

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

KY GIS Conference, 10/01/13

Kevin Holmes, Geographer
US Census Bureau: Philadelphia Regional Office

United States™
Census
Bureau
Philadelphia Region



Sorry folks, America's closed.



Moose out front shoulda told you.

United States™
Census
Bureau
Philadelphia Region



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
- Geographic Sundries



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



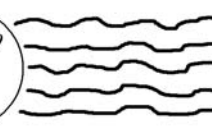
MAF/TIGER: Supporting the 2010 Census

- 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

Sept 7-9
2011
Census

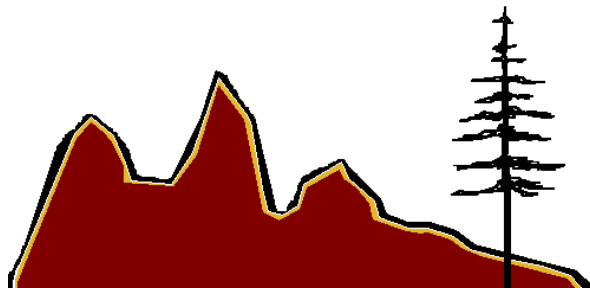


- Objectives

- 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, Pop. Estimates)



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



GSS-I Partnership Program: FY13

- Launched GSS-I Partnership Program in October '12
- Identified ~56 initial partners to participate in 'Phase 1' by providing their addresses, structure points, and/or street centerline datasets
- Goal is to acquire and use partner data for a production test of our process and software first ½ of FY13
- 'Phase 2' expanded production universe



Partner Data Acquisition in FY13

<i>Data as of June 20, '13</i>	Partners Contacted	Address List Acquired	Structure Coordinates Acquired	Street Centerlines Acquired	Use Agreement or Fee Restriction	No Data Available	Data Not Usable
Phase 1	75	61	60	61	6	7	1
Phase 2	180	100*	155	228**	10	10	5
TOTAL	255	161	215	289**	16	17	6

* 5 partners required a signed Memorandum of Understanding prior to address delivery

** Some counties provided multiple partial-coverage street centerline datasets (i.e., cities vs. balance of county)

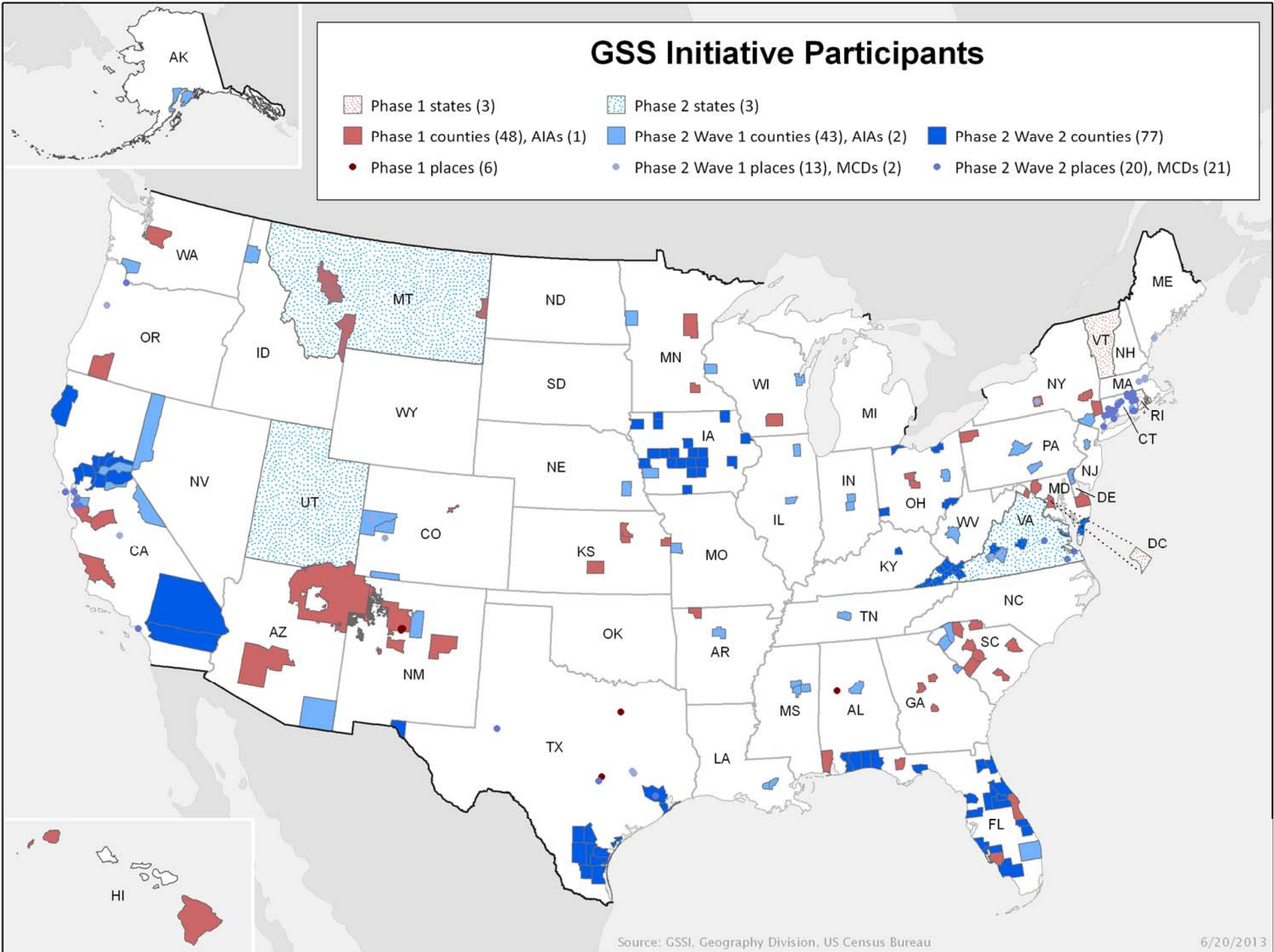
Other Acquisition Notes:

