NPWR Privacy Elements

The key organizational elements that ensure privacy include:

- **NPWR does NOT collect data from citizens**
  - It is a system that was developed among its participating agencies to allow the merging of data in a highly controlled environment using technology that strips "exposure data" (information that identifies an individual) before merged information is released to researchers.
- **NPWR is a "hybrid federated" model for longitudinal data**
  - Rather than build an entirely new "data warehouse" to collect and store data, which would have required the duplication of private data (and redundant infrastructure to support and protect that data), Nevada chose to implement a system that leaves its data where it has always been, secure within the participating agencies' databases.
- **One cannot simply walk up (login) to NPWR to initiate a data merge**
  - Data merges are initiated by vetted researchers who first have completed an application process and whose research questions have been reviewed and validated. Then they are assigned a committee of agency "sponsors" who guide and oversee the process – all in the name of accuracy AND privacy. Each step along the way, from access request to publication of results, must be approved by the sponsoring agency.
- **The NPWR Book of Data Governance requires regular meetings, testing, and assessments of NPWR to include ensuring privacy**

How NPWR Merges Data

- If exposure data is protected behind privacy firewalls of state agencies, how is useful data merged? The State of Nevada's solution to this question was developed by the Nevada Department of Education, the Nevada System of Higher Education, and the Nevada Department of Employment, Training and Rehabilitation, and now managed by the Governor’s Office of Workforce Innovation.
- Using a state-of-the-art probabilistic matching system, NPWR matches data across agencies through a process that de-identifies and cleanses the data to provide the highest possible match rate while maintaining full privacy. All data within NPWR are fully de-identified to ensure that no personally identifiable information ever remains within the system. Each time a researcher requests data, the system will generate a completely new set of unique identifiers for each individual in the data. This feature, developed to comply with state law, ensures that the new data set cannot be linked to the previously requested data set based on unique identifiers.

Key Privacy Protections are as follows:

- Before agency data sets are delivered to the data hub for merging, an algorithm is applied to the data to render key, private information (e.g., name and date of birth) into a string of meaningless numbers and letters.
- The created data sets expire and are destroyed after two weeks of the initial database inquiry, making the data unique for every merge.
- It is important to note that "de-identification" is a "one-way" process that is different than encryption, which is "two-way" scrambling, meaning the data can move back and forth from encrypted to readable data. In simple terms, once de-identified the data cannot be reversed back to identified data.
- In the application review process and in the algorithm and Data Hub processes, there are minimum data threshold requirements to ensure that NPWR does not create data sets that could be used, through a process of elimination, to match de-identified data back to individuals or groups of people.
- The merged data is then reviewed by agency staff for both accuracy and privacy before released to researchers.
- Any data sets that are released to the public are scrutinized again for privacy concerns before released.