DOC-NTIA-OTIA (US)

Transcript

Moderator: Karen Perry
February 17, 2021

0:17

Karen Perry: Hello, and thank you for joining us today for Broadband USA's Monthly Webinar.

0:22

I am Karen Archer Perry, Senior Policy Analyst for NTIA's Broadband USA Program. As a reminder, this webinar is being recorded, and it will be posted on the Broadband USA website after the program.

0:37

Today's webinar focuses on data as a foundation for broadband planning, but before we get into that topic, which is one of my favorites, I'd like to tell you a little bit about the webinar that we will be offering next month at this time.

0:52

Our March 17th webinar will provide an update on the NTIA Grant Programs in the Consolidated Appropriations Act of 2021.

1:02

Congress passed year end Omnibus legislation, that included several broadband and connectivity expansion bills, and next month, NTIA staff will provide the latest information on three new grant programs, including the Tribal Broadband Connectivity Grants, the Broadband Infrastructure Deployment Grants, and the Connecting Minority Communities Pilot Grant program.

1:31

Please join us next month at this time, to learn more about these exciting, new grant programs.

1:37

As I said, today's program will focus on data as a foundation to broadband planning, but the real focus of this session isn't just on the data, it's on tools that you can use to access the data and to use it as part of your broadband plan. You can use it as part of a business case to start conversations with local leaders as to create partnerships and yes, even as part of brand applications.

2:05

I'm joined today by two industry experts in Broadband Testing, Measurement and Analysis, Lai Yi Ohlsen, Project Director at Measurement Labs, and Bryan Darr, Vice President of Smart Communities at Duke Law.

2:19
Before we begin, I'd like to review the details for today's webinar.

2:24
First, we'll open the webinar for questions after the completion of the presentation. As you hear from each presenter, please put your questions or your comments in the question box on the right-hand side of your screen.

2:39
Second, the presentation, transcript, and the video of today's recording will be available on the Broadband USA website within seven days of the webinar under the event's Past Events tab.

2:53
And finally, please do visit the Broadband USA website for information on our programs, including useful guides, products, publications, and other tools that can assist you in your broadband planning, funding, and implementing your Broadband project.

3:14
Broadband planning teams generally come together around a common purpose, to address a gap, or to achieve a vision for better broadband, either to sustain or transform their community.

3:26
There are few key steps involved in starting a broadband plan: Assemble a team, engage stakeholders, conduct an inventory of community assets. You can always jumpstart your broadband plan, and learn a lot from publicly available data.

3:44
Today, what I'm going to do is talk about three of the major, some of the major sources for publicly available data.

3:52
I'm going to highlight some federal tools, as well as some non-federal tools that are available, but really, what I'm going to do is focus on how you can use those tools to add your local insights and use those tools for your broadband plan.

4:11
Three major sources for federal, three major federal sources for broadband data are the American Community Survey, the NTIA Internet Use Survey, and the FCC Form 477 data.

4:24
The first two of those are consumer surveys, where we ask people their opinions and their insights on broadband usage and Internet use and computers.

4:34

The third one is a regulatory dataset, where the FCC as a regulator collects information from internet service providers.

4:43
I'm only going to cover a little bit of information on the first one, the American Community Survey.

4:49
Not because I don't love them all, but really, because we only have a little bit of time today.

4:57
The American Community Survey is a marvelous survey. What the US Census Bureau does, is it surveys three point five million households every year over that, and it does it on a continuous basis, such that, over the course of five years, they've reached 17 million households.

5:17
And by reaching 17 million households over the course of five years, they have a sample size that's large enough for them to make estimates of all 3142 counties in the country of tribal areas of population areas as small as 20,000 people, and of geographic areas known as small as the tract, or a Census block, or a block group.

5:41
What that means is, we can get estimates that are almost at the neighborhood level, and since they do this on a continuous basis, we get fresh data every year, and this survey set is large.

5:57
They cover about 35 different topics: age, race, ethnicity, income, poverty level, number of children in a household, veteran's status, housing type, education level, employment, industry, commute, and they ask information about computer ownership, Internet subscription, cell phone use.

6:19
And our most recent dataset just came out in December of last year, covering the years from 2015 to 2019. So it's very, very recent.

6:31
Since it's a consumer survey, we have to kind of look at how the questions are asked in order to understand the data.

6:38
So they asked, Do you have a computer or not? Do you subscribe to the internet? Do you have a wireline subscription, such as cable, fiber or DSL? Do you have a cell phone? Do you have a cell phone data plan? Do you have satellite? Do you have dial up? Do you have no Internet subscription?

6:57

So, let's look at how that might look in a community.

7:01
We've been working with Idaho County in Idaho.

7:04
So, if you look at Idaho County, if you ask do you have broadband of any type, 69% of the county answers that they do have broadband of any type that compares to a 2% of people in the US, so a little bit less.

7:21
But if you say do you have wireline broadband, cable, fiber, or DSL?

7:26
In Idaho County, it's 37%, much less than the national average of 68.9%.

7:33
And the reason is because many more people use satellite there, 29.3% compared to 6.3% in the US.

7:44
Do you have cell phone as a primary connection, using your cell phone, as your primary connection?

7:50
Only 7.1% compared to 10% in the country?

