July 16, 2018

National Telecommunications and Information Administration (NTIA)
Attention: Mr. Douglas Kinkoph, Associate Administrator
U.S. Department of Commerce
1401 Constitution Avenue NW
Room 4887
Washington, DC 20230

Re: Improving the Quality and Accuracy of Broadband Availability Data
        [Docket No. 180427421-8421-01]

Dear Mr. Kinkoph:

Connected Nation, Inc., respectfully submits the following comments in response to the Notice and Request for Comments on improving the quality and accuracy of broadband availability data. We believe that accurate, robust, and reliable data is vital to both policymaking and decision-making, as it enables the targeted investment of federal, state, and local funding to promote broadband build-out in unserved and underserved areas. We commend the NTIA for taking steps to collect third-party datasets and enlist the public in efforts to refine and update the FCC's National Broadband Map, which is currently based on data submitted via the FCC's "Form 477" reporting process. However, we believe that until more granular and accurate data on service availability is collected directly from the service providers themselves, efforts to improve the accuracy and reliability of the National Broadband Map will ultimately fail to make a significant difference. Third-party datasets, public feedback,

and in-field audits can and should be used as validation tools, but only after a more granular primary source data—at the street or land parcel level of detail—is collected and analyzed.

We are aware that concerns have been raised in the past that more granular data cannot or should not be collected, either because doing so could present an increased burden (and unfunded mandate) on providers, or because the collection of such data would require government access to infrastructure and/or subscriber location data that the government has no business possessing. We agree that these are legitimate concerns but posit that a viable solution exists that would yield a more granular understanding of service availability, while protecting the confidential and proprietary nature of the data that would be used to produce it—a solution described in greater detail below. A piecemealed approach to broadband mapping that does not allow for the collection of better primary source (provider) data will not achieve the end goal of comprehensive, reliable, and granular mapping. We look forward to sharing our ideas and our experiences in the following comments.

Connected Nation's Background

Connected Nation, Inc. (CN) is a national 501(c)(3) non-profit with a 17-year history of working to close the Digital Divide across the United States. Since 2004, we have mapped broadband coverage in 17 states and Puerto Rico, assisted the Universal Service Administrative Company (USAC) in analysis and validation efforts for Phase I of the Mobility Fund, completed mobile drive testing and analysis for the Ohio Department of Transportation, continued to map and validate broadband coverage for a number of states after the conclusion of the federally funded State Broadband Initiative (SBI), and worked with hundreds of local communities on technology action plans focusing on the access, adoption, and use of broadband, particularly in rural areas.

Additionally, we were selected as the designated entity or state agency subcontractor to map broadband availability for 12 states and Puerto Rico during the State Broadband Initiative (SBI) grant program that NTIA administered from 2010 through 2015. As you are likely aware, during this program, we routinely interacted with well over 1,200 unique broadband providers—developing a deep understanding of the challenges related to data collection. In these endeavors, several of our broadband mapping and field validation methodologies were recognized as best practices by the NTIA and were adopted by other states. Furthermore, we have worked locally to bridge the divide between communities and service providers, engaging over 300 communities in broadband strategic planning. We also supported 7 of the 14 winning Lifeline Program pilots in 2012.

Given our experiences, we believe we possess a unique perspective from which to comment on broadband data collection and mapping issues, and we present the following answers to questions 1, 3, 4, and 5 posed in your Notice and Request for Comments.

1. Identifying Additional Broadband Availability Data

> a. *What additional data on broadband availability are available from federal, state, not-for-profit, academic, or private-sector sources to augment the FCC Form 477 data set?*
> b. *What obstacles—such as concerns about the quality, scope, or format of the data, as well as contractual, confidentiality, or data privacy concerns—might prevent the collaborative use of such data?*

Since the conclusion of the SBI Program in 2015, CN has collected broadband availability data in Iowa, Michigan, Minnesota, Nevada, Ohio, and parts of Florida, Pennsylvania, South Carolina and Texas, and will likely resume that work in an additional state later this summer. While the data we have collected in these states is more robust, granular, and reliable than what is generally available at the federal level, the non-disclosure agreements (NDAs) we have executed with service providers govern the disclosure of data supplied to us,

as well as the derivative works created from the collected data—restrictions which would limit what could be shared with the NTIA. At the time that these NDAs were executed, it was not anticipated that the Federal Government might have a need to receive such information.

