



Webinar Transcript: Meeting the Challenge of Measurement



On June 30, 2026, the [National Academy of Public Administration](#) (NAPA), the [Center for Open Data Enterprise](#) (CODE), and the [Bridge Alliance](#) co-sponsored a public webinar on *Meeting the Challenge of Measurement* with the panelists shown above. The Webinar was described as follows:

The public has lost trust in government information on many levels and across the political spectrum. To restore that trust, we need to address the challenges facing government data - including low survey response, privacy concerns, and flawed statistical methods - and consider alternatives. Panelists on this webinar will discuss these issues in federal data, possible government solutions, and the role of complementary data from nongovernment research, citizen-generated projects, and "non-traditional" data sources.

The one-hour webinar recording can be viewed [here](#). CODE has also provided the following transcript, lightly edited for clarity and flow.

Alan Shark: My name is Alan Shark, and I am the co-chair of NAPA Standing Panel on Technology Leadership. The other co-chair is joining us, Theresa Pardo from the University of Albany. And we are delighted to be here.

Whether we like it or not, government is being reinvented before our eyes. And we have a challenge. We can sit here and watch things or we can perhaps analyze things and hopefully make a difference. We have focused last year a lot on data and the state of data and what's going on in the federal government. And I think everyone agrees there's always room for improvement. But there are some headwinds that many of us are facing.

So I am very pleased that we are teaming up once again with Joel Gurin. Joel is one of our more active members of our standing panel, and he came to the challenge saying, I've got some ideas, and we welcome that. Joel is going to serve as a moderator. He is the president and founder of the Center for Open Data Enterprise, better known as CODE. And today, our webinar is going to be co-sponsored by the Bridge Alliance Education Fund as well.

We've got a very interesting agenda dealing with some issues that many, that keeps us up at night that we work on a daily basis. And rather than take up any more time, Joel, we are indebted to you for putting together such a great panel. I'm going to leave it to you to make the introductions.

Joel Gurin: Thank you so much, Alan. Great thanks to both NAPA and the Bridge Alliance. And I am indebted to this amazing panel who you will meet in just a moment.

Today's topic is meeting the challenge of measurement, which has become a very fraught issue and one where we need to seek a lot of creative solutions. The immediate issue that has put pressure on the whole system has been a number of changes that have been made at the federal level to data collections over the last year and a half. One of the things that's become clear as we've all figured out how to deal with this new environment is that there have also been some long-standing issues with government data and some growing issues like low survey response, privacy concerns, and many others.

This discussion is about where we are today, particularly in the U.S. with the federal data system. What are some of the issues that we want to be able to focus on and find solutions to as we not only preserve, but build the system to be even better than it was in the past. And what can we learn from some international experience with novel forms of data collection and analysis that are very promising as we look at this overall picture of national data in the U.S.

Let me introduce the panelists and then I'm going to start asking them questions.

First, Chris Jackson is a senior VP and civil society research lead at the research firm SSRS. That practice is dedicated to pulling together research on understanding and improving American public life, including politics and governance, media and information, commerce and the economy. Chris is an expert on American electoral and public polling. Over his career, he's conducted more than 3,000 published polls across a wide range of research subjects. And most recently, he's been leading the EMERGE Initiative, which is a landmark research program on the impact of recent disruptions to U.S. federal statistical agencies, and which we'll be talking about today.

Beth Jarosz is a senior fellow at the Massive Data Institute at Georgetown, and she's vice president of the Association of Public Data Users, or APDU. She has more than 25 years of experience in understanding and communicating data nuance to help decision makers answer complex questions. She's worked on state and local demography, regional economics, a lot of work advocating for federal and data users. She's worked on issues impacting local government and academia, nonprofit roles, and others. And her recent work has focused on protecting the usability and privacy of public data, including work that she's done with the DataIndex.us project, which we'll talk about a bit today as well.

Francesca Perucci is the Director of Policy and Partnerships for Open Data Watch, one of our kind of colleague organizations. She's an expert in the field of development data and global data initiatives, and she leads Open Data Watch work to scale up data and statistical capacity, technical services, and strategic partnerships around the world. She had decades of experience at the United Nations, most recently as assistant director of the UN Statistics Division, and really has been a trailblazer in the use of data and statistics internationally, including leading the development of indicators for the sustainable development goals, or SDGs.