7:55
And do you have no internet access? It's double, but it is in the rest of the country, 25%, compared to 13.9%.

8:04
I also want to point out the margin of error there. If you look then, at that top line number, 69.4% plus, or -3.2%. That's really pretty small for a rural county.

8:16
And that's because of that large dataset and because of the number of the number of times that the census folks approach people to make sure that they get a high, high response rate, almost 90%.

8:34
But it's not just county numbers we have. The numbers go as small as the track and even sometimes census block level.

8:41
So, when we looked at do you have fixed broadband? That's cable fiber or DSL. It was 37% in the county, but within the county it actually varies quite a bit.

8:53
In the most rural part of the county, it's only 13% for the wireline subscription. You have to go over into that kind of most urban part of the

county in order to get up to 43% of the county. So, very different adoption rate, depending on where you are in the county.

9:19
Similarly, we've got quite a bit of information in this dataset about computer ownership. Do you have a desktop, a laptop, a smart phone?

9:28
And we can combine this information to get information on the digital divide.

9:37
So, again, I'm sticking with Idaho County, and here we're looking at age related information.

9:45
So, I've titled it The Homework Gap, but it's really just data on people under 18, and we're looking at the percentage of people who have no computer and no internet in Idaho County, it's 12.2%.

10:03
And then underneath that, I'm looking at children who have a computer, but no internet, 7.8%, And similar data for adults, 18 to 64, who have no computer and no internet, and then adults who have a computer, but no internet. And then on the right-hand side and translating that information, this data is also available in ACS.

10:28
You can look at it either on a percentage basis, or on a numeric basis, in terms of how many people are implied by that.

10:39
So, it's 1600 children, 1600 adults, and roughly 600 seniors.

10:45
That might not seem like a lot of people, but look at the population of the county. It's 16,000 people in the county. So that's a pretty significant divide in terms of the number of people who are lacking connectivity in that county and therefore are unable to access digital services in that community.

11:09
There are so many tools that you can use to get access to this information and other information.

11:16
The NTIA data explore a number of tools from U.S Census, tools from Esri, Broadband Now, Speedtest by Ookla, Measurement Lab, and the I3 Connectivity Explorer. All the ones that are listed here with an asterisk, are third party tools. Those are not Federal tools, and NTI doesn't endorse those tools but I've listed them here because they're very useful and many of our partners use them. And two of our partners are here today and we'll

be telling you about them. I'm going to highlight two tools here and do a demo of them in just a couple of minutes.

11:53
I'm going to talk to you a little bit about the US.

11:56
Census quick facts tool, and the I3 Connectivity explorer, let me introduce them to you briefly, and I'll show you a little bit about why I picked those two to show you.

12:11
The US. Census Quick Facts tool is really a one stop shop for information to get quick facts on the American Community Survey data. It's very easy to access, and it gives you a broad view of all of the information in the American Community Survey, kind of a snapshot, or a thumbnail, so that some nail view of that data.

12:38
The I3 connectivity explorer, as I mentioned, it's not a federal tool.

12:44
This is a third-party tool, developed by Internet as infrastructure, and it's a place-based tool that combines American community survey data, FCC data, as well as other demographic data and speed test data, including information on schools and libraries.

13:07
Opportunity zones, and it all puts it together in a place-based view. So that you can get information on your community. Carol's going to give me control and I'm going to give you a demonstration of both of these tools.

13:30
First, I'd like to show you the quick facts tool that was developed by the US. Census Bureau, and it truly is a quick way to look at the US. Census American Community Survey data.

13:43
This is the quick facts view of Idaho County, so all I've done is gone to the Quick Facts screen and typed in Idaho County.

14:01
And what it pulls up is a quick view of the American Community Survey data for the county.

14:10
So, all, it's giving me the population data, demographic information, housing information, living arrangements, that information on the computer and information, computer and Internet use, education, health, economy, transportation.

14:38

Income and poverty, business information, and geography information, it truly is quick facts on your community, and it's very easy to get ahold of.

14:49
You also can get comparative information, so you can very easily get information on your community and another community, similar community. So here, I'm looking at Idaho County compared to the United States.

15:08
If I wanted to, I could add the state of Idaho, or I could add a different county in Idaho, or I could add a town. So, it's very easy to get comparative data and look at it side-by-side.

15:22
Truly, quick facts and very easy to get on your computer, on your phone. You can print it out, you can map it. There's a lot of different ways that you can look at the data.

15:34
The other tool that I wanted to show you is called the I3 connectivity explorer.

15:39
As I mentioned earlier, the I3 connectivity Explorer is a third-party tool. It was developed by internet as infrastructure. And I'm going to start in the same place that I showed you before. I'm going to start with the subscription data.

16:00
So, this is the internet, the Wireline Internet subscription data. And as you can see, Wireline Internet subscription data is uneven across the county. We suspected that it might be uneven across the county because of density.

16:16
And we can pull up secondary demographics to kind of explore why it might be uneven, and I'm going to make the guess that it is because of density. So, I'm going to pull up population density.

16:32
And as you can see, the population density is low. So, it's with zero people per square mile out here in this very rural area. But there are still a couple thousand people out here, about 2000 people out in this very rural part of the county.

