

Converting Water System Paper Maps to an ArcSDE Geodatabase

Bryn A. Enright, Tucson Water

Abstract

Tucson Water, a municipal water company in Tucson, Arizona, has recently completed conversion of more than 1600 paper water system maps into an ArcSDE Geodatabase. Prior to the conversion, the entire water system infrastructure was documented by hand on paper maps representing quarter sections of land. Converting from paper to digital geospatial data was a major component of and the foundation for an enterprise GIS. A GIS will allow Tucson Water to better manage, plan, and model its water system and resources. This paper will discuss the major steps involved with the conversion to an ArcSDE Geodatabase and some of the issues encountered during the conversion.

Background

Tucson Water is a department of the City of Tucson, Arizona and operates as a utility for serving both potable and non-potable water to a customer base of approximately 675,000 people located in a 375 square mile area within the Tucson, Arizona metropolitan area.

Over the past several years Tucson Water has been building an enterprise Geographic Information System (GIS). In June 1999, a GIS Strategic Plan was developed to address how an enterprise GIS could enhance and improve Tucson Water by making information more accessible. This plan identified the need to create a foundation for the enterprise GIS by converting the paper-based water system maps, commonly referred to as "valve maps", into digital geospatial data. Each valve map covers a quarter section of land and shows active and abandoned physical assets; pipes, hydrants, valves, fittings, services, etc. A complete set of valve maps is termed a "valve book" and contains over 1600 maps. Over 100 valve books were made available to Tucson Water staff in paper and microfiche format. The system of updating and distributing the valve maps was time consuming and inefficient. Updates to the valve maps were produced quarterly and it is estimated that over 3000 person hours were spent each year replacing outdated valve book pages.

The ultimate vision for the valve map conversion was a GIS that would be used to maintain the valve map information coupled with delivery to Tucson Water staff by either a Web based system to desktops or mobile field devices using wireless modem technology or CDs. The digital information should simplify work processes, allow for rapid distribution of updated information and promote the elimination of pockets of knowledge Department-wide in a timely manner.

Conversion Methodology

Tucson Water launched the Valve Map Conversion project in Spring 2001. Through a competitive process, Tucson Water selected EMA, Inc. and its sub-contractor ESRI to assist with preparation of the conversion methodology. The methodology team consisting of Tucson Water, EMA and ESRI evaluated various approaches and methodologies for valve map conversion. The team's evaluation resulted in the requirements for:

- Conversion team roles and responsibilities
- The software and hardware environment (ESRI's ArcGIS, ArcSDE, ArcIMS and the Geodatabase model)
- The database design
- The data source acquisition responsibilities and preparation of the valve maps for delivery to conversion vendor
- The data capture and conversion specifications
- The data problem identification resolution process
- The data delivery and append procedures
- The quality assurance and quality control procedures

In Spring 2002, the methodology team selected Avineon, Inc. to perform the actual conversion of the existing paper maps to geospatial data. Based on feature count estimates provided by Tucson Water, Avineon proposed a two-year conversion schedule. To test the conversion requirements and procedures established by the methodology team, a one-mile square pilot area was converted. Upon successful completion of the pilot, Avineon progressed into the full-scale conversion.

The water system was divided into seven separate areas (see Figure 1) for the full-scale conversion. This allowed for manageable blocks of data to be passed back and forth between Tucson Water and the conversion vendor. In addition, the smaller datasets would not overwhelm staff and allow for tighter controls on quality. The seven delivery areas varied in physical size and feature density. They were primarily defined by natural boundaries such as washes and mountains to help minimize the amount of edge matching between delivery areas. The first delivery area was sent to the conversion vendor in September 2002. The data conversion and quality checks for the final delivery area were completed in June 2004.

