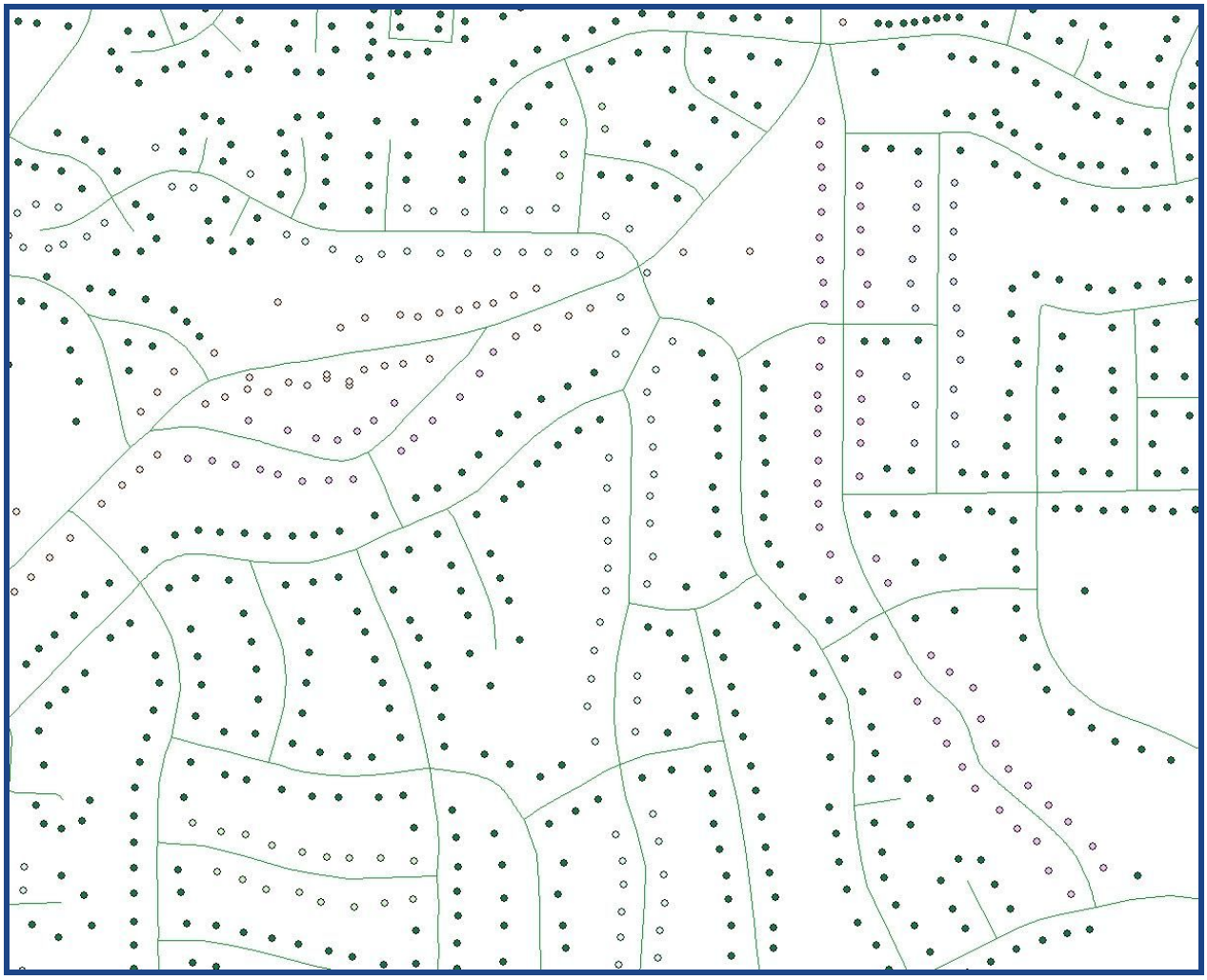
ADDRESSING PROJECT REPORT

Phase II: Data Clean-up and Master Address Repository Creation



MapForsyth
City-County Geographic Information Office

Prepared by:

Matthew Hamby - Addressing Coordinator
Stacy Tolbert - Addressing Analyst
Joseph Sloop, PhD, CGCIO, GISP - Geographic Information Officer
David Neininger - Project Manager

