

# Georgia Broadband Mapping

---

## Data Submission Methodology Report

April 1, 2013



1935 Jamboree Dr.  
Colorado Springs, CO 80920



**Georgia Technology Authority**  
47 Trinity Avenue S.W.  
Atlanta, GA 30334-3404

# Georgia Broadband Mapping

## Data Submission Report (April 1, 2013)

### Table of Contents

<b>1 INTRODUCTION .....</b>	<b>3</b>
<b>2 OVERALL PROJECT STATUS.....</b>	<b>4</b>
2.1 DATA COLLECTION .....	4
2.1.1 <i>Broadband Data</i> .....	4
2.1.2 <i>Community Anchor Institutions Data</i> .....	8
2.2 DATA PROCESSING .....	8
2.3 DETAILED DATA PROCESSING.....	10
2.3.1 <i>General Overview</i> .....	10
2.3.2 <i>Initial State Set-Up Phase</i> .....	14
2.3.3 <i>Provider Processing Phase</i> .....	17
Address Data Processing.....	17
Street Segment Data Processing .....	18
Census Block Data Processing .....	19
Wireless Service Area Availability Data Processing.....	20
Infrastructure Data Processing .....	21
2.3.4 <i>Integration Phase</i> .....	24
Append.....	24
Aggregate.....	24
Deliver to NTIA and Publish to Web Applications .....	25
2.4 DATA VALIDATION.....	25
<b>3 SUBMISSION 6: NTIA SUBMISSION DATA MODEL SCHEMA CHANGES.....</b>	<b>27</b>
3.1 SCHEMA HISTORY AND EVOLUTION.....	27
3.2 NTIA DATA MODEL CHANGES SUBMISSION 2 .....	28
3.3 NTIA DATA MODEL CHANGES SUBMISSION 3.....	29
3.4 NTIA DATA MODEL CHANGES SUBMISSION 4.....	29
3.5 NTIA DATA MODEL CHANGES SUBMISSION 5 & 6.....	29
<b>4 UNIVERSE OF CONTACTED PROVIDERS/NON-PROVIDERS.....</b>	<b>30</b>
4.1 NON-PROVIDERS .....	30
4.2 SHELL COMPANIES.....	32
4.3 RESELLERS .....	35
4.4 NON-PARTICIPATING OR NON-RESPONSIVE PROVIDERS.....	36
4.5 PROVIDERS WITH NO DATA UPDATES .....	37
4.6 PROVIDERS WITH DATA UPDATES .....	39

# 1 Introduction

This report is submitted along with the seventh data submission for the Georgia Broadband Mapping Project. This submission includes all data collected so far per the requirements of the National Telecommunications and Information Administration (NTIA) State Broadband Data and Development Grant Program (Docket No. 0660-ZA29) Notice of Funds Availability (NOFA) and Clarifications to it. Specifically, it includes broadband data collected from broadband providers and Community Anchor Institutions data compiled from various sources for the State of Georgia. The State of Georgia has retained a mapping contractor, primed by the Sanborn Map Company (Sanborn), and supported by Applied Geographics Inc. (AppGeo) to support the Georgia Technology Authority (GTA) related to the Mapping Grant for this project.

This document provides an overview of the Team's progress, processes, assumptions, challenges and improvements needed for each dataset. Section 2 of this document explains overall project status at the time of this submission. This is followed by detailed description of the processes of data collection, data processing, and data validation. We have included with this report a section on detailed data processing for this submission. In the next section, the report documents all changes to the data model so far. The final section provides a list of providers that fall in various categories of participation.

## 2 Overall Project Status

Sanborn was selected as the new Mapping Contractor for State of GA and contract was signed on January 14, 2013. After the kick-off meeting between Sanborn and GTA (held January 23<sup>rd</sup>, 2013), work on the program commenced and letters to providers were sent on February 1<sup>st</sup>, 2013.

In this submission, the following high level statistics represent the degree of participation and data updates from existing broadband providers and newly added providers. The breakdown of these providers in different categories is summarized below and detailed lists are provided in the last section of this report (Section 2.6).

<b>Provider - Data Status - Submission 7</b>	<b>Count</b>
Total Providers Researched/Contacted	203
Non-providers	45
Resellers	16
Total Valid Providers (total participating + non-cooperative but valid providers)	96
Non Responsive Providers	5
Non-cooperative Providers (refused participation)	11
Number of DBAs Represented in Data (not including the 5 companies that had no FRNs)	89
Number of Providers that Supplied Updates	47
Number of Providers – Confirmed No Updates	15
Number of Providers – No updates and no confirmation	22

### 2.1 DATA COLLECTION

This section details data collection related to NTIA deliverables which include broadband data and community anchor institution data.

#### 2.1.1 Broadband Data

Sanborn acquired the most recent NTIA submission data (i.e. Submission 6 submitted to NTIA on October 1, 2012). We moved all the S6 data onto our provider portal and created usernames and passwords for each participating provider so that they could make edits to their service area. Sanborn and GTA held a web conference call to introduce the new team to the Georgia broadband providers and provide online training to them as well. Sanborn also held numerous technical one-on-one calls/training sessions with providers and also explained the program requirements and the type of data we needed them to provide us.

We also created a Technical Requirements document based on the NOFA requirements and changes that have been made to date. This document included specific details on data formats, alternative formats, schema, data

delivery, etc. and was accessible through the Provider Portal and also provided to Broadband Providers as needed.

The Sanborn team then followed up by making calls to all providers identified (including participating providers, those who have refused participation, resellers, and newly identified providers). Sanborn began with the FCC Form 477 Broadband Data Filers as of June 30, 2011. This list had 130 providers for the State of Georgia. During the first year of the program, when Sanborn was the mapping vendor, we already had valid confidentiality agreements with many providers that did not have any end date. New confidentiality agreements were executed with companies that were new in this submission or wanted new NDAs new confidentiality agreements were executed although some companies chose to provide us data without an NDA. We have a total of 86 executed NDAs for the State of Georgia.

According to both our research and lists provided to use by NTIA, there was the potential for Georgia to have up to 203 broadband providers. The breakdown of these providers in different categories is summarized below and detailed lists are provided in the last section of this report (Section 2.6).

Of these:

- 45 companies stated that they do not provide any type of broadband service in Georgia. Many of these are either national carriers without a Georgia presence, or they file 477 reports because they provide VOIP or Video Teleconference services (but not broadband).
- 42 company names turned out to be a DBA or legal holding names for another firm that is listed in another category. So these duplicates were dropped from our list.
- 16 companies are resellers and are not considered part of this program.
- 16 companies may be broadband providers, but they have either indicated they are not willing to provide data, or were completely unresponsive to multiple attempts to contact them.
- 37 broadband providers informed us that there were no changes to their service area so for these providers we downloaded the data that had been submitted as part of the last submission, reprocessed the data, performed validation on it, and we are resubmitting this data.
- 47 Broadband providers submitted either entirely new or partially new datasets for this submission.
- 1 provider submitted Linear Middle Mile data that is not being submitted to NTIA (this provider will make the count of total providers to 204 for GA)

A total of eight new providers have been added to this submission:

- MetroPCS Georgia, LLC
- TruVista
- Northland Cable Properties Eight Limited Partnership

- Trenton Telephone Company
- Public Service Data Wireless
- FiberLight
- Dalton
- Appalachian Valley Fiber Network

All changes and corrections in provider data are documented in the Change and Correction Document submitted with this submission.