And finally, Dr. Stefaan Verhulst is an expert in using data and technology for social impact and has co-founded several research organizations, including the Gov Lab and the Data Tank in Brussels, where he still serves in leadership positions. He's a research professor at the Tanman School of Engineering at NYU, where he focuses on using advances in science and technology, including data and AI, to improve decision making. And he's also editor-in-chief of the open access journal Data and Policy and was recently recognized as one of the 10 most influential academics in digital government globally.

It's an amazing panel. And thank you all so much for being here. I'm going to lead with a few questions for each of you going in, and then we'll see where the conversation goes.

Beth, I'd love to have you begin because you can give us an overview of what's happening in the U.S. and the federal and national perspective. We'd like to start by talking about the challenges that the federal data system is facing. How would you describe the breadth of the federal data system and what it encompasses? And what are some of the major challenges that it's been facing in the last year and a half?

Beth Jarosz: A couple of things that I think most people on this call probably know, but I think are really good level setting. The federal data infrastructure was not coming into 2025 on a super strong footing. We had reliable, consistently reported statistics, but we had seen the erosion in agency budgets, the lack of ability to modernize, and declining survey response rates, just to name a few things. For example, the Bureau of Economic Analysis started suspending data products for state and county level data in 2024 because they just didn't have the staff to keep those things moving forward. So there were some headwinds.

Then 2025 changed the whole landscape. We went from slowly eroding budgets and not having the ability to modernize to the wholesale destruction of entire statistical divisions like the National Center for Education Statistics, the removal of the head of the Bureau of Labor Statistics, major cuts in agency staffing ranging from 10% to nearly 90% of staff, depending on which agency we were looking at, and increasingly rules that constrain the accountability and oversight of those organizations.

Almost all of the statistical agencies had some sort of scientific expert oversight through scientific advisory committees. Those were all disbanded in early 2025. Now we are seeing the administration impose methodological changes on agencies, like a recent order to limit the types of privacy protections that the Census Bureau and BEA can do. That's all on the methods and staffing and mechanics side.

Then we're also seeing things like the removal of data that were collected on sexual orientation and gender identity, on climate, on discrimination, and the termination of collection of those things. And then also things like the wholesale removal of certain data collections like the food security supplement and the agricultural labor survey. So a kind of all-angles and all-sides attack as opposed to just this slow erosion that we had been undergoing.

Joel Gurin: Thank you, Beth. I think for all of us who've been working on government data or with government data for a long time, it's been a sort of startling time. Chris, I'm going to turn to you now. One issue that gets to what we want to talk about today is not only numerically how many federal data sets have been impacted – I'm not sure any of us know, but it's not 50% - but also what impact these changes have on how people trust the system or how much they count on federal data. You've been doing a lot of research on that recently, both from the perspective of actual data users and the public at large. If you could, please share what you're finding and what you think some of the implications are.

Chris Jackson: Yes, definitely. I'm a public opinion researcher and pollster, so I come at virtually every topic asking, What do regular people think? What's the experience of regular people? We've done a number of different studies over the last year and a half, trying to get a sense of what has happened, what the changes by the administration have done to the federal data system, and how they have impacted people who use the system and people in the public at large.

The data user community is a very diffuse population of people from all across the country and lots of different disciplines, from business to academics to local governments to nonprofit. But 93 percent of them in our study have been impacted by what has happened over the last year in some way, shape, or form. That goes from things like having to delay programs that they're working on because the data wasn't accessible, to having to cancel programs because they weren't able to get data access. One person told us that their tenure clock got reset because the data they were using for their tenure presentation was not accessible.

We also expanded the aperture out to the public at large. We've done two surveys with the U.S. general population, that is all American adults, one last summer about this time, and then one in March of this year, about three months ago. We found in both surveys that the large majority of Americans supported the idea of having impartial data to make decisions, and supported the idea that data collection should continue despite any sort of political interest or political interference. We heard that from 70 or 80 percent, depending on the specific question, which in America these days is a super majority.