- 68 partner files contained parcel data
- 110 contained Personally Identifiable Information (PII) that was scrubbed by the Census Bureau
- The 255 partner datasets = **3,939 total governmental units represented**

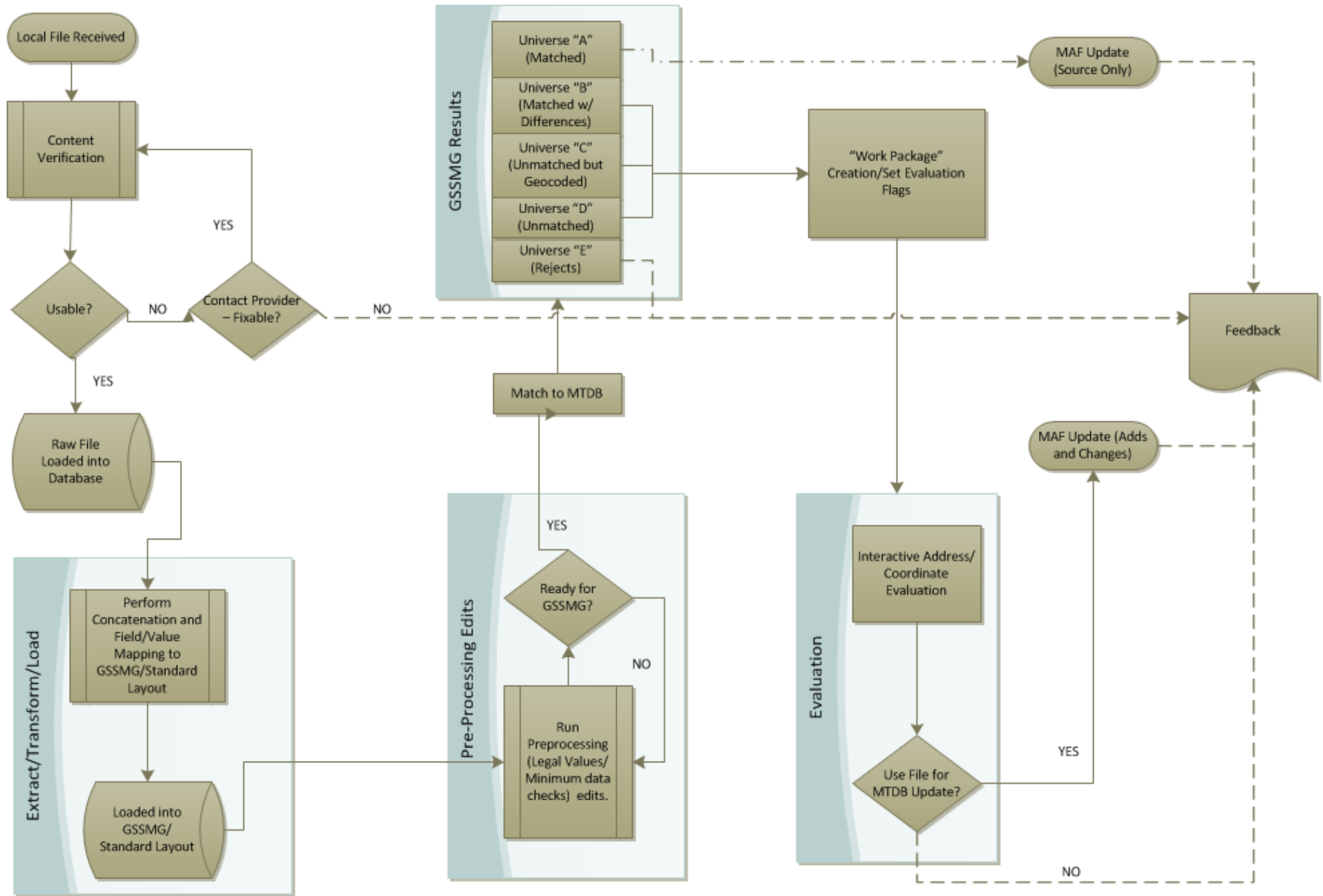
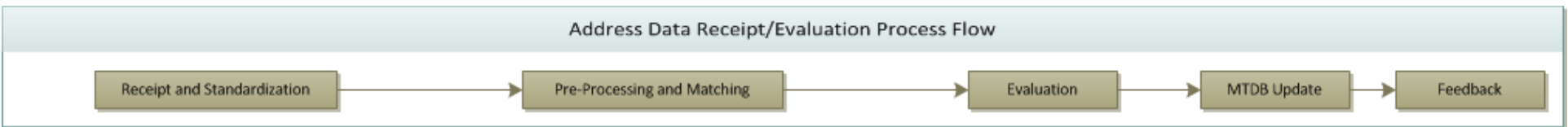


