

HUD Exchange - Data Quality 101: What is Data Quality? 5/5/20

Alissa Parrish: Hi, everyone. Welcome to "Data Quality 101: What is Data Quality?" We are going to go through a couple of housekeeping items before we get started today. So we are recording today which will last approximately 60 minutes. We will post a link to the recording and the presentation slides to the HUD Exchange in the near future.

And during this training, everybody will be in listen-only mode and we do want this to be as interactive as possible. So we encourage you to use the chat box that you see on the right side of your screen. If you don't see it, on the bottom of your screen there's a little three-dot thing that if you click on you should see the chat box.

And when you post things there, please make sure that your to-response is to all the participants so everyone can see them. A quick note about NHSDC, so this session is held -- presented by NHSDC and HUD and we will be -- well, NHSDC will be sending out a survey after the session.

So please make sure to respond to that. This is definitely a trial or a pilot process in the virtual reality. And so we'd appreciate any feedback that you have on that. Before we get started into [inaudible], my name is Alissa Parrish, I'm with ICF. And Meradith, I'll let you introduce yourself too before we get started. Meradith, I think you're still muted.

Meradith Alspaugh: The double-mute has gotten me in all three of the sessions I've done. So I am very sorry about that. Thanks for letting me know. I am Meradith Alspaugh and I am from the Partnership Center.

Alissa Parrish: Great. So we are really excited to have everyone here today and appreciate the flexibility doing this virtually given this current situation. We do want to note that this is a 101 session. So we will be focusing on data quality basics.

There are other virtual sessions offered this week that will touch on more details about data quality, how to measure that data quality and then how to build out a comprehensive data quality management program. Because this is a 101 session, the learning objectives include a [inaudible] of HUD's goals for how CoCs use data to make data-informed decisions and then how important data quality is in that content.

We will spend most of our time together going through the components that make up data quality, how those are defined and then a little bit about how they are measured. We will also spend some time understanding how different key stakeholders and partners play a role in data quality and why each one meets [inaudible] table in a given community.

So again, this is a baseline session about data quality. It's meant to be a high-level overview of data quality, why it matters, what it means and then who's involved in the process. Even though we are virtually presenting this we would really like participant engagement through the chat box.

We will be using some polls throughout this session that we'd encourage you to respond to and we will be leaving room for some pauses to answer questions as they come through the chat. Please, please use it. Please ask questions, please provide suggestions and things that you're doing in your community as it's relevant.

Once we go through the overview of the data quality components that is involved we will end with talking through an action plan to help you think through your next steps related to data quality within your given community. So to our first poll, who is with us today?

So with entity are you presenting -- or which entity are you representing? And you can select all that applies. So if you are representing both a CoC and an end-user participating organization, please select all that apply.

John Panetti: Poll will be closing in 20 seconds.

Alissa Parrish: All right. So we've got some great answers. It does look like HMIS lead administrator is the largest portion but we do definitely have some CoC representation and some HMIS participating organization presentation and funders. So that's great. Thank you.

Our next poll is really a question that will be a short answer of why did you choose this session. So I don't know the character limit on the short answers, John, but quick one-sentence reason why you chose this session today.

John Panetti: The poll will be closing in 20 seconds.

Alissa Parrish: All right. Looks like the poll ended. I do see a couple questions that I wanted to touch on out loud. We will be posting the slides and this session is being recorded. And Meradith, I think I'm going to hand it to you.

Meradith Alspaugh: Yes. Okay. I'm not sure where I'm ushering from. Is that better?

Alissa Parrish: Mm-hmm.

Meradith Alspaugh: I just want to add in here, too, for you folks on the chat, please feel free, if you didn't get a chance to add your reason for participating in the poll, because it looks like it might've closed a little quicker than some of you were prepared for, please go ahead and add that to the side so that we can have [inaudible] perspective on that too.

Okay. So let's move on here. Let's talk -- let's sort of set some baseline standards. Let's talk a little bit about what we're here for today. So we want to talk -- I want to start out talking a little bit about the SNAPS data strategy to improve data and performance. This is something that I know we've talked about at past NHSDC Conferences.