To better understand this challenge, it is important to understand the process that we undertake to map broadband availability at the state level. Confidential provider infrastructure data—and in some cases, subscriber location data—is collected under NDA, analyzed, and processed to create a derivative service area map illustrating broadband service availability by technology type, speed of service, provider name, and serviceable location. The resulting broadband map (made available in both an online interactive format that is searchable by address, as well as a static downloadable PDFs), is published in the aggregate by technology type (e.g., DSL, cable, fixed wireless, fiber to the premises, etc. as individual data layers) based on the information submitted by all known and participating providers and detailed estimates created for non-participating providers. The maps are not currently published on an individual provider-by-provider basis, as the terms of the executed NDAs would prohibit such individual disclosure. However, the interactive map allows a user to query a specific address or latitude/longitude to identify a list of providers that the data indicate serve that specific location. For example, a user can query the map to identify which providers can serve 1313 Mockingbird Lane with 25 Mbps downstream/3 Mbps upstream service and see that AT&T, Spectrum, Munster Wireless, and Google Fiber can provide such service at that address. A user cannot, however, see Munster Wireless's entire individual coverage footprint but rather the aggregated granular footprint for all fixed wireless service providers statewide.

These limitations are due to provisions in typical NDA instruments that allow for the exchange of competitively sensitive or proprietary information and the expectation of confidentiality of this data. Such NDAs are generally mission-critical in states where provider participation in the mapping process is voluntary and/or where there is no state law that

governs the disclosure (or protection thereof) of such information. Providers also typically have concerns about sharing detailed infrastructure and specific subscriber location information, especially with government agencies, out of concern that such information could be used against them in regulatory proceedings or in the competitive marketplace. The standard NDA that CN makes available to providers offers the option for providers to mark any data shared as confidential, as long as service areas derived from the confidential data can be aggregated with other provider data to create a representation of where broadband service is available, and more importantly, where it is not available. We believe that future NDAs at the state level could evolve to allow for the derivative analyses on service availability to be provided to the FCC or the NTIA while still protecting the more sensitive aspects of the data obtained from providers.

Another obstacle that will confront the NTIA in collecting and utilizing state and commercial third-party datasets is that such data is inherently produced via an array of methodologies. In our comments to the FCC on the Form 477 process last September[1], we strongly recommended that any future evolution of the national broadband map should be based on primary source data that is uniformly reported and analyzed according to defined technical standards, and that these standards are enforced across all 50 states, the five territories, and the District of Columbia as a means of minimizing inconsistencies. The acquisition of third-party datasets as instruments to refine Form 477 census block level data will have limited utility, to the degree that such datasets were not produced under uniform standards across the entire country.

With that said, we believe third-party datasets would have great value in validating more granular primary source data that should be obtained from service providers moving

---

[1] https://ecfsapi.fcc.gov/file/109142903105183/CN%20477%20Comments_Final.pdf

forward. Such datasets could highlight areas that warrant further investigation and refinement of provider-reported data, and could be the basis on which it is determined where in-field validation activities/audits are conducted. In-field validation, involving the deployment of network engineers to verify service availability, is a costly exercise and could be done most efficiently if those resources are dispatched primarily to areas where third-party data and crowd-sourced public feedback indicate potential problems or inconsistencies with provider-reported data. Ultimately, this is the best approach to finally achieve a truly accurate and granular broadband map over time: obtaining granular provider data (via a "clearinghouse" process described below) that is compared against third-party datasets and crowd-sourced public feedback—the results of which trigger in-field validation activities in areas that warrant further investigation. Once the investigative process is complete, the map would be refined and the cycle would begin anew with the next reporting period.

Until Congress establishes a pathway for this process to take place, we commend NTIA's efforts to explore ways to improve the FCC's existing Form 477 dataset. We believe the first step in the process should be an effort to quantify and specify the degree to which the Form 477 data are currently overstating the extent of broadband availability. As you know well, Form 477 requires providers to report census blocks where they provide service. Unfortunately, if even one household in a given block is served, the entire block is considered as having service, resulting in a significant overstatement of availability. This is particularly problematic in rural areas where census blocks can be very large. There are over 3,200 census blocks in the country that are larger than the entire District of Columbia (which is about 68 square miles in size), and five blocks that are larger than the entire state of Connecticut. Yet these are the areas where broadband availability is most lacking and needs to be most accurately defined.

