

# DATA QUALITY IMPROVEMENTS FROM STATES USING GENERATE

**Author: Sara Sinani**

CIID has implemented Generate in 9 state education agencies (SEAs), and through this work SEAs have realized a variety of improvements to their special education data in terms of overall data quality, improvements to data governance structures, and ease of federal reporting. This brief is the second in a [series](#) looking at data quality. Here, we will look at some specific ways the implementation of [Generate](#) has helped SEAs realize overall improvement in data quality and as a result, prompted improvements to data processes and procedures.

## Strategies for Improving Data Quality

Through the implementation of Generate, SEAs receiving intensive TA indicated one of the greatest outcomes was in the quality of the data files submitted to the ED Facts system. As they went through the implementation process of creating the Extract-Transform-Load (ETL), and the data review process, several data quality issues were identified. These were issues that, prior to Generate, would have either not be identified or would have been caught much later in the process and potentially required a resubmission of the data to ED Facts.

Generate implementation follows the [CIID Data Integration Toolkit](#) which includes steps for identifying all business rules necessary to program the ETL to move data into Generate to produce the desired ED Facts reports. Through this process the SEA's data practices are reviewed against ED Facts file specifications. For the 9 SEAs that are implementing Generate, this process of business rule and data practice review for creating the ETL revealed data quality issues that are outlined below. As a result, SEA staff were able to identify and correct the issues, make updates to data cleaning processes within the SEAs, and help confidently create finalized ED Facts files ready for submission without worry of misidentifying/representing students and

### States that work with CIID indicated:

Implementing Generate has resulted in real and potential data quality improvements.

counts within the files.

## Strategies for Improving Data Quality

- Several SEAs were incorrectly excluding or including students under their legacy processes for creating ED Facts files. These errors were identified during the comparison of the legacy system files to the Generate-produced data files. In order to support SEAs with higher quality data files, Generate has built-in error checks needed for the files, which made the issues easy to narrow down and correct within the SEA data processes.
- Issues of data duplication were noted during data validation. CIID TA providers and SEA personnel found instances of “final” (cleaned) data that contained duplications of students based on specific program assignments and SEA staff were able to correct the duplications.
- Generate has identified issues with specific elements that had been manually entered and submitted to the SEA rather than being calculated based on date ranges. As Generate pulled the data into respective reports, SEA staff discovered discrepancies when compared to their legacy systems. Because Generate uses fixed elements (such as birthdate, entry/exit date, etc.) to determine age and length of time, the counts within Generate that include such elements, become intrinsically more accurate.

## Improved Data Governance Structures

The Generate implementation process also involves SEA data governance. In many cases, prior to CIID and Generate, SEA special education data was siloed in one department and was not shared with or discussed by any other department areas. With Generate being a cross-team implementation project, it inherently promotes cross-communication within the SEA. Just the act of preparing reports within

Generate has allowed personnel across programs to come together as a team and problem solve data quality issues that often come up last minute. This type of process has necessitated an increase in communication and planning around IDEA data in order to make sure that the data is submitted on time and accurately.

In one example, because of an identified incorrect process for creating a file, the SEA used the Data Integration Toolkit Step 2 and Step 5 as a guide to put together an approved list of business rules, documentation of roles and responsibilities, and a proper communication plan in place for the future. This will ensure that moving forward the process is consistent, standard, and provides more transparency.

## Ease of Federal Reporting

While seemingly unrelated to data quality, having higher quality data in Generate to work with creates efficiency in the report-making process. Having all the reports, quality checks, and production of files in one place has made a big impact on the coordination of efforts needed to submit ED Facts files. For example, SEA personnel involved in ED Facts reporting can review the same reports in the same location, where previously they may have had to run multiple reports separate from one another. In many cases, due to data quality issues prior to implementing Generate, these SEAs previously had to go through multiple

This process drastically decreased the amount of time from starting the review and submitting the files than the previous process, and there is more confidence in the quality and accuracy of the submitted data.

**Observation from a state using Generate**

iterations when creating submission files. With improved overall data quality and an increased emphasis on good data governance and communication, SEAs saw much faster timelines from review to final submission of the data. SEAs using Generate have stated that the ability to do data quality reviews prior to submission is expediting the submission process by allowing the SEA to address data notes and issues on the front end rather than post-submission. As the number of states using Generate grows, we can only expect the number of useful reports and features to grow with them. Looking forward, with the addition of several report features included in Generate (such as student population reports, and data quality reports), more SEAs that begin to use Generate will be able to do better and faster data quality reviews.

## Conclusion

The 9 states that have implemented Generate are realizing improved data quality within their data systems, are updating and improving their data governance structures, and have realized a more efficient and higher quality process for producing, reviewing, and submitting required federal reports. These improvements are not only beneficial to the process, but they are highly beneficial to SEAs and their resources as a whole. As indicated in an evaluation of SEAs who are using CIID resources, including Generate, over three-quarters of Generate implementers reported actual or anticipated cost savings or maintenance (in contrast to anticipated increased expenses with the legacy system).

For more information on Generate or how CIID can help your SEA please contact CIID at [CIIDTA@aemcorp.com](mailto:CIIDTA@aemcorp.com).

## Related Resources

[States Realizing Data Quality Improvements Through Data Integration](#)

[States Improving Data Quality Through IDEA Data Integration](#)

[Data Quality Through Data Integration](#)



Contact Us:

By email at [CIIDTA@aemcorp.com](mailto:CIIDTA@aemcorp.com)

Visit the CIID website for more information at [CIIDTA.org](http://CIIDTA.org)

This content was produced under U.S. Department of Education, Office of Special Education Programs, Award No. H373M140001. Amanda Hoffman and Meredith Miceli serve as the project officers. The views expressed herein do not necessarily represent the positions or policies of the U.S. Department of Education. No official endorsement by the U.S. Department of Education of any product, commodity, service, or enterprise mentioned in this publication is intended or should be inferred.