



# Data Discovery and the NSDI Framework Data Survey

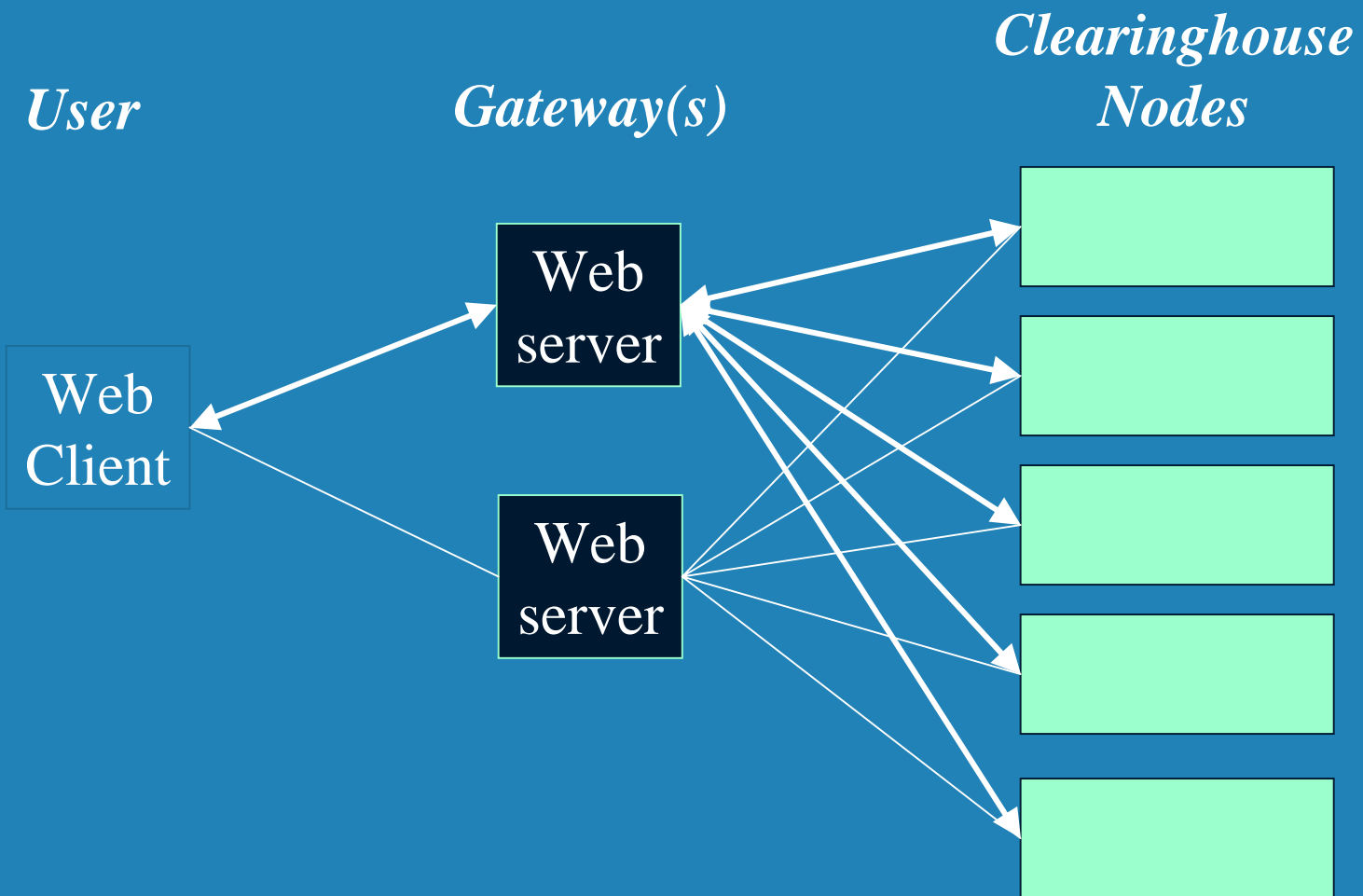
*presented by:*

Tim Haithcoat

**University of Missouri  
Columbia**

*From materials of participants at the  
1998 NSDI Partnership Kick-Off Meeting  
Charleston, SC*

# General Clearinghouse Architecture

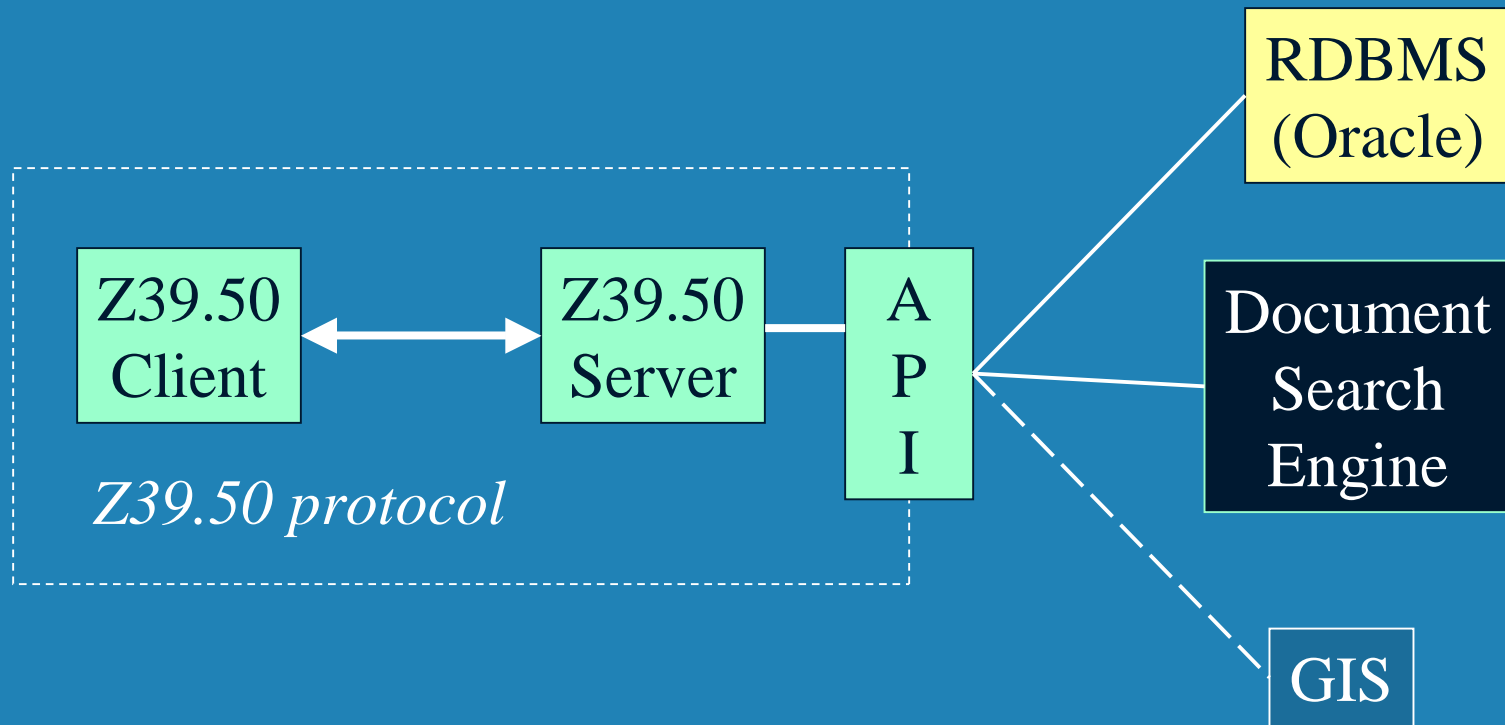


# Who defines a node?



- \* An NSDI Clearinghouse Node includes a public Z39.50 protocol server supporting the FGDC metadata (Geo profile)
- \* A Node can include a Web server to provide forms and organizational reference
- \* Servers may be established institutionally or geographically to optimize sharing of common computer resources

# Z39.50 in practice



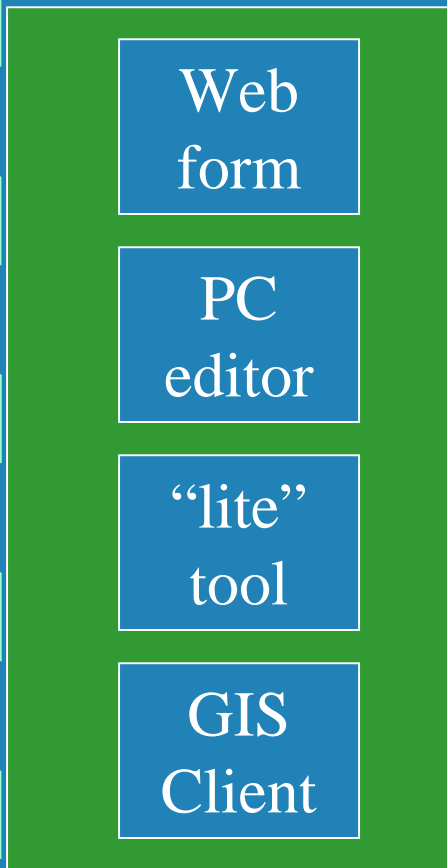
# Why Use Clearinghouse?



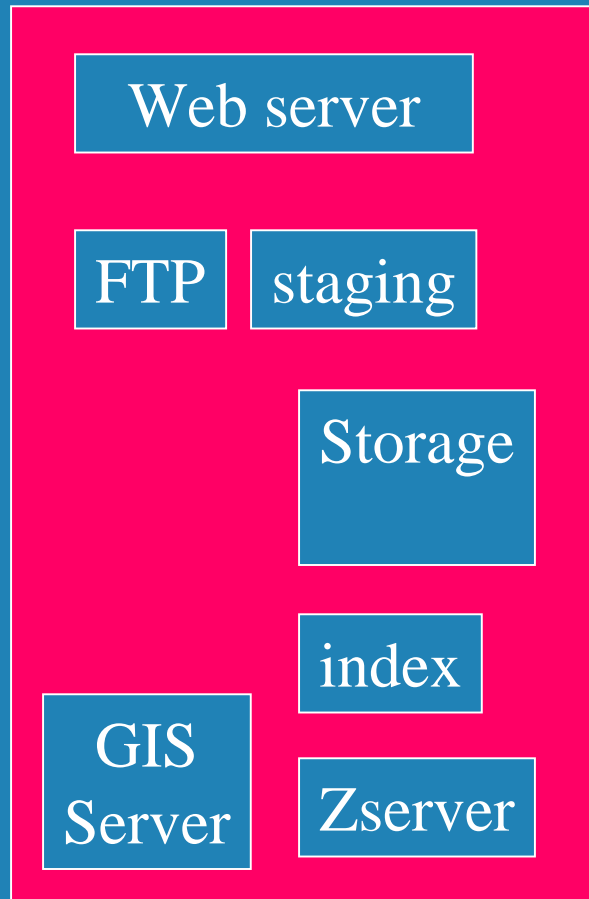
- ✧ Minimize duplication of effort in spatial data collection and processing
- ✧ Provide means to discover data collection requirements, inventory, and quality
- ✧ Support documentation of basic spatial data sets for advised re-use within internal and external applications

# Clearinghouse Architecture

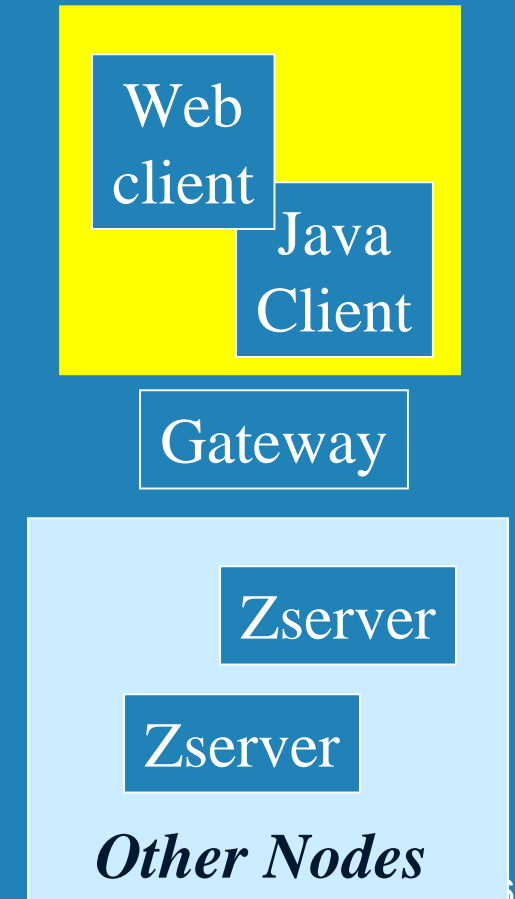
## *Metadata Production*



## *Metadata Publication*



## *User Query*



# Why FGDC Compliant Metadata?


- ✧ To help organize internal investment in spatial data
- ✧ To provide inventory and quality information to catalogs and brokers
- ✧ To provide information for processing and interpretation of data received in transfer (*to accompany a data transfer*)



# FGDC Compliant Metadata

- \* Defined by Executive Order 12906 in April 1994 as formal format for federal use
- \* To be applied to all new data sets, effective January 1995: all legacy data on a schedule
- \* To be used as vocabulary for search engines within the National Geospatial Data Clearinghouse





*Regardless of the approach taken,  
the search of individual  
Clearinghouse Nodes  
looks and feels the same.*

This level of abstraction permits local selection of software solutions to appropriately meet local needs.

# Metadata Exchange

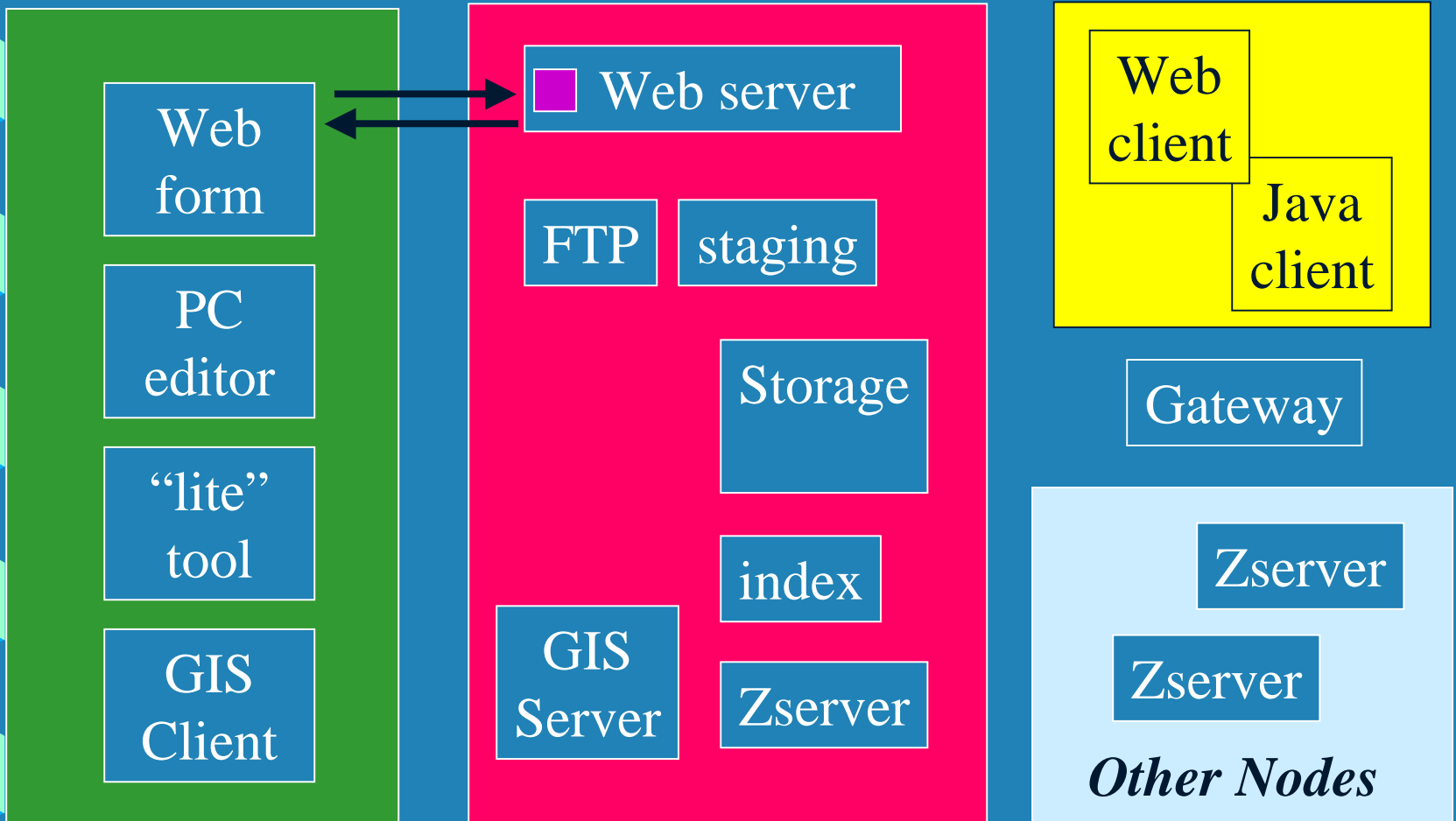
- \* It is desirable for metadata/data systems to be able to use, enhance and report metadata developed elsewhere.
- \* Most systems can **export** metadata in XML (preferred) or as structured text
- \* To use and modify metadata as data change, the ability to **import** it is desirable