If NTIA must prioritize areas where overstatement is most likely to occur, we recommend focusing on blocks that are larger than ½ square mile in area. These blocks comprise 10.68% of all census blocks in the country, but 87.92% of the nation's land area, and encompass 12.76% of the population, or about 39.4 million people.

Recently, Connected Nation conducted a predictive analysis, or confidence rating, of areas where Form 477 data are potentially overstating service availability. This confidence rating process involved the development of a scoring algorithm in which census blocks were scored based on criteria that, in our experience, indicate a lower likelihood of service availability to every household in a given block. The scoring algorithm was based on factors like the size of the block, the household density of the block, the number of providers reporting service in the block, the type of broadband technology reported in the block, and other factors. The result was the creation of color-coded map of Form 477-indicated service availability (Appendix A), with census blocks shaded red, yellow, or green as an indication of the likelihood that service is available to every household within a given block—with red defining areas where we have low data confidence, yellow indicating medium confidence, and green indicating high confidence. This map, and its corresponding dataset, are an example of third-party information that could be made available to NTIA and other federal agencies to prioritize the areas where the existing Form 477-generated National Broadband Map is the least accurate.

3. New Approaches

Are there new approaches, tools, technologies, or methodologies that could be used to capture broadband availability data, particularly in rural areas?

We believe that the creation of an accurate National Broadband Map will only be possible if the following process can be achieved:

1) collecting granular (street address or land parcel-level) data directly from service providers (infrastructure data, subscriber location data, or other geocoded information) on service capabilities;

2) analyzing the submitted data to create footprints of service availability at the street address or land parcel level of detail;

3) comparing those derivative footprints of service availability against third-party datasets and crowd-sourced public feedback to determine areas that warrant further investigation;

4) conducting in-field validation in the areas that warrant further investigation to determine refinements that must be made to the service availability footprints; and

5) updating the map based on those refinements.

As the aforementioned process is completed each reporting cycle—which should be every six to twelve months—the National Broadband Map would become more and more accurate over time.

Unfortunately, there is currently no mechanism in existing law that would allow for this process to take place. Broadband service providers are reluctant to supply infrastructure or subscriber location data directly to a government agency in a way that would allow the agency to generate a derivative, more granular broadband map. And alternatively, any requirement to collect street-level or parcel-level service availability footprints in native GIS format directly from service providers would amount to an increased reporting burden and unfunded mandate that would particularly impact smaller service providers that may not have internal GIS capabilities. Additionally, there is no current integrated mechanism for directly collecting public feedback regarding specific questionable areas on the map, and no mechanism for

conducting field validation activities in areas that warrant further investigation based on public feedback and third-party dataset comparisons.

We believe that a solution exists to address providers' data submission concerns and provide a mechanism for public feedback, data validation, and map refinement. We have called for the creation of a neutral, independent, third-party broadband data clearinghouse—operating under the FCC's or NTIA's oversight—that can act as a firewall for the collection and analysis of provider-submitted confidential data, outputting from that process an understanding of service availability, speed, provider name, and technology type at the street address or land parcel level of detail. Such a clearinghouse could provide assistance to carriers that do not have advanced reporting capabilities internally—minimizing reporting burdens while yielding the government the same type of information it receives via Form 477 today, only at a much more granular level of detail.

The clearinghouse should also have the responsibility of receiving and processing public feedback, implementing field validation and audit processes, and tracking where federal investments have been made to improve access. We believe the establishment of such a clearinghouse will serve the public interest by informing federal decision-making on infrastructure investments moving forward, thus ensuring accountability for those dollars as they are spent. And while the clearinghouse will have responsibility for creating GIS shapefiles for use on the National Broadband Map, the function of designing, operating, and maintaining the map itself can still rest at the FCC or NTIA. The clearinghouse would simply serve as a conduit for collecting, analyzing, processing, and validating the broadband data that underlies the map.

4. Validating Broadband Availability Data

a. *What methodologies, policies, standards, or technologies can be implemented to validate and compare various broadband availability data sources and identify and address conflicts between them?*

We believe that the broadband data collection process as a whole should involve a four-pronged approach that will increase the granularity and accuracy of the map that is produced:

1) Direct Provider Engagement: Working directly, in the spirit of collaboration, with individual broadband providers on the data and information they possess regarding their network capabilities and service offerings is the most beneficial way to begin a broadband mapping process. While some providers do not have the resources or GIS capabilities to produce and/or maintain a location-based set of information on their exact service footprint, we have found that a professional two-way relationship with the providers, with the express intent to avoid increasing their reporting burdens, yields solid baseline data to jumpstart a mapping process.