But when we asked them specifically about how much trust they had in the federal agencies who actually do the data collection, we saw a little less positive numbers. Last summer, before the BLS commissioner was fired, 53 percent of Americans said that they had essentially the top two boxes in the scale level of trust in the federal data agencies. By March of this year, that had dropped down to 47 percent, a six percentage point decline. That may not sound like a lot, but in public opinion data, that's a pretty precipitous decline.

So we are seeing this sort of juxtaposition of a public that recognizes the value, the importance of public data in the abstract, but is increasingly distrustful of the actual institutions and the collection of the data in practice.

Joel Gurin: Thank you, Chris. That's a very clear view of where the public is. It sounds like incidents like the firing of the BLS commissioner could be impacting the trust that people have in government. But I think it's still not clear whether there's been damage to the BLS data collections and what it publishes, although the perception is clearly important.

Chris Jackson: One thing to add is that the impacts the data user community reported to us weren't just about data quality issues. They were also about management. So the fact that the government shut down for a month last fall and data collections from the BLS just weren't published, that itself is a huge impact for a lot of people - even though we think the data from before and after the shutdown is still the same high quality it's always been. We shouldn't just get focused on data quality issues. It's also an administrative and a management challenge that is having a very clear impact on the data user community.

Joel Gurin: Thank you, that's a really important point. I want to segue into talking about where we go from here. I think everybody on this call is very focused and very dedicated

to restoring the resources that have been lost. But there's also a sense that it could be a good time to look at how do we complement or supplement the sources we've always relied on with other kinds of data. We may be forced into looking at different kinds of data sources to solve the problems we have, but that there may be an opportunity to do that in some productive ways.

Beth, I'd like to go back to you on this question. One of the big categories is administrative data – data that comes from government but that is collected and published in a different way than the survey data people are more familiar with. I wonder if you could talk about that a bit and how you see those possibilities.

Beth Jarosz: This is a really nice dovetail from the question about the quality of the data coming out of the system. I feel pretty confident saying that generally speaking, the quality of data that's coming out of government is still meeting the high standards that we have come to know and expect. One of the ways we know that is that there are other benchmarks that we can check against. So for labor data, there are payroll reports. There's housing sales information that states collect. There's unemployment insurance claims data: When you are paid, you are part of an unemployment insurance program in your state, and those wage records are collected by the state.

If we look at those other sources, and organizations have done that during the shutdown, they put together trackers that looked at the historical federal statistics and these other symptomatic indicators and they move in the same direction. As long as they continue to, that's good.

One of the things that this opens up for us is this idea of triangulating toward the truth. I believe that an objective truth potentially exists, but that no one data collection is going to perfectly get at that. So for example, the decennial census in 2010 did a fantastic job of counting the population in a lot of ways but undercounted some groups and overcounted some others. If you look down at the neighborhood level, you'll find housing units in the middle of a freeway in parts of California. So it's imperfect.

If we have multiple signals that are imperfect but are all moving in the same direction, whether they're from state administrative records or from surveys that are being conducted by private sector organizations or from nonprofits or federal data, then it increases trust. It increases our belief in and our ability to rely on those statistics if they are all pointing us in generally the same direction.

Joel Gurin: Thank you. I think those are very important points about triangulating and testing different kinds of data against each other. We're not going to talk about this so much today, but people have talked about using AI to harmonize different data sets. So if the University of Michigan measures something in one way, and the University of Chicago measures the same thing in a somewhat different way, we can now put that data together with AI in ways that might not have been so easy to do before.

One of the things we're going to have to look at going forward - particularly looking at some of the numbers you just quoted, Chris - is restoring trust in data. If we're at a point where people not only don't trust government data sources, but have their doubts about trust in government as well, then the ability to see how government data is echoed by data from other sources could be very important in building back that information and that trust.

Chris, I'd love to get your take on how you see the work that you all do at SSRS, or the work that other groups like NORC do, can complement or supplement the federal data system. How do they go together? What can that kind of work contribute? What are its limitations?