# Web-based “Donation” of Metadata

*Metadata  
Production*

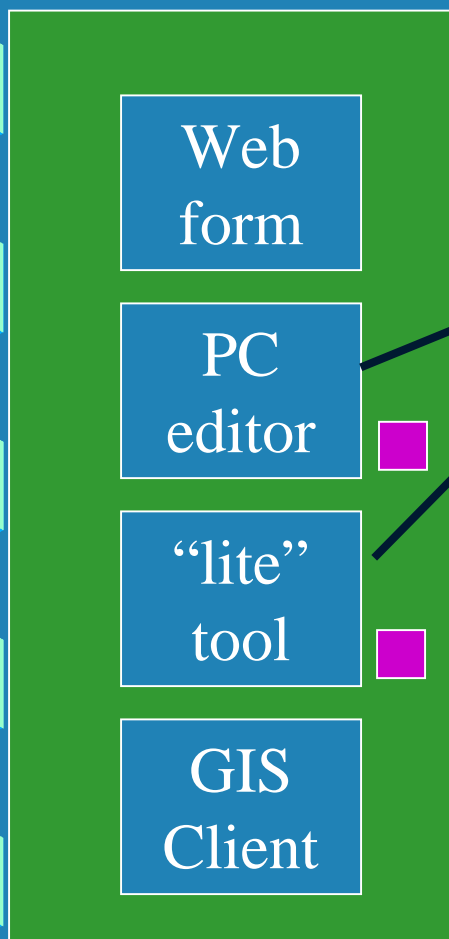
*Metadata  
Publication*

*User  
Query*

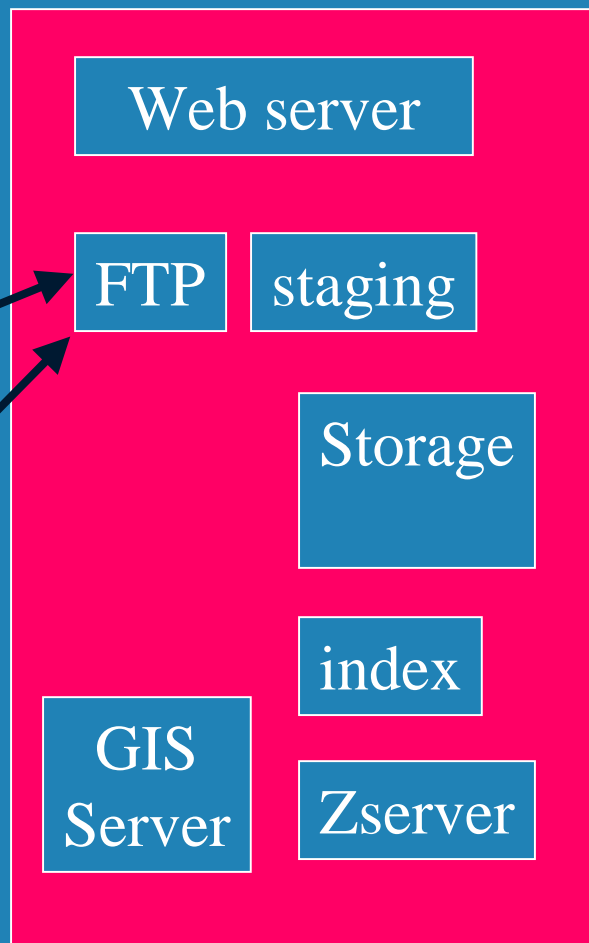


# Publication of PC-Stored Metadata

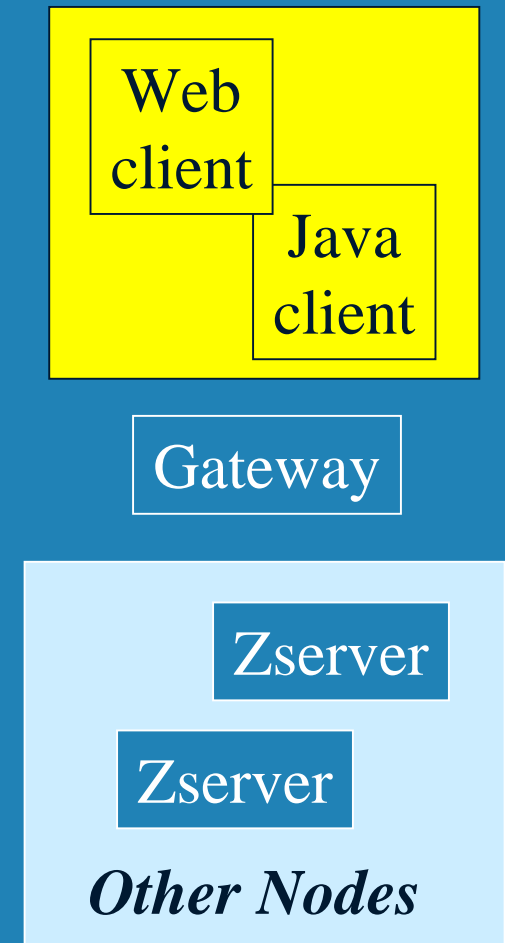
## *Metadata Production*



## *Metadata Publication*

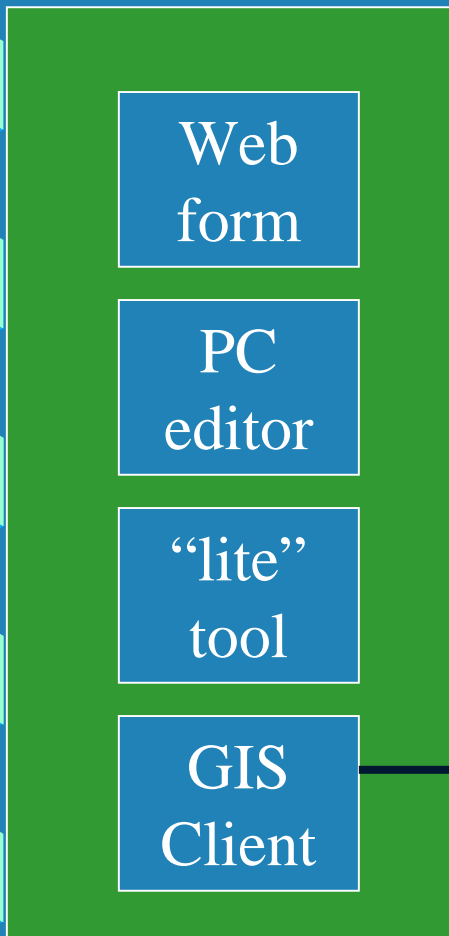


## *User Query*

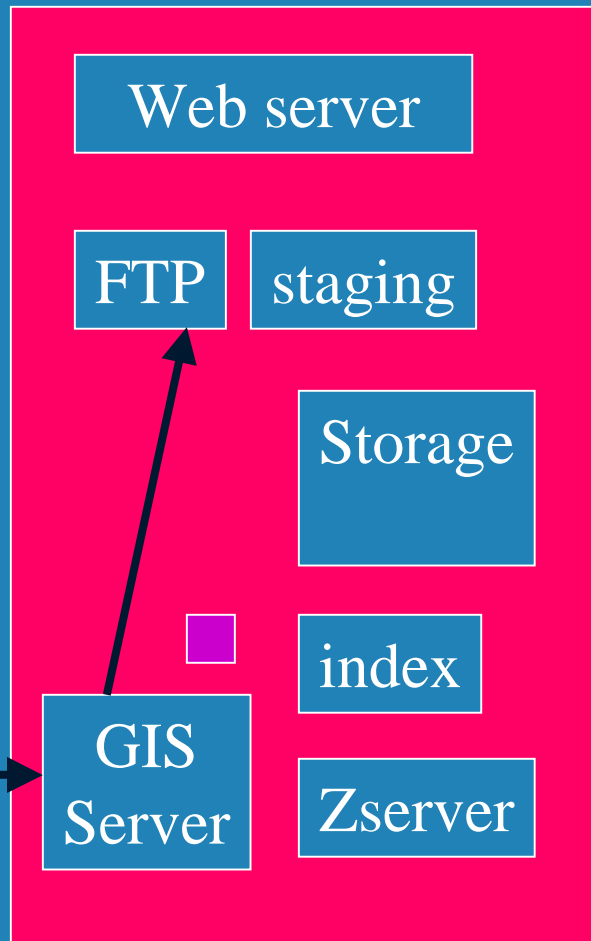


# Server-side GIS metadata

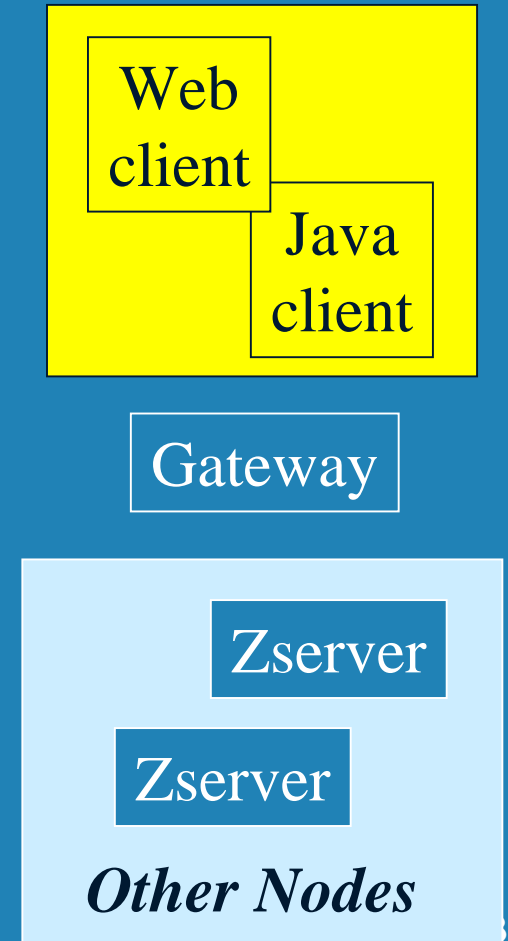
## *Metadata Production*



## *Metadata Publication*

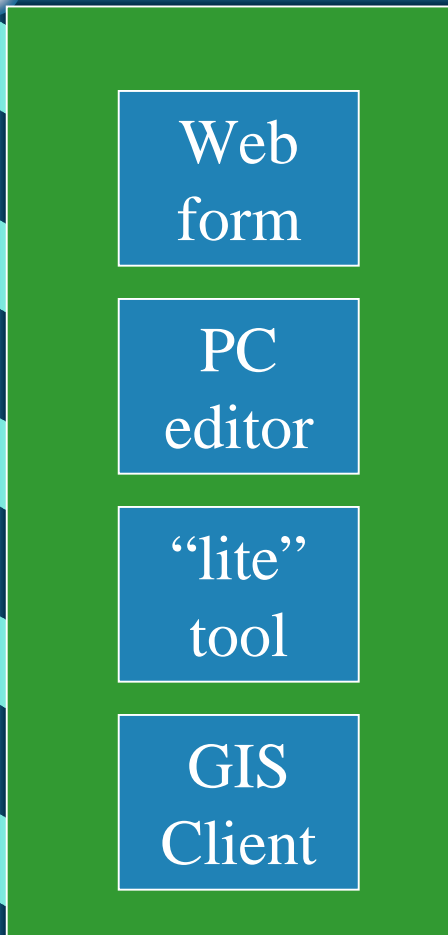


## *User Query*

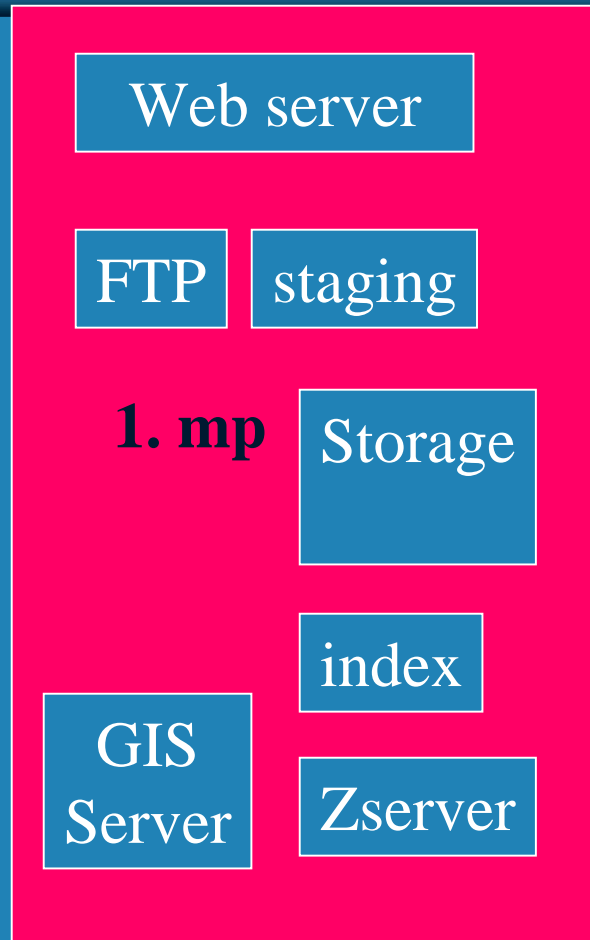


# Validation of Staged Metadata

## *Metadata Production*



## *Metadata Publication*

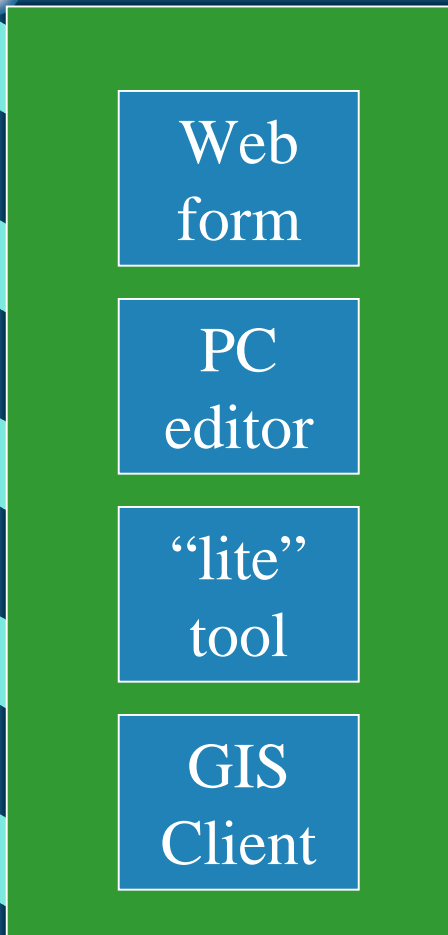


## *User Query*

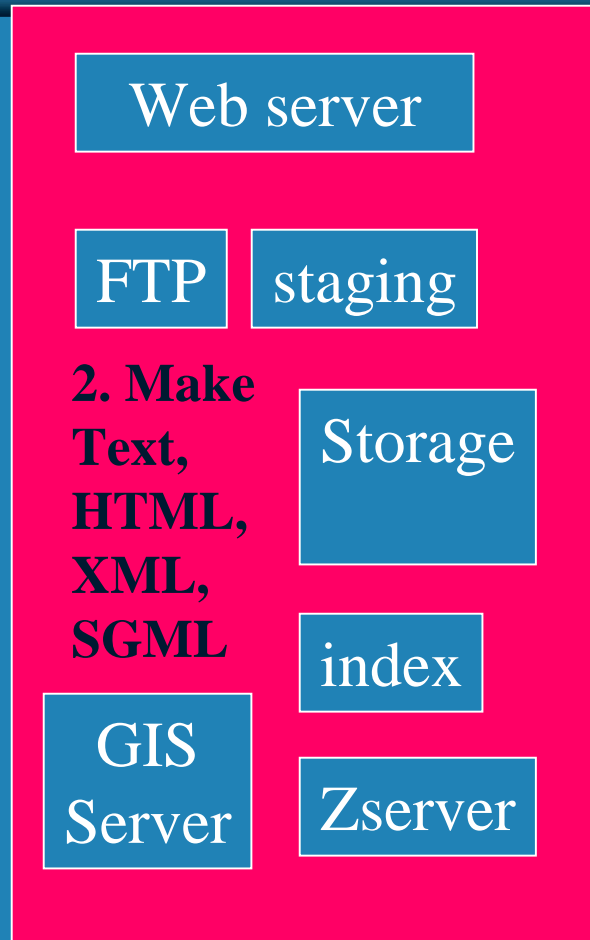


# Create SGML and HTML Formats

## *Metadata Production*



## *Metadata Publication*

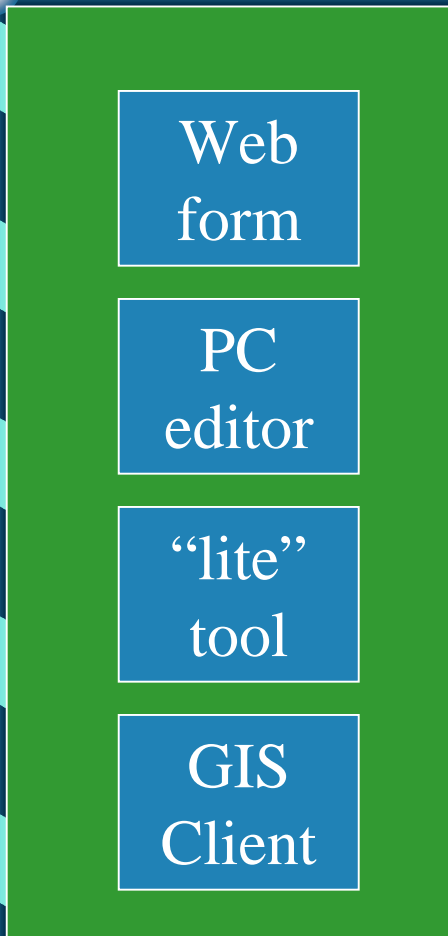


## *User Query*

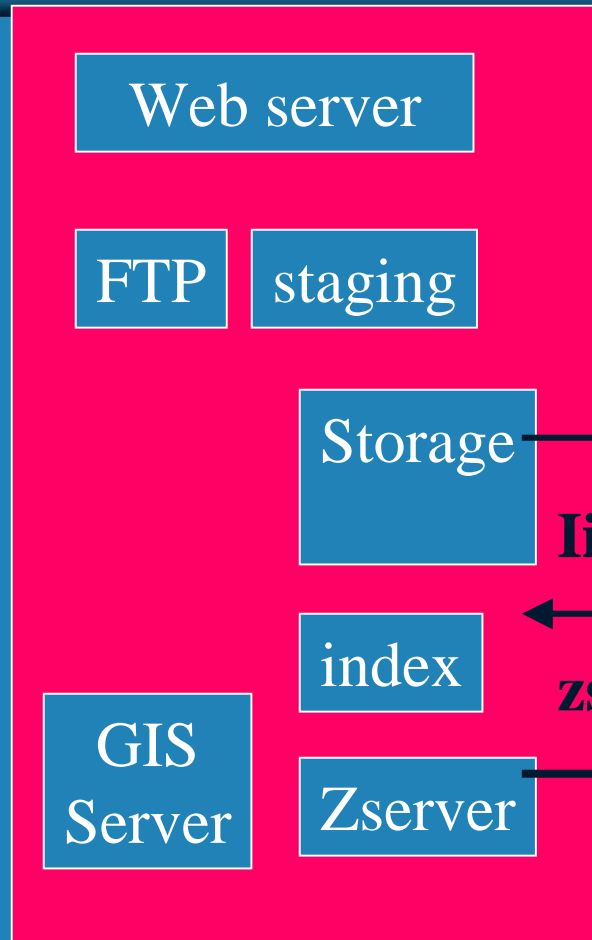


# Activate Z39.50 Server Process

## *Metadata Production*



## *Metadata Publication*



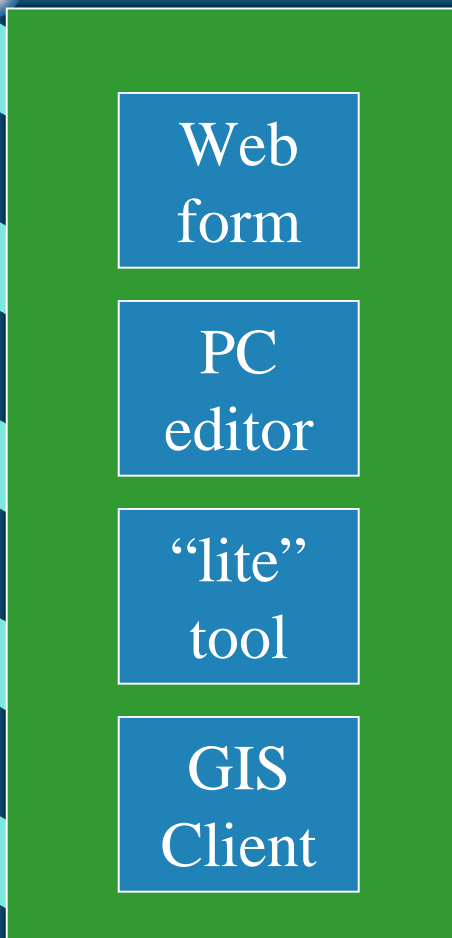
## *User Query*



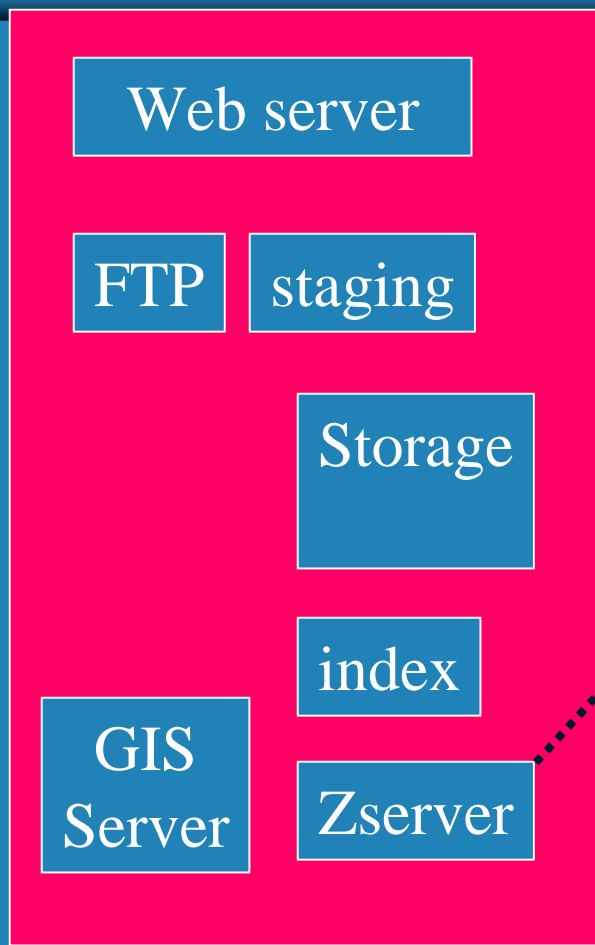


# Register Server With FGDC Registry

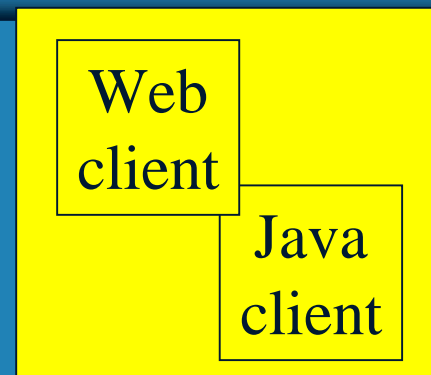
## *Metadata Production*



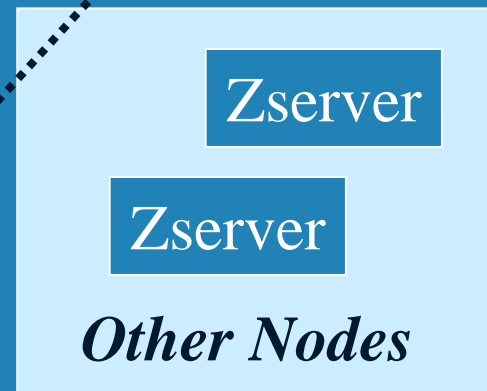
## *Metadata Publication*



## *User Query*

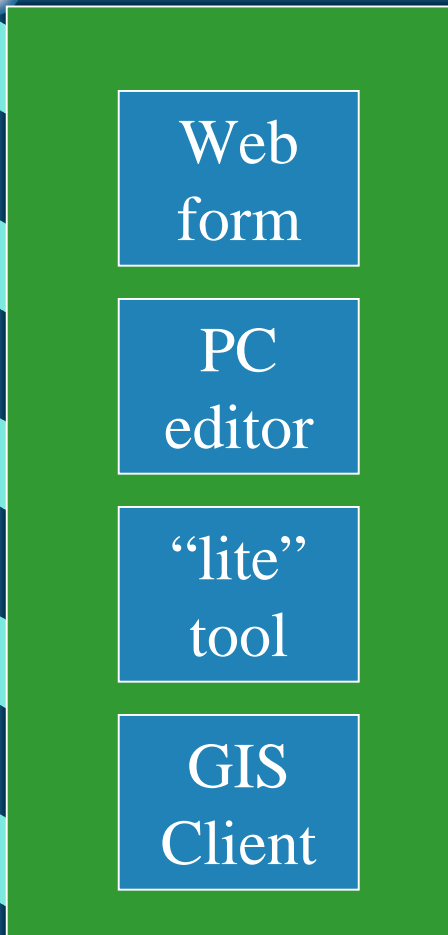


Gateway

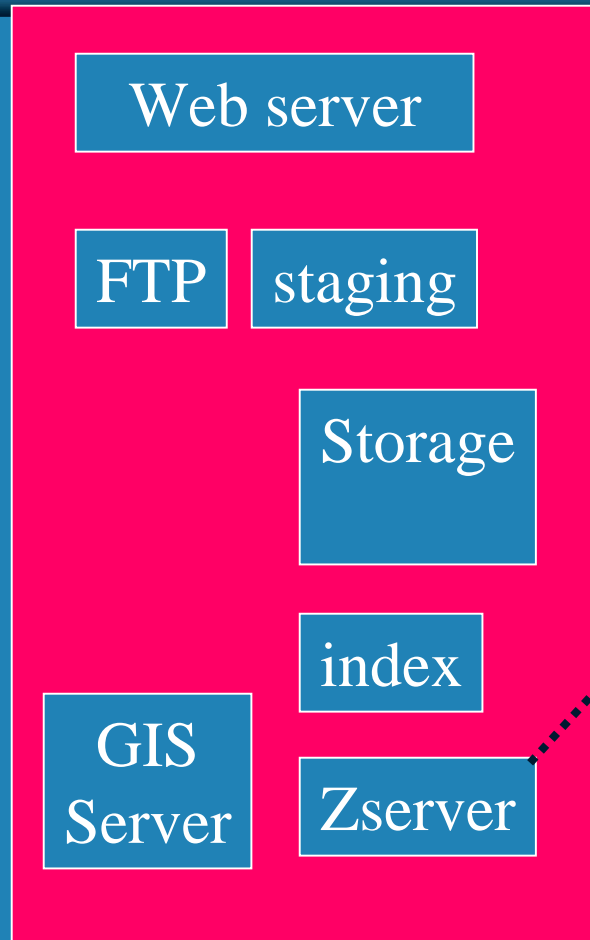


# User Searches Clearinghouse Gateway

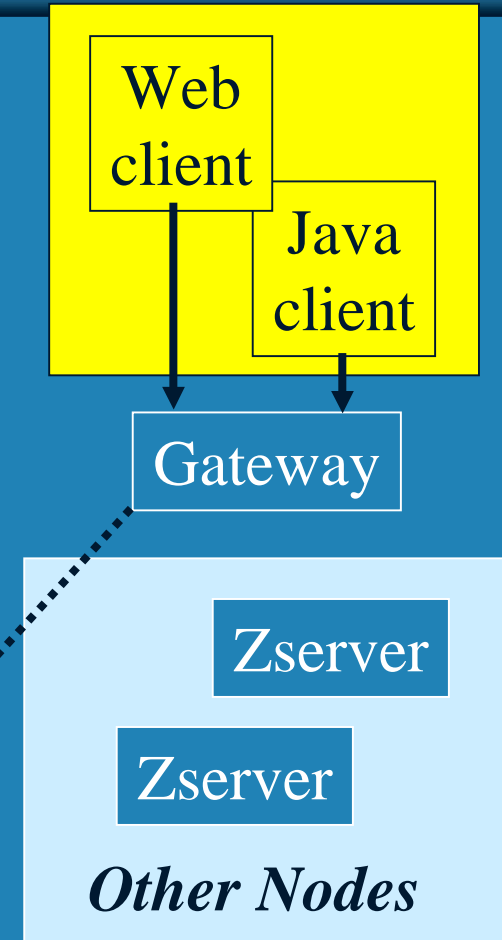
## *Metadata Production*



## *Metadata Publication*



## *User Query*

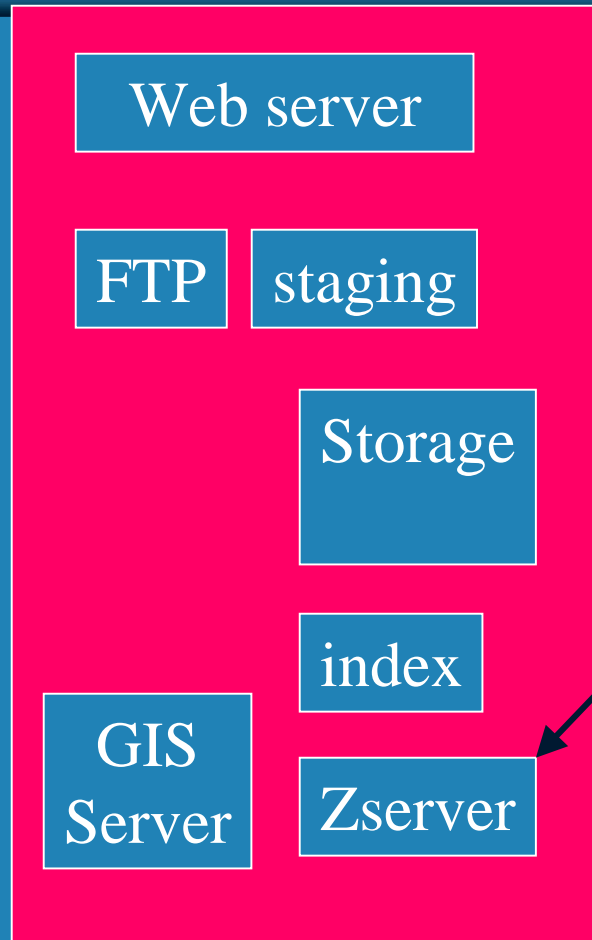


# Search is Sent to One or More Servers

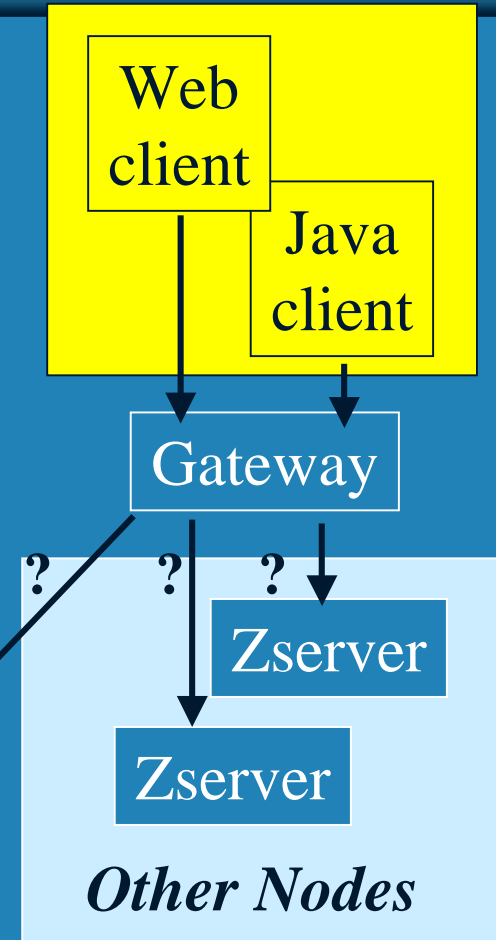
## *Metadata Production*



## *Metadata Publication*

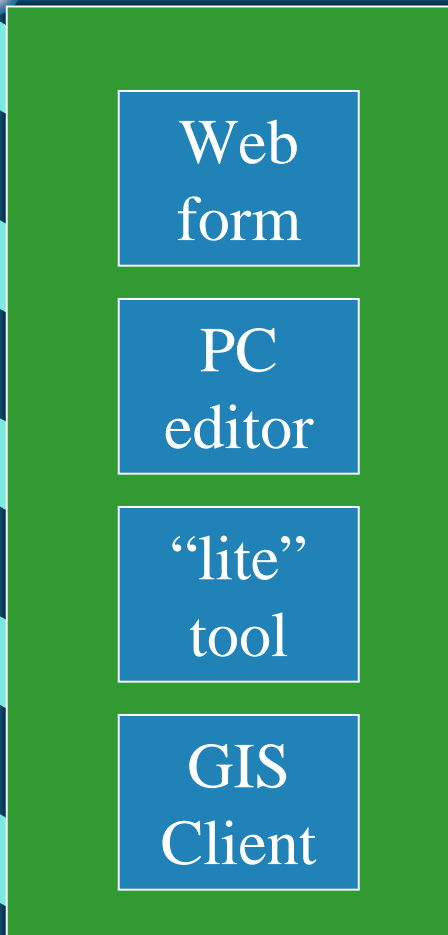


## *User Query*

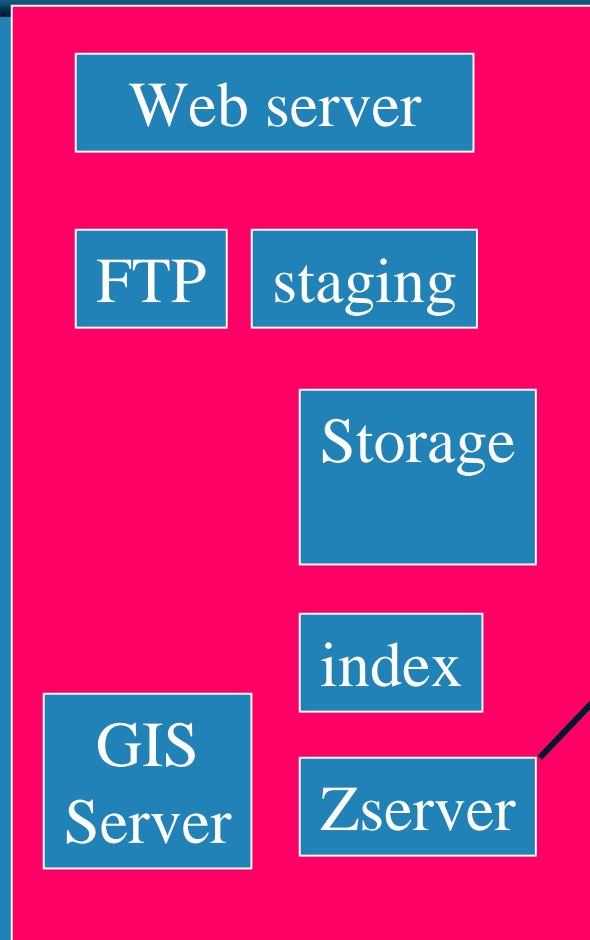


# Results are Returned from Servers

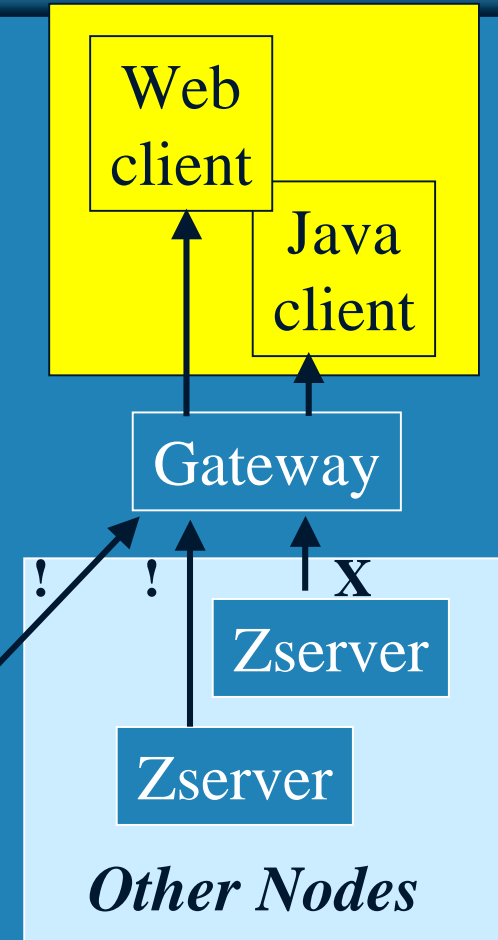
## *Metadata Production*



## *Metadata Publication*

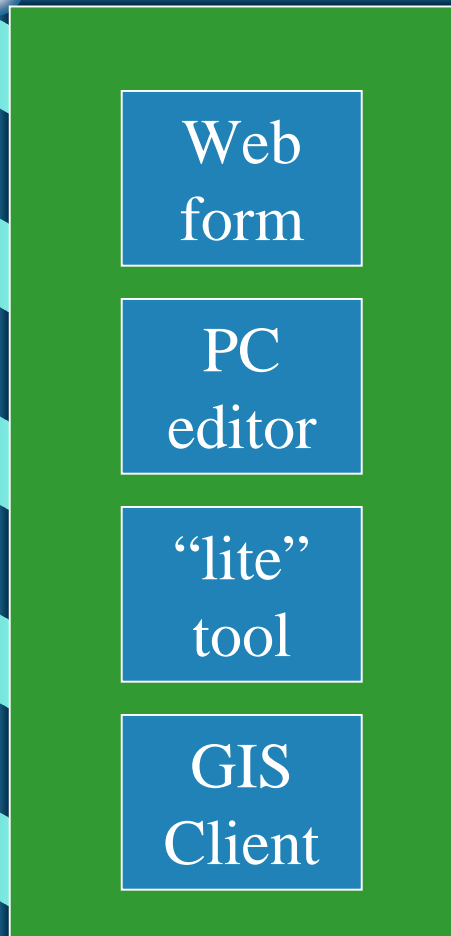


## *User Query*

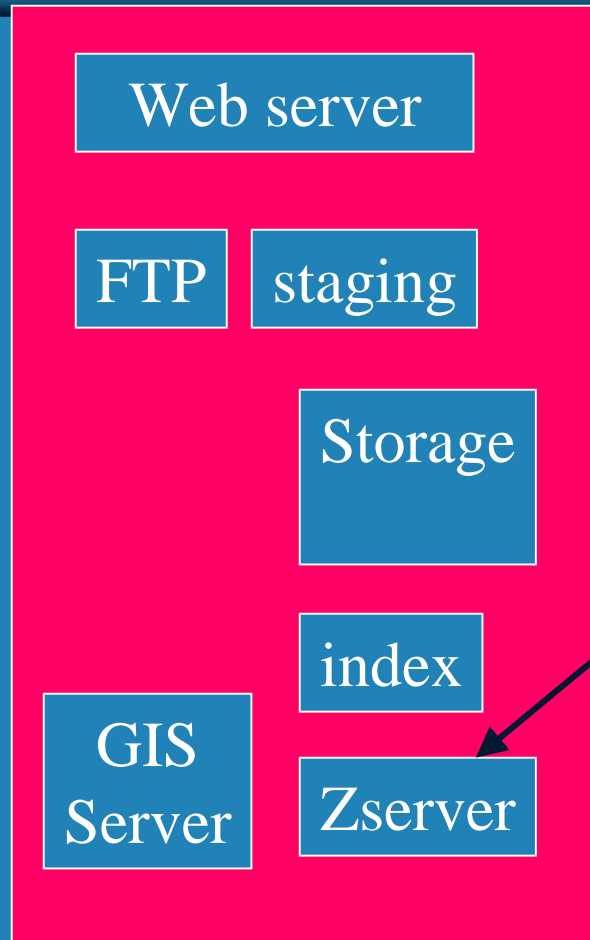


# Document is Requested from Server

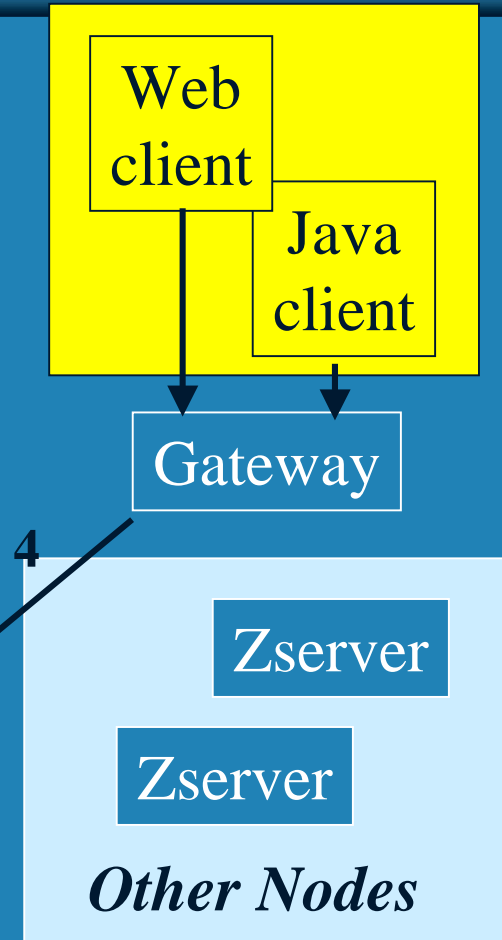
## *Metadata Production*



## *Metadata Publication*

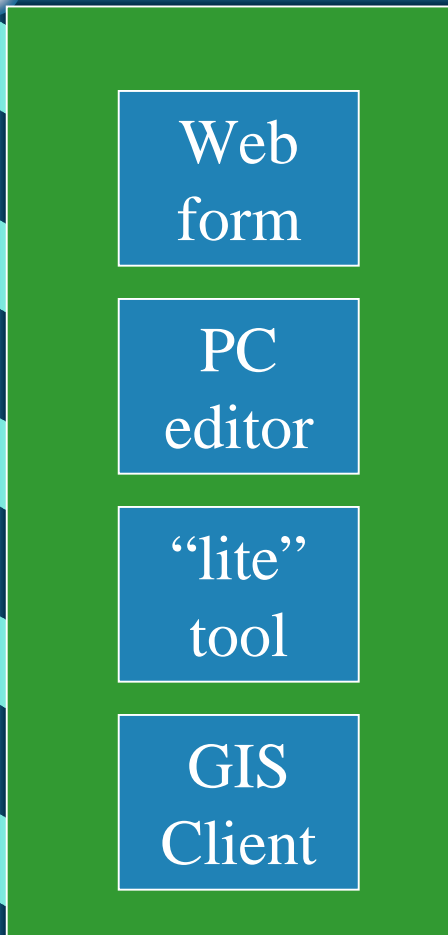


## *User Query*

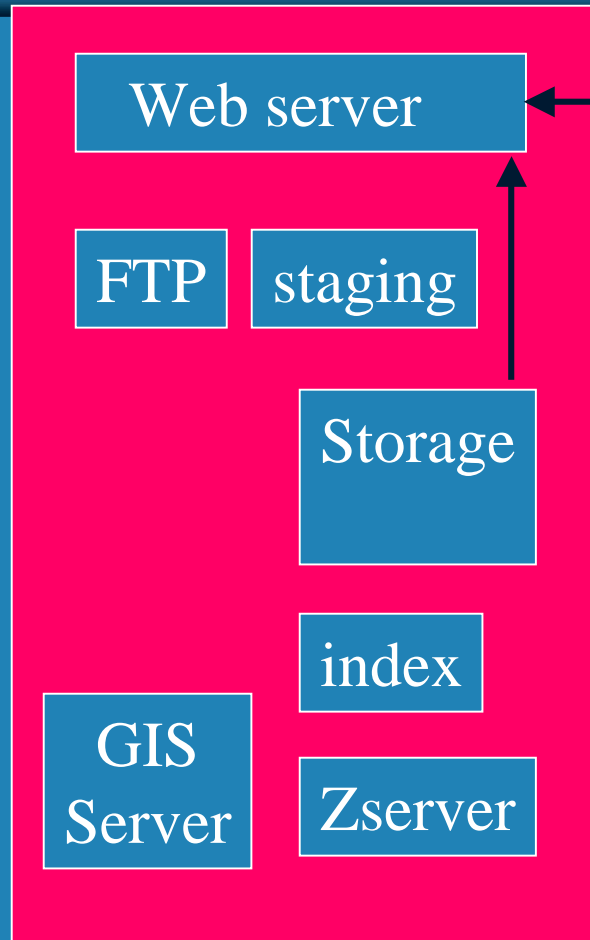


