FINISHED FILE

ATLAS RESEARCH
SAMI COUTOURE/NELS R. BENSON
JULY 30, 2019

PDG B-5 DATA SERIES COMMUNITY OF CONVERSATION
SINGLE ID
2:00 PM CT

Services provided by:
Caption First, Inc.
P.O. Box 3066
Monument, CO 80132
800-825-5234
www.captionfirst.com

******

******

>> EVELYN KEATING: Hello, everyone. We'll get started in about five minutes. Thank you for joining us. If you have any questions for the presenters, please feel free to put those questions in the chat box either before we get started or during the session. Thank you.

Hello, everyone. Thank you for joining us. We've reached our start time. All participants have been muted, and you can press star 6 to unmute and star 6 again to remute yourself. This meeting is being recorded. Sorry. I will record the meeting right now. (Laughing) okay. The meeting is being recorded and we'll make the recording available to you in the next couple of days.

Check out the attachments in the download pod for the slides from today's meeting and other resources. This presentation has been made possible by the preschool development grant from the office of child care, the Administration for Children and Families, US Department of Health and Human Services.

We're excited to welcome our experts from AEM today. And without further ado, I'll pass this over to Jim Lesko. Thank you.

>> JIM LESKO: Thank you, Evelyn. So I'm happy to have AEM join us for our second in a series of webinar and Community of Conversation supporting the Preschool Development Grant and the focus is on data.

Today's conversation will be on single ID efforts within

states. So today, myself, though I'm just doing introductions, so no credit here, am happy to introduce some of my prior colleagues, Jeff Sellers, he's the TA specialist with AEM corporation and Jeff Watson who is also a TA specialist with the AEM corporation. And we are also fortunate because Jeff and Jeff have been able to help us connect with Hayley Young from North Carolina, Department of Health and Human Services, and Deborah Rodriguez, educational statistics director with the office of data quality, Pennsylvania, Department of Ed. And Hayley and Deborah will be joining us periodically during the presentation today to share with us their experiences around single ID. So this community of conversation is invited and we put out this call and announcement to all and we put all the invitation to the PDG B- 5 grantees. We asked them to share this with their partners and consultants. We do have the PDG B-5 TA team on the call today.

We also have as invited participants the federal project officers and regional office representatives swells our PDG birth to five TA partners from the State Capacity Building Center. We do want to remind everyone that this is a voluntary participation. We encourage our grantees to join the calls.

This remains a part of our ongoing universal technical assistance provision made available to the birth to five grantees as a part of the select presentations and resources be made available. We will be posting as Evelyn indicated, this recording on the PDG website which is linked to the child care resource center site.

As she mentioned, this call is -- we've muted everyone. We do encourage you if you do have a particular question to put it into the chat box. We monitor that chat box on an ongoing basis during the call, and we will make sure that we get to them.

I also do want to remind everyone two things. One is if you're on your computer and you're connected through your computer through this. Make sure you turn off the computer speakers. It helps to reduce feedback. Secondly, in the attachment that were made available in addition to the slide deck from today's presentation are a series of resources that Jeff and Jeff have recommended as a part of the conversation today.

These resources are provided through the courtesy of AEM corporation and the presenters. And they are linked to documents located on the state longitudinal data system website.

And so we'll have our presentation on the single ID. We'll have the overview from Jeff and Jeff. We'll stop periodically to listen and have a question with Hayley and Deb and respond to questions and answers from any of you joining the call today. And we'll have closing conversation towards the end. So at this time I'm going to turn this over to Jeff and Jeff. (Audio feedback)

>> JEFF SELLERS: To kick things off, we wanted to introduce a poll question for you. So if you can enter your response to the poll question, what do you think is a benefit to having a single child ID? Enter that into the chat box. Again, a brief little statement about what you think the benefit or the value of having a single child

ID would be for you in your state.

As we go through this presentation, just to kind of give you a heads-up, there will be references to a single ID or single child ID, a unique ID, unique identifiers, all of that is really referring to the same thing and is just referencing a unique way to consistently identify a child. So throughout our presentation today you'll see various ways of it being referenced. Just know it all means the same thing.