Chris Jackson: The United States is fortunate in that we have a very robust sort of private research community. That goes from entirely for-profit enterprises to organizations that are mostly for-profit, like my company, SSRS, to organizations that are more non-profit, like NORC at the University of Chicago, to like wholly academic research institutes like the University of Michigan's in-house research data collection institution. So there's a huge breadth and a lot of expertise and experience in this country.

I think that people like me and organizations like mine can really contribute in three different spaces.

First, someone needs to advocate for federal data, and we can't actually expect the people working for the federal government to do it. In fact, in many cases, they're not allowed to be out there making a very strong values-based case for the importance of federal data. So I think it's left to folks like us, like the folks on this call, to be the ones saying that this is an important public good. And we're the people that can exist between the technical expertise that's required to produce the data and being able to talk in a plain language sort of way that most Americans can understand. So we're sort of that important bridge between those communities.

The second role for organizations like mine is that we can actually afford to be a little bit more innovative and experimental in how we go about doing things, because we don't have to produce the high level of quality and rigor that the federal government does. We can be a bit of a testbed for the federal government, for these federal data collections. And indeed, we've seen that the CDC and the Census Bureau have adopted some of the panel construction practices that were actually first pioneered in the private sector for some of their existing faster turn research programs.

And finally, I think that right now, people in the private research space – using “private” in a very loose sense – can step in to provide data collections that can be a bit of a bridge between some of the data collections we've lost and the time when they will hopefully be restored at some point in the future. I'm thinking particularly about things like sexual orientation and gender identity data, which has been completely purged from government data collections. We just don't know how those communities are being

affected by things right now. And that is a place where we can actually fill in some gaps using private data collections over the short period.

So those are the three ways that I think organizations like mine can be stepping in. But there are probably lots of ways people in the community can be contributing. I'm sure Stefaan and Francesca have other views about ways that we can be involved. It's a bit of a crisis and we all have to be thoughtful and creative about figuring out ways that we can help out.

Joel Gurin: Thank you, that's a really clear delineation of the possibilities. And it also gives us a perfect segue. Now we're going to push a bit more on this question of how we develop new, creative, but really effective ways of looking at what's happening in the country, what's happening in the world. These are not incompatible with or necessarily eliminating government sources, but new ways of looking at things that can help build our common base of knowledge. So Francesca and Stefaan, this now is going to get into your territory. Francesca, you've been doing a lot of work internationally with the UN, and Stefaan, you are now starting to think about how some of the international lessons can apply back in the U.S.

Francesca, let's start with you. You've been involved in a major UN project on citizen data and citizen science. Please tell us what those are, how those are defined. Why are so many national governments in other parts of the world so interested in developing these kinds of methodologies? And are there possible lessons for us here in the US?

Francesca Perucci: Thank you, Joel. It's great to join this conversation.

Let me step back for a second. How did we get to citizen data? I think we can go back to the establishment of the UN Sustainable Development Goals (the SDGs) and the realization that the increased data demands could not be met by traditional data sources alone. The SDGs had this gigantic framework of indicators. There was also an increased demand for data to address the principle of leaving no one behind. That brought attention to what we used to call data disaggregation, but also the realization that data disaggregation, the way we thought of it traditionally, was not applicable. We really needed data to understand the experiences and the realities of all individuals and different groups, especially those that are not normally visible in official data. And we had a focus on the relevance of data and the impact of data, data for change.

So why citizen data? Citizen data are not new. They have existed for quite some time, both within the civil society space and more of the research and academia space, what we normally refer to as citizen science, where projects are more structured around research objectives. We saw an increased call for the recognition of this data, for it to be really integrated into national data systems and civil society organizations, and other partners becoming more active and also producing more examples and good pilots.

Citizen data really responds to a lot of what we've been discussing until now - this whole erosion of trust in data. You provide agency to citizens, so you engage them. Better

representation and inclusion, stronger relevance, and overall, of course, filling the data gaps.

Joel Gurin: Just to make sure that we all understand: When you talk about citizen data are you essentially talking, for example, about surveys that might be conducted by people within their community? Or with citizen science, as we've seen here with a lot of environmental justice projects, people taking sensors and measuring air pollution or whatever it might be? Or is there a different way we should think about what these terms are?

