

Oklahoma Broadband Mapping

Data Submission Report *3rd Submission (April 1, 2011)*

April 1, 2011



Sanborn
1935 Jamboree Drive
Suite 100
Colorado Springs, CO 80920

Oklahoma Broadband Mapping

3rd Data Submission Report

Table of Contents

1	1 INTRODUCTION	3
2	OVERALL PROJECT STATUS	4
2.1	DATA COLLECTION.....	4
2.1.1	Broadband Data.....	4
2.1.2	Community Anchor Institutions Data	5
2.2	DATA PROCESSING.....	7
2.2.1	General Overview	7
2.2.2	Submission 3: Process Modifications	7
2.2.3	Submission 3: Reference Data modifications	8
2.2.3.1	Reference data sources	8
2.2.4	Submission 3: NTIA Submission Data Model Schema Changes	9
2.3	DATA VALIDATION.....	9
2.3.1	Data Validation Conclusions	10
3	APPENDICES	12
3.1	SANBORN QC_VALIDATION LETTER.....	12

1 Introduction

This report is submitted along with the third 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 WA. The State of Oklahoma has retained a mapping contractor, primed by The Sanborn Map Company for doing all work related to the Mapping Grant for this project.

This document is a supplement to the two previous reports submitted with data submissions 1 and 2 on May 1, 2010 and October 1, 2010 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. For this reason, it may be worthwhile to refer to the previous documents, if needed, for more details.

1 Overall Project Status

1.1 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 submission 3, Sanborn started data collection on January 10th 2011 by sending out data update requests and technical data specifications after NTIA did a Webinar announcing final changes for Submission 3. These were sent to a large list of companies which were compiled from FCC 477 list (dated December, 2009) and from a list provided by the Oklahoma UTC. The technical document highlighted the changes from Submission 1 to Submission 2 and requested incremental data only where 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. Providers are participating through the use of the provider portal and are getting used to the process.

Although we sent the technical specifications to all the providers (more than those on the FCC 477 list and many that were non-providers earlier including resellers and non-valid providers), we followed up actively with the providers on the 477 list or those who were already participating, and public providers such as PUDs (public utility districts) who were of strategic interest to the State of Oklahoma. This is because most providers outside of the FCC list were found to be non-providers of broadband.

During this round of the data update, many providers who had refused to participate in the program earlier expressed an eagerness to participate. This validates the importance of the program, not only for the purposes of the government, but also for the providers themselves.

In 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 2. However, we still contacted them after the due date a few times but eventually used Submission 2 data if they did not respond.

As with the second submission, we followed the following protocols:

1. We did not collect data from resellers
2. We have not collected data from satellite providers – we are in the process of formulating a strategy to map coverage from satellite providers and anticipate that we will have some coverage for satellite providers in our next delivery to NTIA (Submission 4, due to NTIA on October 1, 2011).

- 1) Three satellite providers have been identified in Oklahoma – Hughes, Starband, and Wildblue.
3. Affiliates, subsidiaries etc. have been counted as providers. Please note that data for these entities may or may not be reported as a separate FRN if they share the same FRN as their parent company.
4. We have not undertaken any propagation analysis for wireless providers who did not already have their own propagation maps. We are considering doing that for the next submission.

This submission process went smoother than previous submissions. There were a few minor issues that need to be resolved from previous submissions.

- 1) Spectrum: Larger providers are still not willing to provide separate polygons for different spectrums.
- 2) Communication with providers: It would help with data collection if NTIA/FCC held an open forum with the providers for changes that are being proposed for that data collection. This should happen before States start data collection and also providing all change information on an NTIA website to the providers so that they are not questioning the credibility of the request from States.
- 3) Information from NTIA: It would be very helpful to have information on changes in data model, requirements and specifications before the data collection is started. Ideally, in order to meet the next deadline of October 1 (for data good as of June 30, 2011), we would need to send out a data request to providers in the July 1-3 timeframe and giving them 3-4 weeks for preparing data and submitting it to us (given the holidays and the summer, it is important to give providers sufficient time to assimilate all data). Therefore, NTIA would need to get all changes finalized by June 30th so that we can hit the road immediately after that. This lead time allows us to provide more desirable time spans to the providers, and for us and the states to do the right amount of validation. As the process becomes smoother for everyone, we anticipate that this will happen more regularly in the future.

