Transcript for the Make your data count webinar

Hello everyone, I would like to extend a warm welcome to you all to the Making your data count webinar for the Data Exchange.

My name is Rose Pearson and I am with the Data Exchange training team and I will be taking you through this training session today.

Before we begin, "We are meeting today on many traditional lands around the country and I want to acknowledge the traditional owners of those lands and pay our respects to elders past and present."

Ok so let's take a quick look at what we will be discussing during today's webinar.

We will look at data best practices – so what are the data principles that are the foundation of good data, and also looking at what your data is used for

- We will take a look at what constitutes poor and good quality data;
- I will also briefly demonstrate the organisation data quality report and I will talk through that as well as the fact sheet; and finally
- Where you can get help if it is required.

In this webinar we will not be discussing getting an AUSKey or applying for access to the Data Exchange, if you need more information on these items then please visit the Data Exchange website and Training resources tabs and you will be able to find further information on those topics there.

So let's get started.

Quality data is built on a strong foundation of data principles. When we talk about data principles, we're talking about data that is timely, reliable, accurate and complete.

Following these principles will provide us with confidence in our data and allow us to have:

- Visibility and understanding of the outcomes that are being achieved by clients
- The ability to identify combinations of services that lead to the best results (across programs) for clients
- We will also have a view of client pathways within and across services – so we can see which other services clients might be accessing, and we have
- An increased understanding of the client groups demographic profile and complexity

The Data Exchange system and supporting policies are built to support collection of high quality data both relating to the mandatory priority requirements data set, as well as the partnership approach, which includes extended demographics and client outcomes data.

And it's important to remember that these data sets feed into the bigger picture – improving the lives and wellbeing of individuals and families by ensuring their access to quality social services.

So what do we mean when we talk about data quality?

We need to ensure that the data is uploaded and entered into the Data Exchange that it's of good quality, it's accurate and is checked and updated regularly.

- If you have accurate and quality data it then provides your organisation with a true reflection of what is being achieved by the services that you are delivering on the ground.
- It will provide you with evidence of what is working and what isn't working. It highlights if you are servicing the right client target group.
- It will provide a strong foundation of solid data to make wise decisions and support any future growth and implementation of initiatives that you may have.
- Accurate and quality data will also provide strong evidence to support your grant and funding agreement milestones and performance indicators.
- You can go in at any time and check your data via the reports, update it and monitor your progress.

For the funding agencies, all the points mentioned for the organisation:

- Better understanding of what is occurring for clients and if this is expected
- Ability to make informed decisions about policy, funding and programs
- Assists in monitoring program performance.

Now data that is provided by your organisation is valued by the Department and has many uses.

Your data:

- Informs policy development through input into evaluations for the Department to get a better feel for what is happening on the ground
- Is the main data source for the Department's Annual Report
- In the future your data will be used to support the Try Test and Learn project evaluations – where the object of this program is to reduce long term welfare dependence and assist some of the most vulnerable in society into stable and sustainable employment
- In the past, your data has also assisted with disaster recovery planning for Cyclone Debbie which hit our shores in April 2017. Now the data was useful in that we could look at which services were forecasted to be in the path of and be impacted by the cyclone, and where clients could access similar services in other locations that weren't affected by the cyclone
- Your data also highlights the breadth of client pathways across services; and
- Demonstrates the return on Government investment.

Now as a quick reminder: As many of you know there are three different ways to submit data into the data exchange.

XML and system-to-system are applicable for organisation's wishing to use their own client management systems to meet reporting requirements, and your IT vendor or specialist will modify your client management system or CMS software to ensure that the relevant fields are either automatically or manually uploaded to the Data Exchange at regular intervals.

Technical Specifications can be found on the Data Exchange Website. The Data Exchange Helpdesk can assist with technical questions and a staging or sandpit testing environment.

The third option is the free web-based portal, which allows a user to directly enter and report their data.

The web-based portal allows for recording client, service and outcomes data that meets all the data requirements for your program and allows you and your staff to confidentially manage your client information.