I believe this is something that's been talked about on the HMIS System Administrator's calls. There's been a variety of places where we've talked about this, but it really is sort of foundational

to what we're going to be talking about today. So the data TA strategy has been available since 2018.

It serves as a framework for which TA -- by which TA is provided to communities related to data quality and performance. While the goals within the strategy are aspirational and shouldn't be used to monitor projects for compliance we -- HUD does believe, and those of us that were part of this process and these conversations, do believe that they are realistic to achieve.

The strategy addresses all pieces of data quality, completeness, including system coverage, timeliness, accuracy and consistency. The strategy also provides some benchmarks to assist communities in making data-driven decisions based on accurate information.

So CoCs, HMIS leads and organizations all really are going to be working together to review the strategy to set local goals and performance indicators. There is a link to this on the slides which will be posted after the session or you can simply Google it to see more information.

So the three specific strategies -- there's three specific strategies and today we're going to be looking at strategy number two, because it's the most relevant and focuses on data quality. So it's data systems collect accurate, comprehensive and timely data and you can see that here. So there's a table within the framework that lays this out.

So you can see in the green bar there, the characteristic, so the first characteristic is about bed coverage across a continuum for both funded and unfunded projects and HUD recognizes and realizes that the current status, again, this was made in 2018, so roughly current at this point, is at 100 percent of required providers and less than 100 percent of non-required providers are contributing data to HMIS.

HUD would like to see the majority of CoCs in the next 3 to 5 years be at 100 percent; right? So we don't want to see that go down of all homeless service providers. And then for the advanced CoCs in the coming 3 to 5 years, really, the hope would be that 100 percent of homeless providers and non-homeless service providers are contributing data to a shared data environment.

The next characteristic around data quality being timely, accurate and comprehensive. So that's a lot of -- what we're going to be talking about today is really digging into those three points and kind of expanding on them also, but HUD recognizes that the current status is about 100 percent complete but less than 100 percent accurate and there really isn't a timeliness standard that they've established.

However, moving forward the expectation in 3 to 5 years would be that the majority of CoCs have 100 percent accurate data, 100 percent complete data and that projects are directly entering data within 2 hours for crisis response and project start and project exit and that PSH projects would be entering data within 24 hours.

So really looking at like as close to real-time data entry for crisis response projects and then within a day for permanent supportive housing projects. And then the advanced sort of goal here,

the advanced CoCs in 3 to 5 years, would be that 100 percent accuracy as well as customized local data elements being limited to information needed to report, serve and house clients.

So really being mindful about the data that is being collected but that the data that is collect is 100 percent complete and that timeliness meets coordinated entry placement and referral needs. So there's not necessarily a quicker than real-time that they've indicated in this advanced CoC section, but really meeting the needs of coordinated entry; right?

Like if you're using information in HMIS for determining when you have openings in projects or availability of different resources, having that data collection completed in a timely manner is going to be really critical. So let's dig into really talking about what is data quality and I think I'm going to turn it over to Alissa to take us from here.

Alissa Parrish: Great. So we will break down each of the terms, as Meradith mentioned. Components of data quality include timeliness, completeness, accuracy and consistency and I'm sure if you've been in the HMIS world for a while you've heard these terms over and over again.

But we're going to break each down and provide examples of how to measure them. So what is data quality? It's really speaking to the reliability and comprehensiveness of any community's data and obviously, in our world we're talking about HMIS data quality. So reliability and comprehensiveness can be broken down even further into four core components which include the ones on the screen.

So timeliness, completeness, accuracy and consistency. And we will go ahead and move forward with the slides. Yeah. You can see the requirements. Real quick before we get into the defining the terms, wanted to take it back to the requirements for data quality.

So the requirements for HMIS data quality, based on the 2004 Data and Technical Standards, and in their regs -- in those regs there's really a single sentence that talks about data quality. And so it's up here on the screen, PPI collected by a covered homeless organization must be relevant to the purpose for which it is to be used. To the extent necessary for those purposes, it should be accurate, complete and timely.

So [inaudible] data and technical standards about how each community measures [inaudible] accuracy, completeness and timeliness varies slightly. So I think we are on our next poll here. So talking -- let's start with our strengths within the community. So which data quality component is your community doing well?

