Data Refresh: Breathing New Life into a GIS Database

By Ruth Craven and Dr. Will Shepard

Data Refresh: Breathing New Life into a GIS Database

Geographic Information Systems (GIS), a component of what is known as "spatial technology", has origins in simpler systems such as Computer Aided Design (CAD) and Automated Mapping/Facility Management (AM/FM) systems. However, in order to fully reap the benefits of spatial technology, many utilities and other organizations find it necessary to augment and increase the quality of their existing data; in other words, to refresh their GIS data.

A History of Spatial Data Acquisition

In its infancy, from 1950 to the early 1970's, spatial technology addressed cartographic applications such as mapping and charting of land, water and demographic data. By the late 1970's and early 1980's, spatial technology had begun to mature, and utility companies were amongst the first organizations to adopt CAD or AM/FM. At the time, these systems were at the leading, and often bleeding, edge in the utility industry. Primary applications included mapping, facility inventory, map-book production and sometimes construction print production. In these early systems, spatial technology was usually a departmental solution, addressing only a limited set of needs. Spatial technology matured in the late 1980's and early 1990's, fueled by the proliferation of personal computers, client/server technology and scalable architecture. The late 1990's saw the growth of more sophisticated spatial applications, and the technology became known as "GIS" to emphasize the geographic focus, and saw the promotion of increased enterprise use.

Utilities that adopted spatial technology in the early years were challenged by the data resources they had – or didn't have – to populate the system:

- **Existing maps**: How accurate were the maps? Were the maps maintained? What was the map coverage?
- **Facilities data sources**: What data was available? Was it available digitally? Was the data consistent across the enterprise? How complete was the available data?
- **Land data**: Was commercial land data available?

AM/FM systems were usually populated by digitizing source paper maps, performing field inventories and incorporating other digital data sources, such as transformer databases. Not unlike today, the quality of the resulting AM/FM data was directly proportional to the quality and completeness of the source data or field inventory, the ability to associate various facilities data and the scope of the data conversion. For example, attribution from source maps often could not be related to nearby facilities and hence was converted as annotation, instead of attribution. In addition, connectivity was not a priority, since it was not necessary for map production.

Given that the value of spatial data for applications was often not understood during initial data conversions, current GIS data holdings often do not support advanced use of the data. If the scope of the facilities attribute population encompassed only what was needed to support mapping and facilities management, then the data may not have included all of the facilities, attribution and connectivity required for advanced applications such as outage management, asset investment, gas distribution integrity, advanced metering, dashboard applications or integration with an enterprise asset repository.

Refreshing GIS Data

Today, GIS is a mainstream enterprise system that is recognized as an enabling technology for other operational and energy delivery systems. Utilities recognize the enterprise value of spatial data and are eager for advanced applications and integrations. However, lacking data or poor quality from initial conversion efforts will hamper moving forward with these initiatives. Successful implementation of these initiatives will require refreshing GIS data through additional collection and reconciliation of existing data problems. Advanced applications and integrations, with their supplemental data requirements, include:

- GIS often owns the relationship between a customer's premise and the feeding transformer. Applications such as outage management systems (OMS), transformer load management (TLM), and network analysis require the premise-transformer relationship, circuit connectivity and integrated customer data.
- Applications receiving network/circuit connectivity from GIS require high data quality. For example, OMS requires that the nominal state of the network accurately reflects the field conditions. If it doesn't, outage projections and statistics will not be useful.

- Analysis tools available with graphic work design (GWD) products require location accuracy in order to provide useful results, such as accurate pole locations for guy placement analysis.
- Along with circuit connectivity, engineering analysis programs require valid stock code attribution, for example, device and conductor characteristics.
- Mobile applications such as service calls and trouble restoration utilizing spatial data are more successfully implemented when there are unique location identifiers, unique device identifiers, and accurate positional data.
- Transmission and distribution siting requires additional land data layers in order to do multiple realistic path sitings for the approval process.

Refresh activities to acquire this data include capturing additional facilities through field collection (for example, facilities that were not originally included in the inventory because they were not needed for maps, but that are now important for applications), updating of currently populated but dated attribute values, populating attribution from floating annotation and improving the positional and topological accuracy of facilities.