User support task cards and eLearning modules are available on the Data Exchange website under the Training resources tab, to help users navigate the portal and enter their data effectively.

The Data Exchange allows organisations to enter their data as often as they wish to within the relevant reporting period.

- So, you can enter data daily, weekly or monthly during the specific reporting period
- Entering data early and regularly is in your organisation's best interest as it will ensure that the data quality in your reports is optimum as it will be up to date and relevant.

So, if your organisation is only entering data every 2 or 3 months – the question to ask is if this frequency of upload is providing your organisation and the Department a true indication of what is happening with your clients and services?

- Entering data regularly will ensure that you are meeting your grant agreement milestone and KPI requirements
- And it will also ensure that you avoid delays with the data being submitted on time to the Data Exchange as sometimes, increased traffic may result in slower than average performance uploads
- Also submitting your data regularly will allow you enough time to not only check the quality of your data, but afford you the time to fix your data if you have found errors in the upload or data entry.

So it makes sense to submit your data in on a regular basis because if it doesn't get submitted and accepted to the Data Exchange within the reporting period you will have missed the cut off date and you will not be able to enter it in at a later stage. Your data will not count and your organisation and the Department will miss out on this valuable information.

So the key message is to get your data in early for it to count.

Now we are going to look at what is means to have good and poor quality data.

With anything we want to make sure that we are hitting the mark in regards to not only submitting our data regularly and on time, but also that we are submitting quality data.

So we want to be data smart. That raises the question, what does poor quality data look like compared to good quality data? What do you think?

What we are going to do now, we are going to look at some examples and you can sort of think to yourself is this what is happening in my organisation with my data?

So the first example is that of unidentified group clients

I would like to remind everyone that unidentified or group clients should only be recorded when it is not practical or feasible to collect individual client details such as large-scale community events.

Group clients are clients that do not have an individual client record in the data exchange, so we do not have a first and last name, a date of birth, gender, address details made up of town/suburb/region, state or postcode. There are no details recorded for indigenous or disability status or CALD information made up of Country of birth and main language spoken at home.

With many programs, there may be situations where unidentified group clients will exist and the expected amount allowable is provided by the policy area and is set out in the programs description found in Appendix B. I would encourage all of you to take a look at that information if you haven't already done so.

Now because these demographics aren't collected it means that there is no valid Statistical Linkage Key or SLK created for the client and the data cannot be used in reports. When we think about SLKs, an SLK is useful when we want to view a client's pathway across services and across organisations. It allows us to get a picture of the different types of services clients are accessing. Now remember that the Department does not see your client information, as before it's transferred to the Data Exchange it is de-identified.

Also whilst unidentified group client numbers can be used in the Annual Report it doesn't provide that beneficial background information and this is a missed opportunity to demonstrate the outcomes you are achieving with individual clients as you can only record group outcomes against group settings.

Now in your reports you will get output information about how many unidentified clients attended how many sessions. However, in regards to demographic information, this is what a report looks like when only unidentified group clients are recorded. There is no information available. There is no age, gender, indigenous or disability status recorded at all.

Compare this to when we do have demographic information available from recorded individual clients. Just to reiterate, the Department understands that there are going to be instances in program delivery where due to the numbers of attendees at a group event such as a program BBQ or school information session that it is going to be impractical to collect data for all clients in attendance, however, if you find that this is a regular occurrence for your organisation and program then speak to your funding arrangement manager and let them know what is happening.

Let's now take a look at incorrect date of births. So poor quality data regarding date of births limit the usefulness of the data and can potentially impact the data of a program. Age groups demonstrate part of the standard demographic profile for clients required for many government programs and is of particular importance to programs that target age-specific cohorts. The example on the screen highlights a high number of individual clients over 105 years of age. This occurs as a result of entering an estimated date of birth of 1900, but this is not the correct way to enter date of birth's as this will skew the data and make it inaccurate and unreliable.

Where a client doesn't know their date of birth or does not disclose it, it is acceptable for an estimated date of birth to be recorded**.** So if the client thinks that they are around 30 years of age then the date of birth that should be recorded by an organisation would be 1st January 1988.