16:52
Whereas over here, in the more urban part of the county where we have a 43% adoption rate for wireline services, remember that’s cable, fiber, or DSL.

17:04

There's 17 people per square mile, and it's more like, 5000 people.

17:12
Here, nine people per square mile and another 5000 people.

17:18
Here, more like one person per square mile, so you can see the density varies across the county, but there's still, you know, a number of people that we need to pay attention to.

17:30
We also kind of hint from the previous screen that perhaps our cell coverage might be a little uneven across the county and maybe might be low and that might even be lower than what we might expect from some of the national averages. And so, let's take a quick look at that.

17:47
The national average for cell phone coverage is 66%. That might be lower than you might have expected.

17:57
And here in the more urban parts of the county, it's 54, 40, 46, 45%, and out here, it's as low as 35%.

18:12
And down here as low as 21%.

18:16
Why do you think that is, and it might be an availability issue? So, let's look at the FCC Form 477 data on wireless coverage, and that information is right up here.

18:27
I'm going to look at it on a table basis. I could look at it on a map basis, but I find the tables are work useful and insightful.

18:37
So, the way this table reads is, I'm looking at the name of the provider, the technology they offer, the number of blocks they serve, and then, what percentage of that county that implies?

18:50
So, our two major providers for this are AT&T mobility, and that's two. Because it's with different technologies, and I'm focusing on their LTE service because it's the LTE service that will carry data.

19:05
And for LTE service, that's covering 50% of the of the county and then the next provider that covers that provides LTE service is islands cellular, which covers 34% of the county.

19:21

So, if one thinks that perhaps you could just put out a lot of hotspots as a way to provide data surfaces and filled the digital divide in Idaho County, you know, that might work in 50% or 34% of the county. But it's certainly not going to work in all of the county.

19:41
A lot of people look at the FCC form for 77 wireline data and that's really important dataset as well. So, let's take a quick look at that data set before we leave this tool.

19:56
The FCC, we know that the most important carriers, one of the most important carriers out here are the satellite carriers. They're listed here at the top because they provide the broadest coverage.

20:10
They're recovering again, here's the carrier. Here's their technology, the blocks covered, and the percentage of covered. So, they're covering 187 and 87%.

20:25
I'm going to get rid of the satellite coverage carriers, just so I get a better look at the wire line coverage.

20:32
After the satellite coverage Carriers', the next coverage comes from a number of fixed wireless and DSL carriers.

20:40
But when you look at the fixed wireless carriers, you see, fixed wireless is a point-to-point service. So, if you're getting, 23% of the blocks covered with fixed wireless, we know that not all of those blocks are covered. You know that it's only a point-to-point service covering a portion of those blocks.

21:00
Similarly, when you look at a DSL service that's advertised with a max of 60 meg down or 115 meg down. That might be an optimistic view in a rural area where you might have long loops that perhaps have been out there for a little while.

21:17
So, when you look at and you look at the service this way that really brings to mind why two things are important.

21:26
One, why community input is important and the other is why speed tests are important. And that's one of the reasons why I think that it's great to look at federal broadband data.

21:36

But what's even more important than federal broadband data is you, it's really important to put your community input into that broadband data and why it's important for you to use that data to tell your story.

21:56
That is why the data itself is only important in the context of you building your business case and you building your plan.

22:08
It's one of the reasons why I'm very excited about our next two speakers. Let me go back to the slides so that we can take a minute and introduce those next two speakers.

22:35
First, let me make one correction in something that I said during my talk, and that is to call Grangeville, which is on the eastern side of Idaho County. Urban, is a little bit of an overstatement.

22:47
Grangeville is a town, not really exactly an urban area, but it is the most concentrated part of the county, but that is the reason why community input is important, when you're on the ground, when you're in the community you know what the data means, and you can tell your own story with it, but it's also one of the reasons why speed tests are important. And the two most popular speed tests are from Ookla and from M Lab, and we have experts from both of those firms with us.

23:20
Lai Yi Ohlsen is the project Director of Measurements Lab which is a fiscally sponsored project of Code for Science and Society as the Project Director of Code for Science of M Lab.

23:34
Lai Yi oversees the M-Lab's mission to measure the Internet which is to save the data to make it universally available, accessible, and useful. She works to engage researchers, support policy, and advocacy work and promote a healthier internet for all.

23:57
Previously, she worked to defend and promote human rights online at a company called Eco Light.

24:07
Lai Yi?

24:11
Lai Yi Ohlsen: Hello. Thank you, Karen.

24:13
Hello, my name is Lai Yi Ohlsen. I am the Project Director of Measurement Lab, which is a fiscally sponsored project of Code for Science and Society, a non-profit that supports open science and open data

projects. Thank you to Karen and NTIA for having me. It is a rare opportunity to get to talk to hundreds of people from my apartment.

24:37
So, Measurement Lab's mission is to measure the internet, save the data, and make it universally accessible and useful. And before I get into any of the details about how we do that, I want to discuss why we do that.

24:51
Measurement Lab was founded in 2008 with the idea that we needed more information about how the internet was working for users.

24:59
It seems straightforward to say, but it's actually a really difficult problem, because for one thing, how do you measure something that is, by definition, always moving and growing? And furthermore, how do you choose which perspective or perspectives to measure the internet from? Because, after all, it's a network with many moving pieces.

