Assembling a Statewide Road and Address Database

Howard Ward, TerraSystems Southwest, Inc.
Curtis White, Global Systems Modeling, Ltd.

Prepared for GIS Cooperative Meeting
November 8, 2013

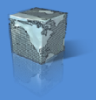
Overview

Need

Progress

Work Ahead

Database



Need

- AZ Wildland Fire Program
- AZ Broadband Mapping Program
- AZ 9-1-1 GIS Map Development Projects
- AZ All Roads Project (MAP 21 or ARNOLD)
- US National Address Database (NAD)

TSSW

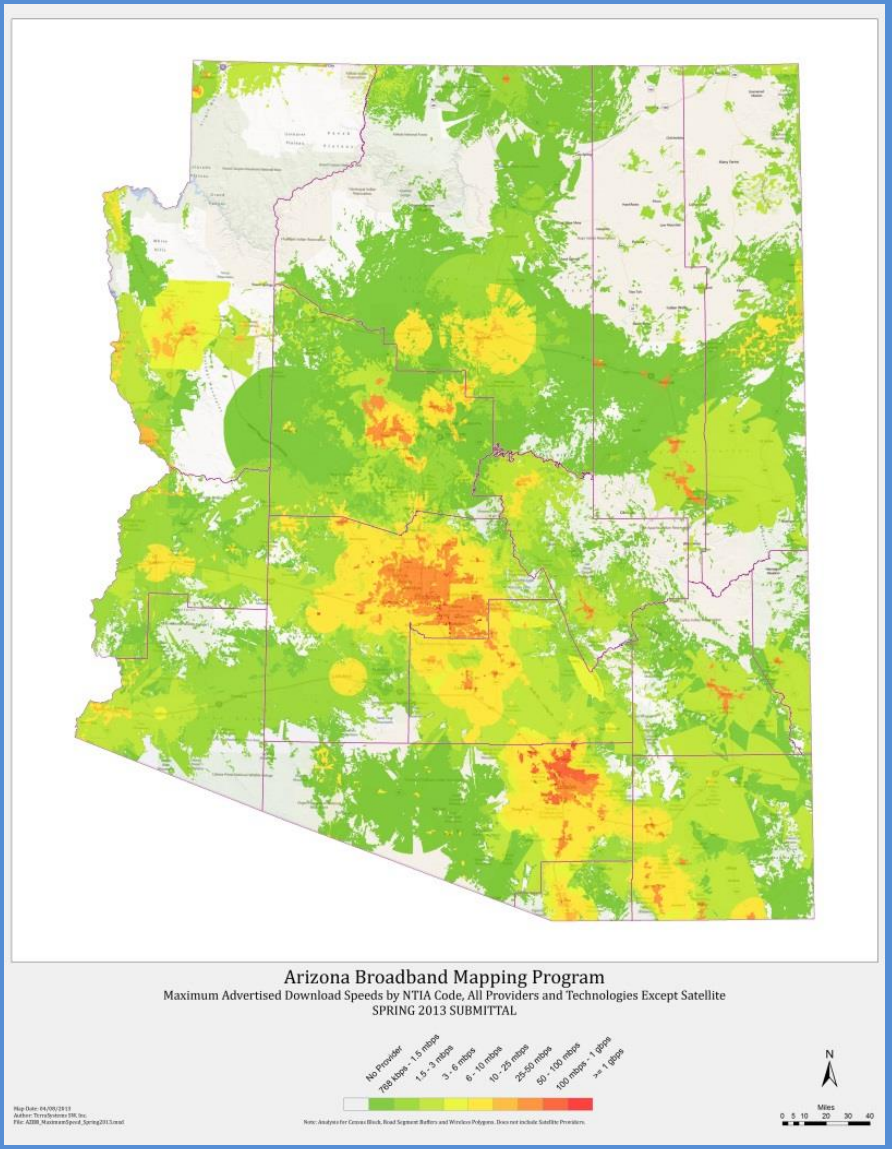
AZ Wildland Fire Mapping/Response

- Need current roads and addresses for emergency response coordination in wildland fire situations.
- Local parcel, address and road network data has been collected by the State over the past two years.
- This data is being provided to the National Fire Center.

TSSW

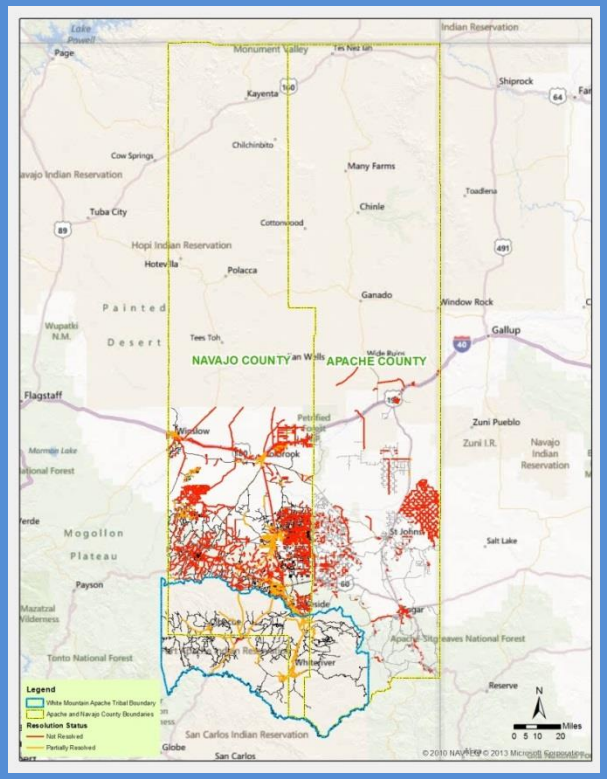
AZ Broadband Mapping Program

- The State is in **the 4th year** of this bi-annual collection and mapping process funded by the American Recovery Act.
- There are currently **75 BB providers** in the database as well as hundreds of points representing public and private community anchor institutions.
- To accurately map this information, the **AZBB Program needs better road and address data for geocoding** provider service addresses, especially in rural areas.



TSSW

AZ 9-1-1 GIS Map Development Project

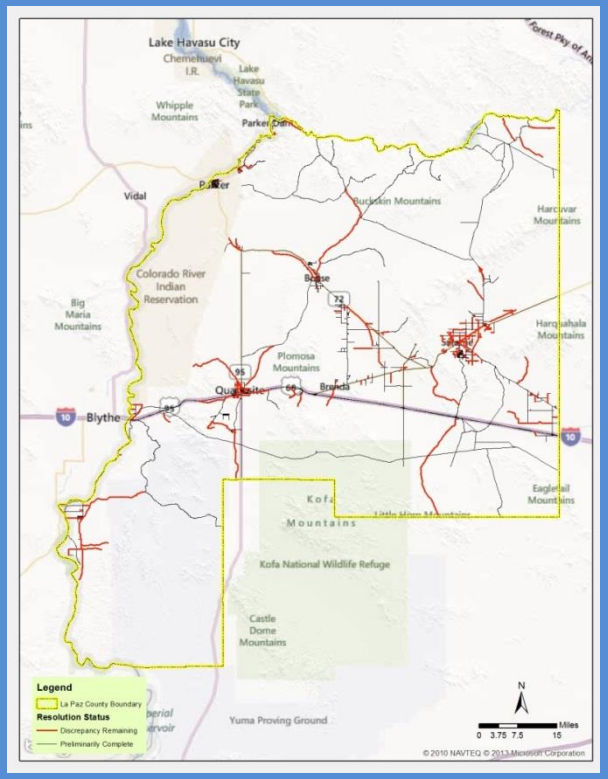


Navajo and Apache County

Three counties are without a road network suitable for E 9-1-1 services.