So we are getting some comments here. Reduction in multiple IDs, continuity, data linking across systems, longitudinal data and duplication of efforts. The linking of longitudinal data and reducing duplicate accounts. Several references to being able to establish a longitudinal look of the data and of the students and children over time. Reducing duplicate counts between programs so that challenge that many times that comes about in trying to uniquely identifying a student or child and the ability to know all the services that child's being provided.

Let's see possible linking to other state agencies as well. Good. Looking at other state agencies. Good. These are all good comments relating to that single child ID. So, again, kind of overarching looking at the ability to track a child over time, longitudinally, linking to other program areas, linking to other agencies. So we're going to be talking about some of that in our presentation today and seeing some of the benefits from both North Carolina and Pennsylvania.

So just a couple of quick comments. What is a unique ID system? These are kind of characteristics of such a system and things to be considered as you adopt such a system.

First, the system assigns a unique ID to each child as they enter into the program. Again, that is fairly straightforward but, again, as -- there are considerations that need to be considered going into that. Again, this -- what we're describing here is really a process on the front end that would assign a unique identifier to a child as they are entering into receiving some of these services.

Another really a consideration and really an agreement that would need to be made is that everyone who is providing a service everywhere where that child exists in the system, that unique identifier would need to be assigned to that child. So there needs to be agreement and coordination between systems, between services, program offices, in assigning and using that unique ID.

As the child progresses through the system, if you will, with the various services they are provided, the ID would need to follow that child as they go. And this last bullet is really just something to keep in mind as far as what the benefit of having that unique identifier for that child preK-12 to have in place a process that could actually follow the ID would follow that child into the K- 12 system and beyond. So the continuity of service and the ability to track that child longitudinally would be within the context of early childhood but could extend through their entire educational career.

So why have a single ID? What's the benefits? Several of

you have entered those benefits into the chat box. Just a few here. One is the ease -- to ease the process of that child as they've transitioned from a given -- from program to program or even from agency to agency. So as that identifier follows that child, the ability for them to transition would at least the process of transition would be eased.

The facilitation to have an integrated process -- Hayley will talk more about that as far as what they're currently doing in North Carolina, but it is part of the challenge when it comes to integrating data, in other words, bringing data from multiple sources, multiple programs, multiple agencies is to accurately connect that data to the child so having a single ID helps in the integration process.

Another area to minimize the need for that matching process, the algorithms that come to play. An integrated data system would leverage but having that single ID would minimize the need for that and probably increase the ability to accurately match those student records together.

Along those lines obviously you would have fewer matching errors. You may needless manual intervention to reconcile those errors. The last two bullets here are really kind of around related to data quality. The first one here, fewer uncertainties when reporting on or evaluating program enrollment and impact. These represent some high stakes kinds of reporting and evaluation. You definitely want a high degree of confidence in the quality of the data and the establishment of those linkages to that child as though having that single ID kind of helps in your confidence when you're doing the evaluation, when you are reporting out on the performance of a given program or child. That single ID helps in that degree of confidence.

Then lastly, more accurately count children and the services they are providing. Someone made a comment about the ability to accurately know all the services that a child is receiving. And so again that single ID would help facilitate that as well. So much of this, what I've been talking about here really does address that idea of establishing that single ID on the front end as a child is brought into the system and starts to receive services.

Another approach is after the data has been collected on a given child is to look at how to integrate that data into a single system and to be able to compile that and so Hayley is going to share a little bit about how North Carolina has been able to do that and to leverage a single ID kind of after the fact, if you will, and to expand their abilities in being able to look at the different aspect of a child and the services they provided.

So with that, Hayley, I will hand it over to you.

>> HAYLEY YOUNG: Thank you, Jeff. So -- we can actually advance to the next slide. Thank you. So I'm going to talk a little bit about North Carolina's early childhood integrated data system which uses a single unique identifier to bring together data from multiple early childhood programs in our state. For us our early childhood data system consists of data from various programs in the

areas of early education, health, and social services.

