

# 21st ANNUAL PA GIS CONFERENCE

The Penn Stater Conference Center Hotel  
State College, PA

June 3-5, 2013



## The Geographic Support System Initiative (GSS-I): Working with Partners for Enhanced Data Management

PAGIS: June 4, 2013

Kevin Holmes, Geographer  
US Census Bureau: Philadelphia Regional Office

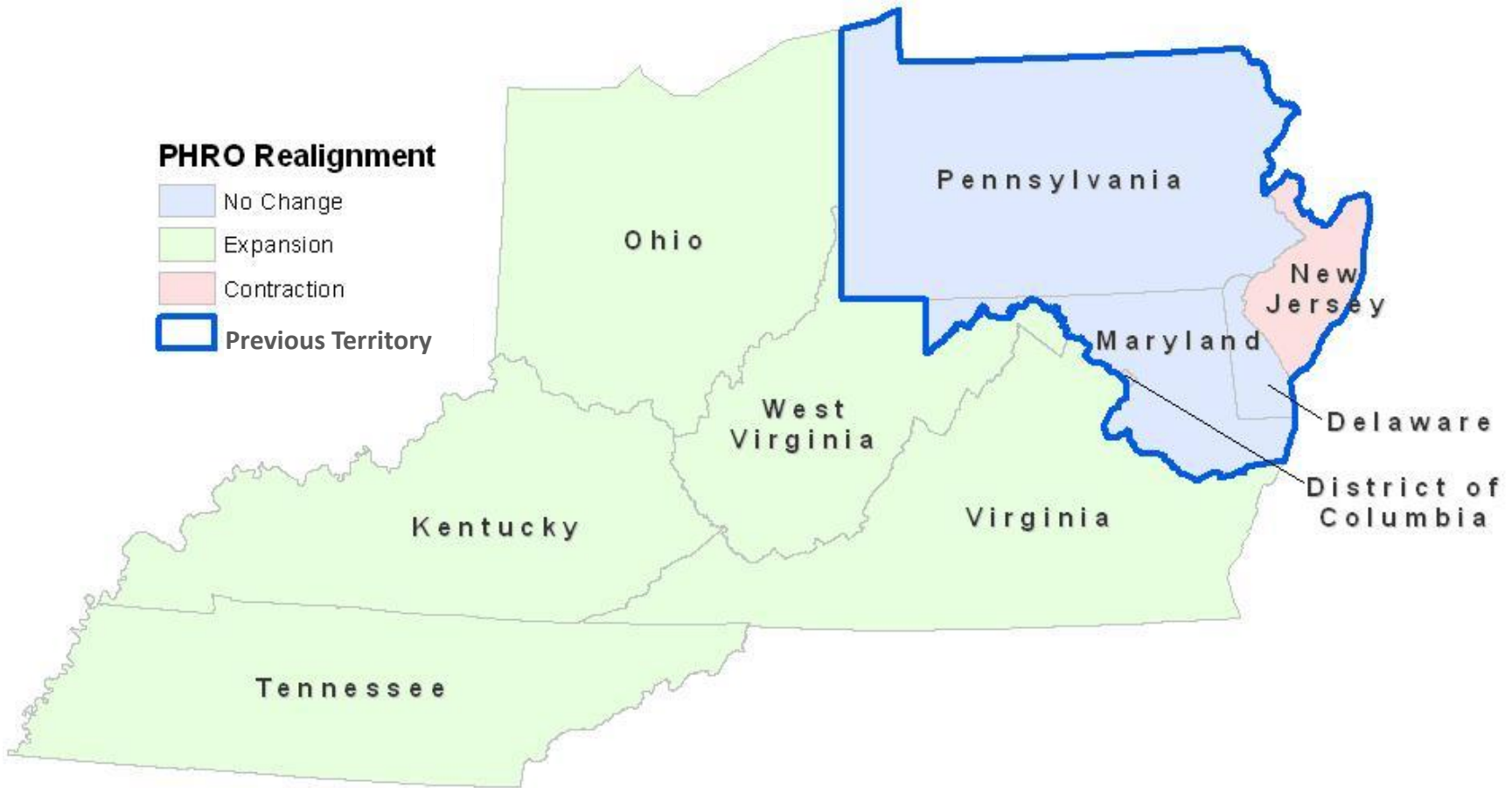


# Overview

- MAF/TIGER Database (MTDB) recap
- Supporting the 2010 Census (MTAIP)
- Supporting the 2020 Census + (GSSI)
  - Inception
  - 2019 Targeted Address Canvassing
  - GSS-I News/Updates
  - Minimal & Optimal Guidelines
  - Next Steps
- ESRI Community TIGER



# Changes in PHRO Territory: January 2013



# The MAF/TIGER Database (MTDB)



- **M**aster **A**ddress **F**ile (MAF)
- **T**opologically **I**ntegrated **G**eographic **E**ncoding and **R**eferencing (TIGER)
- Developed to support 1990 Census
  - Digitized from USGS quads





# The “TI” in TIGER

- **T**opological **I**ntegration amongst all linear feature classifications
- Efficient storage (national coverage!)
- Helps prevent tabulation data gaps (open polygons)
- Poses Conflation “Rubbersheeting” challenges

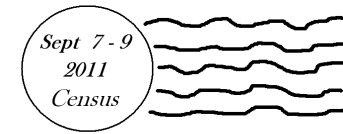


# MAF/TIGER: Supporting the 2010 Census

- MAF/TIGER Accuracy Improvement Project (MTAIP): 2003-2008
- Address Canvassing: Spring 2009
- Geographic Partnership Programs
  - Local Update of Census Addresses (LUCA)