# Web Server can also Provide a Browse List

## *Metadata Production*



## *Metadata Publication*



## *User Query*





# User Interfaces for Clearinghouse

- ✧ Search through a Gateway is done via downloaded search forms
  - HTML-based query forms
  - Java-based query forms

# Query Components



- \* Spatial Search
- \* Temporal Search
- \* Field or Full-text Search
- \* Server Selection

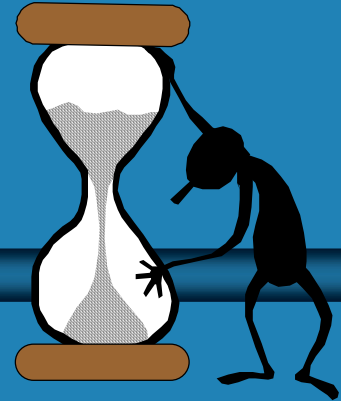




# Spatial Search

- ✧ Selection from a list of place names for which coordinates have been derived
- ✧ Entering coordinates to search directly
- ✧ Entering coordinates from a map

# Temporal Search



Search for data content timeframe  
Single date (before, equal, after)  
Range of dates (within range)

*Time period search uses YYYYMMDD  
to characterize the content or reference  
date*



# Field and Full-text Search

Spatial metadata may be searched based on contents of individual fields or full-text word search

Fields may be numeric, text or picklist

Full text supports wildcard search (e.g. geo\*)

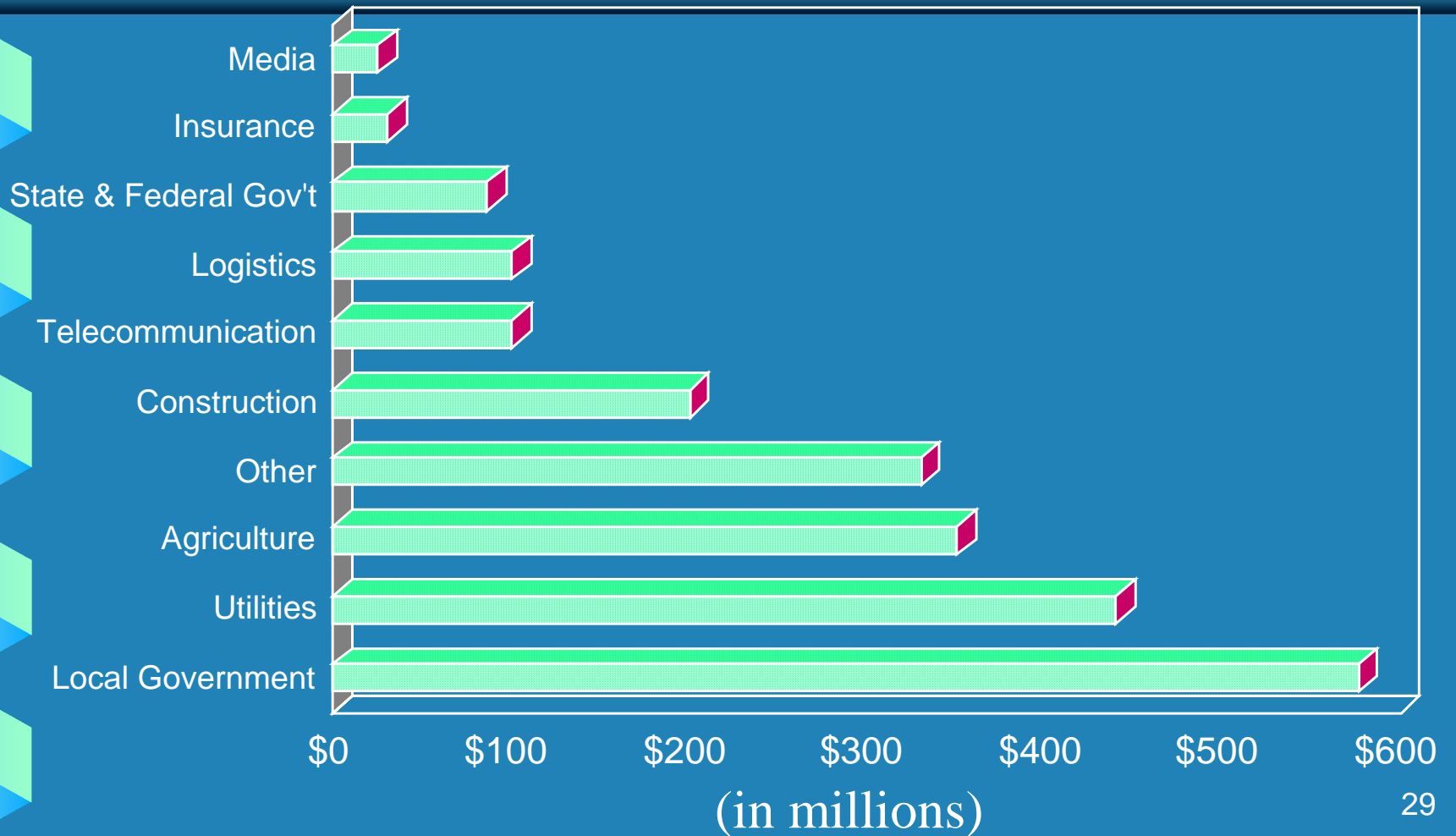
Fields may be ANDed or ORed together



# Server Selection

- \* User may scroll the list of servers and select one or many (currently over 90)
- \* User may select to search ALL servers
- \* Near-future server selection can be assisted by 'footprint of service' to limit search to most relevant servers

# 1998 North American GIS Market

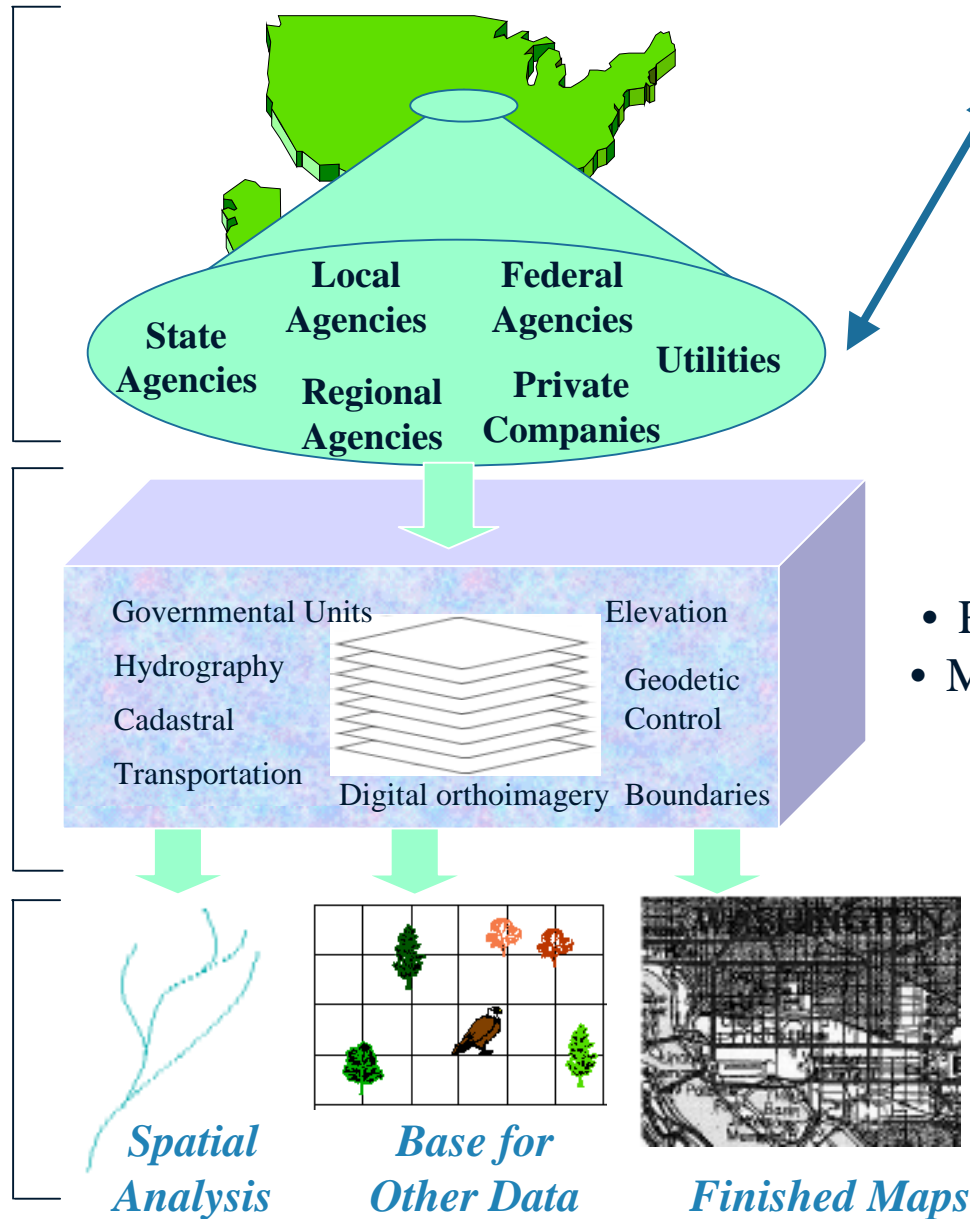


# Framework Concept

Within a geographic area, organizations that produce and use geospatial data

Collaboratively build and maintain commonly used “framework” data

To provide a common base for many uses



## *Organizational Functions*

- Data development, maintenance & integration
  - Data access
- Data management
  - Coordination
- Executive guidance
- Resource management
- Monitoring & response



# What is the framework?

- \* Data backbone for the NSDI
  - Data
  - Procedures
  - Technology
- \* A way to facilitate:
  - Production of geographic data
  - Use of geographic data
  - Reduction of operating costs
  - Improve service
  - Decision making

# Business Context

“to be widely used and useful”

- ✧ Avoid restrictive practices:
  - No exclusive distribution arrangements that inhibit timely and equitable access
  - Avoid restrictions on reuse or redistribution of data
  - Charges set to recover cost of service only.
- ✧ Availability in nonproprietary format(s).
- ✧ Conform to standards.