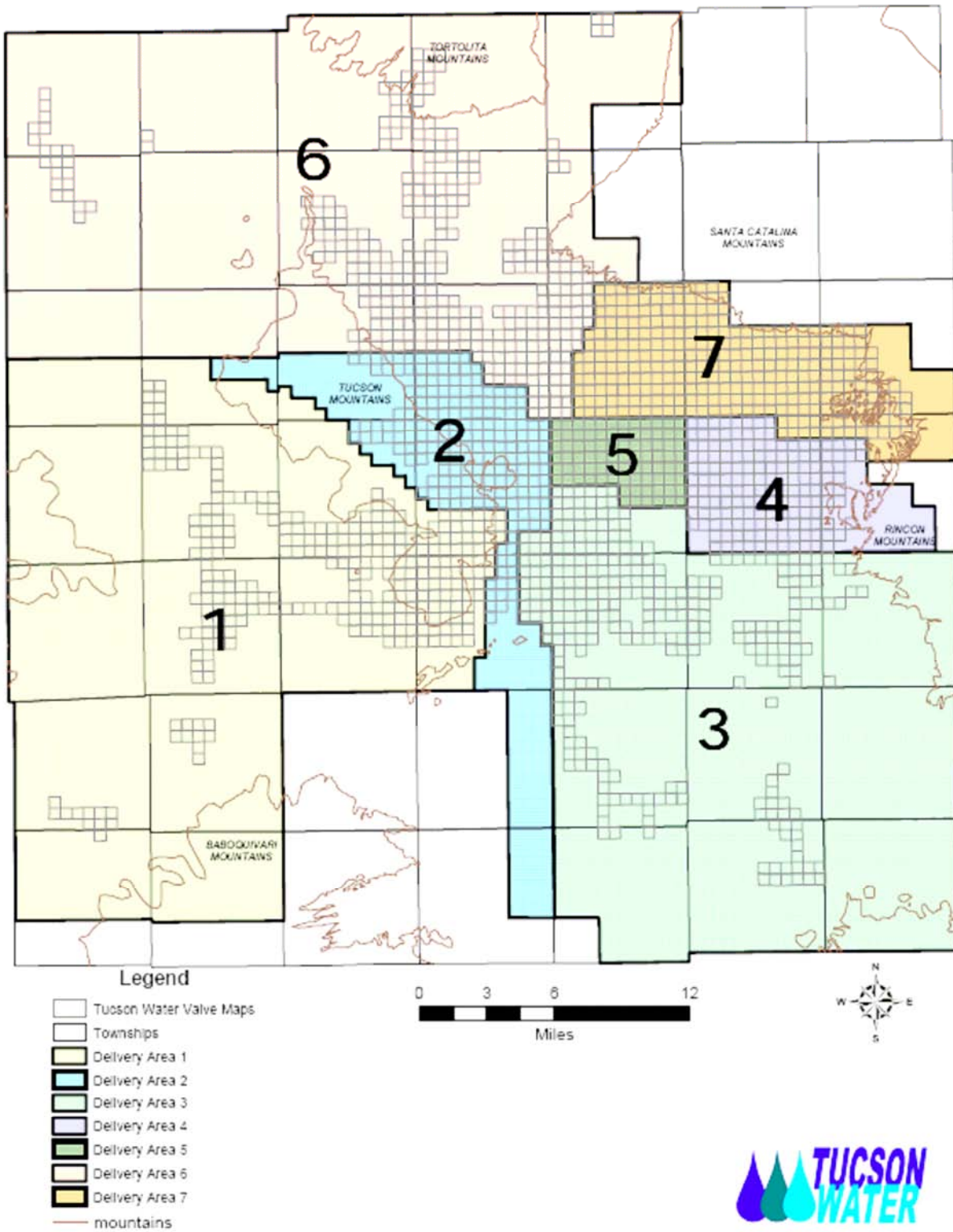


Figure 1 – Delivery Area Map

Data Acquisition and Scrubbing

One of the first and most critical steps in the valve map conversion project was preparing the maps for conversion. Tucson Water acquired, organized and reviewed all necessary source materials for delivery to Avineon. While Avineon performed the conversion, Tucson Water carried out the data acquisition and scrubbing for the next delivery area. The source materials comprised the basis of feature and attribute completeness and correctness. These materials included:

- A GIS parcel base provided by Arizona Pima County Department of Transportation
- A water service location layer geocoded from the addresses in the customer billing system
- “Scrubbed” versions of the paper valve maps

Avineon digitized the water system features using Pima County’s GIS parcel base as reference. The parcel base will also be used in Tucson Water’s new GIS. In the paper-based mapping system, Tucson Water maintained its own parcel base. Using the county maintained data will provide a more accurate base map as well as reduce mapping times by eliminating the need to digitize parcel information.