9-1-1, AZBB and ADOT teaming with **Apache, La Paz and Navajo counties** to create the data.

All three counties **should be complete by the February 2014**.



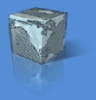
La Paz County

TSSW

AZ DOT All Roads (MAP-21) Initiative

- In 2012, the US Department of Transportation expanded the responsibilities of the State under the HPMS program to submit an all-public-road-network by June 15, 2014.
- The all-public-road-network is intended to be a fully functional Linear Referencing System (LRS) which can support several initiatives.
  - Be the foundation for Statewide roadway data
  - Validate the annual Certified Public mileage report
  - Support traffic data for safety analysis and,
  - The National Bridge Inventory (NBI) inspection and reporting.

TSSW



Progress

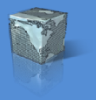
The data vintage from the counties ranges from two years to two months in age.

Some agreements with providing agencies are pending.

We received recent updates from several counties but not in time for the most recent AZBB submittal process.

County	Data Vintage	No. of Features
Apache	**pending**	**pending**
Cochise	Aug - 2012	17,615
Coconino	Feb - 2013	46,492
Gila	Jan - 2013	4,098
Graham	Aug - 2013	1,469
Greenlee	Feb - 2013	3,076
La Paz	**pending**	**pending**
Maricopa	Feb - 2011	195,167
Mohave	Feb - 2011	37,911
Navajo	**pending**	**pending**
Pima	Aug - 2012	58,874
Pinal	Aug - 2013	41,377
Santa Cruz	Feb - 2011	6,395
Yavapai	Aug - 2013	53,578
Yuma	Aug - 2012	15,710

TSSW

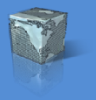


Progress

We are comparing our geocoding results from the latest data collection cycle with various other road network sources to track the benefits of this approach.

Source	Total Records	Matched	Tied	Unmatched
Navteq	2,533,035	2,509,848 (99.1%)	5,818 (0.2%)	17,369 (0.7%)
Census 2009	2,533,035	2,052,279 (81.0%)	6,137 (0.2%)	474,619 (18.8%)
Census 2010	2,533,035	2,184,281 (86.2%)	11,765 (0.1%)	336,989 (13.7%)
Local Append 1	2,533,035	2,103,407 (83.0%)	233,031 (9.2%)	196,597 (7.8%)
Local Append 2	2,533,035	2,327,405 (91.9%)	198,099 (7.8%)	7,531 (0.3%)

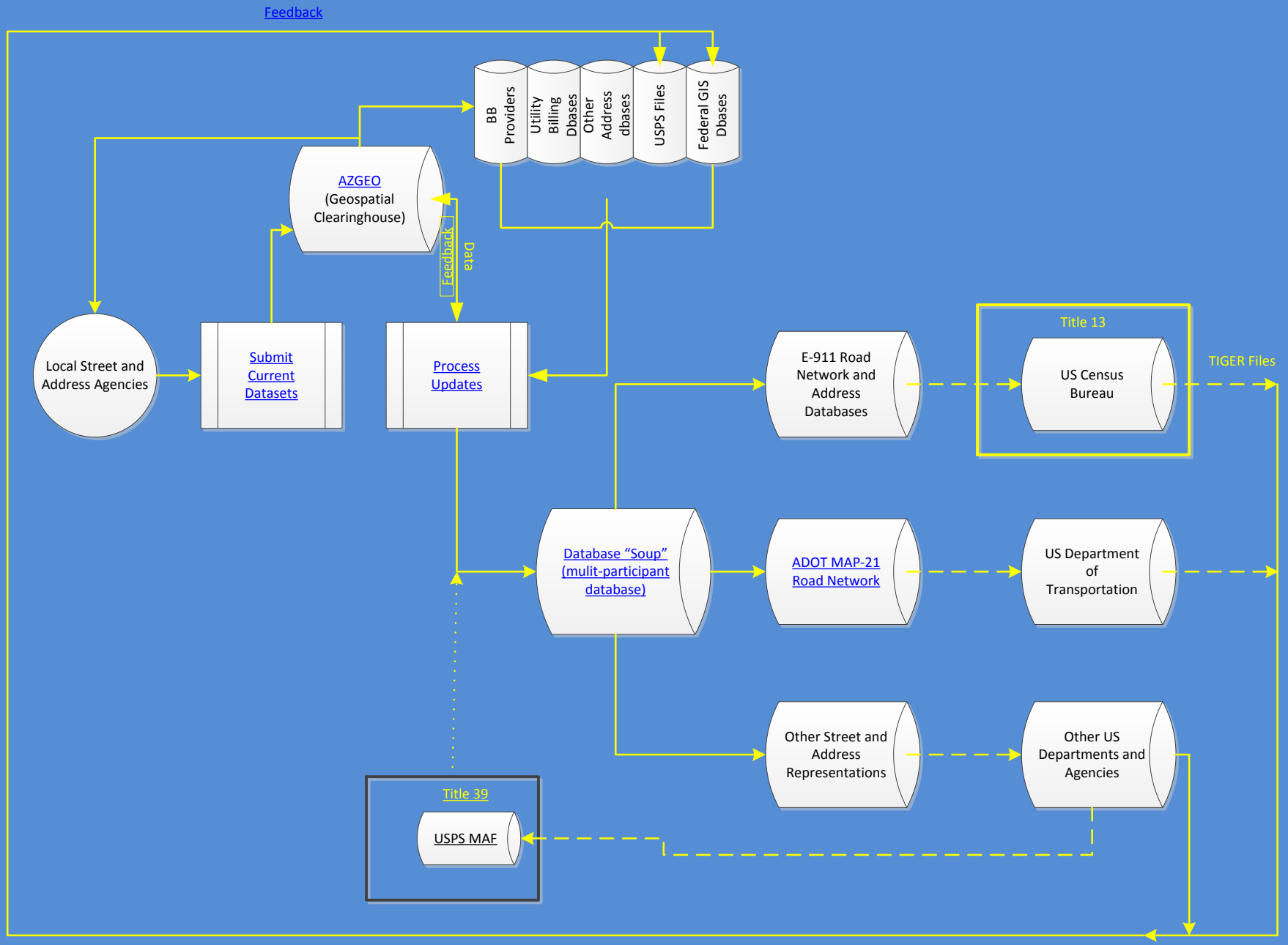
TSSW



Work Ahead

- Quality assure data being produced for the three remaining counties.
- Develop formal outreach and feedback process.
- Develop workflow, database(s), procedures and tools to assemble and maintain a statewide street network and address database.

The Work Ahead



Road Network Hierarchical Framework

A proposed Database Design to accommodate multiple representations of road networks, including linear and polygonal features.

Curtis White

Global Systems Modeling Ltd.

November 8, 2013

The Problem

- **E9-1-1**: Two distinct needs:
  - Location: Typically done by geocoding an address against a road network
  - Dispatch / Routing: Typically done by reference to a road network.

  Problem: These road networks have different requirements in terms of level of detail.
- **AZ Broadband**: Needs to report some data based on road network elements.
  - Census TIGER/Line data is the preferred road network data source for reporting.
  - Proprietary (NavTeq) and local road network data sets yield much better results in geocoding address-sourced data.

  Problem: Discrepancies between TIGER/Line and local data makes it difficult to match them up. Proprietary data cannot be used for reporting; spatial features are inaccessible for exporting.

global systems modeling limited

The Problem

There are numerous other GIS applications that rely on road networks. These may be classified as those that are:

- Road network status related: Planning, construction, maintenance (pavement, signage, etc.), removal (abandonment).
- Road network usage related: Traffic analysis and control, maintenance (e.g. weather related), vehicle dispatch and routing, closest facility, location allocation, origination-destination matrix, demographic and other analysis.
- Background related: Base maps.

global systems modeling limited

The Proposed Solution

Create a methodology that allows all applications to share the same spatial feature representation, as appropriate to their "level of detail", while allowing each application to have its required attribute data.

