry. In addition, we have collectively struggled to develop design processes that blend sound IT and GIS design methods in building data models. In many respects developing a GIS data model is just like designing any other IT system.

Tailoring Troubles

When it comes to GIS design, however, most project teams have concentrated on the information needs and corresponding database schema. We have not focused on output from the system, and this has been problematic. Firstly, many output (mapping and other) requirements are complex and cause rework later in the project lifecycle; annotation, multiple display scales, feature representation choices, symbols, cartography and other factors are important aspects of data-model design. Secondly, it is difficult to validate the suitability of the design for a specific purpose without an understanding of the output requirements within the context of a â€˜use caseâ€™ or â€˜business processâ€™. As a result, it has been difficult for individual project teams to tailor template data models to their own specific situation. This is in many respects an unfortunate side effect of concentrating on consensus in database schema. While this is true of our data-model projects, it is also true of many of the geo-standards of the last several years; consensus required simplification of goals relating to standards processes, resulting in less useful implementation guidance and less standardisation.

Invitation

While a lack of end-user and business context has been manageable in many implementations, we are now moving towards new possibilities for publishing maps and Web services on the internet and our design methods and definition for a â€˜Data Modelâ€™ needs to be revisited. We are broadening our thinking about data models to include:

- information products (maps, globes, reports, queries/results)
- layers, views, tools, Web services, caching models, geoprocessing models
- datasets and data models
- metadata.

The GIS community is invited now to start thinking about how a template design might look for own organisation and industry, how we would document the various parts of the data model and how we could share tools and methods across user communities. It will be especially important to understand what end users of these systems will consider â€˜usefulâ€™ and find ways of providing practical, implementation-oriented templates and design methods. If we are going to realise the potential of internet geographic information services we need to move quickly to support content developers with these template data models and design methods.