Avineon captured critical billing information from the water service location dataset to populate a unique key in the water service feature class in the Geodatabase. This unique key will be used to link the Geodatabase to the customer billing system and provide current customer information as well as consumption histories that are used by Tucson Water’s hydraulic modelers.

Each paper valve map was reviewed and scrubbed (see Figure 2). The main objective of the scrubbing was to provide a clean map for Avineon. Some of the steps in the scrubbing process are included below:

- Extents of engineering plans and work orders were highlighted for clarification because the installing document and install year are attributes carried in the new Geodatabase.
- Well, booster, reservoir, and pressure reducing valve names were scrubbed to match the updated naming conventions.
- Piping was added on well, booster and reservoir sites to connect them to the transmission mains outside the property. The paper-based valve maps did not show the piping inside the fence line of the production facility.
- Text and buildings not associated with Tucson Water were crossed out so they would not be converted.
- Changes in map standards such as different symbology were clarified.

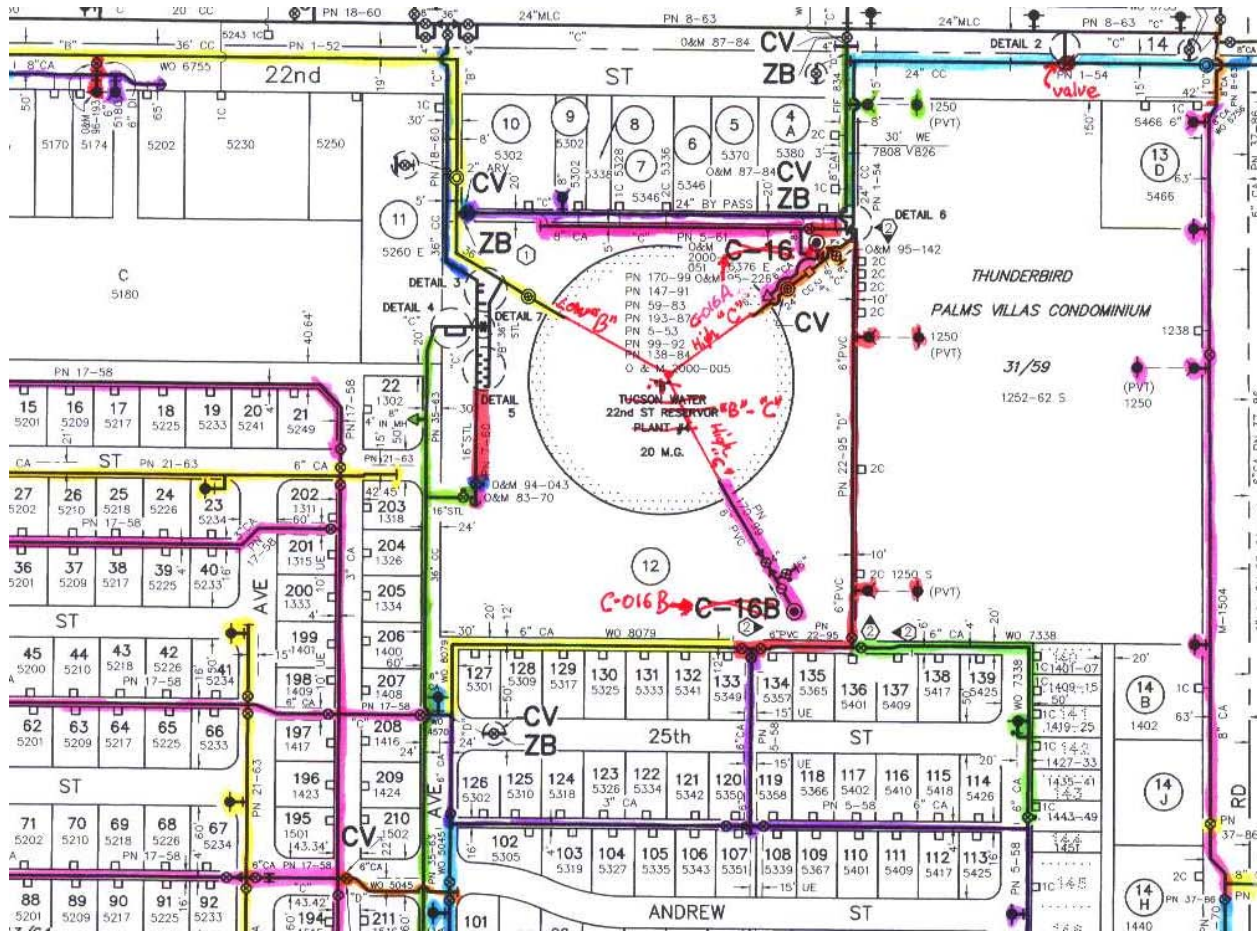


Figure 2 – Sample scrubbed map

When one of the seven delivery area reached the data scrubbing stage, the valve maps in that area were “frozen”. In other words, all new engineering plans, work orders and other edits to those maps were shelved until after the conversion. After the paper valve maps were scrubbed, they were scanned. Tucson Water mailed the original scrubbed version to Avineon and kept the scanned file as a record of the transmittal.

During the conversion process, it was necessary to continue using the paper-based valve maps. Tucson Water capitalized on the concept of the scanned valve map. Any Tucson Water employee could view a scan of a valve map at their desktop or in the field on a laptop. This new, cheaper form of distribution helped reduce reproduction costs. The number of staff requiring updated maps for their valve books or microfiche sets was cut in half. Until the conversion was completed, this became the preferred method for viewing the valve maps.

Quality Assurance and Quality Control

To ensure that all data automated into the Geodatabase was accurately converted, a series of quality assurance and quality control review processes were performed at three different stages of conversion: data capture, data review, and data append. A team composed mainly staff from the Tucson Water Mapping/GIS section who had previously maintained the paper Valve Maps performed the quality control testing. This team's core strength came from their knowledge and experience mapping the water system on the paper valve maps.