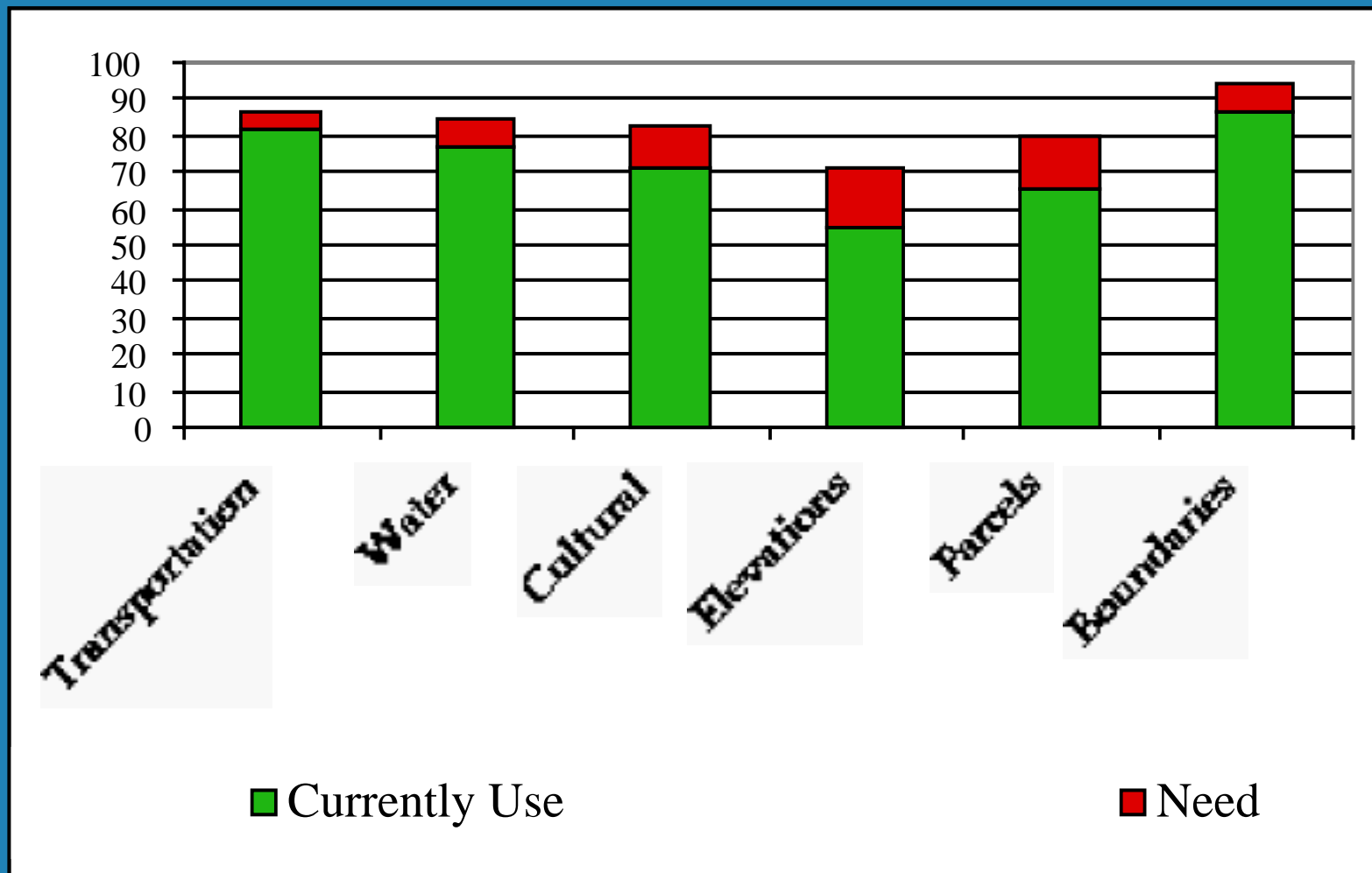
# What Layers of Information are Currently Included?

- \* Geodetic control
- \* Digital orthoimagery
- \* Elevation data
- \* Transportation
- \* Hydrography
- \* Governmental units
- \* Cadastral (*reference system and public parcels*)

# Commonly Used Data Themes

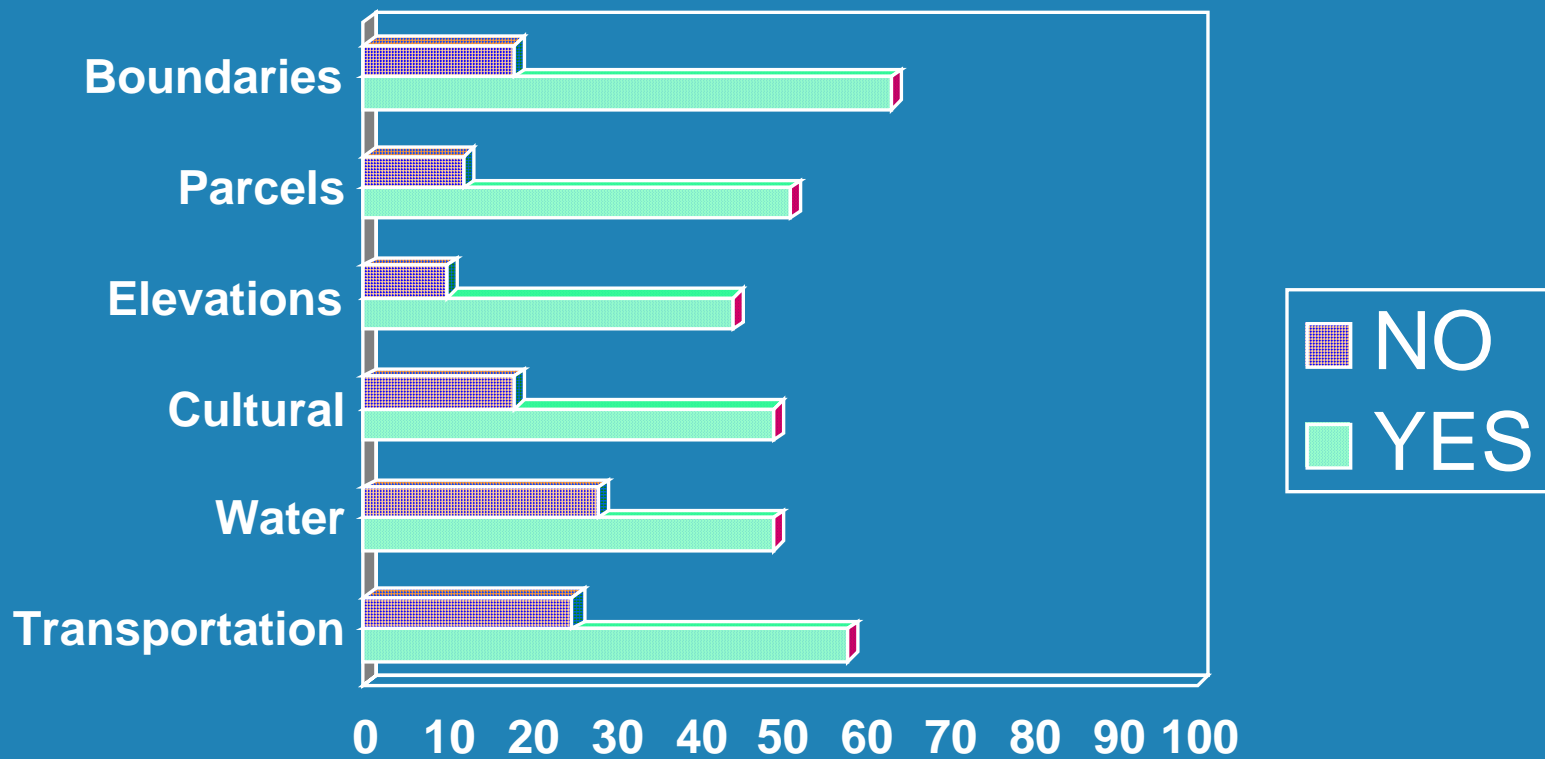
Do you use these themes of digital data?

(If no, do you foresee that you might need them?)



# Commonly Used Data Themes

If you use digital data, do you have a substantial need for better data?





# Efforts to Extend Framework

- \* NSGIC Framework Data Survey has been completed. The results are posted on the FGDC website in an ArcView apr format as well as other database formats.

[home](#)

[what's new?](#)

[search](#)

[NSDI](#)

[metadata](#)

[clearinghouse](#)

[standards](#)

[framework](#)

[stakeholders](#)

[FGDC](#)

[funding](#)

[publications](#)

[data](#)

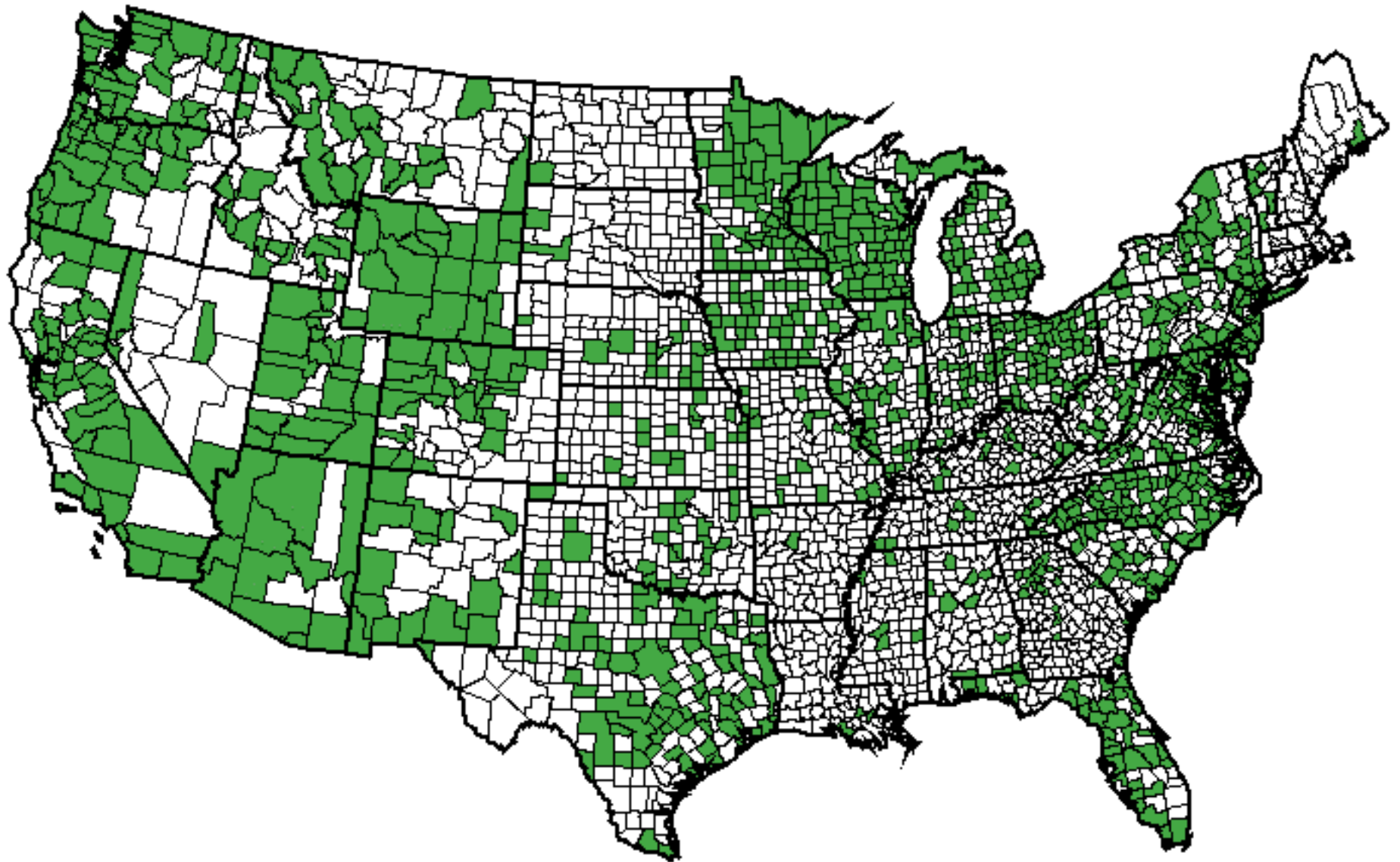
## Framework Survey Download Page

It is highly recommended that you download the complete package that includes everything. However, you can download components of the survey, if desired.

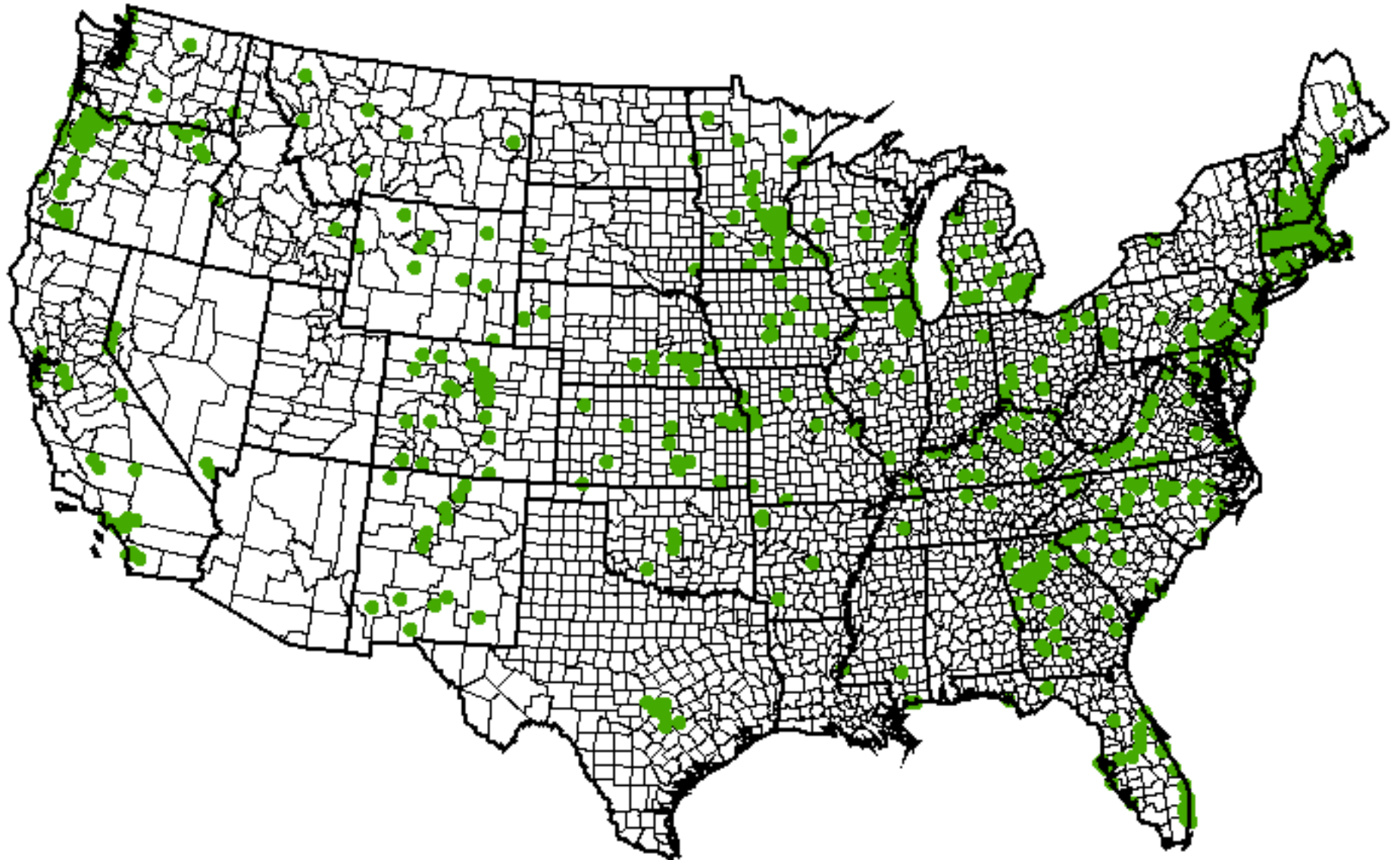
- **Download the Complete Survey Package** [survey\\_pre.zip](#) (size 9 MB) that includes:
  - survey12 directory (49 MB uncompressed)
  - docs/survey.doc (Microsoft Word version of the survey questions and coding options)
  - docs/readme.doc (help file for survey data)
  - docs/survey\_diskette (original diskette version of the survey)
  - all results files (dbf files and shape files).
  - ArcView 3.1 Project File with custom interface to analyze the survey results)

- 
- Download only the help file for the survey results in Microsoft Word format: [readme.doc](#)(size 216 KB) [readme.txt](#).
  - Download only the survey questions in Microsoft Word format: [survey.doc](#) (size 253 KB) [survey.txt](#).
  - Download only the digital survey questionnaire: [survey\\_disk.zip](#) (size 349 KB)
  - Download only the raw survey results (pre\_raw.dbf) in zip format: [pre\\_raw.zip](#) (size 766 KB)
  - Download only the auxiliary results from write-in answers (pre\_other.dbf) in zip format: [pre\\_other.zip](#) (size 31 KB)
  - Download only the auxiliary computed survey (pre\_sum.dbf) in zip format: [pre\\_sum.zip](#) (size 18 KB)

