

Project Update Webinar: Customer Experience

Questions & Answers

As of September 3, 2024

	Question	Answer
Models for a Data Concierge Service		
1	Is the data concierge real people?	Yes, real people would be a component. There could be a chat bot or some kind of AI component, but there will be real people available.
2	How would you get started? By topic, e.g., workforce? In terms of building it.	<p>NCSSES: It's thought as more of a workforce issue. We're looking at different components and skill sets that individuals would need in order to provide that technical assistance within a National Secure Data Service (NSDS).</p> <p>NORC: We'll get there eventually but in terms of things we're looking at are job descriptions for similar organizations that have data concierge services and what kind of background and education they're looking for in the resources when they go out and hire. The report is believed to get there but hasn't crossed that bridge yet.</p>
3	What was the most surprising finding, if any, that came from interviews?	The most surprising finding as we were going through the interview process was a consensus on the novice user side that the resources weren't there to try to navigate the complex nature of the federal statistical agency data. Some had success using the data sets they knew, but they know there's more data out there. They're just not aware of the breadth of data available, so if there were a service to help people expand their knowledge in terms of the data that are available it could improve evidence-based research and the data people are using.
4	Has the team looked at metadata standards like DCAT 3?	Not necessarily for this project but the transition to DCAT 3 is known and with the standard application process which will likely be a shared service within the NSDS. DCAT 3 is being looked at in terms of converting to DCAT 3 and being consistent with DCAT 3 for our metadata standards for restricted data and confidential data.
5	What thinking has been done towards the data governance required, especially in these cases of linked files (cross-agency)? Does documentation just append across the files being linked?	This is something that's larger than just a data concierge service. This is something that we try to deal with within the larger Federal statistical system. Better documentation of data sets and variables are needed. For example, ones that can be used for linkage. Different techniques are also needed for linking to break some of those agency boundaries that we have regarding linking so those topics are being explored in other projects as well for the National Secure Data Service. Privacy preserving technologies that look at record linkage is being looked at and these project updates will be offered in the future Also, how to link data using different methods. This is being explored, but it's bigger than this project.
6	How will the data concierge have enough knowledge of all of the data to give good answers?	The data concierge service will have multiple components, including individuals with knowledge of data, tools to assist concierges in finding relevant data and information, and a list of referrals to send users to experts that can answer their questions when needed.
7	Have you estimated what the cost of PII data anonymization activities would entail?	PII anonymization activities can vary widely in terms of cost depending on method. Different methodologies are being explored for the NSDS demonstration project.

	Question	Answer
8	Is it part of your remit to build a running prototype of the non-person part of such a service? For example -- a common metadata repository?	Metadata repositories currently exist for both public and non-public federal data. How to best utilize and coordinate between these repositories is under discussion.
9	Is there an interface so commercial Search Engines can interface with the Data Concierge Service? How do you intend to make the Concierge and datasets discoverable from Google?	The concierge service is in the conception and modeling stage. The build of the actual service will be in a future project.
10	How did you select interviewees?	Organizations were contacted to identify an individual to interview. Organizations were targeted to represent a cross-section of academic institutions with robust research programs, state data centers, and organizations with a strong data focus who have provided leadership in data sharing initiatives and development of data sharing legislation. Additional respondents were identified by individuals interviewed.
11	Some of the pain points identified by the personas seem fuzzy regarding the scope of our efforts vs. the responsibilities of the user; where would the concierge draw that line?	Users interfacing with an NSDS would likely represent a wide range of levels of expertise. As a result, some users may require more support than others. Guidelines would be provided to both concierges and data users, however, on what services are in scope for a data concierge and what are not.
12	Does the team anticipate making an application programming interface (API) available so entrepreneurs and companies can build solutions that tap the NSDS?	We have not yet explored this option.
13	Are you leaning toward the Connector Model only vs. the Full-Model? I'm in a very small statistical unit, and we would potentially struggle with providing staff time for technical assistance.	We have just started evaluating the pros and cons of each of the models. The decision as to which model to utilize, or which elements of the models, will be decided within a larger forum that will include input from statistical agency leadership and Interagency Council on Statistical Policy (ICSP) input.
Federated Data Usage Platform		
14	Will you be able to distinguish uses of restricted-use versions of data from the public use versions?	Yes, we plan to include filtering options for restricted vs. public use data sets and some other levels.
15	When will you expect the prototypes to be available?	We expect both projects to conclude in spring 2025. We hope to potentially provide a sandbox where users can come in, play with the prototypes, and give feedback. That's yet to be determined, but we are thinking about how to involve the stakeholder community and potential data users so they can test these prototypes for continuous feedback.
16	In the User Research, were there any commercial data providers/brokers interviewed who may want to enhance/republish/enrich the data?	No. We focused on state and local data creators, policy makers, and academic researchers
17	Why are you building two different prototypes?	We were very interested in the different technical approaches to the solutions for a data usage platform. We wanted to explore that since there's a lot going on with machine learning and AI. We also wanted to understand the variety of stakeholders. We decided to support two efforts and compare the prototype results at the end.
18	Will I be able to test the DUP prototypes and provide feedback as a user?	We hope to, but the timeline is TBD.
19	Do the experiments include creating incentives for downstream users to better attribute data usage?	We have not yet explored how to include incentives to help facilitate that; however, that is a great idea and something we can discuss with the groups.