And we've got our timeliness, completeness, accuracy and consistency as our options and then maybe once you answer write in the chat box while you're doing well on this component. What is it that's bringing you success here?

John Panetti: Poll will be closing in 20 seconds.

Alissa Parrish: Okay. So we didn't have as much of a response right at this time, but it does look like completeness and consistency were the highest responses of those who are doing well.

Curious if anybody had thoughts on why they're doing well, but I don't see any responses in the chat.

So we will go onto the next poll which is the opposite of strengths, which component of data quality is your community struggling with right now? So same choices, timeliness, completeness, accuracy and consistency. Then the same question I have in the chat box of why do you feel like you're struggling in that?

I see one response of completeness is easiest, we'll get into that. Also, the distinction between completeness and accuracy so that data can be complete but not always accurate, absolutely true.

John Panetti: Poll will be closing in 20 seconds.

Alissa Parrish: Accuracy seems to be the struggle here, although, in the chat box we're also seeing some comments about why completeness is a problem, HMIS utilization, timeliness is a struggle. Great. Thank you. So Meredith, I think I'm going to kick it back to you to go through some of these data quality components.

Meredith Alspaugh: Sure. And I saw a couple of questions about some of these specific individual components. So hopefully we can make sure we cover some of those questions as we're going through. You guys are writing a lot in the chat, it's tough to keep up with.

So timeliness, so with regards to that we are talking about the degree to which the data is collected and available when it is needed. So there's a couple of different ways to look at timeliness. First of all, there is -- in the HMIS data quality framework, there is a question.

So that's in both standalone HMIS data quality reports that vendors were asked to make available, it's also one of the first six questions in the CoC APR as well as the ESG CAPER. So you can look at that particular question that goes into sort of a little more detail about the timeframe in which that data was entered into HMIS relative to the data collection point date.

So looking at, if you know, the project start was on the 1st, how many days did it take for that data to be entered into the system. You also may have some local reports or some custom reports that have been created for your system. Those are going to vary based on different software.

Understanding and reviewing the timeliness of data for all phases of the client's project activity is really helpful to understand, because you may also be finding -- you may be able to find places where a lack of timeliness is affecting data quality. So another big one that we look at a lot with timeliness is annual assessment collections.

So you know that the annual assessment has to be completed within the 60-day window around a person's anniversary data. Getting that information recorded in a timely manner ensures that that information is going to be accurately reported in the APR. It's not going to do your project any good if you enter that information late and you're not able to reflect that information in the APR for your project.

It may also be useful to look at which parts of the system need to have different timeliness standards than others. So it's probably not always realistic to say that you're going to have real time data entry all the time; right?

Like if you have a crisis project, like an emergency shelter, getting that real-time data collection is going to be much more important and critical, because you have those clients that are accessing those resources immediately.

Theoretically, you have staff there that are able to record that information in HMIS in real-time whereas a permanent supportive housing project it may be the case that a client leaves their unit or leaves their apartment and you don't even know it happened.

So it may not be realistic to have that real-time data collection because you're not out there every single day necessarily but setting some sort of realistic standard and timeline requirement around when that data should be collected is going to be really helpful and knowing the difference between the two, because I think that when you have unrealistic expectations for some projects, they're not going to necessarily strive to achieve those because they know they're not going to do real-time or they may not even be able to do 24 hours based on your geography or based on different circumstances.

So having some realistic timelines is going to be really helpful to get to see those improvements over time. I think also one of the questions I think I saw was how do you encourage timely data collection, I think that might be a place where you work with your CoC and talk about establishing some parameters or some expectations for projects about what the timeliness standards are and then having some way, again, within the data quality report or some other custom report, to measure and monitor the performance of different projects in your community.

Again, would love to see -- I know that timeliness, I see this isn't necessarily a top priority for overburdened agency staff, but I think if there are others that have had success with getting the timeliness of their data collection improved upon or how they're monitoring for that, I think it'd be great if you could add that into the chat over there on the side.

Moving on to talk about accuracy and I liked the comment that someone made here -- is accuracy next? Sorry, my slides are out of order. Completeness is next, thanks. Sorry about that. So talking about completeness this was still related, that completeness is easy, accuracy is hard.

