



## Oklahoma Broadband Mapping

---

# Data Submission Report

*4<sup>th</sup> Submission (October 1, 2011)*

**October 1, 2011**

---

Sanborn  
1935 Jamboree Drive  
Suite 100  
Colorado Springs, CO 80920

# Oklahoma Broadband Mapping

## 4<sup>th</sup>Data Submission Report

### Table of Contents

<a href="#">1 INTRODUCTION.....</a>	<a href="#">3</a>
<a href="#">1 OVERALL PROJECT STATUS.....</a>	<a href="#">4</a>
DATA COLLECTION.....	4
<a href="#">1.1.1 Broadband Data.....</a>	<a href="#">4</a>
<a href="#">1.1.2 Community Anchor Institutions Data.....</a>	<a href="#">6</a>
DATA PROCESSING.....	8
<a href="#">1.1.3 General Overview.....</a>	<a href="#">8</a>
<a href="#">1.1.4 Processing of Provider Data.....</a>	<a href="#">8</a>
<a href="#">1.1.5 Conversion Process of Data from Census 2000 to 2010.....</a>	<a href="#">10</a>
<a href="#">1.1.6 Submission 4: ReferenceData Creation.....</a>	<a href="#">11</a>
<a href="#">1.1.7 Submission4: NTIA Submission Data Model Schema Changes.....</a>	<a href="#">11</a>
DATA VALIDATION.....	11

# 1 Introduction

This report is submitted along with the fourth data submission for the Oklahoma 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 formal and informal 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 OK. The State of Oklahoma has retained a mapping contractor, The Sanborn Map Company to perform the work related to the Mapping Grant for this project. Data from the previous submission is now publicly accessible via the Oklahoma Broadband Program (<http://www.ok.gov/broadband/>).

**This document is a supplement to the three previous reports submitted with previous data submissions on May 1, 2010, October 1, 2010, and April 1, 2011 respectively.** Therefore, it builds on the document provided with those submissions. Rather than repeat the contents of the previous report, this document makes incremental updates on various topics where changes have been made in the methodology or reiterates the methodology used. Please refer to the previous documents for further details.

# 1 Overall Project Status

## DATA COLLECTION

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

### 1.1.1 Broadband Data

For this submission, Sanborn began data collection efforts on July 12<sup>th</sup> 2011 by sending out data update requests and technical data specifications to all providers. This incorporated all the NTIA changes released on June 30<sup>th</sup>, 2011. These were sent to a large list of companies which were compiled from past collection efforts, the revised FCC 477 (dated June 30<sup>th</sup>, 2010), a list of providers from the NTIA's Wireless Internet Service Providers Association (WISPA) and from any providers that were identified through other sources such as web research, planning meetings, etc. We then actively followed up with the providers and as we had discovered in the past, many of the providers listed on the FCC 477 list are either resellers, or not involved in the actual delivery of broadband. (Many are VOIP or teleconference service providers that utilize existing broadband connections.)

In our technical document, we highlighted the transformation of data from Census 2000 to Census 2010 and given that change, we requested all providers to submit data in the Census 2010 format. Due to the change in census geography new data was requested whenever possible. Sanborn also uploaded the final data for each provider in NTIA format to the Sanborn Provider Portal. The providers were encouraged to use the provider portal and update their information on it. More and more providers are participating through the use of the provider portal and appreciate the ease of being able to see and validate their data through this process.

During our solicitation for data updates, we told providers that if we didn't hear from them by a certain date, we would default to using their data from Submission 3. However, we still contacted them a few times after the due date but eventually used Submission 3 data (converting it to census 2010) if they did not respond.

We followed the same contact and follow-up protocols as the previous submissions. The following are some of the important changes or no changes:

1. All census blocks and road segments are mapped based on 2010 census data set. Any data submitted in 2000 or 2009 format was converted to 2010 for this submission.

2. We continued to not collect data from resellers
3. We are submitting data for satellites in this submission based on NTIA clarifications. All satellite providers who have provided speed, FRN number and other technology information have been mapped to serve the full state. At present, Oklahoma received acceptable files from two companies (Hughes and Wildblue). We anticipate receiving coverage from the other satellite providers (Starband and Stratos) in our next delivery to NTIA (Submission 5, due to NTIA on April 1, 2012) both of which did not provide adequate attribute information in order to be included in this submission.
  - 1) Four satellite providers have been identified in Oklahoma – Hughes, Starband, Wildblue, and Stratos.
4. Due to NDA restrictions and our inability to accurately flag service by “category of end user”, address points are not included in this submission to NTIA for any commercial provider.
5. Some providers did not submit middle mile elevation or backhaul capacity, particularly when they asked us to reuse previous submission data. Wherever possible, we went back to providers to obtain that information, but it is not available for every record.
6. Terrestrial Mobile Wireless and Terrestrial Fixed Wireless (licensed and unlicensed) were treated as wireless coverage and were delivered as a shape file. In cases where a provider served using the same technology and spectrum but with different speeds, overlapping areas were removed and the higher speed was assigned.
7. If a cable based wireline provider provides both DOCIS 2.0 and DOCIS 3.0 service to the same area, the block or road was listed only once with a technology code of 40.
8. Providers were only willing to indicate on a general level if they served business, residential or both, so we did not get any providers that broke down the type of service by block. Only if the provider stated they only serve business to business customers did we fill in the “category of end user” with a code of 2, otherwise this field was left blank.
9. The submission 4 provider data model is currently based on the NTIA data model as of 6/30/11.