Data Capture

To maintain quality and consistency during the conversion stage, Avineon and Tucson Water collaborated to create a conversion specification document. This document contained all the rules for data capture and unusual or confusing scenarios that might have been encountered. Tucson Water and the vendor also shared a web-based project tracking application with a mechanism to post problems and resolutions regarding the map content. During the conversion project, over 750 problems or questions were posted and resolved in the web tracking application (see Figure 3).

PIRF Form

Last Question Number Used: 742 Find By Question Number 739
Find A Valve Map Number

Problem Identification And Resolution Form

Baymont Section:

Question Number:	738
Question Type:	Conversion
PIRF Type:	Rule
Feature Type:	Easements
Delivery Area:	6
Priority:	Medium
Valve Map #:	NW031314
Problem Location:	Entire map.
Question:	The entire valve map is showing R.E. along the roads and parcels, normally we do not capture easements that run along a
Suggested Resolution:	
Open Date:	15-Jan-2004

Tucson Section:

Status:	Resolved
Resolution:	No, that rule does not apply to Road Easements. Please capture.
Actual Close Date:	21-Jan-2004

This is the last Question Number Tucson has resolved: 739
Number of Questions Unresolved: 3
List of Unresolved Questions:

[Add Attachment](#)

Responsible Persons:

ABI Contact:	Mike Veigl
Tucson Contact:	Bryn Enright

Done? Check this box if question has been incorporated to map

[PRINTING OPTIONS](#) [EXIT FORM](#)

Figure 3 – Problem tracking form

Automated Checks

The quality control team used the trace tools available in ArcMap to find disconnected network features. The feature attribution checks were run with a series of out-of-box and configured batch validation tools from ESRI's Production Line Tool Set (PLTS) GIS Data ReViewer software. The first batch check examined fields tied to attribute domains and created an exceptions list for those values that did not adhere to the range or coded value domains. The second batch check compared attributes against values in the PLTS knowledge base. The knowledge base was a set of tables with additional rules to which the data must apply. Tucson Water configured the knowledge base with rules for valid attribute combinations and conditions. For example, a valid attribute combination for water main feature class required that all mains with a lifecycle attribute value of 'abandoned' must have a 'no' value for the Boolean attribute describing whether or not the pipe is pressurized. Condition rules check attributes by using SQL queries. For example, a condition rule on the water main feature class required all PVC mains to be less than or equal to 12" in diameter.