This data system was initially built in using race to the top grant funding back in 2012. And the system has been live since 2016. As you'll see on this slide, the idea behind our data system was to collect information in a single data system to give us better insight into how early childhood services are being used across our state. And by being able to look at that data, we would be able to answer key program policy questions to allow us to make better decisions about the way that services are being used in our state.

And just to give you a little bit of a sense of the way that unique identifiers work for us, we use a technology platform called E-Scholar to identify -- I'm sorry -- to do unique identifier assignment for all of our data. If we could go to the next slide, please.

So in terms of the programs that are currently part of our data system, you'll see on the left-hand side we have many programs. One thing that I didn't mention yet is that our data system is for children ages zero to five. All of these programs that are currently in our data system for the most part fall under our department of health and human services but we also incorporate data from a couple other data sources outside of the department of Health and Human Services. On the right-hand side you'll see other data sources we're looking to incorporate. And some of these programs are in active development now for integration. So in particular, we are working actively to integrate home visiting data into our ECIDS. We're also actively working on enabling integration to our education data and our K-12 data system.

Next slide, please. As I mentioned this system has been live since 2016. So we feel proud that you can now say that we have actual integrated data which has allowed us to look at unduplicated counts of where children are being served and by what programs they are being served by.

We're also able to do things like link data to answer questions that in the past could not be answered. For example, we can use the unique identifier for part C data and link that data to part B data to look at what that link data looks like.

We also have a research request portal which is available on our website. So this is a place where researchers can go in and submit an IRB approved research request for child level data. And just to emphasize this point, we have many different pieces of that research request process in place to ensure that all of our data is highly protected and only shared if a vigorous research request is made.

Next slide, please. So what we're really excited about with the preschool development grant is we actually have the opportunity to look at the data system that's been live since 2016 and has been in what we're thinking of in a maintenance phase of the system. So the lights are on and everything is running. But what we would like to do now is to expand and improve upon that system. So the PDG gave us a chance to really evaluate the way that the system looks now and

what we can do to improve it going forward.

On this slide are some of the initiatives we're engaging in this year and years to come to expand and improve upon our ECIDS. The first bullet is realigning our data system to make sure we're current to our initiative in early childhood. As an example we as a state have just released our early childhood action plan which is a strategic plan to reach quantitative targets by the year 2025 state-wide. That plan actually looks at children ages birth to eight. As I mentioned earlier, our data system only looks at children zero to five. So that's one opportunity where we could potentially expand our data to look at that full range of early childhood and be aligned to a state strategic initiative.

As I mentioned, we're also working on integrating home visiting data in our state. Home visiting data right now in North Carolina is collected in a variety of ways, depending on the model that you're looking at. So what we're working on now is aligning data from those different models and being able to integrate that data into ECIDS so that we can look at the full range of services that children are receiving including home visiting.

Our state has also just conducted a data user survey. So this a survey to look at where there are gaps in our early childhood data, and then using that information we can make decisions about whether to expand the data sources that are currently part of our ECIDS.

Then I kind of alluded to this a few minutes ago, but we are also actively working on connecting to our K- 12 data system through what we're calling the education longitudinal data system. So this is what we're referring to as a system of systems which would include our ECIDS which is the early childhood integrated data system, but it would also include our PW20 system which is that broader K- 12 data.

We're also looking at enabling data visualization and exchanged reporting and just continuing to improve on that.

Next slide, please.

>> JEFF SELLERS: Hayley, real quick. There's a couple questions that have come in.

>> HAYLEY YOUNG: Yep.

>> JEFF SELLERS: So I just wanted to make sure that we get these asked for you. First question is, have you experienced different barriers when trying to integrate data outside of HHS compared to integrating data within HHS? If so, what are those -- how did you work through them with the different stakeholders?

