Forum Guide to Data Quality

National Forum on Education Statistics




Forum Guide to Data Quality

National Forum on Education Statistics

National Cooperative Education Statistics System

The National Center for Education Statistics (NCES) established the National Cooperative Education Statistics System (Cooperative System) to assist in producing and maintaining comparable and uniform information and data on early childhood, elementary, and secondary education. These data are intended to be useful for policymaking at the federal, state, and local levels.

The National Forum on Education Statistics (Forum) is an entity of the Cooperative System and, among its other activities, proposes principles of good practice to assist state and local education agencies (SEAs and LEAs) in meeting this purpose. The Cooperative System and the Forum are supported in these endeavors by resources from NCES.

Publications of the Forum do not undergo the same formal review required for products of NCES. The information and opinions published here are those of the Forum and do not represent the policy or views of NCES, the Institute of Education Sciences (IES), or the U.S. Department of Education (ED).

November 2023

This publication and other publications of the Forum may be found at the websites listed below.

The NCES Home Page address is http://nces.ed.gov

The NCES Publications and Products address is http://nces.ed.gov/pubsearch

The Forum Home Page address is http://nces.ed.gov/forum

This publication was prepared in part under Contract No. 91990021F0031 with Quality Information Partners, Inc. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. government.

Suggested Citation

National Forum on Education Statistics. (2023). *Forum Guide to Data Quality* (NFES 2023-086). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Technical Contact

Ghedam Bairu

(202) 245-6644

ghedam.bairu@ed.gov

Foreword

The Forum is pleased to present the *Forum Guide to Data Quality*. The purpose of this document is to provide best practices from districts and states for maintaining quality data as these agencies regularly review and revise their methods for working with data, as well as their expectations for all staff who are responsible for education data and their quality. This document also highlights the U.S. Department of Education (ED) efforts to improve data quality. The information will help districts and states seeking to improve or further develop their policies and practices related to data quality.

Publication Objectives

This resource aims to help agencies

- understand how quality data are defined;
- maintain data quality; and
- improve and develop policies and practices related to data quality.

Intended Audience

The primary audience for this publication is staff schools, LEAs, SEAs, and the federal government whose responsibilities include any aspect of collecting, reporting, analyzing, or using education data.

Organization of This Resource

This resource includes the following three parts:

- **Part One** includes three chapters that cover key aspects of data quality.
    - **Chapter 1** defines data quality and discusses the importance of quality data.
    - **Chapter 2** discusses building a culture of quality data in an education agency.
    - **Chapter 3** focuses on best practices for agencies to achieve quality data.
- **Part Two** offers case studies from local and state education agencies.
- **Part Three** offers tip sheets for key stakeholders who help to ensure quality data.

National Forum on Education Statistics

The work of the Forum is a key aspect of the Cooperative System. The Cooperative System was established to produce and maintain, with the cooperation of the states, comparable and uniform education information and data that are useful for policymaking at the federal, state, and local levels. To assist in meeting this goal, NCES within IES—a part of ED—established the Forum to improve the collection, reporting, and use of elementary and secondary education statistics. The Forum includes approximately 120 representatives from SEAs, LEAs, the federal government, and other organizations with an interest in education data. The Forum deals with issues in education data policy, sponsors innovations in data collection and reporting, and provides technical assistance to improve state and local data systems.

Development of Forum Products

Members of the Forum establish working groups to develop guides in data-related areas of interest to federal, state, and local education agencies. They are assisted in this work by NCES, but the content comes from the collective experience of working group members who review all products iteratively throughout the development process. After the working group completes the content and reviews a document a final time, publications are subject to examination by members of the Forum standing committee that sponsors the project. Finally, Forum members

review and formally vote to approve all documents before publication. NCES provides final review and approval before online publication. The information and opinions published in Forum products do not represent the policies or views of ED, IES, or NCES. Readers may modify, customize, or reproduce any or all parts of this document.

Working Group Members

This online publication was developed through the Cooperative System and funded by NCES within IES—a part of ED. The Data Quality Working Group of the Forum is responsible for the content.

Chair

DeDe Conner*, Kentucky Department of Education

Members

Jon Bridges*, formerly of the Beaverton School District (OR)

Deb Came, Washington Office of Superintendent of Public Instruction

Raifu Durodoye, Regional Educational Laboratory--Northwest

David Hand, NOLA Public Schools (LA)

Laura Hansen*, Metro Nashville Public Schools (TN)

Georgia Hughes-Webb*, West Virginia Department of Education

Peggy Jones*, Pasco County Schools (FL)

Allen Miedema*, Northshore School District (WA)

Hope Moffett*, Brandywine School District (DE)

Adrian L. Peoples*, Delaware Department of Education

Linda Roska*, Texas Education Agency

Cheryl L. VanNoy, Saint Louis Public Schools (MO)

Consultant

Bridget Thomas, Quality Information Partners

Project Officer

Ghedam Bairu, National Center for Education Statistics

* Working group members marked with an asterisk also contributed case studies or real-world examples to this guide.

Acknowledgments

The Data Quality Working Group would like to thank everyone who reviewed or otherwise contributed to the development of the *Forum Guide to Data Quality*, including the following:

Core Reviewers

Stephen Gervais, San Bernardino City Unified School District (CA)

Linda Jenkins, Arkansas Department of Education

Ryan Kuykendall, DeSoto County Schools (MS)

Bradley McMillen, Wake County Public School System (NC)

Melanie Stewart*, Milwaukee Public Schools (WI)

Case Study/Vignette Contributors

Molly Abend, Chandra Haislet, and **Dawn Hubbard,** Maryland State Department of Education

Becky Hampton, Metro Nashville Public Schools (TN)

Todd Nelson, Union Public Schools (OK)

Jon Wiens, Oregon Department of Education

* Reviewers marked with an asterisk also contributed case studies or real-world examples to this guide.

Contents

Glossary

Business rule: a framework or set of constraints under which an organization operates, and the expression of those constraints as a mathematical or logical assertion governing how data can be entered, used, or calculated within a data system

Data check: used to confirm that data were collected according to standards and guidelines

Data element: the most basic unit of specific information that can be defined, measured, and stored

Data file specification: a technical document that establishes rules for how each piece of information should be submitted, and provides guidelines on handling specific situations, such as missing data

Data governance: refers to a formal and comprehensive set of policies and practices designed to ensure the effective management of data within an organization—encouraging robust data security, definition, collection, access, quality, and disposal

Data note: descriptive text associated with reason code that describes data quality issue or warning found through data check

Data stewards: those individuals within an organization who are responsible for one or more data collections, implementing data governance policies and standards, and maintaining data quality and security

Reason code: code used to identify a data quality issue or warning found through a data check

Validation: a process that ensures that data agree with expectations for reasonable values and accepted norms

Part One:
Data Quality

Chapter 1: Introduction

The collection and use of education data have grown exponentially in recent decades. As local and state education agencies (LEAs and SEAs) use increasing amounts of data to answer questions, identify concerns, and make decisions, quality data are critical. Quality data help educators at all levels to make informed decisions and improve teaching and learning. Some examples include teachers using formative data in classrooms to guide instruction, administrators using data to justify funding, LEA staff using data to evaluate programs, and SEA staff using data for federal reporting. Quality data result from the efforts of stakeholders throughout the education system who understand why particular data are collected and used; who adhere to agency policies and best practices for collecting, maintaining, and using data; and who promote a culture of quality data, in which everyone who has a role in student outcomes understands their responsibilities in the data process.