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 doing the PR around this data collection and contacting the relevant agencies to request them to fill in data. This has been a slow process and we are getting to a point of diminishing returns with this effort. The numbers of community anchor institutions that have responded so far is provided below:

Category	Name	Total	Total with Broadband Information in Submission 3
1	School - K through 12	1959	103
2	Library	209	66
3	Medical/healthcare	441	0
4	Public Safety	1793	8
5	University, college, other post-secondary	79	16
6	Other community support - government	489	18
7	Other community support - nongovernmental	15	1

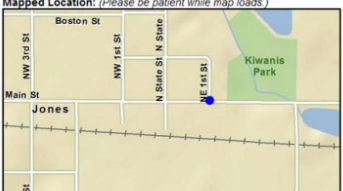
Broadband Mapping - Home

Please select the institution from the list. Jones County
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? [All 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: morgan

Organization: appgeo

Title:

Phone:

Email: test@appgeo.com

Please send me email updates about the Broadband Mapping Process.

Internet | Protected Mode

Community Anchor Institution: Crowd sourcing Portal

1.2 DATA PROCESSING

1.2.1 General Overview

In general, the submission 3 processes followed the same basic approach that was used in Submission 1 (s1) and Submission 2 (s2). As mentioned before, the submission 1 and 2 process documentation was included with those submissions and may be worth looking at for details if needed. The following sections outline the modifications made to the initial processing in order to meet the submission 3 requirements as defined by NTIA.

In summary they can be divided into the following three categories:

- Process Modifications
- Reference Data Modifications
- NTIA Submission Data Model Schema Changes

1.2.2 Submission 3: Process Modifications

Based on NTIA feedback and information provided in NTIA webinar sessions, the submission 3 data processing workflow was changed minimally to support the new NTIA submission requirements:

1. All census blocks are mapped based on 2000 census blocks. Any data submitted in 2009 format was converted to 2000 for submission. During processing a 'hybrid' census dataset (2000 IDs with 2009 line work) was used to take advantage of the improved 2009 line work. Prior to submission to NTIA, all features were mapped back to the 2000 census blocks. The Reference Data section below contains additional details.
2. For consistent representation the state road reference data used was 2009 Census Tiger Line IDs (TLIDs). Other data sources (non-TLID features, or 2000 TLID features) were mapped to 2009 TLID features.
3. Overview was removed completely from submission data due to the fact that all maximum advertised up/down speeds are being reported in blocks, roads, and wireless features.
4. Due to our NDA restrictions, address points and last mile points will not be submitted to NTIA. As mentioned before, Qwest requested that their address points be submitted to NTIA for blocks greater than 2 square miles. However, they could not provide the end user category and hence this data was not submitted but reprocessed data (address points reprocessed to street segments) are being submitted.
5. Some providers did not submit middle mile elevation. Wherever possible, we went back to providers to obtain their middle mile elevation information.
6. Terrestrial Mobile Wireless and Terrestrial Fixed Wireless (licensed and unlicensed) were treated as wireless coverage and were delivered as a shape. In cases where a provider served the

- same technology and spectrum with different speeds, overlapping areas were removed and the higher speed was assigned.
7. The submission 3 Provider data model is currently based on the NTIA data model as of 1/13/2011.

1.2.3 Submission 3: Reference Data modifications

This section describes the reference data used in submission 3.

BLOCK REFERENCE

For s3, a hybrid block dataset (2000 IDs with 2009 line work) was used to take advantage of the improved 2009 geometry. The data was set up as follows:

- 2009 BlockID suffix is dropped and the blocks are dissolved (by Block ID) to produce data with 2000 BlockIDs and 2009 shape geometry
- Block size (AREA) is calculated combining the 2000 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 REFERENCE

To take advantage of the 2009 geometry improvements, 2009 Tiger Line IDs (TLID) were used for data processing in s3. Any non-2009 TLID (i.e. 2000 TLID or other) submitted by providers were mapped to the 2009 reference data. The data was set up as follows:

- The GT2SMI (Greater Than 2 Square Mile) indicator is set to True when:
 - o The 2009 road segment is completely within a hybrid 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.

OVERVIEW REFERENCE

This dataset was dropped completely for this submission.

