ERC Supplemental Data Process

Supplemental data refers to additional data not included in the Repository. Procedures are available to enable researchers to add supplemental data to the Repository for use by a specific research project. If requested and approved in the proposal, each state agency whose data is involved will need to approve each request. The supplemental data will be de-identified prior to uploading to the ERC data Repository. In cases where student records within these supplemental data must be matched to corresponding Texas ERC student data records, the TEA and the THECB will provide alternative ID mapping services. In some cases, the researcher will be charged a cost for processing and manipulating such data (19 Tex. Admin. Code § 1.18 (d)). Researchers must cover the additional costs. Ideally, supplemental data requests should be made at the time of the proposal as to allow for a seamless review by the ERC Advisory Board. In the event the researcher(s) determines that supplemental data is needed, they should review the Project Data Amendments section. After the review, the researcher(s) should discuss their needs with the Director and complete the appropriate Supplemental Data request form (see forms tab). Requests must be submitted to both the Director and the IT Coordinator. It should be noted that to comply with the requirements of FERPA, supplemental data files must have a destruction date that corresponds with the end of the dependent research project, plus five years for publication purposes.

There is a process on the de-identification of the ERC

The ERC data contains two <u>alternative</u> identification numbers ID1 and ID2. The background to the ID1 and ID2 as alternatives to state assigned id or SSN, is that PK-12 students are NOT required to have an SSN. Texas is supposed to educate every student. So, this is why we even have and ID1 (and ID2). The ID2 is an alternative for SSN. The ID1 is the alternative to the state assigned number if no SSN is reported, the state assigns an ID to all PK-12 students. Technically, at higher education, students are supposed to have an SSN or the equivalent for international students. The IDs have been put through a strong algorithm, so that none of these IDs are real IDs.

To use the ERC data system to determine the impact of an intervention, researchers need to be able to identify the students who received the intervention under question. An external source organization holding the intervention data can share the specific data with TEA. The intervention data must be sent directly to TEA by the external owner, not the researcher, via the secure file sharing program. This data is referred to as *supplemental data* because this data supplements the ERC longitudinal data system. Prior to transmitting data to TEA, the project researcher must obtain permission from the external owner to use the data and allow submittal of the data to TEA. This is part of the proposal process on getting approval from the ERC Advisory Board to add supplemental data. TEA will need the external source to send the treatment group data containing personally identifiable information (PII) of the treatment group along with any program data that is to be used in the analysis such as timing, amount, scores, and other qualities of the intervention received.

When the PII data of the treatment group is sent to TEA, TEA will match the external data to the PII data they already have on all PK-12 public school students and employees of Texas Public schools. To have the most number of accurate matches, TEA would like to receive the following PII data for each treatment group member: social security number (SSN) or state assigned id, first name, last name, and date of birth. If SSN is not available, including the high school the student attends or which school the student graduated, and student's high school graduation year will assist in increased record matching. The identified and matched supplemental data is then sent to the Texas Higher Education Coordinating Board (THECB). The THECB will match the students who are enrolled in higher education in Texas

(public, private, and for-profit institutions). The next step is the de-identification process. The THECB matches ids to a master file. They create an "age-as-of-September-1" and create an alternative ID1 and/or ID2, delete the name, DOB, and SSN or state assigned id. The de-identified data is securely transferred to the ERC. The de-identified supplemental data is then placed into researcher's individual research folder accessed only by researchers connected to the project. TEA and THECB destroy the original personally identifiable data when the project ends.