Francesca Perucci: The collaborative on citizen data was established to address exactly that. What is a common framework that brings us all together? To work and collaborate we need some agreed normative framework, right? And so the terminology is important.

What we define as citizen data are those initiatives that are co-created with citizens engaged throughout the data value chain. So it's not just a survey that originates from one small civil society organization and then is used by, you know, who knows whom. It has to be a continuous engagement throughout and citizen control and ownership over the data to some extent.

That's the idea of the collaborative. The collaborative has developed a framework for the different actors to come together. But also the collaborative on citizen data was recognized by the UN Statistical Commission. For those of you who are not familiar with that mechanism, it's the intergovernmental process within the UN.

Having this recognition was very important for all those different partners that were working on citizen data. It gives legitimacy, but it's also a sort of a door that opens for citizen data to become part of the national and global data systems.

The collaborative has focused on developing the normative foundation, giving guidance to translate all these principles into practice, coordinate the various activities, coordinate the different initiatives and work, identify the methodological gaps where there was need for capacity building or further guidance, and share experiences across the system.

Joel Gurin: Thank you so much. There have been citizen science projects in the U.S., where communities are suffering from environmental injustice, for example, that want to gather their own data as a counter to the official government line that says everything is okay. But what you're talking about, Francesca, is work on citizen data and citizen science that is really done in many cases together with national statistical offices. They are seeing this not only as another different source of information, but as a source that can really help build national statistical understanding.

Francesca Perucci: Yes, citizen data initiatives can take very different forms, serve different purposes, and be at very different scales. But what they have in common is this focus on impact. They're initiatives that are co-created by those who are really concerned with specific issues.

Now, in terms of being within the official statistical systems or outside, there are cases where they've been very impactful in both situations. They can stay outside the official system and still be very impactful when it comes to addressing needs for delivery of services or bringing up certain aspects of human rights being violated within certain communities, et cetera. And then there are cases where they are co-created with national statistical offices, or even national statistical offices take them on board afterwards.

They can fill the data gaps, but they can also signal where data are missing, where there are important emerging issues or invisible issues in official statistics that can then be taken up within different programs. Even some of the foundational programs, surveys or censuses, can integrate some of those issues once they are signaled by specific initiatives.

Joel Gurin: Thank you, that's really clarifying. I think what you're all describing is what's the ecosystem of national data look like going into the future. National government data is essential, but it's not the whole thing. When we start looking at this as an ecosystem of different kinds of data that can work together, that can give perspective on each other, it becomes a very dynamic situation that hopefully will have some promise.

Stefaan, now I want to turn to you. You and your colleagues at the GovLab have been doing groundbreaking work on what you're calling non-traditional data. You've just published a really terrific blog post on this. I'd love to hear you talk about how you define non-traditional data, why you think it's important, and particularly how you see the relationship between what you're seeing internationally and how it might apply in the US as well.

Stefaan Verhulst: Thanks, Joel. Before answering, I do want to emphasize the critical situation that we are in. As you know, have described the current situation as a data winter. While we have an AI summer, we actually see the decline of data, both in terms of trust, but also in terms of accessibility and in the fact that some data, especially official data, has kind of lost its prominence in today's conversation as well.

At the same time, while we have this crisis, we need to recognize that as a result of digitalization, we've seen this massive datafication of society where we have a lot more data about society, typically collected by the private sector. So when we talk about non-traditional data, we talk about data that was digitally generated, quite often using new instrumentation methods, such as cell phones, satellites and drones, for instance. The

data typically were collected for a particular purpose, but given the nature of digital data, they can also be reused for other purposes, including official statistics.

One of the key opportunities and key challenges that we have focused on is how to reuse this massive supply of data that has been generated as a result of digital instrumentation and digital transactions. How do we start using this supply for public interest purposes, including for improving and complementing, but not replacing official statistics. And so we've been for the last ten years focusing on that.

We've also developed a whole range of data collaboratives to have access to that data, because obviously the data is typically privately held. And so the question is, how do we actually develop new partnerships? We've done work on that internationally. One of the first reports on that topic that I wrote was in 2017 with PARIS21, where we focused on accessing private sector data for official statistics.