Prepared For:

Forsyth County Board of Commissioners
Winston-Salem City Council
Forsyth County Manager
Winston-Salem City Manager

PURPOSE OF REPORT 3

INTRODUCTION 3

PROJECT METHODOLOGY 3

DATA USED DURING CLEAN-UP 4

PROJECT AREAS 4

ADDRESS CLEANING AND VALIDATION PROCESS 5

rGA MAR TOOL 7

RESULTS OF PHASE II 9

NEXT STEPS 9

PURPOSE OF REPORT

The purposes of this report are to provide a general background and overview of the Addressing project and comply with Section III (a)(v) of the City/County Cooperative Services Agreement for MapForsyth.

INTRODUCTION

The Addressing project began in October 2013 with Phase I: Strategic and Implementation Plan. The strategic plan provided MapForsyth an understanding of the current state of county addressing and identified opportunities for workflow improvements. Phase I served as a foundation for Phase II of the Addressing Project, which creates a clean authoritative dataset and loads this data into a Master Address Repository (MAR). Phase II is nearing completion, and the MAR will contain all of the official situs addresses in Forsyth County, NC. Phase III of the project includes system integration with the MAR.

The purposes of the Master Address Repository are to:

- Provide an authoritative Addressing dataset;
- Avoid duplication of effort;
- Limit the number of individuals who can edit addressing data (Address Points and Centerlines); and
- Increased integration between enterprise applications and business systems.

Project cost for Phases I and II totaled $191,720 with Phase I being $25,000 and Phase II totaling $166,720. Phase III costs are to be determined per each system integrated. Some systems will have minimal MAR integration costs because there are no proprietary constraints. However, the cost to integrate proprietary systems will depend on the third party vendor requirements and the work involved with the MAR integration.

PROJECT METHODOLOGY

The project methodology for Phase II consisted of gathering data from GIS (Address Points and Street Centerlines) and from the four highest priority databases as identified by the GIS Steering Committee during Phase I of the project. All of this data was compared, system/dataset to system/dataset, to identify the authoritative and

correct address. In some cases, addresses were validated through historical documents such as recorded plats, City Council meeting minutes, and County Commissioner meeting minutes.

DATA USED DURING CLEAN-UP

1. Address Points and Street Centerlines (GIS Feature Class)
    a. These datasets are maintained by MapForsyth's Addressing Team using Esri Software (GIS).
2. Address Information from the following Systems
    a. Tax (County System)
    b. City Works (City Business System)
    c. CSR (City 311 System)
    d. Cayenta (Utilities' Billing System)

PROJECT AREAS

Due to the large number of addresses within the county (+180,000), the county was divided into six (6) areas (Figure 1). The areas were based on natural breaks such as highways, corporate municipal limits, and so forth.