2) In-Field Data Collection and Validation: Telecommunications engineers deployed in the field are able to refine broadband service areas, test networks, catalogue, photograph, and map infrastructure assets, and validate new and disputed service areas. If small service providers do not have the resources to actively participate in the data collection process, field engineers can work with them to develop service area maps based on infrastructure locations, propagation modeling, and other techniques to produce service availability footprints with significant accuracy and granularity.

3) Desktop Research: Research and review of current federal filings and databases, spectrum licenses, broadband service advertisements, permits, and other available

information can provide another level of validation that can help direct field validation activities to refine broadband data. In many rural areas, such research can help resolve ambiguous, confusing, or conflicting information.

4) <u>Consumer Feedback</u>: Once a broadband map is published, collecting public feedback is essential to the ongoing refinement of the map, as described at length above. At the local level, such feedback—when permission is granted for it to be shared—can also showcase where demand exists in or near a provider's service territory—information that can be helpful in fostering better relationships with local providers while helping to close coverage gaps.

b. *Do examples or studies of such validation exist?*

Yes, absolutely. In one recent example, Connected Nation undertook a detailed broadband data collection, validation, and expansion feasibility study for the government of Walton County, Florida. We began by engaging local service providers in a discussion about the accuracy of the FCC's Form 477 data for the county to determine areas of overstatement, as well as areas where services exist but were not being reported. Those discussions yielded the discovery of valuable information—that additional broadband providers (not reflected in the Form 477 data) were indeed offering service within the county, while others were reporting service where none actually existed. Armed with this information, our network engineers then spent several weeks driving every publicly-accessible road in the county—cataloguing and photographing telecom infrastructure assets and determining the true extent of each service providers' capabilities. The attached map (Appendix B) is the result: a street-level service availability map indicating fixed services available at 25 Mbps /3 Mbps (in purple) overlaid on top of the FCC's Form 477 availability data for the county (in gold). While there are areas of understatement—areas where service actually exists that were not reported via Form 477—the

larger takeaway is the significant overstatement of service overall. In just this one county, Form 477 data indicated that approximately 42,938 people have the ability to access 25/3 service, while in actuality, only about 33,822 people have such access—a net overstatement of approximately 27%, or 9,116 people.

c. *What thresholds or benchmarks should be taken into account when validating broadband availability, such as bandwidth, latency, geographic coverage, technology type, etc.? How can conformance to such standards be used to evaluate the accuracy of broadband data sets? How could those standards be used to improve policymaking, program management, or research in broadband-related fields?*

Standards should include, when possible, crowd-sourced network speed tests as one validation tool (among many) to gauge whether maximum advertised service speeds are actually being achieved within a given service area. Aggregated testing data should be compared against the speeds reported by a service provider as available in a particular area. Because individual tests can be affected by factors such as network interference or poor signal strength on a user's Wifi network, for example, it is important to collect a statistically significant sample size of speed test data. Additionally, speed test reliability may be negatively impacted if few or no customers in a given area are actually subscribing to a provider's maximum available speed, as plans that are limited to slower speeds would artificially skew test results.

Robust infrastructure validation, as described in the Walton County, Florida, example above, is also key in determining the validity of a reported service territory. Identifying outside plant (OSP) and critical equipment (e.g., DSLAMs, fiber nodes, remote terminals, etc.) can often greatly improve the quality and accuracy of provider-reported data. Simply put, where fixed wireline OSP does not exist, neither does the service (regardless of whether it is fiber to

the premises, DSL, or cable modem service). Regarding fixed or mobile wireless services, the lack of fiber or microwave backhaul infrastructure and/or vertical assets in a given area provides a strong indicator that service may not be as widely available as advertised. Seasoned telecom engineers are trained to spot and identify such infrastructure and are an invaluable resource to any serious validation effort. Because of the cost involved in deploying such personnel on the ground, third-party datasets and crowd-sourced public feedback can be used to effectively target their efforts.