I think that's a really good point to make. When we're talking about the completeness, in particular, related to data collection not necessarily about bed coverage just yet. You know, having something in the field is certainly something that people are good at doing. You can enter something into your field just to make sure that there is something there, but that doesn't necessarily mean it's always right.

So it's generally easy to measure completeness. We have a variety of different reports that have client doesn't know, client refused, data not collected where you can measure that information. I would also put out a flag that we've seen, when talking about data quality and looking at data

quality in our community, that sometimes there are people that mark the indicator for client doesn't know or data not collected but actually do have the data.

So I just want to flag that as something that you might be -- maybe should be looking at in your community also. There may have been a record that was created without a Social Security number, for instance, so it was marked as client doesn't know or data not collected. It gets updated, but then that flag doesn't change and then you're reporting worse data quality than you actually have.

If you actually have that information, it'd be great to make sure that you're actually getting credit for it. So there's a number of different reports, like the APR, the CAPER, probably other reports in your system that are also getting at that completeness sort of point there.

Now, when thinking about coverage, that's our other struggles here looking at bed utilization and bed coverage rate and we know that a lack of bed coverage in HMIS can really impact your understanding of your homeless service system and that it's a challenge to increase that coverage rate.

This is not a new issue that folks have been -- you know, that's just coming onboard, people have been struggling with this for a long time, like getting those non-HUD-funded or non-required agencies to get onboard.

So I think when thinking about strategies for increasing coverage we've seen a lot of communities have success when really helping those non-HUD or non-required agencies understand the value of the data and making the data collection manageable and easy for them to do.

I think also the other thing that we've seen is making sure that there are technological solutions for some of those other systems or for some of those other agencies when we're talking about small, maybe poorly-funded, as many unfortunately are, they may not have the resources for Internet or reliable Internet or high-speed Internet, they might have limited computer skills, they might not even have functioning computers; you know?

We've seen a lot of those things. So I think when looking at ways to be strategic in offering potentially additional resources, additional training, additional support or even looking at the data that's being collected for your particular projects, like is there somewhere you can give some flexibility for those non-required agencies that could be participating in HMIS.

So moving onto look at accuracy, again, like we've said, accuracy is hard to -- can be hard to measure, because your system doesn't know what it doesn't know. So this one can be really challenging, but a couple of places where you can look at some of this data quality is looking at, for instance, I think Alissa mentioned this when we were talking about this session beforehand, the data quality table for the LSA.

Like there are some important flags in there that could be easily addressed. So making sure that there's one and only one head of household per given household, making sure that the date of birth doesn't equal the project start, making sure that clients under 18 are not veterans.

So looking for those different pieces of information that are logically inconsistent. You can also look at the HMIS data quality report for some of this information too. So again, when we're talking about this comparison between completeness and accuracy, I think of Social Security numbers a lot, because people want to put something in the Social Security field and they know it's a nine-digit number.

So they put in some series of repeating numbers. Well, the HMIS data quality report has some data quality sort of issues that we've programmed into it that look for things that are not possible. So it looks for a double zero in the middle; right? Like that's something that's not possible.

There are certain rules around Social Security numbers that can help you identify a real Social from a fake Social. So that's one way to think about how you can improve the accuracy, similarly with dates of birth, there's some data quality checks that are built into that table for the PII in the HMIS data quality report.

So looking at those existing types of reports and then also, again, looking at ways that you can build out your own report. Another fun thing that we've seen is that sometimes people will put mom's Social Security number for her kid's Social Security number because they don't know the kid's Social Security number.

So we know that's not right. So how can we look for ways in our system to identify duplicate Social Security numbers or duplicate whatever it might be. Obviously, birthdays can be the same, but there may be some other situations where you can look at different sort of customized reporting options in your system to look for some of those indicators that don't make sense.

Similarly, again, with the HMIS data quality report, you can look at clients who are inactive, clients whose records haven't been touched in street outreach or night by night shelters in some period of time. That can be a really good indicator that someone's not recording their exits.

That's going to get you on timeliness as well as accuracy, because having those people in the projects is going to be showing your longer lengths of time homeless, it's going to be showing more active clients. So there's a variety of different ways that you can look at that data for different purposes.