Visual checks

The Tucson Water quality control team did the visual checks manually. For each delivery area, the team had 25 working days to review the converted data. For every map in a delivery area, the team did a one-to-one comparison of the original valve map with the converted data to make sure that every feature and annotation on the map was captured. The team reviewed tens of thousands of features in over 1600 valve maps. The team also verified the accuracy of several attributes that were captured along with the map features. The types of feature and attribute errors flagged by the team included:

- Missing features
- Duplicate features
- Incorrect feature shape or placement
- Missing or incorrect feature attribute values
- Missing or incorrect annotation

Because of the 25 day timeframe, it was impossible to check the attributes for every feature. For the smaller groups of features that numbered less than 1,000 such as the wells and boosters, every attribute was reviewed. For the larger groups of features that numbered in the tens of thousands, approximately five to ten percent of the attributes were reviewed.

Acceptance Criteria and Error Correction

The errors found during the automated and visual quality control checks were logged and used to calculate an accuracy rate. The contract with the vendor defined a set of acceptance criteria for the converted data. If the accuracy rates met the acceptance criteria, the delivery was accepted. Overall, the quality of the converted data was very good and delivered on or ahead of schedule. Only the second deliverable was sent back to the vendor for minor corrections.

When a deliverable was accepted, the quality control team revisited the errors and made the corrections. After three delivery areas, the quality control team was satisfied that Avineon would continue to produce a product with minimal errors and questioned the logic of visiting the errors twice. In order to streamline the process, the decision was made to correct the error at the same time that it was discovered during the visual checks. The corrections were made in a version of the Geodatabase. Using the “Show differenced between two versions” tool from the ArcGIS Developer Samples, the original database delivery was compared to the versioned instance and a report was created documenting the differences between the versions (see Figure 4). This report was used to validate the assumption that the Avineon data deliveries would continue with their trend of minimal errors.

```
ADMIN.MAIN: OBJECTID 920
Row has been updated in ADMIN.QC and not changed in SDE.DEFAULT
* Field LASTEDITDATE changed:
  - value in SDE.DEFAULT: 8/1/2003
  - value in ADMIN.QC: 12/2/2003
* Field DOCUMENTNUMBER changed:
  - value in SDE.DEFAULT: FIF 847
  - value in ADMIN.QC: PN 078-1997
* Field INSTALLYEAR changed:
  - value in SDE.DEFAULT:
  - value in ADMIN.QC: 1997

ADMIN.MAIN: OBJECTID 924
Row has been updated in ADMIN.QC and not changed in SDE.DEFAULT
* Field LASTEDITDATE changed:
  - value in SDE.DEFAULT: 8/1/2003
  - value in ADMIN.QC: 12/2/2003
* Field DOCUMENTNUMBER changed:
  - value in SDE.DEFAULT: FIF 847
  - value in ADMIN.QC: PN 078-1997
* Field INSTALLYEAR changed:
  - value in SDE.DEFAULT:
  - value in ADMIN.QC: 1997
```

Figure 4 – Version difference report generated by Versioned Difference tool

Data Append

The quality control activities were performed in a SDE Geodatabase separate from the production database. Prior to appending, Tucson Water compared the delivery database with the production database using ESRI’s Schema Validator tool. The schema validator looked for differences between the two databases in coordinate systems, coded value and range domains, attributes, subtypes, feature classes, and feature datasets. Following the append process, the quality control team used the ArcMap trace tools again to ensure network connectivity between the delivery areas.

Post-processing

One of the last steps in the valve map conversion project was the data post-processing. This step gathered and populated information required in the GIS that was not available on the original valve maps. Some of this information was assigned programmatically using existing spatial datasets while some required manual updates after researching work orders, engineering plans and other documents. The post-processing efforts concentrated on attributes that would be used to integrate the GIS with other key systems at Tucson Water:

- Computerized work order and asset management system
- Customer billing and information system
- Water quality database
- Hydraulic modeling software
- Document management system

Through integration, the GIS will become more than just a digital mapping system. GIS technology is a fundamental way of tying together these disparate systems because virtually all data has a spatial component. By leveraging technology and automating processes, Tucson Water can enhance business practices to create efficiencies.