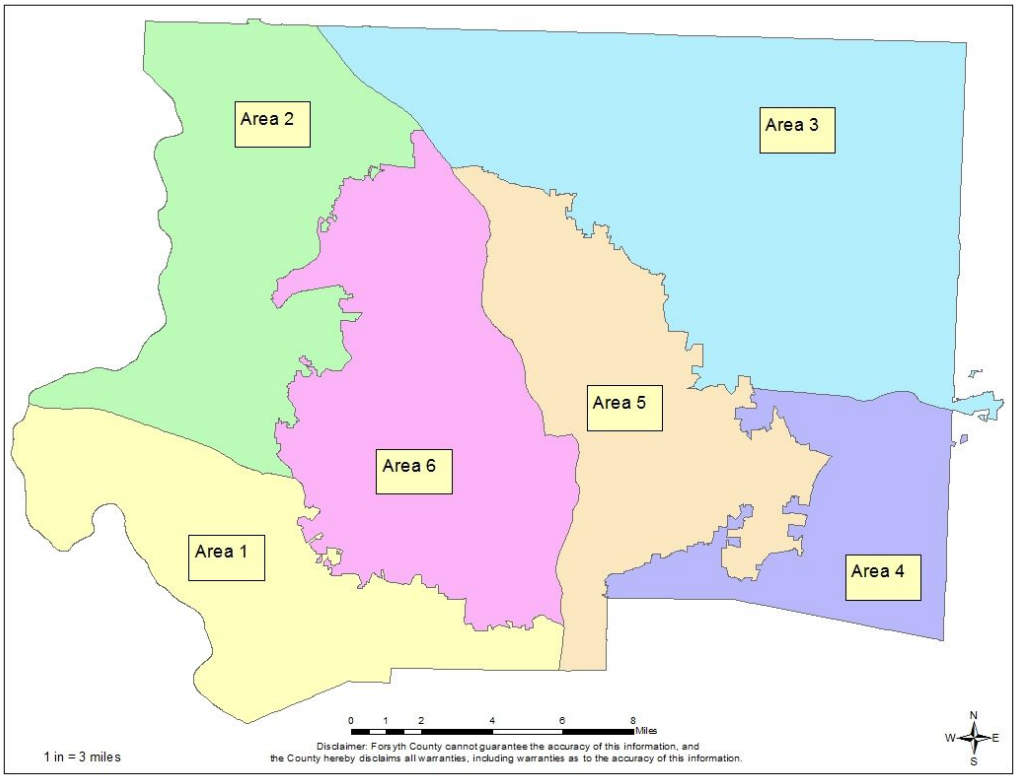


Figure 1: Addressing Project Areas

ADDRESS CLEANING AND VALIDATION PROCESS

The collected data was sent to Spatial Focus[1] for cleaning and processing. Spatial Focus ran a spatial analysis in which all data sources were compared to see how many addresses within each system matched. If an address did not match in all systems, it was flagged as an anomaly. Anomalies were sent back to MapForsyth's Addressing Team for review. The Addressing Team reviewed all anomalies and submitted to Spatial Focus a resolution for each. Spatial Focus used that information to correct the data used to build the MAR (Figure 2).