Data quality is increased when it is emphasized from the point of entry into the collection system through use for reporting. LEAs and SEAs have different methods of reporting; many LEAs report data through datasets to the SEAs, while others use statewide collection systems. SEAs report through datasets for federal and state reporting. The focus of this publication is best practices to ensure quality data reporting.

What Are Quality Data?

Quality data are complete, valid, and accurate. These qualifications are individually important, but data must first be both complete and valid to be accurate. Complete data include all expected information. Valid data must have been evaluated against certain defined rules to ensure the correctness of the data both in structure and content. While these three qualities are not the only components of data quality, they provide a key foundation upon which other aspects are built.

Completeness

Data that are considered complete meet two qualifications. The first qualification is that all expected records or respondents are entered into the collection system and included in the datasets used for reporting. For example, an LEA might check that all schools expected to respond to a data collection submitted data, and an SEA might check that required data have been submitted by their LEAs. The second qualification is that each record or response includes all expected information. For example, a record for a student discipline incident in which the student was suspended that did not include the number of days of the suspension would

be lacking critical information. Student enrollment data, which indicate whether a student's name was, is, or will officially be registered at a school or schools, show the importance of completeness to quality data.[1] At the SEA level, completeness means that for each submission, the SEA has data from each LEA (and within each LEA, data from each expected school) and each individual data element within each required file is submitted.

For data to be considered complete, zeros, missing data, or data anomalies must follow the guidelines established by the file requirements or specifications. A file specification is a technical document that establishes rules for how each piece of information should be submitted, and provides guidelines for handling specific situations, such as missing data. Some data collections may require, for example, that any anomalies are explained with a data note or reason code.

Data notes and reason codes are especially useful for large datasets, in which a single data anomaly can lead to multiple reported errors. Some collections require that an agency submit an action plan for fixing missing or anomalous data issues before data are considered complete.

Validity

Valid data conform to expectations for reasonable values and accepted norms.[2] For example, when collecting staff and student phone numbers, valid data typically do not include letters, only numbers. Validity also encompasses formatting; for example, data standards often specify the format for entering the month, day, and year in date fields. Consistency in whether items such as months are stored as text or as values is another formatting issue that can affect data validity. Additionally, valid data do not contain duplicates or redundancies, which can affect calculations.

The Oregon Department of Education has front-end validations that are checked upon data loads. These typically are restricted to ensuring the submission conforms to the file format and business rules for the collection (that is, the data are valid and internally consistent). Once all data are submitted, data stewards review the data for additional quality checks, including reasonableness and consistency. Individual rows of data are flagged, with a reason for the flag entered, and those rows are uploaded in an audit system for school districts to review. This gives them the option to confirm that the data are correct or to correct any errors that are present. This audit process has significantly improved data quality for the state.

Validity and completeness should be checked against business rules as data are submitted. Business rules guide the creation and management of quality data. A business rule is a framework or set of constraints under which an organization operates and the expression of those constraints as a mathematical or logical assertion governing how data can be entered or used within a data system. For example, an agency that only serves students between the ages of 5 and 21 may have a business rule that states that values for the data element Student Age must fall within the range of 5 to 21 (that is, $5 \leq$ Student Age ≤ 21).

Accuracy

Accurate data pass edit-checks, follow data quality rules, and contain no known errors. In many cases, the data are certified by the appropriate party.[3] Accuracy follows validity because data cannot be accurate if they are not valid. Data accuracy is specific to each data collection and refers to whether data adhere to pre-set standards and expectations. Examples of expectations:

1 U.S. Department of Education, National Center for Education Statistics, Common Education Data Standards. *CEDS Elements v. 10*. Retrieved August 28, 2023, from https://ceds.ed.gov/elements.aspx?v=10.

2 National Forum on Education Statistics. (2020). *Forum Guide to Exit Codes*. U.S. Department of Education. Washington, DC: National Center for Education Statistics.

3 ED*Facts*. *Maximizing EDFacts Data Quality: A Comprehensive Approach*. Retrieved August 28, 2023, from https://www2.ed.gov/about/inits/ed/edfacts/edfacts-data-quality-process-overview.pdf.

- **Reasonableness**–Values do not fall outside of an expected range and are comparable to similar collections. For example, a state might compare data for schools with similar demographics to establish parameters for what is considered reasonable.

The collection of student guardian data in Milwaukee Public Schools (WI) attested to the importance of accuracy. To avoid the risk of sharing secure data with people who should not have access, the LEA ensures that legal documentation identifies a grandparent or other adult as guardian prior to their information being entered into the student information system (SIS).

- **Consistency**–There are no large, unexpected changes in data from what was submitted in the prior collection period. For example, communities with stable industries would expect stable student enrollment. Unexpected and unexplainable changes may indicate a data accuracy issue.
- **Summation to totals**–The sum of elements is equal to subtotals and totals, and combined subtotals are equal to totals. For example, the sum of individual educational attainment categories should equal total membership.
- **Internal consistency within a file**–Values entered for individual elements are consistent with related elements. For example, age is consistent with birthdate.
- **Consistency across files**–Data submitted across multiple files are aligned. For example, daily attendance codes for individual students should align with the aggregated totals for the school, which are then reported to the LEA.

Accuracy also must be viewed in the context of a particular situation. Large shifts could occur across time periods and may not be errors. For example, Milwaukee Public Schools (WI) created a virtual program in one of the district's schools, and the enrollment skyrocketed after the first year. This shift was striking, but the data were both valid and accurate.

Ensuring Accuracy

Since home addresses can be written in different formats and with varying abbreviations, it is possible to have multiple versions of the same address appear in the student information system (SIS). To avoid the possible confusion this may cause with residency, some districts use an address database to ensure that each address only appears in one form in the SIS. Rather than typing an address, staff members select an address from a prompt.

Data collections should be reviewed and certified for accuracy. Certification is the process of revalidating data after all edits have been reviewed and resolved and ensuring that the data can pass all business rules. Data that do not pass all business rules may be certified, provided that an accepted reason is given. For example, LEA data provided to SEAs may be incomplete when data collections are affected by natural disasters. An action plan that accompanies the data can explain why the agency cannot provide the data and any necessary steps planned to ensure that the data are available for the next collection. For more information and best practices that education agencies can adopt before, during, and after a crisis, see the *Forum Guide to Planning for, Collecting, and Managing Data About Students Displaced by a Crisis* (https://nces.ed.gov/forum/pub_2019163.asp).

A Framework for Data Quality

One model that provides a more detailed structure–and that builds upon the foundation of completeness, validity and accuracy–is the Data Quality Framework established by the Federal Committee on Statistical Methodology (FCSM).[4] In response to the rapidly changing world of data sources and analysis methods, the FCSM established a Data Quality Analysis Working

4 Federal Committee on Statistical Methodology. (2020). A Framework for Data Quality. FCSM 20-04. Federal Committee on Statistical Methodology. September 2020.

Group in 2020 to provide practical information on identifying and reporting data quality for federal agencies. The working group established a quality framework that provides an inventory of the elements (that is, domains and dimensions) of data quality with a review of identifiable threats to each dimension of data quality. The FCSM Data Quality Framework provides a common foundation upon which federal agencies can make decisions about the management of data products throughout their lifecycle by identifying and mitigating key data quality threats, evaluating trade-offs among different quality dimensions where necessary, applying accepted methods at an appropriate level of rigor, and accounting for and reporting on the quality of data products and outputs.

While the FCSM Framework was designed with federal agencies in mind, it provides a structure that is useful at many levels. It divides the concept of data quality into three key domains: utility, objectivity, and integrity, and each of these domains is composed of multiple dimensions. Agencies or stakeholders working to establish and maintain data quality—especially when working with staff newer to the concept—may be able to use the components and definitions provided by the framework to expand understanding of the necessary concerns involved in a data quality process and culture.