25:17
There are a lot of potential answers here, but from the beginning, our goal has been to represent the user's experience of the internet, and how effectively it supports them day-to-day.

25:29
So, for the past 12 years, we've hacked away at this problem because we know that collecting and publishing data about how the Internet is working is essential for its future.

25:39
If we're going to make plans about the Internet's improvement, we need to be able to model its behavior over time. And that model needs to be based on a method and data that is available to users, policymakers, advocates, and researchers to ensure that everybody's making decisions with the same information.

26:00
All to say, we do this because we want to support a healthy internet. And we know that the health of the internet can be taken for granted.

26:07
The pandemic has only underscored that what works for some does not work for all, and that there continues to be a need to provide evidence for experiences that would otherwise be anecdotal.

26:21
So, moving from the why, and to the how, I'm going to get a bit into the details here, because, one, I believe they can be useful to you and your work, but, too, because one of the things that measurement lab tries to do is be fully transparent and our methods, and part of that is making sure you know what we mean when we say we measure the internet.

26:41
To that end, we run a network of about 500 plus servers and 130 plus metro locations globally.

26:48
On each of those servers, we host a suite of open-source measurement services, one of which is NDT, which stands for network diagnostic tool.

26:59
NDT reports the upload download and latency metrics of a connections, single stream, bulk transport capacity, which is a standard defined by the Internet Engineering Task force, and provides an effective benchmark for a user's performance.

27:16
As you might already know, the internet is made up of interactions between clients and servers.

27:22
So, we host the server side but to get a full measurement, we depend on a community of developers who integrate our open-source NDC code into the client side that is beside that the user interacts with.

27:35
It is from these integrations, that users of the internet can run tests.

27:39
A popular example of this, of this experience is if you Google, how fast is my Internet. An NDT test is what pops up, because it is embedded into the search page.

27:50
So, when users run tests, they are running it against one of our servers, which are placed in off net locations, which means that they are measuring to a server that is placed outside of their access network and within a transit provider that appears with their access network.

28:07
This assesses the approximate path to content that is hosted outside of the access network.

28:13
Again, the server placement is referred to as off net. And a simple way to think about it is that we measure the inter part of the internet.

28:25
So, when a user runs a test to a server on our platform, the measurements are shown to the user, and archived in our public database.

28:33
And here, you can see the user interfaces of some popular entity client integrations, which show the user the upload, download, and latency metrics.

28:40
Each individual test result, however, can only tell you so much about the user's connection to the internet.

28:46
To fully understand what these numbers mean, it is important to be able to put each of these tests into a broader context.

28:55
Fortunately, this is made possible by what has grown to be the largest open internet measurement dataset available today.

29:03
Because of a large community of NDT integrations, we've been able to collect around two billion rows of NDT data and it's now available today.

29:13
And now that we've covered the how of what we do, I'll talk about what that is, what insights that our data makes available.

29:22
So, for starters is to re-iterate. When users use one of our tests, they are measuring to an off-net server, meaning you have data about how well users can access content that is hosted outside of their access network

29:33
Our additionally, the NDT dataset can be broken down by a ton of a service provider number, time of day, day a week, IP subnets, and even different destination transit providers.

29:45
So, it's a great resource for researchers, looking for an extended, detailed view of beyond a summary, and we highly recommend taking advantage of this capability, because it's often possible to overlook specific nuances when looking at a generalized view.

30:02
For example, one provider might be providing high quality service, but the local average could appear to be quite low and aggregate, which would thereby give you an incomplete picture of the kind of investment needed in that area. Or, perhaps, peak usage times differ greatly from non-peak usage times. But the non-peak usage times could skew the average to give the impression that the performance is higher than it is when people need it the most.

30:28
Our most recent blog post dives into these nuances in greater detail, but these are some examples of ways in which being able to break down M Labs NDT data is to a customized granularity is essential for its full use.

30:42

The importance of collecting data from more than one source. If you are looking to specifically study the last mile or the performance of a specific ISPs network, we recommend you compare our data with other data sources, and think of M Labs NDT data as a benchmark for performance.

30:59
And so, with those potential use cases outlined, I now get to do my favorite part, which is to remind you that the data is free. All of our data is available for free.

31:10
To manage public access, we use Google's BigQuery tool, which is a platform that allows for the analysis of large datasets. To get access to the data, all you have to do is sign up for our mailing list, called M lab Discuss. And from there, you can query the data, either through the BigQuery web interface, using SQL or through a command line interface, using the language of your choice.

31:32
Either way, you can download the data in any format that works for you, CSV, JSON, et cetera. And most importantly, this means that you can share the data with anyone. So, all stakeholders are able to start from and make decisions with the same information, and all the information that on how to access the data, can be found on our website under the data tab. I think that's in the chat, but measurement lab dot net.

32:00
So, we know that large amounts of data can be difficult to process. So, we really lean on visualization tools to help us understand and share potential insights.

32:11
In the next month, we will be relaunching our own NDT statistics dashboard which I unfortunately don't have a screenshot for, but will provide a useful entry point to the data by giving summary views that tell you what percentage of tests were able to reach relevant performance thresholds in a given region. So, for example, what percentage of users in a city are able to make a video call or an audio call or meet the FCC standard of broadband?

