Whitepaper: GIS Data for NG911

WHAT'S INSIDE

The quality of geographic information system (GIS) data arguably is the most important factor in the effectiveness of Next Generation 911 systems. Low-quality data will result in a high number of misrouted 911 calls and an inability to dispatch the appropriate response to an emergency incident in a timely manner.

However, developing high-quality GIS data is a time-consuming, complicated and costly process, which is why some agencies opt for quick-fix solutions. This whitepaper explains why doing so is a bad idea.

Don't Take Shortcuts When Developing GIS Data for NG911

Geographic information system (GIS) data is a foundational component in the migration to, and continuing operation of, Next Generation 911 (NG911) systems. But developing local GIS data so that it aligns with NG911 standards is a laborious and time-consuming process that can take months or years to complete.

The data-development process typically requires the coordination of several GIS professionals across multiple jurisdictions—they may not work for public safety answering points (PSAPs). Once local GIS data is developed, all datasets must be validated with neighboring jurisdictions to ensure that there are no overlaps or gaps within each of the five required datasets:

- Site/structure address points
- Road centerlines
- PSAP (i.e., call-routing) boundary
- Emergency service boundary—i.e., emergency medical services (EMS), fire/rescue, and law enforcement
- Provisioning boundary

It is imperative that the GIS data used to route 911 calls be spatially and topologically accurate and fully attributed to a standard database schema. The highest-quality GIS data is sourced directly from the local jurisdiction. Statewide NG911 deployments depend on local GIS capabilities for developing and maintaining these datasets to NG911 standards. The lengthy development time coupled with varying GIS capabilities in a local jurisdiction can delay the entire NG911 migration process. While alternate sources exist for acquiring the initial GIS data, none of these sources can support the complexities required within the NG911 call-routing process.

The National Emergency Number Association (NENA) has developed, and continues to develop, standards and best practices regarding creation and maintenance of GIS data to support NG911 service. These standards are meant to increase the likelihood of a positive outcome for the 911 caller by increasing call-routing accuracy, reducing call transfers, and providing actionable data with the delivery of the voice call to the PSAP, thus reducing response time to the incident.

Deploying substandard GIS data into the NG911 ecosystem contradicts the goals of NG911 by introducing undocumented errors into the location databases, or omitting locations from them. This potentially causes a 911 call to route incorrectly, thus significantly increasing response time to the population that NG911 is designed to assist.

MissionCriticalPartners

The reality of time, effort, and cost can be overwhelming and may lead a 911 entity to seek a quick fix. However, sacrificing long-term viability for short-term gains creates a situation that is contrary to the reasons for investing in NG911. Moreover, a quick-fix solution can cause the community to lose confidence in the 911 system, because performance is no better, or even worse than, the legacy solution, and would require additional cost to implement corrective solutions.

You Get What You Pay For

Some states are considering, or even promoting, the use of open-source, U.S. Census Bureau Topologically Integrated Geographic Encoding and Referencing (TIGER) data for 911 call routing. The old adage, "you get what you pay for," holds true for TIGER data. Free, nationwide GIS data is a very appealing option. However, the GIS community recognizes, as does the Bureau, that the TIGER data does not meet basic public safety requirements, nor the established NG911 standards, due to incomplete data attribution, poor spatial accuracy, incomplete coverage of the PSAP's jurisdictional footprint, inaccurate street names and address ranges, and a lagging data update schedule.

The Bureau's mission is to serve as the nation's leading provider of quality data about its people and economy.[1] To achieve this mission, the Bureau created a nationwide master address file (MAF) originally based on the U.S. Postal Service (USPS) delivery sequence file of all residential addresses. The MAF was coupled with the TIGER spatial database, and in 2010 the TIGER/Line Shapefiles GIS dataset was created.