>> HAYLEY YOUNG: Yeah. There are a couple of thoughts there. So the first is that what we've experienced when working with an outside agency, the really important thing is to develop really strong data sharing agreements from the beginning. I think that that really is what is the most time consuming part of actually getting to integrated data for cross-agency data sharing. That's the first thought, ensuring that the structure is in place to ensure that you have really strong data sharing agreements and you have all the right people at the table to develop those.

The second thought is that, depending on the data source that you're looking at, you may be connecting to, for example, a data warehouse that lives within a state agency, but for certain programs, you may also be looking at potentially receiving a data extract from a third-party data vendor. One example would be Head Start data which sometimes that data is actually stored in a third-party data vendor, for example, Child Plus. In that case, you would have a different process where you would need to actually work with a third-party data vendor to get a data extract to get it integrated into your data system. So those are a couple thoughts for that question.

>> JEFF SELLERS: Thank you. Do you see the second question there as well?

>> HAYLEY YOUNG: Yes. In terms of expansion priorities, I would say that in my experience one thing that's important is to think about some of those state priority areas where you can use data to support some of the strategic initiatives that your state is putting forth.

So I mentioned our state has released a strategic plan to improve early childhood outcomes in our state. And so I think it's really important to think about as you're developing a data system like this, the way that you can design it so that it is able to support some of the key questions that you have to support priority initiatives.

And then I see we've got another question. Are we working with data vendors to automatic data extract and PPIs?

So we recently -- there are a couple of states that actually have started conversations around using an API, for example, with Head Start data. We currently do not use APIs for our data extracts, but it is something that we're exploring further for some of those trickier data sources where we do have to get a data extract.

>> JEFF SELLERS: Very good. Thank you. Next slide.

>> HAYLEY YOUNG: Yeah. So I think we went through a couple of -- I think some pieces of this slide already. But for states that are sort of just starting off in this process, I've listed a few things that I think are important to consider. First of all, who are the end users of your data system and how do you want them to be able to access data?

For example, is it important for your state to have public facing reports and data visualizations online for anyone to access, and what would those look like? You know, for us what we've been considering is what are ways that we can expand the types of users that would be able to access our data better, for example, with policy making around early childhood.

Another thing to consider is the type of data system that you want in place. There are a few different types of data system models that you can use. In North Carolina we're currently under a federated data system, which means that all of the source data is still living in the source data warehouses where data was initially stored. So we just access that data as needed for our data system.

There are other versions of models. So, for example, you could have a data warehouse where all of your data is stored and managed in one place.

I also think that sustainability is a very key consideration, especially in terms of technology. So thinking about not only what are the upfront costs of developing a data system like this, but also what are the ongoing costs you should expect. We kind of touched already on short and long-term priorities. But considering what strategic initiative you want to align your system to so that you can support those initiatives using data through this type of system.

>> JEFF SELLERS: Okay.

>> HAYLEY YOUNG: That's it for me. Now I think I'm turning it over to Jeff.

>> JEFF WATSON: The other Jeff.

>> JEFF SELLERS: And to wrap up -- right. I'll turn it over to the other Jeff. But I just wanted to thank you, Hayley, for the North Carolina update, if you will. Also, too, Hayley mentioned this ECIDS or ECIDS. Again, in case you're curious if you're not familiar with it, it is an early childhood integrated data system. It's an opportunity to take existing data systems, and it could be applied in a variety of ways, either in early childhood or even beyond and put a process to be able to extract those child records and then establish a linkage of those records to uniquely identify a given child and to attach, if you will, related data and information for them so that you can in turn run reports and different kinds of information as Hayley described that they're doing in North Carolina. That's really the reference to ECIDS is integrated data system specifically related to early childhood.

So with that, I want to hand it off to my partner here, Jeff Watson to continue the process of our presentation.

>> JEFF WATSON: Great, thanks Jeff. Thank you, Hayley. That was a wonderful overview of what you're doing there in North Carolina. So let's just talk briefly about a couple of different models. So there are a couple ways to go about integrating across IDs. One is to have a centralized system where you're essentially requiring all vendors to use a specific and unique ID for a student.