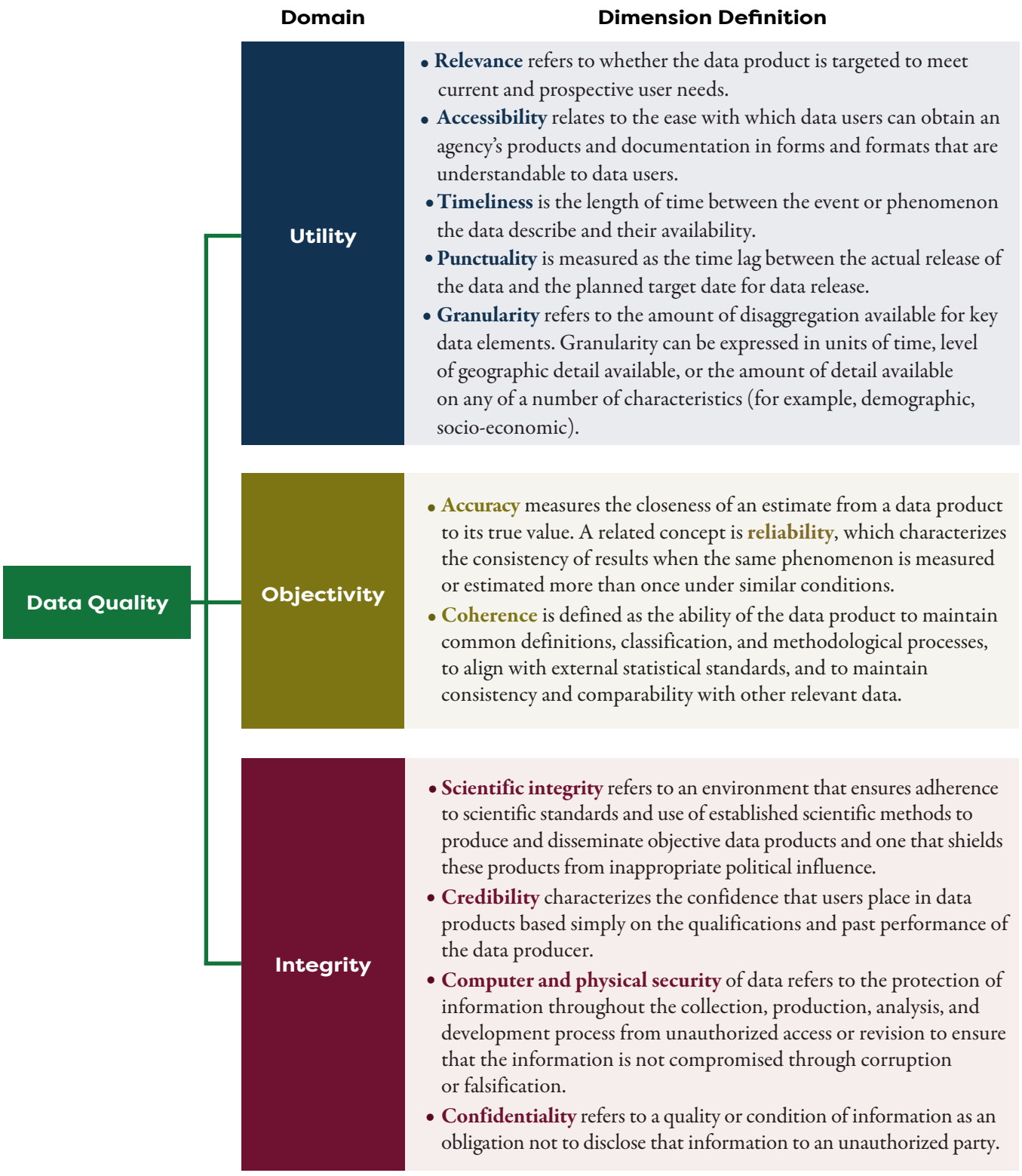

Figure 1: FCSM Data Quality Framework

Why Are Quality Data Important?

Education agencies rely on high-quality data for all aspects of their operations, from supporting students in the classroom to planning, funding, and providing efficient programs and services to students. Quality data are also essential for communicating information about education to stakeholders such as parents, community groups, boards of education, policymakers, and others. Although data collected by schools and LEAs are used to inform local decisions and meet local needs, many also are required for federal and state reporting. Federal staff rely on relevant, accessible, timely data collected at the right level of detail, or granularity, to make decisions. The quality of these data is essential to ensure that schools receive the appropriate resources, comply with laws and regulations, and make effective internal decisions about how services are provided to students.

Quality data allow agencies to track and use key information, meet reporting requirements, and shape and support their data strategy. Quality data help

- teachers make the right decisions about their students' instructional needs;
- district leaders assess the effectiveness of systems and services in aiding student development;
- principals track student and teacher progress to make sure goals are being met and problems that impede progress are identified;
- district personnel apportion staff or other resources equitably;
- state departments of education plan and manage effective programs;
- state lawmakers make informed decisions about public education programs and funding;
- researchers evaluate the impact of education programs; and
- parents, policymakers, and other stakeholders understand how resources are making a difference in education

Because practices related to data collection and use change over time, education agencies must review their policies and procedures regularly to account for new requirements, changing technologies, and shifting data needs. This guide offers considerations and best practices for agencies to ensure that their policies and procedures help improve data and encourage a culture of data quality.

Data Requests from Stakeholders

Clear, specific data requests provide a stronger basis for communication between requestors and data stewards. When requests for information in West Virginia's SEA go beyond what is regularly available, an initial data request process with a brief form allows state data staff to pull the data for the requestor, if possible and in compliance with standard privacy protections. The form clarifies the request and allows the data team to keep track of requests to inform planning for additional public reporting. Beyond this level of information is a tier for researchers who want access to restricted-use or suppressed data. In these cases, researchers must submit a more official and detailed proposal application.

More information about data requests can be found in the *Forum Guide to Strategies for Education Data Collection and Reporting (SEDCAR)* (https://nces.ed.gov/pubs2021/NFES2021013.pdf).

Chapter 2: Building a Culture of Data Quality

A culture of data quality is, among everyone who has a role in student outcomes, a shared understanding of the importance of data quality and commitment to ensuring data quality. In an agency with a strong culture of data quality, staff members and stakeholders demonstrate a common belief that good data provide critical support for teaching, learning, planning, governance, and strategy. The culture stems from an agency's intentional and integrated strategy to ensure that stakeholders understand the purpose of the data and to give stakeholders a chance to review, discuss, and support quality data and the role of data in evaluation and decision-making. These stakeholders can include students and parents, school office staff, teachers, data stewards or coordinators, technology support personnel, principals, school board members, superintendents, U.S. Department of Education (ED) staff, and staff of state and local education agencies (SEAs and LEAs). In the culture of data quality, these stakeholders all play a part in supporting data integrity, which allows trust in the data standards, credibility, and security of the data.

> Collaboration is the foundation of a culture of data quality both within and across agencies.

Key Steps to Build a Culture of Data Quality

1. Engage all stakeholders.
2. Establish a data quality team whose members possess the qualifications necessary to manage the agency's data collections and have access to effective data tools.
3. Offer ongoing professional development so that all staff stay informed about current requirements, changes, and new collections or procedures.
4. Support school-level staff who enter data into established systems but do not work with data regularly.
5. Regularly demonstrate the importance of quality data, for example, by demonstrating how quality data are tied to program funding.