The proposed solution is a hierarchical framework with explicit feature associations between levels that allow generation of the required features while maintaining application-specific data.

Roughly corresponds with several existing hierarchical road network structures (Census, HPMS).

global systems modeling limited

The Proposed Solution

This solution fits very well with our “intuitive” way of thinking about road networks.

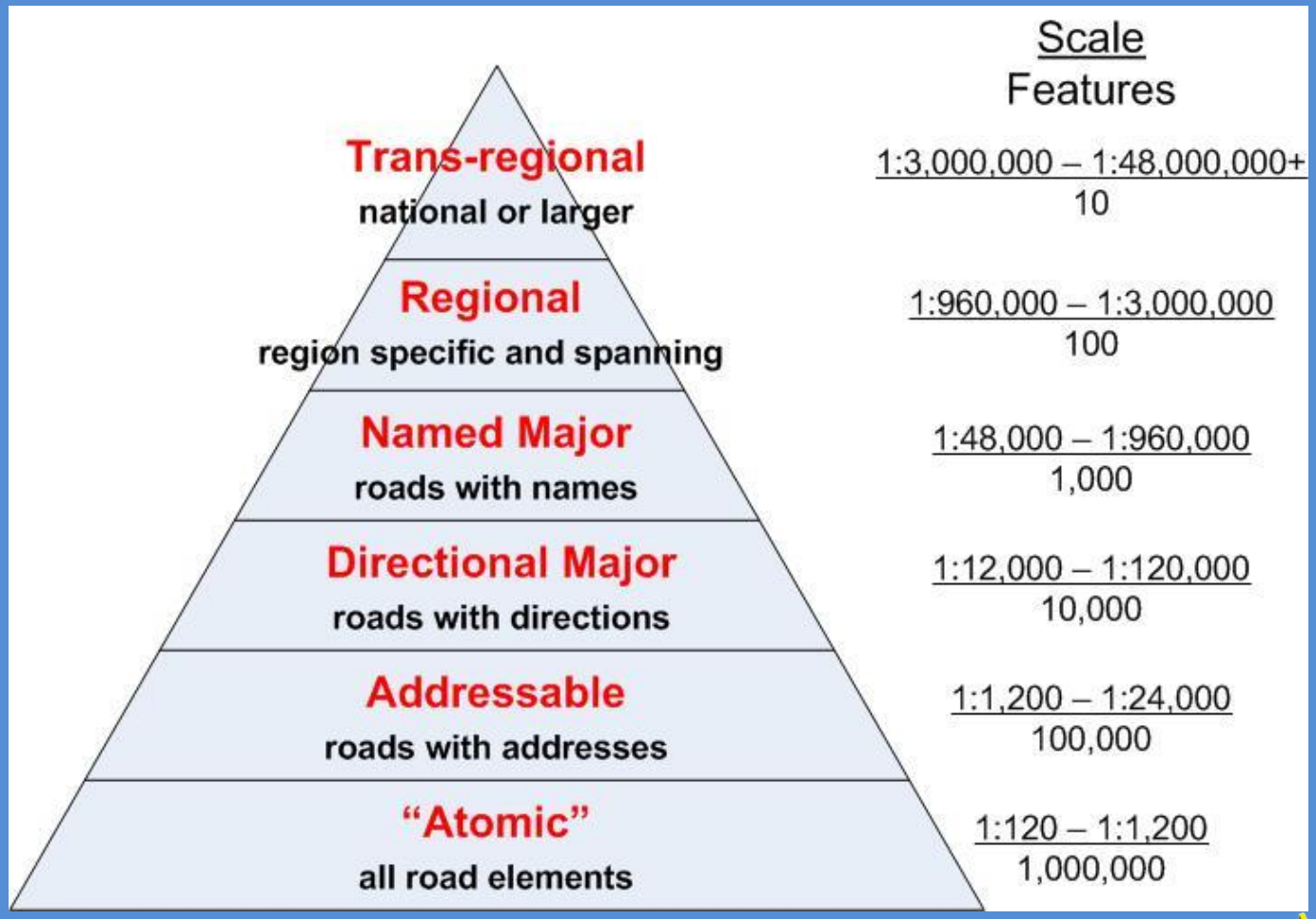
It basically looks at road networks based on the geographic “scale” required by an application:

- “Coarse” resolution applications require fewer road elements, e.g. Interstate and major highway maps.
- “Medium” resolution applications require more road elements, e.g. bus route maps.
- “Fine” resolution applications require the most road elements, e.g. city street maps.

Think of “scale-dependent” visibility for layers in ArcMap.

