

*National Forum on Education Statistics EDFacts
Modernization Working Group:
Feedback on EDFacts Modernization Plans
July 2021*



Members of the Forum establish working groups to develop guides in data-related areas of interest to federal, state, and local education agencies. The Forum has historically provided the U.S. Department of Education (ED) feedback and support on the data collections that impact state and local education agencies (SEA and LEAs), such as the Civil Rights Data Collection. In this case, the group convened to provide feedback to potential plans of ED to modernize *EDFacts*, including guidance on how SEAs map their student information systems to federal reporting requirements. These mappings are the backbone of any modernization of ED but also provide one standard data pathway that other SEAs can benefit. One part of the output of the working group is this written feedback on its plans to modernize *EDFacts* from SEAs, the *EDFacts* data submitters.

The recommendations and opinions included in this document do not necessarily represent the policies or views of ED, the Institute of Education Sciences (IES), or the National Center for Education Statistics (NCES).

Working Group (State Education Agency) Participants

DeDe Conner, Chair, Kentucky Department of Education
Georgia Hughes-Webb, West Virginia Department of Education
Laura Boudreaux, Louisiana Department of Education
John Keller, Indiana Department of Education
Kristen DeSalvatore, New York State Education Department
Charlotte Ellis, Maine Department of Education
Linda Roska, Texas Education Agency
Jonathan Wiens, Oregon Department of Education
Debbie Yedlin, Arizona Department of Education

Federal and Contractor Support

Ghedam Bairu, Ross Santy, Barbara Timm and Kelly Worthington, NCES
Beth Young, Quality Information Partners
Jim Campbell, AEM Corporation

EDFacts Modernization Context

ED is considering the modernization of the EDFacts collection system. The current system is 17 years old and the basic way the data are being collected (aggregate data files) has not changed over this time. No decisions have been made yet about the details of EDFacts modernization.

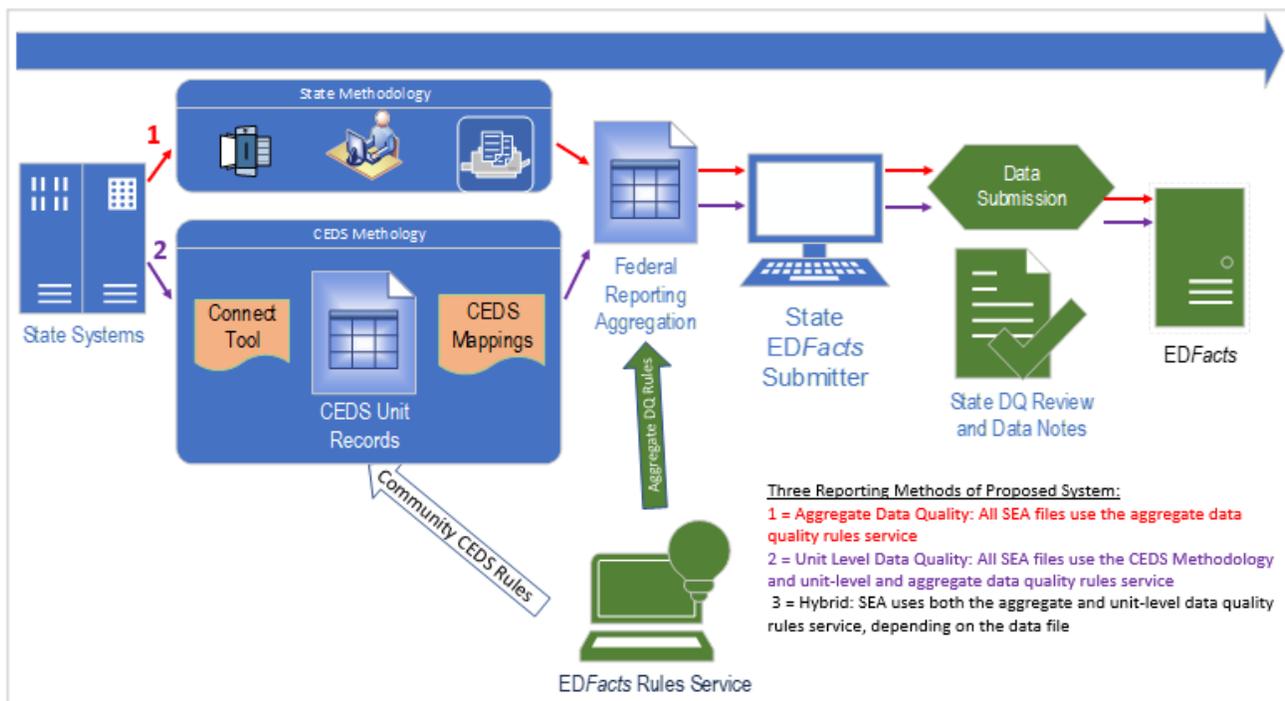
Current Problems Identified with EDFacts Collection:

1. The lag in receiving post-submission data quality review from submission results in significant burden on the states.
2. There are delays in publishing the data because of the data quality issues found post-submission and the time it takes to resolve those issues.
3. There are multiple places and methods for EDFacts data quality review, and this adds confusion and increases the burden on the states.
4. States need to repeat explanations year after year, even within the same school year reporting, which adds to the state reporting burden.

ED's Priorities for EDFacts Data Modernization:

1. Pre-submission data quality review: Move the data quality review to before the data are submitted.
2. Common Education Data Standards (CEDS) Mappings: Provide clear guidance for states on how elements in CEDS can be used and aggregated into EDFacts data groups, categories, and permitted values.
3. Data Comments: Organize state data comments and send to ED as the data are submitted.

The figure below outlines the basic proposal from ED, offering three different ways for states to submit their EDFacts data.



ED developed a free *EDFacts* reporting tool. Generate uses the CEDS data model as standardized data definitions for the purpose of transforming unit record data into the aggregate data sets required for *EDFacts* reporting. As of this writing, ten states have submitted files to *EDFacts* through the Generate tool and six additional states are in the process of installing and configuring this tool for their own reporting purposes, or have implementation on their current roadmap in the next two years. Details around the progress of Generate in developing and piloting *EDFacts* files can be found at <https://edfacts.communities.ed.gov/#program/generate>.

Using Generate is not required for federal reporting, but has improved the quality of data submitted by those states and has resulted in several conversations with federal program offices resulting in greater clarity and specificity in the file specification. It is only through this collaborative work between states and ED that a thorough understanding of the requirements can be achieved.

Working Group Feedback

The working group met multiple times with staff from ED in preparing the feedback below. This feedback is meant to provide ED with the information they need to move their modernization planning forward with input from the states.

1. File Due Dates

One of the fundamental aspects of *EDFacts* modernization is moving the data quality reviews from post-submission to pre-submission. This allows states to receive their data quality review results as they prepare the files, and solves the problem of post-submission data quality review timing and burden on both SEAs and ED. With this change the total amount of time SEAs spend on their files will be lessened, since the data quality review will be completed as the files are being built or immediately upon their pre-submissions.

This change moves the SEA resource pull to pre-submission, which means more time is now needed by SEAs up front to run through the file processing, all the rules, etc. Some of the current due dates are hard for some SEAs to meet now, as the time from data availability at the SEA level and the file due dates are too close. Therefore, the change to add more work for SEAs up front during the pre-submission process cannot be considered without pushing file due dates back.

Working Group Recommendations: Due dates should be pushed back on some files to allow SEAs the full amount of time needed to run through the new pre-submission process after the data become available to the SEA. The need for more time in pre-submission to utilize the new aggregate data review is true for all states, regardless of the methodology they are using (CEDS or state methodology). SEAs understand that sequencing must be taken into consideration with due dates. ED should leave the system open as long as possible so states can have ample time to complete files (see 2b below)

2. EDFacts Modernization System

There are several issues with the current EDFacts system and processes that have been identified that can be improved while modernizing EDFacts.

Working Group Recommendations: Below are several aspects of the current system that should be considered while modernizing.

