CALPADS Update Flash #221

Date: April 8, 2022

To: Local Educational Agency (LEA) Representatives

From: California Department of Education (CDE) —
California Longitudinal Pupil Achievement Data System (CALPADS) Team

Post Process Redevelopment Project

Since its implementation in 2009, the amount of data that LEAs must submit to CALPADS has increased significantly to meet new and ever-changing state and federal reporting and accountability requirements. In addition, the more recent use of CALPADS data for operational purposes, such as for statewide test registration, has increased the frequency of LEA data submission to CALPADS. While significant technical upgrades have been made to CALPADS, such as moving into a cloud environment, the process by which data are ingested into the system has not been updated since 2009. This has resulted in increasingly degraded performance, with CALPADS users experiencing unacceptable wait times for data to be posted to the system. Accordingly, the CDE initiated the Post Process Redevelopment Project to significantly decrease the time it takes for files to be posted to CALPADS. To accomplish this objective, there will be some necessary changes to current processes, which will be highlighted in this and upcoming CALPADS Flashes and included in updated documentation.

This Flash:

- Summarizes the overall objectives of the project and the key changes that needed to be made to meet project objectives.

- Describes the changes to the Statewide Student Identifier (SSID) assignment process, and the SSID replacement process.

Project Objectives and Required Changes

The objectives of the Post Process Redevelopment Project are to:

- Significantly reduce the time it takes to validate data within a file being submitted; and

- Significantly reduce the time it takes to post files to CALPADS.

In order to accomplish these objectives, the following system changes were necessary:

- Replace the single-lane file posting queue with a multi-lane file posting queue.

- Reduce the number of validations that occur upon file input by focusing on validating only data within the incoming file rather than comparing the data in the incoming file against data in the CALPADS Operational Data Store (ODS).

- Post a file only if it passes all input validations.

- Eliminate partial-post, and auto-post files only if it passes all input validations.

- Move validations that were previously conducted on input, but which require comparing data in the incoming file against data in the CALPADS ODS, to display after the file has posted in new data discrepancy reports, which will also include validations that will eventually become Certification Validation errors.

- Redesign the Statewide Student Identifier (SSID) assignment process to exclude the comparison of demographic data in the incoming file to demographic data in the ODS.

- Run anomaly detection throughout the day to provide immediate feedback regarding Multiple Identifiers (MID) anomalies.

Highlighted Changes: SSID Assignment and SSID Replacement Processes

Currently, the acquisition of SSIDs is integrated in the submission of the Student Enrollment (SENR) file. Under the current process, LEAs upload an SENR that must include a local student identifier (ID) assigned by the LEA, and the SSID, if it is known. When SSIDs are unknown (for example, for new or transferring students), the SSID field is left blank. Based on the demographic data provided in the SENR, CALPADS conducts a search in the CALPADS ODS for students with blank SSID fields, and provides a list of potential candidates that match the demographics provided in the file. LEAs make SSID selections (existing or new) for students from the list, and then post the SENR records with the selected SSIDs for the selected students.

As described in the previous section, in order to increase performance of the file upload process, the search for potential matches against the CALPADS ODS has been eliminated as part of the new post process. Instead, CALPADS will now assign a new SSID to students with a blank SSID field (except when the local student ID matches a local student ID for the same LEA in the CALPADS ODS). Understanding that elimination of this functionality may result in an increase of MID anomalies, changes to local business processes that can mitigate the creation of MIDs are described below. In addition, the CDE will be implementing additional functionality following the initial release of the redeveloped post process, to enable LEAs to more easily search for SSIDs prior to submitting files to CALPADS.

New SSID Assignment Process

CALPADS will perform the same process regardless of file size for the new auto-assignment of SSIDs. That said, here are two workflow options LEAs may take depending on the volume of blank SSIDs. While Option 1 is preferable, it may not be feasible for LEAs that experience large numbers of new and transfer students on an ongoing basis. LEAs may also pursue both options, depending on the volume of transfers they experience in any given time period.

Option 1: Low Volume Transfers/New SSIDs – Student Demographic Search

1. Upon enrolling in an LEA, perform a student search in CALPADS Online Maintenance based on demographic criteria.
2. If an existing SSID is found, manually update the SSID in the Student Information System (SIS).
3. Generate the SENR extract from the SIS and upload to CALPADS.
4. If an existing SSID is NOT found, proceed with creating a new SSID by selecting "Create New SSID" from the Online Maintenance screen and CALPADS will assign a new SSID.
5. Download the SSID Extract to populate the SIS with the new SSIDs, as described below.

Option 2: High Volume Transfers/New SSIDs – SENR Batch File and MID Resolution

1. Generate the SENR extract from the SIS and upload to CALPADS, which includes blank SSID fields for new and transferring students.
2. For records where the SSID is blank, CALPADS then:
    a. Assigns **an existing SSID** where there ARE ODS records with a matching Local Student ID at the same LEA;
    b. Assigns **a new SSID** where there are no ODS records with a matching Local Student ID at the same LEA; then
    c. Automatically posts the file.
3. Once the SENR is posted, in order to avoid MIDs from being submitted to the Test Operations Management System (TOMS) in the nightly feed of CALPADS enrollments to the test vendor:
    a. **Resolve MIDs on the same day that the SENR records are posted** through MID Resolution, which is now updated every time new records are posted to the database. (NOTE: MIDs are now only matched on **exact** first and last name and date of birth, but will be enhanced in the near future to include probabilistic matching. The MID Resolution screen has been enhanced to display Parent/Guardian information to help identify the correct student.)
    b. If MIDs are not resolved on the same day, CALPADS Administrators should share MID reports with Testing Coordinators to ensure that students with possible MIDs are not tested until the MIDs are resolved.
4. Only once MIDs are resolved, download the SSID Extract to populate the SIS with new or replacement SSIDs that result from resolving MIDs, as described below.