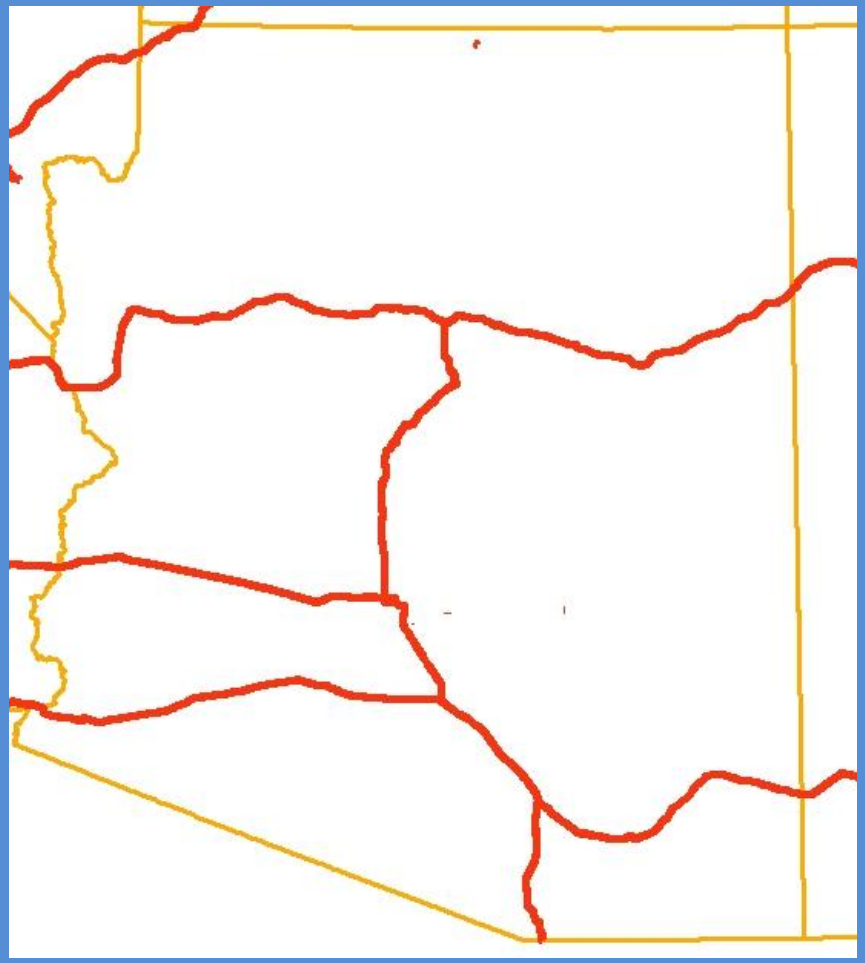
global systems modeling limited

The Proposed Hierarchical Framework



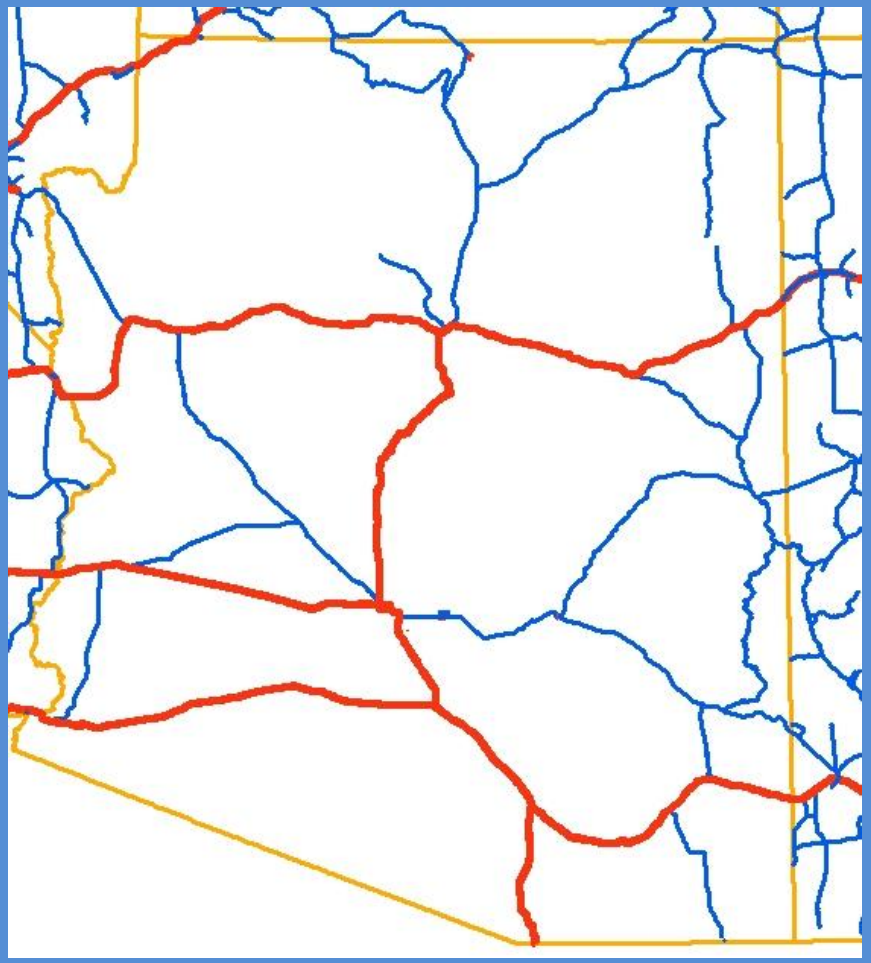
global systems modeling limited

Example: Trans-regional Interstates



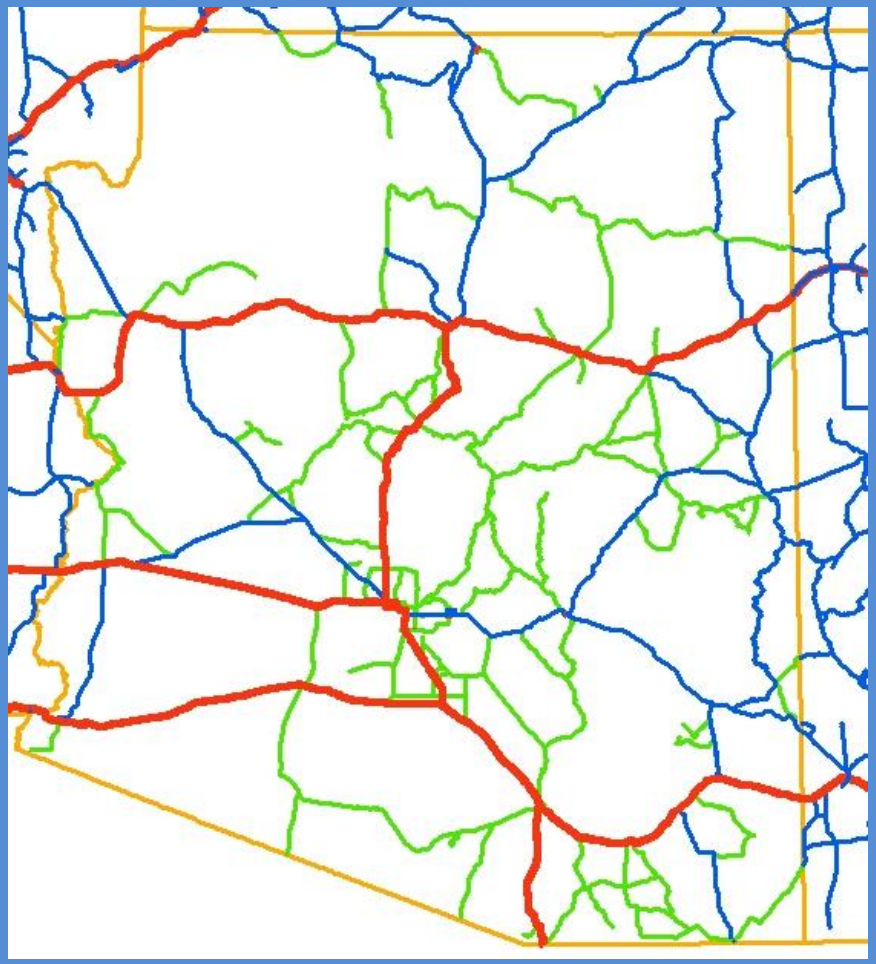
global systems modeling limited

Example: Trans-regional Adding US Routes



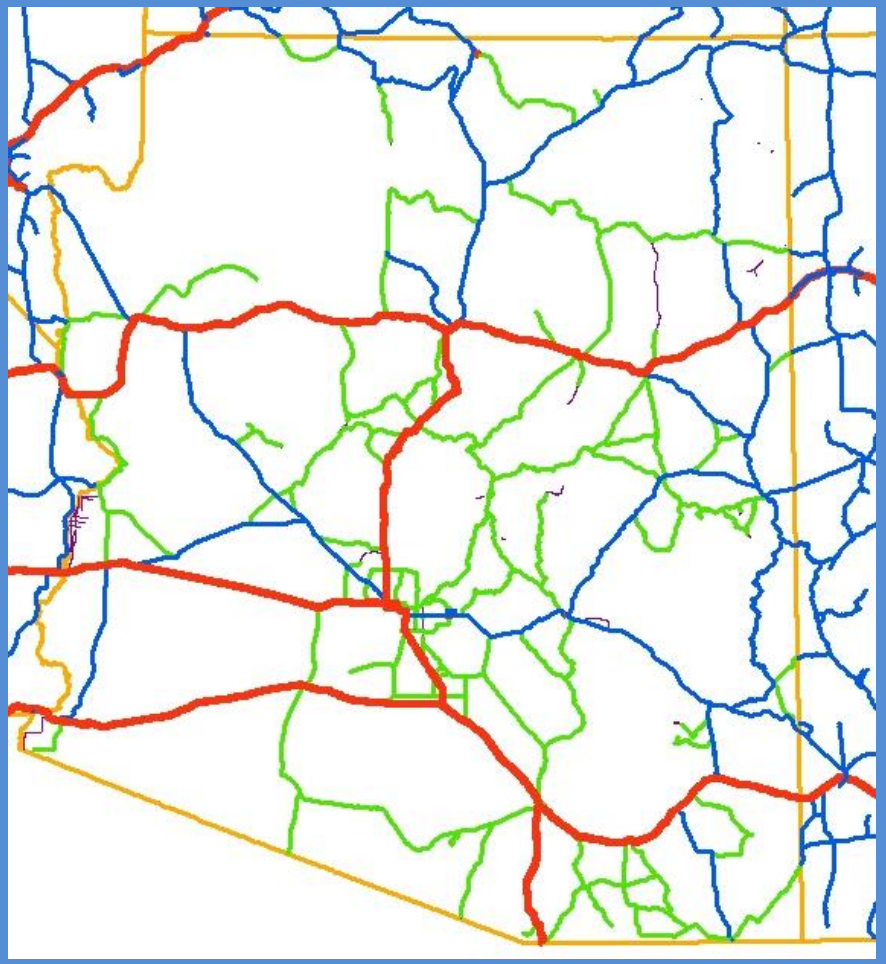
global systems modeling limited

Example: Regional
Adding Arizona (State) Routes



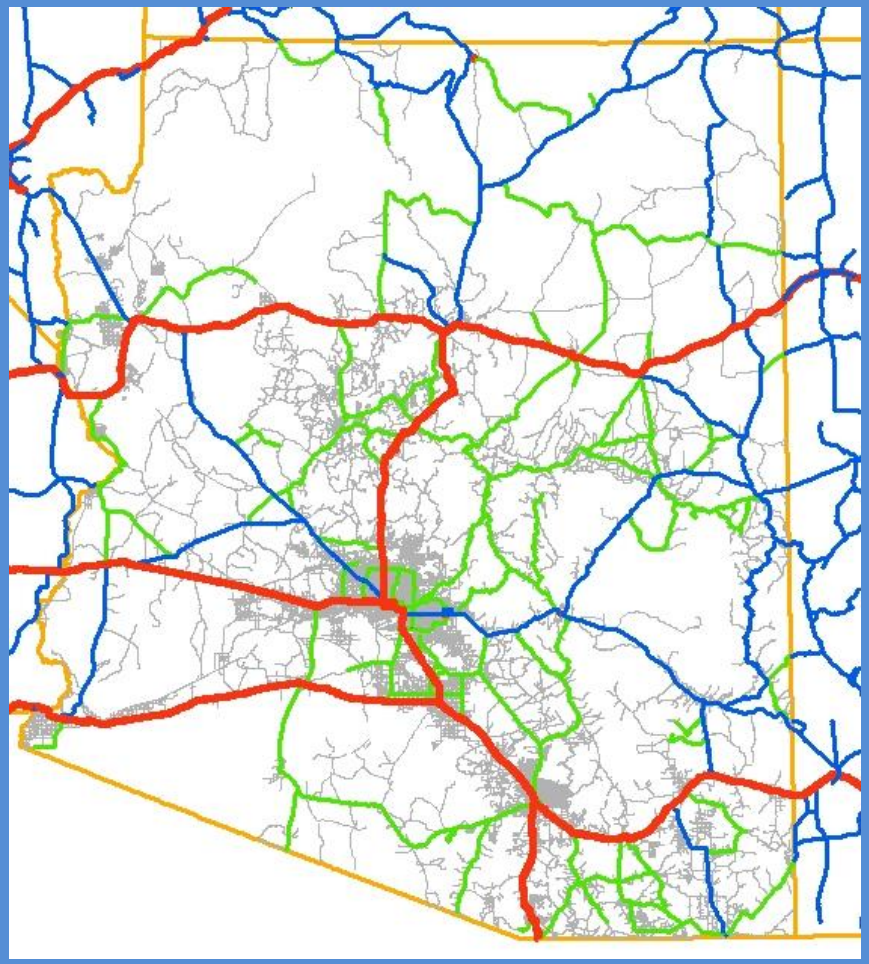
global systems modeling limited

Example: Regional
Comparing ADOT & Census TIGER/Line



global systems modeling limited

Example: Addressable



global systems modeling limited

Problems with the Proposed Solution

There are many "squishy" aspects to the proposed hierarchical framework approach:

- Level names and scale ranges are indicative not definitive.
- Within each level, there are multiple representations possible, ranging between "coarse" and "fine" resolution, depending on application needs.
- There are no definitive standards regarding how a spatial feature needs to be represented at the finest ("Atomic") level. Among many questions:
  - How many vertices are required for a "straight" road segment?
  - How many vertices are required for a "curved" road segment?