In general, the submission 7 processes followed the same basic approach that was used in previous submissions made by Georgia. This document provides further details on the following topics:

1. Submission 7 Processing Assumptions
2. Reference Data Creation
3. Processing of New Provider Data
4. Quality Control Checks
5. Improved Validation Techniques
6. NTIA Quality Control Scripts
7. NTIA Submission Data Model Schema Changes

Based on NTIA feedback and information provided in NTIA webinar sessions, the submission 7 data processing workflow is created with the following assumptions to meet NTIA submission requirements.

1. All census blocks and road segments are mapped based on 2010 census data set.
2. Due to our NDA restrictions, address points and last mile points are not being submitted to NTIA.
3. Terrestrial Mobile Wireless and Terrestrial Fixed Wireless (licensed and unlicensed) were again treated as wireless coverage and were delivered as a shape. In cases where a provider served the same spectrum with different speeds, overlapping areas were removed and the higher speed was assigned. The exception to this rule is where a provider is using the same spectrum, but delivering different underlying technologies such as 3G, 4G, or 4G LTE. In this case a continuous polygon is being created that represents the area that is offered for both 3G and 4G even if these polygons overlap.
4. If a cable based wireline provider can provide both DOCSIS 2.0 and DOCSIS 3.0 service to the same area, the block or road was listed only once, with a technology code of 40.
5. Providers were only willing to indicate on a general level if they serviced business, residential or both. None of the providers broke down their type of service by block. Where providers stated or we knew from local sources what kind of end users the provider served, we filled in the "category of end user" with a code of 1 or 2. For the remainder, we made an assumption of 5.
6. The submission 7 Provider data model is currently based on the NTIA December 2012 data package.

7. The weighted average speed alternate format which was originally part of the NOFA was not provided per NTIA's direction.
8. Wireless coverage was provided with unique shapes for each spectrum utilized.
9. If provided, Franchise Area was captured during the ingest process, and kept for use during the validation process. These areas are not being submitted to NTIA.
10. All Provider data and Community Anchor Institutions (CAI) locations were clipped to the state's boundary.
11. Records dropped during data processing will have an associated reason code. Dropped records were maintained in a separate similarly formatted dataset and given to the providers so they had an opportunity to correct any issues. Records without required attributes were not submitted to NTIA.
- 12. In this submission, for landline broadband, we removed blocks and roads that are in water only census blocks. We communicated this to providers at the beginning of data collection to make sure they would let us know if they really served on blocks that were water only and no population (i.e. situations of docks or houseboats or other unforeseen situations).**
- 13. It was not clear to Sanborn what criteria were used by previous vendor to identify blocks that were less than 2 square miles (census published numbers vs. calculated areas from feature geometry). Sanborn has always used the census published numbers for the area cut-off by adding land and water area. This created some changes in the data where in some cases roads needed to be submitted instead of blocks and vice-versa.**
14. Where providers told us to reuse data from the previous submission or did not respond to our data request, we are resubmitting data that was submitted in S6. Because of the removal of water polygons and the 2 sq. mile criteria for blocks, we have some differences in counts for roads and blocks even when no data updates were submitted by a provider.
15. In the final stages of processing this submission, we noticed that some providers are delivering street segments that appear to be new roads that have been constructed since the 2010 data was created, but they are not in the official Census 2010 geometry data. These roads were dropped from the submission, but we are going to look into a process to add these roads to the next submission if they can be verified as accurate. Some guidance on this from NTIA may be useful so all states submit the data consistently.
16. In this submission, we also found that some providers were using street segments that collapsed multiple census streets into a single segment. We have used manual processes to select roads in the census data for such providers.

17. This submission is being made based on the NTIA data model as of Dec, 2012 provided by NTIA on the SBDD site.

### 2.1.2 Community Anchor Institutions Data

In this submission, Georgia is supplying a substantially increased CAI dataset. The state has outsourced telecommunications services for internal needs to AT&T. GTA was able to obtain a list of all locations serviced under this contract and the level and type of service provided at each location. This new data provides almost 2000 additional CAI datapoints across 260 customers.

## 2.2 DATA PROCESSING

We started with the following base data:

### Census Blocks:

For Submission 7, Census 2010 data was utilized. The data was set up as follows:

- Block size (AREA) is calculated combining the 2010 land area (ALAND) and water area (AWATER)
- AREA is converted from square meters to square miles to calculate square mileage (SMI).
- If the SMI of a block is less than or equal to 2, then the less than or equal to 2 square mile indicator (LE2SMI) is set to true.

### Road Segments:

2010 Tiger Line IDs (TLID) were used for data processing for this submission. The data was set up as follows:

- The GT2SMI (Greater Than 2 Square Mile) indicator is set to True when:
- The 2010 road segment is completely within a block that is NOT less than 2 square miles
- Only minimum and maximum address ranges and a single zip code for each road segment is maintained.

All data received went through the following processing steps:

**Triage:** All new data was quickly reviewed to understand what was received, and in what format. We also made sure we had all the required components for NTIA's data model, such as their FRN and advertised speed information. We also screened for any known issues that we might have seen before (such as Excel 2003 spreadsheets that cut off at 32k rows.)

**Ingest:** At this time the data is actually brought into our systems. Each provider is set up with a unique file geodatabase to store their information. Record counts of what was received is logged so that we can validate we did not drop anything in processing.

**Data Processing:** This is where the data goes through a number of ETL routines to convert the raw proprietary information into a format similar to the NTIA format. The exact routine utilized depends on how the data is received:

1. When a wireline provider submits a service boundary, we select all the blocks and roads that are within and cross that shape.
2. If a wireline provider submits a customer address list, the points are geocoded, and then the appropriate block or road segment is selected.
3. If a wireline provider submits block and road information using Census data, we make sure everything is formatted to the appropriate specifications
4. If the wireline provider submits any type of road or line data that does not directly correlate to the TIGER data set, we convert the lines to TIGER by selecting the road centroid and spatially selecting the closed segment in our data set. If the road is in a block less than 2sqmi, then the block is selected. Some manual cleanup is also applied to make sure we do not accidentally drop any road segments that should have been processed.
5. Wireless provider data is formatted to ensure that there are no overlapping polygons with the same technology type unless the provider is offering different speeds such as 3G and 4G over the same area. In addition the data is cropped to the state boundary.
6. After each round of processing, we make sure that we only keep unique records. A unique record is defined as having a one of a kind combination of FRN, Block/Road ID, and technology type. If there are multiple records with different speeds, but all else is equal, then we select the maximum of the advertised speeds.

**QC Review:** All data is then sent to a different analyst to perform a thorough quality control review on the processed data set. Record counts are compared to what was submitted. The QC staff also makes sure the ETL scripts and routines populated all of the right fields.

**QC Change Detection Review:** Data is then sent to another team for a second Quality Control Review. In this step the data is not only double checked against what was originally submitted, but it also brought up inside standardized MXD templates that allow us to make sure our results make sense. This step involves comparing the new data set with prior submissions, developing change maps, and looking for any possible technology or speed anomalies. At this stage we also begin our validation process. This includes looking at the provider data in comparison to things such as speed test results, franchise boundaries, siting information, and feedback from the planning surveys.

**Provider Review:** Processed data is all posted to a customized web application we refer to as our Provider Portal. All providers were notified once their data was available in the site, and they were always given time to review the data and respond. In this site, providers can log on and visually see their processed data in a map format. It also allows them to overlay their raw data (boundaries or addresses provide in any format) to help them validate that we did indeed process things correctly. The provider portal also has a suite of markup tools that will allow the providers to edit their data, including adding or removing service areas, and making changes to the data attributes.

**Comment Processing:** All comments and feedback received from the provider portal, is then reviewed and applied to the processed data set. This updated data set goes back through our QA and QC processes, and if time allows, back out to the Provider Portal, for the provider to review and sign off on.