The GIS data used by the Bureau needs only to be accurate enough to serve its mission. The expense of creating and maintaining highly accurate GIS data on a nationwide scale is cost prohibitive for the Bureau based on its needs. So, the fiscally responsible data development-and-maintenance plan for the Bureau is to only achieve geospatial accuracy sufficient to support its mission.

Until 2018, the Bureau claimed a spatial accuracy for TIGER line files at plus or minus 25 feet. Upon completing the 2018 migration of the TIGER files from ASCII[2] fixed tables—think automatic location identification (ALI)-type data—to GIS shapefiles, the Bureau removed all claims to an accuracy standard. Further, the 2018 revision of the TIGER Technical Documentation replaced such claims with the following official statements:[3]

> No warranty, expressed or implied, is made with regard to the accuracy of the data in the TIGER/Line Shapefiles, and no liability is assumed by the United States Government in general, or the Census Bureau specifically, as to the positional or attribute accuracy of the data. (Pg. 1–2)
>
> In order to maintain a current geographic database from which to extract the TIGER/Line Shapefiles, the Census Bureau uses various internal and external processes to update the MAF/TIGER database. While it has made a reasonable and systematic attempt to gather the most recent information available about the features each file portrays, the Census Bureau cautions users that the files are no more complete than the source documents used in their compilation, the vintage of those source documents, and the translation of the information on those source documents. (Pg. 2–8)
>
> The Census Bureau cannot specify the spatial accuracy of features changed or added by its field staff or through local updates, features derived from the GBF/DIME Files (TIGER's predecessor in 1970 and 1980), or other map or digital sources. Thus, the level of spatial accuracy in the TIGER/Line Shapefiles makes them

"No warranty, expressed or implied, is made with regard to the accuracy of the data in the TIGER/Line Shapefiles, and no liability is assumed by the United States Government in general, or the Census Bureau specifically, as to the positional or attribute accuracy of the data."

TIGER Technical Documentation, 2018 Revision

[1] https://www.census.gov/about/what/census-at-a-glance.html

[2] American Standard Code for Information Exchange

[3] https://www2.census.gov/geo/pdfs/maps-data/data/tiger/tgrshp2018/TGRSHP2018_TechDoc.pdf

> unsuitable for high-precision measurement applications such as engineering problems, property transfers, or other uses that might require highly accurate measurements of the earth's surface. (Pg. 2–8)
>
> Some street features may have a misclassified MTFCC. This means that there could be gaps in features in the primary roads or the primary and secondary roads shapefiles, if a segment of the feature was misclassified as an S1400 (a local neighborhood road, rural road, or city street) instead of an S1100 or S1200. (Pg. 3–41)

When asked by MCP about the use of TIGER data to support 911 service, the Bureau stated:

> "… we do not release an address file, but only an address range file. The term 'address range' refers to the collection of all possible structure numbers from the first structure number to the last structure number, including all numbers of a specified parity in between. Address range overlaps, gaps, odd/even reversals, and low-high orientation reversals may exist in the data. With the exception of overlaps, these may be valid. Further, we do not have address ranges for the entire country, but focus on residential areas. Therefore, many commercial areas are not covered.
>
> While the Census Bureau provides this free data, we cannot provide guidance on whether the data meet the needs of an organization or how they use the data."

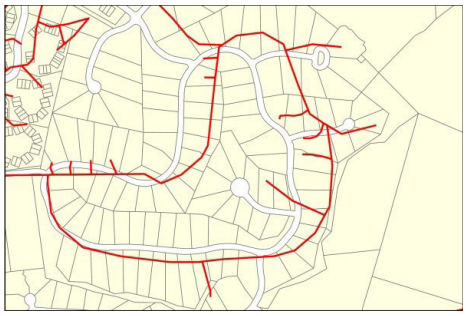


FIGURE 1: Positional Inaccuracy of TIGER GIS Shapefiles

Why Tiger Data Doesn't Work

In its official documentation that accompanies the TIGER/Line Shapefile data, the Bureau states multiple times that the TIGER data is not of good quality. Figure 1 provides a random sample of TIGER road centerlines overlaid on county-maintained parcel data in a suburban area.