  - New Construction (NC)
  - Participant Statistical Areas Program (PSAP)
  - Redistricting Data Program (RDP)
  - Boundary & Annexation Survey (BAS)



# The Census Address Summit

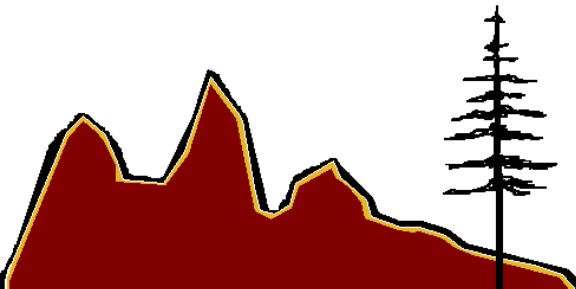


- Objectives

- Education on the benefits of a targeted address canvass
- Shared understanding of the definition of an address
- Learning how various stakeholders work with addresses
- Brainstorming and discussion of pilot projects



Geographic Support System Initiative (GSS-I)



Census Address Summit September 7-9, 2011



# MAF/TIGER: Supporting the 2020 Census +

- Geographic Support System Initiative (GSS-I)
  - Integrated program utilizing partnerships for:
    - Improved address coverage
    - Continual spatial feature updates
    - Enhanced quality assessment & measurement (QI)
  - May allow for a targeted address canvassing operation in 2019 (cost avoidance)
  - Better quality throughout the decade to support intercensal surveys (ACS)



<http://www.census.gov/geo/www/gss/>





# GSS-I Partnership Program Update

- Launched the GSS-I Partnership Program in October 2012
- Identified 56 initial partners to participate in 'Phase 1' by providing their addresses, structure points, and street centerlines
- Goal is to acquire and use partner data for a production test of our process and software first half FY13