There also may be date of births that are not appropriate for the program activity, so for example you might be delivering a program that is targeting 12 – 18 year olds. So you can look at your data and see that there have been individual clients recorded outside of these age groups – and again this may skew the data for a program.

Another data issue that we are seeing is the high instances of not stated responses for demographic questions. The not stated option is found in the listing of disability status and indigenous status, main language spoken at home and country of birth. Now with these fields the information collected is as the client self identifies – so if the client is asked the question and they do not identify with any of these items then the response selected would be None.

Not stated means that the client either hasn't been asked the question or they have chosen not to disclose. If the client doesn't self disclose - could you ask the question in a different way? If the client is not comfortable in providing that information then the response would be Not stated.

These fields can be updated in the Data Exchange at any time so your data quality can improve over time with these changes and updates, and remember that asking these demographic questions is mandatory but it is up to the client to disclose.

On the screen you can see the screen shot for no paired scores. Now it's something that can been seen in the client outcomes report available to those organisations who are participating in the partnership approach.

A paired score means that a client has had a pre and a post outcome assessment for the same domain, for example under the family functioning domain under the circumstance SCORE. The example on the slide shows the number of clients with sessions, the percentage of that number assessed and finally the percentage of clients that have had a partial assessment.

If there are no paired scores, or they have had a partial assessment, this means that a pre and a post score for the same domain has not been recorded and this is where it shows up under that particular heading under the percentage of clients only partially assessed. Now the reason for this could be varied and could be that:

- The pre could have taken place towards the end of a reporting period and sufficient time hasn't passed for a post to take place.
- Or, an organisation has just opted into the partnership approach and a client has been accessing services for some time so it might be too late to record a pre score as they started accessing services in the last reporting period which is now closed.
- Or, a third reason could be that a pre and post score has been recorded for different domains. So for example a pre has been recorded in family functioning and then a post recorded for physical health, so no outcomes can be measured for the client.

Now these are valid reasons as to why a paired score hasn't occurred and these should be explained to your funding arrangement manager. Over time it is expected that the percentage of clients only partially assessed should decrease.

When we think about the key message for this particular section:

- Get your data in early so you can quality check it before the close of the reporting period. Use the time in the reporting period to check for errors. This is especially important for organisations that are using bulk uploads.
- If you upload your data early you can make sure that the upload is working correctly.
- To check and test the submitting of data please contact the Data Exchange Helpdesk for access to the staging or testing environment and we've also got information on our website under IT Access as well.

No we are going to look at a bit of information in the new report of the Organisation data quality report and how that can assist you. It can assist organisations in checking, improving and maintaining their data quality. Now this is a standard report and is available to all organisations using the data exchange. The report highlights:

- Key data quality issues for an organisation;
- How data quality is changing over time and whether it is improving or declining;
- How data varies across outlets, to help organisations identify different data quality practises
- And also throughout the report it links to key documents to assist in the improving of data quality.

Now on the screen here you will see the different sheets or pages that are available in the report and what I would like to highlight because in my demonstration I will be using the training report, the training report does not contain this first section here this first sheet of a User guide, that is a new implementation. Just to make you aware of that.

I am going to pause my screen for a moment and I am going to load that training report so that we can demonstrate that for you.

So what you have on the screen ok are the different sheets available in this report and as I mentioned the Organisation data quality report is available to all organisations regardless if they're in the partnership approach or not.

So you'll notice that we have some different sheets or pages within the reports and the functionality for the reports are the same across all the reports so, how you would bookmark, how you would filter, how you would set up a story, it is actually the same functionality across there.

So, what we will do is go into the first one here, the data quality health check and we will see what that can tell us and again as mentioned, this is a training report so the data here is not live data or real data.

So we can see some information in this data quality health check sheet, were we have the statistical linkage key, not stated demographic values and also unidentified clients. You will notice we've got percentage of clients with low quality SLKs for the last 6 months, now what I will do is just direct your attention to the bottom left hand side here of the screen and you will notice that this is for clients entered between and it's a six month rolling period (between the dates on the report).