32:37
That is one option that we provide. But because the data is open, it can be ingested it to any visualization tool of choice. Like D three, Tableau, Data Studio are shown here, et cetera.

32:47
If you want to ingest into your tool automatically, we are currently beta testing a statistics API that your site can pull data from as it becomes available. And actually, NTIA, who incorporates our data into their national broadband availability map, is one of our beta test users for that tool.

33:07
So, we have some data studio dashboards available on our website, as well as examples of exploratory analyzes. Here and on the last slide, we're looking at dashboards that were created for the Measuring Libraries Broadband Networks Project that we worked on with Simmons University and Internet two to give you a sense of the kinds of visualizations that you can create on your own.

33:28
Here we're showing a work in progress collaboration with Marconi Society where we map NDT data in relation to demographic data. The internet as infrastructure tool that we looked at before does this as well. And as mentioned before, NTIA's inbound tool is using our data to provide more mapping as well. Essentially, our data is available through several third-party visualizations, including our own, and they can be a really good place to start your research.

33:59
So, it's also entirely possible that when you dig into the data and your area of interest, you'll find that you simply need more data to be statistically competent. And M-Lab has thought a lot about this need and has tried to make it as easy as possible to do so.

34:14
As we talked about before, anyone can integrate an NDT client, so you can set up that test on your state, city, county website, and start collecting data directly that way today.

34:25
It is worth noting here that because we are a public dataset, and we want users to feel as comfortable as possible suddenly as possible. So, when they submit their measurements, the only information we collect about the user beyond the results is their IP address, which can be geo located at the city or county level, depending on the size of the region.

34:45
If you need more fine-grained geolocation, which is likely, if you're wanting to get at the Census block level, it is possible to use additional geolocation methods and have that metadata sent to your own local storage.

34:58
You will, of course, need to think about your own privacy policy but M Lab will never collect any additional metadata. So, you don't need to worry about the additional geolocation being published publicly.

35:10
TestIT, SpeedUp America and Piecewise are three great examples of custom integrations that were developed for the specific purpose of local data collection and have been used as a part of as part of a public engagement or a citizen science campaign.

35:26
One of my favorite examples of this is when the Michigan Moonshot Project used piecewise to ask students to measure the Internet and fill out a survey about their connectivity as homework, which gave them more data about the digital divide in their school districts.

35:45
So, lastly, I also did want to at least mention our solutions for measuring productivity on an automatic and recurring basis. This is practically another talk in itself. So, for the sake of time, I won't talk too much about them. But essentially, we developed a way to run tests on a recurring basis, either through on-premise devices or from the Chrome browser as an extension. These are solutions designed for folks who want to eliminate the issues of self-selection bias and are good options for feasibility studies or measuring community anchor institutions, as we did recently in the Measuring Libraries Broadband Networks Project.

36:22
One of the most exciting features of this solution is that it allows you to run any open-source test client. So, for example, during the libraries project we ran plus community speed test dot net client alongside NDT.

36:33
Essentially this just makes it a lot easier to get a lot of data from a variety of sources, and the

36:41
So, I'm going to end by re-emphasizing our mission and motivation.

36:46
We get e-mails every day from users who share that their internet performance is impacting their lives in a significant way.

36:54
We don't have all the answers, but we do know that any pathway to a viable solution will require verifiable data driven evidence that supports users' testimonies. M Lab is here to help you. All the leaders, and advocates, of your communities use this data, because we see the potential, it has to improve a resource that we all depend upon.

37:16
We've covered a lot of ground today, so I do want to encourage you to reach out to support at measurement, lab dot net, or to myself, directly, if you have questions. Thank you for your time, today, and, again, thanks to NTIA for having me.

37:40
Karen Perry: Well, I want to remind you all to put your questions in the question box. And I also want to remind our presenters, that when you get ready to answer the questions, we'd better be just as quick as Lai Yi was, because the questions are coming in like crazy. And we're only going to have a few

minutes to answer them. I'd also like to introduce Bryan Darr, our next speaker, he's from Ookla. He is the Vice President of Smart Communities Ookla. And he works to bring performance and coverage data both to the public sector and to the telecom industry. Bryan grew up in the cellular industry.

38:14
He started with Cellular one in 1985, after discovering that customers were having difficulty traveling with their cell phones, He founded Mosaic in 1988. And then when Mosaic was acquired by Ookla, he joined Ookla, and he's now serving as the Vice President of Smart Communities at Ookla.

38:43
Bryan, take it away.

38:46
Bryan Darr: Karen, thank you very much.

38:47
Really appreciate the opportunity to present today, and thanks to everybody at NTIA for including us, a little bit of information about Ookla and who we are, for those of you who may not know.

39:02
We are the most popular speed test app across the globe.

39:07
We do have many ways for people to take a speed test, our browsers, mobile applications and more.

39:12
We have hundreds of millions of users and 10 million or more tests per day is what we were seeing even before the pandemic, it's even more now. 34 plus billion tests to date. So, we have an enormous amount of information.

39:29
There's already been mentioned today about privacy concerns, we are GDPR compliant, as well as, as for the new California laws, that went into effect last year.