  - How do we handle differences in curve representations in different data formats, e.g. parametric curves in File Geodatabases vs. non-parametric curves in Shapefiles?
  - How do we handle complex (multi-lane) intersections?

Standards

Arizona 9-1-1 Office:

- Added requirement for including driveways of 100 feet or longer.
- Modified accuracy standard for road centerlines so that they must fall within the “roadway” as judged by NAIP 2010 ortho imagery.
- Added an attribute field to the “parsed” street name elements.

NENA (National Emergency Number Association) - numerous

FGDC (Federal Geographic Data Committee) - numerous

FHWA (Federal Highway Administration)

- HPMS (Highway Performance Monitoring System)

Creating the “Atomic” Level

The most complex and time-consuming tasks are:

- Creation of the road elements at the “Atomic” level
- Creation of the associations between road elements at one level and those at the next higher level which allow us to “compose” coarser resolution features from finer resolution features.

We believe that creating specific editing and analysis tools can greatly accelerate this process.

global systems modeling limited

Composing Higher Levels

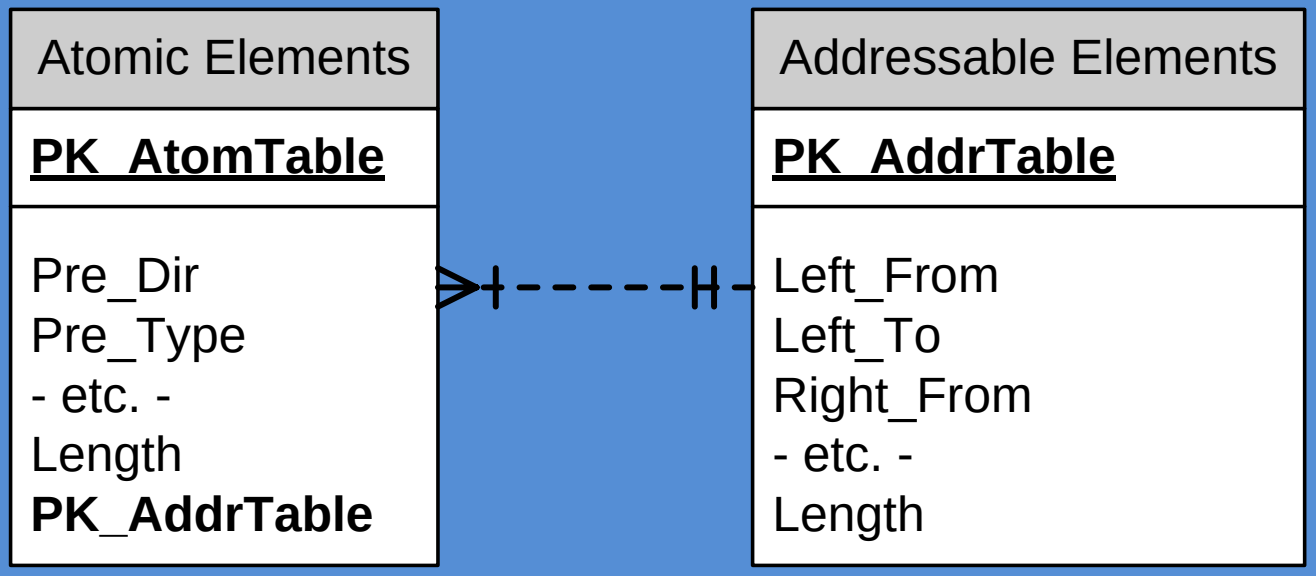
Based on the Relational Database requirement for Primary Keys (unique identifiers) in each data set.