**Data Append:** After all of the individual data sets are processed and approved, we run an append process which merges all of the individual provider data sets into one geodatabase. This is also the point where our team will do any final transformations to get our working data model into the latest NTIA publishing format.

**Submission Comparison Check:** Starting with this submission an additional check was added to our quality review process. An application was written that compares the individual provider's unique data that is stored in their unique file to that which is stored in our final appended file and the NTIA submission data. Any variation in each of these data files in thoroughly investigated and resolved. This was done to assure no data loss or data transformation issues. We also compare this submission to the previous submission, review any variations and assure that the changes found can be documented as being requested by the provider.

**Final QA/QC:** A series of quality checks are run on the final appended data sets to ensure it is ready for submission to NTIA. We also run the latest version of the NTIA receipt tool at this time. If any issues are flagged as failing they are reviewed and corrected. All warnings are also reviewed and either corrected or documented in the attached document which explains that we have validated this data and it should be accepted. Any last issues are corrected, and the data is sent to the state for their review.

## **Submission to NTIA**

## **2.3 DETAILED DATA PROCESSING**

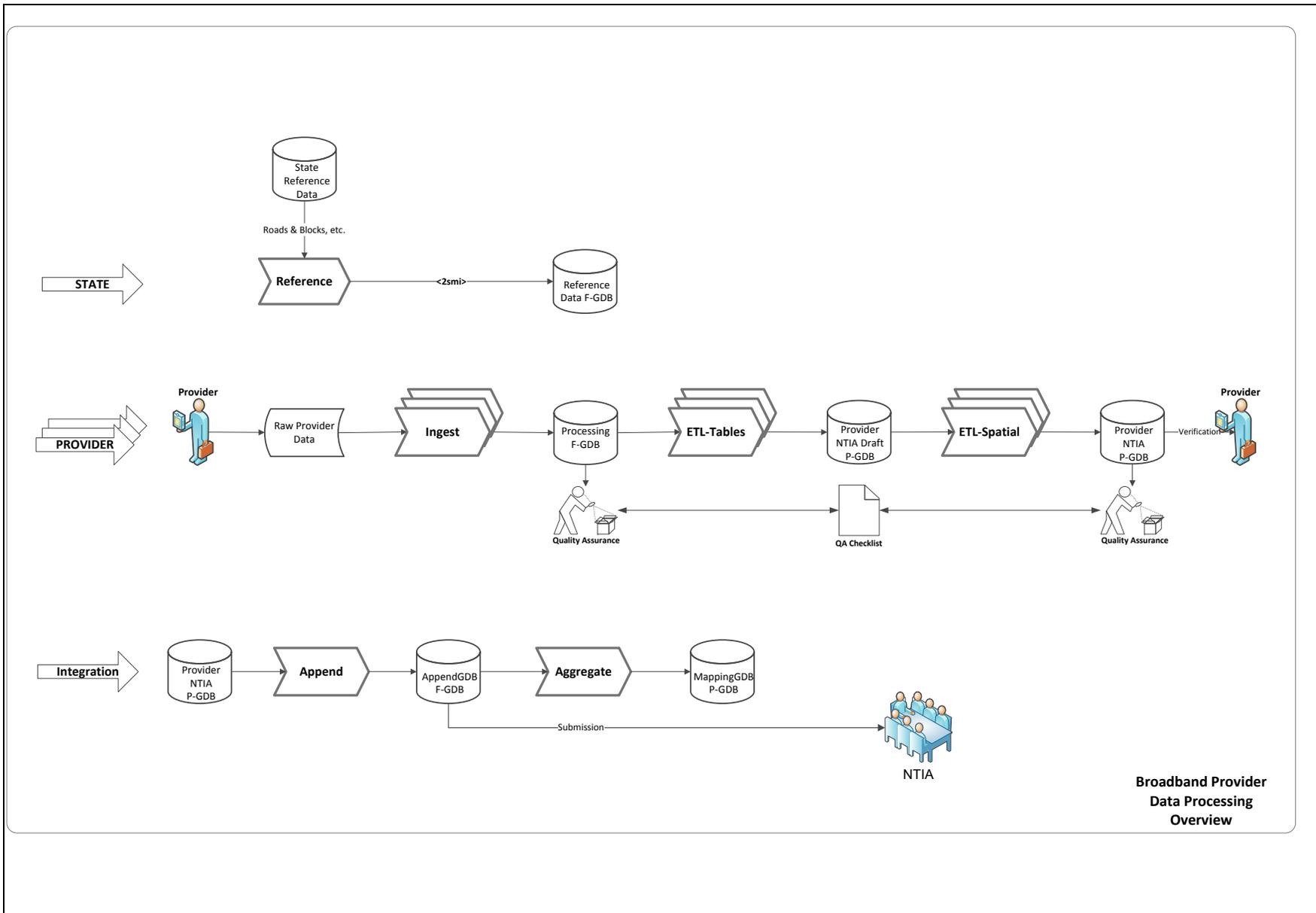
### **2.3.1 General Overview**

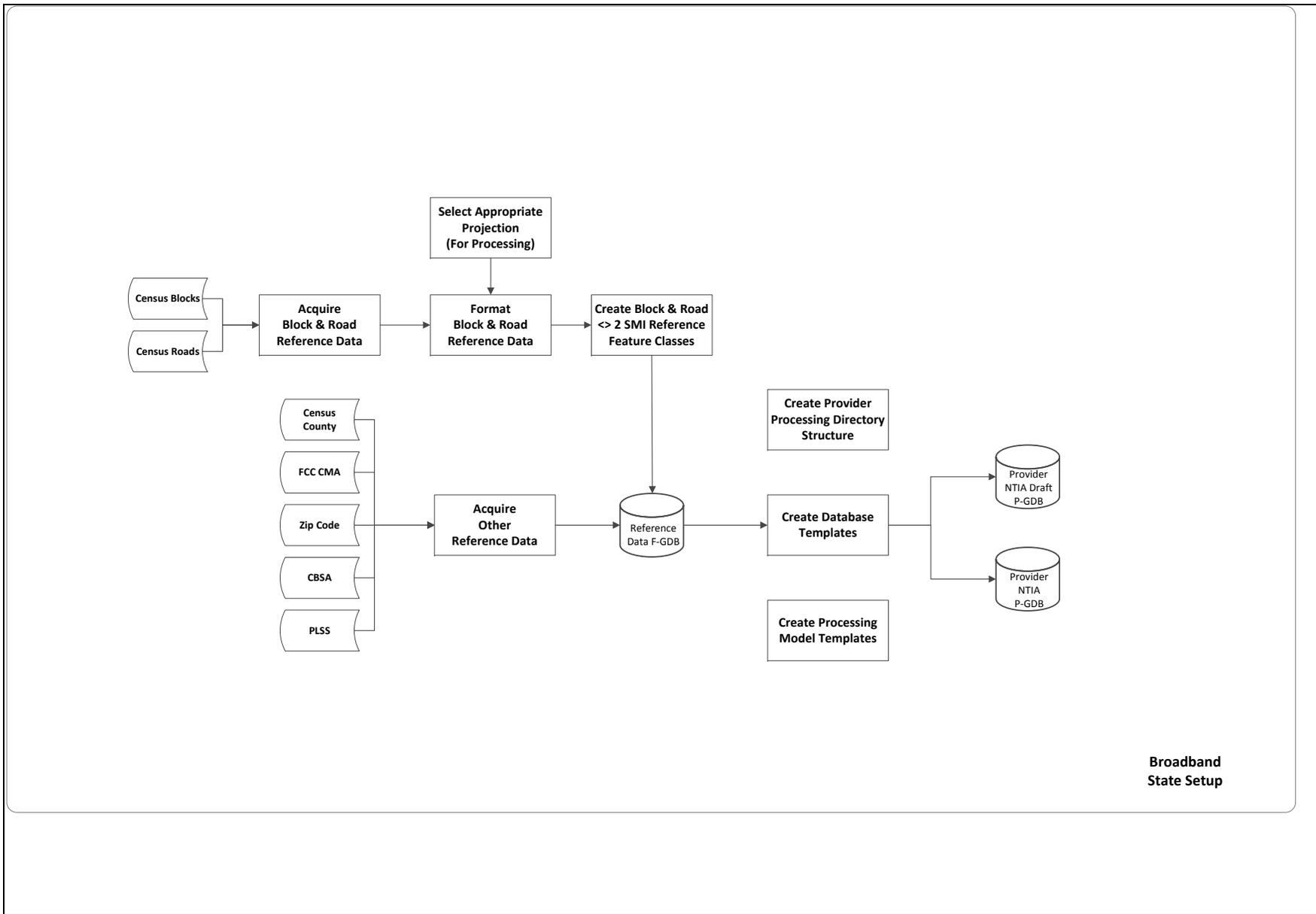
The Data Processing Team is tasked with receiving data directly from each State's broadband providers and carrying that data through a series of spatial and database transformations that result in a standardized, aggregated database