Moving on here to consistency, so when we're talking about consistency we're talking about the degree to which the data is equivalent and the way it is stored and collected and I think that this is also a really, really hard one to measure.

Consistency, there's not always a report to be able to run on this, but this is one where maybe you can look at your training methods, your training resources and look at are you training organizations the same way, do all of the users understand the data elements in the same way?

This would be a really good opportunity for those of you that are system administrators having those refresher trainings for folks to help them remember what all of these different pieces of the data mean.

The data standards change every couple of years. It's a lot of information to remember and if that's not really your primary focus, you might always not think about the difference between -- or understand the difference between friends and family permanent versus friends and family temporary.

So having those ongoing regular trainings and refreshers can help ensure that that consistency is as high as it can be.

Having trainings also that are targeted towards specific user types, so an entry -- a data entry person at a shelter as opposed to maybe a case manager in a Rapid Re-housing project, just making sure that they're understand what their role is in the system and how that data impacts them can also help improve their consistency with the collection of those different pieces of data.

I see a question over there, anyone with success stories. Oh, that was yours, Alissa, I'm sorry. Yeah. We would love to see the success stories on measuring consistency, because again, I think that's one that is a little harder to grasp but absolutely very important to make sure that the data is meaningful that we're collecting in HMIS. Alissa, I'm going to pass it back to you.

Alissa Parrish: All right. Thank you. So we're on our next poll because we are going to shift from defining the data quality elements to who within your community or which entities within your community need to be involved in measuring data quality. So who is currently involved in your data quality process in your community?

And again, this is one where you can select all that apply. [inaudible] CoC, as many [inaudible], others [inaudible]. We also saw some responses [inaudible].

Meradith Alspaugh: Alissa, you're cutting out a little bit, we can't really hear you.

Alissa Parrish: Is it any better right now?

Meradith Alspaugh: It is a little better right now. Yeah.

Alissa Parrish: Okay. So we also got some responses directly in the chat about who's involved. So HMIS lead and the participating organizations work together. CoC's involved a little bit but not significantly and if agencies are involved, it's at their own choice. Well, we're going to go through the different stakeholders involved and why they should be involved.

So for data quality to be fully addressed within a community, each of the stakeholders need to be involved and then also understand their roles and responsibilities related to HMIS and data quality.

Most communities see that each of the stakeholders are involved in some way but maybe not with clearly defined roles and responsibilities to ensure that everything that needs to be addressed is addressed and also that the lines aren't getting blurred between and among the stakeholders in their roles related to data quality.

So Saul [ph] mentioned that the CoC is kind of involved, HMIS lead and organizations work together. Don't know if there is governance in place that really codifies those relationships, but we will go through each of the different stakeholders and how, in an ideal world, they would be involved in data quality.

So we will start with the CoC. So the CoC is the entity that can serve as both the incentivizer and the enforcer of data quality. So they have both the carrots, which is often money and the stick, which is also often money and the CoC helps the community understand how high data quality supports the other efforts of the CoC in addressing and ending homelessness in the community.

It also helps -- the CoC helps make the connection on how they are making data-informed decisions and also adjusting the homeless service system to best serve clients using that data in the system.

The CoC also empowers the HMIS lead to do the day-to-day work of implementing your comprehensive data quality management program and communicates to the community why the HMIS lead's role is so important in data quality efforts. The partnership between the CoC and the HMIS lead to provide consistent messaging about the importance of data quality and how it fits into the larger community of priorities is really essential.

I know that we have a lot of HMIS leads on this call. And so in communities where the CoC is not as largely involved, I think you are feeling kind of that pressure or that strain of you being the only one involved in that work. And so the CoC's role is really crucial in this in connecting it all back to the larger CoC goals and processes and then also helping with enforcement [inaudible].

So the HMIS lead -- so while the HMIS lead carries out the day-to-day work of monitoring the data quality in HMIS we do need to have, again, the support of the CoC to enforce those efforts and help the community understand that it isn't just the HMIS lead that's asking the providers to focus on data quality, it really is the CoC and how that is related back to their overarching goals.

So while the majority of the work of implementing a comprehensive DTMP falls on the HMIS lead. The other stakeholders that we've talked about so far, the CoC and the others here, in the next couple of slides, they are essential to making sure that a data quality plan is successful.

