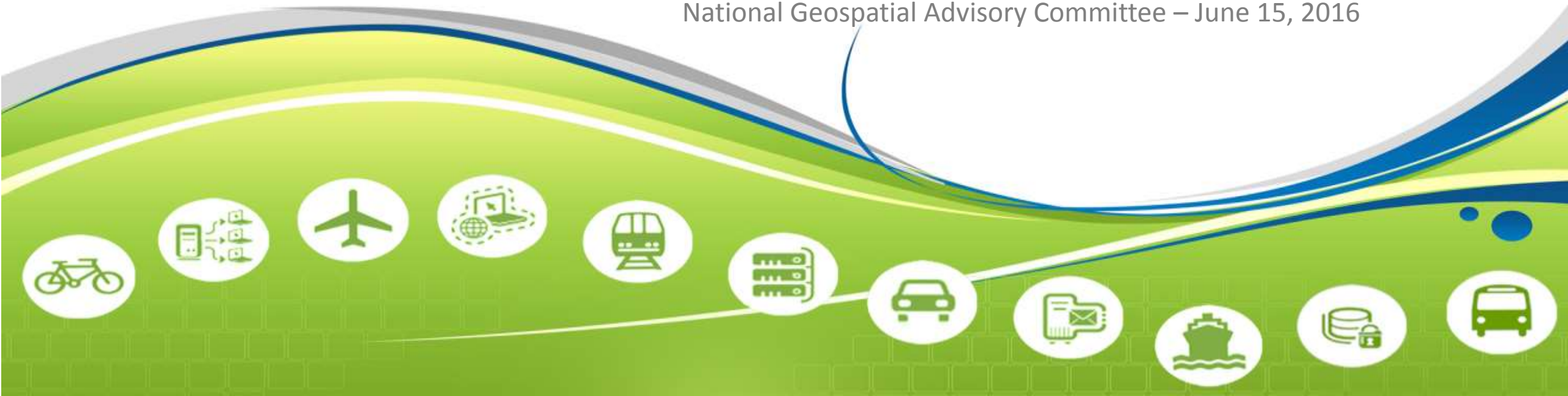
OCIO Office of the DOT Chief Information Officer

Update on National Address Database Pilot

Steve Lewis, U.S. Department of Transportation

National Geospatial Advisory Committee – June 15, 2016



Minimum Content Guideline – 3 Components

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.)