To promote a culture of data quality, it is useful to provide agency staff with answers to and additional information about the following questions:

- *How will the data be used?* Demonstrating the connection between data collections and the day-to-day operations of classrooms and the school helps stakeholders who are involved in data work recognize how high-quality information supports their work and improves education outcomes.
- *What data are collected?* An overarching understanding of the data collections helps contextualize the data and identify potential quality issues, such as duplicate or outdated data collections and fields.
- *When are the data entered, checked, and validated?* Data staff often need to meet multiple reporting deadlines throughout the year. Ensuring that these staff know when the tasks related to each data collection must occur helps them to meet deadlines and control their workflow. General information about the data entry, review, and

> To avoid inaccuracies in data once it moves beyond its origin in the LEA, LEAs can clarify how the data will be maintained within external systems after being pulled from their student information system (SIS). For example, one LEA asks how data such as "name" and "gender" will be maintained, and only provides these data if a system can accept corrections to these elements (and others) that are not static. The LEA's reasoning is that incorrect or invalid data are typically worse than no data: outdated data is invalid but may be treated like it is valid if this issue is not considered.

validation process can help staff understand how data are processed.

- *Who enters, validates, and manages the data?* Clear roles and responsibilities around data collection help ensure that data are entered accurately and on time. Similarly, clarity around how data will be validated and managed moving forward is critical. Leaders should ensure that staff members understand how their diligence and careful practice when entering and reporting data help the agency provide services and meet reporting requirements.
- *What are the consequences for poor data quality or inaccurate reporting?* Clear processes for auditing data and known consequences for inaccurate reporting are essential for dedicating sufficient resources to quality reporting. For example, questionable data may be suppressed in public reporting or not used for making important performance decisions such as accreditation ratings.

By knowing this key information, agencies create confidence in their data and lay the groundwork for implementing the best practices described in Chapter 3.

Data Quality Culture Throughout the Education System

Data quality must begin at the data's sources, the LEAs and the schools from which the data are collected. LEAs often take the first steps toward creating and ensuring quality education data. LEA staff, including school staff, initially collect the data that are used for administration, instruction, and operations. These data are entered in local data systems and may be shared later with other LEAs, the SEA, and federal agencies for additional planning and reporting.

Agencies with a culture of data quality understand that data quality can be strengthened by strong collaboration within and across local, state, and federal levels. High-quality, valid national or state level data are crucial for education leaders who make decisions about programs, policies, and funding. Leaders also need to understand how local results compare nationally or at the state level, as well as how current data compare to those collected earlier or under different circumstances.[5] Though federal, state, and local education agencies serve different roles in the data process, they have responsibilities that should complement and support one another.[6]

Agencies support data quality when

- data definitions are aligned, understood, and consistent between LEA, SEA, and federal education entities;
- SEA staff work with LEA data coordinators to establish sound, practical procedures for collecting, managing, and reporting data while remaining sensitive to data burdens;
- LEA and SEA offices responsible for data reporting cooperate to identify and remove redundancies and to consolidate requests;
- LEAs and SEAs work together to create new data elements and indicators when needed;
- LEAs and SEAs work together to retire the collection of data elements when they are no longer needed or relevant;

5 Spiegelman, M., & Merlin, J. (2023, August 30). *Using Federal Education Data to Inform Policymaking: Part 1-Benefits and Advantages*. NCES Blog. U.S. Department of Education, National Center for Education Statistics. Retrieved from https://nces.ed.gov/blogs/nces/.

6 Spiegelman, M., & Merlin, J. (2023, August 30). *Using Federal Education Data to Inform Policymaking: Part 2-Challenges and Opportunities*. NCES Blog. U.S. Department of Education, National Center for Education Statistics. Retrieved from https://nces.ed.gov/blogs/nces/.

- ED works closely with SEAs to plan and improve data collection processes to ensure quality data at the federal, state, and local levels, while simultaneously aiming to reduce data burdens on SEAs and LEAs;
- LEA leaders work with their data coordinators to make sure that modifications to data collections (such as definitions, calendars, or reporting) align with mandated reporting requirements and support timelines, are well understood, and are properly communicated; and
- LEA data coordinators collaborate with school staff responsible for data entry to provide input to student information system (SIS) vendors and identify efficiencies in data collection.

School Level

The data collection process technically begins at the school level, with data collected and entered by school staff, clerks, teachers, and others. The foundation of data quality includes effective training that ensures that those entering data understand the requirements for doing so accurately, as well as validation edits and local review. Best practices at the school level include

- developing tools to enable efficient review of data (especially during collections), which can identify data that do not conform to the business rules or expected values and formats;
- implementing controls to prevent common errors during data entry (for example, specifying the preferred format for date fields rather than allowing for free-response data entry);
- including training on the business rules under which data are collected as part of the professional learning provided on using data systems;
- working with district staff to regularly review aggregate data—at least in key areas like enrollment, attendance, discipline, and finance—to monitor for anything that looks anomalous or out of place;
- making data quality monitoring and feedback tools available so school staff notice when data quality issues occur and correct them quickly;
- assessing the root causes of data quality issues, with a goal of preventing them from reoccurring, rather than just fixing them; and
- cross-training staff on data systems and collections so that when schools are facing staff absences or attrition, they are still able to collect and input data.

LEA Level

To develop a culture of data quality in an LEA, it is best practice that district staff work with schools to help staff understand the purpose of the data collected and, where possible, connect the data to a specific instructional program or goal. It is also important for school personnel to understand the reasons behind the collection for data that may not be related to instruction, such as data required to qualify for financial resources or data required for the Civil Rights Data Collection (CRDC). Program funding is tied to accountability measures that, in turn, are based on data collected by schools and districts. Therefore, it is critical that districts provide staff with the resources needed to produce quality, on-time data.

Districts must adhere to deadlines set by other agencies as well as their own data needs and schedules. Creating and communicating a district data calendar will help track the times when reports are due, when the corrections windows open and close, and when schools must provide data to meet these deadlines. A district data calendar will also communicate timelines for making modifications to collections to district leadership. Additionally, LEAs often do a

certain amount of "managing upward" when working with their SEA. SEAs may not be directly familiar with the front-facing data entry systems in LEAs, and managing expectations and communicating with the SEA can contribute to an improved culture of data quality.

> When Beaverton School District (OR) began implementing a new Multi-Tier System of Support (MTSS) module provided by the district's student information system (SIS) vendor, an IT systems analyst, a teacher on special assignment, and the data analyst responsible for reporting outcomes collaborated to refine dropdown menus in the MTSS module's teacher interface. Their work also established shared data definitions for data entry fields and choices in dropdown menus. This collaborative effort improved the data entry interface for users (in this case, teachers), ensured professional development training with teachers, supported clear understanding of the data to be entered, and aligned data collection with reporting requirements for evaluating the effectiveness of reading interventions and supports for elementary students.

Districts also must ensure that data and data systems adhere to policies and regulations set by state and federal authorities as well as their own internal policies. District staff can more easily meet these policy and regulatory requirements if they are confident in the data they receive. To that end, district personnel often are responsible for training data collectors and for ensuring that the data collected are of high quality. Districts can help schools identify the authoritative sources for data and reduce data collection and reporting burdens by determining when data are duplicated in different school and district systems. Additionally, districts can implement automated quality controls to ensure data accuracy by helping staff members identify and correct data errors while entering them.

Districts can benefit from including school personnel in the creation of district data policies, regulations, and processes. For example, a process for transferring data from a school to a district should involve

- staff responsible for developing relevant reports and information;
- representatives from the IT and data teams;
- representatives from the schools involved; and
- data stewards at multiple levels.

> In the Northshore School District (WA), data leaders recognize the value of having the communications team involved with any messaging regarding data collections that are sent to families. The district receives better results when asking families to update data when the district's communication professionals are involved.