And that's different from, I think, what is happening in North Carolina where vendors are actually using kind of their own individual IDs, but then it's federated up a layer. And those IDs are merged across.

There's also another flavor or another twist on that where your IDs within your individual systems may not be unique, and that can be very problematic. I think that's a situation where we would all quickly agree to avoid at all costs, because then you have to worry about another layer of potential confusion about which child record goes with -- how do you match across systems? How do you match children's records across systems?

So I will mostly be talking about the first two in this slide.

So under a single unique ID system when a child comes into one of your programs that's under that umbrella of early childhood,

your system will need to check to see whether or not that child is in the system. And if not, then assign an ID. And that process is a critical process. You want to make sure that you're only assigning IDs when in fact a child is not already in the system. And so that needs to be designed with care, implemented with care, and managed with care.

That ID, once assigned, should follow the child as they exit and enter into other programs and receive additional services. And so all of that care that goes into that initial assignment needs to be replicated at each entry point and to each program to make sure that when you look up a student, you're in fact matching the student that's standing right in front of you with the records from that student's past.

So if you do that correctly then, all of your systems have each child assigned to unique ID. And that ID then is what you would then map on to your other agency data or your cross-agency data, might be a better term, such as K- 12 data or workforce data or post-secondary data.

So that's kind of the overview of having a single unique ID. I would like to note that it is still possible to have matching errors in that process. And so, for example, if a child receives services at one point in time and then exits and then comes back into the early childhood system after some period of time and presents the program, the second program with a slightly different set of information, perhaps they moved, perhaps the parents divorced and there was a name change, or perhaps even the child is spelling their name differently now and they've gone from Billy to Bill or William to Bill. And they present their information, their individual information differently, it is possible to sort of have the situation where the same child gets two different IDs. Likewise, it would be possible, at probably less likely, to assign two children to the same ID.

So, again, the point there, I think, to think about is that while this is a technical solution at the core, it's very important to involve stakeholders, end users from the beginning so that as you design your system, the process is aligned to how people go about the work of assigning and intaking children into their programs.

If you're thinking about a multiple ID solution, then each child will receive an ID for each program as they go through or from one program to the next in your early childhood umbrella. And those records will need to be matched with some sort of technical intervention.

And that matching process will probably rely a great deal on the child's demographic and name data. And so if you think back to Jeff Sellers' opening comments, there is some inefficiency being introduced here, because each program sin taking a child and collecting and executing data entry for all of that same information, so you do lose quite a bit of efficiency on the intake side, but then more importantly you're going to have to build a layer where you have a process for matching across programs based on that demographic information.

So when you do end up matching across multiple IDs, you are essentially making a guess about what record goes -- which records go together. There are two types of errors you can make when you're making a matching decision. A false positive occurs when two different children are linked under the same ID. So you've decided that these two children -- these two records, although they're different in a little bit, in some minor way, they actually represent the same child. So then you decide to represent them as the same child.

So you're at odds with reality at that point.

The other type of error would be to -- would be failing to connect records that should be connected or matching records that should be matched. And so an example of that might be you have one child that's going by William in one system, but a couple years go by and then he goes by Bill. Everything else looks pretty close, but those two first names are pretty different, so you decide not to match them, in which case you've made a false negative.

An example of false positive occurs often with twins where all of the demographic information may be identical except for one letter of the first name, Tim and Jim, for example. So sometimes you'll see as you build out our technical layer exception rules for dealing with false positives that occur from identical twins.

So things that can introduce error into that matching process, again, different name spellings, emerging nicknames, data entry errors, entering a date of birth, May 3, 2018, if you transpose the 0 and 3, May 30, 2018. Sometimes people identify race and ethnicity differently over time. Parents names' change, guardians change. Mobility between service areas, in and out of service areas, in and out of state, back in to state, all of those things can kind of create a lot of uncertainty when you're looking at demographic variables.

We can go to the next slide now. So one of the key things that I thought Hayley touched on really well is the need to get people involved. I'm sorry, I can't advance the slides from my controls. So -- there it goes. Thank you.