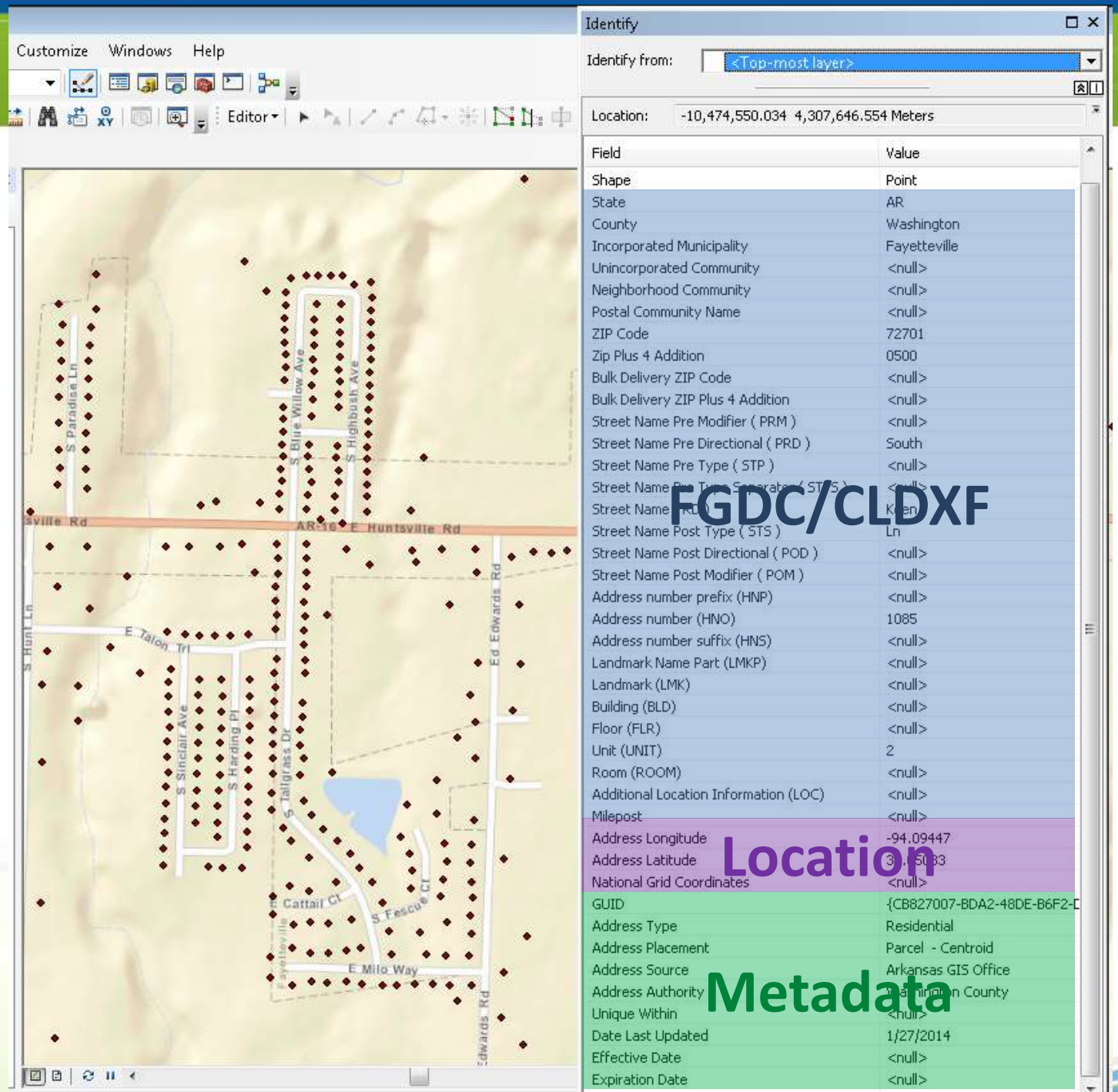
Review of the Minimum Content Guideline

- **Round 1:** NSGIC/Census project steering committee
- **Round 2:** All Summit attendees
  - Received 11 sets of written comments
- Guideline was revised/refined in response to each round of comments

Feedback on the Minimum Content Guideline

- Overall, feedback has been *mostly* positive:
  - "The '*low barrier to participation*' is likely an excellent idea to encourage greater data coverage."
  - "Simplicity in parsing roll up tools is critical."
  - "… CLDXF maintains the applicable components of the FGDC and PIDF-LO standards while addressing the needs of NG9-1-1…"
  - "… the summary captures what is critical yet allows flexibility so that data can be updated and upgraded iteratively."
- From a position paper released by NSGIC in April:
  - "As a point of emphasis, we strongly concur with the direction and recommendations made in the *National Address Database Draft Minimum Content Standard* (v8, March 2016) document under development by the USDOT."

NAD Schema



OCIO Office of the DOT Chief Information Officer

Pilot Participants Compiled Into NAD Schema

“Have Not” Status

- Goal was to find agencies (likely counties or tribes) that haven’t yet created their addresses
- Wanted entity that was **interested, motivated**, and **willing to work with us**.
- We did not want to create addresses that will then sit on a shelf.

✓Jackson County, AR

AGIO was a helpful partner, they want to finish statewide addresses by plugging few remaining holes

Jackson County, AR - Data Sources

- Countywide **E911 Address List**
  - 18k records
  - Some missing zip/city info
  - Some basic data scrubbing needed
- **Countywide centerlines** existed
  - No data scrubbing needed!
- Countywide **parcels**
  - 79% had some address info
  - Data standardization was needed
    - E.g., for city name, address field, etc.