GSSI Initiative Participants

- Phase 1 states (3)
- Phase 2 states (3)
- Phase 1 counties (48), AIAs (1)
- Phase 2 Wave 1 counties (43), AIAs (2)
- Phase 2 Wave 2 counties (77)
- Phase 1 places (6)
- Phase 2 Wave 1 places (13), MCDs (2)
- Phase 2 Wave 2 places (20), MCDs (21)



Address Data Receipt/Evaluation Process Flow



Minimum Address Guidelines: Address 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: Address 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)



Total Partner Addresses Received (6/20/13)	6,896,785			
Duplicates Identified within Partner Datasets		226,697		
Total Unduplicated Partner Addresses	6,670,088		100 %	
Total Partner Addresses <u>Matched</u> to Master Address File (MAF)	5,720,449		86%	
Addresses in Same Block as MAF		5,366,058		94% of Matched <i>(81% of Total)</i>
Addresses in Different Block than MAF		280,106		5% of Matched <i>(4% of Total)</i>
Partner & MAF Addresses Ungeocoded		74,285		1.3% of Matched <i>(1% of Total)</i>
Addresses w/ Same Address Class as MAF (residential, etc.)		3,794,773		66% of Matched <i>(57% of Total)</i>
Addresses w/ Different Address Class than MAF		113,515		2% of Matched <i>(1.7% of Total)</i>
Addresses w/ No Address Class Identified		1,812,161		32% of Matched <i>(27% of Total)</i>
MAF Addresses <u>Newly Geocoded</u> via Partner Information	73,240		1%	
Total Partner Addresses <u>Unmatched</u> to MAF	949,639		14%	
Unmatched Class Residential		431,939		45% of Unmatched <i>(6% of Total)</i>
Unmatched Class Commercial/Other/Unspecified		517,700		55% of Unmatched <i>(8% of Total)</i>