The HMIS lead does need to be empowered by the CoC to carry out the work and the participating organizations should know that any requests to address data quality that come from the HMIS lead are fully backed by the CoC.

So the HMIS lead should be seen by that participating organization as the entity that assists them in understanding and addressing data quality issues by providing technical assistance, training

opportunities, some cheerleading and motivation along the way. I don't -- I think sometimes we talk about like good cop/bad cop.

The CoC can serve as the bad cop and the good cop as in both the enforcer and the incentivizer and the HMIS lead is really that support person or support entity that helps the participating organization succeed with data quality. I didn't see a lot of responses in our poll about the HMIS vendor being involved in the data quality process within communities, but they do have a role to play or can have a role to play.

So the entity that holds the contract with the HMIS vendor does have the direct line of communication with the vendor, but any stakeholder that's involved in HMIS or data quality within your community can provide feedback on how HMIS can be improved to ensure effective and efficient data entry into HMIS and then that person or entity that holds the contract with the vendor should ensure that the vendor and the software is compliant with HUD standards.

Any feedback about system features or functionality that could improve these or [inaudible] increase data quality can be provided to the vendor and any responses from the vendor can and should be communicated back to the requesting stakeholder and the larger community.

So participating organizations, we know that they have a very large well-defined data quality since they are the ones entering that data into the system, but in order to understand the role that they play in HMIS data quality they do need to understand how the CoC's overall goals and data quality plays a part in the [inaudible].

So when participating organizations don't have a connection to the larger CoC and the CoC's priorities, data quality suffers because there's not that connection between the work that they're doing by entering the data into the system every day and the larger goals of addressing and ending homelessness across the community.

Also, what we've seen is that the organizational culture doesn't include an understanding of the importance of data quality and how the organization fits into the larger homeless service system. The end-users don't really connect how what they do every day fits into a larger picture.

There is an organizational culture of data quality that can be fostered and developed and it is a culture and it does start at the top. So making sure that participating organizations know where they can find resources to a system and their data entry efforts and then also make sure that HMIS lead is responsive to those [inaudible] is important.

Local and state funders, while HMIS is required for HUD and [inaudible], as we all know, state and local funders often play a huge part in the homeless service system within any given community, which means that they can help ensure a comprehensive HMIS by requiring that their grantees enter that data into HMIS and then also abide by the HMIS' data quality plan.

The reason why is because HMIS does have the ability to be a really robust and comprehensive data set about those at-risk of and experiencing homelessness in the community, but this does

often require the partnership with local and state funders that provide grants for homeless service providers in their community.

So I think sometimes there's a language barrier. So if we can find ways to partner with funders and speak their language and also show the benefits of HMIS and how it can be if you meet their reporting requirements for their grantees can be really beneficial. And we are actually at our action plan.

So our next step and kind of thoughts from you all about what your plan is for the next step that you will take within your community. So after going through the definitions and components of data quality and then the stakeholders who are involved, what is your next step and what will you take back to the community to go through and improve your data quality in HMIS?

So is it a what, is it a who, is it a when, is it a why, is it all of this? We would love some interactivity in the chat box to hear from you. And Meradith, I don't know if you have anything.

Meradith Alspaugh: It's okay. Mm-hmm. I was thinking about some questions when you were talking about that and the different sort of stakeholders and the roles that they play and I was thinking about a lot of the folks on the call are system administrators and I was thinking about the fact that all communities really are using data in such a different way these days and making dashboards and making these custom reports and making these visualizations for data and I am wondering if, in your community, are you doing dashboards or visuals on data quality?

Are you helping your communities see this data in a different way? I guess I'm just really curious about how you're even currently, if you are, sort of monitoring for this and thinking through what some different options might be for that.

I think that just one specific example of something that we did in my community at one point, we just made a list and like listed agencies or project names of who had the best of any particular data quality indicator and I think putting just a list out publicly gave people incentive to be at the top of the list.

We didn't have to give people prizes, we didn't give people things, just sort of an acknowledgement of here's where you are in our system made a big difference for a lot of communities and we really saw -- or a lot of agencies and we saw agencies strive to make those improvements and our community, we were sending data quality reports every month.