ready for submission to the State and to NTIA. The key phases that make up this critical process are:

1. Initial State Set Up Phase
2. Provider Processing Phase
3. Integration Phase

Each of these phases is described in detail below and represented graphically on Figure 1.





### 2.3.2 Initial State Set-Up Phase

An overview of the Initial State Set-up phase and processing is shown in Figure 2. For every submission for each State, the team completes a series of setup steps to ensure that necessary reference data is acquired from the State or US Census Bureau and then prepared to support later spatial analysis. The results of this set up phase are maintained in the **Reference Data Personal Geodatabase**. The following steps are followed to accomplish this phase:

1. Acquire Census Block and Road Reference Data
2. Select Appropriate Projection for Processing
3. Format Block and Road Reference Data

#### Blocks

- For each Census block compute the total area by summing the water area and land area values
- Convert total area to square miles (from square meters)
- Compute and populate column to indicate that the block is < 2 square miles in area (isl2smi)
- Add NTIA standard columns to the block reference data set
- Rename columns to Block\_\* for easier identification during processing
- Output all blocks into the state reference database

#### Roads

- Add NTIA standard columns to the block reference data set
- Rename columns to Road\_\* for easier identification during processing
- Perform a spatial join (intersect with lt2sqmi) then compute and populate column to indicate that the road segment is in a block that is greater than 2 square miles in area (isgt2smi)
- Output all roads into the state reference database

#### 1. Create Block and Road Greater Than or Less Than 2 Square Miles Reference Data

The size of each Census Block is the key factor in determining the spatial format (block or road) for final submission, thus the initial setup phase includes categorization of reference data according to area in square miles. Provider service coverage data for areas represented by blocks that are less than 2 square miles is processed as block features. Provider service coverage data for areas represented by blocks that are greater than 2 square miles is processed as Tiger Line road segments features. Therefore, we output blocks that are less than 2 square miles in area. (**lt2smi**)

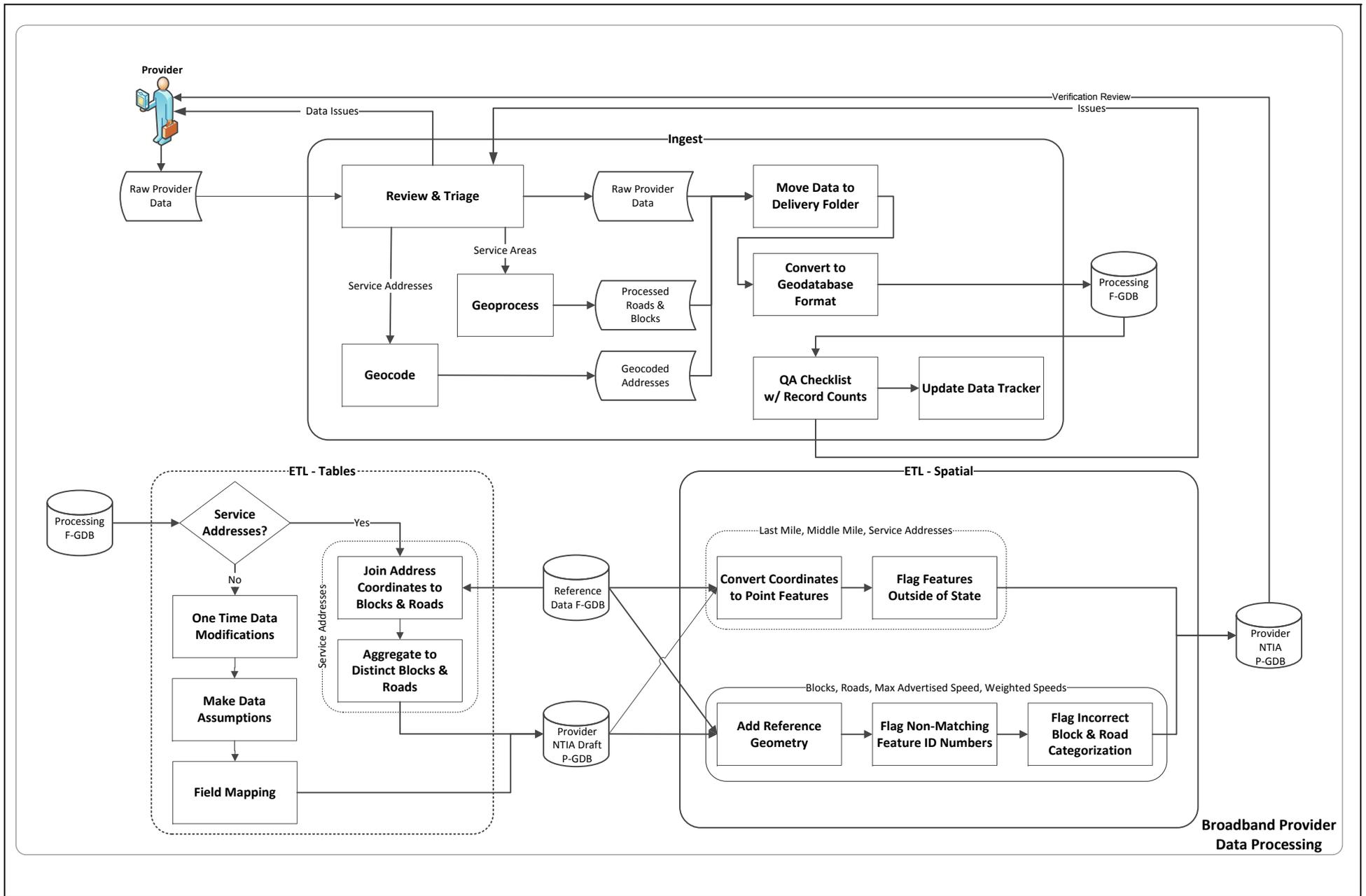
#### 2. Create Database Templates

To ensure consistency across providers and to increase the potential for process scripting, template geodatabases are created. In addition to the Reference Data File Geodatabase described above, there are four additional geodatabases used in processing:

- Provider Processing File Geodatabase
- Provider NTIA File Geodatabase
- NTIA Append Geodatabase
- Mapping Aggregation Geodatabase

### **3. Create Processing Model Templates**

After these steps are complete, the processing of individual provider data sets can begin.



### 2.3.3 Provider Processing Phase

The Provider Processing Phase is described below in details and also shown in the Figure 3 above. Following the initial set up, the team processes the provider data to bring it from its “raw” format to a standardized file Geodatabase format. This process is repeated and checked until each data set is successfully carried through each step.