GSS-I Next Steps : FY14

- FY13 Phase 1 & 2 feedback
- Identifying FY14 GSS-I partners
 - Estimated ~300 partners
 - Criteria-driven (*Targeted Address Canvassing research, data quality issues in MAF/TIGER, existing BAS participants, whole states*)
- Community TIGER Integration
- WFS Acquisition Testing
- Cooperation w/ other Federal agencies (USGS)



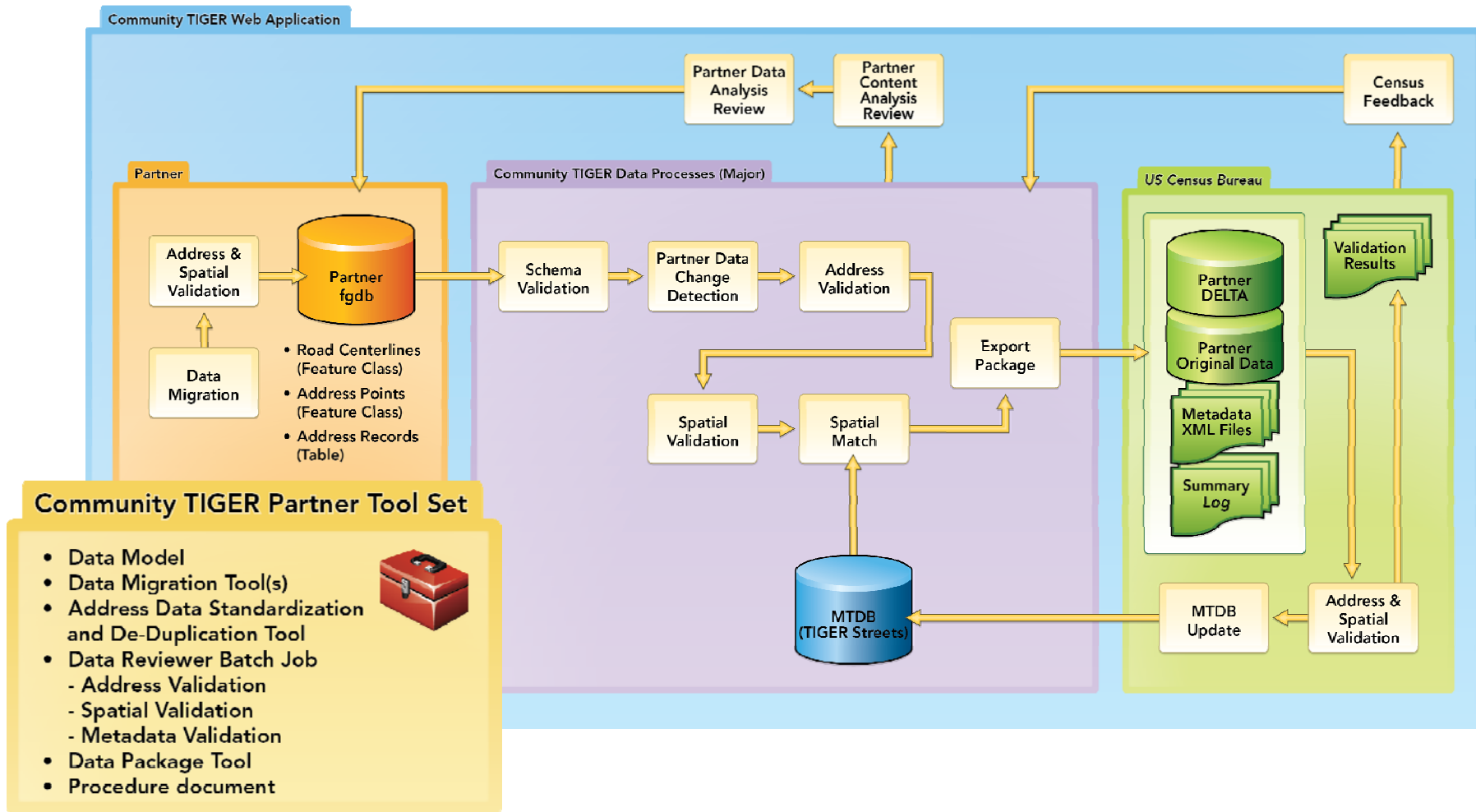
Community TIGER



- 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
- Local govt. GSS-I Phase 2 beta testers



Community TIGER – Partner Toolbox



Address Data Migration

The screenshot displays the QGIS Data Designer interface for an ETL workflow titled "Step1_PartnerDataLoader_ETL". The main workspace shows a sequence of transformers: an "Attribute Renamer" transformer, followed by five "Attribute Value Mapper" transformers, and a "String Concatenator" transformer. The "Attribute Renamer" transformer is currently selected, and its configuration dialog is open.