Does not use Foreign Keys in any data set.

Instead, we create separate association tables for each composition requirement. This follows Object Oriented Programming philosophy of "separation of responsibilities". It also allows us to maintain, if desired, attributes related to the association itself, which are completely separate from the attributes of the participating (associated) objects.

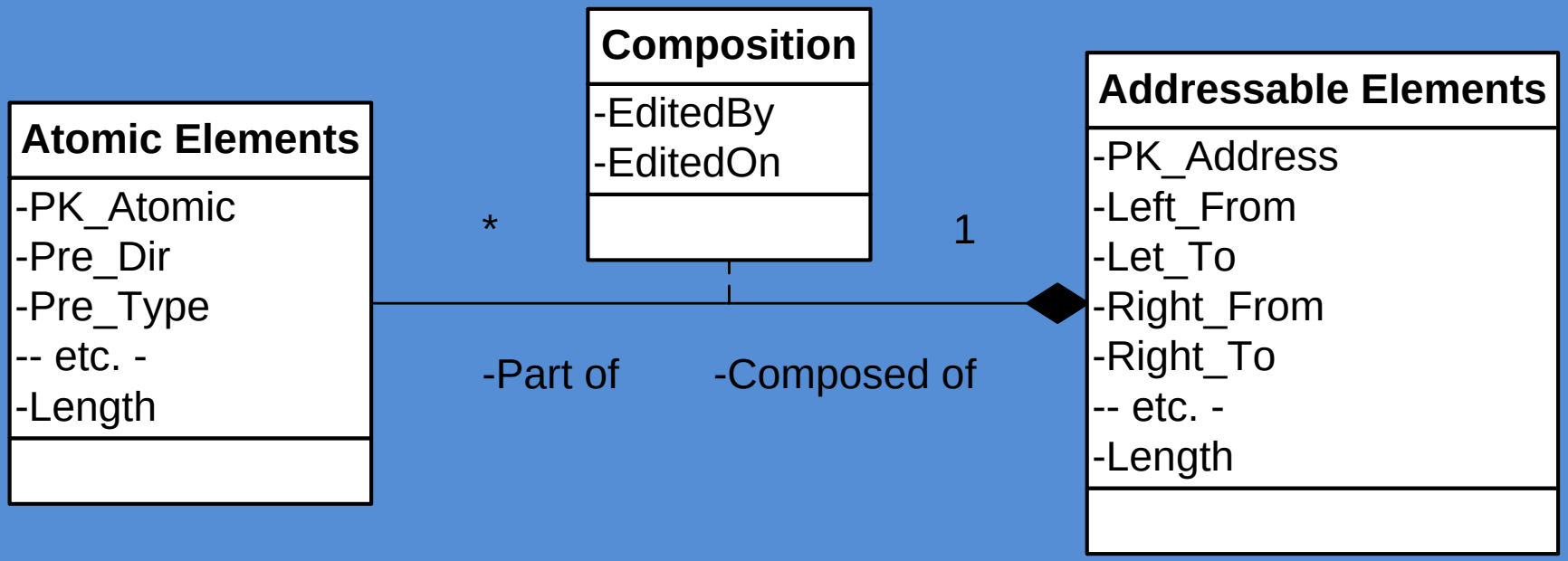
Composing Higher Levels

If we allowed Foreign Keys, then we'd have to modify the Atomic Level table structure every time we created a new composition of elements to higher level representations.



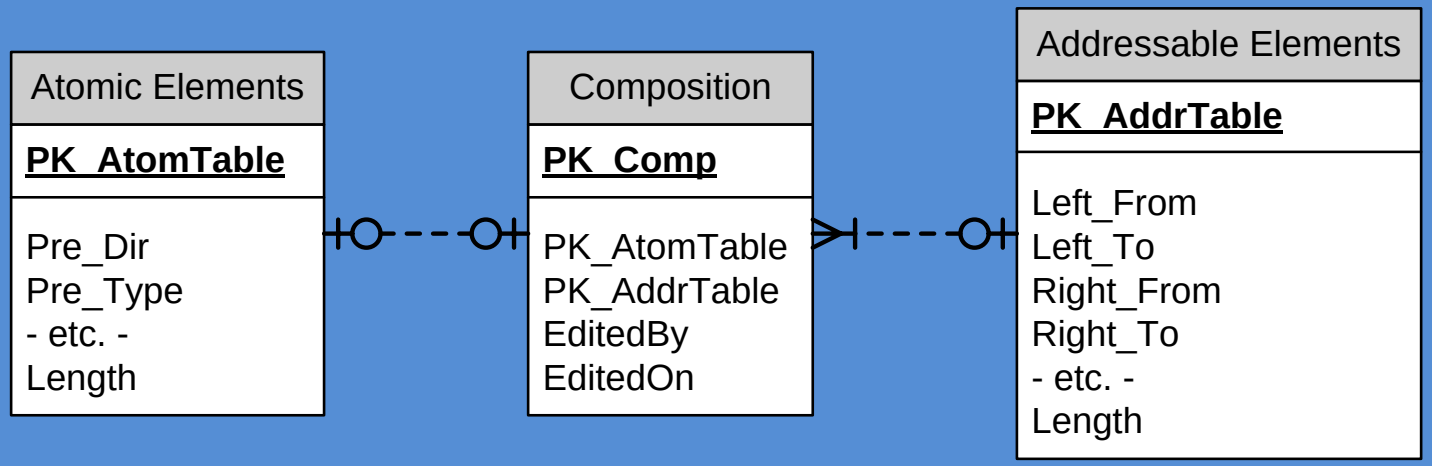
Composing Higher Levels

Object Oriented Programming paradigm dictates a separate association object relating the two other objects.



Composing Higher Levels

The new association Object becomes an association Table in the RDBMS. It stands between the two road network feature class tables. Each composed road network has its own, separate composition table. Changes to one of them has no effect on any others, nor on the Atomic Elements table.