We got at a higher rate of male open indicators, we got a higher rate of data cleanup just by taking a look at a list -- it was really as simple as putting it in a list and seeing where people were. So it looks like there were some thoughts and feedback that came in here. "If your CoC is not committed to data quality how can you begin to build buy-in?"

That's a really good question. I don't know how your CoC isn't bought into data quality. So I'm struggling to come up with something, but I think if others are struggling with that issue or have overcome that issue, that would be really helpful to chime in there.

Meradith Alspaugh: Meradith, you asked about dashboard and [inaudible] -- I think using graphics and interactive dashboards that let people see data at a high level and then those who want to dig further into the data through those interactive dashboards is a really tangible step that you can take to build buy-in both at the CoC level and the other stakeholder level.

It's pretty -- it's not threatening, it's an easier digestible way to really show data and I've seen that be successful in communities for multiple types of stakeholders.

Meradith Alspaugh: Yeah. I think Doug's got a good point here, data presentation is important. From what he's seen, the cycle is present aggregate performance data to CoC leadership and boards, then relate that lack of data quality means that the data presented is not reliable, then build momentum to work on data quality. I think that's a good point.

Alissa Parrish: And I think going back to the question about if your CoC is not committed to data quality, you've got -- I mean, the system performance measures continue to be very large pieces, I know, for every year and then as communities are really HMIS to build out their binding lists for the purposes of coordinated entry, it's a goal to serve those that are most vulnerable in your community the data in HMIS has to be really accurate to make sure that you truly are prioritizing those most vulnerable in your community. So like think of that --

Meradith Alspaugh: [inaudible]. Sorry Alissa, I didn't mean to cut you off.

Alissa Parrish: No. Go ahead.

Meradith Alspaugh: I also wanted to tie back to what Alissa was saying about the role of the CoC here in this process, the CoC does have that carrot; right? They have the ability to allocate funding and communities have to have a process by which they determine how they're going to allocate funding.

Is there something about data quality in your CoC's processes for allocating funding? Like if you've got a project that has the best outcomes in their APRs but you're questioning the consistency or the accuracy of their data, how helpful is that?

So maybe -- and I am reluctant to say like take money away from anyone, but look at how those funding decisions are being made and look at how you could use data quality to help strengthen the case for certain indicators to be valued different, weighted differently, whatever that might be.

That's great. So Doug writes also that they have support from a university research center. They want to use HMIS data. So they're assisting with several TA projects, including data quality. I think that's a really cool example of getting buy-in in a different way. Like our data is meaningful; right?

Like we -- I think, as data people know, our data is meaningful, but I don't know that that always translates well to everyone else, but to tie that back to like other work or other opportunities, other ways to use this data is going to be really critical and I think even, too, again, CoCs are

doing pretty amazing things with data these days and doing warehouses and integrating other data systems in with your data and looking at child welfare or justice data, health data, whatever it might be.

In order to get really meaningful impacts and results from those, you're going to have to make sure that your data quality is as high as other system's data quality, like that you're able to contribute and have reliable data with all of these efforts and research projects and data integration projects that you're working on in your community.

Great. It looks like it's slowing down -- go ahead. It looks like it's slowing down a little bit in the chat, so we can move forward.

Alissa Parrish: And we just wanted to -- there will be links in the slide deck once it's available to these resources that dig more into your data quality. So they are on the HUD Exchange and the links are right there in the presentation slide deck. There is a question about what industry or field does data quality go?

Meradith, I don't know if you have an answer to this. I was going to say pretty much what you had said earlier about as we see more data integration projects and warehouses across country, like data quality on all sides of those data warehouses is really, really vital to make sure that we're matching data accurately and that the data is really telling an accurate story of the clients that are in those warehouses or integration projects.

I don't know of any specific field that's been touted as doing it really well. Meradith, I don't know [inaudible].

Meradith Alspaugh: I don't know that I do either, but I would agree, the integration projects I've seen go together. I mean, I think both systems have to clean up their data; right? Like it's inevitable that there's going to be something that someone has to clean up in their data, but yeah, I can't say that I can point to any one in particular and say there's is the best.