The "Attribute Renamer Parameters" dialog shows the following configuration:

- Transformer Name: AttributeRenamer
- Attributes To Rename:

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 workflow also includes several "Attribute Value Mapper" transformers with the following descriptions:

- Attribute Value Mapper (Street Suffix Direction):** Maps directional values (N, E, SE, NW, etc) to the accepted values in the output file. May need to check that the values the contributor uses are included in the table.
- Attribute Value Mapper (Status):** Maps contributor values for development status to Community TIGER values. May need configuration.
- Attribute Value Mapper (Residence):** Creates default via the field RESIDEN. No configuration needed.

The "String Concatenator" transformer is configured to concatenate the fields FULLADDR and RESIDEN.

The Navigator on the left shows the project structure, including the "AddressPoint" feature type and its attributes:

- ADDPKEY [char(20)]
- ADDRCLASS [Integer]
- ADDRESSNUMBERTYPE [char(2)]
- ADDRNUM [char(20)]
- ADDRNUMSUF [char(5)]
- ADDRRANGE [char(50)]
- ALTUNITID [char(10)]
- ALTUNITTYPE [char(10)]
- CAPTUREMETH [char(30)]
- CENSUSADDRESSCATEGORY [char(1)]
- CENSUSADDRESSQUALITY [char(1)]
- CITY [char(50)]
- CONTRIBUTORID [char(38)]
- DELIVERYID [char(38)]
- ESN [char(5)]
- FEATUREID [char(38)]
- FULLADDR [char(250)]
- FULLNAME [char(125)]
- geodb_old [Integer]
- LASTEDITOR [char(50)]
- LASTUPDATE [date]
- LATITUDE [double]
- LATITUDE_CT_GEOCODE [double]
- LONGITUDE [double]
- LONGITUDE_CT_GEOCODE [double]
- MSAG [char(35)]
- MUNICIPALITY [char(100)]
- OBJECTID [Integer]
- PLACENAME [char(250)]
- POINTTYPE [char(30)]
- PREADDRNUM [char(5)]
- PSAP [char(4)]
- RESIDENTIAL_FLAG [char(1)]

The Transformer Gallery at the bottom left shows various transformer categories, and the Log window at the bottom right is currently empty.