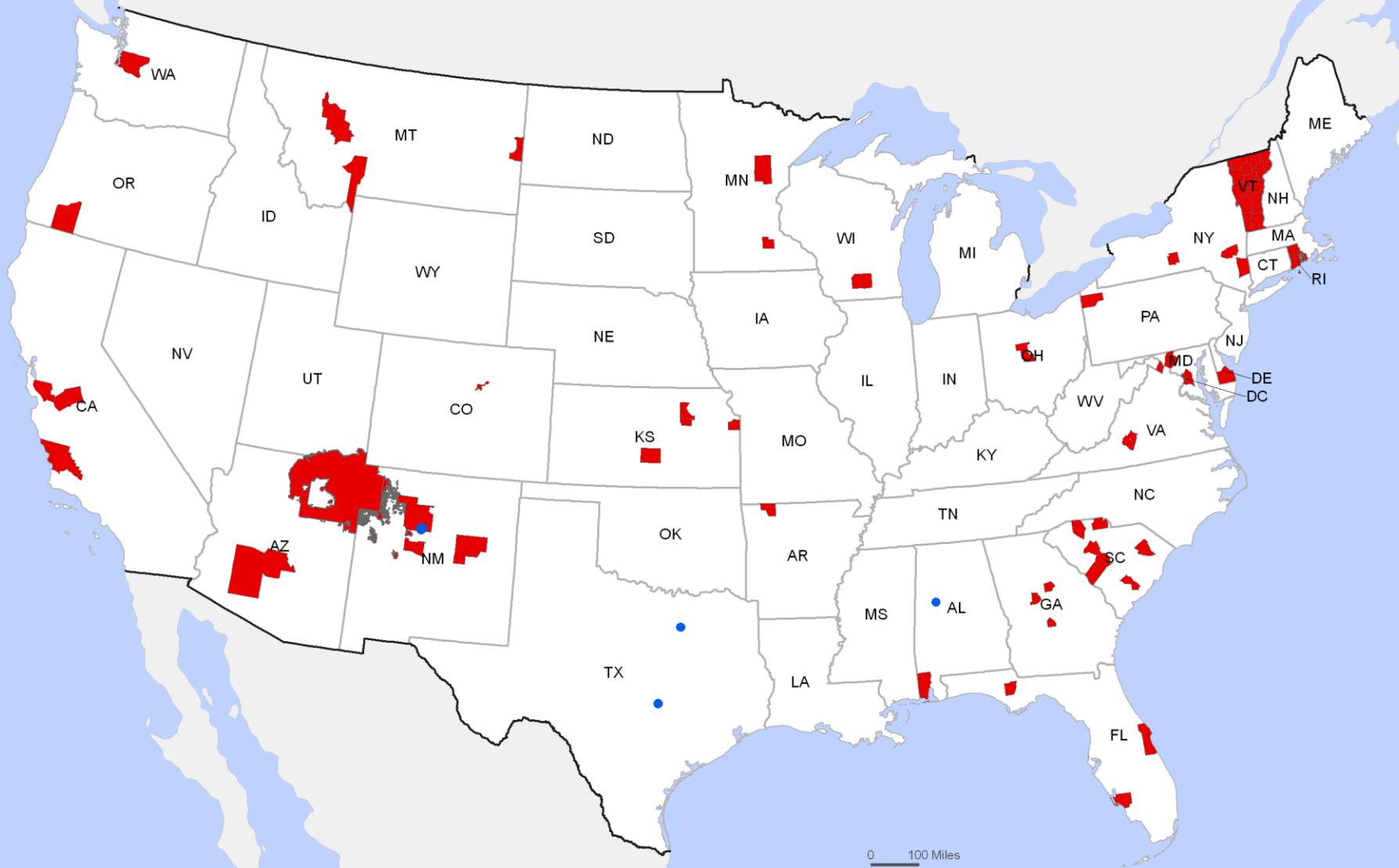
# GSS-I Phase 1 Partners

TR0161TA2430	Navajo Nation Reservation	
CO01097	Mobile	AL
PL0177256	Tuscaloosa	AL
CO05015	Carroll	AR
CO04013	Maricopa	AZ
CO06047	Merced	CA
CO06079	San Luis Obispo	CA
CO06085	Santa Clara	CA
CO08031	Denver	CO
ST11	District of Columbia	DC
CO10005	Sussex	DE
CO12009	Brevard	FL
CO12013	Calhoun	FL
CO12071	Lee	FL
CO13207	Monroe	GA
CO13235	Pulaski	GA
CO13237	Putnam	GA
CO15001	Hawaii	HI
CO15007	Kauai	HI
CO20061	Geary	KS
CO20091	Johnson	KS
CO20155	Reno	KS
CO20161	Riley	KS
CO24021	Frederick	MD
CO24033	Prince George's	MD
CO27001	Aitkin	MN
CO27131	Rice	MN
CO30031	Gallatin	MT
CO30049	Lewis and Clark	MT
CO30109	Wibaux	MT

CO35019	Guadalupe	NM
CO35043	Sandoval	NM
CO35061	Valencia	NM
PL3506970	Bernalillo	NM
PL3517960	Corrales	NM
PL3563460	Rio Rancho	NM
CO36027	Dutchess	NY
CO36039	Greene	NY
CO36109	Tompkins	NY
CO39041	Delaware	OH
CO39101	Marion	OH
CO41029	Jackson	OR
CO42039	Crawford	PA
ST44	Rhode Island	RI
CO45003	Aiken	SC
CO45035	Dorchester	SC
CO45041	Florence	SC
CO45063	Lexington	SC
CO45071	Newberry	SC
CO45083	Spartanburg	SC
CO45091	York County	SC
PL4819000	Dallas City	TX
PL4857176	Pflugerville	TX
CO51019	Bedford	VA
ST50	Vermont	VT
CO53033	King	WA
CO55025	Dane	WI
CO54037	Jefferson	WV



# Phase 1 Participants



0 100 Miles

# Minimum Address Guidelines: Matching

A. In order to perform a match to existing MAF addresses, the submitted record must include:

- Complete Address Number
- Complete Street Name

and AT LEAST ONE OF THE FOLLOWING:

- Address Coordinate
- ZIP Code
- Postal City and State
- Census 2010 Tabulation State, County, Tract and Block Code