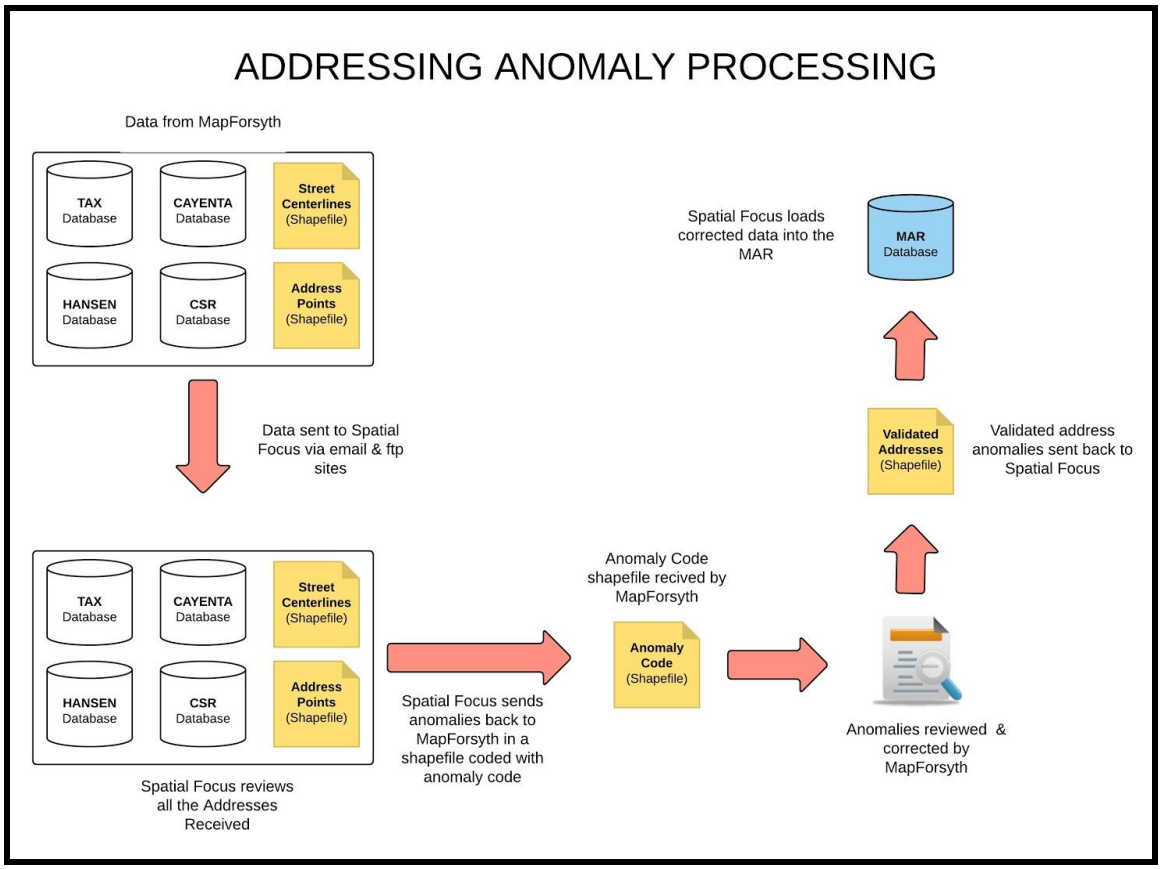


Figure 2: Addressing Anomaly Processing

Table 1 shows the different anomalies found during the cleanup process. The anomaly review process identified 22 different types of address anomalies. The anomaly codes

[1] Spatial Focus is the vendor selected during a bidding process for Phases I and II of the Addressing Project. They specialize in the area of Addressing and Master Address Repositories (www.spatialfocus.com).

varied in each of the six areas. Areas with more dense development (based on the number of residential addresses) contained more out of order addresses than less dense areas. Street name spelling codes occurred more in the rural areas than the urban area. For example, Area 2 has streets with hyphenated spellings such as Vienna-Dozier Road and Lewisville-Vienna Road. The anomaly process revealed that many of the data sources/databases did not use the hyphenated spelling. In Area 2, Vienna-Dozier Road alone was spelled wrong 152 times in one database and 137 times in another database.

Anomaly	Code	Pilot	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6
Conflicting Address Source Information	1					•	•	•
Address is geocoded	2	•	•					
Spelling Anomaly	4	•	•	•		•	•	•
Ambiguous or Inconsistent place name	6	•						
Street Type Anomaly	8	•	•	•		•	•	•
Address far from street	10	•	•		•			•
Address Number is Zero or Null	12	•	•	•	•	•		
Address # not in range for segment	15	•	•	•	•	•		
Address out of order	17						•	•
Ambiguous Address Location	18	•		•		•		
Streetname Predirectional Anomaly	21		•	•				
Address and Access do not match	22		•	•	•	•	•	•
No Strucutre with that address	23		•	•				•
Missing Centerline	25		•	•	•	•	•	•
Address uses alias	27		•	•				
Nearest named street lacks range	31		•		•			
Address number parity is incorrect for side	33		•	•		•		
Modifier Error	35			•				
Long Fishbone	46				•			
Outside Study Area	54				•			
Address street cannot be located	63							•

Table 1: Spatial Focus Anomaly Chart

Before the anomaly review process started, a review of a portion of Area 1 was used as a Pilot Area. This Pilot Area was used to train the Addressing Team on the review process and help to get a general idea of the time needed to review Anomalies. The Pilot Area data is included in the Area 1 data. Most of the addresses in this area were coded as “Address # not in range for segment” or anomaly code #15. This identified addresses in which the range for the centerline did not matchup with the address point location. This creates problems with the availability of address numbers along a segment of road. Staff was to adjust the centerline range to include the address number in some cases. Some of these anomalies were generated due to the process the tool used to find the address along the centerline.

When the County was completed, staff reviewed the codes for each area. The “Spelling Anomaly” code was the highest in Area 2. Of the 499 spelling anomaly codes, 46% (228) were hyphenated road names in which the business system data did not contain the

hyphenated spelling.

Kernersville is within Area 3 and Area 4, which generated a number of anomalies due mostly to the address points supplied by Kernersville and the address points from the official County point layer not sitting atop each other. The other large number of anomalies in Area 4 were anomaly code #33 “Address number parity is incorrect”. Most of these addresses were assigned using Kernersville’s Addressing Policy. The anomaly review was designed around Forsyth County’s Addressing Policy; therefore, it flagged these addresses as anomalies.

