

OGC's column for the September, 2001 issue of GeoWorld

Spatial Catalogs & Registries Deliver Framework

Louis Hecht

With respect to the National Spatial Data Infrastructure (NSDI), "Framework" refers to basic data layers -- transportation, hydrography, cadastre, boundaries of governmental units, elevation, digital imagery, and geodetic control -- which provide a base on which to collect, register, integrate and analyze other information.

In the new world of the Spatial Web, how will a user (or the software components serving the user) find on the Web the most recent, or the "best," or the most accurate, or a particular small subset of transportation framework data for a particular region? The answer evolves:

1. In the beginning, (before the Web became widely used) there was "The Clearinghouse," FGDC's online card catalog to the US library of spatial data sets in which data holders would register standard metadata describing their holdings. Data seekers could search the Clearinghouse, find a data set, and download the data using ftp, or order the data on tape or CD.
2. Then the Open GIS Consortium (OGC) delivered the OpenGIS Catalog Services Specification for standard interfaces enabling open access to catalogs of spatial data (such as the Clearinghouse). It became clear that there would be not one Clearinghouse, but many, and they could be recursive, and they would be online and could interoperate like one big virtual Clearinghouse. The specification provides for discovering data sets, but it also provides for discovering specific data about specific geospatial features.
3. Now Web-based distributed computing is maturing. More and more of people's time at a computer will be spent on the Web, and remote Web sites will provide not only the data but also the software functionality (services) that users require. Often the services invoked in a work session will be hosted on multiple sites by various companies or organizations. Such services will interoperate by means of standard interfaces. In most cases the user won't know or care where the services are coming from.

But you will need to know and care if you are responsible for the geoprocessing needs of an organization. You will need to know the difference between *catalogs* and *registries*, and you will need to know those catalogs that contain your data and those registries that define your metadata.

A registry records types and a catalog records instances. For example, somewhere on the Web there will be a registry defining the standard types of geospatial features contained in a transportation layer. ISO TC/211's metadata standard (ISO/CD 19115 Geographic information - Metadata, currently in a draft version) will be in a registry, and it will

contain the authoritative schema for transportation data and metadata. But perhaps you work for a county highway department, and your schema, though based on the ISO standard, is particularized to accommodate 1) the schema used by the state highway department and 2) the schema used by the transportation department of the major city in your county. Both of those metadata schemas will be in online registries. Your county's transportation data will be a collection of instances of such data, and the data will be described in all its uniqueness in at least one catalog.

How will a query work? Imagine that an engineering firm is preparing a bid to repair a bridge in your county. The firm needs to know the details about the road supported by the bridge. Your county's data is online in a catalog, and the metadata describing your county's data is structured according to a schema in the aforementioned registries. The query will go to the catalog first, discover the schemas that must be accessed to make sense of the data, go to the registries to get those schemas, and then complete the query, returning to the engineer information about the road that crosses the bridge.

The transportation data returned to the engineer is Framework data. Its dictionary of feature types may be slightly non-standard by the US DOT standard, but the differences will be documented in machine-readable text in the registries. The Web's XML standard will play an important role here: XML is for the unambiguous, self-documenting transfer of information between systems. OGC's Geography Markup Language (GML) is XML for spatial information -- Spatial Web infrastructure layered on Web infrastructure.

Alas, there's no room left on this page to explain catalogs and registries for *services*, or registries (and in some cases catalogs) for Spatial/Temporal Reference Systems (coordinate systems, geoids, etc.), symbols, gazetteers, semantic translators, and Information Communities. Be we can explore this further in future columns.

-- end --