A data refresh initiative is not a simple effort. However, the results of this undertaking can enable successful implementation of advanced GIS applications, facilitate benefits realization of such applications and provide additional ROI from the GIS. Complexities and considerations of undertaking a data refresh initiative are not unlike initial data conversion or migration. **(See sidebar)**

The good news is these complexities can be resolved with a thoughtful, organized approach. A data refresh effort requires thorough planning and a project approach to allow the enterprise to reap the benefits of improved spatial data and to recover data refresh costs. **(See sidebar)**

Advanced Applications of Spatial Data

GIS implementation experts delineate different phases of implementation. These phases define increasingly sophisticated spatial data usage, and the expansion of advanced applications and integrations. Each phase allows the utility to realize increasing benefits and efficiencies in the energy delivery process (refer to Figure 1 and Table 1). A phase 1 GIS implementation is comprised of the standalone GIS along with other similarly standalone enterprise systems, such as work management systems (WMS), inspection and maintenance systems (I&M) and computerized maintenance management systems (CMMS).

Depending on the scope of a standalone GIS implementation as well the original data conversion scope and current data quality, GIS data may not support phase 2 and higher phase implementations. If current data cannot support the higher phase implementations, a data refresh initiative is required before the GIS can be leveraged to obtain additional ROI through advanced applications and integration with other systems.

Phase 2 implementations involve the integration of standalone systems. From the perspective of spatial data, GIS is integrated with other enterprise systems such as WMS and CMMS, as well as graphic work design (GWD), outage management (OMS), customer information (CIS) and distribution planning (DPS).

A phase 3 implementation brings together data marts owned by the various energy delivery systems for asset optimization. Distributed data is integrated in order to optimize capital expenditures and maintenance expenses. GIS continues to be an enabling technology in phase 3 implementations, such as OMS, interruption reliability reporting, enterprise asset management, and network planning. In the case of OMS, outage causes can be visualized geographically and analyzed using equipment maintenance and characteristics, and failure history. Interruption reliability reporting utilizes data from the GIS, OMS and CMMS. For enterprise asset management, CMMS obtains assets and spatial references (such as GPS coordinates, routing or survey grids) from the GIS. Finally, for network planning, data from GIS and CIS, together with historical performance of assets, can be used to optimize load and reliability.

DATA REFRESH COMPLEXITIES AND CONSIDERATIONS

- For what applications and integrations is the data refresh targeted?
- What positional accuracy is required for targeted applications, and beyond? How will existing coordinate data and new GPS position data be reconciled or coexist?
- How will the data refresh be funded?
- Should a widespread inventory be conducted? Should the data be collected along side normal field work? Can the data be collected in small areas as part of summer internships? Or all of the above?
- How will the data be maintained going forward?
- Can in-house labor perform the data refresh activities? Is contract labor needed? Or both?
- What data can be collected? For example, some conductor characteristics may not be discernible.
- Will existing GIS data be extracted or will the inventory return 'from scratch' data? If extracted, what is the mechanism to extract the GIS data?
- What is the mechanism for incorporating data returned back into the production GIS data mart?
- What data reconciliation efforts will need to be performed, i.e. meshing the new data with existing data? What software and effort will be needed?
- Will normal day-to-day posting of as-built data be frozen or will data coming back from a field have to be 'reconciled' with data posted since the original data was extracted from the GIS?

Phase 4 integration combines real-time field automation with systems and repositories that house and apply spatial data. Examples of real-time systems include substation automation (SA), distribution automation (DA), advanced metering infrastructure (AMI) and meter data management systems (MDMS). Spatial facilities data, integrated with customer data, can be used to justify and plan AMI and MDMS implementations. In addition, AMI/MDMS can update the GIS when a meter has been installed on the network and notify OMS of a service outage.

These phases of implementation are not successive; they do not require that a utility complete a lower phase of implementation before elements of higher phases. However, all phases of implementation benefit from (and to a certain degree require) a foundation of accurate and complete spatial data. If higher phase implementations are built on a foundation of poor data quality, then their success is at risk. All implementations, irrespective of phase, should evaluate current data quality as part of project planning. Implementation of advanced applications may require further data collection and improved data maintenance strategies.

Summary

For some utilities, the current GIS data holdings may not be good enough to support advanced GIS applications and integrations. This occurs because advanced functions were not foreseen during initial data collection, either because data was not available, or because the data have not been maintained in the GIS. The good news is that refresh initiatives can update this data. While this is a significant undertaking, the benefits gained from strategic, enterprise applications typically outweigh the costs involved. Data refresh can breathe new life into an old GIS database.

