2023 Fall Session 1: CCSQ D&A Roadmap & Vision

1. **Q - Is the plan to move off SAS completely or just expand to include other options?**
   A - We intend to carry our current version of SAS Viya, which is 3.5, and maintain it through 2025. SAS Viya 3.5 reaches end of support in February of 2025, and we will continue to support and maintain it through 2025. As of today, we are looking at ways of migration and moving off the SAS standard platform.

2. **Q - What is the timeline for a move off SAS?**
   A - We are committed to keeping SAS 3.5 online and operational through 2025. There are no current plans to move beyond that.

3. **Q - Is the plan to move off SAS 9.4 and keep SAS Viya? Or to eliminate SAS altogether. What about the new SAS Workbench and Visual Studio (VS) capabilities?**
   A- For the CCSQ environment and the Centralized Data Repository (CDR), we do not have 9.4 except for a few exceptions. The exceptions are contract specific and whether that is maintained and funded is different from CCSQ. We plan to keep the environment online and available through 2025 with reduced onboarding.

4. **Q - How do we get reports to our participants?**
   A - This depends on the community that is being presented to.You can use the QualityNet portal used by the lines of business as a static, depending on what program you are in and whether pages need to be presented. If it is dynamic reporting, QuickSight has that capability and in the future some other reporting capability to provide that insight and data visibility, functionality, etc. QuickSight does offer the ability to embed dashboards in existing authority to operate (ATO) systems. If you have an existing system that you operate, and your users retrieve reports from today and there is now a capability within QuickSight to be able to embed dashboards directly.

5. **Q - Is the plan for CCSQ to have 3-5 people at each of the 15,000 nursing homes and even more people at every hospital to get HARP accounts so that they can access QuickSight reports?**
   A - QuickSight offers the ability to embed dashboards in existing systems. For any file transfer, report distribution, or other related question or need, please fill out a [request form](#) ticket with your use case and we are happy to work with you.

6. **Q - Will this Databricks (DBX) SQL Notebook solution resolve the slow performance of pushing SAS data tables to DBX?**
   A - We are working on a solution to this specific use case of SAS to the Databricks SQL where there is a slow performance in the connector or the driver for that connectivity.

There is a common library language, a Phyton script, which is being put together to help remediate that impact. Users that are purely in Databricks using notebooks and R would not have that performance issue.

7. **Q - Any plans to add Medicare Advantage claims data to CCSQ CDR? Any plans to add to the CCSQ CDR any International Classification of Diseases, Tenth Revision (ICD-10) diagnosis code, ICD-10 procedure code, or Food and Drug Administration (FDA) National Drug Code (NDC) classification (grouper) tools to assign a code to a general category and help standardize processing of claims data?**
   A - Yes. We are working on Part C data coming into the environment. It will be secured, and users will need the necessary business justification and data use agreement (DUA) entries, etc., to get access to it. We are currently in negotiations with the Part C team to figure out the best way to get access to it; whether that is a direct connection, in or to the Integrated Data Repository (IDR) or a separate file load to make it available in its raw format is to be determined. Our intent for future dataset and in negotiations to expand to that dataset. For ICD-10, if there is an author of those owners we will take this as an action and get your specific feedback. If it is available to our community, we can consider adding it to our repository. For the FDA, if it is a public data or a non-proprietary data, we can look to add it to our environment. If it is a proprietary we must work towards inter-agency agreements to that information. As for grouper tools, a lot of those groupers are extremely specific in what they are using and grouping. There is not a plan at this time to at this to use that functionality, but the use case is something we should explore. We currently do have the Medicare Severity Diagnosis Related Groups (MS-DRGs) grouper.

8. **Q - Will there be group Databricks training or online training available?**
   A - Yes. Part of the Portal creation is in improving and increasing our training tool sets. The Databricks center is working very closely with us and providing relevant training materials. The Quality Data Analytic Services (QDAS) team is making training more specific to how it works in the CCSQ environment. In addition to those trainings, we are developing any content and documentation is necessary to make users all successful including videos, white papers, brown bags, office hours and a dedicated support channel where users with a unique need can bring their issues up for assistance.

9. **Q - Will Minimum Data Set (MDS) data be made compatible with Databricks prior to Ambari decommission?**
   A - We are aware that there is currently an issue with two fields in the mds_asmt_fed table in the assessment_iqies schema that are causing some issues in Databricks. We are currently working with the Databricks vendor and data source owner to identify a resolution.