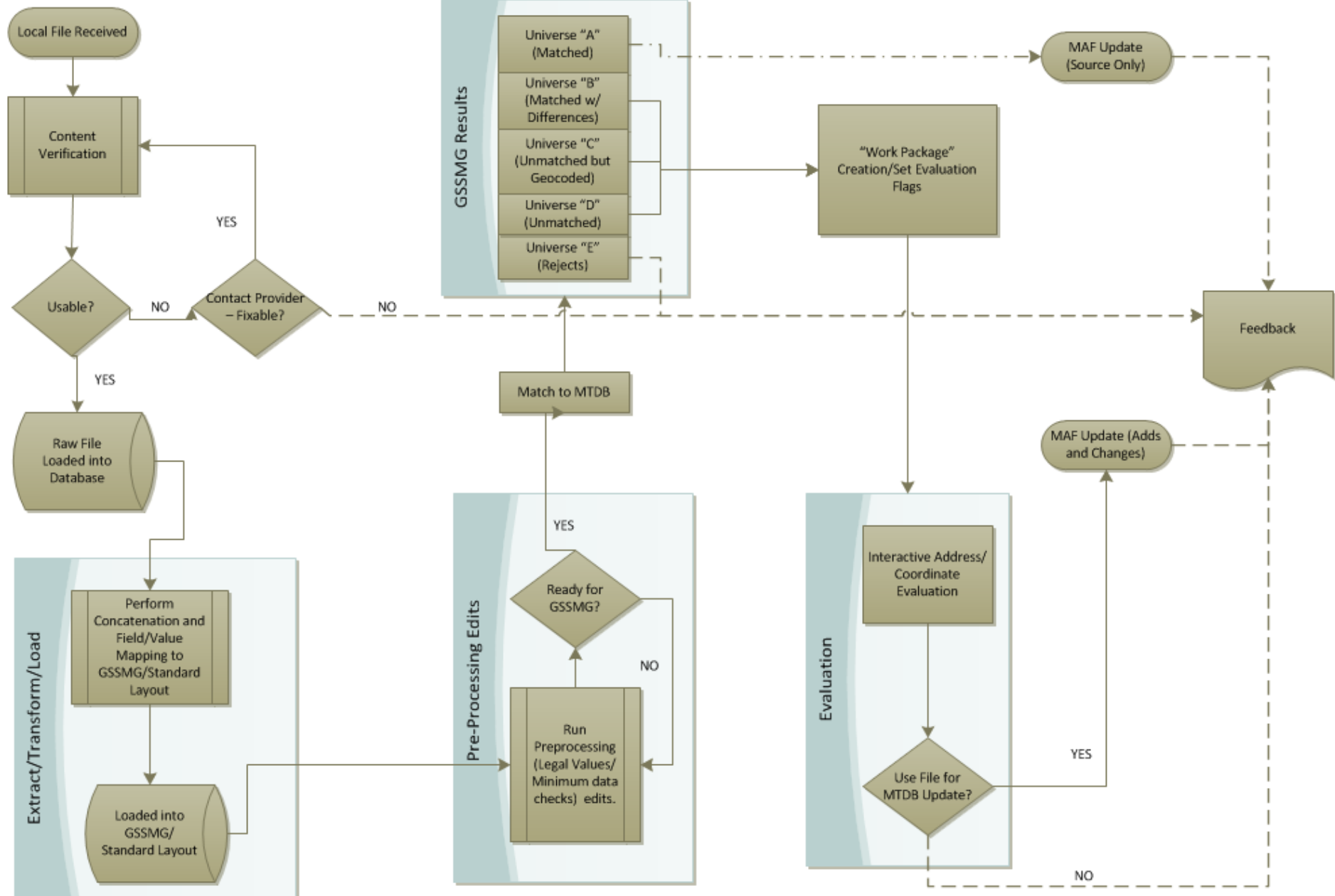
# Minimum Address Guidelines: Geocoding

- B.** In order to update the location (geocode) for an existing MAF address, the submitted record must meet the requirements of “A” above, but this action requires either:
- Address Coordinate or
  - Census 2010 Tabulation State, County, Tract and Block Code
- C.** In order to ADD new records to the MAF, the submitted record must meet the requirements of “A” above, and additionally must include an Address Feature Type indicator identifying the address as residential, commercial, utility, etc.

# Minimum Address Guidelines: Etc.

- D. Within Structure Identifiers (Apt 3, etc). If not available:
- # of units
  - Multi-unit structure flag
- E. Group Quarters:
- NAME (i.e. Shady Acres Retirement Home)
  - TYPE (i.e. Hospital, Prison, College Dormitory)
- F. Cannot process only Non-City-Style Addresses, such as:
- Rural Route Addresses (i.e. RR 3 Box 725 Anytown, NC 28999)
  - Post Office Box Addresses (i.e. P.O. Box 12374 Anytown, NC 28999)
  - General Delivery (i.e. General Delivery Anytown, NC 28999)
  - Location Descriptions (i.e. Brick House at intersection of 1st and Main Streets)

# Address Data Receipt/Evaluation Process Flow



# Preliminary Processing Results

For Two States and 45 Counties...

	National	
Total Records	6,829,527	
Duplicate Records	226,408	
Total Unduplicated	6,603,119	100.0%
Matched to MTDB	5,655,678	85.7%
Unmatched	947,441	14.3%
Unmatched Residential	430,421	6.5%
Newly Geocoded	72,630	1.1%
Same Block	5,304,000	80.3%
Different Block	278,003	4.2%
Same Class*	3,731,084	66.0%
Different Class*	112,483	2.0%

~ 10% of the total records had no address classification code.