Areas 5 and 6 are almost entirely within the municipal limits of Winston-Salem, which has over 58,500 more addresses than the other four areas combined. These areas provide a contrast to the more undeveloped areas like Area 2 and 3. Areas 5 and 6 had more “Address Out of Order” and “Address and Access do not match” anomaly codes. Some of these anomalies were due to road closings, lots with access from two streets, or apartment complexes where multiple points with the same address are spread out. Overall both areas had low anomaly percentages.

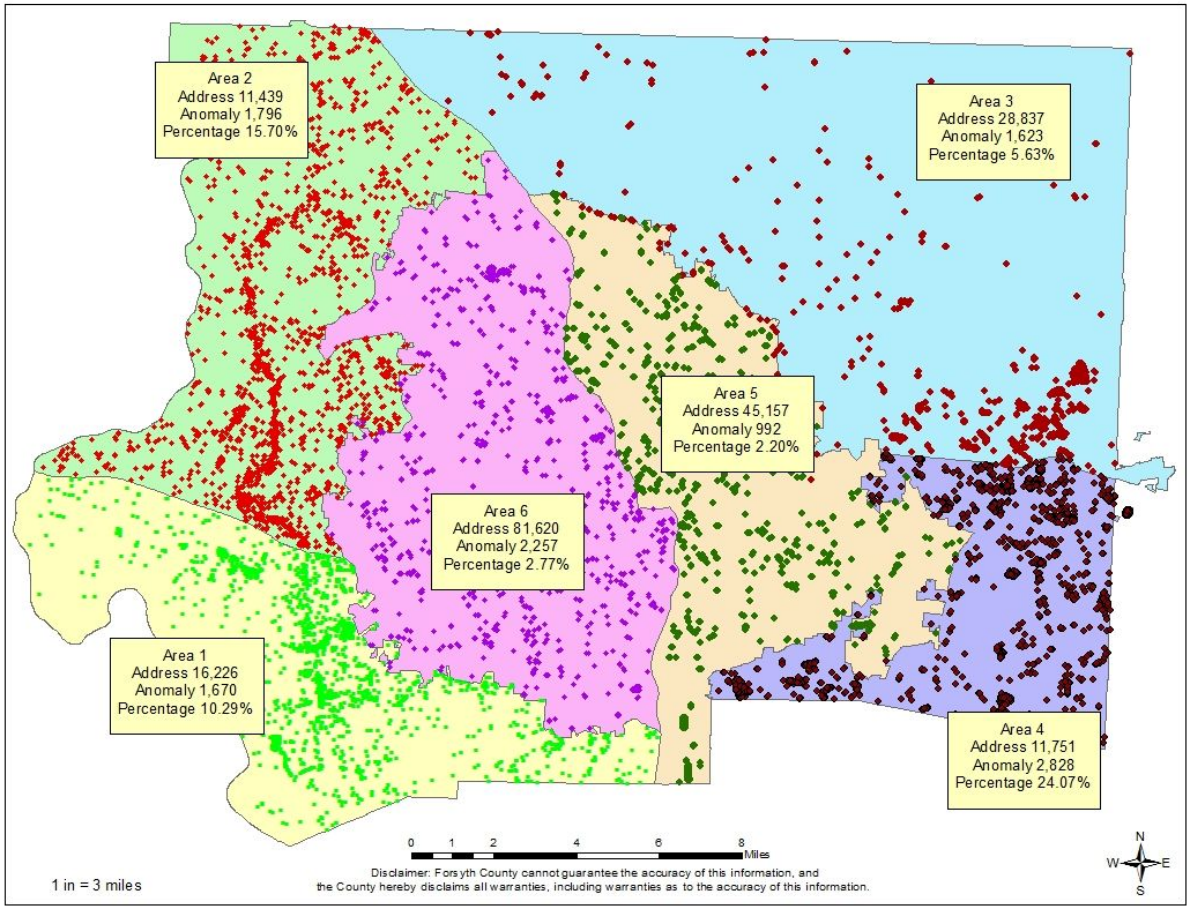


Figure 3: Address anomalies per project area

rGA MAR TOOL

The majority of Address data has been verified and validated. Now there needs to be a way to continually update the MAR and the Address GIS layers at the same time. In order to maintain a GIS database along with the MAR, MapForsyth purchased a tool embedded in the ArcGIS mapping software we currently use. This tool allows addresses and street centerlines to be added, retired, and edited in the GIS database and in the MAR simultaneously, thus preventing potential errors (Figure 4).

