



Left: Pepco power plant in Benedict, Maryland [www.aceinc_umces.org].
Below: Pepco customers in the D.C. area [www.sprol.com].

*“Which technology breakthroughs in the last decade have had the greatest impact on the state of the geospatial industry in 2006?”
—An enterprise-wide GIS*

Avineon

A ROAD MAP TO IMPLEMENTING AN ENTERPRISE-WIDE GIS
Xiaodong Hong, GIS Manager



Increasingly, utility companies are choosing to implement GIS so as to improve their facility, outage, and other emergency response management. I was involved in designing and implementing an enterprise-wide ArcFM GIS for the Potomac Electric Power Company (PEPCO) in Washington, D.C., and in this article I describe how this impressive geodatabase was implemented.

The PEPCO GIS was built on ArcGIS 8. Unlike its predecessors, ArcGIS 6 runs completely on a relational database called “geodatabase,” which facilitates the storage of both spatial and non-spatial information and allows multiple users to access and edit the data. The geodatabase is thus the foundation of an enterprise-wide GIS.

PEPCO selected the Alexandria, Virginia, based Avineon to implement an enterprise-wide GIS in ArcGIS, on the ArcFM 8 platform. The Avineon team was tasked with designing an Oracle geodatabase and migrating PEPCO’s legacy CAD data to it.

Even though the focus of this paper is on a GIS information system designed for an electric company, the same approach can be applied to implementing enterprise-wide GIS at other utility companies. We learned some valuable lessons with regard to GIS implementation with currently available ArcGIS software from which other organizations may benefit.

GIS IMPLEMENTATION: SOME MOOT POINTS

A typical GIS implementation using the ESRI ArcInfo Coverage data format (supported by ArcInfo v. 7.2 and older) consists of three phases: data conversion specification, data migration, and application development.

When designing a data conversion specification, one needs to determine the functional requirements of the GIS system by collecting information from the organization’s different departments and then transform these requirements into a database structure.

For instance, if the users specify that a transformer needs to have nominal and operating voltage and rated KVA, then these requirements should be incorporated in the point attribute table (PAT) of the device coverage.

A data conversion specification typically contains table and item definitions, as well as business rules that need to be followed during data conversion. The specification should also clarify symbol requirements for each feature.

Due to limitations of the ESRI Coverage data format, relationships cannot be implemented at the data structure level. As a result, the relationship rules are typically enforced at the application level; in other words, the GIS will function properly only with a set of applications. Conversely, if the data structure is to support the desired GIS applications, the data conversion specifications must be developed simultaneously with the applications.

Many organizations would normally have converted their data to CAD formats (AutoCAD DWG

or Microstation DGN) before implementing the GIS. The target GIS data format would most likely be ArcInfo coverage of spatial and non-spatial data.

tenance tools present several challenges.

GIS maintains the geographical locations of each topology feature stored; but retaining their accuracy during data editing in GIS is often problematic. Hence, while maintaining topological integrity, GIS applications also need to be able to enforce certain business rules—such as the use of valid values, meaningful relationships to other features, and up-to-date links to data records in external databases.

In a traditional GIS, the coverage data format only allows single users to access and edit the data. Library and ArcStorm can overcome this limitation, but they still lack versioning and rollback, the two mainstay modules of multi-user information systems.

The relational geodatabase data structure introduced in ArcInfo 8 provides a solution that supports a multi-user, enterprise-wide GIS. Its basis is ArcSDE which facilitates data editing, versioning, and rollback by multiple users simultaneously.

The major plus of a relational data structure is that it enables users to define objects as well as behaviors (Zeiler 1999). Moreover, they can define true feature classes such as “Switch,” “Fuse,” and “Primary Electric Line;” not just “point,” “line,” and “polygon”.

These new feature classes introduce new elements to the GIS implementation, and a new approach is called

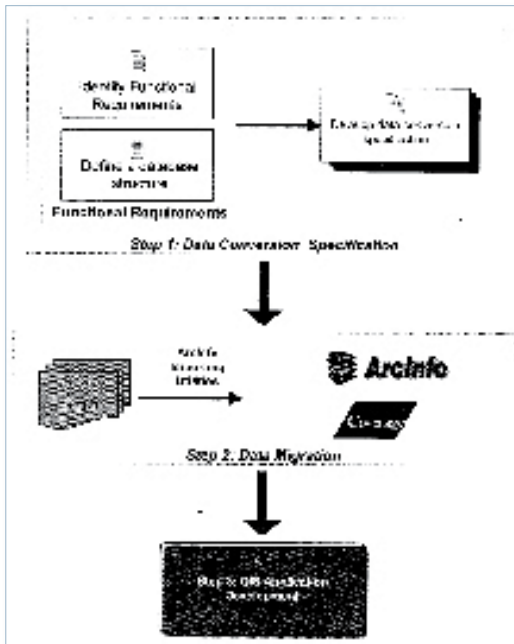
for. That approach is the “enterprise-wide” approach

ENTERPRISE-WIDE GIS

The GIS implementation plan for PEPCO assumed that sufficient funding had been allocated for the project; that ArcGIS and ArcFM 8 would be the GIS software used by the project; that CAD data were available for migration to the GIS; and, last but not the least, that the implementation will be carried out by a multi-professional team comprised of a GIS data modeler, GIS manager, and representatives from departments that will use and maintain the GIS.