And so I would encourage everybody to really think hard about data governance. Data governance is a process and an organizational structure that's designed to manage an organization's data assets for the purpose of achieving their organizational strategic goals.

When you do data governance, you will establish roles and responsibilities and best practices related to data management. You can see how a lot of what we've talked about really falls under that. You're going to need to have a pretty prescriptive process for intaking students or children and their families. When you do data governance well, you'll get better definition about organizational processes, roles and responsibilities. You'll be able to communicate better about who does what and who are the subject matter experts for the different parts of data that you're collecting.

As a result, you'll have better documentation, your data quality will improve, you'll reduce the inconsistencies and how data is collected and managed. You'll probably have improved

coordination between your programs, maybe even between your technical IT groups and your functional groups as well as your vendors. You'll have improved collaboration and hopefully you'll see a reduced workload that's associated with data cleaning and data documentation while the overall utility of the data will improve.

So there are resources out there for data governance. And then I'll just go to one more slide, do a quick time check. I think we're okay.

Privacy considerations is our next slide. Again, this has come up before, so I'll just reiterate some of these themes. Of course, you'll want to take care not to reinvent the wheel and leverage existing structures in your state that can help you with developing good data sharing agreements, memorandums of understanding, making sure you're following local and state, and federal laws, FERPA. You'll want to reach outside of your team and make sure that you're leveraging other teams that may exist in your state so that you can avoid reinventing the wheel if possible.

I'll stop here and pose another question to the audience and use the chat window and tell us what do you think are the biggest challenges in your state to establishing a single unique student identifier?

>> JEFF SELLERS: While we're getting response to the poll question, there was an additional question asking what an API is. Can you briefly describe that?

>> JEFF WATSON: Sure. An API is a relatively new approach to moving data between applications. When I say relatively, it's not brand new. APIs have been around for long enough to actually become the industry standard for moving data in many cases. People start with APIs and then build out the rest of their applications around that.

There is quite a few resources on the web around APIs because they are so common. So I don't know if that's sufficiently detailed answer. Feel free to follow up that question if you have any follow-up questions. Essentially an API stands for application -- application protocol interface is the acronym. It's essentially a way to send data between two ways, by the way, between different applications.

So you can think of it as a way of integrating data across. You could send data both ways, which is a little easier than how you might do with traditional ETO tools.

Okay. So let me just scan through the answers for where the biggest challenges are. So I'll start at the bottom, unduplicated counts and being able to track across multiple systems, easier to collaborate with other agency, easy ability to -- I'm not sure -- I see my chat window wasn't scrolled all the way down.

So challenges, yes, relevant agencies to agree on having a unique ID. Absolutely. You do need to make sure you have conversations early and often and develop those MOUs and develop the intra-agency agreement to move forward with that approach. Whether to build it internally or buy off the shelf, what do I recommend?

(Laughing) I wouldn't recommend anything right now. But I would recommend considering that question carefully and ultimately make sure you talk to your stakeholders and have a vision for what you need your system to do and who your stakeholders are before you get too far into that decision. Start with your stakeholders. Start with your end users and clearly articulate what data you need to collect, how you're going to merge across IDs, and how all of that fits within your organizational strategic goals. Interdepartment jurisdiction, lack of legislative mandate, indeed, politics, absolutely.

The problem of silos is still very much a real problem. Silos exist. And we run into them every day. Establishing the data governance structure and updating Legacy software. Indeed. There's a lot of work to do here. Again, I think it's really important to focus on what do you need to do -- what are you trying to accomplish, and where can you start, and what are some maybe other things that are longer term efforts?

With that said, those are great answers. I'm going to hand it over to Deborah Rodrigues from Pennsylvania Department of Education, so we can hear about your experience, her program's experience with single child identifiers in Pennsylvania. Deborah?

>> DEBORAH RODRIGUES: Thank you, Jeff. I have along me, I'm going to share as we go along here, Michelle key from the office of child development and early learning. That is a Department of Human Services office which is also a Department of Education office that was the very first agency we shared data with and having a foot in both agencies enabled that.