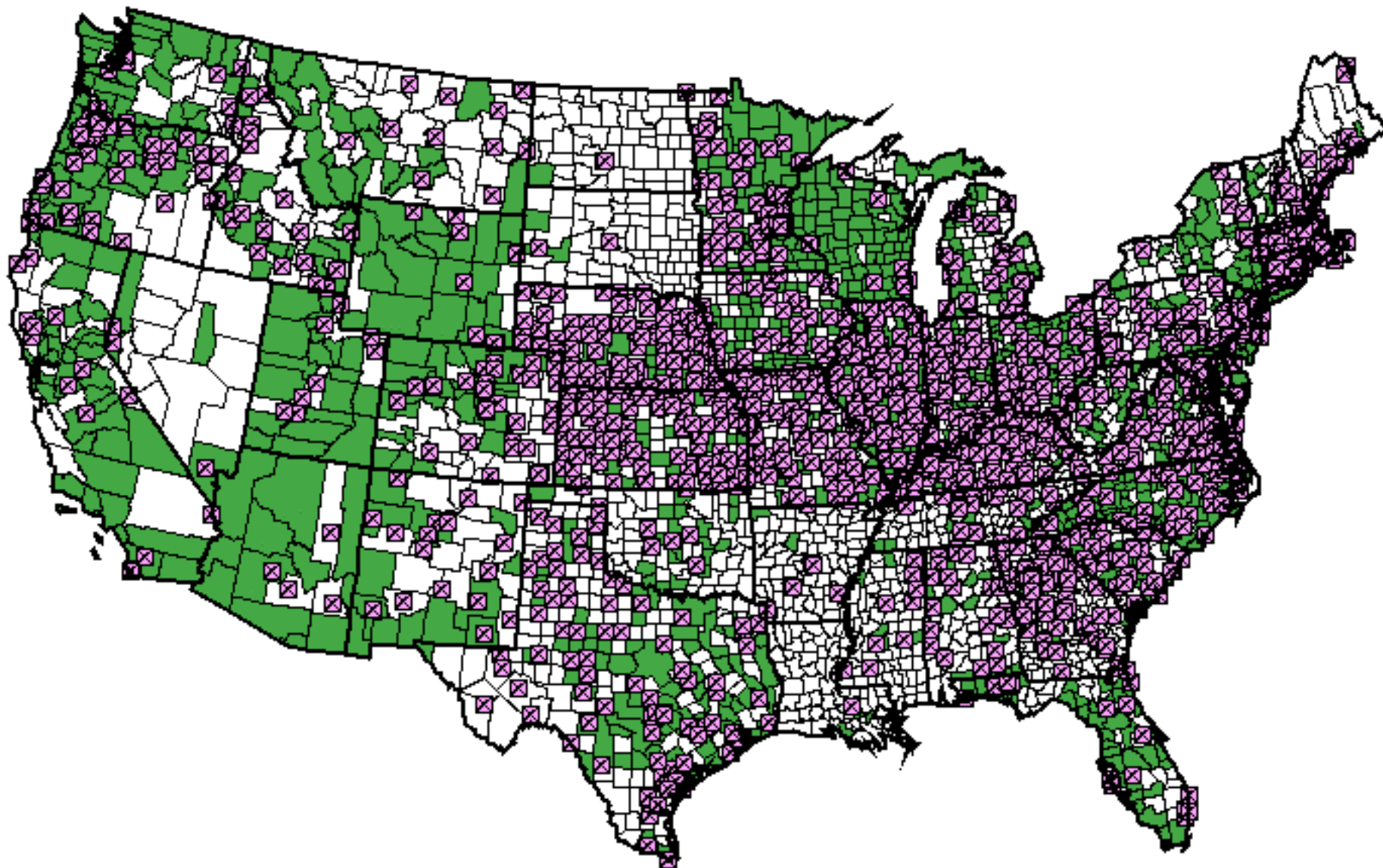
# Counties with Digital Data



# Cities with Digital Data

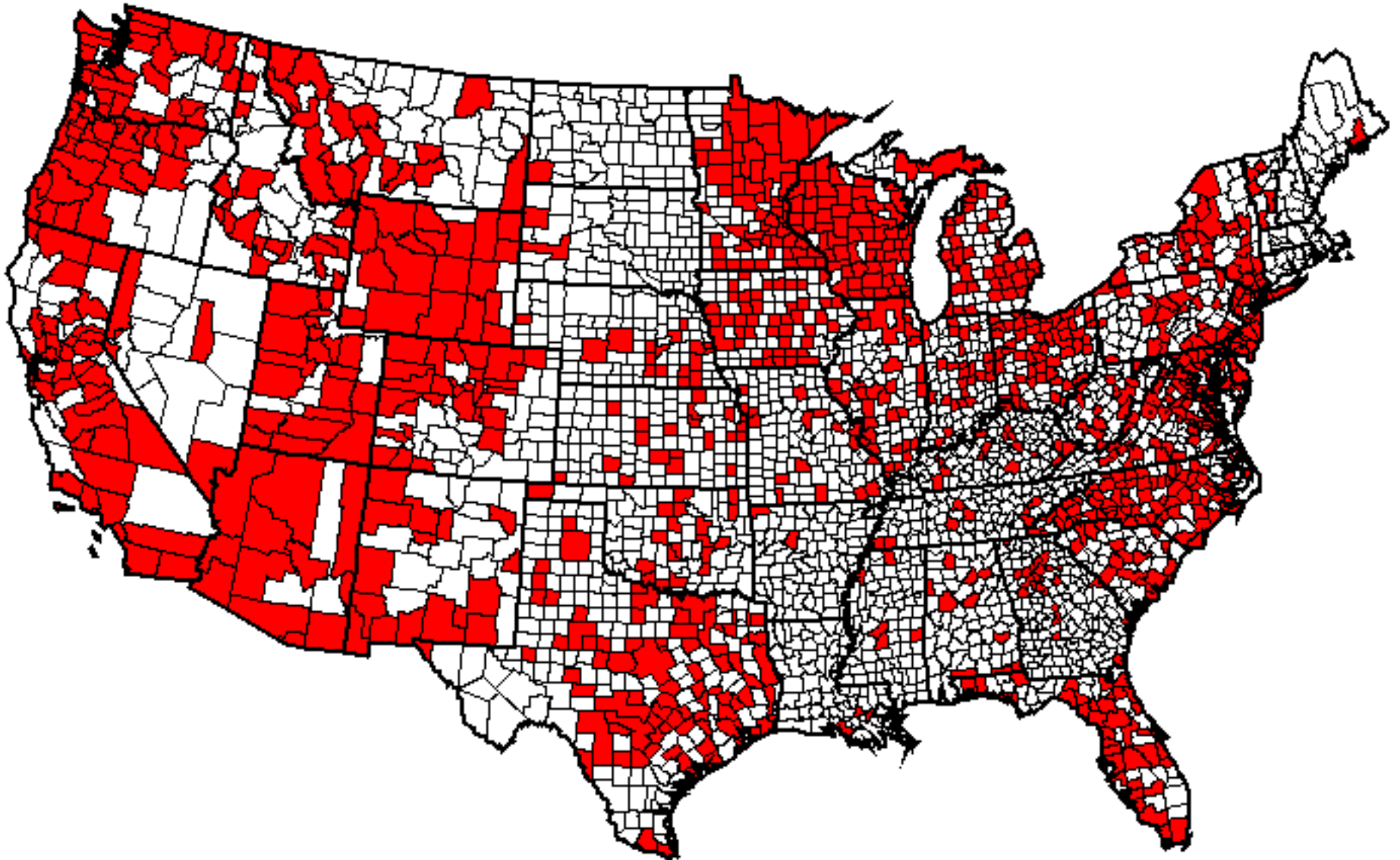


# Inventoried Sites w/wo Data

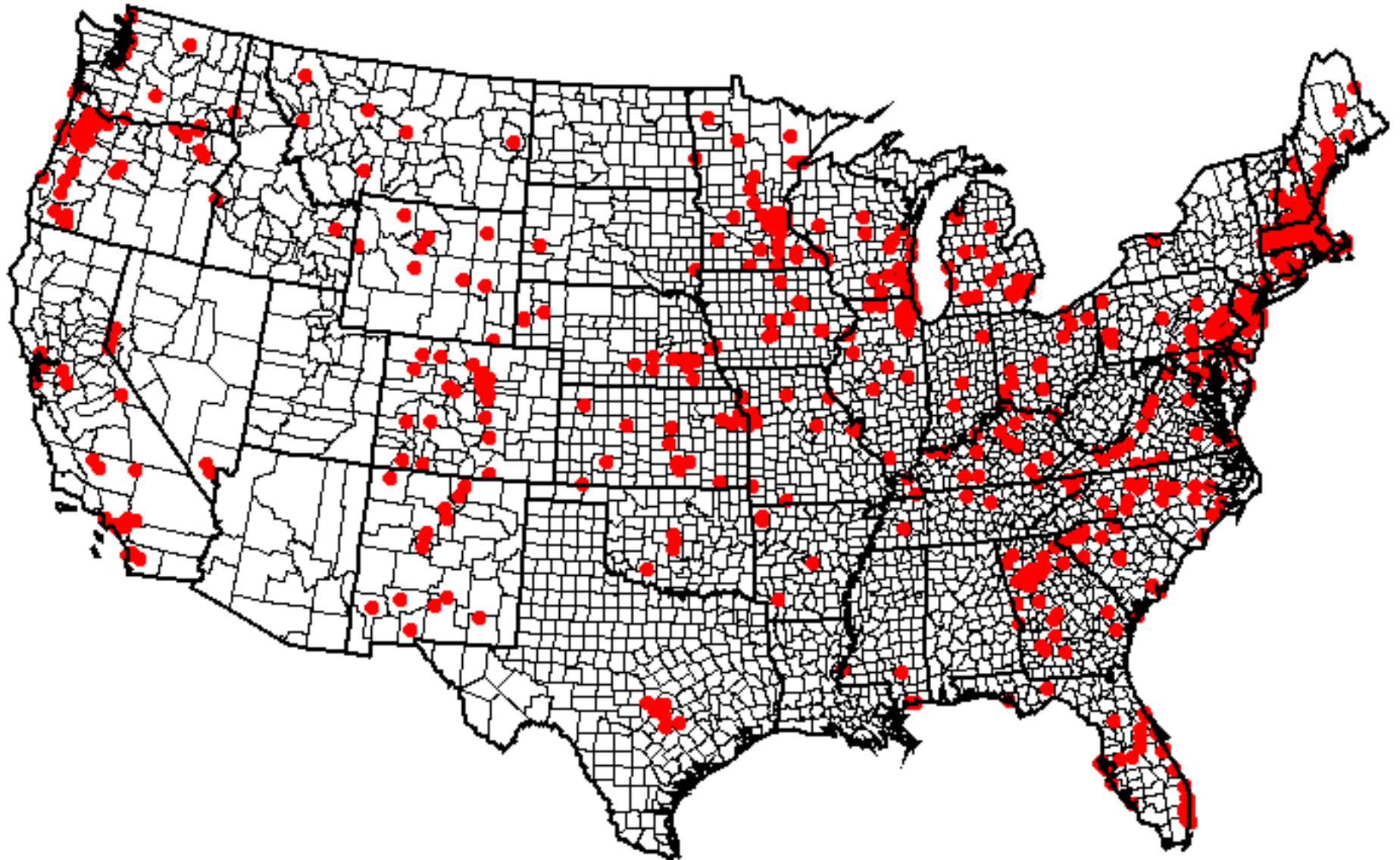




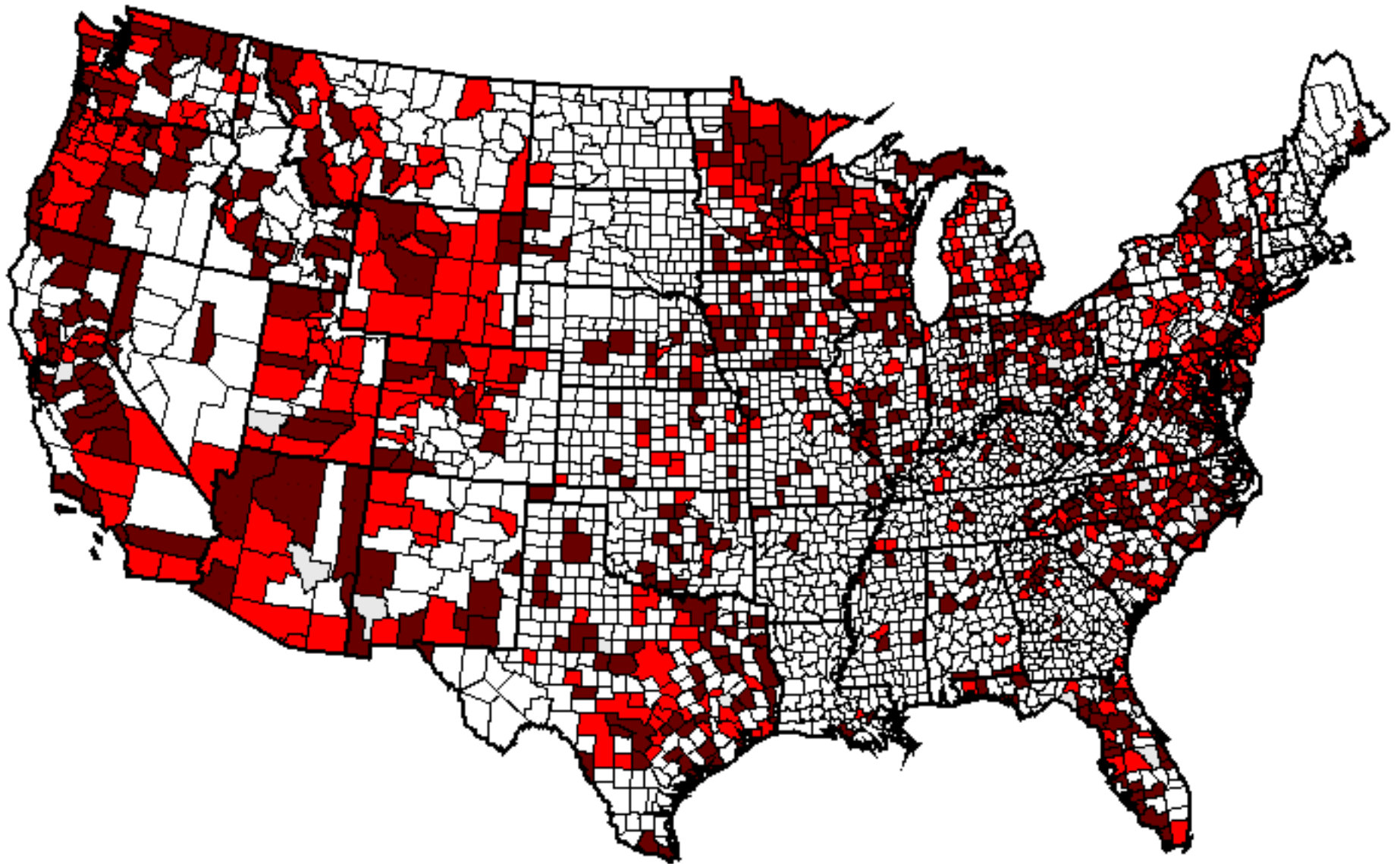
Do you create, update, integrate, and/or distribute digital geographic data?



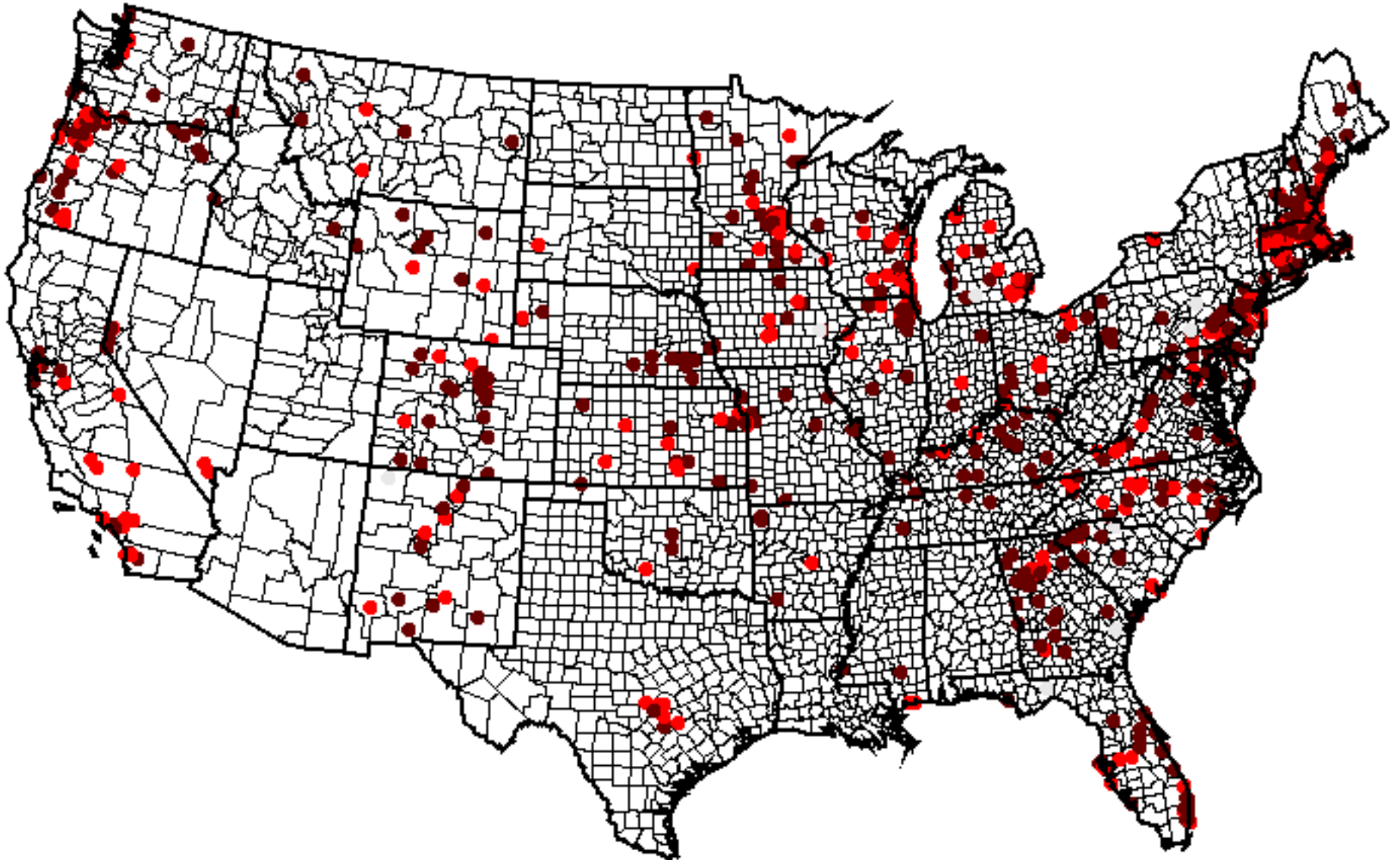
Do you create, update, integrate, and/or distribute digital geographic data?



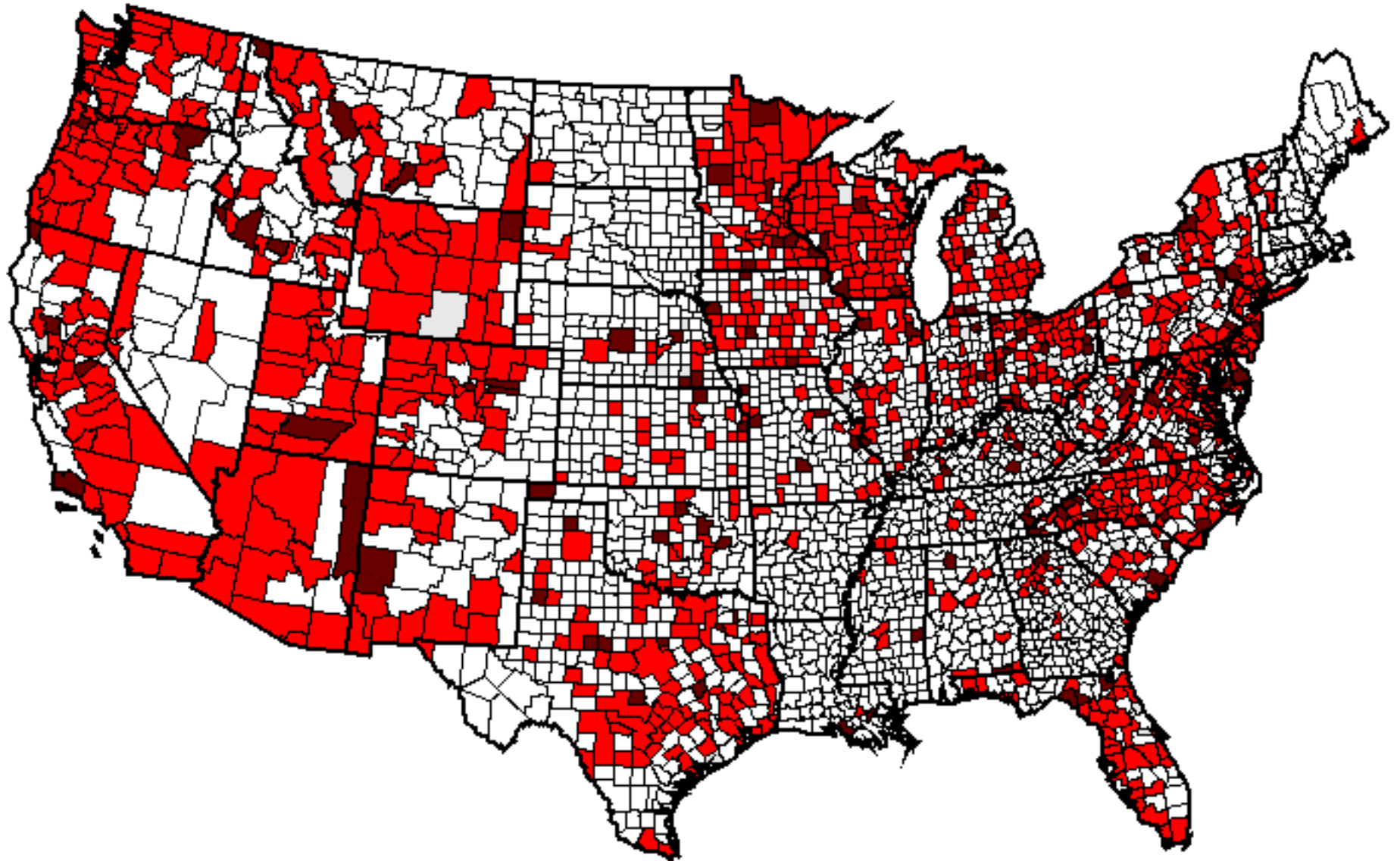
Does your organization participate with a geographic data coordinating council or group?



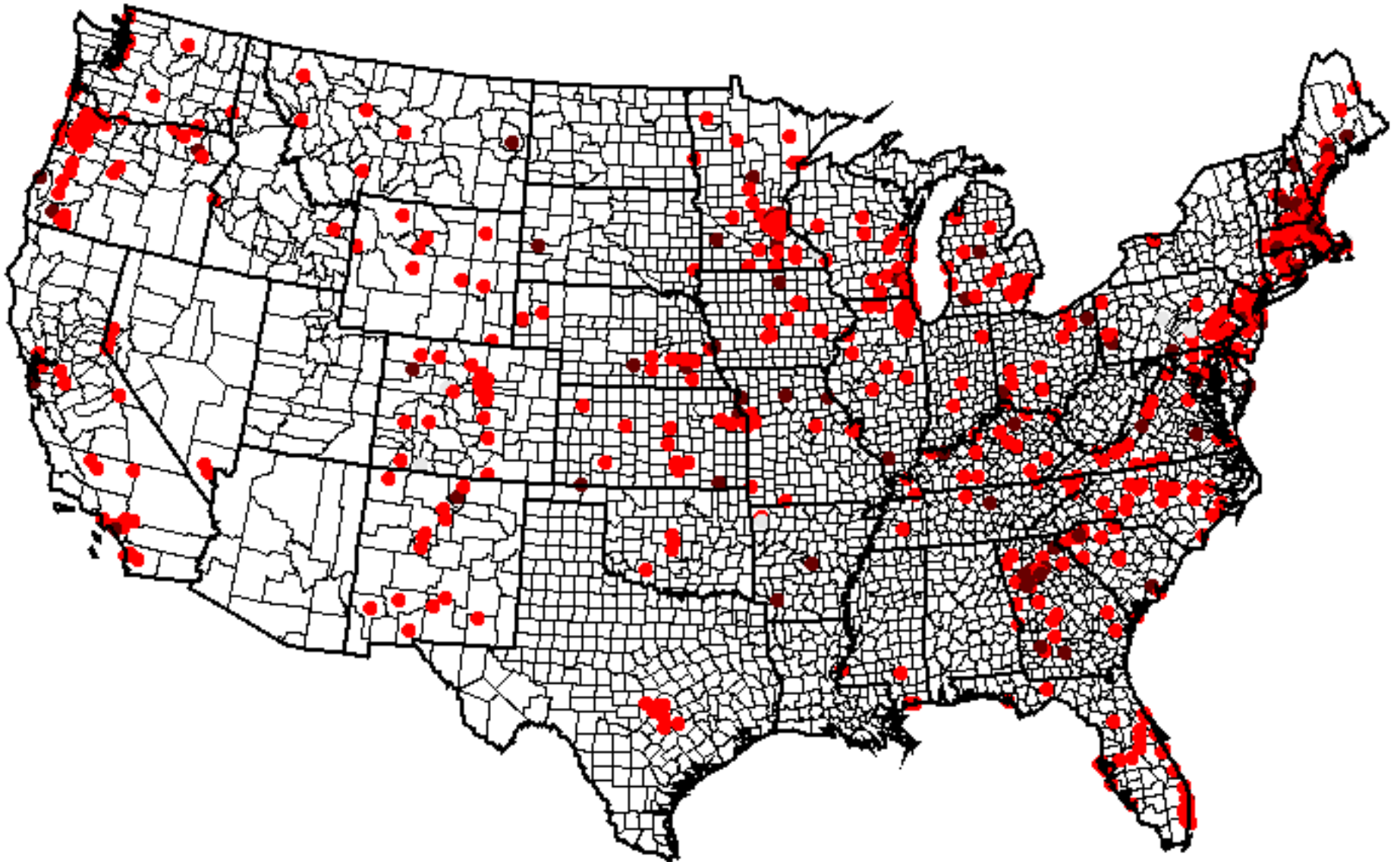
Does your organization participate with a geographic data coordinating council or group?



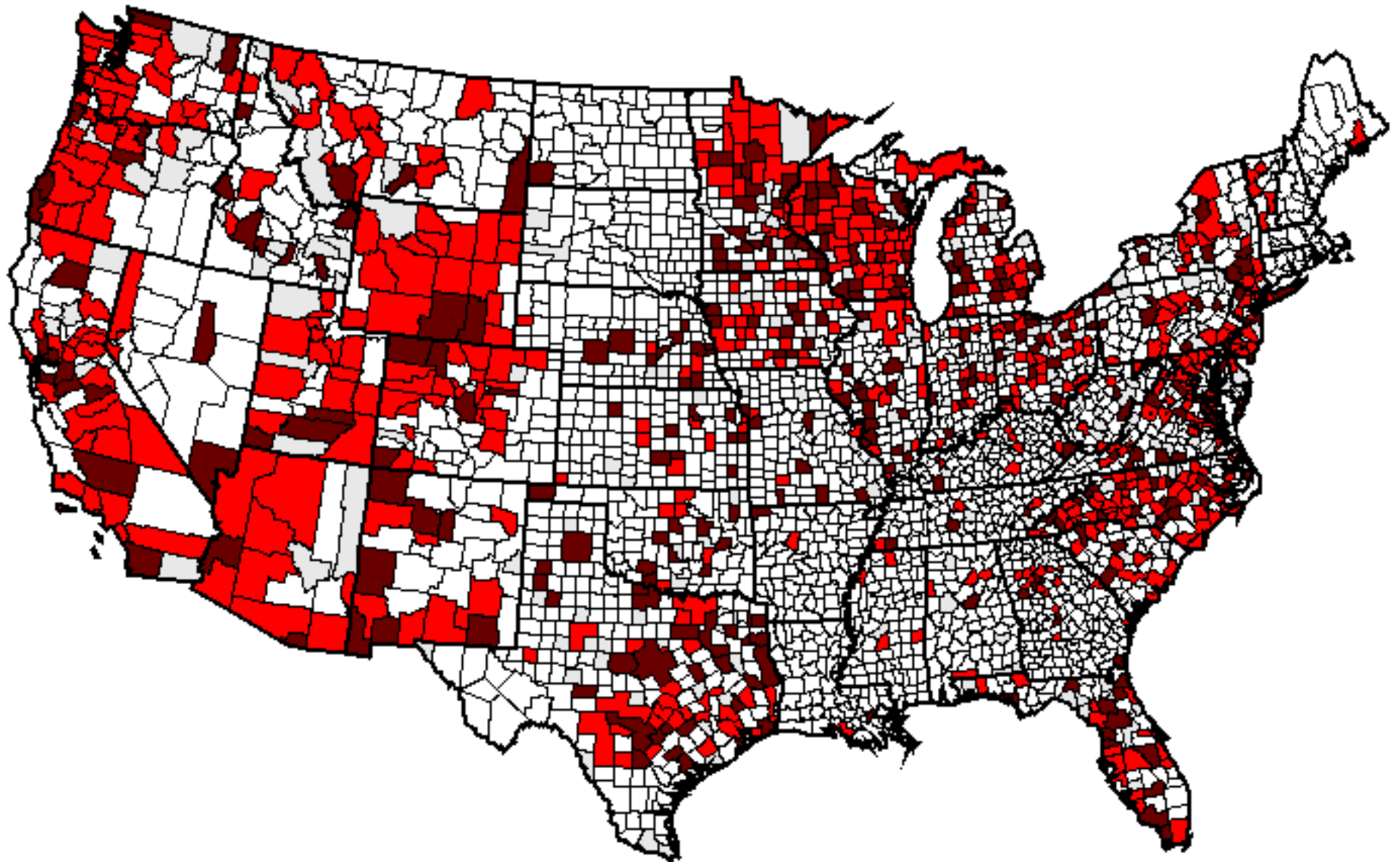
Does your organization share your data with other organizations?