Lessons Learned

While having the water system data in a GIS creates efficiencies and opportunities in several of the business workflows, it also comes with its own trials and tribulations. In preparation for conversion, Tucson Water conducted a project risk analysis to assess what could go wrong and develop appropriate responses. Unfortunately, not all risks were identified.

Parcel Base Rectification

In an effort to eliminate redundant work and streamline the water system mapping process, Tucson Water chose to use Pima County's GIS parcel base instead of maintaining a separate parcel dataset. At the same time as the conversion project, Pima County was in the process of adjusting the GIS parcel base to fit a more accurate orthophoto base. As a result, there were many areas where the converted water system data did not correctly overlay the parcel base (see Figure 5). Tucson Water's strategy was to update these areas manually during the post processing stage. However, Tucson Water did not accurately assess the magnitude of the parcel shifts and the labor required for the manual updates. Furthermore, additional flights for updated orthophotos with higher resolutions are extending Pima County rectification process and Tucson Water's efforts to keep up with the shifts.

- Lesson learned: carefully consider risks of relying on outside data sources.

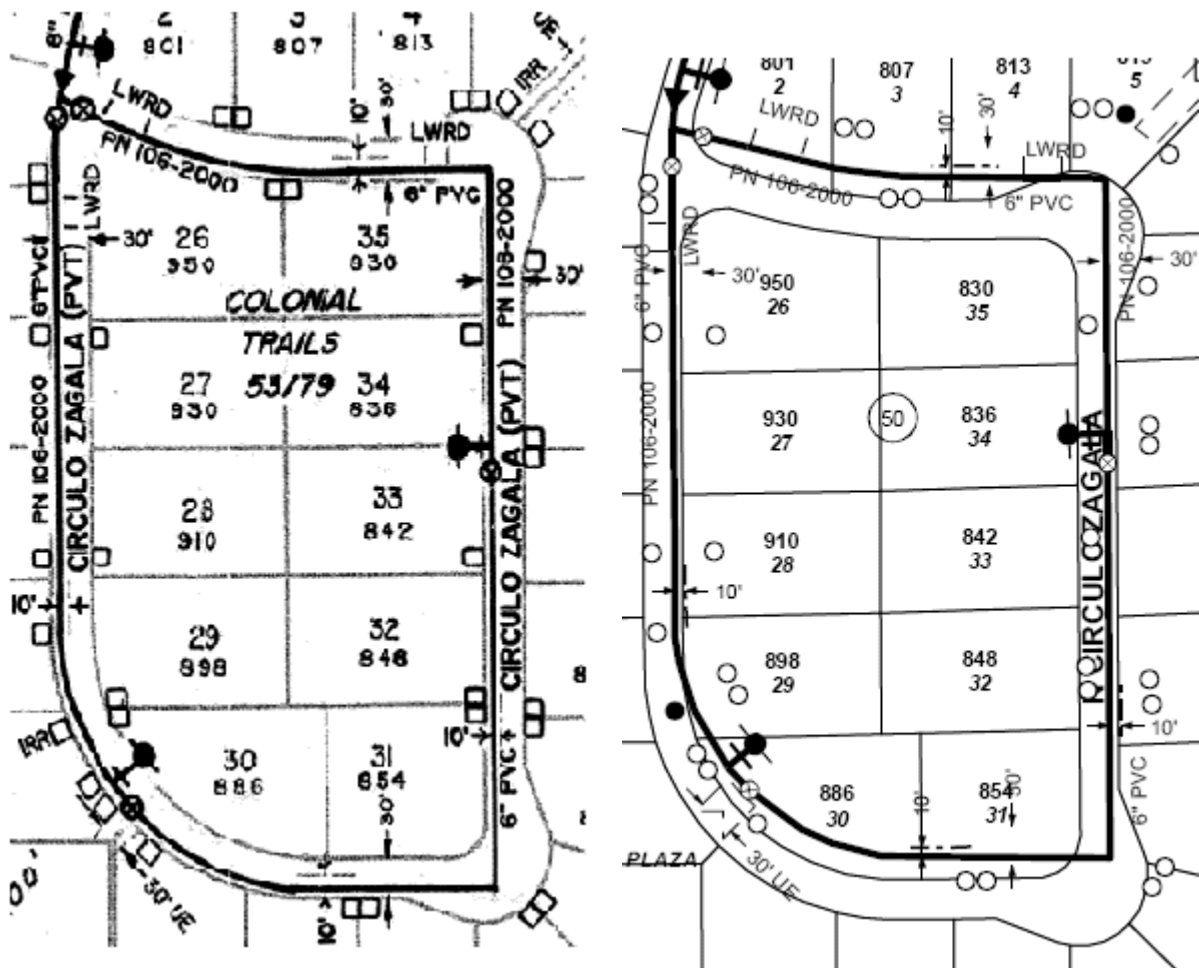


Figure 5 – Original valve map (left) and new valve map showing parcel shift (right)

Labels Versus Annotation

The original valve maps contained a considerable amount of annotation. In another effort to eliminate redundant work and streamline the water system mapping process, street names and parcel addresses were not converted into Geodatabase annotation. Instead the valve maps annotation would be reproduced by labeling the available GIS street and parcel base from Pima County. Unfortunately, after conversion, it was discovered that neither the ArcMap label engine nor the ESRI Maplex extension could provide acceptable label placement at the scale of the valve maps. One of the risks identified during the risk analysis was the adoption of the ArcGIS software product and Geodatabase architecture which was relatively new in the marketplace and new to Tucson Water, but there was no strategy for this particular labeling problem. Currently, Tucson Water is in the process of adding and placing street name and address annotation to the Geodatabase.

- Lesson learned: carefully weigh software options and test assumptions regarding functionality.

Checkplots