39:39
Our clients call the major telecom providers across the US and we have an enormous server network that you see represented here on the map. This is just our US map.

39:52
We have over 13,000 servers scattered around the world, operated by our IS partners and you can see here over 1600, every state, pretty much all of the US Territories, 293 million speed tests captured in the United States last year, from 58 million different unique devices. A key half of this,

because location has been such a big concern and understanding size location, 154 million of those tests were captured with GPS precision location.

40:31
States, many states have been capturing their own tests, trying to better understand where they're, where their networks are strong or weak. We encourage this very much. We have a speed test custom that I'll talk about a little bit more, that people can roll out on their own websites.

40:48
But I think it is important to understand that the volume of data that you're actually going to be able to capture is, is likely going to be far less than what are already being generated with the major platforms. And you can see some numbers here. We actually are seeing, you know, 60 times more tests.

41:07
And in some cases, some of these platforms have just recently been rolled out, they tend to spike.

41:14
Get a lot of activity early on, when the media attention is high, and as the media attention drops off, their usage drops off.

41:21
And the major platforms, like Google speed test, going to be, you know, just in the amount of data that it's capturing on a regular basis.

41:30
Where does this data go?

41:32
It rolls up into a number of different platforms that we have. This is our primary one called Speed Test Intelligence Portal.

41:39
You're looking at an aggregate on the map of all of the information that we've been able to plot by co-ordinates.

41:46
But the graphs themselves include all of the data that we're able to capture and then locate, which is not all of it going to be completely precise, but as Lai Yi mentioned, we can, we can aggregate that by state, by city.

42:04
For all of those tests, the vast majority of those can be aggregated, at least by zip code. But a little over half of our tests can be can be even more than that from a geolocation standpoint.

42:19
All in this portal as we roll down, you can each individual ISP name and told test.

42:24
This was for just shy of a year off of the state of Georgia. So, you can see there's millions of tests count for extending AT&T network, the major suppliers in that state.

42:33
But tens of thousands, even for smaller providers, all the way down through the page, you see in the sample here.

42:44
So, for the tests, that can be geo located.

42:48
Let's break these down. We've set this up where each layer is going to represent tests that fall into a different bucket. All of the tests that you see here, this is Northern Illinois. This is the Chicago area.

43:02
And these are tests that were less than five megabits per second.

43:08
Then we've got all the tests, between 25 and 100, then all the tests that are north of 100.

43:14
So, when we layer all of these together, what kind of a picture is that?

43:19
Now, remember, we're looking at just a subset of our data.

43:21
These are the tests that can be geo located with a precise latitude and longitude and so wherever you see the red showing through on the map which represent the test's less than 25 megabits per second, we are clearly not seeing a single test that even meet the 25 to 100 megabit per second threshold. And so, anywhere you see red on this map, these are areas that need to be inspected, to see if broadband investment is worthwhile.

43:50
Karen had mentioned, you know, some of the other data sets that you can use SEC 477 people commonly use, you know, bouncing our data against sap and other speed test data against that.

44:02
What you're seeing here is, is the footprint for Frontier's DSL Service area and the speed that have only been captured on frontier and this is, by the way, only a three-month sample.

44:15
What you're looking at here, of the speed test for frontier instead of Illinois, and very few tests, you know, meet at 100 megabits per second.

44:23

And you can see that there's usually in a cluster, you'll see one test that it meets at 25, and the rest of them fall below in a given cluster, pretty common for DSL, as you get far away from wherever their connection point is, in or out of the residences. In businesses that they're not going to be able to experience the same level of speed, the further you get away.

44:46
Karen mentioned tribal lands. There is another webinar coming next month.

44:52
There is significant funding that has been made available to improved broadband in tribal lands, and just showing you a couple of states here that have a significant amount of tribal land within their boundary of South Dakota and Arizona.

45:05
And you can see the speed test clusters that we have across these areas can give excellent indication of, you know, the kind of difficulty that folks living within these tribal zones are experiencing as their speeds are commonly below 25 megabits per second.

45:24
Rolling these speed tests into a zip code presents a picture, what you see here.

45:32
So, we've got an aggregate of information. This is six months' worth of tests across one county, in New York, Chenango.

45:42
Part of the county, as, you know, seems to be fine, North, West part of the county, seems to be great deal of difficulty, but with all tests that we can, we're fine location where that test was taken.

45:59
Then you end up seeing, then you end up seeing a mirror.

46:08
And looking at this, north-east Section, 46, you can see a cluster, we're all together, you've got a good duty there, that decent coverage, you get out to the East, and you've got problems.

46:23
When we go to the next, down here, where the accounting seems to be fine, rather, the zip code seems fine. There are in the north-west portion of 13733 zip Code that does not have good coverage consistently. All the tests are below 25 megabits per second.

46:40
Idaho County, Idaho, was used in Karen's presentation.

46:45

And I wanted to go ahead and add to that a little bit here.

46:48
We've got, most of this is wilderness and national forest, and you can see, from the light green area that it's really going to be mostly uninhabited.

46:58
But when we take and look at it by Zip Code, some of these zip codes are extraordinarily large.

47:04
What's that really going to tell you?