But suffice it to say that the U.S. has also experimented tremendously in that space. When we are looking at the demise of kind of the data ecosystem in the U.S., it's also important, as Chris was saying, to recognize that there is a data ecosystem out there that perhaps can also be leveraged and to learn how you can start leveraging non-traditional data.

Some kind of inspirational examples are, for instance, the efforts that Professor Raj Chetty has done with regard to opportunity insights, where initially they were starting to link IRS data with educational data, but then started complementing it with non-traditional data, including geospatial data and consumer credit data. And it has given us an opportunity that typically statistical agencies would probably not be able to develop because they would not have this breadth of data and the velocity and the veracity of the data that other data sources bring to the table.

It goes back to what Beth was saying: that we actually need a variety of data sources, not just the traditional data sources that official statistics are producing, but a variety to have a better sense of today's society.

Now, there are many other examples, Joel, like Meta's social connectedness index using Facebook friendship indications. You have MasterCard with the inclusive growth index, which is using payment data. We have seen Zillow developing insights with regard to housing stock. And of course, ADP has payroll data that complements employment data in a far more timely and quite often more accurate fashion as well.

These are just a few examples of the potential, Joel. These typically are kind of outside the statistical agencies. But even the census, with its own experimental data products, has also shown that using non-traditional data can actually massively improve and, I want to underscore, complement official statistical data. You still need the ground truth at

the end of the day, as Chris was referring. We still need federal data in order to make sense of the non-traditional data sources as well.

Joel Gurin: Thank you. No, that's very clear. That gave us a great picture, both of what's developing and, as you said, how it can be complementary, but not a replacement for some of these national statistics.

I'd like to go around the virtual room twice on two questions, having had this conversation we've just had so far.

My first question is this. Everybody agrees very strongly that these other kinds of approaches we're talking about cannot replace federal data, or certainly cannot replace all of it. For example, we must have a reliable U.S. Census. But data.gov, the country's centralized portal to federal data, has more than 300,000 data sources, and we know they're not all equally important. How do we prioritize? How do we know what to fight for? What are the things that the federal government in this country absolutely must do?

And then for the second round, I'd like to go around and ask each of you, What do you think is the greatest potential for some of the other kinds of data we're looking at? For example, Stefaan, you mentioned that some of these data sources may be more timely, and may bring a timely kind of counterpoint to the more statistically rigorous but slower moving statistics that the government collects.

Let's start with the essential federal role. Beth, do you want to lead us off on that?

Beth Jarosz: Sure. There are so many ways to answer this question. On the one hand, there is the legislative mandate side. There are certain things that are written into law that say you will use this data to allocate funds or to plan for something or to evaluate a program.

So that's one potential answer. But more fundamentally, there's a role for the federal government in setting standards. So even in places where the data collection actually happens at the county or the state level, the federal government is often the organization that gets them all together.

For example, vital statistics, birth and death records, a lot of health data set the stage for that. Go back to the late 1800s, early 1900s, states were collecting data, but they were each doing it in their own way. And over time, there was an ecosystem built up to say, okay, we're going to classify diseases this way. And states then implement them in a way that makes sense for their state. Maybe they collect more data than the federal government needs. Maybe they only collect the minimum. That gives a lot of flexibility, but it allows us to compare nationwide. I think other domains could potentially benefit from models like that.

The other place where the federal government is absolutely needed are places where there's no other institution that's going to step in and do it. So things like climate and ocean monitoring, you know, there are lots of needs for those kinds of data. Air quality: Do you really want your air quality monitoring to change from administration to administration because a state governor has changed? No, you want to know, is it a bad air quality day and should people who are at high risk stay indoors? And you want that monitoring to be consistent.

So those are the two places where I would put the emphasis: Places where the federal government sets the standards, and places where there's a critical need that no other institution is going to fill.

Chris Jackson: What Beth said is absolutely the right thing. The only thing I would add is I think we do need to also be reaching out to the data user community on a regular cadence. What are they using? How are they using it? How essential is it? We can use that also to make a determination. So it's not just the top looking at data collections and saying what seems important, but the user community also feeding back up to us what they're relying on, what they need.