By including everyone involved in a data collection in the planning stage, district data stewards can more effectively create a collaborative environment that fosters a culture of quality data. In this environment, the people responsible for all aspects of the reporting cycle will carry out their work with a full understanding of required tasks and why they are important. Collaboration and communication are the foundation of a culture of data quality both within and across agencies. When guidelines are developed within a process that considers multiple viewpoints and needs, districts and schools can create real-world procedures that enhance the quality of information across the state as well as within the district.

SEA Level

SEAs provide the foundation for a culture of data quality statewide. Their roles and responsibilities in collaborating with LEAs are critical to achieving quality data. Given the complexity of collecting data from multiple LEAs and reporting data to ED, SEAs play a key role in fostering collaboration among stakeholders at all levels and implementing data-quality controls and processes.

SEAs set standards and expectations for data collection and reporting among LEAs and SEA offices. Standards define data elements and outline proper data entry procedures and business rules, and they should reflect state and federal requirements as well as local and state information needs. SEAs can help LEAs adhere to standards in a number of ways, including

> The Oregon Department of Education holds a monthly Data Collection Committee (DCC) meeting, which is open to all data submitters and information system vendors. Attendees discuss any changes or pending changes to data collections. If the SEA is proposing a new collection or a major change to an existing collection, they ask that members of this DCC form an informal working group to work with the data steward to discuss and resolve any issues that arise. SEA leaders note that this helps make school districts a partner in the work.

- providing tools such as data standards[7] that help LEAs follow data collection procedures;
- publishing a calendar of data collection due dates, including data correction opportunities;
- conducting audits to ensure the accuracy and completeness of data;
- establishing a data governance process that includes LEA input in the development process for new and modified collections;
- working with LEAs to locate and correct data errors within data correction windows;
- offering training materials and manuals covering data definitions and collection procedures;

> Data leaders in Milwaukee Public Schools (WI) have monthly meetings with the SEA to remain up-to-date on data quality, learn about changes, and troubleshoot issues. Leaders also attend a quarterly meeting that is statewide for large districts.

- providing "train-the-trainer" opportunities to create local experts on the use and functionality of the SEA data system;
- facilitating helpdesks to respond to LEA questions and issues;
- sharing data collection changes, deadlines, and details through newsletters[8] or other existing communications;
- providing clarity on how data quality impacts agency processes, functions, and responsibilities; and
- working with districts to resolve data quality issues that originate from other districts and may cause conflicts when reported to the SEA level.

The SEA also is an important resource for districts. Like schools and districts, the SEA relies on high-quality data and can support districts with information on data elements, collection timelines, and alignment across collections.

Federal Level

ED's Data Strategy[9], published in 2020, established strategic goals for advancing data capabilities and envisions agency-wide outcomes. It established an ambitious vision: to realize the full potential of data to improve education outcomes and lead the nation in a new era of evidence-based policy insights and data-driven operations.

In its 2023 update, the Strategy included 16 new guiding principles within four key areas:

- Strengthen Agency-Wide Data Governance
- Build Human Capacity to Leverage Data

7 For an example of state data standards, see https://education.ky.gov/districts/tech/sis/Pages/KSIS-Data-Standards.aspx.
8 For an example of a state data newsletter, see https://education.ky.gov/districts/tech/sis/Pages/KSIS-Newsletters.aspx.
9 U.S. Department of Education. (2020). *Data Strategy*. Retrieved September 1, 2023, from https://www.ed.gov/sites/default/files/cdo/ed-data-strategy.pdf.

- Advance the Strategic Use of Data
- Improve Data Access, Transparency, and Privacy

Like both ED's inaugural Data Strategy and the ED's updated Data Strategy, all work carried out under the updated Data Strategy will be informed by the three guiding principles of ethical governance, conscious design, and learning culture.[10]

Proposed changes to data collections are shared through the U.S. Office of Management and Budget (OMB) to allow SEAs and LEAs an opportunity to comment on the visibility and impact of changes. ED encourages comments and considers impact on quality of data when reviewing.

Forum Guide to Reporting Civil Rights Data
https://nces.ed.gov/forum/pub_2017168.asp

This guide presents a variety of effective methods through which local education agencies (LEAs) report civil rights data to the U.S. Department of Education's Office for Civil Rights, including examples of how state education agencies can voluntarily help their LEAs with Civil Rights Data Collection (CRDC) reporting.

Forum Guide to State Education Agency Support for Local Education Agencies in Civil Rights Data Reporting
https://nces.ed.gov/forum/pub_2023026.asp

This supplement to the original guide, published in 2023, provides several detailed case studies from states that currently support their LEA reporting.

As ED collects data from states and districts, it utilizes multiple processes and tools to ensure data quality. These data quality efforts include those targeted toward the largest collections, CRDC and ED*Facts*. These ED collections have allowed states and districts to report data through file uploads with data notes, receive data quality feedback from the system, and allow for states and districts to revise files and resubmit. Recently, the ED*Facts* modernization process was launched to bring the reporting system in line with technology, data tools, and data needs that had changed significantly since ED*Facts*' inception. Modernization will improve ED reporting systems to support the growth of state data systems, provide business rules to be tested ahead of reporting, and provide immediate feedback in a more detailed and yet easy to use interface, thus allowing these data to be submitted, reviewed, and released quicker and with more accuracy.

10 U.S. Department of Education. (2023). *Data Strategy*. Retrieved November 8, 2023, from https://www2.ed.gov/about/offices/list/opepd/ocdo/ed-data-strategy.pdf.

Chapter 3: Best Practices for Collecting Quality Data

Agencies can implement best practices to ensure that they are collecting quality data. Best practices focus on the policies and processes for collecting data as well as the people responsible for the data. A focus on policies and processes can help agencies ensure that they have established standards and clear expectations for data collection, data quality checks, and data management. A simultaneous focus on people can help agencies foster communication, provide effective training, and ensure that all staff understand the purpose of the data collection and its use.

Establish Data Standards and Guidelines

Standards and guidelines encourage respect for accurate and useful data. They define data elements as well as outline proper data entry procedures and business rules, and they should reflect state and federal requirements as well as local and state information needs.

> ***Forum Guide to Data Governance***
> https://nces.ed.gov/forum/pub_2020083.asp
>
> This Forum guide highlights the multiple ways that data governance programs can benefit education agencies. It addresses the management, collection, use, and communication of education data; the development of effective and clearly defined data systems and policies to handle the complexity and necessary protection of data; and the continuous monitoring and decision-making needed in a regularly shifting data landscape.

Agencies benefit from establishing guidelines for how data should be submitted and from ensuring that the guidelines are available to all staff who work with the data. Guidelines may include file requirements that provide information on rules for the submission process and procedures for how to handle situations such as missing data. These are useful even if datasets are small and informal. For example, establishing a standard way to represent a missing value in a dataset saves staff the time of seeking out the answer and helps to avoid situations where staff may enter different values.

Standards and guidelines are part of data governance, which refers to the formal and comprehensive set of policies and practices designed to ensure the effective management of data within an organization. Data governance is a foundational part of the data collection and reporting process, and it is concerned with the people, processes, and technology tied to maintaining safe and effective data management, use, analysis, and communication. Many state education agencies (SEAs) and local education agencies (LEAs) may find it useful to use the Common Education Data Standards (CEDS) to support accuracy in data entry and to facilitate data sharing.

> **Common Education Data Standards (CEDS)**
> https://ceds.ed.gov/
>
> CEDS is a set of common data terms and definitions that allow data to be shared across platforms and among varied data teams and education levels and includes data models and several online tools to support users. CEDS serves as an important data management tool for many education agencies.