Sources

The following sources were used in the development of this article:

- Milestones of GIS, the Geospatial Resource Portal (http://www.gisdevelopment.net/history/1950-1960.htm)
- Integrated Delivery Framework: An Integration Tool for Utilities, Presented at ESRI Electric and Gas Users Group 2006 by Tom Helmer, Enspiria Solutions, Inc

Figure 1. Evolution of GIS

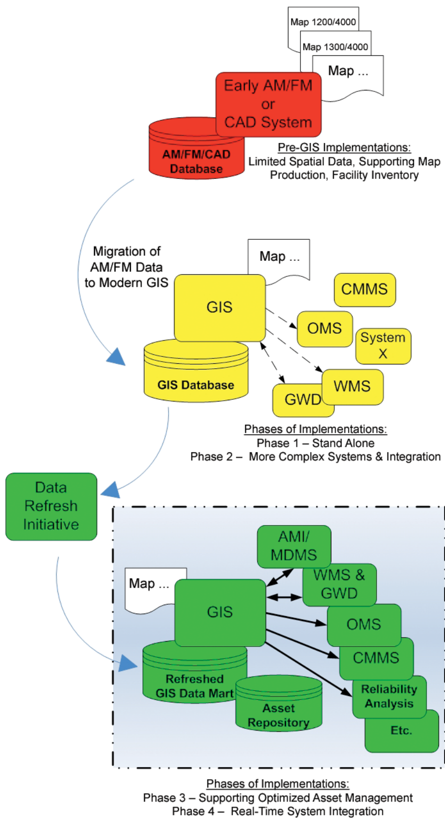

Table 1. Phases of Systems Implementation and Integration

Phase	Phase Activities	Key Features	Benefits
1	Stand-alone deployments of energy delivery systems: GIS, CIS, GWD, WMS, CMMS, MWM, SCADA and DPS.	Data redundancy from data maintained in multiple places results in data conflicts.	Work becomes automated.
2	Integration of systems for workforce optimization, e.g.: CIS→GIS GIS→OMS GIS↔GWD/WMS GIS↔CMMS	Integration is aided by modern enterprise integration infrastructures. GIS is used to correlate information from disparate systems.	Work processes and data maintenance costs are optimized.
3	Integration of individual system data marts for asset optimization.	GIS is a key enabler for the further integration of enterprise data, resulting in implementation of an enterprise data repository.	Asset information (e.g. characteristics, performance and location) is used to analyze and optimize capital expenditures and maintenance activities.
4	Integration of near real-time field automation.	Information from real-time systems such as AMI, MDMS, SA, & DA are leveraged by operational and engineering systems.	Field information (e.g. service outages, feeder outages, device outages, loads) is used in near real-time by the enterprise.

DATA REFRESH PLANNING AND MANAGEMENT ACTIVITIES

- Perform a strategic implementation plan and data audit to target applications and integrations that will benefit from new or improved data and to target data efforts that will offer the biggest-bang-for-the-buck.
- Explore cross-organizational funding and sponsorship during strategic planning of the data refresh initiative.
- Design and implement business processes to enable ongoing data maintenance; without this, new data will become stale. Ultimately, the organizational culture must be changed such that data quality is everyone's job.
- Pursue qualified vendors if contracted labor is to be used. While the techniques for data population have not changed drastically since early AM/FM implementations, data conversion and migration resources are now well established and experienced resources are available. Supplemental digital data sources are also more readily available.
- Keep in mind that the more 'natural' data exchange is between the GIS and the collection tools, the easier data import/export methodologies will be.
- Include a pilot phase in the data refresh initiative, including back-end processes necessary to bring refreshed data back into the production system and to reconcile it with existing data.

About the Authors

Ruth Craven *is a Senior Project Engineer with Enspiria Solutions. She has 28 years of utility software engineering and GIS experience. She holds a bachelor of science in computer science.*

Dr. Will Shepard *has extensive experience in utility GIS, graphic work design, and enterprise integration solutions. He received his Ph.D. in Geography, an M.S. in Geography with GIS Certificate, and a B.S. in Mathematics from The University of Georgia and is a Certified Project Management Professional (PMP).*