Alissa Parrish: So we didn't have any question slide on here, but we do have some time for questions if there are others. The data issues that are specified in the HMIS standard reporting glossary, that falls into the accuracy category, Meradith.

Meradith Alspaugh: Yeah. I sent a response back to that one. So there -- like in -- specifically, I was thinking of in the Q2 and the personally identifiable information, we call out those data issues around like rules for valid Social Security numbers, rules for valid dates of birth, those kinds of things. So in the regard, we're looking at accuracy in that particular section.

Alissa Parrish: There was a mention up above about if there's a push for timeliness, then completeness suffers; right?

Because we -- there's a balancing act between staff that are working directly with clients and there's the data entry side and there's the timeliness standards, then there's also a completeness standard and accuracy standard and I think there's different models or ways in which you could

start to address that, whether it's getting the baseline information into the system, when because of the timeliness and then building on that [inaudible] over time so you at least have that framework in the system for a client record [inaudible] collect more and add to that record.

Meradith Alspaugh: Yeah. I think that's a good point. I mean, collecting that data and starting with timeliness, like if you're entering that data as close to real time as possible, your odds for accuracy are going to go up, because either that person is sitting right in front of you and you can ask them double-check that Social because I don't think that's right or whatever the issue might be, you can check yourself and your data collection and entry in real time whereas maybe if you write something down on a piece of paper and you say I'm going to come back and do this later today or I'm going to come back and do this at the end of the month when I do my stuff and you get to your piece of paper and you're like I don't know if that was a two or a seven, so now you don't know that your data is accurate.

And so starting with those smaller pieces of timeliness that you can control to help build upon and improve your completeness and your accuracy and your consistency over time are going to be really important.

Maddie [ph], there's some resources on the HUD Exchange for a data quality monitoring plan. I think there's also the data quality management plan session. Is there another one of those or was that -- were those already completed, do you know, Alissa?

Alissa Parrish: There's one more tomorrow that will be highlighting some different tools available to look at roles and responsibilities. So some of the stuff we talked about today with the stakeholders and how to involve them, enforceable agreements between the different entities, of the sample data quality plan and some other tools.

So I think that's the last data quality 201. That session is tomorrow. I know they did one yesterday and I think there was a bed coverage one that was either today or tomorrow as well.

Meradith Alspaugh: Yeah.

Alissa Parrish: There was a question about the recording and the slides. So there isn't any recording or slides currently available on that link on the [inaudible]. I think those will all be populated after all of the sessions are completed. So maybe not end of this week, maybe next week, but soon after these are done, but none are available like right now. Do you see the question from Teague [ph], Meradith?

Meradith Alspaugh: Mm-hmm.

Alissa Parrish: So the clients are going in anonymously directly into your HMIS, is that what you're saying?

Meradith Alspaugh: I mean, I -- that's a tough one, because you are to be holding to your state and local laws that require certain factors for privacy or security and I'm not sure that I have a good answer for that question right now. I guess when we're talking also about data quality,

we're thinking about creating duplicate clients and avoiding that so as to minimize the amount of bad data in your system.

So I guess my thoughts on having anonymous clients is is there a naming convention, is there some sort of identifier unique beyond their name and date of birth and Social that you can make sure that they are the same client when you're going back and entering information about that particular client so that you're not creating a fake client when they're in shelter and then a fake client when they're in housing and then that same person getting another fake client when they're in a different housing project.

I mean -- so if there was some way to minimize the anonymous data damage, I suppose, on your system, but I don't have anything else specific to how you would deal with those particular clients. Did you have something, Alissa?

Alissa Parrish: I mean, I was also thinking about if your community has that data quality plan, are these projects included in that plan and how does that affect the overall data quality of the system? Because just like you said, Meradith, the duplicate or the potentials of duplicate client records is an issue?

Meradith Alspaugh: Yeah. That's a good, hard question to end on there, Teague. So thank you. I wish we had a better answer for you off the top of our head. But yeah, it is after 4:00 o'clock. So I think we probably should wrap up here. Thank you all so much for joining us today and hope you're able to participate in some of the other sessions later today or tomorrow. I guess there's no more today, tomorrow.

Alissa Parrish: Thank you.

Meradith Alspaugh: Thanks so much.

(END)