- a. One Stop Shop – When the new pre-submission data review process/system is created, it should be in a central location or portal where an EDFacts Coordinator, or other SEA staff such as the state assessment director, can go to provide metadata, review data quality checks, and provide comments, regardless of the type of data being checked. Having one portal will provide both SEAs and ED efficiencies in the time spent reviewing data and providing metadata comments and data quality comments. For example, it would allow the EDFacts Coordinator to assist their state assessment director in completing the metadata survey if they both have access to the same portal. Even if there are data groups that must be audited later for monitoring purposes, it would still be easier to have 90% in one central location.
- b. Data Review System – The aggregate data quality rules system should stay open as long as possible so the SEAs can use it to check their data anytime. The SEAs could use the system in slow times to test out new aggregations they have on the SEA side, build a new CEDS database, work out issues they had during a previous submission, etc. This allows SEAs time to correctly prepare for when the current data are ready for submission. The Working Group understands that ED would need time to make year-to-year rollover changes. However, even keeping the prior year data review system open would allow SEAs to test the majority of their data, understanding there may be some changes to the current year reviews. Once the current year data review system is open, the file specifications, edits, etc. from ED cannot change. In other words, the data review system would need to have the flexibility to manage multiple years of data with different layouts/edits as needed.
- c. Submitting Files with Errors – Files should be submittable even with errors and comments that address the errors. If edits are required before file submission, then there could be issues with files being submitted on time. Currently, SEAs can submit their files with data errors since most edits are only checked after submission. If files cannot be submitted with comments that address the error until all edits are resolved, it could dramatically impact the timeliness of submissions. This could be resolved by allowing files to turn off edits.
- d. Edits/Comments
 - i. Transparent Edits – As is done now with the business rules repository, all edits that will be applied to the data by ED should be shared in advance. This would allow SEA work to be more efficient. If edits in the ED's data review system use manipulated data (for example, aggregated data or ratios), ED should provide the full details of which reported data items were used.
 - ii. Comment History – With one system, there should be a way to keep SEAs from having to repeat answers to data quality reviews and metadata (such as Charter schools, yes or no) each year. One solution is to pre-slug the responses from prior years and have the SEA actively agree there are no changes. This would be more efficient than the SEA answering the same questions, and providing the same edit responses, year after year.

- iii. Apply State-Level Edits – The current edit reports can be burdensome when the error is in an LEA or school (it can be like finding a needle in a haystack through the list of data errors). There should be edits applied on a state-side level (for example, cannot compare school level counts to state level counts because of out-of-state placed kids. Currently, this must be explained multiple times every year). If the data are aggregated and already available in the system, then the SEA can complete the review process before it gets to final submission.
- iv. Ability to Turn Off Edits – States need to be able to turn edits off at the state level as they use the new system. When turning off edits, states would provide ED with the reasons why the edits are not applicable for them in order for the file to be submitted.
- e. SEA Submitters – Some states work with a true *EDFacts* “Coordinator” who manages or coordinates with all SEA data owners to submit, and review, their data through *EDFacts*. In other states the coordinator has actual data ownership over much of the data submitted. The system should be able to support both types of SEA *EDFacts* management, allowing for the *EDFacts* Coordinator to set up the users for their state based on their state needs. For example, in some states the state assessment coordinator would submit the metadata but in other states the *EDFacts* Coordinator is the most suited to that task. The *EDFacts* Coordinator would assign their SEA staff to the system and level of access (read only, edit, submission, etc.). At submission there could be a note indicating who from the SEA was a part of the submission (as an example, if an *EDFacts* Coordinator submits the assessment metadata, they could indicate they did this in consultation with their assessment coordinator).
- f. Using Unit-Level System to Stop Duplication of Effort - Generate should be able to create reports for data that are already in the system. For example, membership and poverty quartile files are the same data looked at in different ways at different times of the year. Thus, Generate should be able to use data more than once (create different reports from same data) when needed without the SEA expending effort to tell the system they are the same data.

3. Privacy and Security of SEA Unit-Level Data

For some SEAs, if any part of the *EDFacts* reporting process (the business rules service, for instance) involves the SEA moving student or staff-level data onto a third-party site, even if ED does not have access to this site, it is a security issue. Moving unit-record level data off the SEA servers is an automatic trigger for SEAs, causing more burden for most SEAs. *EDFacts* Coordinators are not typically the person, or group, in the SEA that makes these rules, so even if the *EDFacts* Coordinator understands the process and believes it is not a risk, they are not the ones that need to be convinced. Therefore, if any part of the system includes moving SEA data to a third-party server, most states would not be able to participate.

Working Group Recommendations: ED should find a way to implement the business rules without moving data out of the state’s secure environment. Aggregate versus unit level are two different conversations – anything with unit-level data outside the SEA servers is problematic if not critical. A cloud-based system where states can take advantage of the rules engine in their own environment would work. An SEA’s own Generate instance could retrieve the standard business rules and run them against their data in their own secure environment.