New SSID Replacement Process

In the current system, when an SENR file is submitted with an existing SSID that has been retired, LEAs receive a warning that the SSID has been retired and that CALPADS will replace the retired SSID with the retained SSID. In the new system, CALPADS will continue to replace retired SSIDs with retained SSIDs, however, LEAs will not receive a notification. Therefore, LEAs will need to download the Replacement ID Extract to view the retired SSIDs that were replaced.

Loading New and Replacement SSIDs into the SIS

The SSID Extract is now downloaded by a date range that is based on the Enrollment Start Date and **not** based on job ID. Therefore, the SSID Extract may include SSIDs that already exist in the SIS, which may cause errors when trying to import the SSID Extract file into the SIS. To avoid this problem, LEAs may want to import the SSID Extract into the SIS using the Local Student ID as the key, which will then update blank SSIDs in the SIS. However, since each SIS works a little differently, it will be important to work with your SIS vendor for specific instructions on how to resolve any errors.

Importing the SSID Extract file into the SIS using the Local Student ID as the key, *may* also update SSIDs in the SIS that have been retired or replaced due to a MID merge, or cause an error. Therefore, it is also important to work with your SIS vendor to understand how to resolve any errors, or identify which SSIDs have been replaced.

Summary of Suggestions for Changes to Local Business Processes

- Since the SSID Candidate List will no longer be generated upon uploading the SENR to CALPADS, LEAs should perform a manual search of students in CALPADS to determine whether or not a student already has an SSID, then manually update the SIS with any existing SSIDs prior to generating the SENR extract from the SIS.

- Since the new SSID assignment process may result in more MIDs, MIDs should be resolved on the same day the SENR is posted in order to avoid MIDs from being sent to TOMs in the nightly feed, and prior to downloading the SSID Extract to import into the SIS. To accommodate the need to resolve MIDs on the same day the SENR is posted, MID processing will occur on an ongoing basis throughout the day, and not just overnight.

- LEAs should be aware that the MID matching process will now only produce matches when there is an exact match on First Name, Last Name, and Birth Date; probabilistic matching will be implemented in a future release.

- Since users will no longer receive a warning when a retired SSID submitted in the SENR has been replaced with the replacement SSID, users should regularly download the SSID Replacement extract to determine which SSIDs have been retired within the SIS.

- Since the SSID Extract can only be generated using a date range based on Enrollment Start Date, and is no longer available by Job ID, there is a potential for the extract to include SSIDs that are already in the SIS, making it important to work with your SIS vendor who may have to modify import criteria to accommodate existing SSIDs in the import file.

Mitigating the Creation of MIDs

The CDE recognizes that the revised SSID assignment process will likely increase the number of multiple identifiers assigned to a student due to the fact that the SSID demographic matching process has been eliminated. While we recognize that this negatively impacts downstream processes, such as test registration, the CDE feels that improving the efficiency of the post process in this initial release outweighs the risks of an increase in MID anomalies, particularly because the majority of MIDs are generated in the Fall as compared to the Spring.

In the short-term, LEAs can take the specific actions outlined in SSID Assignment Process options 1 and 2 to mitigate the creation of MIDs. The CDE also plans to release additional functionality this May to help LEAs reduce the creation of MIDs, followed by more permanent solutions in the short and mid-terms, as described below.

Immediately Mitigating the Risk of MID Creation

The increase of MIDs that will initially occur could impact Initial and Summative ELPAC testing as well as Smarter Balanced Assessments; therefore, it is important to mitigate MID creation to the extent possible. Following the implementation of the redeveloped post process, the CDE recommends following the steps outlined in Option 1 and Option 2 described above. The key actions that LEAs can take to mitigate the creation of MIDs that are included in these options include:

1. Modifying local business processes to include a manual student search in CALPADS Online Maintenance to search for possible SSID matches for the student and manually enter those SSIDs into the SIS prior to generating and submitting the SENR to CALPADS. **Do not send a SENR file with a blank SSID field for students transferring from public schools within California, if it can be avoided.** While we recognize this process is manual and not ideal, especially for larger LEAs, it is an option immediately available to avoid MID creation.

2. Reviewing and resolving the MID anomalies on the same day that the SENR was posted if blank SSID fields are included in the SENR file. Doing so will minimize the number of MIDs being sent to TOMS in the nightly feed, and ensure the correct/permanent SSIDs are imported into the SIS.

3. Sharing MID Anomaly reports with assessment coordinators, if MIDs are not resolved on the same day that the enrollment file is posted, so they do not test the students until the MIDs have been resolved.

Planned Short-term Solution (1 month)

Shortly after post-redesign implementation, in the May-June timeframe, the CDE plans to build an SSID Match Process in CALPADS that is separate from the File Submission process which would allow LEAs to upload the SENR file specifically for SSID matching against demographic data in the CALPADS ODS. This upload would allow LEAs to generate Report 0.0 – *SSID Candidate* List which they can review to identify potential matches. Once the matches have been determined, the LEA would populate the SIS with the matched SSIDs and then upload the SENR to the file submission portal.

Planned Mid-term Solution (2–6 months)

In the mid-term, the CDE plans to implement a student search Application Program Interface (API) that will enable vendors to implement a search feature within their SIS that will allow for searches for transfer students to acquire and populate the SIS with their SSIDs, without requiring logging into CALPADS.