Agencies can help staff understand and adhere to standards in multiple ways, including

- providing tools that help staff follow data collection procedures;
- publishing a calendar of data submission due dates;
- conducting audits to ensure the accuracy and completeness of data;
- working with staff to locate and correct data errors;
- offering training materials and manuals covering data definitions and collection procedures;
- facilitating helpdesks to respond to questions and concerns;

- sharing data collection changes, deadlines, and details through newsletters or other existing communications; and
- establishing the primary source of the data and avoiding redundant collections or storage where multiple versions of the data can exist, causing inaccuracies.

At the federal level, the U.S. Department of Education (ED) provides multiple data quality resources for data submitters that specify ED standards and guidelines, including

- standardized reporting and data validation instructions for data submission;
- a standardized data dictionary;
- tools built into the submission system to check for errors and to review metadata about datasets; and
- guidance documents to assist with data submission.

Implement Data Checks

Data checks are used to confirm that data collection adhered to standards and guidelines. They "ensure the completeness, validity, and accuracy of any new data, and ensure these practices are integrated into the broader schedule and system for data cleaning and auditing."[11] These data checks can be invaluable to ensuring accuracy. For example, in Pasco County Schools (FL), checks identified a student's home incorrectly being mapped to the middle of the Atlantic Ocean after the district added longitude and latitude coordinates to students' addresses.

Explain the Purpose of the Data and Promote Data Use

Agency staff who understand the purpose of the data they collect as well as how those data will be used to support students are more likely to enter the data carefully and precisely. Agencies can provide staff with information about the data's intended purpose and how the data meet the needs of data users. Data leaders can increase staff members' recognition of the key components of data quality by maintaining focus on the "why" of the data: In short, the ultimate purpose of the data collected and used is to serve students and support their needs. Information on the purpose of the data and how the data will be used can also help information technology (IT) staff who are responsible for developing the technology that enables data collection, storage, and interoperability.

Data Quality Dashboards

The Rhode Island Department of Elementary and Secondary Education uses data quality dashboards to provide data managers with a clear and private view of their own data to help identify and correct any errors.[12]

Tools and reports such as data dashboards and visualizations increase the visibility of agency data and encourage data use. In turn, data visibility and use demonstrate the importance of data and encourage staff to take ownership of data quality as well as to use the data to inform decision-making. The more staff use data, the more likely they will be to consider data quality.

11 U.S. Department of Education. Office of Elementary and Secondary Education. *Data Quality Component 2: Data Quality*. https://oese.ed.gov/resources/oese-technical-assistance-centers/state-support-network/resources/data-quality-component-2-data-quality/.

12 U.S. Department of Education. Institute of Education Sciences. National Center for Education Statistics. (2021) *SLDS Webinar Summary Improving Data Quality: State Strategies to Enable Data Use*. Retrieved September 1, 2023, from https://slds.ed.gov/#communities/pdc/documents/20386.

Data Visualization

The *Forum Guide to Data Visualization* (https://nces.ed.gov/forum/pub_2017016.asp) offers practices to help education agencies communicate data meaning in visual formats that are accessible, accurate, and actionable for a wide range of education stakeholders.

Additional in-depth information is provided in the Forum Data Visualization Online Course (https://nces.ed.gov/forum/dv_course.asp).

Capture Data Context

Metadata, or data that describe the data, are an essential part of data collections. Metadata provide data users with the appropriate context to understand the data's purpose and significance, including any temporary or permanent changes to data definitions or collection practices that affect their use. For example, metadata are especially useful when there are changes in data from one year to the next, such as the widespread adoption of virtual education during the coronavirus disease (COVID-19) pandemic.

For more information on how metadata can be used by education agencies to improve data quality and promote a better understanding of education data, including information on how to plan and successfully implement a metadata system in an education setting, see the *Forum Guide to Metadata* at https://nces.ed.gov/forum/pub_2021110.asp.

Provide Technology and Technical Support

Data validation tools help ensure accurate data at the time of entry and allow staff to verify and correct data as frequently as needed. Automated data checks at the field level can help prevent data quality issues, and reporting tools can help with identifying issues. Available tools must be used to be effective, as it is easier to correct data errors routinely throughout the year than to wait until the end of the school year. Some agencies offer these checks nightly.

Outdated systems can hinder data quality. Older systems may not have the same features and functions that can prevent, identify, and mitigate data entry errors. In Delaware's student information system (SIS), outdated fields and duplicate fields have sometimes caused data quality issues when data sources do not align with the data system location where staff members enter updated data. Some codes were available on multiple screens, and the location where schools and districts updated them varied.

Collecting and managing data in a single system can help to eliminate the issue with multiple sources of information and confusion regarding which source is accurate. Even if data are collected in multiple systems, recognizing the authoritative source for the data is crucial. Interoperability solutions can help reduce redundant data entry and issues created due to data entry errors. These interoperability solutions will help create a more efficient situation where each data element has a single system in which the data element is managed, and all other systems become consumers of those changes.

Agencies can use the functions of their data systems to increase data quality by

- utilizing selections or dropdowns to enhance clarity and avoiding text fields wherever possible;
- applying business rules or required formatting to data fields;
- considering the workflow of the task and allowing for streamlined and organized data entry;
- providing warnings or notifications when data outside of established limits or standards have been entered; and

- ensuring interoperability to avoid redundant entry of the same data in multiple systems.

Foster Communication Among Agencies

Many education data collections involve multiple agencies, such as the school where the data are initially documented; LEAs and SEAs where the data are collected, reviewed, and used; and in some cases, the federal government. Maintaining open lines of communication can improve data quality by providing a way for staff involved in the process to clarify needs and requirements.

Beaverton School District's (OR) Information Technology (IT) Department convenes school staff responsible for entering data in the district's student information system (SIS) 6 times a year. Since user groups have different responsibilities for data entry and quality assurance, meetings are role specific (that is, elementary office assistants, middle and high school office assistants, middle school registrars, and high school registrars). These regular meetings provide users with training and support prior to data entry efforts (for example, documenting immunization day exclusions or entering parent opt-out requests from state testing) as well as the opportunity to share best practices in job-alike roles. Agendas, handouts, and recordings of the meetings are posted for future reference.

Create Workable Calendars and Timelines

Stakeholders rely on quality data that are available when needed. Effective data calendars and submission timelines include consideration of when data can be collected, the time needed to check and adjust the data, and when the data can be certified. For example, quality enrollment data may not be available immediately at the start of the school year.

Communication can also help to ensure that new or updated data collections are implemented in ways that give agencies time to prepare. For example, new or modified data elements, data collections, and reporting frequently require changes to the SISs used by LEAs. Early communications involving the SEA, LEAs, and vendors are essential for making these changes effective for all involved. Communication between the LEA and the SEA is crucial to ensure that LEAs can provide quality data. SEAs provide support to LEAs when implementing a new collection by engaging the SIS vendors about changes and providing enough time for LEAs to alter systems as needed to capture or reformat the data. For example, in Texas, an initial step in the data governance process in developing new data collections is review and approval by a committee that is composed mainly of LEA staff but that also includes SIS vendors, regional education center support staff, and SEA stakeholders.

Train and Support Staff

Skilled staff with the tools and information needed to enter data correctly are essential to data quality. Agencies can support staff who enter or work with data by offering and encouraging staff to attend professional development and training sessions that clarify data standards and practices. While many agencies offer training to new staff, it is important that existing staff also have the opportunity for ongoing training to keep their data literacy skills up-to-date and ensure that they understand the purpose of the agency's data collections.