10. **Q - Where is Python being used here for analytics and which libraries /what version? What is the Volume of Data that we are dealing with here?**
   A - The CDR has about a petabyte of data in it between our standard datasets and the Workbench data. The volume of data is based on your DUA and what information you have. Python is being used throughout the Databricks notebook capabilities and there is a mixture a of R and Python. Which language is used depends on your specific analytic use case. If you have a specific question about what the right tool is, please open a [request form](#) and we can help you figure out. Libraries are open and there are no restrictions yet about which libraries can be leveraged because it is all running within our Federal Information Security Management Act (FISMA) boundary.

11. **Q - Besides MDS data, is OASIS data going to be compatible with Databricks?**
A - Yes. We are working with the Databricks vendor to determine what the issue is with Parquet files. The current issue is related to the Parquet format. In the meantime, converting those files from Databricks to Delta tables would resolve the issue. In addition to that, continual improvement with the product, which OASIS and CASPER and ASPEN all fall within that program component, an area is that they are constantly in migration of that dataset to Internet Quality Improvement and Evaluation Systems (IQIES). Eventually, that all ends up in the MDS. 2028 is the final date for the completion, but it is progression between now and then.

12. **Q - Are any of the large text-based datasets such as the Statements of Deficiencies (2567s) available in the CDR holdings?**
A - Yes. We currently have a contractor working on the 2567s for natural language processing and sentiment analysis and other output. Users need a specific data access request to see if they qualify for that data. Users can fill out the [CCSQ Data and Analytics Request Form](#) for assistance.

13. **Q - Any chance you could provide the big picture? CMS has Chronic Condition Data Warehouse (CCW), Integrated Data Repository (IDR), and Centralized Data Repository (CDR). What is the future of CCW and IDR?**
A - We cannot give details on those processes. Each one of those serves a specific community of purpose. CCW is typically for the researcher community. All the data is transformed modeled and made available to that community. They pay for a license to get into the CCW for access. The DUA process grants approvals for specific data access. We primarily service our Quality Improvement Organizations (QIOs), the Quality Improvement Networks (QINs), the dedicated analytic contractors directly for our programs. We have that closed community and it does not fit the CCW model. IDR has a limited dataset compared to our quality datasets. We are looking to establish a connection between our CDR and the IDR and expect that to come online sometime next year. Currently, they are prioritizing moving the IDR to their cloud solution, the Snowflake. We are engaging early next year for us to connect our Databricks or our CDR environment to the Snowflake/IDR environment. Once we get that connection between the CDR and the IDR established, it will change that dynamic next year. We do share a few datasets with the CCW, they author those files, and they are currently available within the CDR.

14. **Q - Do you offer other analytical tools in SAS Viya, such as SAS Visual Analytics, SAS Text Miner, etc.?**
A - We have Visual Analytics, however only one group is using it on any relative scale. There is a limited license for that tool for the environment and there are no plans on expansion and usage of that tool. If you feel like the tool is needed to complete your work, fill out a request form and state your use case and from there, a team member will help you figure out the correct solution space. Our enterprise solution for dashboarding and visual analytics is QuickSight. If you or your team has any reporting or dashboarding needs, please reach out and we will work with you and then the Star team in in getting you the access that you need. Databricks libraries allow visualization of the content as well.

15. **Q - Will Quality Reporting Document Architecture (QRDA) collected from the Quality Payment Program (QPP)/Merit-based Incentive Payments System (MIPS) program be available in CDR?**
A - QPP Uniform Data Systems (UDS) data is currently available in the CDR. Unfortunately, we do not have any current communication with the QPP team to receive additional datasets now, but if there is a dataset that you are interested in or that you need similar, submit a [request](#) and we will work with that team to see if that can be added to future datasets.

16. **Q - You mentioned that you would decommission Ambari after you got it to work better with SAS. Is it working better with SAS now?**
A - Ambari will be decommissioned on Sunday, December 31, 2023, except for one use case. If you are currently using SAS to access Ambari, we need to work with you to get you off it. The [macros](#) for the exception use case will be available on Monday, November 27th.

17. **Q - When will R have access to the CDR?**
A - That direct access is not available today. There is a workaround if you need some assistance, please [reach](#) out.

18. **Q - Could you guide us to get coding help to develop MDS quality matrix?**
A - You can send an [email](#) to the QualityNet service desk and open a request. If you need any advice input, clarity, or recommendations, the team can help you with that.

19. **Q - Are PRS 2.0 tables available through Databricks or SAS?**
A - This will be coming soon to the CDR. It is scheduled for Q1 of 2024.