We added 6 new providers in this submission – SkyBeam/Jab Broadband (terrestrial fixed wireless unlicensed), AirLink Internet Services (terrestrial fixed wireless unlicensed), Cellular Network Partnership / CNP (terrestrial fixed wireless licensed), Ozark Telephone Company (aDSL), Seneca Telephone Company (aDSL),

and Dobson Telephone, McLoud Division / Dobson Telephone Company (aDSL).

In this submission, 41% of the providers submitted new or updated data whereas for 59% of the providers we reused data from their previous submissions. Jab bought Partnership Broadband and also bought out Rhino as of July 1, 2011.

### 1.1.2 Community Anchor Institutions Data

The community anchor institutions data continues to be crowd-sourced through the online data gathering application created by the Sanborn Team. The State of Oklahoma is undertaking the PR around this data collection and contacting the relevant agencies to request their participation in filling out the data survey. This has been a slow process and we are getting to a point of diminishing returns with this effort. The State of Oklahoma is also preparing to implement additional planning tasks to try to increase these numbers for future submissions. The current totals for community anchor institutions that have responded so far through this submission are provided below:

Category	Name	Total	Total with Broadband Information in Submission 4
1	School - K through 12	1956	112
2	Library	210	71
3	Medical/healthcare	444	29
4	Public Safety	1793	8
5	University, college, other post-secondary	79	16
6	Other community support - government	490	20
7	Other community support - nongovernmental	16	1

Broadband Mapping - Home

Please select the institution from the list. JONES LIBRARY

If you do not see your institution on the list, please select "Other".

Location Address of Institution (no P.O.Box): 111 E. MAIN

City: JONES Zip: 73049

Update Address on Map

Mapped Location: (Please be patient while map loads.)

If needed, use this tool to place the address point in the correct location on the map.

Does the Institution subscribe to Broadband Service at this location? Yes ☒ No ☐

Who is your Broadband Provider? Air Speed, LLC

What type of technology is used for your Institution's Broadband Transmission?

What is the DOWNLOAD speed advertised by your Broadband provider?

What is the UPLOAD speed advertised by your Broadband provider?

Are you currently physically located at the Community Anchor Institution address provided above? ☒

Yes ☒ No ☐

Updated By

Name: morgen

Organization: appgeo

Title:

Phone:

Email: test@appgeo.com

Please send me notifications about the Broadband Mapping Process.

Internet | Protected Mode

Community Anchor Institution: Crowdsourcing Portal

## DATA PROCESSING

### 1.1.3 General Overview

In general, submission 4 processes followed the same basic approach as previous submissions except for the conversion of Census 2000 to Census 2010. The following sections outline the modifications made to the initial processing in order to meet the submission 4 requirements as defined by NTIA.

In summary they can be divided into the following categories:

- Processing of Provider Data
- Conversion from Census 2000 to Census 2010 format
- Reference Data Creation
- NTIA Submission Data Model Schema Changes

### 1.1.4 Processing of Provider Data

All data received went through the following processing steps:

1. **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 row).
2. **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.
3. **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 provider submits a service boundary, we select all the blocks and roads inside that shape.
  - 2) If a provider submits a customer address list, the points are geocoded, and then the appropriate block or road segment is selected.
  - 3) If a provider submits block and road information using Census data, we just make sure everything is formatted to the appropriate specifications.
  - 4) If the 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, than 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) After each round of processing, we make sure that we only keep unique records. A unique record is defined as having a unique 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.
4. **QC Review:** All data is then sent to a different analyst to perform a through quality control review on the processed data set. Record counts are compared to what was submitted. The QC staff also made sure the ETL scripts and routines populated all of the right fields.
5. **QA Review:** Data is then sent to another team for Quality Assurance Review. In this step the data is not only double checked against what was originally submitted, but it is also brought up inside standardized MXD templates that allow us to make sure our results make sense. This often involves comparing the new data set with prior submissions, as well as looking for any possible technology or speed anomalies.
6. **Provider Review:** Processed data is all posted to a customized web-mapping tool we commonly refer to as the Provider Portal. All providers were notified once their data was available on the site, and they were always given 3-5 business days 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 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.
7. **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.
8. **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.
9. **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 NTIA receipt tool at this time. Any last issues are corrected, and the data is sent to the state for their review.
10. **Submission to NTIA.**

