Project Update Webinar: Privacy Preserving Technologies Questions & Answers

	Question	Answer
Privacy Preserving Technologies		
1	What are the plans, if any, for a phase two for the PPT projects?	Right now, we are trying to figure out some next steps with high-impact use cases, and it's believed that this will be part of our conversation. We currently do not have a direct plan, but it is on our radar. After implementing these two PPRL projects, we will take the time to discuss our next steps.
2	Please describe the time the steps did take, such as getting to a signed agreement.	It took a long time to think through the timing and the steps needed for the data-sharing agreement. This is partly why, when looking at the periods of performance for both PPRL projects, we did these for two years, knowing that the linkage of value would not take nearly that long. We did this to build in time, and it probably took between 8 and 9 months in terms of timing with the data-sharing agreements. The steps needed were separate for both, but we had sort of initial templates within the parent agency and then between the two Federal statistical agencies of where to start. However, the project leads took a first stab at it in consultation with others. I worked very closely with our office of general counsel regarding reviewing and the written language. Involving our IT specialists and understanding the entire FedRAMP ATO process was time-consuming. We had to learn how to navigate the necessary approvals and identify which groups needed to be consulted. Regarding steps, it was this idea of the multidisciplinary team involving parties. There is a definite review process. Questions were raised by many different reviewers in the review process, which made the agreements that much stronger and able to note with that data flow diagram where there were concerns. For example, what if somebody's covariate pattern is somewhat identifying? And that was really where we thought, let's separate that from any external sharing in terms of the CLKs or the hash tokens going into the cloud. That came out through the review process, and there were concerns about whether we were to attach the covariate data with all the encrypted tokens.
3	Were there obstacles regarding whether NHIS, as respondents, consented to the linkage?	We only included respondents who consented to linkage. There were linkage eligibility criteria on the site for NHIS. I don't think that was an obstacle. I believe that NCHS has a long-standing data linkage program. Determining linkage eligibility and creating adjusted weights based on linkage eligibility helps with some of that non-response bias. They're very well versed in that, and I think they carried over to this project as well.

	Question	Answer
4	Has there been an evaluation of the quality of this linkage relative to the other methods?	We discussed this during the data-sharing agreement process, emphasizing transparency in our approach to assessing data quality. The linkages have not yet been completed. We are hopeful that is coming in the next few months. However, we do have some plans on how we will spend time assessing data quality. Part of it is what I was showing with these anticipated overlaps and going in with this outside knowledge in terms of benchmarking to see how we do in terms of how the populations overlap. We also have other ideas for looking into data quality and ensuring that anytime we're talking about those data, we address any sorts of data quality concerns, strengths, limitations, or analytic considerations when working with these types of data.
5	Who will get to see the link data?	That was also part of the data sharing, which is really thinking about governance. Again, I think it will depend somewhat on the quality of some of the measures we use to assess that. But we would like to make it available through a restricted-use access program where there would have to be an application in both NCSES and NCHS linkage, and both would have to be part of the review process. Similarly, with the NSF/NCSES data, a review process would be in place to access those data.
6	Examples are all within the government. Is this assumed to be generalizable to public user requests? Or will there be test runs for specific outside requests?	Starting within the government was just how we anticipated doing it. In thinking of the vision of the NSDS, I think that this idea of working with outside users involving states and territories is really the direction that we want to go. So, going back to that first question, what will phase two look like? It may take that shape or form. We are just not sure yet, but it's a great question. I think our team is also very interested in pursuing it.
7	What are the thoughts on data sharing agreements, allowing access to two agency data sets when the data was collected under different laws?	That's where those data-sharing agreements come in, as well as having the right people at the table to have those conversations. When we thought about the data sharing agreements, we did try to think of them as modular in some sense that there would be different sections that could be applicable depending on the data type, probably the governance over the data, and that's where we're really hoping that at least some of the work that we've done can serve as a starting place. It doesn't erase the need for those conversations and coming to an agreement to create these data-sharing agreements.