Again, these dates are incorrect because it is a training report but the report that if you had a look at it today, it would actually display todays date and six months prior to that date, it just rolls over every day so you have the opportunity of updating the information they are using from this report.

Now what we can see is if we can think about the percentage of clients with low SLKs over the last 6 months we then have this graph or column graph of the different states. So if your organisation is delivering services in different states, you can see where there maybe issues there with the different outlets of that of delivering those services.

So we might select this particular state here of Queensland and it's telling me the percentage of clients with low quality SLKs for Queensland is around the 30.51%. We will tick that off (confirm our selection) and we will see the information has now changed, it is now bringing up for me or displaying those areas where the outlets are located where those issues are very high and so again we can have a look at that information and just find out why there.

We have the same information or premise for not stated and unidentified clients and now what we might do is move then to the next sheet we can get a further idea of what is happening with these particular areas. I just select this box here where it says data quality health check it will then display the next lot of sheets for me.

So I can select my icon of what I want to look for and which SLK sheet, this sheet is telling me then for this filter that we have selected which is the outlet state of Queensland there that we've got that percentage of 30 to around 31 percent of clients with low SLKs down the bottom here on the left hand side. We have a subset of contributing factors so this is the information of telling me what is making that SLK of low quality and remember we need a strong quality SLK so that we can actually have a look at the pathways clients are, they're actually accessing the services your delivering and that other organisations are delivering.

So, what is making this percentage so high?

- A high number of estimated date of births;
- A high number of unknown genders;
- Missing first or last name; and
- Percentage of clients with pseudonyms.

So that's really interesting and so we'll have a look at where we can find information that we can update that information for our clients in a moment but if we go towards the right hand side here you'll notice this sheet has here a "Why is this important" hyperlink.

If you select the hyperlink, it will take you to the Data Exchange website, It will take you to the information that will back up as to why this information is important or on other sheets it will show you how you can fix this information and I will demonstrate that for you as well.

We can see by the table on the top right hand outlets in this area that are delivering services. Remember that for your report this will actually show for your organisation, we can see there is a number of outlets in this organisation with a very high percentage of low SLKs.

If we right click and export the data we can export that into an excel spreadsheet to make it easier to have a look at. To click that off, use this button here for the full screen and that information will make it a full screen to make it easier to navigate and have a look at what I can do when we are thinking about the functionality of the reports. You now notice that I am on the top here of the column, I can now actually sort this out from lowest to highest numbers again and clients and that's again is a functionality across all the reports.

I am going to close this off now and I am back to my statistical linkage key sheet. What you can also see here is the distribution of outlets for each activity for clients entered in the last six months. I can now have a look at this activity and I can click on here.

It is telling me outlet 33 has a percentage of low SLKs, outlet 15 and outlet 6 is giving me a good indication, this particular activity I am delivering has issues entering this data. It is indicating that this is an actual activity issue of not being able to enter the data correctly or it could just be an outlet issue and so that is what that graph is telling you there.

Let us now have a look at the estimated date of birth information because we want to have look at how we can fix up this information. We are looking at the birthdate detail sheet, on the left hand side here at the top, we have clients with an errant date of birth.

We have a definition of what this means, so it means that the date of birth is equal greater or equals the session date. This would be highly unlikely that somebody would attend a session on with this particular

date of births and the client would be over 110 years old. As of the session date, we do have some older clients but again it is about you investigate that information and that data to see if it is correct.

What you will notice and what I absolutely love about this report is, we have this graph on the right hand side and you will notice here you have a column here for client ID's. These ID's match up what is in the data exchange and what is in your system. Here, client 1968 has a date of birth that is on or after the session date.

Remember client information such as date of birth, gender, disability and indigenous status can be updated at any time and is not locked down to a specific reporting period. It is only session information that is locked down to a specific reporting period, I can go in, update this client 1968, and update that date of birth there and that will be a stronger SLK.