Agencies are increasingly offering time for data stewards, information technology (IT) staff, program staff, and others to discuss data collections in meetings where they can offer support to one another and ensure that all aspects of an agency's data collection procedures are aligned. It is also useful to involve these staff in discussions about changes to procedures and updates in policy. When monitoring errors, agency leaders also should communicate the need for data staff to submit promptly, so that datasets can be corrected or SISs updated.

The ubiquity of data means that even staff who may not work with data systems on a day-to-day basis nevertheless interact with data in other ways; for example, data often inform many of the programs offered in schools. It benefits agencies to keep these staff informed about data

collections so that they understand the importance of the collections and how data are used. Often, agencies achieve this by publishing data collection updates in school newsletters or sharing key information in staff meetings.

Minimize the Burden of Data Collections

When implementing new data collections or revising existing collections, a best practice is to consider the burden of the data collection on the agencies and staff impacted. While many data collections are mandated by law, others are not, and agencies should consider the best methods and timing for obtaining needed data. To minimize the time and administrative effort spent, it is useful to weigh the potential value of the new data against the time and resources needed to collect them. It also can help to clearly communicate why the additional burden is necessary by explaining why the data are needed and their intended value for instructional support, decision-making, and compliance with state or federal requirements.

Agencies also can collaborate to promote data quality and reduce reporting burdens by ensuring that new data elements added to existing or new collections are essential and do not duplicate data collected elsewhere, and periodically reviewing data collections to remove elements that are no longer needed. For example, in Texas, by state statute, agency collections are reviewed for duplication with existing collections as an explicit step in the data governance process.

Align Data Collections with the Agency's Data Strategy

A comprehensive data strategy is a robust, integrated approach to using data to deliver on a mission, serve stakeholders, and steward resources while respecting privacy and confidentiality. A data strategy enables education agencies to leverage data to improve education, increase agency effectiveness, facilitate oversight, and promote transparency. Data strategies encompass data principles and practices such as governance, access, privacy, security, dissemination, and use by internal and external stakeholders.

The *Forum Guide to Strategies for Education Data Collection and Reporting (SEDCAR)* (https://nces.ed.gov/pubs2021/NFES2021013.pdf) provides timely and useful best practices for education agencies that are interested in designing and implementing a strategy for data collection and reporting, focusing on these as key elements of the larger data process.

Part Two:
Case Studies from State and Local Education Agencies (SEAs and LEAs)

Education agencies must maintain a focus on data quality to ensure timely, accurate information. The specific goals, processes, and outcomes of different agencies may differ, however, depending upon the nature and needs of each individual state education agency (SEA) or local education agency (LEA). The following case studies offer examples of how different agencies work with data quality and create a data quality culture.

Kentucky: A Culture of Data Quality

The Kentucky Department of Education (KDE) supports 171 school districts in which there are over 1,500 schools, with a total enrollment of almost 640,000. Kentucky uses a statewide network and statewide systems to support schools and districts. With any statewide system, establishing standards for use and reporting are the foundation of quality reporting.

KDE's data governance group promotes data quality agencywide. Agency staff members understand that data reporting requires the team to work together and think beyond their individual areas and consider the overall needs of the agency, districts, and communities. All requests for new data collections and changes to existing data collections go through this group.

Creating a Clear Data Quality Process

Early in the process, KDE learned the importance of engagement between data controllers and the office leadership team to accurately represent the office team's data governance priorities, maintain awareness of office needs, and gauge impact as change control requests are reviewed. To achieve these aims, this work must be assigned to an individual who is connected statewide and who approaches data governance from a statewide perspective while still accounting for their needs and the needs of their office.

As KDE's data governance team matured, the agency made changes to the process to ensure that data controllers reported directly to the lead for each office. Data governance is not a full-time position for anyone in the agency; instead, KDE promotes the perspective that all staff should consider governance a part of their everyday work. Data controllers in each office make up the data governance committee; these individuals work directly with the data stewards in their respective offices to promote best practices and advance work within KDE and with LEAs, focusing on data quality.

Lessons Learned

As the KDE data governance committee has developed and refined its data governance processes, lessons learned along the way have allowed the committee to establish, promote, and adhere to practices that help emphasize data quality. Examples include the following:

- **Data standards are documented to support data collections.** The standards are reviewed and updated as needed, at least annually. Standards set the foundation for data quality and must be communicated through trainings and other opportunities. Documentation of standards helps to ensure consistency so that data can be extracted from 171 districts in the same manner. Examples of KDE data standards are available on the Kentucky student information system (KSIS) standards web page (https://education.ky.gov/districts/tech/sis/Pages/KSIS-Data-Standards.aspx).
- **All data collection changes must be vetted through the data governance committee before implementation.** Each request requires support from office leadership before the committee considers it. Vetting requests through the committee helps to ensure cross-agency awareness and gauge impact of the potential change.
- **Requests for new data collections are vetted through the same change control process.** As data governance members review each request, they must consider the impact to the districts. Each request must include justification for the need as well as any associated regulatory or statutory requirements. As they review requests, data governance members consider five key factors of data quality:
    - Are there legal or regulatory requirements?
    - Are there rewards or penalties associated with those requirements?
    - Does the team agree that the request adds value?
    - Will districts see the value and validate the data?
    - Is there a defined KDE data steward who will inspect the quality of the data?

  Without these factors, data quality may be compromised; therefore, the value of the collection would not justify the added burden of the data collection.
- **Data stewards are responsible for hands-on work associated with the program and data.** The data steward is the person who
    - holds expertise regarding what the data element means;
    - trains and answers detailed questions about that data element;
    - regularly inspects the level of quality to ensure it meets not only their office needs but all the others that depend on the data;
    - works with external advisory groups to provide feedback and input to help define better solutions; and
    - proactively communicates timelines to district contacts, with instructions on how data should be validated.
- **Use of data can bring issues to light and result in increased quality.** Agencies can create tools or validation reports that can help illuminate common errors so data can be updated before key reporting time frames. For example, data visualization tools can make data more actionable and help to bring anomalies to light.
- **A centralized data request process ensures that data requests, both internal and external, go through the same process.** Requirements are documented and approved, and the data are tested by more than one person before they are shared. This focus on consistency and efficiency has increased the quality of data being shared.
- **Public use data are available on the School Report Card** (https://www.kyschoolreportcard.com/) **and Open House** (https://openhouse.education.ky.gov/) **sites.** Datasets are standardized to increase usability. An added benefit of creating public datasets is that requestors can be pointed to public data that already have been vetted for quality, reducing demands on staff time and resources that would be needed

to respond to individual requests.

- **Federal reporting through CEDS Generate** (https://ceds.ed.gov/generateInfo.aspx) **has provided KDE an additional validation tool at the state level.** It validates historically used files against current reporting and aggregation standards, which can improve consistency and help to refine reporting and improve processes.

Moving Forward

KDE's model may differ from those of other states, but data quality remains its central focus. With a statewide student information system (SIS), districts do not send or sync data directly to the SEA and errors do not eliminate data if they do not match certain parameters. KDE can export the data needed to meet federal and state reporting requirements. The agency does not have a formal written data quality plan; rather, much like data governance, data quality has become part of agency culture. Data controllers, data stewards, districts, and data requestors work together with a mindset of continuous improvement.

Delaware: A Forum for the SEA and LEAs

To improve education data governance in the state of Delaware, SEA and LEA representatives joined together to form the Delaware Education Data Forum (Delaware Forum) in 2016. The Delaware Forum is modeled on the National Forum on Education Statistics and is designed to promote collaboration between SEA and LEA participants. It is co-chaired by an SEA and an LEA representative to ensure equal representation between SEAs and LEAs on agenda-setting and decision-making. Collaboration between the SEA and LEAs and a focus on data quality at the student level have been key to the Delaware Forum's effectiveness.