# Next Steps

- 'Phase 1' feedback
- Planning for @ 300-500 'Phase 2' partners
  - Criteria-driven (*Targeted Address Canvassing research, data quality issues in MAF/TIGER, existing BAS participants, whole states*)  
**\*Centre & Cumberland Counties, PA**
- Community TIGER Integration
- Phase 2 feedback flowing starting in summer 2013



# Community TIGER



- Proof of Concept collaborative project with ESRI
- Web (cloud) based data exchange and data management portal
- Phased and iterative project
- Leverages COTS technology, existing systems and proven workflows
- Utilizes and builds upon the next generation ESRI Community Maps



# Address Data Migration

The screenshot displays the Step1\_PartnerDataLoader\_ETL software interface. The main workspace shows a data flow diagram with several transformers: Attribute Renamer, Attribute Value Mapper (Street Suffix Direction), Attribute Value Mapper (Series), Attribute Creator, Attribute Trimmer, String Concatenator (FULLADDR), Case Changer, and UUID Generator. A callout window for the Attribute Renamer transformer is open, showing a table of attributes to be renamed.

Old Attribute	New Attribute	Default Value
SUFFIX_DIR	ST_SUFFIXDIR	k
STREET_TYP	ST_TYPE	k
STREET_NAM	ST_NAME	k
STATUS	STATUS	k
PREFIX_DIR	ST_PREFIXDIR	k
OBJECTID	ADDPKEY	k
OBJECTID	SOURCEID	k
Apt_No	UNITID	k
ADDRESS	ADDRNUM	k
SUFFIX_DIR	ST_SUFFIXDIR	k

The interface also includes a Navigator on the left showing a tree view of data sources and features, and a Transformer Gallery at the bottom left. The status bar at the bottom indicates the current transformer is AttributeRenamer [AttributeRenamer].

## Community TIGER Partner Tool Set

- Data Model
- Data Migration Tool(s)
- Address Data Standardization and De-Duplication Tool
- Data Reviewer Batch Job
  - Address Validation
  - Spatial Validation
  - Metadata Validation
- Data Package Tool
- Procedure document



# Address Data Validation

ctapp2 - Remote Desktop Connection

dr\_plugerville.mxd - ArcMap

File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help

1:2,573

Data Reviewer Select Data Check

Table Of Contents

Layers

- RoadCenterline
- AddressPoint

**Batch Validate**

Features to Validate

- Selection Sets
- Current Extent
- Definition Query
- Full Database
- Changed Features Only

Batch Job

Add from File...  
Add from Product Library...  
Remove...  
Validate All  
Workspace...

Batch Jobs

- Default (Click 'Validate' to update)
  - AddressPoint Attribute Validation
    - AddressPoint ADDRESS is NULL
    - AddressNumberType is NULL
    - AddressPoint ST\_NAME is NULL
    - AddressClass is NULL
    - AddressPoint ST\_TYPE is NULL
    - AddressPoint STATE is NULL
    - AddressPoint CITY is NULL
    - AddressPoint ZIPCODE is NULL
    - AddressPoint STATUS is NULL
    - AddressPoint CaptureMethod is NULL
    - Addresspoint FULLNAME is NULL
    - AddressPoint FULLADDR is NULL
    - AddressPoint RESIDENTIAL\_FLAG is NULL
    - AddressPoint FEATUREID is NULL
    - AddressPoint SOURCEID is NULL
  - Domain Checks
    - AddressPoint Must Have Valid Domains
  - Topology Check
    - Duplicate Geometry
    - AddressPoint Has Invalid Geometry

Validation Status: Not validated

Run Cancel

Reviewer Table

General

Drag a column header here to group by that column.