If we go down to 9992 it is the client id, it is telling me that this client is greater than 110 years old at the session date. I can see here the date of birth that I have entered is actually the 01/01/1900. It provides an opportunity for me to go in an update that information but if I am doing a bulk upload as my way of uploading data to the Data Exchange or system-to-system, I need to check then are the codes are correct as to the upload methods. If I am using, or if I am entering my data manually via the free web base portal, am I entering information correctly?

So again having a look at what information needs to be adjusted and also adjusting the practises at your outlets there. If we go to the top right hand side here, the How do I fix this? This will take you to the find and edit a record task card. So it provides you not only what the issue is but also how you can fix this issue, which is absolutely fantastic.

So on your screen you will have the Data quality fact sheet, so when we think about why is data quality so important? We have a fact sheet to provide further information for you. So the way that you can get this is – it's on the Data Exchange website under the Policy guidance tab. You will need to scroll down to the heading of The importance of Data quality. This fact sheet contains a table of the top 10 items for data quality checks – why they're important and what the consequence is of poor data.

So this is what it looks like and let's look at an example of the incorrect date of births. Now the document will explain why a correct date of birth is required where possible, and what the dangers are of an unrealistic estimated data of birth – that it can skew the data for a client cohort or program, and this could affect then all the data collected for that program and affect the organisation in meeting their grant or funding agreements That is available to everyone so if you haven't already done so, to print that out for yourselves.

So let's look at some key takeaways from this presentation today:

- It is about uploading your data regularly. It provides an up to date and accurate picture of what is happening at your organisation in relation to your outlets and programs that are being delivered and what your clients are achieving.
- Check your data regularly for quality, use this new report, the Organisation data quality report to check and maintain your data quality that way you can fix it straight away before the end of a reporting period – so that your data is accurate. You don't need to ask your funding arrangement manager about this, as this is something that you can check yourselves, but obviously if there's any issues then please contact them and discuss that with them.
- Aim for high quality SLKs. High quality SLKs are required so that we can actually follow clients journey through the different services. This is to ensure that your date of births, first name and last name and gender information is correct, so we can see as a group what types of services they are accessing and captured correctly the first time, and the report that I demonstrated before is a great tool to help you in that situation.
- Limit the number of 'Not stated' responses – use only for genuine large group events where it is impractical to collect and record client information (Host note: the information should be that this is missed opportunity to collect demographics data for clients)
- Limit the number of unidentified 'group' clients as well, and again check the Appendix B for your program and under the heading 'Can unidentified group clients be recorded for this program

activity? " because you will find an expected percentage amount for the program that you are delivering, or you can speak to your FAM for clarification

- Remember that you can update your client records at any time, now that includes the first and last name, the gender, the indigenous or disability status, the address or country of birth and main language spoken at home. These are not restricted to reporting periods.

- And that you can only update session information such as when it was held, what activity was delivered at the session, and who attended during the relevant reporting period. So when that reporting period is closed it is too late.

- Now this new report, the Organisation data quality report will actually assist you with those key takeaways.

- So when we think about where to get more help if required?
- Your first point of contact for assistance will be:

- Your funding arrangement manager – they will be able to assist you in providing guidance on issues regarding data entry issues, staff absences or grant variations.

- They will also be able to discuss with you any milestone or KPI concerns that you may have

- You then have your IT vendor or specialist especially if you are using system to system or bulk upload and there are issues or errors in your data upload, speak to your IT vendor or specialist as they will need to review the coding used to abstract information from your client management system to the Data Exchange

- Also too, you can speak to our helpdesk because they will be able to provide further assistance with other technical issues and provide your IT vendor or specialist with a staging testing environment so they can actually test the programs that they have been designed to use

- And there is also the Data Exchange website. Just type in DSS DEX and the website will display, and you will find training and policy support material located in there, and we strongly encourage you to subscribe to our website so that you do not miss out on the latest Data Exchange news.

So when we think about the key message here that quality data is needed for you to meet your grant agreement.

Well everyone, we have come to the end of our webinar information today. If you have any further questions after this webinar please send them through to our helpdesk at the email address that is shown on your screen as they will be and we will be happy to answer you, we look forward to that. But we would like to thank you for your time and we'll see you next time. Thank you.