That can help us prioritize, and also help us figure out where we need to improve our processes, improve our deliverables and publications. We never entirely know how the data that's being produced is being consumed. I think building a process that allows us to do that in a more consistent and rigorous fashion would let us be much smarter about what we're actually producing.

Francesca Perucci. I struggle with thinking in terms of federal government data versus other data. I see this as more of an issue of the overall data ecosystem.

The challenge in this country is that the statistical system is quite unique. It's federated along sectoral lines. There are other federated statistical systems, but they're more like geographical, regional, sub-regional level.

So that's a challenge because I think the future is about integration and investing in integration, better technical approaches to interoperability, better metadata, open data principles, good data exchange agreements across the different sectors and across the different agencies.

My hope is that we could rethink the way we approach the data system in this country, looking much more broadly at the data ecosystem in its entirety and investing in both integrating the sectors and integrating the different types of sources.

But I agree with Beth. There has to be an authority to set some of the norms and hopefully, you know, agreeing with international standards because those are developed by statisticians and experts altogether.

So ideally, if I had a magic wand, and we could have a better, perhaps a little bit more centralized system where you have an independent authority that can work at this level and bring and act as a steward, in a sense, right, to bring all these different data actors together.

Stefaan Verhulst: In Europe they have this concept of high value data sets. And I think that's really what it comes down to. From my perspective, and that also relates to the whole discussion about public trust, the question then is who determines what are high value data sets?

High value data sets are directly connected with high value questions. And then the question is, of course, who formulates those questions? And here, I would argue that in the same vein as Francesca has been arguing for citizens' data, we need to actually engage citizens far more in formulating and prioritizing what are the questions that matter to them. Then, what are the high-value data sets that can answer those questions in a way that is rigorous and trustworthy?

I would say today, in a crisis perhaps, it might also be an opportunity to really engage with citizens more around the questions that matter. And then, as a result, create a value proposition for why they need a federal data system. Because a system like that can answer their questions in a far better way and in a far more trustworthy way than otherwise would be feasible.

How do you determine what must be federal and what's not federal or what's official and not official? Well, let's engage citizens more on actually determining what they need to know. And then you will also have an understanding of what high-value data sets are and who should produce them.

Joel Gurin: Thank you. You've emphasized the great importance of beginning with what is the question we're trying to answer, not with "I have data, what can we do with it?" That's not a very productive way to go, although it tends to be the way we often do.

Alan, do we have questions from folks on the webinar?

Alan Shark: One of the questions that comes up has to do with citizen buy-in and trust. What can our respective organizations do to actually reach out? We're the experts talking to each other about how we do this. But again, we are dealing with the lowest level of trust, and it keeps going down. We'd like to be able to say decisions are based on data - to me, this is the greatest challenge. And so the question is, what can each organization do to reach out to consumers, letting them know why data is important? How do we improve things? It's almost like a public challenge to all the organizations that have a stake or an interface with data.

Joel Gurin: Great questions. Let's open it up to anybody who'd like to respond.

Chris Jackson: I think one of the big challenges that we face as a space, and as a sort of technocratic class in general, is that we have adopted this very technical, specific, precise language. We've done so for a lot of good reasons, but it has really distanced us from regular Americans. When we talk about things like data, they don't even know what the heck we're talking about. So I think one of the big challenges that we as a community face is figuring out how to talk like regular people again and to bring stories to them that relate to their lives, not expect them to meet us where we are. Not to have them figure out all the intricacies of federal data collection, but instead meet them where they are with stories that affect their lives that help them understand the value of the work that we're involved in.

The big challenge is that we've trained generations of statisticians and methodologists and procedural experts, but we haven't really trained a generation of communicators. And I think that's something that we have to do. As I said earlier, we have to be the ones making that case. We can't expect the federal data workers to do it because in many cases they're statutorily not allowed to.

I think the big place that we need to be active right now is figuring out how we tell stories in plain language about things that are really affecting people's lives in very practical terms and then equipping our people to go out and do that work. Go out and annoy our friends and our family and everyone else. For example, Denice Ross likes to talk about how counting bats is actually good because it helps reduce insect populations and reduces use of pesticides, which means farmers can be more productive, which means children have less birth defects and all this other stuff, which all goes back to counting the number of bats that are out there. We need to use stories of this sort. I think is the big step that we all need to do as a community.