1.2.3.1 Reference data sources

The following data sources were used as reference data sources for submission 3:

BLOCK REFERENCE DATA: 2009 CENSUS BLOCKS

No changes from previous submission.

ROAD REFERENCE DATA: 2000 CENSUS TIGER LINES

No census 2000 TIGER line data were used for this submission.

ROAD REFERENCE DATA: 2009 CENSUS TIGER LINES

No changes from previous submission

OVERVIEW REFERENCE DATA: 2009 CENSUS COUNTIES

This data has not been included in Submission 3

1.2.4 Submission 3: NTIA Submission Data Model Schema Changes

The data model released on January 13, 2011 contained the following changes from the s2 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. State of Oklahoma Public Utility Districts or PUDs), the ProviderType is considered “Other” (value = 3)
- In the CAI feature class, the field BBService has been 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 values must be changed to code u (for Unknown) – nulls are not allowed.
- Three new fields have been added to the CAI feature class. Wherever possible, these values have been populated in the CAI data.
 - Public Wifi (Y, N, or U)
 - URL
 - CAIID

1.3 Data Validation

Sanborn has continued to perform the same validation on the data as the previous two submissions and listed below (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 WA, 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
 - a. 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 was done by sending them 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. A sample of a letter is provided in Appendix A in this document. This helps by making this process a little more targeted for the providers and allows them to hone into issues.
- 4) Speedtest data collection and other data collection for verification
 - a. We continue to use speedtest data and community anchor data crowdsourced for validation purposes.
- 5) 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 continuing to gather data and this process will be performed throughout Submission 4.

1.3.1 Data Validation Conclusions

We continue to believe that we do not have sufficient information to alter provider data and we have been careful not to do so unless there are obvious errors such as incorrect block numbers, or unidentifiable street segments, etc.

Data validation involves working with providers to improve the data and we are dealing with issues as they arise. This activity continues to be a challenging activity. There is no complete truth sometimes and different pieces of evidence are collected and pieced together to point discrepancies that are explored in more detail. Commercial datasets are of limited value and often self-reported by the companies and subject to the same errors that we get from providers directly, and sometimes

exaggerated by the fact that there are different vintages and resolution and hence the comparison is not easy. Speed test locations are also sometimes incorrect and similar issues exist with all crowd-sourced data.

There is no absolute truth exists and that data validation cannot change data arbitrarily based on only one evidence or two. Hence it takes a long period of time to fully address a reported issue.

2 Appendices

2.1 Sanborn QC_Validation Letter

March 9, 2011

**Broadband Mapping Services
State of Oklahoma**

Re: QC and Validation of Provider Data

Dear Provider:

As part of the Broadband Program, the Sanborn Map Co. is performing a QC and validation of the data received from you by comparing your data to publicly and commercially available broadband datasets. This includes exchange boundaries for DSL, MediaPrints for Cable and Fiber and others as deemed necessary. We are also using Speedtest.net data for some speed validations.

If you are receiving this notification, it is because we have found certain issues that need your assistance. Screen shots of the issues are provided below along with a table denoting the issue found. We would appreciate it if you would please review these issues quickly and go to the provider portal and note the correction that needs to be made since we need to finalize your data to be submitted to NTIA.

If you need any further clarification after reviewing the issue, please contact Bridget Marcotte at (503) 228-8708 x 306. Please note: if we do not receive a response from you with what correction needs to be made, Sanborn reserves the right to change the data if needed.

Thank you very much for your assistance providing answers on the issues noted below.

Sincerely,
The Sanborn Broadband Mapping Team


QC and Validation Issue(s) Encountered

Please make all corrections on the provider portal link provided below. For confidentiality, your login and passwords were sent during the last submission, in another email.

<http://beta.appgeo.com/OklahomaBroadbandProviderPortal/>

Issue found:

Issue Category	Description/Screen Shot
Part of the data is extending outside Media Prints boundaries	
Part of the data is	

extending outside of Telephone Exchange boundaries	
Spatial Outliers - data which is off by itself and not consistent with other data spatially	Areas within the red circles are examples of spatial outliers in your data 
Independent Validation point showing there is NO service in this area	Block Numbers Affected:
Middle mile has missing/invalid elevation	
Invalid Max Advertised Speeds	