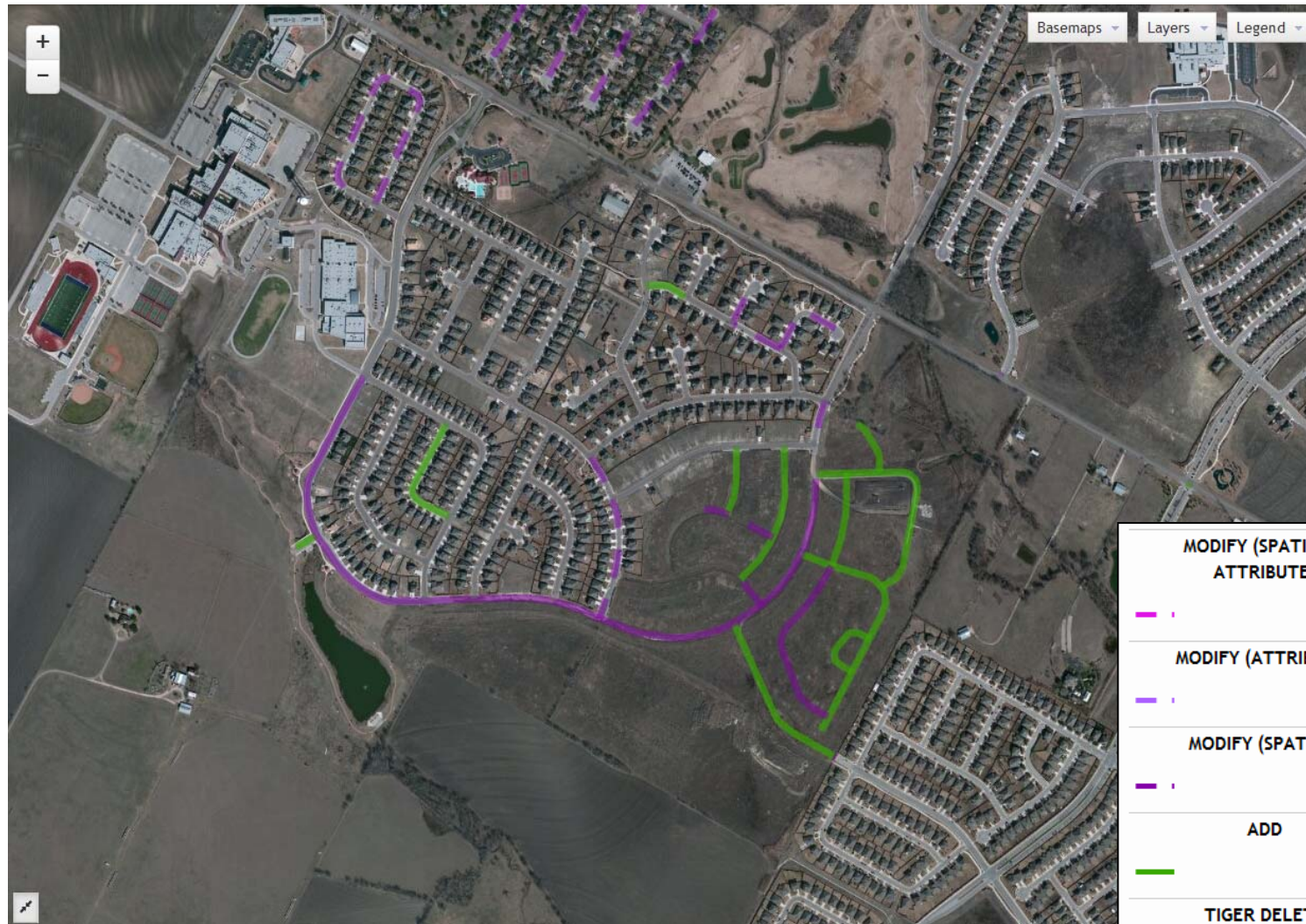
Address Data Validation

The screenshot shows the ArcMap interface with a 'Batch Validate' dialog box open. The dialog box has several sections:

- Features to Validate:** Includes options for Selection Sets, Current Extent, Definition Query, Full Database, and Changed Features Only.
- Batch Job:** Includes buttons for 'Add from File...', 'Add from Product Library...', 'Remove...', 'Validate All', and 'Workspace...'.
- Batch Jobs:** A tree view showing validation rules:
 - 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

- Reviewer Table:** A table with the following columns: VERCODE, SESSIONID, CHECKTITLE, ORIGINTABLE, ORIGINCHECK, NOTES, PARAMETERS, SEVERITY, REVIEWSTATUS, IDENTIFIER, REVIEWTECHNICIAN, REVIEWDATE, CORRECTIONTECHNICIAN, CORRECTIONL. The table contains 15 rows of data, all with a severity of 1 and a review status of 'Invalid'.


Spatial Matching with TIGER



Spatial Data Reviewer

The screenshot displays the ArcMap Spatial Data Reviewer interface. The main map area shows a network of roads with green diamond markers indicating review points. The 'Table of Contents' on the left shows two layers: 'RoadCenterline' and 'AddressPoint'. 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 rules. Several rules are circled in red:

- Road Centerline Must Have Valid Domains
- Road Centerline Attribute Validation
 - Road Centerline From Left Address Must Not Be Set To 0
 - Road Centerline From Right Address Must Not Be Set To 0
 - Road Centerline To Left Address Must Not Be Set To 0
 - Road Centerline To Right Address Must Not Be Set To 0
 - 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
- Topology
 - Road Centerline Has Invalid Geometry
 - Topology - Overshoot and Undershoot Check
 - Road Centerline: 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

The 'Batch Validate' dialog also includes options for 'Features to Validate' (Current Extent, Definition Query, Full Database, Changed Features Only) and buttons for 'Add from File...', 'Add from Product Library...', 'Remove...', 'Validate All', and 'Workspace...'. The 'Run' button is visible at the bottom right.