Some of the problems that have plagued the TIGER data files include:

- Poor positional accuracy—up to three miles for jurisdictional boundaries
- Gaps and overlaps in line segments
- Duplicate and missing features
- Missing attribution

A 2014 study by the University of Southern California GIS Research Laboratory using verified addresses in California, found that the GeoLytics geocoding solution built using the Bureau's TIGER data was only 8 percent accurate at locating addresses to the parcel or building level.[4] Forty-six percent of the locations in the sample dataset did not geocode (i.e., find a location match) in any category. General location—i.e., city center or ZIP code–was provided for 52 percent of the tested records. Applying these findings to 911 service logically results in a scenario whereby the vast majority of 911 calls are default-routed based on the policy routing rules defined within the NG911 system. With only eight percent of addresses being located using the Bureau's data, as many as 92 percent of emergency calls are at risk of being misrouted.

Several studies from 2003 through 2017 provide similar results. TIGER data for the study areas has an average horizontal positional accuracy of 167 feet. Rural geocoding within the study areas averaged a horizontal positional error factor of nearly two miles. Urban accuracy was far better, but still averaged a horizontal positional error factor of 498 feet. The use of TIGER data for NG911 call routing most likely will cause a regression in location information for the majority of 911 calls.

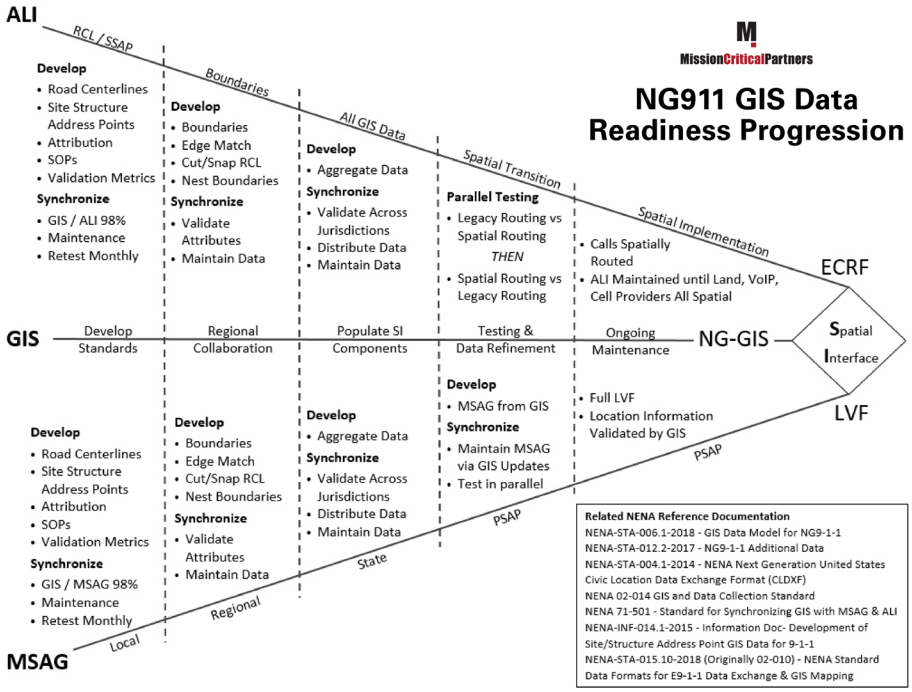
[4] https://spatial.usc.edu/wp-content/uploads/2014/03/gislabtr10.pdf

Development of the NG911-required polygons typically involves extensive local knowledge and research of local response areas to ensure that the appropriate emergency response is sent to an incident. Available TIGER boundary data does not include the datasets required for NG911, and any TIGER boundary files used likely would require significant manipulation and verification before being used in an NG911 solution. Further, a geospatial database that includes a mix of locally developed data and TIGER data more often than not would require considerable editing at boundary lines to ensure topological accuracy between jurisdictions.