Latency has long been discussed as a key performance indicator related to broadband services. It certainly impacts service quality—particularly on real-time interactive applications such as two-way video communication. Historically and even today, however, it has a far smaller overall impact on user experience than other infrastructure-related issues that impact networks, such as backhaul capacity. With that said, and with an eye toward the future, there will be an increased need for ultra-low latency connections to support IoT applications that will continue to evolve as 5G mobile technology becomes a reality. Therefore, it will become increasingly important to measure latency as well as bandwidth as coequal factors in the evaluation of broadband accessibility.

5. Identifying Gaps in Broadband Availability

  a. *What data improvements can the government implement to better identify areas with insufficient broadband capacity?*
  b. *What other inputs should NTIA seek to inform data-driven broadband policy- and decision-making?*

CN's comments to the FCC regarding its Form 477 data collection in September 2017 identified key issues with the current reporting process. We believe any future data collection and mapping effort must address these challenges, and the best way to ensure a granular and accurate map is not to piecemeal the approach, but rather to offer a consolidated solution that takes into account all the relevant factors confronting the problem: the need for increased

granularity and accuracy of the maps, the need for validation of reported data, provider industry concerns with regard to reporting at a more granular level of detail, and the imminent need to more effectively guide the expenditure of federal grant and universal service funds.

Other issues with the current Form 477 process include:

1) Form 477 requires providers to report census blocks where they provide service. Unfortunately, if even one household in a given block is served, the entire block is considered as having service, resulting in a significant overstatement of availability. This is particularly problematic in rural areas where census blocks can be very large—some being larger than the entire state of Connecticut. Yet these are the areas where broadband availability is most lacking and needs to be most accurately defined.

2) Since some providers rely on third-party vendors to compile Form 477 data and the filings are primarily in .csv (comma-separated values) format, providers that do not have GIS capabilities have no way of visualizing their service territories to ensure accuracy, resulting in overstated and understated reporting.

3) Some known providers from the SBI Program years, including some that have been confirmed as still in business and operating as broadband providers, are simply missing from the Form 477 datasets, meaning that they are likely not filing as required.

4) Fixed wireless coverage during the SBI Program years (when properly mapped) was developed from propagation modeling based on tower locations, predicted signal levels, and other technical parameters. Under Form 477, however, fixed wireless coverage is reported by census block as any other type of fixed service, indicating areas as served where there may actually be no service for miles.

It should also be noted that missing data and inaccurate filings may also have the effect of understating service capabilities—putting the providers themselves at risk for overbuild, since Form 477 data (and the FCC's National Broadband Map that is produced from it) is the only viable resource at the federal level to direct subsidies and grants toward areas lacking adequate broadband.

Taking into consideration these lessons learned, we recommend the following:

1) Any future mapping effort must prioritize the accuracy and granularity of the maps themselves to ensure that the nation's broadband landscape is fully understood at the street address or parcel level of detail. Census block data is not sufficiently granular as we look to solve the broadband gap in rural and other insular areas of the United States.

2) Obtaining the sufficient level of granularity requires the protection of providers' proprietary and confidential information. Such protection is needed to safeguard critical infrastructure from vandalism, sabotage, or worse, and to preserve the confidentiality of competitively sensitive infrastructure and subscriber information, which should remain closely held.

3) Any future mapping effort must be premised on uniform technical standards in reporting to eliminate the inconsistencies inherent in state-by-state reporting. Such uniformity in reporting will provide decision-makers the high level of confidence needed to more effectively target federal funding for broadband deployment projects.

Conclusion

We appreciate the opportunity to provide our perspective on these issues in response to this Notice and Request for Comments. We have long enjoyed a strong working relationship with NTIA and look forward to our continued engagement with you in the years to come.