Creating the "Atomic" Level

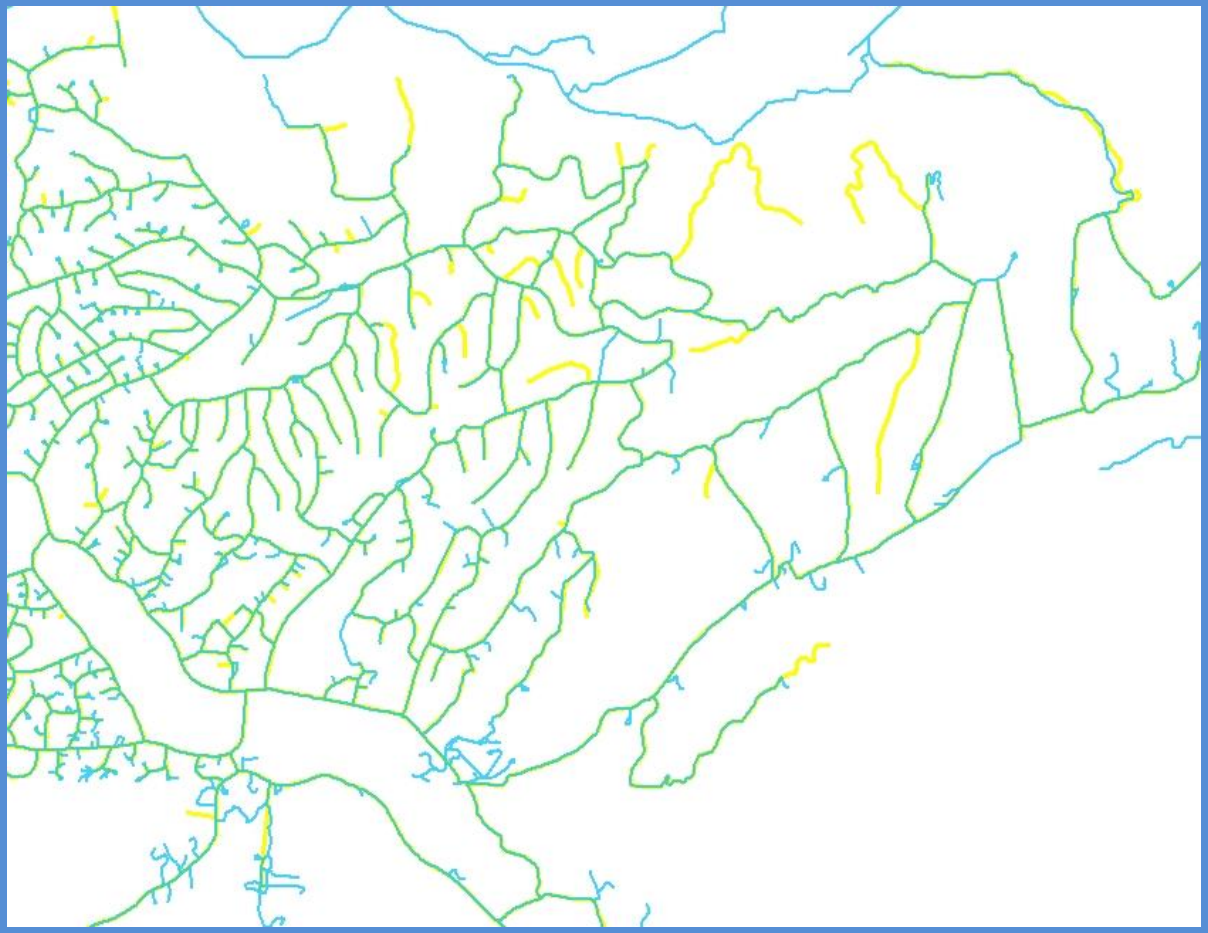
Two data sources are the most likely candidates for use in creating the "Atomic" level:

- US Census Bureau's TIGER/Line "Edge" data that is tagged as road features
- Local data sets from each E9-1-1 participating agency

These data sets have a comparable level of "fine" resolution. We wish to identify the "best" element from one of the data sets and use that to populate the "Atomic" level feature class. Our criteria for "best" is based on E9-1-1 standards (spatial accuracy, driveways of 100 feet or greater length).

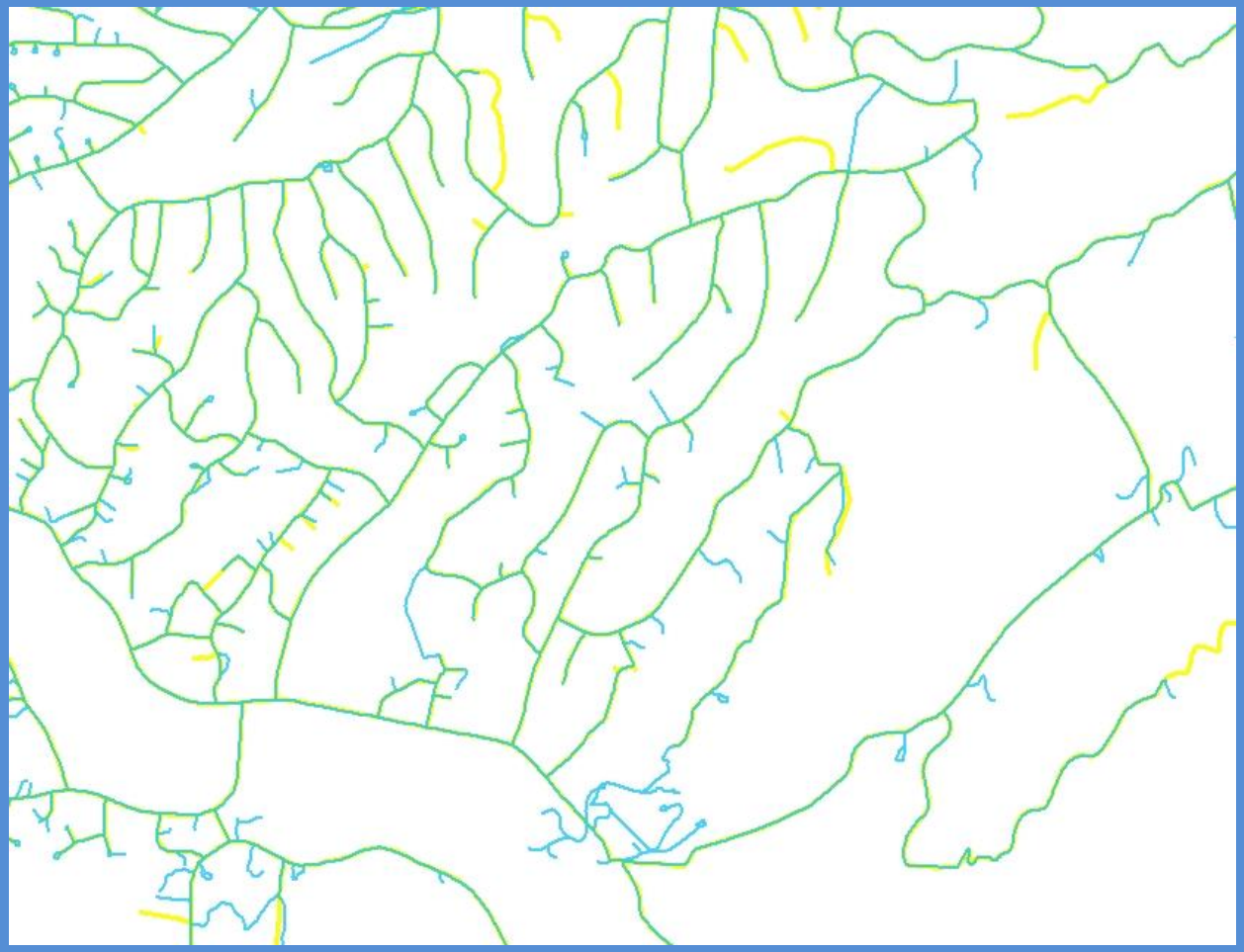
Creating the “Atomic” Level

Here we overlay sample data from Santa Cruz County and TIGER/Line data for that County for 2013.



Creating the “Atomic” Level

TIGER data is in blue, SCC data is in yellow; lines that appear green show a close match spatially between road elements.