The original contract with Avineon required checkplots to be produced for each valve map at 1"= 200' scale. The checkplots and appropriate source map would be compared and the quality control team would flag any discrepancies. As mutually agreed upon by Avineon and Tucson Water, this requirement was abandoned in favor of performing the visual checks on-screen with the PLTS Data ReViewer software. In hindsight, a combination of the on-screen and checkplot review would have provided better quality control. The on-screen review gave the quality control team the ability to zoom in and view the data at a larger scale. The reviewer missed many overplots that were only visible at the smaller scale of the valve map.

- Lesson learned: implement a thorough change control process for modifications to the project scope.

Integration With Maintenance Management System (MMS)

The GIS and the MMS are integrated through a primary/foreign key relationship. The primary key is generated in the MMS and posted back to the GIS. In cooperation with the MMS vendor, scripts were written to perform this process on the business tables, which required the Geodatabase to be in an unversioned state. Following conversion, the Geodatabase was placed in production mode and the number of versions quickly grew to over 100. At this point, unversioning the database became a cumbersome procedure, not to mention the numerous quality control checks that had to be bypassed. Tucson Water is now in the process of re-writing the scripts to perform the data transfer on a versioned instance of the Geodatabase utilizing stored procedures, multiversioned views and SQL. Information on this approach can be found in the ESRI technical paper [Working with the Geodatabase Using SQL, February 2004](#).

- Lesson learned: prototype production environment business workflows.

Conclusions

To fully realize the benefits of an enterprise GIS, the data core must be as complete and accurate as possible. A well-thought out and planned conversion project is essential to ensuring data quality and completeness. The objective of this paper was to highlight some of the key areas that require focus and attention for a successful conversion. This paper also acknowledges that even with proper planning and management, the end result is not always perfect.

Acknowledgements

Special thanks to Robert Czaja, Karla Borbon-Dausinger and Sandy Elder for their valuable assistance with this paper.

Author Information

Bryn A. Enright
System Analyst
Tucson Water – Mapping/GIS
PO Box 27210
Tucson, AZ 85726
Bryn.Enright@tucsonaz.gov
Phone: 520-791-5080 x1240
Fax: 520-791-5288

Paper prepared for the 25th Annual ESRI User Conference, July 25th-29th 2005, San Diego, CA