VERCODE	SESSIONID	CHECKTITLE	ORIGIN	ORIGINCHECK	NOTES	PARAMETERS	SEVERITY	REVIEWSTATUS	IDENTIFIER	REVIEWTECHNICIAN	REVIEWDATE	CORRECTIONTECHNICIAN	CORRECTIONDATE
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_SUFFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_TYPE: Invalid	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_PREFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_SUFFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_TYPE: Invalid	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_PREFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_SUFFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_TYPE: Invalid	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_PREFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_SUFFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_TYPE: Invalid	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_PREFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_SUFFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_TYPE: Invalid	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_PREFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_SUFFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_TYPE: Invalid	<Null>	bala3756	Wednesday, ...	<Null>	<Null>
	1	AddressPoint	AddressPoint	Domain Check		[Connection1: ...]	1	ST_PREFIXDIR:	<Null>	bala3756	Wednesday, ...	<Null>	<Null>

Notepad



# Spatial Matching with TIGER



# Spatial Data Reviewer

The screenshot displays the ArcMap Spatial Data Reviewer interface. The main map area shows a road network with green diamond markers. The 'Reviewer Table' at the bottom left contains the following data:

RECORDID	OBJECTID	SUBTYPE	REVIEWERCATEGORY	CATEGORY	REVIEWERCODE
1	1		1	Feature Record	228
2	2		1	Feature Record	228
3	3		1	Feature Record	228
4	4		1	Feature Record	228
5	5		1	Feature Record	228
6	6		1	Feature Record	228
7	7		1	Feature Record	228

The 'Batch Validate' dialog box is open on the right, showing a list of validation checks. Several checks are circled in red:

- Domain Checks
  - Road Centerline Must Have Valid Domains
  - Road Centerline Attribute Validation
    - Road Centerline From Left Address Must Not Be Set To
    - Road Centerline From Right Address Must Not Be Set To
    - Road Centerline To Left Address Must Not Be Set To 0
    - Road Centerline To Right Address Must Not Be Set To
    - Road Centerline From Left Address Must Be Odd
    - Road Centerline From Right Value Greater Than To Ri
    - Road Centerline To Left Address Must Be Odd
    - Road Centerline From Right Address Must Be Even
    - Road Centerline To Right Address Must Be Even
    - Road Centerline From Left Value Is Greater Than To Le
    - Roadname Must Not Be NULL
    - RoadCenterline ROADTYPE is NULL
    - RoadCenterline RoadClass is NULL
    - ST\_PROV Must Not Be NULL
    - ZIPLEFT Must Not Be NULL
    - ZIPRIGHT Must Not Be NULL
    - RoadCenterline\_LIFECYCLESTATUS is NULL
    - RoadCenterline\_FeatureID is NULL
    - RoadCenterline\_SOURCEID is NULL
    - RoadCenterline\_CTDeltaFlag is Not NULL
- Topology
  - Road Centerline Has Invalid Geometry
  - Topology - Overshoot and Undershoot Check
  - RoadCenterline: Vertices Must Not be within 1.0 M
  - Topology - Orphan RoadCenterline Features
  - Duplicate Geometry
  - Road Cutback Check (< 45 deg angle)
  - Pseudo Node Check
- AddressPoint Attribute Validation
  - AddressPoint ADDRNUM is NULL
  - AddressNumberType is NULL
  - AddressPoint ST\_NAME is NULL
  - AddressClass is NULL
  - AddressPoint ST\_TYPE is NULL
  - AddressPoint STATE is NULL
  - AddressPoint CITY is NULL

Buttons in the dialog include 'Add from File...', 'Add from Product Library...', 'Remove...', 'Validate All', and 'Workspace...'. The 'Run' and 'Cancel' buttons are at the bottom right. The validation status at the bottom is 'Validation Status: Not validated'.