Available TIGER boundary data does not include the datasets required for NG911, and any TIGER boundary files used likely would require significant manipulation and verification before being used in an NG911 solution.

It is necessary to coordinate the development of all five required NG911 GIS datasets that will be used in the Next Generation Core Services (NGCS) components of the NG911 system. Emergency service boundary polygons must correspond with the appropriate locally defined boundaries and attributes. Adjacent PSAP and provisioning polygons must have coincident edges without gaps or overlaps. Road centerlines must break at the boundary of each polygon and snap to the road centerlines in neighboring PSAPs.

Poor data quality within one of those datasets will lead to degraded quality in the remaining four. Adjacent jurisdictions creating and maintaining GIS data using industry-standard data-development methodologies will create higher-quality data that will not correspond to the Bureau's TIGER data used by neighboring jurisdictions, which contains inferior attribute and spatial information. There is no substitute for the necessary investment in the development and constant maintenance of local GIS data to support NG911 service. The process to develop GIS data for use in the NG911 environment is long and arduous. However, providing quality data for the NG911 ecosystem will enhance each PSAP's life-saving mission by improving call routing and reducing emergency response times.



Help Wanted

GIS professionals cannot prepare for the migration to NG911 without direction from, and the assistance of, the 911 community. GIS road centerlines and address points must be improved in tandem with legacy 911 data in the ALI and master street address guide (MSAG) tables. Once the quality of GIS data exceeds the accuracy of the legacy data tables for routing 911 calls, the GIS data should be put into service as the primary call-routing engine.

The migration to NG911 cannot be rushed and still be successful in improving location accuracy and accurate call routing. The best, i.e., most accurate, GIS data is sourced from the local jurisdictions. The promotion and funding of local data development by national, state and local public safety leaders is critical to the success of local GIS solutions and the implementation of NG911 systems.

It is not necessary for jurisdictions just beginning a GIS program to start from scratch. Many sources exist for developing or improving GIS data. The United States Geological Survey (USGS) is a great resource for GIS data. Road centerlines may be digitized from USGS orthoimagery. Road centerline data also is available—the same data used in vehicle Global Positioning System (GPS) navigation systems— from commercial sources. Grant money is available every year from multiple federal sources for developing and improving the GIS data necessary for NG911. It is imperative to retain the focus on improving emergency response when choosing a GIS data-development solution to support NG911 service.

The use of TIGER data for NG911 call routing most likely will cause a regression in location information for the majority of 911 calls.

Conclusion

It is imperative that the GIS data used to route 911 calls be spatially and topologically accurate and fully attributed to a standard database schema. The highest-quality GIS data is sourced directly from the local jurisdiction. To this end, NENA has developed, and continues to develop, standards and best practices regarding creation and maintenance of GIS data to support NG911 service.

Deploying substandard GIS data into the NG911 ecosystem contradicts the goals of NG911 by introducing undocumented errors into the location databases, or omitting locations from them. Nevertheless, due to the time, effort, and cost of developing high-quality GIS data, public safety agencies may be tempted to adopt quick-fix solutions, such as the open-source TIGER data that is available from the U.S. Census Bureau free of charge. However, this should be avoided because the TIGER data *does not* meet basic public safety requirements, nor the established NG911 standards.

There is no substitute for the necessary investment in the development and constant maintenance of local GIS data to support NG911 service, because doing so will enhance each PSAP's life-saving mission by improving call routing and reducing emergency response times.

Mission Critical Partners
690 Gray's Woods Blvd.
Port Matilda, PA 16870
Phone: 888.8.MCP.911 or 888.862.7911
Fax: 814.217.6807
MissionCriticalPartners.com

May 2020 | Doc # NTWKSWP040520

MissionCriticalPartners