#### 1. Review and Triage

The ingest process begins upon receipt of raw data from the State’s providers. Data is routed for initial geoprocessing based on data format. Any issues that surface during the initial triage will trigger a request for additional information and clarification from the Provider.

#### 2. Data Processing

The following sections of the document described the workflow for each dataset.

### Address Data Processing

#### 1. Ingest Address Data

If service address locations are supplied by the Provider, these data will **be geocoded** to the reference data resulting in address point data in a Geodatabase format. We have an 85% geocode rate as a cut-off – and if such a rate is not achieved through an automated geocoding, we then will manually geocode the data. The geocoding is done using a geocoder based on the US Census TIGER 2010 street centerline data. These data are then also stored in the **Delivery Folder** according to provider name and delivery date. Address points are then imported to the **Processing File Geodatabase** along with all other data supplied by that Provider on that date. A set of **quality control checks** are performed on the address data including:

- Visual review to confirm address locations are within the state boundary and there are no outliers that are incorrect.
- Required attribute fields are present.
- Latitude and Longitude are appropriate (not inverted).

#### 2. Extract, Transfer, Load (ETL) Tabular Address Data

Address point data requires additional **geoprocessing** to convert the points to either block or road data in preparation for submission to NTIA.

One time Data Modifications are as follows:

- Address points are re-projected to allow spatial join with State reference data.
- Using state reference data, address points are spatially joined to block polygons.

- Resulting blocks and roads are aggregated according to the project specifications.
- The output from these tabular ETL processes is the Provider NTIA Draft File Geodatabase.

### 3. Extract, Transfer, Load (ETL) Spatial Address Data

Reference geometry is now added to the address data (which has been aggregated into block and road records during the tabular ETL processes). Reference geometry is added by spatially joining State reference data.

- Non-matching Feature Identification Numbers are flagged.
- Inappropriately categorized data is flagged.

These processes result in the **Provider NTIA File Geodatabase** in NTIA format with all attribute and spatial elements.

- Address point features are added to the Provider NTIA File Geodatabase.
- The Provider NTIA File Geodatabase is given to Providers for review and verification via the web-based Provider Portal.

## Street Segment Data Processing

- Ingest Street Segment Data
  - a) Street segment data are typically supplied by the Provider as tabular data.
  - b) These data are stored in the **Delivery Folder** according to provider name and delivery date.
  - c) Data is then converted to a **File Geodatabase format**.
  - d) A series of **quality control checks** are performed on the street segment data.
- Extract, Transfer, Load (ETL) Tabular Street Segment Data
  - a) One-time Data Modifications
  - b) Data Assumptions
- The Team assumes all data is 2010 TIGER Line Identification Numbers.
- Data table fields are then mapped to fields in the NTIA Draft Geodatabase format.
- The output from these ETL processes is the creation of database in NTIA format that does not yet have associated spatial features. This is the Provider NTIA Draft File Geodatabase.
- For block and road data, if advertised speeds have not been provided for all the records, the team verifies that a Maximum Advertised Speed feature class exists with this information based on an approved alternate geography.

- When Maximum Advertised Speed is provided in an alternate format per the Clarifications to the NOFA (e.g. by franchise area, service area or CMA) – the team has not back-filled the speed information into the block dataset because providers have told us that they may not be able to provide at those speeds at the individual block level.
- When the same provider provides on the same block with multiple technologies of transmissions, providers have been asked to provide two separate records for the same block representing the two technologies of transmission.
- Where TLIDs are present, data may or may not be present in the Segment\_Min\_Address etc. fields – by joining on TLIDs, such data can be populated in future
- The Minimum and Maximum Address values have been calculated using From and To Address on the left and right of street segments (as in the TIGERLINE data) and often the minimum can be more than the maximum based on the direction of digitization.
- Extract, Transfer, Load (ETL) Spatial Street Segment Data
  - a) Reference geometry is now added to the street segment data through attribute join using the TIGER Line ID number and available State reference data.
  - b) Many-to-one relationships are resolved so that every record has an associated shape and a valid TIGER Line ID number.
  - c) Non-matching Feature Identification Numbers are flagged.
  - d) Inappropriately categorized data is flagged.
  - e) These processes result in the Provider NTIA File Geodatabase in NTIA format with all attribute and spatial elements.
  - f) The Provider NTIA File Geodatabase is given to Providers for review and verification via the web-based Provider Portal.

## **Census Block Data Processing**

1. Ingest Census Block Data
  - a) Census block data is supplied by the Provider.
  - b) These data are stored in the Delivery Folder according to provider name and delivery date.
  - c) Data is then converted to File Geodatabase format.
  - d) A series of quality control checks are performed on the census block data including:
  - e) Data is checked to verify whether vintage is 2000, 2009 or 2010.
  - f) Required attribute fields are present including Provider Name, DBA, FRN, Record Number, and Census Block Number
2. Extract, Transfer, Load (ETL) Tabular Census Block Data
  - a) Data Assumptions

- o Team calculates the State, County, Census fields from the full block ID
  - o When Maximum Advertised Speed is provided to us in an alternate format per the Clarifications to the NOFA (e.g. by franchise area, service area or CMA) – the team has not back-filled the speed information into the block dataset because providers have told us that they may not be able to provide at those speeds at the individual block level.
- b) When the same provider provides on the same block with multiple technologies of transmissions, providers have been asked to provide two separate records for the same block representing the two technologies of transmission.
  - c) Data table fields are then mapped to fields in the NTIA Draft Geodatabase format.
  - d) The output from these ETL processes is the creation of database in NTIA format that does not yet have associated spatial features. This is the Provider NTIA Draft File Geodatabase.
3. Extract, Transfer, Load (ETL) Spatial Census Block Data
- a) Reference geometry is now added to the census block data through an attribute join using the Census Block ID numbers and State reference data.
  - b) For records that came in as Census 2000 or 2009 format, they are compared to a dissolve reference set that uses the 2010 geometry and the 2010 block numbers.
  - c) Non-matching Feature Identification Numbers are flagged. Team puts unmatched records in a separate 'dropped records' table, once for each feature type.
  - d) Inappropriately categorized data is flagged.
  - e) These processes result in the Provider NTIA File Geodatabase in NTIA format with all attribute and spatial elements.
  - f) The Provider NTIA File Geodatabase is given to Providers for review and verification via the web-based Provider Portal.

## **Wireless Service Area Availability Data Processing**

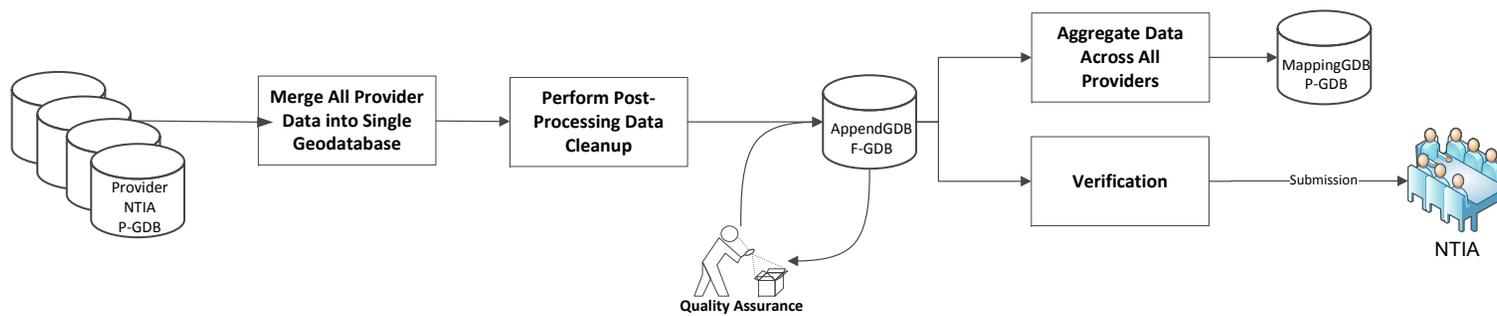
1. Ingest Wireless Service Area Availability Data
  - a) Wireless Service Area Availability data are supplied by the Provider typically as shapefiles or tables.
  - b) These data sets are also stored in the Delivery Folder according to provider name and delivery date.
  - c) Data is then converted to File Geodatabase format.
  - d) A series of quality control checks are performed on the census block data including:

- Visual review to confirm whether features are within state boundary
  - Required attribute fields are present including Provider Name, DBA, FRN, Technology Transmission, Spectrum, and Maximum Advertised Downstream Speed
2. Extract, Transfer, Load (ETL) Tabular Wireless Service Area Availability Data
    - a) One-time Data Modifications
      - Project data to WGS-84 Web Mercator
      - Spectrum data is translated from correspondence if not included in data.
    - b) Data table fields are then mapped to fields in the NTIA Draft Geodatabase format.
    - c) Topology is run on the data for each provider for no overlaps.
    - d) The output from these ETL processes is the creation of database in NTIA format. This is the Provider NTIA Draft File Geodatabase.
  3. Extract, Transfer, Load (ETL) Spatial Wireless Service Area Availability Data
    - a) Spatial ETL process does not occur for Wireless Service Area Availability Data. Draft data is copied into Provider NTIA File Geodatabase.

## **Infrastructure Data Processing**

1. Ingest Infrastructure Data
  - a) Infrastructure data are typically supplied by the Provider as a table or shapefile.
  - b) These data are then also stored in the Delivery Folder according to provider name and delivery date.
  - c) Data is then converted to File Geodatabase format.
  - d) A series of quality control checks are performed on the infrastructure data including:
    - Required attribute fields are present including Provider Name, DBA, FRN, Technology Transmission, Latitude, Longitude
2. Extract, Transfer, Load (ETL) Tabular Infrastructure Data
  - a) One-time Data Modifications
  - b) Data Assumptions
    - Team calculates the State, County, Census fields from the full block ID
  - c) Data table fields are then mapped to fields in the NTIA Draft Geodatabase format.

- d) The output from these ETL processes is the creation of database in NTIA format that does not yet have associated spatial features. This is the Provider NTIA Draft File Geodatabase.
- 3. Extract, Transfer, Load (ETL) Spatial Infrastructure Data
  - a) Points are created from Latitude/Longitude data
  - b) Features outside of state boundary are flagged
  - c) These processes result in the Provider NTIA File Geodatabase in NTIA format with all attribute and spatial elements.
  - d) The Provider NTIA File Geodatabase is given to Providers for verification via the web-based Provider Portal.



**Broadband Provider  
Data Integration**

### 2.3.4 Integration Phase

The Integration Phase combines the individual Provider data sets into a single geodatabase for submission to NTIA.

#### **Append**

The Provider Processing phase produces a separate geodatabase for each Provider in NTIA submission format. The Append step compiles all of these data sets into a single geodatabase with rows for each provider, technology type, end-user category, etc. for each data set. This results in many duplicate (a.k.a. stacked) block and road features.

Following the compilation, quality checks are performed to ensure that all Provider data was transferred successfully into the appended geodatabase that contains the following feature classes:

- Area availability
- Census Blocks Less than 2 Square Miles
- Last mile
- Maximum Advertised Speed
- Middle Mile
- Road Segment Availability

Refer to the submission data model schema in this document for details about the columns contained in these feature classes of the append geodatabase. An additional post-processing performs final data validation steps across the full submission. For example, all data elements that have domain values and are null are set to the domain value that indicates that data was not provided. The append geodatabase is the format that is submitted to NTIA.

#### **Aggregate**

Following the Append step, the Aggregate step creates another geodatabase that is used for mapping and analysis purposes. For each data set in the append geodatabase, identical features from all Providers, technology transmission types, End-User Categories, etc. are rolled-up into a single, aggregated feature class. The duplicate blocks and roads are removed and the attributes are combined into a single row that describes the feature. This process produces a geodatabase that contains only block and road feature classes. Descriptive attributes for these feature classes include:

- Census Block ID number (blocks) or Tiger Line ID (Roads)
- Number of providers providing service to the feature area
- Maximum advertised downstream speed tier for that feature area
- Maximum advertised upstream speed tier for that feature area
- The number of providers providing service at each speed tier (to easily calculate served/underserved area)
- Number of providers providing service for each technology of transmission type

This data set also undergoes a series of quality control checks to ensure that no data was lost during the aggregation process.

## **Deliver to NTIA and Publish to Web Applications**

A copy of the **Append File Geodatabase** is generated to be used in the provider portal web-based application. When verification feedback is received, the individual provider geodatabases are updated. After verification is complete, the Append process, including QA steps, is executed again and then submitted to NTIA.

## **2.4 DATA VALIDATION**

Sanborn has validated data using the following steps:

1. QC of the data at various steps – this includes when data is received (triage), when it is processed through the various processing steps discussed above, etc.
2. Spatial checks against public and commercial datasets
  - a. For GA, we used Pitney Bowes exchange boundaries for validation. Any wireline providers whose areas fell outside the exchange boundaries were marked up in an issues database and the information about this was relayed to the provider. We will continue improving the data where providers did not respond in the next submission.
3. Speedtest data
  - a. For this submission we used the FCC speed test data for validation. We geocoded the data, used the IP to reverse engineer the provider name and used it to check speeds where possible. We used the deadspot data to identify areas of no providers in our feedback to providers.
4. GTA provided a huge amount of local and previous knowledge in validating data. Sanborn created an Online Data Verification Tool for GTA to review the data. Issues identified by Sanborn through the above methods were already available for GTA to review and further feedback from GTA was reported to providers and data corrections were made where possible. In some cases, providers were non-responsive and we tabled some changes for next submissions. We feel that we addressed most major issues with the previous submission in this submission. Some examples of these verification discussions include:
  - a. Zayo Group, LLC. This is a fiber provider and their service was marked as business only. All blocks that were obviously only serviceable by their long-haul fiber were removed. Several attempts were made to contact them with these changes but no responses were forthcoming. Further improvements may be needed on this and we hope to engage the provider in the next submission.

- b. Knology, Inc. / WOW. We removed their fiber offering from S6 based on feedback from GTA, discussions with the provider and using the guidelines developed for fiber mapping.
  - c. Comcast – Based on several communication attempts regarding claims of 100 mbps service in several rural areas of GA that are disputed by locals, Sanborn facilitated a meeting between Comcast and GTA. Comcast explained that they have made a huge investment upgrading their systems and a few counties (Hart and Elbert) were downgraded in speed based on that conversation. Both Comcast and GTA were agreeable to that resolution.
  - d. AT&T – There was some discrepancies in the total number of roads served by AT&T in past submissions vs. Submission 7. It is unknown to us what methodology was used in the previous submission that resulted in a submission of 33,289 roads to 10,810 roads. We have confirmed with AT&T that the data we have submitted for S7 correctly displays the actual road segments that are currently served by AT&T rather than a filling in of gaps.
5. Verification by providers – processed data is uploaded on our Provider Portal for providers to review both the outcome of data processing and any issues that we found in the third-party and GTA validation. Issues pertaining to a particular provider are highlighted and shown in the portal for those providers only. Issues that are global and cannot be assigned to a particular provider are shown to all providers (e.g. there are no providers in this area, or we tried to get service here and heard x from A provider, y from B provider, etc.). We have also made several additional calls to providers who have issues.