47:06
But then, you overlay additional information on top of that, where you can precisely locate these tests, and we get the averages in the background for the zip Codes, but then ultimately, you can see where, again, where the red and dark red is showing through, we're not seeing any indication that these areas have are reaching the speeds that meet the FCC minimum.

47:29
So from our standpoint, using data is going to help you make better decisions, a small investment in better data, and that can be not only in purchasing some data, but spending time on taking whatever data that you have, and comparing it against these other federal datasets. The ones that Karen mentioned allows you to target the areas where broadband is needed the most. You can avoid overbuilding and harming for existing businesses. Some of these smaller businesses aren't reporting Form 477.

48:01
Being able to see that data speed test can avoid actually harming smaller businesses.

48:08
And then you can use your demographics to do the most good, for the most people, with whatever amount of money is available, at a given time, We know that whatever comes out at the state or federal level, is not going to fix it for everybody right away, So how can we get the most people online as quickly as possible?

48:27
It can also even help prioritize fiber in the rider, got to have that middle mile fiber, in order to be able to ultimately reach that final mile, and connect homes and businesses.

48:39
This is a product called Cell Analytics, which actually captures cellular signal strength from the wireless operators, individually, across the entire country and across the world.

48:51

And it's roughly relates, as you can see, the different colors, to the number of bars that you get that’s not an absolute.

48:57
Just a way to better understand, that you do have stronger signal strengths in the blue to green areas that are shown on the map, and the red to orange are poor.

49:09
In this particular instance, the State of Georgia is using this solution to help understand where to deploy busses that have been outfitted with WIFI LTE devices as well as direct families that need to keep their students connected to the classroom.

49:27
We also have a tower dataset.

49:29
Wireless is going to be, in some cases, the only way to really reach truly remote areas, and understanding what tower's may be available, in order to attach those, those wireless assets, can, we think, can be extremely helpful, as well.

49:43
Couple of examples of what states are already doing with our data. The state of South Carolina actually rolled out an entire atlas, state level maps, county by county level maps.

49:53
What you see here in the green is areas that meet that FCC minimum based upon the data that we provided to them.

50:00
And then, in that red in the orange areas that show through, there are areas of high-density population that do not meet the FCC minimum of 25 down in three up.

50:13
West Virginia recently rolled out their map in November.

50:16
This is a brand-new map is created, not only with our data, with a lot of these other sets, such as eligible areas, their own Kids Connect zones, which is a state project, by overlaying all of this data. Then you can then present this information in an aggregated form to the public.

50:33
That allows them to understand the areas that have good connectivity and help you express to people. You know, we're working on this where we're identifying the areas that need help.

50:48

This is what I mentioned before, the speed test custom. You can roll this out on your websites for free.

50:54
We do have an enterprise version of that does have a cost associated with it. But if states and local governments want to roll this out, you can do it for free. You can collect the speed test directly from your own site. Include a survey so that people can report their service address.

51:11
When someone doesn't have any connectivity at all, they can't take speed tests, they can't take ours, they can't take M Labs, they can't take anyone's.

51:19
So, providing people a way to actually make themselves heard, by providing a survey that they can take, once they get to somewhere that has internet access is important.

51:30
As well as those who do want to report what type of services, they're getting at a given address.

51:35
One last thing I would like to mention is that we have an SDK that's currently in testing with a major school system that is going to allow you, too, are actually incorporate our software into whatever educational software has put into the tablets.

51:50
Chromebooks, whatever has been distributed in a given area so that it can help troubleshoot when kids are having difficulty staying connected to the classroom. And ultimately, we've made this massive investment in our society now to get a lot of kids for them but a lot of kids better connected than they were by providing them with Wi-Fi hotspots and various pieces of equipment.

52:12
Ultimately, utilizing this to offset the homework gap could be a big win for everyone looking forward.

52:21
I very much want to thank NTIA again and Karen for inviting us to participate in this today. Thank you so much. We look forward to your questions. Reach out to me directly. Reach out to Ookla.

52:32
You can also, at our website, in the chat, I think Karen has put the link to where you can download our apps. We certainly appreciate your time and attention today. Thank you.

52:52

Karen Perry: Yes, ask all our speakers to come on, and we're going to answer a sampling of the wonderful questions that we've gotten, and we're going to answer them in kind of a rapid-fire way. I'm going to answer one of them really quick.

53:05
A number of folks asked if you need to account to log on to I3, and also if we had, if there was data in the system for all of the US. You do need an account. But the account is free, so you do need to ask for an account. It's a free account, and there is data for all of the US. And I believe Puerto Rico as well. So, just ask for a free account, and you have data for everything.

53:34
And then, one of the big questions that people asked, and Bryan, I think you had a really good answer for this when we talked earlier, was about the difference between doing a speed test from a direct connection to an Ethernet port in your house, or a Wi-Fi connection, or on your cell phone. And what are you testing when you test in those different places?

54:11
Bryan Darr: When you run a speed test through a browser and Lai Yi mentioned this as well, what gets reported back is a latitude and longitude, but it's an approximate one based upon based upon just a regional area.

54:26
And it depends upon the service that's used to translate that into a geographic location. But generally, we find that the most granular area that you'll that you'll see reported back, is a zip code.