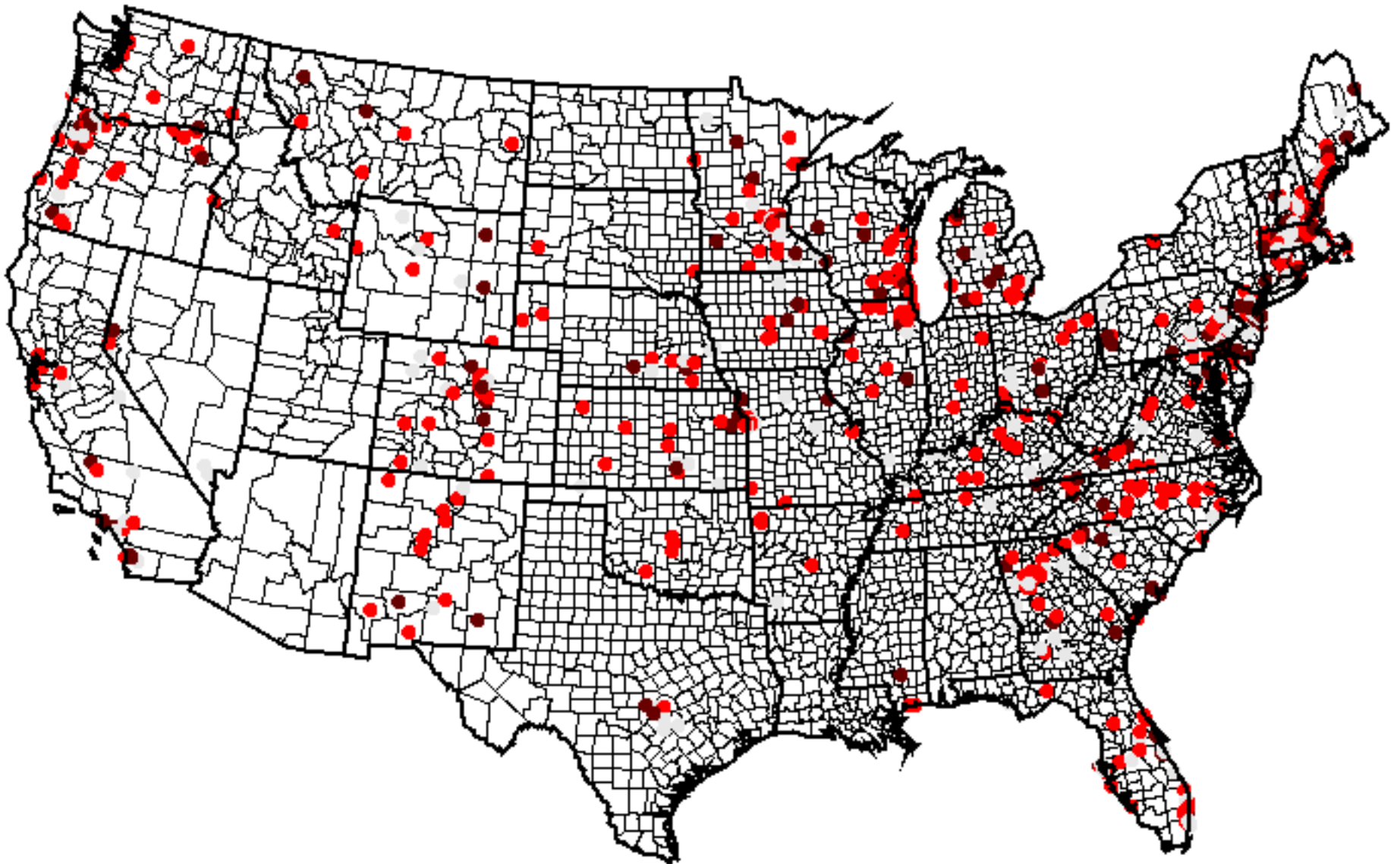
Does your organization share your data with other organizations?



Do you create, update, integrate or distribute  
framework transportation data?

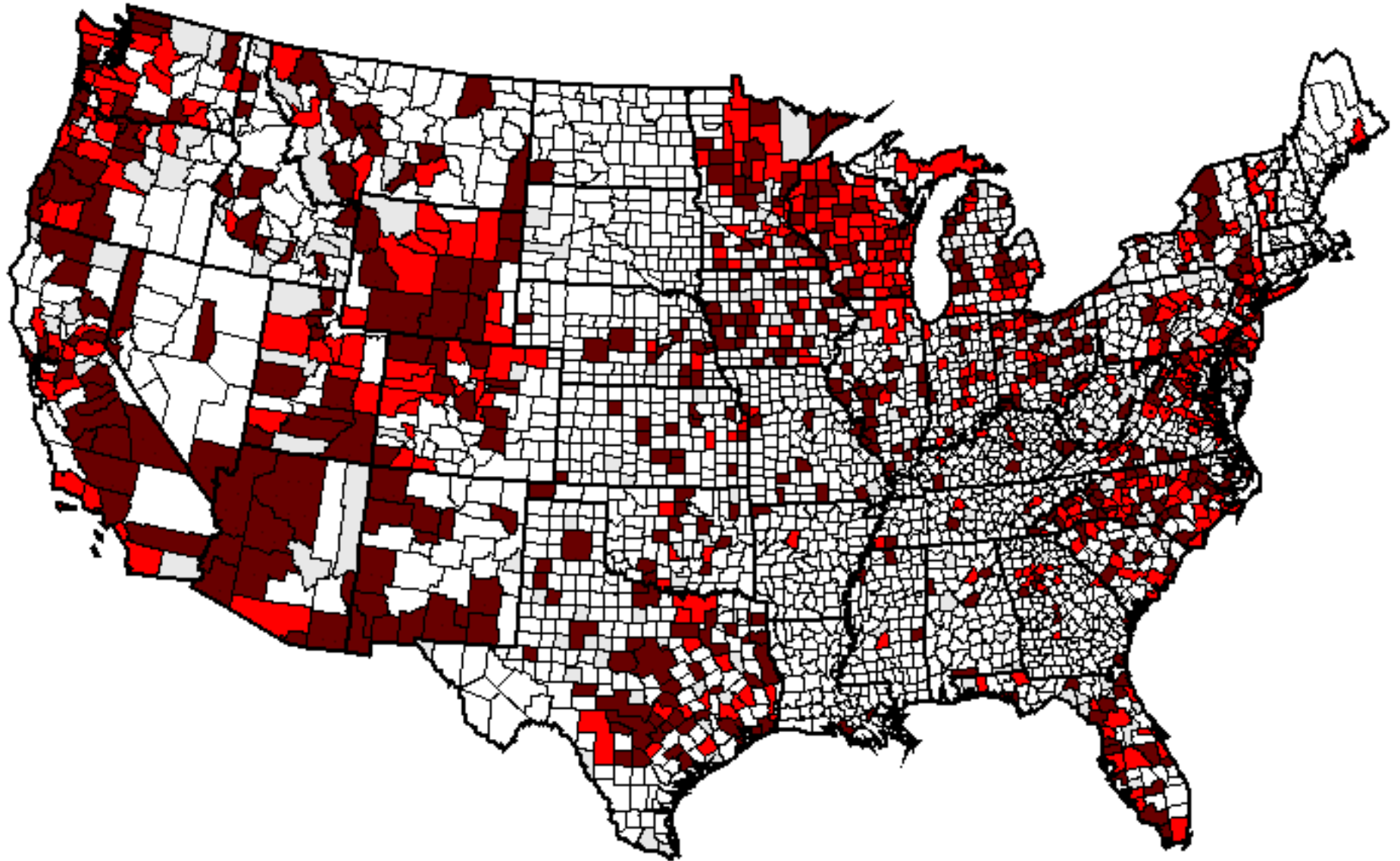


Do you create, update, integrate or distribute  
framework transportation data?

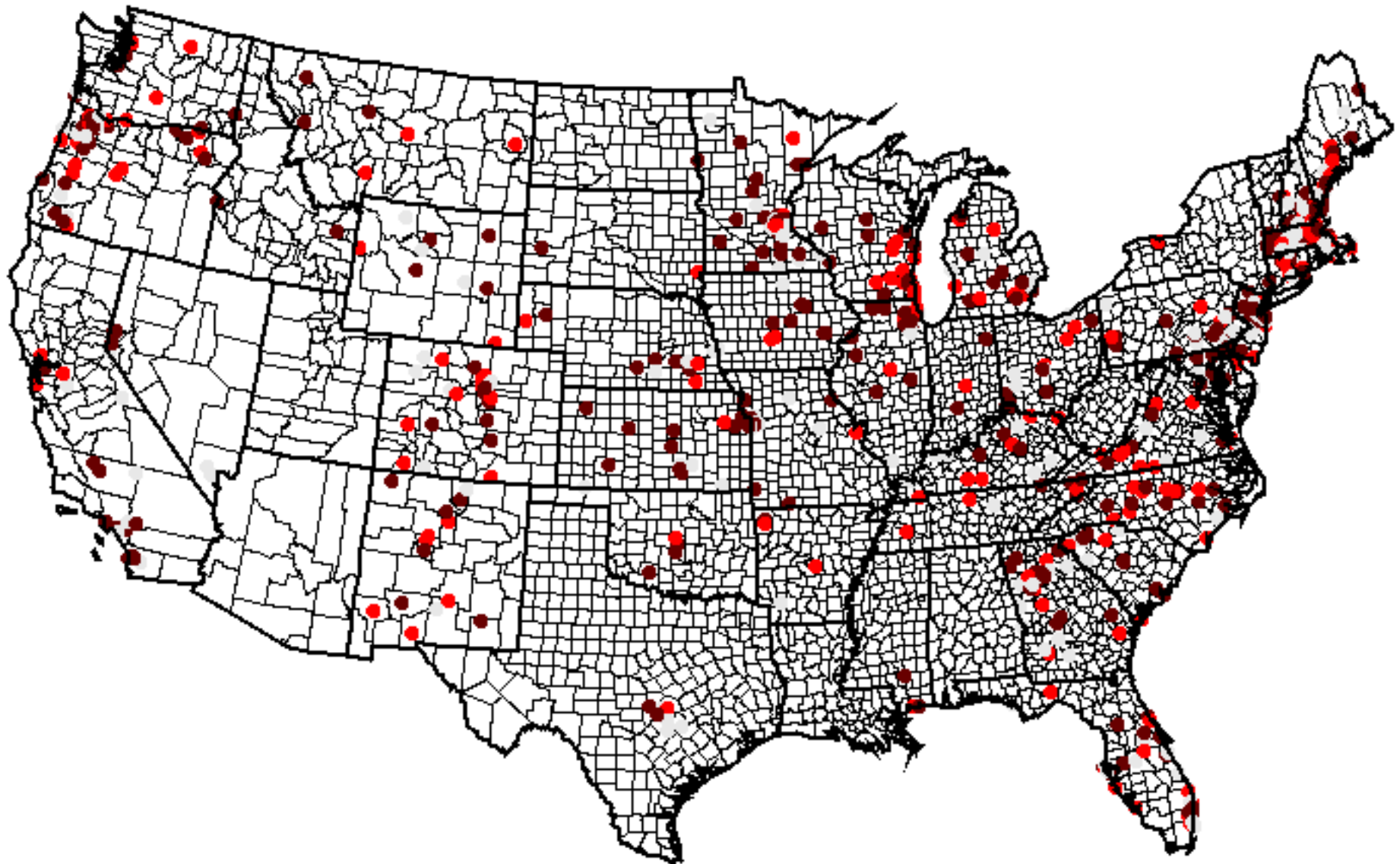




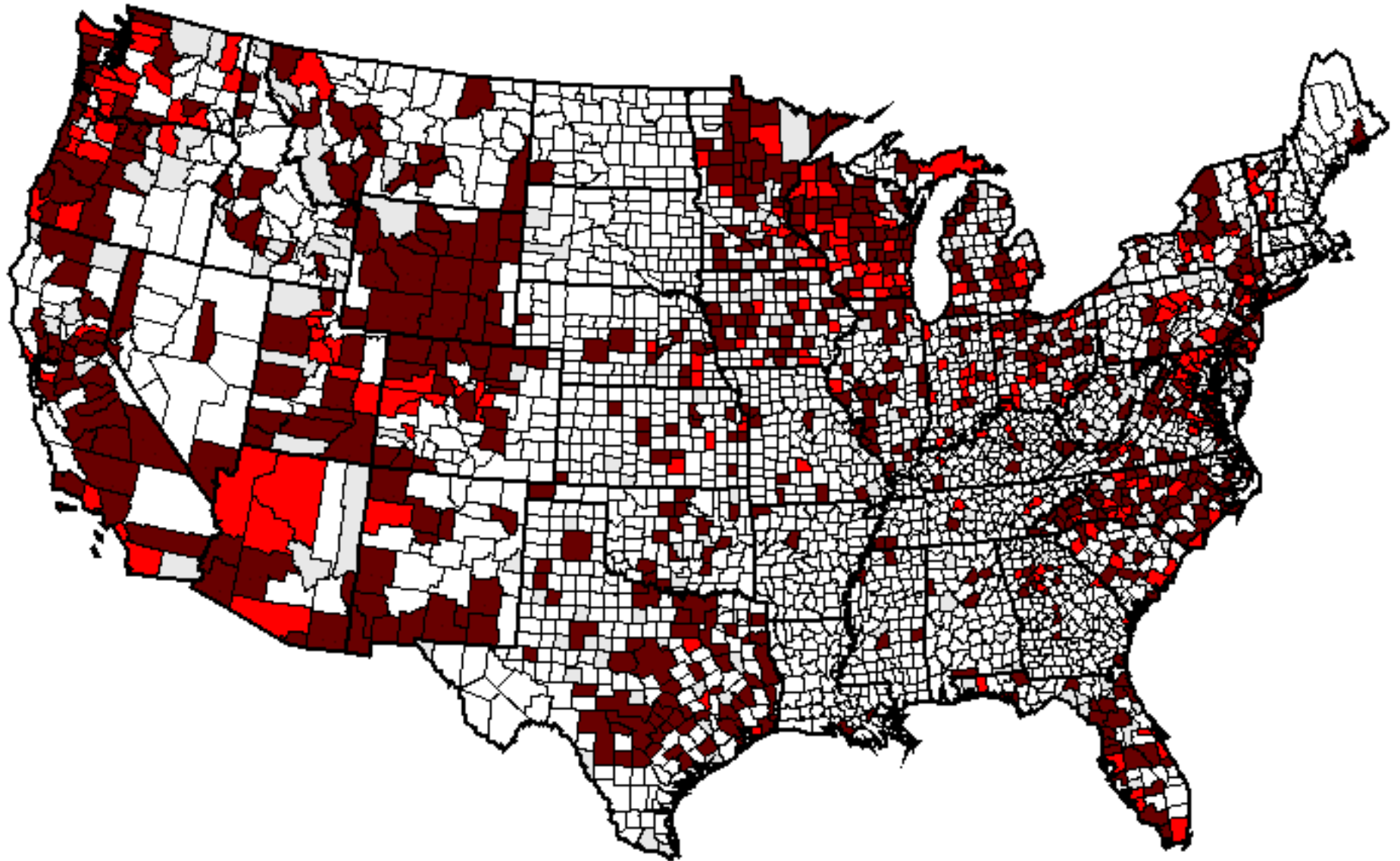
Do you create, update, integrate or distribute  
framework hydrography data?



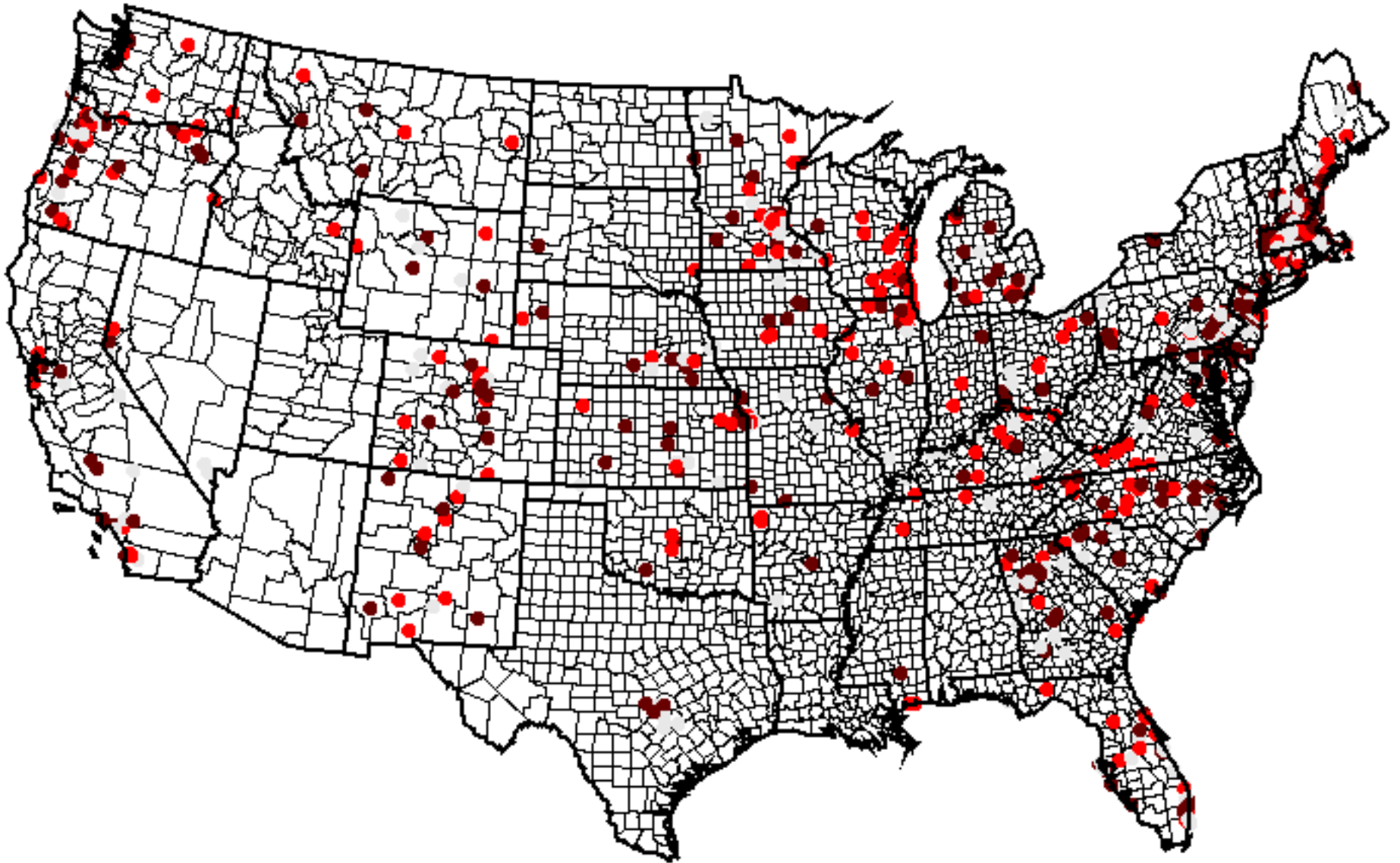
Do you create, update, integrate or distribute  
framework hydrography data?



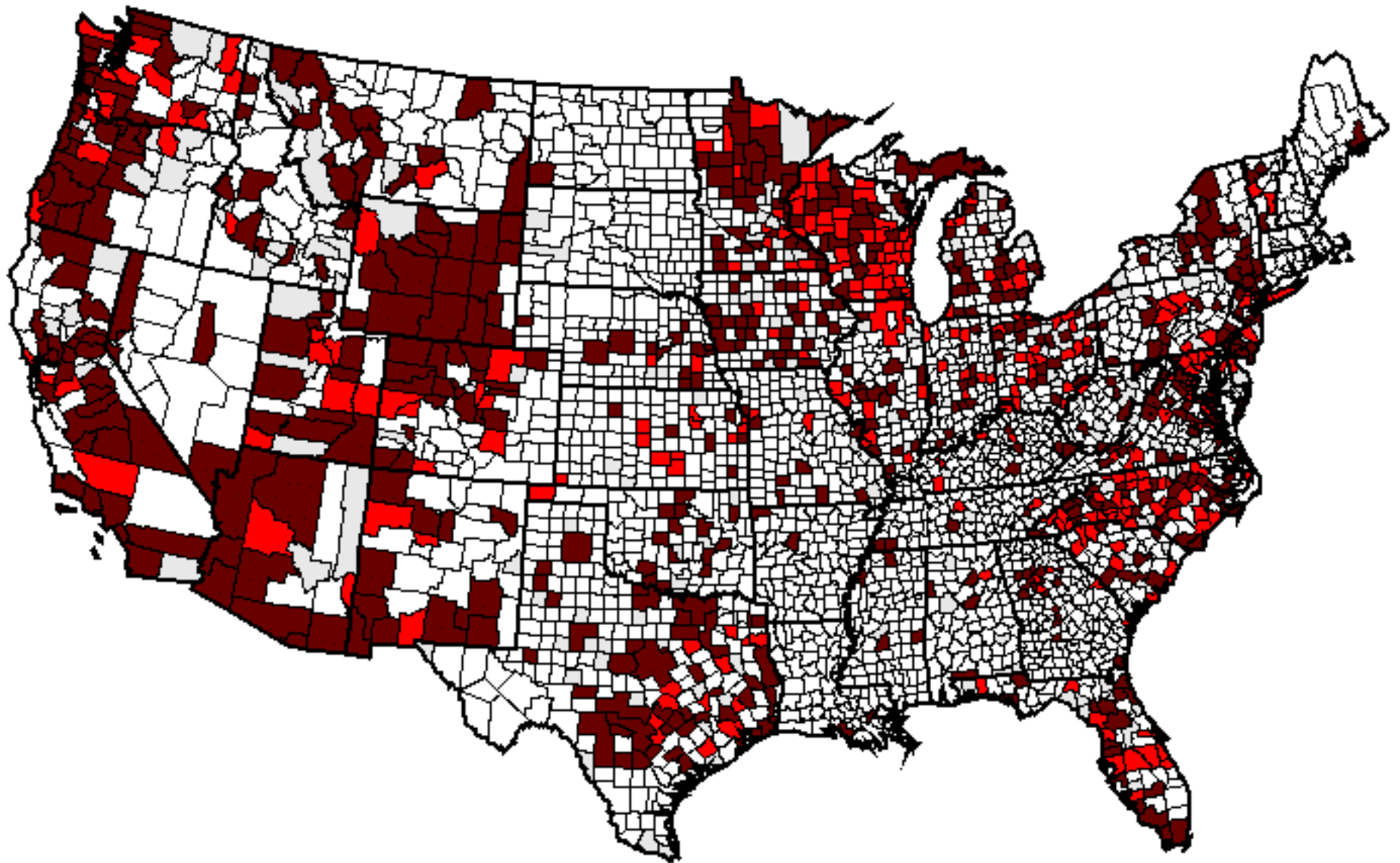
Do you create, update, integrate or distribute  
framework elevation data?