I also have Tim went who is going to talk about our sharing with data around foster children. And finally I have Missy Conyer who is our PA secure ID expert to answer any questions people may have.

Pennsylvania began individual student education ID in 2008 with our K- 12 students. We are also using the E-Scholar unique ID product. It is a separate system from our state longitudinal data system. The only purpose of unique ID is to create and store each student's individual student ID.

This is an ID that should follow them from the very first a time they enter education or educational services from early childhood through post-secondary into adulthood. We try to help our agencies and institutions understand this is much like a social security number. It should not be changed. It should be one number per one person. And obviously that enables us to share data accurately.

It is -- we developed a link. Again, Pennsylvania has a federated system. Each agency is maintaining their data in their own system. We have a -- Michelle will talk about how we create PA secure IDs for our early childhood students. We have, what we call, a bridge. And it basically links the education system with the early childhood system to allow aggregate data to return, so the early childhood program office can see how their students are going to

evaluate their programs once their students get into third grade and above where we have state testing.

It's automated in the fact that you can upload a file, and if there are no children in the system that look like the data you entered, it automatically assigns the ID to the child. If there is a student in the system who exactly matches the record you uploaded, it will give you that student's ID back, not create a new one. However, it is a manual process in that there are going to be a number of records every time you upload a large file where it may or may not be the same child or there are multiple records. We call that a near match, and that takes manual intervention to look at the records and determine whether or not it is the same child. And then assign that number or find another -- create a new number for the individual.

There are some policies that really help. These policies, I'm sure, are more difficult to enforce on the early childhood side, but it is essential to enter a student's legal name. There are a few exceptions where that may not be possible. But 99% of the time if it is possible to verify that name by some legal document, no nicknames should ever be used. We've already heard what happens when William has been entered as bill.

Last name, first name, and while middle name and middle initial are not required, they are highly encouraged and recommended because the more information in the record, the better the system will work, and the less likely two students will share the same ID.

We first did this in 2008. I'm surprised at the number of times a name that looked very unique actually had someone else with the same name in another part of the state. Again, birthday has to be accurate.

One of the lessons we learned early on is even though a child might be eligible for services, before they are born, don't allow a due date to be entered as a birth date because rarely is the birth date the due date, and that causes problems down the road.

Another place we've had issues is when a student goes into post-secondary education. There aren't quite the same guidelines around collecting information from students on the post-secondary level, and sometimes it became difficult to get the legal name of the student.

Also, what we found in post-secondary, a hole we had to plug early on was because birthday was not required, some institutions were just entering 1-1-1900 for everyone's birthday and because birthday has such a significant part in determining if it's the same child, it always assigned a new number preventing us from linking the student K- 12 with their post-secondary. We obviously plugged that hole.

We highly recommend -- and this was not something we did in the first year of assigning a PA secure ID, it wasn't until the second year we actually started using that ID to collect data. We highly recommend that this information be shared in a confidential manner the same way you would share something like social security number

when a student moves from one educational institution to another. That could be transferring from one school district to another school district or on their transcripts going into post-secondary education. That's obviously the place where the receiving institution can be sure that they have the right number for the student that they are receiving.

I'm going to pass on to Michelle who is going to talk about the early childhood.

>> JEFF WATSON: Sorry to interrupt. I would like to do a quick reminder on our time. We are at 2:54, so we have about five or six minutes left.

>> MICHELLE: All right. I'll be fast and for sure on time. This is Michelle. I'll just add a few things to what Deb already shared. So we do -- we have a lot of people entering data into our early childhood system which is Pelican, so while we recommend they use legal name and date of birth, that doesn't always happen. So I think our biggest takeaway from this whole process of having a unique ID is our data quality issues. We do have headquarter staff that process our PA secure ID near matches. But our children in our office also get a unique ID that links to our Department of Human Services I would say our lesson learned is we have data quality issues and we're trying to alleviate that with children having duplicate IDs. I'll pass it over to Tim since we're running out of time. Then we'll get to the questions.