Stefaan Verhulst: Two quick reactions. The first one is that, together with two colleagues, we published an article which was also published in *Le Monde*. And the basic argument there was that one of the reasons why citizens have lost trust in official statistics is because official statistics typically pushes out averages. And if you ask people, are you an average? I think most people will consider not themselves to be an average. So the question is, how do we go beyond averages? And how do we include lived experience in a far better way to complement the averages that typically are being produced and that people are focusing on. That also relates again to what Francesca was saying, because citizen science could bring in this kind of lived experience and complement the averages to make it more real for people.

The second element: I would argue that especially as we are experimenting with new data sources, and even kind of reinvigorating the system, we need to reestablish a social license for using data. We've been working on how you get a social license for the work that official statistics are doing beyond saying we have a mandate, because that's not

sufficient anymore in today's world. And here we need new kinds of deliberative platforms. For instance, during COVID, we developed the first ever citizens' assembly around data here in New York, which was called the Data Assembly.

We need a lot more of that kind of activity where we actually engage with citizens in a meaningful way. Not just are you for or against, but what are some of the questions that you care about, some of the conditions and some of the output that you would like to see moving forward.

Francesca Perucci: Very quickly, I think this is fundamental. Traditionally, citizens have been seen as respondents. So the official statistics are seen as not relevant in many cases, right? They have this passive role. People have a passive role. It's so important to engage citizens and rebalance that.

At Open Data Watch, we've been doing a lot of work on intersectionality, really trying to get to a better understanding of how you can use data to reflect the lived experiences of individuals. So that's also a way to regain trust, to really show the data value and show the impact, the data that are relevant to those that are normally the respondents. It's conceptually changing the way we look at data, both in the official statistical system and what we can consider the broader data ecosystem.

Joel Gurin: Thank you, Francesca. It reminds me that a lot of the groups in the U.S. that are working on community control of data or community engagement around data are using the phrase, "Nothing about us without us." Don't collect data on us: Engage us in what we need to know as a community and then let us figure out together how we learn what we need to know.

Now to wrap up, I'd like to go around and ask you each in two or three sentences, what are the most important points that you hope folks listening take away from this conversation?

Stefaan Verhulst: First of all, I think there's a massive variety of data out there, but it will require investment in new notions and new roles of data stewardship in order to leverage that. Typically we focused on the analytical aspect, but we forgot the stewardship aspect of data that is necessary to have access to it and use it in a responsible way. Without investing in data stewardship, trust is not going to get rebuilt and we're not going to be able to innovate in responsible ways.

Francesca Perucci: I'm very optimistic after this conversation. I think the work we've all been doing is producing some results. What I hear here resonates with what we do at the global level and what I see happening in the US. So I feel like we are going in the right direction. The challenge really is to bring the balance: Have better governance and engage people. And what we call in the citizen data space, the responsible and ethical production and use of data and really engaging people throughout the data value chain.

It's very important to get to the end and demonstrate impact that will help us rebalance these powers and also regain trust and make data really powerful and useful to people.

Chris Jackson: We do not have to convince the public that impartial, high-quality data is important. They already recognize that. But people are being impacted by what is happening. Those impacts oftentimes are much more of the drivers of declining trust than anything else. So we as a community need to be thinking about how we actually help people in their day-to-day lives with practical solutions that help them do the things that they're trying to do rather than things that we're trying to get them to do.

Beth Jarosz: What others have said echoes what I would say. The oversight is particularly important. And I'll put a finer point on that. We had some oversight mechanisms in place, and I think a first step would be to bring those back. Bring back the scientific oversight, bring back the community input and oversight processes that we had, and then build upon those to help both make sure that community voices are included in the decision making and also to rebuild trust.

Joel Gurin: Thank you so much. I want to again thank NAPA, thank the Bridge Alliance, and tremendous thanks to all of you on this panel. We're in serious times with serious issues, and this is a conversation that definitely will continue. Thank you all again for such a thoughtful and insightful discussion.