Thank you for your commitment to improving the quality and accuracy of broadband availability data, which is truly key to closing the Digital Divide once and for all. Like you, we believe everyone belongs in a Connected Nation.
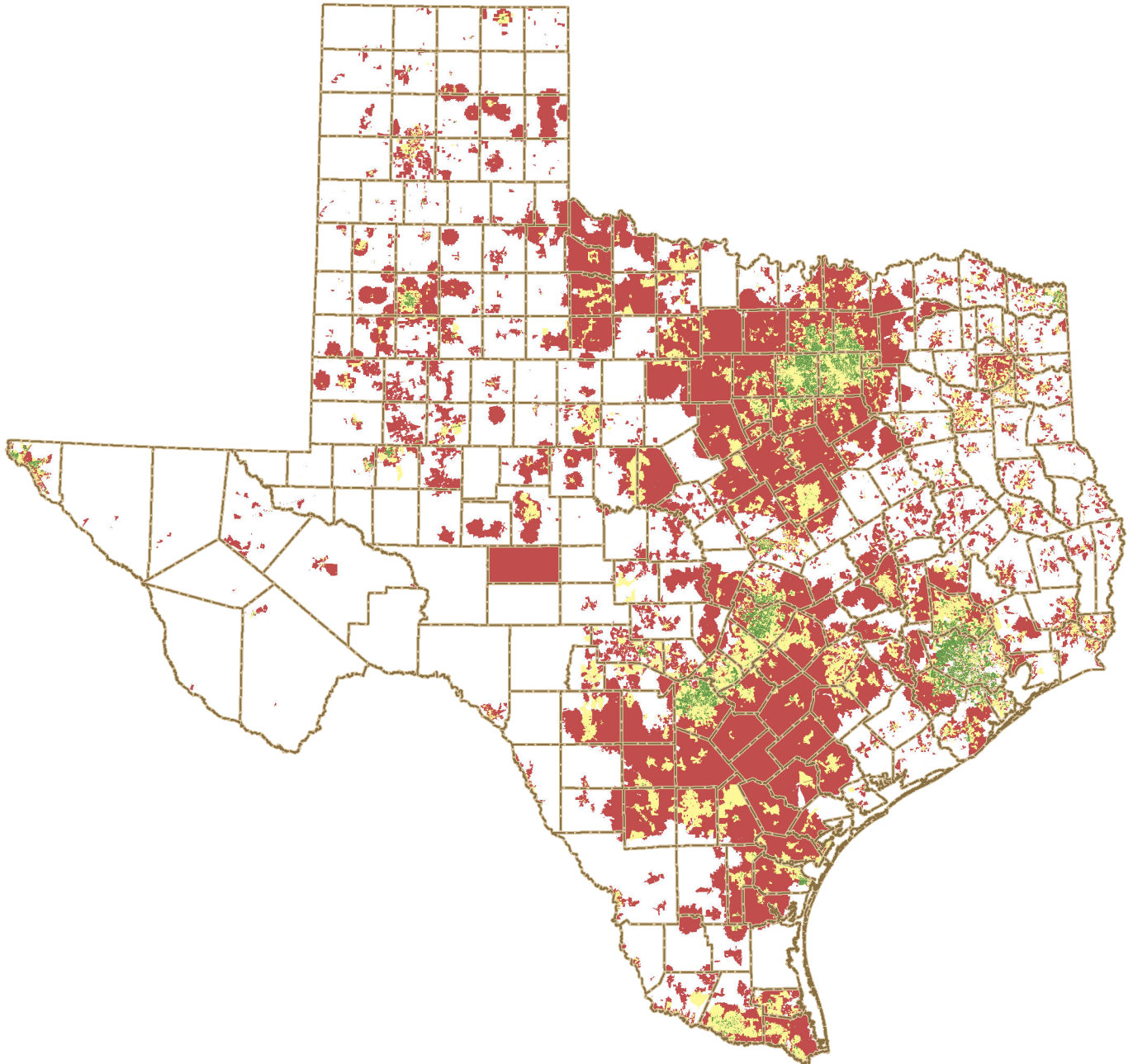
Respectfully submitted,

Thomas W. Ferree
Chairman & Chief Executive Officer
Connected Nation, Inc.