>> TIM: Tim what I've been doing with the PA secure ID and DHS is working with them to identify the foster students. So DHS has been providing a file to the Department of Education. It is their APCAR file which is a file that they submit to the federal government two times a year.

And basically we pull the demographics off the app cars file and try to match them up with PA secure ID here on our end to see if we can find an ID in our system that DHS has already identified as a foster student. And once we identify that match, we provide that information back to DHS and we also keep that information for ourselves so that we can share some reports with DHS and provide them with the analysis that they need for their foster students versus students that are not in foster care.

Once again, that's done with all the demographics. Not every single child gets matched about 90% of them do. Then we have a manual process where we have to look at each individual ID student demographics to try to identify that student.

>> MICHELLE: Did you just want us to answer the questions? How do you want us to proceed? We only have three minutes.

>> JEFF WATSON: So let's take the last one first because it's probably a yes or no. Does Pennsylvania have a state-wide student ID system? I assume that's a K- 12 student ID system.

>> DEBORAH RODRIGUES: It is the K- 12 system but it is where we are maintaining the records for the early childhood students as well.

>> JEFF WATSON: Okay. Excellent.

>> DEBORAH RODRIGUES: Just the ID record not all the detailed records that the Department of Human Services has.

>> JEFF WATSON: How does Pennsylvania handle situations with a child has two last names and sometimes that gets recorded differently depending on how you do the data entry?

>> DEBORAH RODRIGUES: That's where we've strongly encouraged the legal last name. If it's a hyphenated name, the system will accept it. It is obviously a problem. It causes near matches. The one thing that we did a few years ago which greatly helped us is we created a linkage between our PAID system and our state longitudinal data system, when you upload a record into the LSDS the PA secure ID and the student name, anything else on that don't match, then the record won't upload. Your file will fail. So that kind of forces and date of birth as well. That kind of forces the correct information.

>> JEFF WATSON: Right. That makes sense. Well, thank you very much for sharing that. I would like to use our remaining minute to touch on some highlights that came up all the way through today's discussion.

So things to consider as you move forward with the single ID system. Our early childhood program already established in your state somewhere? What is the overall design of your state's integrated data system? What is the scope that you need out of your unique identifier? Are there already people in your state doing matching procedures? And a few more. So I think today's discussion's been very rich and smattered with both a little bit of theory but also a whole lot of practice. And I personally appreciate that very much. I would like to thank Deborah and her team as well as Hayley and her team, Jeff and Jim, and everybody else for joining us today.

Jim, did you want to wrap us up?

>> JIM LESKO: Yes. Just -- let's see. So as I mentioned previously, there is a list of resources that have been posted. Thank you to Jeff and Jeff for making those resources available to us. If there are some questions, I think we can get to most of the questions today. If you have further questions following data, you can either email those questions to us at package B5TA team you see the email link there. You would love to hear your feedback and suggestions. You can actually click on this survey monkey link on this PowerPoint slide. It will take you immediately to the survey link. We would love your feedback.

We will have additional webinar events and Community of Conversation through the end of September, and your feedback is very helpful to us to customize and make sure we're meeting your direct needs.

And you will, for those of you who registered, we will be sharing the link to the recording and an additional link to the survey monkey if you can't get to it today. We would love to still hear your feedback. I do want to thank Jeff and Jeff and Deb and Hayley for making your time available today. I know that there's some

really fascinating information. I know it was very helpful of the I can imagine many of you can will be doing some follow-up connections. If we can assist you in any way with those connections, please reach out to your PDB B-5 TA specialist. We'll be able to facilitate those linkages and or answer your questions. Thank you to Evelyn behind the scenes our webinar manager. Thank you so much for making sure that everything worked as it needed to.

And we'll end. We'll have a -- the next in series of data webinars by the way will be on August 8th. That will be on data privacy. Thanks so much, Jeff and Jeff, and Deb and Hayley. Bye.

(Meeting concluded at 3:03 PM CT)

******

******