### 3 Submission 6: NTIA Submission Data Model Schema Changes

This section of the document describes the strategy that was used for the development of the specific data schema used for the NTIA submission 7 provider data. The current data model schema is as of Dec 2012 and as posted on the SBDD site.

#### 3.1 Schema history and evolution

In submission 1, NTIA asked the National States Geographic Information Council (NSGIC) to comment and provide a spatial data model that can provide a common format for data submitted to NTIA. The initial NSGIC data model released had a number of flaws that clearly needed to be resolved.

NSGIC released version 2 of the data model close to the submission 1 delivery date. The new model had improved functionality and conformed more closely to the NTIA submission requirements. The NSGIC version 2 model was used as the basis for our internal processing models and for submission 2.

After submission 1, NTIA took ownership of the submission data model, but did not release any changes until mid-August. The NSGIC version 2 was used as the basis for our internal processing models. The submission 2 NTIA data model is similar to the NSGIC version 2 model.

To retain as much of the NSGIC v2 /NTIA spatial data model as possible, the relationship between the provider data and the output specification is kept as simple as possible. Here are a few key NTIA submission data model design considerations:

- Submission feature class names reflected the names in the NSGIC v2 specification
- Column data types are based on the NSGIC v2 specification
- Where possible, field names retained the naming conventions of the NSGIC specification
- All road segment address information used the NSGIC specification of a single min, max, zip for each feature
- The data schema for wireless data follows the NSGIC specification for submitting a single feature per spectrum
- To retain Provider Source Information the ID field is calculated as State Name Abbreviation “\_”, Short Name. The ID field exists in the NSGIC v2 data model, but not the final NTIA submission 2 delivery model. This column is used during processing and was dropped during final processing, prior to submission to NTIA.
- Any Overview records that were not submitted using State-County codes were not delivered.

### 3.2 NTIA Data Model Changes Submission 2

During the processing of provider data for submission 2, a number of issues were raised about the data model requirements proposed by NTIA. A number of specific errors, such as typographical errors in domain values, or inconsistency surrounding processing of null values, etc., were documented and forwarded to NTIA for response. Based on changes made to the NTIA data model, some data processing procedures were required to populate the current NTIA data model. The following is a list of specific data processing changes that were implemented:

The following are the rules for removing records for the final NTIA submission:

- Basic Assumptions:
  - Remove any record that has a Maximum Advertised speed that did not meet the definition of broadband
  - MaxAdv is only required in wireless
  - MaxAdv can be null in blocks/roads
- Criteria for removing records from Blocks/Roads (wireline)
  - Remove records with invalid MaxAdv speeds
- Criteria for removing records from Wireless
  - Remove records with invalid MaxAdv speeds
  - Remove records with null MaxAdv speeds

In addition, the following processing changes are now performed during post-processing and before the final NTIA submission:

- There is a new feature class called State Boundary. These shapes were prepped and added to the reference datasets for each state. For NTIA submission output, these were moved into their own feature class.
- The Blocks table has the Block ID separately defined as State, County, Tract, and Block ID. The provider data as processed include the full 15-digit FIPS code, which has been parsed to populate these fields.
- ID columns no longer exist. They have been dropped from the final processed data.
- The Middle Mile, Overview, and Wireless tables all have a field called StateAbbr (2 character alphabetic code). The final publishing script created and populated the StateAbbr field.
- In the Roads table, the Ref\_Values are used for Street info and Zip Code. Because the processing produces a null value for Ref\_City, the City field is populated with Div\_City.
- TransTech was converted from string to small integer.
- Any record with a TransTech value of X, Y, or Z was dropped.

Any other field with a value of X, Y, or Z was set to null.

- Any Elevation with a -9999, -9998, or -9997 was set to null.

- Any FRN generated during processing (those starting with 00000000\_\_) were converted to a value of 9999.
- In the Blocks and Roads tables there are new fields called Reseller. Because only data from actual providers was accepted, this field was set to 'No' for all records.
- Block geometry was converted from hybrid geometry back to 2000 Census geometry.

### 3.3 NTIA Data Model Changes Submission 3

The data model released on January 13, 2011 contained the following changes to the submission 2 data model:

- A new field was added to several feature classes called Provider Type
  - Provider Type is "Short Integer" and has domain values of 1, 2, or 3 (1= Broadband Provider, 2=Reseller, 3=other)
  - Most providers are calculated to be "1" (Broadband Provider). In some cases (e.g. Public Utility Districts), providers are considered "Other" (value = 3).
- In the CAI feature class, the field BBService was modified:
  - In S2 if the information was not known, the field was left blank (null)
  - In S3, if we do not have the information, Null's must be changed to code U (for Unknown) – nulls will not be allowed.
- Three new fields have been added to the CAI feature class. Wherever possible, these values have been populated in the CAI data.
  - PublicWifi (Y, N or U)
  - URL
  - CAIID

### 3.4 NTIA Data Model Changes Submission 4

The data model released on June 30, 2011 and it contained the following changes to the submission 3 data model:

The Category of End user field was added back in to the block and road tables.

- In addition the domain values were changed. 1 still represents residential, but a 2 now represents all non-residential uses.
  - This field was not required, and for many providers, was left blank since the data was not provided.

### 3.5 NTIA Data Model Changes Submission 5 & 6

No changes were made to the data model for submissions 5 and 6.

The latest data model released was released on August 8, 2012 was very similar to the previous data model. No substantive changes were noted and changes related to allowable speed and technology of transmission combinations. Most of

these combinations have exceptions to them and hence were not being completely disallowed by NTIA.

## **4 UNIVERSE OF CONTACTED PROVIDERS/NON-PROVIDERS**

According to both our research and lists provided to use by NTIA, there was the potential for Georgia to have up to 204 broadband providers.

Of these:

- 45 companies stated that they do not provide any type of broadband service in Georgia. Many of these are either national carriers without a Georgia presence, or they file 477 reports because they provide VOIP or Video Teleconference services (but not broadband).
- 42 company names turned out to be a DBA or legal holding names for another firm that is listed in another category. So these duplicates were dropped from our list.
- 16 companies are resellers and are not considered part of this program.
- 16 companies may be broadband providers, but they have either indicated they are not willing to provide data, or were completely unresponsive to multiple attempts to contact them.
- 37 broadband providers informed us that there were no changes to their service area so for these providers we downloaded the data that had been submitted as part of the last submission, reprocessed the data, performed validation on it, and we are resubmitting this data:
- 47 Broadband providers submitted either entirely new or partially new datasets for this submission:

### **4.1 Non-providers**

45 companies stated that they do not provide any type of broadband service in Georgia. Many of these are either national carriers without a Georgia presence, or they are out of business or have been purchased by other companies or they file 477 reports because they provide VOIP or Video Teleconference services (but not broadband).

1. 360 Networks
2. Airespring, Inc.
3. American Fiber Network, Inc.
4. Bellsouth Long Distance, Inc.
5. Bluebird Wireless Broadband Services, LLC

6. Broadcore, Inc.
7. BroadRiver, Inc. & BroadRiver Communications Corp.
8. BullsEye Telecom, Inc.
9. Cbeyond Communications, LLC
10. CIMCO Communications, Inc.
11. City of Augusta
12. City of Manchester
13. City of Milledgeville
14. City of Statesboro
15. Coastal Broadband
16. Convergence Technologies, Inc.
17. EagleNet
18. EnerSphere Communications LLC
19. Enventis Telecom Inc.
20. eVolve Business Solutions LLC
21. Harbor Communications
22. HCE Media, LLC / Ridge Networks
23. Light Tower Fiber Long Island LLC
24. LightEdge Solutions, Inc.
25. Lintel, Inc.
26. MainStreet Broadband
27. Netlogic, Inc.
28. Qwest Communications International, Inc.
29. Reliance Globalcom Services, Inc.
30. Reynolds Cable TV Inc.