DATA MODEL DESIGN REQUIREMENTS

The data model is the heart of the enterprise-wide GIS. It

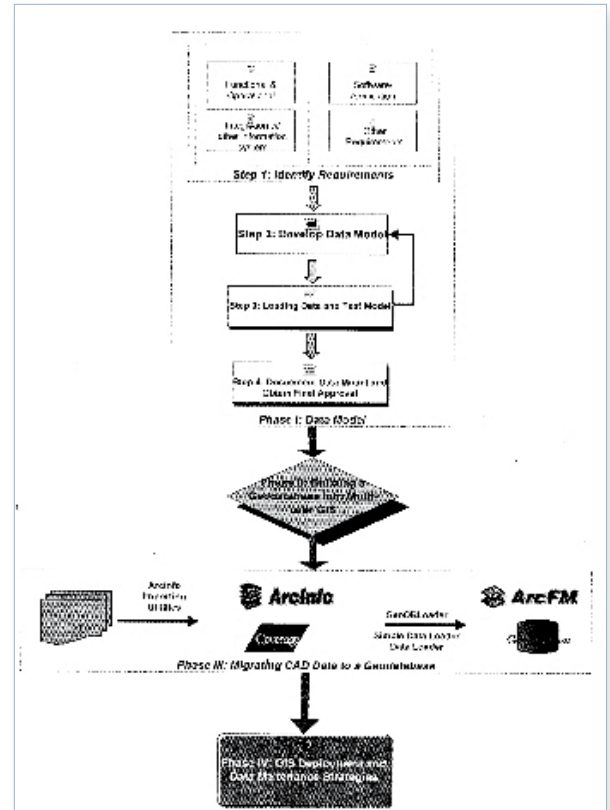


A traditional implementation plan

Additional attributes are usually stored in external databases. For example, a transformer in the device coverage may contain a link to an external Customer Information System (CIS) database. Attributes stored in external databases can be easily migrated to the GIS using ArcInfo’s import tools.

The development of GIS applications typically begins after the first GIS data delivery. During this process, GIS operations such as data query and data analysis, data maintenance, and enforcement of business rules are of primary concern.

ArcInfo and ArcView have many built-in functions to support data query and data analysis, and they are fairly easily tailored to the organization’s specific needs. The data main-



An enterprise-wide GIS implementation plan

must meet the organization's operational requirements and, it must also support the requirements of the software selected for the implementation. Above all, it must be capable of integrating with other information systems, such as the Outage Management System (OMS) in the case of PEPCO. Typically, the model must also support meaningful symbol display and data maintenance.

The first step in building the model is gathering information on daily operations and data management functions. Next, a list is compiled of all the information database systems with which the GIS will interact, and a decision is made as to how and when the linkages will be created. The third task would be to determine which GIS software functions and applications will be required. For instance, PEPCO has chosen ArcFM as the primary interface to their GIS database. For this interface to work, system-wide feeder management and tracing tools must be incorporated in the model.

To support the organization's data approval and data maintenance processes, the model also needs to include items that record these stages as well as the revisions of features that may be undertaken.

For example, PEPCO has a business process in place that records a feature from the "proposed" to the "approved" stage. To facilitate the same process in an enterprise-wide GIS, an item called "Construction Status" was added to PEPCO's data model.

Many legacy CAD systems contain feature attributes in various formats. To ensure that CAD attributes are migrated

correctly, the data model must include certain temporary items that hold CAD information together. The "temps" should be deleted after the data are fully migrated and validated in the Geodatabase.

Another important consideration at this stage is how well the model supports the displaying of symbol and coincident features.

In general, the more comprehensive the data model is, the better the GIS implementation will be. The PEPCO data model as built on the ArcFM 8 data model developed by Miner & Miner within the ArcGIS 8 framework.

SOFTWARE

ArcInfo 8 offers two options: to generate the data model in ArcCatalog or to develop an UML data model in Visio. The latter has the advantage that it can be used to generate complete geodatabases instead of one feature at a time, as is the case with ArcCatalog. Further, the UML model provides a graphical representation of the data and is thus more intuitive. On the minus side is that generating the database structure from an UML data model takes a long time.

However, the 8.1 version of ArcInfo incorporates tools which can detect schematic errors in repository, thereby shortening the generation process, and there are well designed GIS UML models on the market by ESRI and Miner & Miner which have been shown to shorten the learning curve and development time, while also ensuring the data model's compatibility with ArcGIS and ArcFM software, the software of choice for enterprise-wide GIS.

TESTING THE DATA MODEL

Once the geodatabase structure has been generated, several tests are performed to ensure that it meets the requirements. In the process, feature and object classes, as well as domain values and the relationship generated are confirmed and test data are loaded in to the geodatabase to inspect the behavior of each feature class. Other tests focus on the performance of the database, on software compatibility, on the suitability of symbology, and on integration with other systems, if this is a requirement.