CONNECTED
NATION ℠

Texas

Texas Broadband Service

Fixed Broadband at Least 25 Mbps Download/3 Mbps Upload
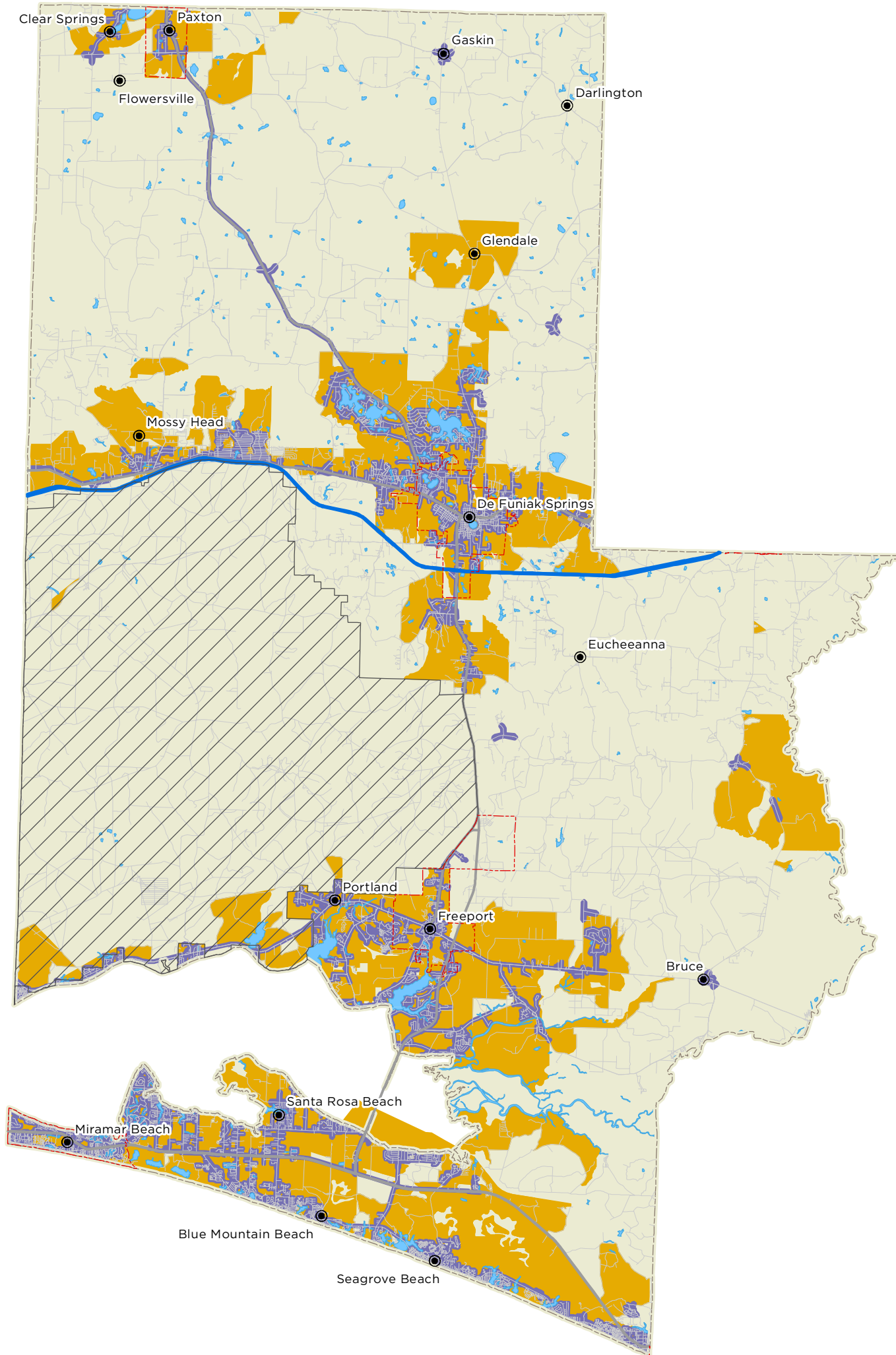

Published June 14, 2018

A low confidence rating does not mean that CN believes there is not broadband available in that census block or that the provider-filed data is incorrect, only that it warrants investigation to confirm whether or not broadband is available to all locations within the census block.

Legend

🟥 Low Data Confidence

🟨 Medium Data Confidence

🟩 High Data Confidence

Data Source: FCC Form 477 Broadband Deployment Data as of December 31, 2016, released November 16, 2017.

Broadband Service Comparison to FCC Form 477 25 Mbps Download/ 3 Mbps Upload

Walton County
Florida

Published May 8, 2018

CONNECTED NATION℠

Clear Springs
Paxton
Gaskin
Darlington
Flowersville
Glendale
Mossy Head
De Funiak Springs
Eucheeanna
Portland
Freeport
Bruce
Santa Rosa Beach
Miramar Beach
Blue Mountain Beach
Seagrove Beach

Symbology

◉ City	🌊 Water
▬ Interstate	⬚ Eglin AFB
▬ US Road	⬛ CN Validated Broadband
▬ Local Road	⬛ FCC Form 477 Broadband
▭ Municipal Boundary	⬛ Unserved Areas

N

0 1 2   4   6   8
Miles