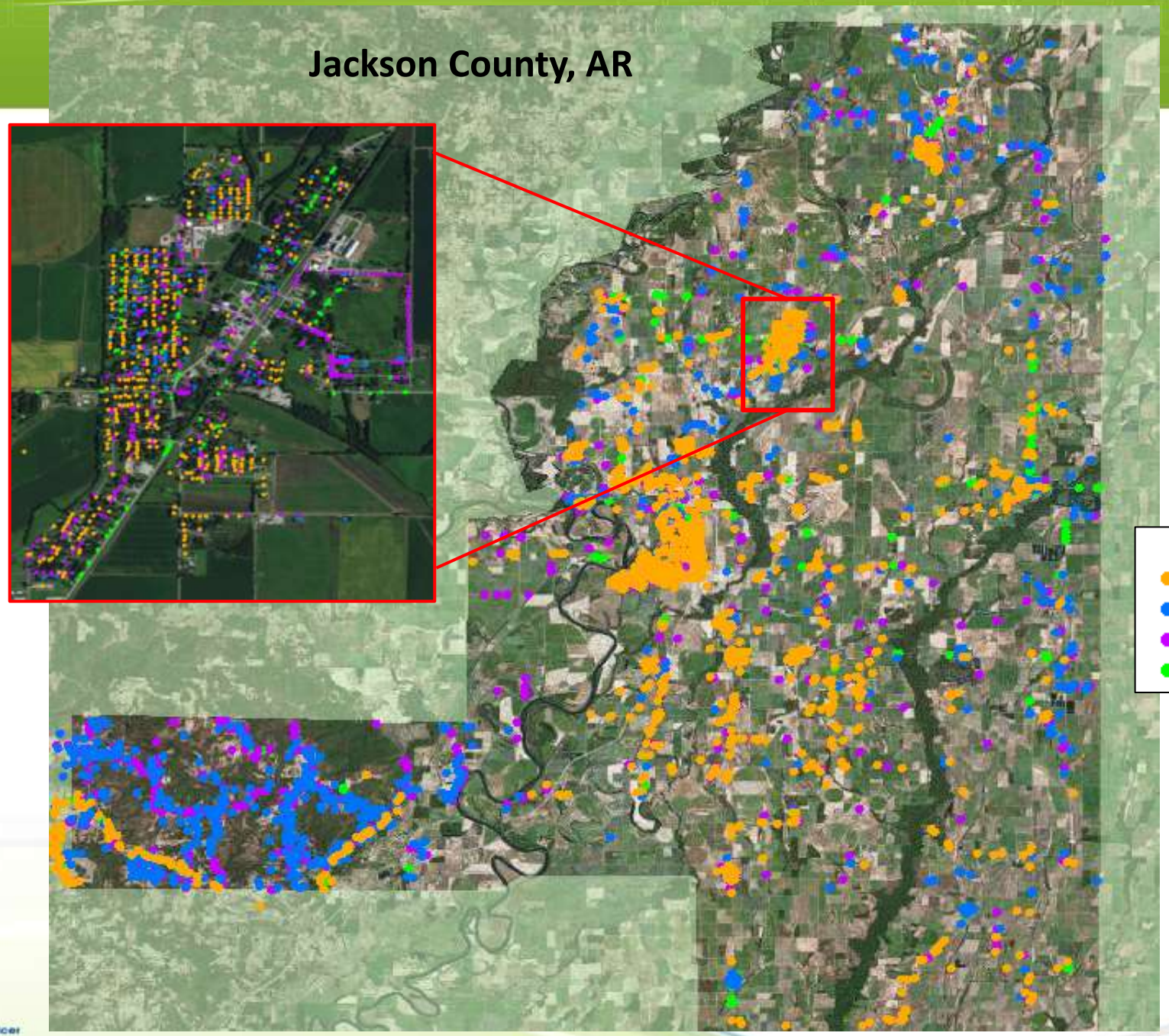
Jackson County, AR Geocoding Approach

- 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

Source	Total Records Matched	% Matched*
Melissa Data	7,073	38%
Parcel Centroids	1,700	9%
County Centerline	4,112	23%
Census/Tiger Centerlines	1,347	7%
Totals:	**14,232**	**77%**

Jackson County, AR



Preliminary Pilot Findings

- **Tribal participation** is going to be a **challenge**
  - Lots of outreach, lots of interest, but no contributed data
  - Gila River data is part of AZ statewide collection
- Data sharing agreements to make data **publically available could be a challenge**
  - AZ has yet to provide clearance for public release
- **Aggregating existing statewide/have collections was straight forward**
  - Five additional states have volunteered to ETL their own data for inclusion in the pilot NAD database
- **The schema will likely evolve**, but needs to remain consistent with leading address schemas to allow for streamlined ETL

Pivoting from Pilot to Development

Goal: Compile address data from 30 states into version 1 by December 2016

Philosophy: Follow the Digital Services Playbook

Digital Services Plays

#3 Keep it simple!

- Broadest participation possible
- Lowest barrier to entry

#4 Methodology

- Agile approach – quick responses to change, continuous development & customer engagement

#13 Default to Open

- Fork code, reuse parsing from GitHub

Architectural Considerations

#8 Choose a modern technology stack

#9 Deploy in a flexible hosting environment

- Cloud First
- First priority: DC, NJ, OH, UT, and VA
  - Push vs. Pull
  - Extract, Transform, Load (ETL)
  - Feedback mechanism
  - Preferred Model: Local to State

What's Next

Pilot Phase

- Make available data with AZ, AR & Boone County, MO
- Finalize report Q3 FY16

Development

- Choose platform
- Initiate work with states that are prepared to develop ETLs (no cost)
- Identify funding to launch
- Launch Data Challenge for "have nots"

Data Challenge

- Goal: develop an app to gather crowd sourced address information
  - Must collect the items identified in the minimum content guideline
- App can be used by
  - Local police and firemen
  - Real estate agents
  - Boy Scouts and Girl Scouts
  - FEMA Corps
  - Public
- Resulting address information would be used as “seed” data for local governments with no data and for QC/QA of existing data

Contact Info

Steve Lewis

Chief Geospatial Information Officer

U.S. Department of Transportation

(202) 366-9223

steve.lewis@dot.gov