Spatial Data Sharing Options

- After Data Validation and Spatial Checks, Option to Share Data with:
 - Census Bureau, Community TIGER, Esri Community Maps
 - Census Bureau, Community TIGER
 - Census Bureau



Data Review

Community TIGER: Contribution Management

Active Contribution

Contribution History

Notifications **0**

Lilin's Test

Upload Contribution



Upload Data

✓ Received January 31st, 2013



Preview Your Data >

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

Upload revised data



Finalize Contribution

Review your map



Feedback from Census

Not submitted yet

Data Analysis

[View details](#)

Automated QC Results

5
Errors

9
Warnings

15
Messages

[Download as file gdb](#)

Layers

RoadCenterline	99%	2	5	6
AddressPoint	98%	3	4	9

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

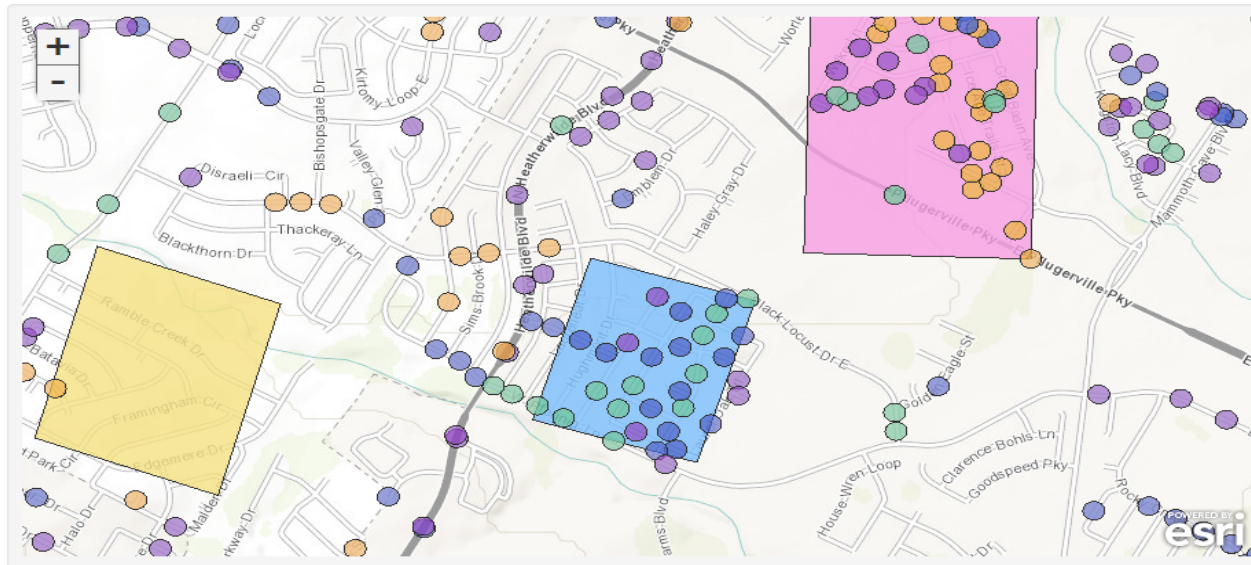
Notifications 0

Lilin's Test

Contribution History

[Back to Contribution History](#)

Census Feedback



- 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	Approve		
2	Add	Accepted	Approve		
3	Add	Accepted	Approve		
4	Modified	Accepted with modifications	Approve		
5	Modified	Not used	Approve		



A La Carte: Geographic Products

TIGER/Line Shapefiles

Full detail, extensive attributes, current vintages

TIGER Geodatabases

Full detail, limited attributes, current vintages

TIGER/Line Shapefiles
& Geodatabases w/
Demographic Data

Full detail, limited attributes, limited vintages

Cartographic Boundary
Files

Generalized detail, limited attributes, limited vintages

KML Prototype Files

Generalized detail, limited attributes, limited vintages

TIGERWEB – WMS &
REST map services

Detailed, extensive attributes, current vintages



A La Carte: Geographic Input

ACTIVE
PENDING
TBD

Structured

- GSS-I
- Yearly Boundary & Annexation Survey (BAS; 2015)
- More as 2020 approaches (2017+)

Informal

- Tract-level change detection alerts → PHRO
- Highlighting Change AOIs via Community TIGER
- VGI?





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

Questions?

kevin.j.holmes@census.gov

215.717.1830

United States™
Census
Bureau
Philadelphia Region

