Building a National Address Database

Presented by
Steve Lewis, Department of Transportation
Mark Lange, Census Bureau

July 13, 2017
Topics

- Building the NAD
- FGDC Address Theme
Building the NAD
First Step: The National Address Database Summit

- Held April 8-9, 2015 at the Maritime Institute in Linthicum, MD
- Funded by DOT/Bureau of Transportation Statistics
- Objective: To identify and discuss possible options for developing a National Address Database (NAD)
- Broad stakeholder representation
  - Government (Federal, State, Local, and Tribal)
  - Private Sector
  - Non-Profits and Trade Organizations
How the NAD Would Work

- "Vetted" Crowdsourcing
- Data quality and error feedback
- Federal
- State 1
- State 2
- Local 1
- Local 2
- Local 3
- National geocoder

Data flow
Pilot Launch October 2015

- DOT Funded
- Included State and Local Partners
  - Arizona
  - Arkansas
  - Boone County, Missouri
  - Jackson County, Arkansas (“have not” partner)
- Goals
  - Identify the “haves” and the “have nots”
  - Explore workflows and understand best practices for address roll-up
  - Compare Data Schemas
  - Develop minimum content guideline and associated data schema
  - Explore geocoding options for “have nots”
NAD Minimum Content Guidelines

The Address Itself
- Address Number
- Street Name
- Subaddress
- City/Town/Place
- County
- State
- Zip

Geographic Location of the Address
- Lat/Long
- National Grid Coordinates

Metadata About the Address
- Address authority
- Address source
- Address date
- Unique ID
- Type
  (residential, commercial, etc.)
- Placement
  (rooftop, driveway access, etc.)
NAD Schema
Pilot Participants Compiled Into NAD Schema
Jackson County, AR Geocoding Approach

- Multiple geocoding sources were used:
  - Melissa Data (commercial geocoding service)
  - County parcels
  - County road centerlines
  - Census road centerlines

- If an address wasn’t matched in one source, the next source was used

- Achieved a 77% overall match rate from the 18,469 records
## Final Jackson County Geocoding Results

<table>
<thead>
<tr>
<th>Source</th>
<th>Total Records Matched</th>
<th>% Matched</th>
<th>Cumulative % Matched</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melissa Data</td>
<td>7,073</td>
<td>38%</td>
<td>38%</td>
</tr>
<tr>
<td>Parcel centroids</td>
<td>1,700</td>
<td>9%</td>
<td>47%</td>
</tr>
<tr>
<td>County centerline</td>
<td>4,112</td>
<td>23%</td>
<td>70%</td>
</tr>
<tr>
<td>Census/TIGER centerlines</td>
<td>1,347</td>
<td>7%</td>
<td>77%</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>14,232</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Pilot Findings

- **Tribal participation** is going to be a **challenge**
  - Lots of outreach, no explicitly contributed data

- Data sharing agreements to make data **publicly available** could be a challenge

- Aggregating existing statewide collections was **straight forward**

- **The schema will likely evolve**, but needs to remain consistent with leading address schemas to allow for streamlined ETL
Since the release of the minimum content guidelines and schema, 15 additional address programs volunteered to develop their own ETLs

- District of Columbia
- New Jersey
- Ohio
- Utah
- Virginia
- 9 additional counties and 1 city from Missouri (Locals Helping Locals)

Recently received data from Colorado, Indiana and Montana (not yet loaded into NAD)

Maryland, Massachusetts, New Mexico, New York, North Carolina, Tennessee, and Washington are in the queue

Seeking other volunteers through NSGIC
22.3 Million Addresses

Legend
- Data from NAD Pilot
- Data Received, Not Integrated
- States in Queue
What’s Next

- Continue the Coalition of the Willing
- Choose platforms for development and production
- Identify funding for continued development
  - NHTSA has provided funding to hire a contract developer resource
  - Still no dedicated, sustained funding
- Make the data available!
- Launch Data Challenge for “have nots”
FGDC Address Theme
Federal Geographic Data Committee (FGDC)
Address Theme Roles and Responsibilities

- Executive Champions
  DOT – *Steve Lewis*
  Census Bureau – *Tim Trainor*

- Address Theme Leads
  DOT - *Steve Lewis*
  Census Bureau – *Lynda Liptrap*

- Address Theme Subcommittee
  DOT – *Steve Lewis*
  Census Bureau – *Mark Lange*

- NAD Dataset Manager
  Census Bureau – *Matt Zimolzak*

- NAD Technology Operations Manager
  DOT - *Steve Lewis*
Address Theme Subcommittee

16 Federal Members
Census Bureau*
Consumer Financial Protection Bureau
Department of Education
Department of Energy
Department of Homeland Security
Department of Housing and Urban Dev.
Department of the Interior
Department of Justice
Department of Labor
Department of Transportation*
Department of Veterans Affairs
Federal Emergency Management Agency
National Technical Information Service
Office of Management and Budget
U.S. Postal Service
Social Security Administration

17 Non-federal Partners
State of Arkansas
State of California
District of Columbia
Commonwealth of Massachusetts
State of Michigan
State of Minnesota
State of Montana
National Association of State 911 Admins.
State of New Jersey
State of North Carolina
State of New Hampshire
National States Geographic Info. Council
State of Oregon
Oakland County, Michigan
City of Philadelphia
URISA
State of Vermont

* = Co-chair
Address Theme and NAD Goals

- Establish Address Subcommittee - done
- Develop Theme Definition – done
- Finalize Subcommittee Charter – in progress
- Gather and Refine NAD User Requirements – done
- Establish Theme Community on the GeoPlatform – in progress
- Assess the DOT Pilot NAD Database
- Develop a Theme Strategic Plan
Address Theme and NAD Team Contacts

Steve Lewis
steve.lewis@dot.gov
202-366-9223

Lynda Liptrap
lynda.a.liptrap@census.gov
301-763-1058

Mark Lange
mark.lange@census.gov
301-763-2660

Matt Zimolzak
matthew.a.zimolzak@census.gov