global system modeling limited

Creating the "Atomic" Level

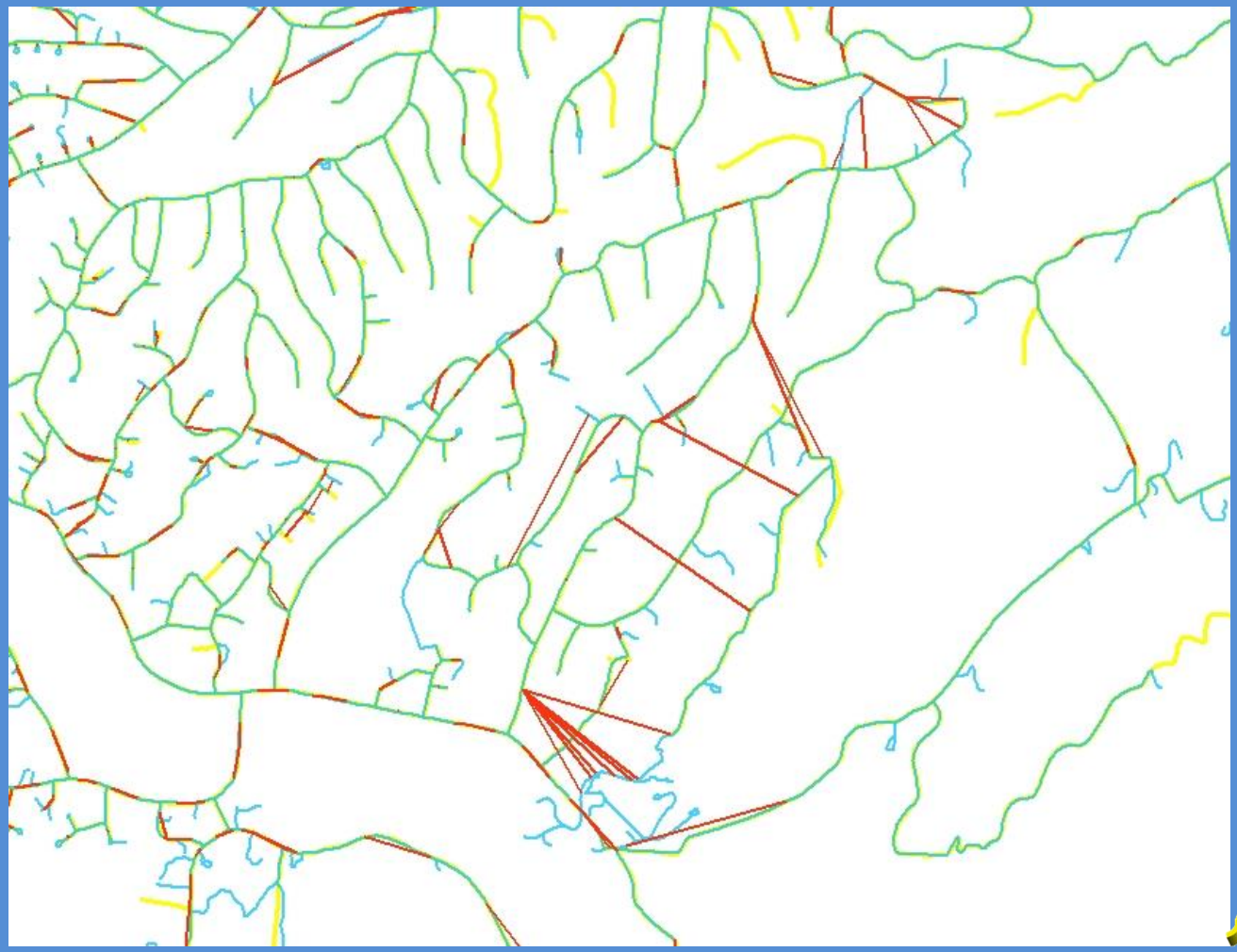
The standard way of creating an association table in ArcGIS is to use the Relationship Class. The biggest drawback to the Relationship Class is that we can't visualize it.

Realizing that associations are basically binary relations (an object related to only one other object) allows us to implement an association table as a standard Polyline feature class. The line will represent the association.

We use the mid-point of a road element to represent that road element. Then by drawing a line between two mid-points we have created an association that we can see!

Creating the “Atomic” Level

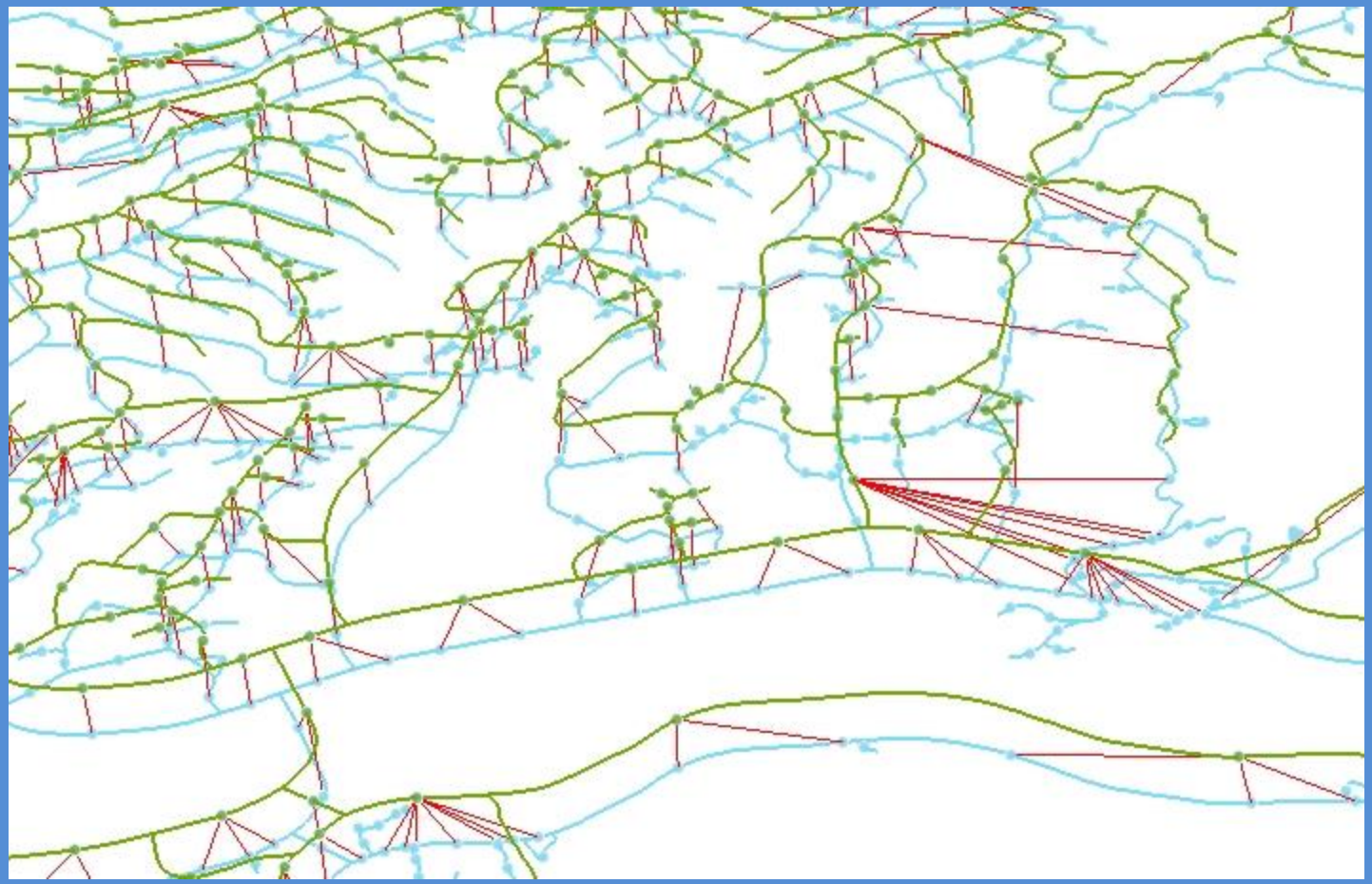
Doing this in 2D is only shows us gross mistakes or one-to-many relationships, but it doesn’t show “identical” relationships very well.



global system modeling limited

Creating the “Atomic” Level

Doing this in 3D gives us a better view of all the relationships.



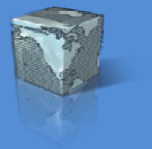
global systems modeling limited

Questions?

Howard Ward – TerraSystems Southwest, Inc.
hlward@terrasw.com

Curtis White – Global Systems Modeling Ltd.
cwhite@globalsysmod.com

TSSW



global systems modeling limited