During this testing phase, users will, for the first time, see and "feel" the data without having to actually understand the complex terms of the data model.

DOCUMENTATION

Because a GIS data model is developed interactively, all successive new configurations must be documented to keep track of the changes.

BUILDING A MULTI-USER GIS

We distinguish between a personal geodatabase running in MS Access and an enterprise geodatabase running under ArcSDE on Oracle, MS SQL, or DB2. The enterprise geodatabase, which is a relational database, supports multiple-user editing through versioning.

CAD DATA MIGRATION

Data migration is a labor-intensive process and often it is the most expensive step in the entire GIS implementation. The key to data migration is accuracy. For data accuracy to be maintained, data conversion specification should be generated for the data model well before full production starts.

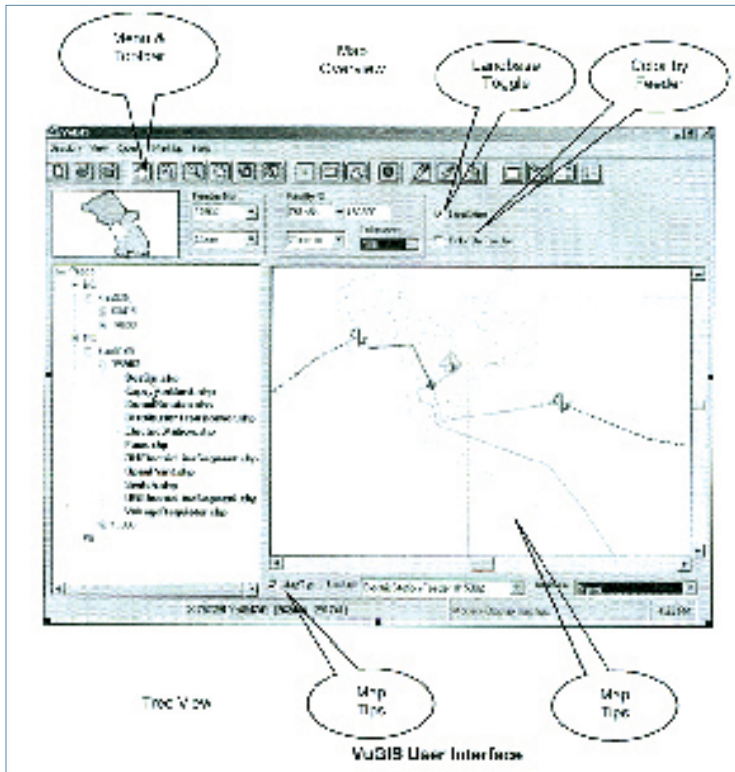
In addition, data validation in both CAD and GIS are critical to the success of data migration. Data validation routines are typically developed in CAD using AutoLisp, MDL, or VBA, and in GIS using AML.

There are many ways available to migrate CAD data to a geodatabase. In the PEPCO project, the ArcInfo Coverage

on the level of data accuracy stipulated.

GIS DEPLOYMENT AND DATA MAINTENANCE STRATEGIES

GIS deployment has to do with data access and data sharing throughout an organization, both of which can be achieved via a stable LAN connection or in the wireless mode.



format was used as a bridge, i.e., the CAD data were converted into an ArcInfo coverage, then coverage was "processed" within the geodatabase structure, and, finally, it was loaded into a personal geodatabase.

Although the target of the data migration is often the more portable and more manageable personal database, the deliverables can be received in either the personal or the enterprise geodatabase, depending

Access and sharing via the LAN is facilitated by ArcSDE 8 and such Internet technology as ArcIMS. The wireless connection may be sufficient to retrieve data from a central GIS database, and it usually requires a large and dedicated bandwidth. However, the unpredictable performance of current wireless technology is still of concern. The other option is to use lightweight data-viewing tools on mobile computers. The

downside of this approach is the need to daily download the GIS data from the server to the mobile computer.

PEPCO chose a viewing product called VuGIS developed by Avineon. VuGIS can run on a 266 MHz and 64 MB RAM computer and provides query and redline functions, in addition to viewing.

Finally, once access and deployment questions have been resolved, one has to ask: who, when, and how will the GIS data be maintained. A good starting point are the existing data maintenance procedures many organizations have for CAD data. However, for a multi-user, enterprise-wide GIS, the data maintenance responsibility will have to be decentralized, while maintaining a centralized data approval process.

SOME ROADBLOCKS AND SOLUTIONS

- Frequent model changes can lead to project delays and even unsuccessful implementation—freeze the model.
- In ArcInfo 8, symbols can be assigned to a feature based on a combination of up to three attributes. This effectively eliminates the need for a symbol item in the database and the application to reassign symbol values. But, one still has to validate the data model whether it has suitable attributes from which to generate distinct symbols for each feature class.
- Incremental data loading. To facilitate data migration Avineon developed a data-loading tool, the GeoDBLoader, which transfers data from one geodatabase to another in a batch process.

Zeiler, Michael. 1999. Modeling our world. The ESRI guide to geodatabase design. Redland, California: ESRI Press.