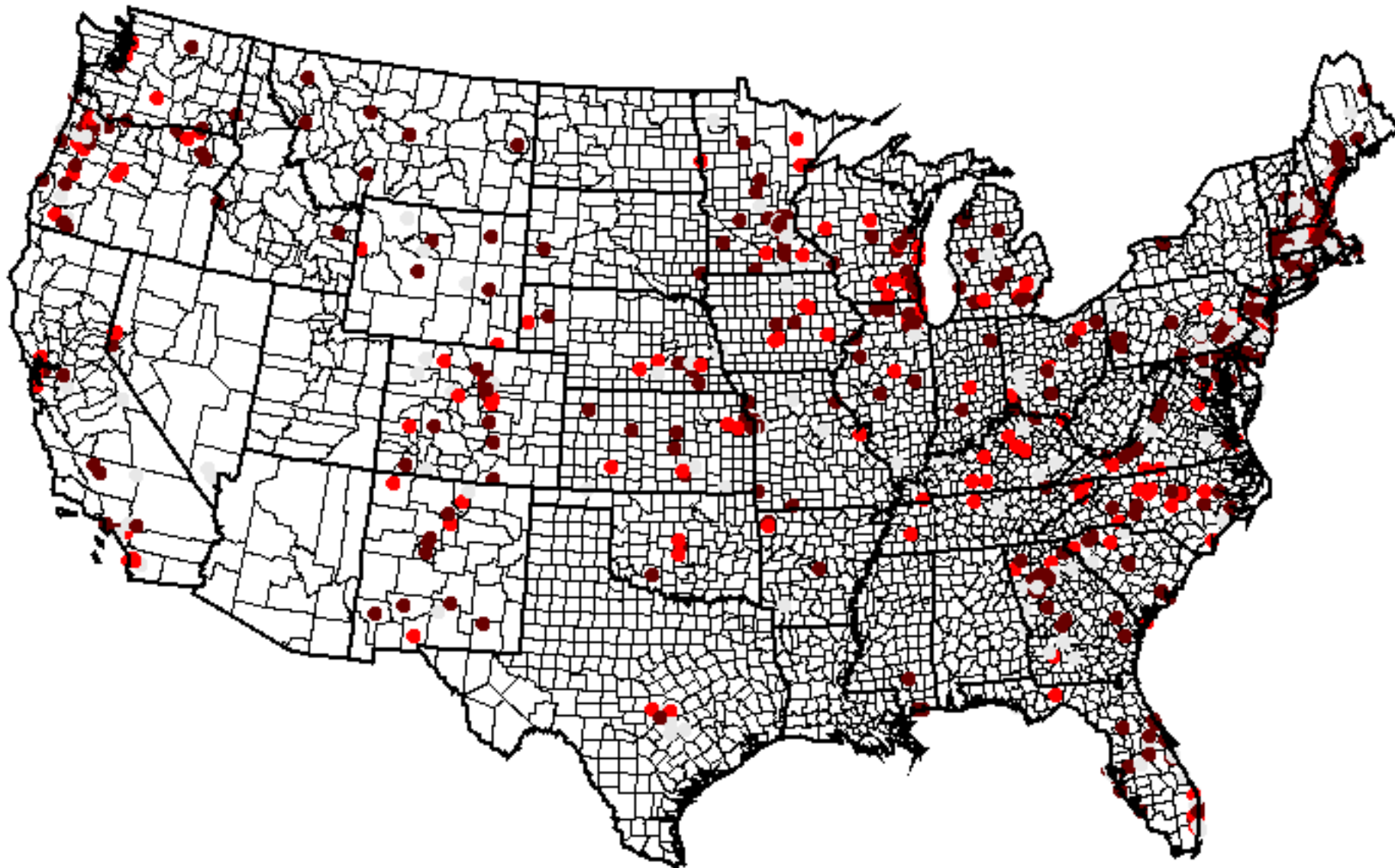
Do you create, update, integrate or distribute  
framework elevation data?



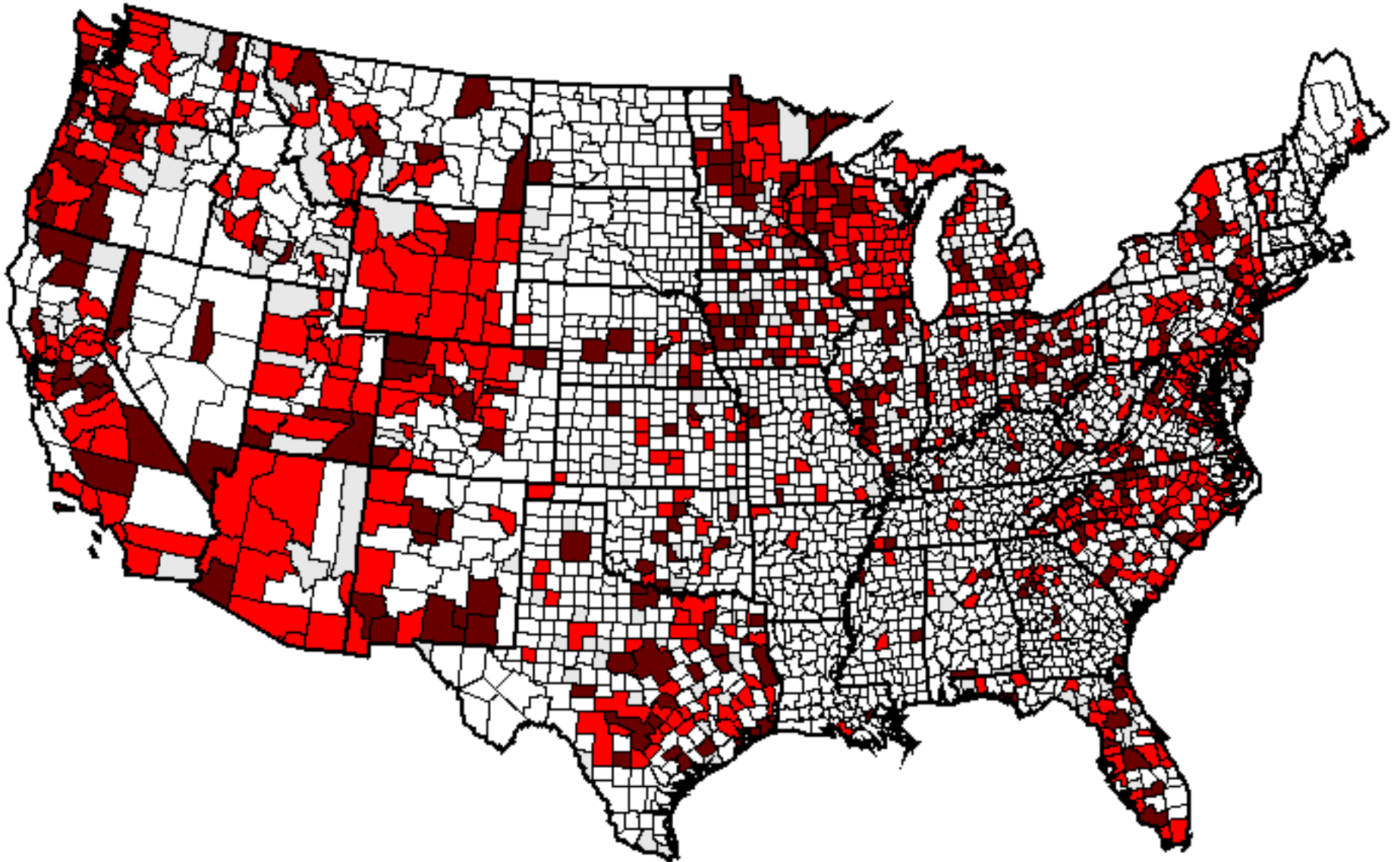
Do you create, update, integrate or distribute framework digital orthoimagery data?



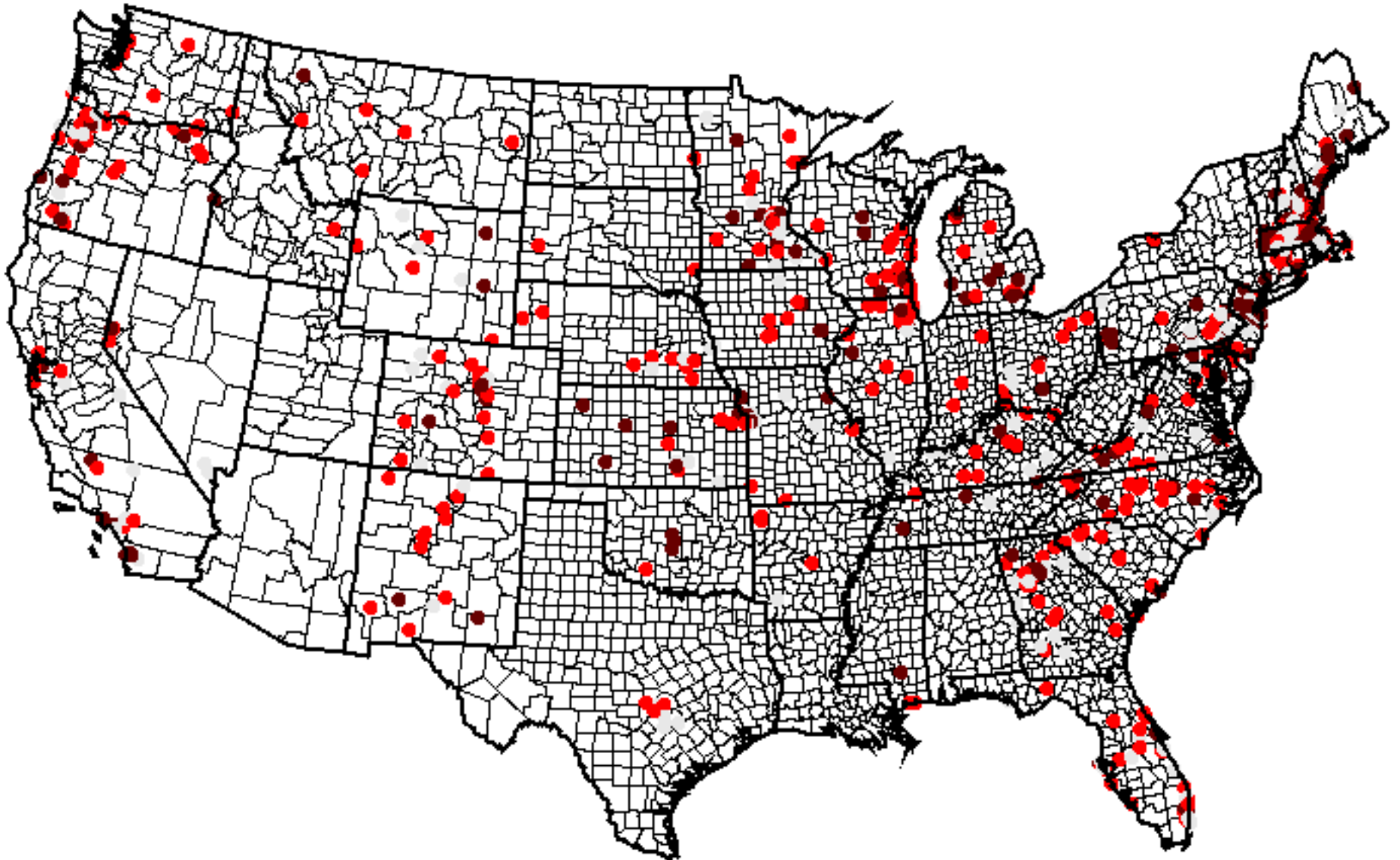
Do you create, update, integrate or distribute framework digital orthoimagery data?



Do you create, update, integrate or distribute framework government units (boundaries) data?

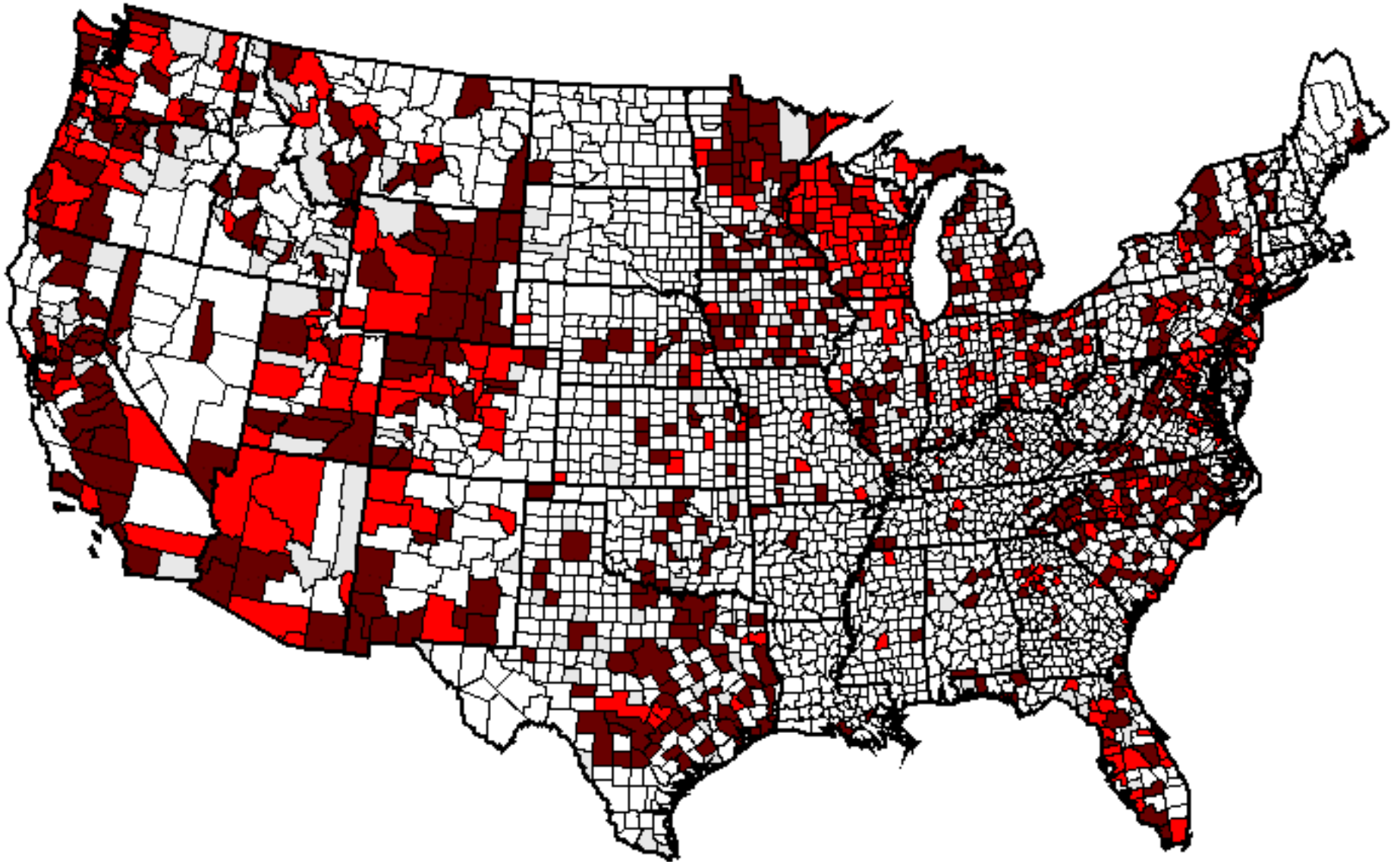


Do you create, update, integrate or distribute framework government units (boundaries) data?

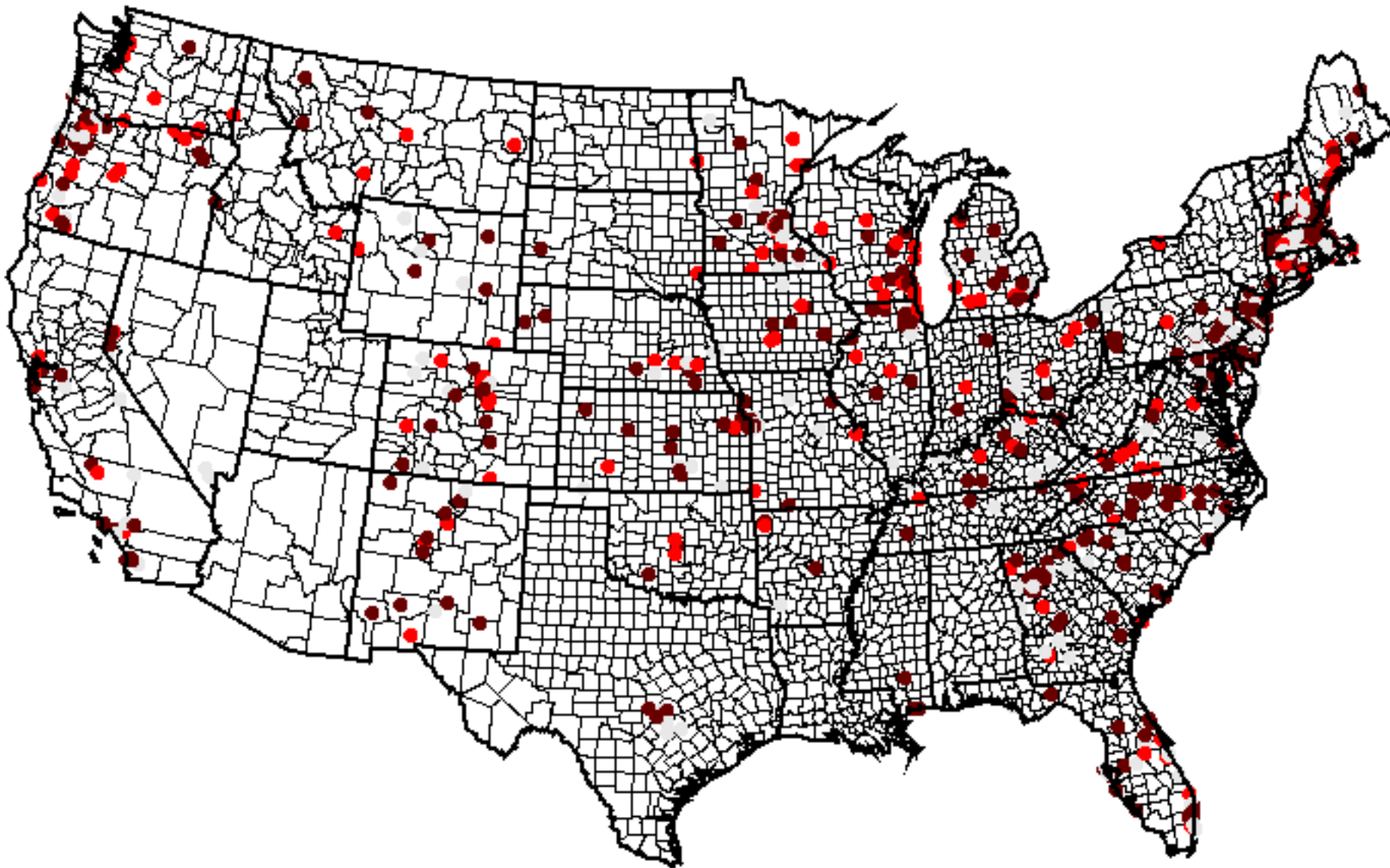




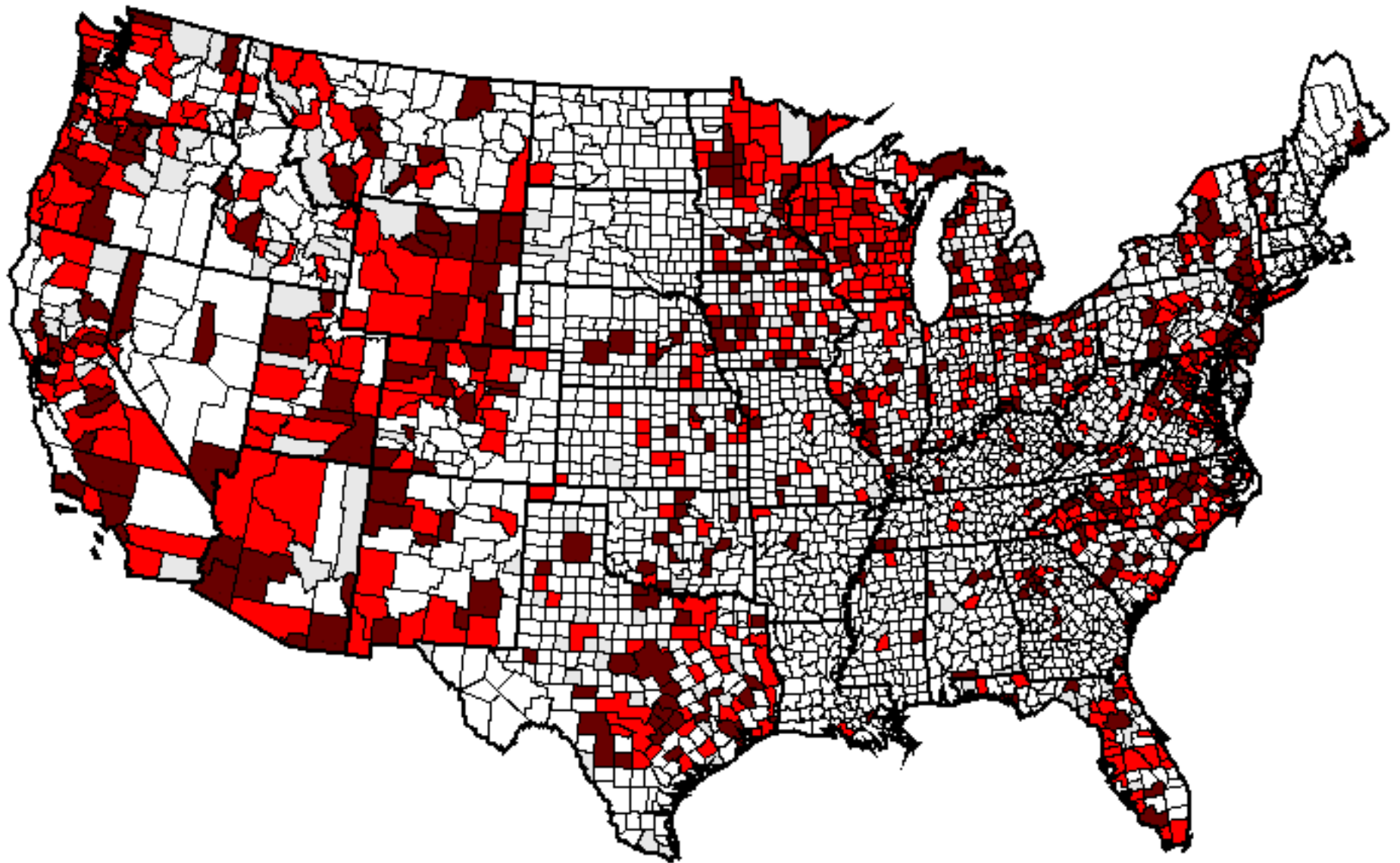
Do you create, update, integrate or distribute framework geodetic control data?