4. Current System Investment

All SEAs will be starting in a different place in terms of their data structures, federal reporting processes, and support systems. Some SEAs already have a well-thought-out and historical federal reporting system that works for them. SEA federal reporting systems include complex data manipulation processes, and if SEAs are already processing their data to an individual-level ED*Facts* dataset, how do they find benefit in moving this process to Generate? There is an understanding of the value of using unit-level rules and data quality checks since SEAs do this with their LEAs, but right now, some SEAs do not have the plans, resources, or political will to make changes to their current federal reporting system. SEAs have spent significant resources on their current systems and could be in the middle of multi-year contracts with vendors to support their current system. Some may have vendors/contracts that support more than federal reporting for the SEA, and some may have been so far ahead of the curve that they are now behind. For SEAs that would need to shift their programming from reporting the aggregate to conforming their system to CEDS, they see too large of a cost and time barrier to change the systems they have now.

Working Group Recommendations: There are several different recommendations to support states that do not currently see a way forward to making a system-wide change:

- a. Support Aggregate Pre-Submission System - Allow aggregate reporting and aggregate pre-submission rules service for those that will not be changing their systems. This is part of the proposal now but continue to dedicate ED funding and support for this reporting option moving forward.
- b. Provide SEAs Start-up Funding – Many SEAs would consider changing to the CEDS methodology if they had funds to get this started. The states that have made the move already noted that the effort is in the beginning, creating plans to change their systems and moving their system to be CEDS compliant. They reported it was a significant amount of work and they received grant funding to get more staff onboard for this period.
- c. Get Vendors Involved – If ED could find a way to get more SEA vendors to incorporate the Generate business model it would help all states considering making this move. ED could create a vendor system that would allow vendors to submit test files that go into the file checker.
- d. Partial Participation – It would help more states transition over if ED allowed states to submit some files using the CEDS methodology and some using the state methodology. Meaning, allow states to transition in file by file over a period of years. This would address the issue of the large upfront investment needed to do it all at once.
- e. Where Does ED Start – The SEAs need to know from ED where they plan to start. ED should prioritize files that historically have more errors (for a discussion point), identify ED*Facts* files that are more problematic, and start with those (comparability of data).

5. Post-Submission Edits Still Needed

From experience with federal reporting and collecting data from LEAs, SEAs know that sometimes there are edits that do not get discovered or developed until after the data are submitted and reviewed. However, applying edits to data months after release would undermine the pre-submission process. Some balance needs to be found between these.

Working Group Recommendations:

- a. Provide additional detail for how ED will handle across school years data quality checks for SEAs that make it past the validation process. ED should provide for resubmission in this new environment and explain why edits might be run after the data have already gone through the new system.
- b. Provide all business rules that will be applied across files and across school years for those states that can run these validations in systems that store longitudinal data.
- c. Year-to-year edits - The comparison often is between the current year and the original file submitted from the prior year, rather than any resubmissions. The system should allow comparisons to files that have been resubmitted with revisions.

6. CEDS Connections (Unit-level to Aggregate Reporting)

Currently, CEDS provides Connections for every *EDFacts* file. These Connections include the CEDS data elements needed to create the file, a crosswalk from the *EDFacts* permitted value and the CEDS option, and some analytical notes. However, they do not include how to pull all the data together or the *EDFacts* business rules that will be applied to the data. This is very important to modernization, and it is also the most difficult to standardize amongst states. Providing unit-level to aggregate-level mapping and rules in CEDS Connections would help all states, not just those using the CEDS methodology. These Connections could also touch upon state issues that are not directly discussed with *EDFacts*, such as a state having different or more permitted values (such as race/ethnicity) and how these would be rolled up into *EDFacts* reporting.

Working Group Recommendations:

- a. Create CEDS Connections for all *EDFacts* file specifications that explain in detail the process to go from unit record data to aggregate data. As a “pilot”, create one sample unit record to aggregate Connections that explains in detail how Generate aggregates the unit record data. Use proven development processes for these Connections such as the CEDS Open Source Community. Start an *EDFacts* Working Group made up of *EDFacts* Coordinators and their business/technical staff together to create and/or modify CEDS Connections for each *EDFacts* file specification. These Connections will show the process for transforming unit record data into aggregate data for category sets and totals provided to ED through the *EDFacts* reports. In this process, any differences in interpretation will be documented and provided back to the working group to compare their SEA process to the Generate process and look for discrepancies. This process will allow ED to provide any additional clarity to file specifications, and for SEAs to challenge their own processes. Also, these CEDS Connections need to be kept up to date as part of the *EDFacts* file specification release process.
- b. Involve CEDS in the process for updating or releasing new *EDFacts* file specifications. This will ensure existing CEDS element names are used where possible and/or new CEDS elements will be created as early as possible. Through this process, creating the CEDS Connections prior to *EDFacts* file specification release will show with specificity how all aggregations will occur for these newly released file specifications.