# Data Review

## Community TIGER: Contribution Management

Active Contribution

Contribution History

Notifications **0**

Lilin's Test

### Upload Contribution



#### Upload Data

✓ Received January 31st, 2013

#### 2 Preview Your Data >

- [Review data analysis results](#)
- [Preview your map](#)
- Optionally upload revised data to resolve issues

Upload revised data

#### 3 Finalize Contribution

Review your map

#### 4 Feedback from Census

Not submitted yet

### Data Analysis

[View details](#)

#### Automated QC Results



[Download as file gdb](#)

#### Layers

<a href="#">RoadCenterline</a>	99%	<b>2</b>	<b>5</b>	<b>6</b>
<a href="#">AddressPoint</a>	98%	<b>3</b>	<b>4</b>	<b>9</b>

### Content Analysis

[Show on map](#)

Suggested TIGER updates to be sent to Census:

RoadCenterline(adds)	12
RoadCenterline(modified)	22
AddressPoint(adds)	20
AddressPoint(modified)	5



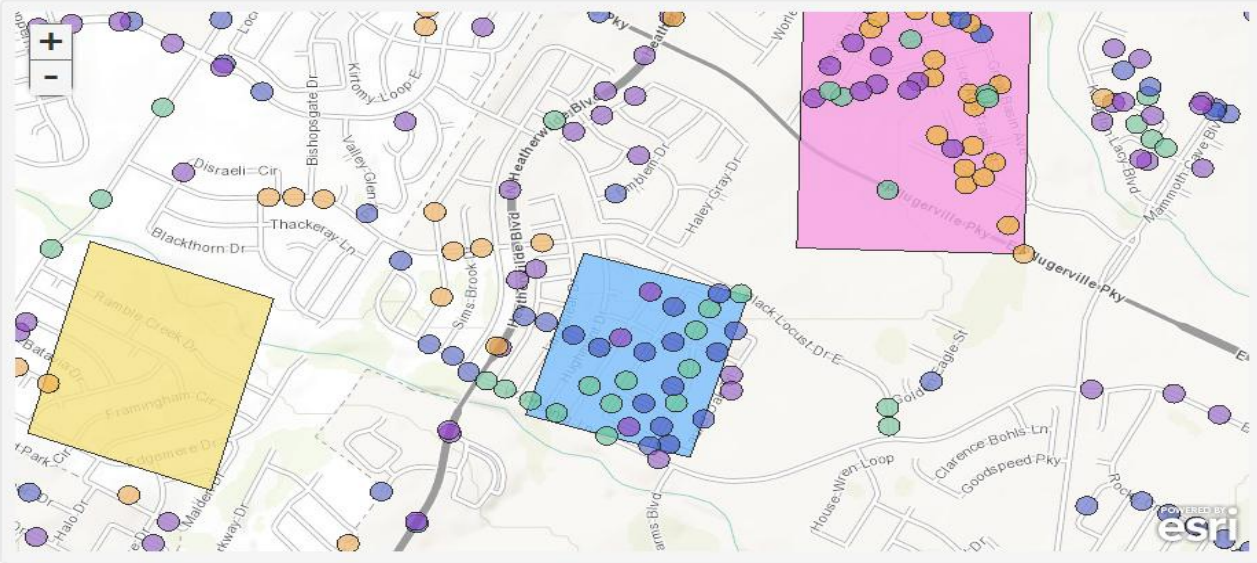
# Contribution Feedback

Community TIGER: Contribution History

Active Contribution | Contribution History | Notifications 0 | Lili's Test

Contribution History [Back to Contribution History](#)

Census Feedback



Map legend:

- Modified (spatial)
- Modified (attribute)
- Modified (both)
- Added

hide table

All 59 | Adds 32 | Modified 27 | Delete 0

ID	Proposed Submission Type	Census Acceptance Status	Actions
1	Add	Accepted	<a href="#">Approve</a>
2	Add	Accepted	<a href="#">Approve</a>
3	Add	Accepted	<a href="#">Approve</a>
4	Modified	Accepted with modifications	<a href="#">Approve</a>
5	Modified	Not used	<a href="#">Approve</a>







<http://www.census.gov/geo/www/gss/>

# Questions?

[kevin.j.holmes@census.gov](mailto:kevin.j.holmes@census.gov)

215.717.1830