31. RGW Communications, Inc.
32. Shentel Converged Services, Inc.
33. SkyWay USA
34. Smartresort Co., LLC d/b/a/ Beyond Communications
35. Suburban Cable Inc.
36. Telefonica USA, Inc.
37. Telovations, Inc.
38. University Corporation for Advanced Internet Dev.
39. VectorLink
40. Wandering WiFi
41. Windjammer Communications LLC

The companies listed below; do not provide broadband service consistent with the NTIA definition of broadband availability.

42. Broadstar, LLC d/b/a PrimeCast
43. DirectPath
44. FPL FiberNet LLC
45. Stratos Offshore Services Company

#### **4.2 Shell Companies**

The following 42 company names turned out to be a DBA or legal holding names for another firm that is listed in another category. So these duplicates were dropped from our list.

1. Accucom Telecommunications
2. AGL Networks , LLC

3. Airimba and Windchannel Communications
4. American Fiber Systems, Inc.
5. ATC Broadband LLC
6. Birch Telecom, Inc.
7. Blue Ridge Telephone Company
8. Board of Water, Light & Sinking Fund  
Commissioners
9. Broadwing Communications, LLC
10. Business Telecom, Inc.
11. Camden Telephone & Telegraph Co., Inc.
12. Cellco Partnership
13. City of Monroe
14. City of Tifton
15. ComSouth Telenet, Inc.
16. Covad Communications Company
17. Depot Street Communications, Inc.
18. DoveTel Communications, LLC
19. DSLnet Communications, LLC
20. Dycom Holding, Inc.

21. ETC Communications, LLC
22. Flint Cable TV, Inc.
23. GEORGIA RSA # 8 PARTNERSHIP Limited Partnership
24. Habersham Electric Membership Corporation
25. ITC Globe, Inc.
26. James Cable LLC
27. KLiP, LLC
28. Madison River Communications, LLC
29. MediaStream
30. Nelson-Ball Ground Telephone Company
31. New Cingular Wireless Services, Inc.
32. Northland Cable Properties Seven Limited Partnership
33. Northland Cable Properties, Inc.
34. Northland Cable Television, Inc.
35. NuVox Communications, Inc.
36. Plant Tifnet
37. Quincy Telephone Company

38. US LEC of Georgia Inc.
39. Valley Cable TV, Inc.
40. ViaSat Inc.
41. Wideopenwest
42. WiTel Communications, LLC

### **4.3 Resellers**

The following 16 companies are resellers and are not considered part of this program.

1. American Telephone Company LLC
2. Birch Communications, Inc.
3. Broadview Networks, Inc.
4. CONEXIZ Corporation
5. Digital Agent, LLC
6. Georgia Business Net
7. Global Crossing North American Networks, Inc.
8. Greenfly Networks, Inc.
9. Interglobe Communications, Inc.
10. Metropolitan Telecommunications of Georgia, Inc.
11. Net2Atlanta
12. New Edge

13. Reallinx, Inc.
14. Smart Choice Communications, LLC
15. South GA Governmental Services Authority
16. Wholesale Carrier Services

#### **4.4 Non-Participating or Non-Responsive Providers**

16 companies may be broadband providers, but they have either indicated they are not willing to provide data, or were completely unresponsive to multiple attempts to contact them.

1. airBand Communications, Inc.
2. Albany, Water, Gas and Light Commission
3. Brightlan LLC
4. City of Cartersville
5. Electric Power Board
6. Georgia Public Web, Inc.
7. Gosuto Wireless Internet
8. Kennedy Cablevision Inc.
9. Netlink IP Communications
10. One Ring Networks
11. Parker Fibernet
12. Sunesys
13. SyncGlobal
14. The Seimitsu Corporation

15. Verizon Communications d/b/a Verizon Business  
Glob

16. Wireless Hometown LLC

#### **4.5 Providers with No Data Updates**

37 broadband providers informed us that there were no changes to their service area so for these providers we downloaded the data that had been submitted as part of the last submission, reprocessed the data (to remove water only polygons and to reprocess for the different methodology for 2 sq mile calculation), performed validation on it, and we are resubmitting this data:

1. AL-GA Wireless Broadband, LLC
2. Alltel/ Allied Wireless Communications Corp
3. ATG Communications, LLC
4. Brantley Telephone Company, Inc.
5. Bulldog Cable Georgia
6. Bulloch Telephone Cooperative
7. Cavalier Telephone LLC or Talk America
8. Chickamauga Telephone Corporation
9. City of Cairo, GA
10. City of Camilla d/b/a South Georgia Gov't Svcs.
11. City of Dublin
12. City of LaGrange
13. City of Moultrie

14. City of Thomasville Utilities
15. ComSouth
16. Depot Street Communications, Inc
17. Fort Valley Utility Commission
18. Glenwood Telephone Company
19. Hargray of Georgia, Inc.
20. Kings Bay Communications, Inc.
21. KitePilot Wireless Internet
22. MonroeAccess.net
23. Nextlink Wireless, Inc.
24. NuLink Digital
25. Pembroke Telephone Company, Inc.
26. Plant Telephone Company
27. Plantation Cablevision, Inc.
28. Progressive Rural Telephone Co-op., Inc.
29. Quitman Wireless
30. Skycasters
31. South Georgia Regional Information Technology Auth

32. Southeastern Services, Inc.
33. StarBand Communications, Inc.
34. Viasat/WildBlue Communications, Inc.
35. Wave2Wave Communications, Inc. & RNK
36. Wilkes Telephone & Electric Company
37. XO Communications, LLC

#### **4.6 Providers with Data Updates**

A total of 47 Broadband providers submitted either entirely new or partially new datasets for this submission:

1. AI-CALL, Inc./ATC
2. Appalachian Valley Fiber Network
3. AT&T of Georgia
4. AT&T Wireless
5. Bright House Networks Information Services
6. CenturyLink/CenturyTel
7. Charter Communications
8. Citizens Telephone Company, Inc.
9. Clearwire
10. Cogent Communications Group

11. Comcast Corporation
12. Cox Communications, Inc.
13. Cricket Comm/Leap Wireless International, Inc.
14. Dalton Utilities
15. Darien Telephone Co., Inc.
16. DeltaCom, Inc./EarthLink Business
17. ELBERTON, City of
18. Ellijay Telephone Company
19. Fairpoint/GTC, Inc.
20. FiberLight, LLC
21. Frontier Communications of Georgia, Inc.
22. Hart Telephone Company
23. Hughes Communications/HNS License Sub, LLC
24. iWispr LLC
25. JamesCable (Waycross Cable) d/b/a MediaStream
26. Knology, Inc.
27. Level 3 Communications, LLC
28. Mediacom Communications Corp & MCC Georgia LLC

29. MegaPath
30. MetroPCS Georgia, LLC
31. North Georgia Network Cooperative, Inc
32. Northland Cable Properties Eight Limited Partnersh
33. Pineland Telephone Cooperative, Inc.
34. Planters Rural Telephone Cooperative, Inc.
35. Public Service Data Wireless
36. Public Service Telephone Company/Flint Cable
37. Ringgold Telephone Company
38. Sprint Nextel Corporation
39. TDS Telecomm
40. T-Mobile
41. Trenton Telephone Company
42. TruVista
43. TW Telecom of Georgia L.P.
44. Unite Private Networks, LLC
45. Verizon Wireless
46. Waverly Hall Telephone, LLC

## 47. Windstream Georgia Telephone