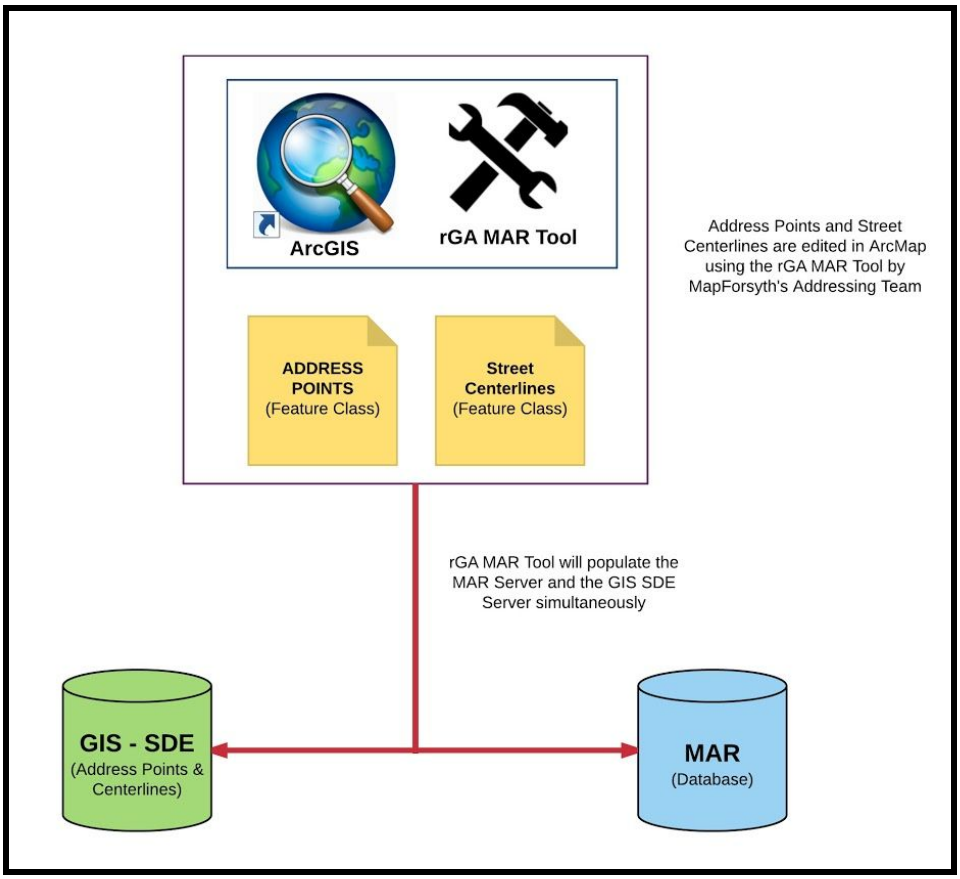


Figure 4: rGA MAR Tool Workflow

RESULTS OF PHASE II

The MAR will contain the official spelling of streets and address numbers so that any system using the MAR for verification will be creating accurate data with the correct address number and street spelling. The MAR is the official address database in the same way the Tax Office has the official database that stores all the values for the real and personal property in Forsyth County.

The results of Phase II gives MapForsyth the ability to take the anomalies found in each business system to the appropriate department where the data is stored. Each department can then decide what corrections to make to their data.

NEXT STEPS

The next steps for the Addressing Project is Phase III: Integration with existing systems. The original scope of integration includes the following systems:

a. Tax (County System)
b. City Works (City Work Order & Permitting System)
c. CSR (City 311 System)
d. Cayenta (Utilities' Billing System)

Since there is no magic bullet for integrating business systems with the MAR, each system integration will be treated as an individual project. This allows impacted departments the time to develop the project scope, work with any third party vendor on cost estimates, and request the funding needed for project completion. MapForsyth will work with those departments and City or County IT staff to assess and develop the project scope and funding requirements.

MapForsyth, in conjunction with the GIS Steering Committee, will continue identifying and prioritizing the list of system integrations with mission critical systems being the first to integrate.

In summary, Phase II of the Master Address Repository project is on target for completion and the final invoices from Spatial Focus have been received and are being vetted for accuracy.