54:39
But you can roll that up to a county. You can roll it up to a city or a state, of course.

54:43
But if you're looking for a very precise location, one of the important things about having a mobile application test is that, with that native mobile application, it can access the operating system of the mobile device and actually capture the latitude and longitude from the GPS chip. So, that gives you a very precise location, you know, within a few hundred meters, at most, on the vast majority of tests of where that test was captured.

55:10
Did you want to add anything Lai Yi?

55:12
Lai Yi Ohlsen: Yeah, I would just add that when you're measuring from your browser, you're going all the way to the server, like we discussed. And, so, it's going to include all of the hops between you and that server, and when you're measuring from your browser, it's going to include the hop from

you to your router. And then, when you're measuring from the Ethernet, it's not going to include that hop. So, that would be a difference, as well.

55:35

Karen Perry: A number of states are putting in kind of statewide testing regimes that include speed tests. And some of them are trying to make, make a choice as to whether or not they might want to standardize on Ookla or standardize on M lab. And I don't want to pit the two of you against each other, but one of the questions was [differences between Ookla and M-Lab data/platforms]

56:03

Lai Yi Ohlsen: I would love to hear Bryan's answer but I have a feeling we'll both say that both tests are very useful if we would rather use both and get a more complete picture of the conductivity in your area. And in terms of the differences, we have different platform architectures. And you can, I'm sure read more about those on our websites. And we talked a little bit about today, but in general, the more data, the better, and I would encourage use of both to compare.

56:39

Bryan Darr: Yeah, and to add my two sense into that, yeah, we would certainly encourage people to use multiple data sources.

56:46

And the more data that you have, the better.

56:49

Our platforms are different. Our methodology is different.

56:53

It's not apples and bananas necessarily.

56:56

But there are some pretty significant differences that people need to be aware of.

57:00

As Lai Yi mentioned in her presentation, what is being tested as they've got, groups of servers, and specific locations there,

57:12

They're measuring the internet in a larger scope. In some respects, you can, with our application, choose, whatever server you want. You can choose a server on another continent if you prefer, but typically, when you run an Ookla speed test, then what's happening is it's looking for the closest server to you, and because we have such a widely distributed server system, you're going to really be more testing what the local provider is actually, ultimately providing to that end user.

57:44

Because a lot of those tests will be on net, Lai Yi, you mentioned earlier, that you've got on net and off net, they're doing, virtually all off net testing. Ours is a bit of a mix, depending upon who the customer is and who their ISP is, but they're always going to be hitting a server nearby.

57:59
So, particularly in compliance questions, states and local governments are trying to get the ISPs that have been provided with subsidy money to ultimately, you know, hold them to what they have promised to do.

58:14
Then understanding what they are supplying on a local level is an important distinction, as opposed to what the experience is for the user of being able to hit faraway websites.

58:25
Karen Perry: We have a question for Lai Yi. Can M-Lab make open-source speed test data available with Geo coordinates.

58:35
Lai Yi Ohlsen: Yeah, So, I'll pick apart this question a little bit. So, all of our data is available openly, and as mentioned before, it'll Geo locate using the IP address. So, in a sense, yes, but the IP address is limited in the way that it can go down to about the city or county level is what we say. We have a really excellent blog post on this, I can dig in deeper, but it's like a Yes and No. And that you might want to collect more specific geolocation data. And in that case, that's when you would use one of the additional methods referenced in the presentation where you can have a custom integration that would use either mobile or HTML five or Google's Location picker to geolocate, and then you would save that data. But it would never be published to M Lab, so it's a yes and no, and I'm happy to go into further detail in all of it.

59:27
Karen Perry: This one is, I think this one’s for me, and it's about the American Community Survey data. And I'm going to do it from memory.

59:33
It asks, how long the American Community Survey has included questions on computer and internet use.

59:41
And, because the survey had to run for five years, before it had enough data, in order to produce any of the data, there was five years of surveys before they shared any of it.

59:59
The first year. And then, there were three more years before we had any, so that we've had three years where they produced data.

1:00:08

And it was five years before that, that they collected it. So, we've had three years of data that they've shared, and then five years that they collected the data before they shared it. So, we've had three years of data that they shared, and they've been asking the questions for five years before that. So, I believe it's been three years that we've had the data. I hope that was clear.

1:00:28
They've been asking the question for eight years, but we've had three more years, three years since we've had the date.

1:00:34
I'm pretty sure that it's accurate. Let me just see if we have another. And I think we have to wrap it with that. So, I'm sure that there are more questions. We never answer all the questions on the data webinar, because people always have such good questions and there's going to be even more questions when we have next month's webinar, because next month's webinar is going to be about the NTIA Consolidated grant program. We're really excited to be able to tell you more about that program and I'm sure a lot of people are going to tune in for it, as many people know.

1:01:06
You also can come to our website for more information about our grant programs and as many of you know, we're working diligently to develop the rules and the guidelines for these three new grant programs. We're excited to have the opportunity to do this work, and we look forward to sharing it with you as it becomes available. Please, do check with the Broadband USA website for updates and contact and contact us if you have any questions at broadband USA at NTIA dot gov.

1:01:45
Thank you and we look forward to sharing more with you next month. We really appreciate you tuning in to this webinar, and please come back next month!