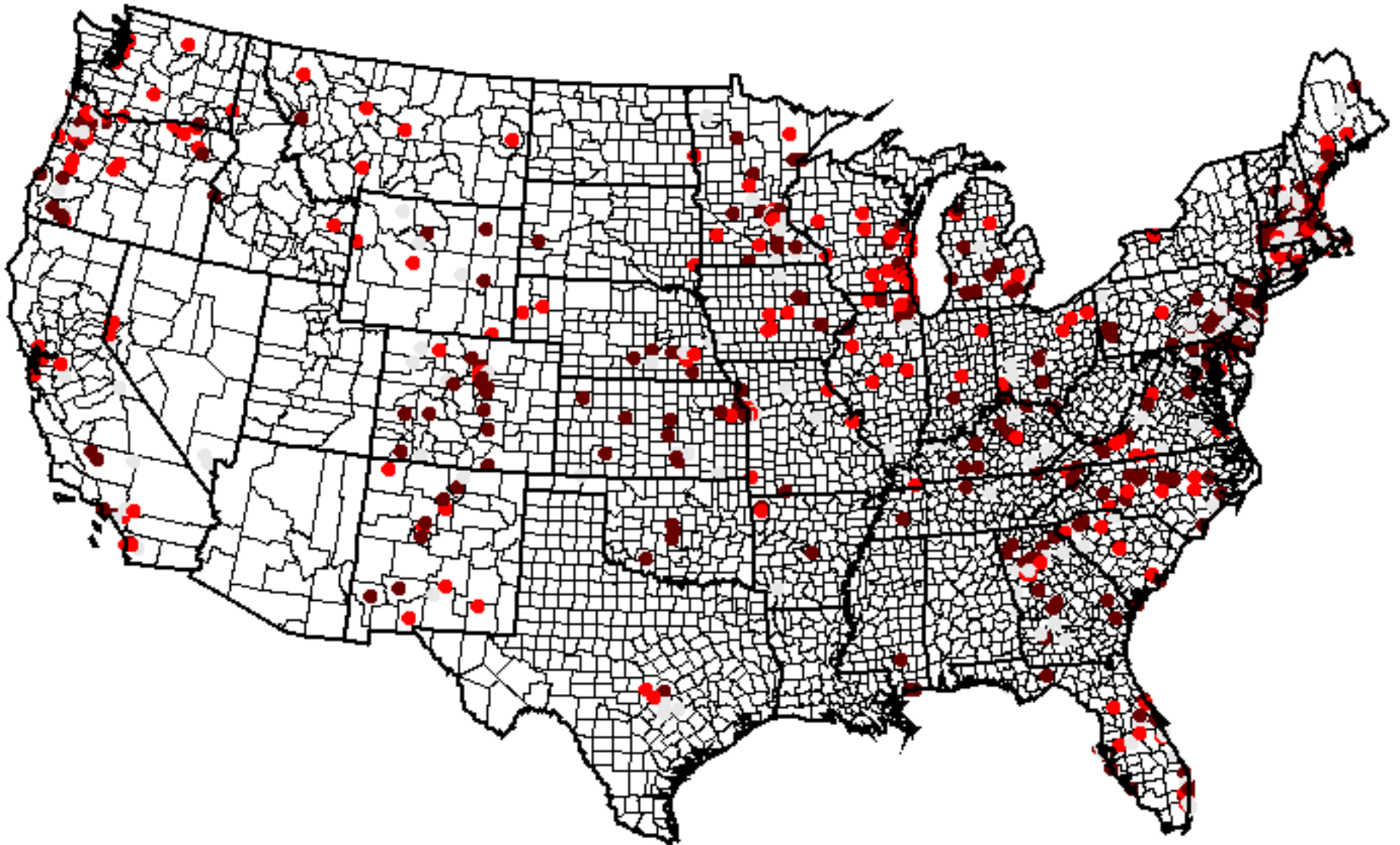
Do you create, update, integrate or distribute framework geodetic control data?



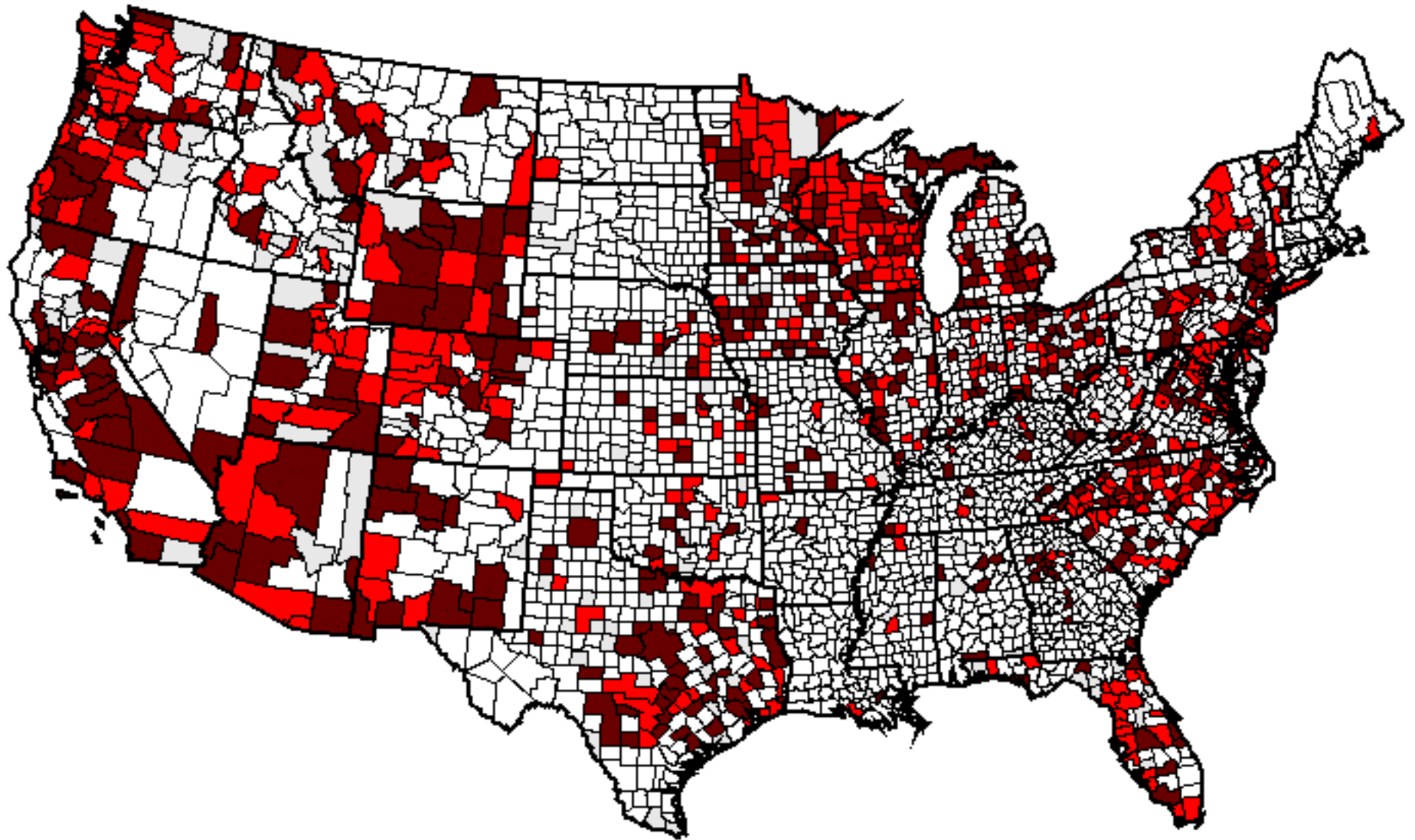
Do you create, update, integrate or distribute  
framework cadastral reference system data (PLSS)?



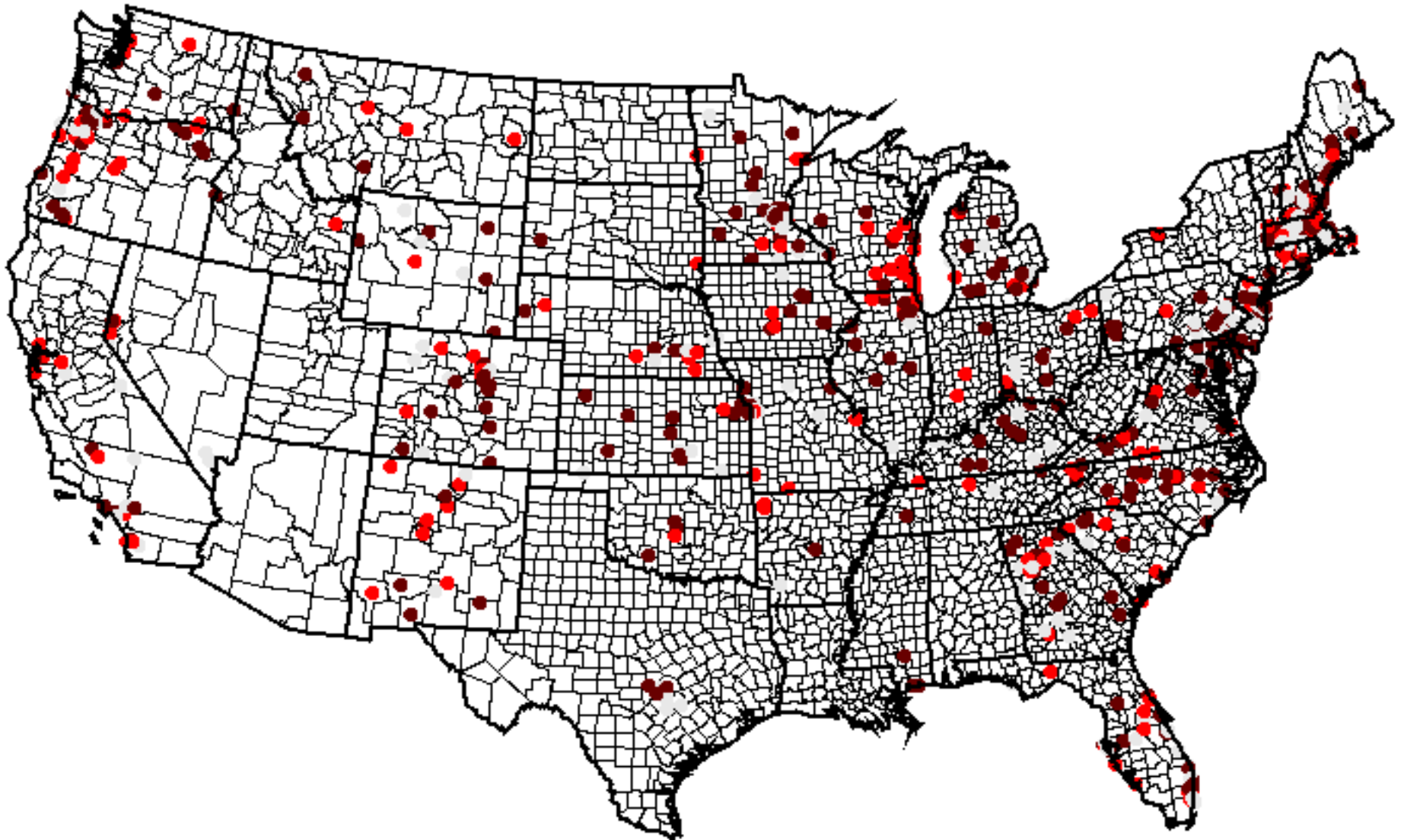
Do you create, update, integrate or distribute  
framework cadastral reference system data (PLSS)?



Do you create, update, integrate or distribute  
framework cadastral publicly-owned units data?  
(military reservations, national forests, parks, etc.)

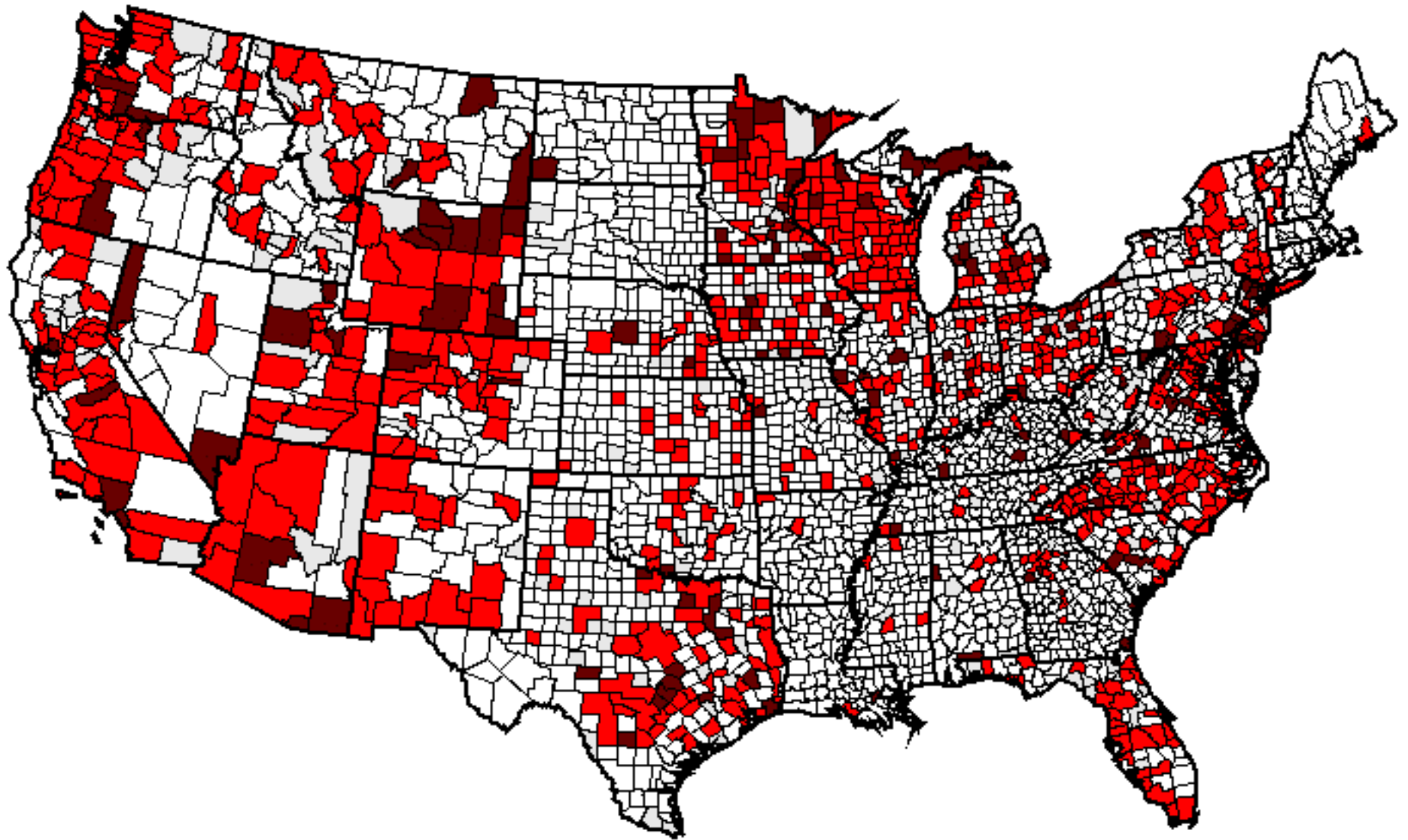


Do you create, update, integrate or distribute  
framework cadastral publicly-owned units data?  
(military reservations, national forests, parks, etc.)



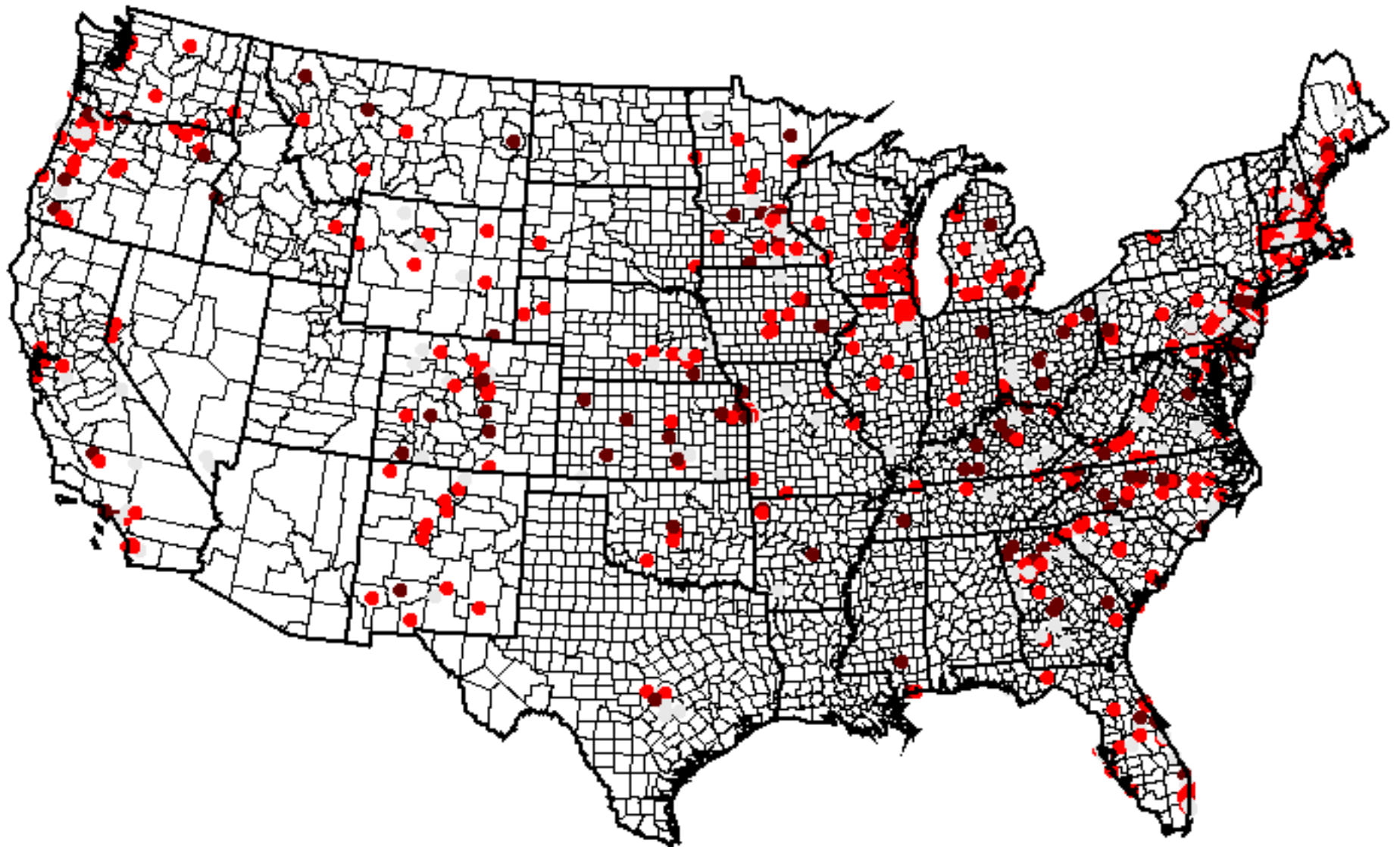
Do you create, update, integrate or distribute data that describe privately-owned parcels?

( This is not a framework layer but is of interest and value to many users )



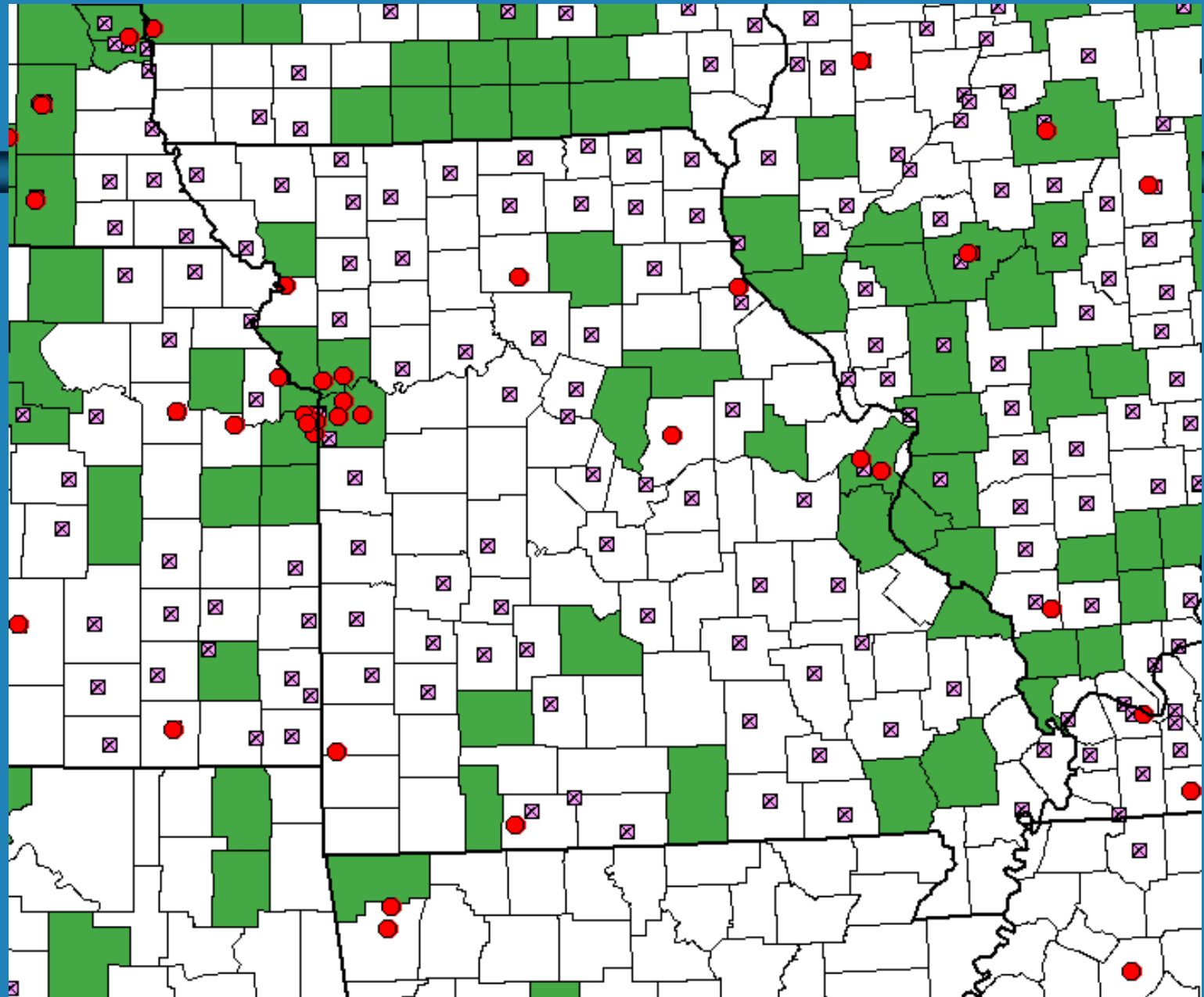
Do you create, update, integrate or distribute data that describe privately-owned parcels?

( This is not a framework layer but is of interest and value to many users )

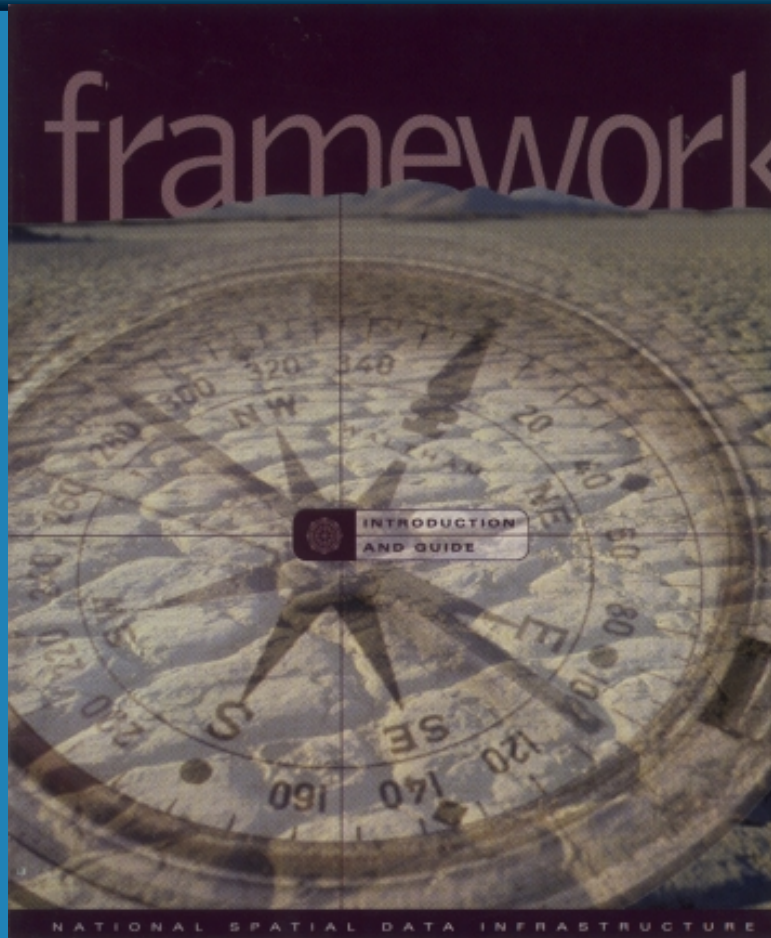




?



# Guide on “How to get started” Framework Introduction and Guide



<http://www.fgdc.gov/framework/frameworkintroguide/>