### **1.1.5 Conversion Process of Data from Census 2000 to 2010**

Due to the changes in census geography, all providers were asked to submit new data. In those instances when a provider A) submitted new data in Census 2000 format, or B) instructed us to reuse their last data submission, we had to convert the blocks and roads into 2010 format.

#### **Basic 2000 to 2010 Conversion Process:**

1. For the blocks, take the 2000 block ID, and select all the corresponding 2010 block id's
  - 1) using census crosswalk table – not an actual spatial process, since this was faster
2. Look at the new 2010 block ids, and filter on greater than or less than 2 sq. miles.
  - 1) If less than or equal to 2 --> bring in the 2010 geometry and add that record to the blocks table
  - 2) If greater than 2 --> select any roads in that area – spatial select (using roads gt2 table)
3. For the roads, take the 2000 or 2009 TLID and try to match it to the 2010 TLID's
  - 1) If there is a match, add that record to the roads table
  - 2) If there is not a match, select centroid of existing 2000/2009 segment, and select closest 2010 road
  - 3) If the road is now in a block LT2, select the block(s) instead and drop the road
4. Remove any duplicate records in both tables
5. Run some automated checks to catch missed features (i.e. add le2smi blocks surrounded by roads that have not already been added)
6. Manual review (QC) and corrections.
  - 1) There will be some blocks that are selected inappropriately (especially at town edges for CT providers, where we know their franchise ends at a town line.)
  - 2) There are some holes in the census crosswalk table
  - 3) The road conversion process may only select one portion of the road if it has now been broken into multiple segments

#### **Assumptions**

1. If a road was in an area greater than 2smi in s3, and due to census re-drawing, is now in an area less than 2smi, we will grab blocks (le2smi) on both sides of that road and add them to the provider data:
2. If a new 2010 block, that is less than 2smi, is completely surrounded by roads and/or blocks served by that provider, than we will add the block to the provider service area.

### 1.1.6 Submission 4: ReferenceData Creation

This section describes the reference data used in submission 4.

#### BLOCK REFERENCE

For s4, Census 2010 data was utilized. The data was set up as follows:

1. Block size (AREA) is calculated combining the 2000 land area (ALAND) and water area (AWATER)
2. AREA is converted from square meters to square miles to calculate square mileage (SMI).
3. 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 REFERENCE

2010 Tiger Line IDs (TLID) were used for data processing in s4. The data was set up as follows:

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

### 1.1.7 Submission4: NTIA Submission Data Model Schema Changes

The data model released on June 30, 2011 contained the following changes from the s4 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 is not required, and for many providers, was left blank since the data was not provided.

## DATA VALIDATION

Sanborn has continued to perform the same validation on the data as the previous threesubmissions(details in previous reports). Some minor updates to the validation process are discussed below.

- 1) QC of the data at various steps
- 2) Spatial checks against public and commercial datasets

- a. For OK, we continued to use the following datasets for validation:
  - i. Exchange Boundaries: for DSL boundaries
  - ii. MediaPrints: for Cable boundaries
  - iii. Speedtest.net data
- 3) Verification by providers
- 4) In this Submission, along with the standard verification by providers using the Provider Portal, we also identified for providers issues that they needed to focus on regarding the findings of our validation team. This also included validation and feedback we received through our website – this submission we have incorporated and integrated several feedback tools in the Interactive Map and information sourced from users is evaluated with respect to provider data and any noted discrepancies are passed back to the provider for correction. In addition, in this round, we incorporated any feedback provided by NTIA for Submission 3. All of these were done by sending providers a letter that identified issues using screenshots and explaining to them what the error was and then asking them to go fix those errors using the secure provider portal. If providers disagreed with the feedback, we have documented their response.
- 5) Speedtest data collection and other data collection for verification
  - a. We continue to use speedtest data and community anchor data crowd sourced for validation purposes.
- 6) Planning workshops and local validation
  - a. During this submission, local validation was undertaken by an independent group, the Center for Spatial Analysis at the University of Oklahoma (OU). OU performed an independent survey gathering data points from CAI's and the GIS community for the State of Oklahoma. Within Sanborn's validation process, OU's points were compared against provider's data. Those data points found in question were taken back to the providers for correction. OU is increasing their efforts to gather more data points and this process will be performed throughout Submission 5.