The Delaware Forum's mission is to increase confidence, consensus, and clarity in the state's education data. The group wants data to be accessible and comprehensible and for varied stakeholders to use the data to make better decisions. A founding principle of the Delaware Forum is the ability of every to see the student-level data that are aggregated by the SEA. This allows greater understanding and more nuanced conversations between LEAs and the SEA. In the past, LEAs might see that aggregated data for their location were incorrect, but they did not have the detailed information (such as business rules or an explanation of which data were included) to allow them to see where inaccuracies were located.

Since the creation of the Delaware Forum, the SEA has been committed to having student-level data reviewed by the LEAs. This has required an adjusted timeline. Whereas in the past, little time existed between student data collection and aggregation, the SEA now has added weeks into the schedule to allow LEAs to view the data, identify problems, and work with the SEA to fix the problems before they aggregate the data. The LEAs have committed to intensive reviews of student-level records. Because of these shared commitments and responsibilities, the Delaware Forum is a powerful force for improving data quality.

LEAs Collaborate via the Delaware Forum

The Delaware Forum provides opportunities for LEAs to work together, learn from each other, and strategize the best way to handle key processes. More experienced LEA members can support and train newer members on what types of errors to watch for and how to validate data. The group members also discuss how to standardize processes across LEAs whenever possible and brainstorm how to handle business rules for complicated topics such as attendance and absences. By collaborating to develop solutions that meet the needs of different LEAs with different board policies, LEAs increase alignment between data collections.

SEA leaders support this collaboration among LEAs because it leads to improved data quality. It also helps to reduce burdens on LEA staff. In many smaller LEAs, the responsibility for

reviewing LEA data and submitting them to the SEA falls on an individual data steward or a small team. The Delaware Forum provides a venue for these individuals or teams to collaborate with staff in similar positions in other LEAs who may have solutions to common problems, more efficient processes, or best practices to share. Time is set aside during Delaware Forum meetings for LEAs to meet on their own (without SEA presence). Whenever possible, LEAs work together to identify solutions and make improvements to data processes.

Working together, LEAs at the Delaware Forum also identify data issues and bring them to the attention of the SEA. For example, LEAs identified a business rule that captured students who had passing grades in core subjects on their student record but did not differentiate between students who earned failing marks and those who were missing mark reporting records entirely. This made data validation more time-consuming for LEAs. In another case, LEAs identified an issue in which International Baccalaureate (IB) scores were needed for a college- and career-readiness indicator, but these scores did not come out until after data for the indicator were collected. In each of these cases, the SEA was able to go back and correct data, align timelines, or adjust business rules to ensure data accuracy and to increase the efficiency of data quality review.

More Effective Data Sharing Between LEAs and SEA

The ability of LEAs to review student-level data, as well as increased communication between the SEA and LEAs, allows for more accurate and useful data sharing between the agencies. The SEA gives LEAs downloadable student-level data, and LEA members of the Delaware Forum discuss best ways to check and validate the data. In small districts or charters, data teams may take the aggregate measure and review student by student. Larger districts may do spot checks, or they may look at data for targeted students. Others may do a pivot table, create an aggregate, and then—based upon their knowledge—check to see if aggregates meet expectations.

Interactions between LEAs and the SEA also can solve problems across the state. For example, the SEA calculates 14 or 15 aggregates for accountability for which LEAs are provided individual records. This can lead to hundreds of thousands of records. The Delaware Forum is intentional in its focus on the attitudes of "divide and conquer" and "a rising tide will lift all boats." If an LEA raises an issue that reveals something systematically wrong with the data, the state will fix it for all LEAs in the state. Thus, though LEAs may focus on different areas, problems raised and solved ultimately benefit everyone.

These relationships also allow the SEA to communicate desired changes and the logic underlying them. For example, the SEA uses School Courses for the Exchange of Data (SCED) codes to identify if students are taking the courses necessary each year to be considered on track for graduation. Delaware encourages all LEAs and charters across the state to report data using SCED codes for this measure to be comparable, but SEA leaders found in the data that the codes are not used universally, which may lead to underreporting at the SEA level. With these data—and the impacts of them not being complete—in hand, the SEA can better explain the problem to LEAs and charters and redouble its efforts to encourage all agencies to use SCED.

Increased Trust in Data Quality Increases Data Use

Involving LEAs directly in the data quality process and providing insight into how the SEA level aggregates are created has increased LEAs' level of trust in the data. In the past, when aggregated data from the SEA did not align with LEA data, LEAs were unable to see business rules, data sources, and processes. Therefore, they were unable to determine whether the aggregates were correct, and they were hesitant to use the aggregate data. Now, LEAs' critical role in the review process means that they can question data that do not align with their expectations and identify the source of discrepancies. As a result, LEAs are less likely to dismiss the SEA aggregate data. Their increased trust in the data leads to greater data use for tasks like decision-making, planning, and program design.

Going through the validation process builds confidence in the data and emphasizes the importance of training and consistent standards. LEA staff can recognize that a clear focus on data quality at the outset helps avoid later issues. The on-track attendance indicator is an example of how greater confidence improves data use. Prior to reviewing this indicator as part of the data validation process, many LEA staff were unfamiliar with the definition of "on-track attendance" and the business rules guiding the measure. Data validation for this indicator functioned as a learning tool for new state accountability requirements. For example, some principals working on school success plans reviewed student-level attendance, which showed them the links between absences and academic struggles. LEA staff across the state now have a better understanding of critical numbers of absences, the key metrics involved in attendance measures, and why attendance is included in the accountability metric. By working with the data, they can see the risk factors.

Increased trust in and understanding of the student-level data has led educators in Delaware to use data in additional ways. Though public accountability data is not available until October, school administrators used cleaned and embargoed data for strategic planning by the end of the summer. These data let them see which individual students need interventions, which was not possible with the aggregate data available to them in the past. Additionally, program offices are using data to implement statewide initiatives, such as those related to career and technical education (CTE). Because these offices have more confidence in the data, they are using them for more purposes.

Moving Forward

Delaware has made great strides in data quality through the work of the Delaware Forum. Increased communication among LEAs, as well as between LEAs and the SEA, has brought more stakeholders into the data quality process and given them ownership over the data that they provide and validate. Leaders of the Delaware Forum want to do more and are thinking about ways to improve the data quality process further. At the SEA level, leaders have committed to improving communication with LEAs and providing more support to LEAs for large data collections such as the Civil Rights Data Collection (CRDC). As LEA data staff become more advanced in their knowledge of things like logic models and business rules, SEA and LEA leaders are poised to increase the sophistication of data collection and reporting in the state and to increase data use across many different areas. Finally, though the majority of LEAs participate in the Delaware Forum, the group's leaders aim to ultimately expand participation to all LEAs and charters in the state.

Metro Nashville Public Schools (TN): Continuing Dedication to Data Quality

Metro Nashville Public Schools (TN) has 22 data quality specialists responsible for monitoring data quality, identifying the root cause of data quality issues, and guiding and coaching the staff at each school to implement solutions to the issues. The Office of Data Quality and Integrity also has four managers and two analysts, giving the district far more staff directly devoted to data quality and governance than most education agencies.

Part of the mission of the Data Quality and Integrity Office is a focus on building capacity in other staff who are involved in the process of collecting, entering, and using data. The Data Quality and Integrity team members identify errors and discuss why the entry was an error and how it impacted other data systems and operational processes. By helping staff understand where the data are used and seeing how they fit into the larger picture, the data quality team has increased the understanding and investment of leaders and staff across schools and district offices to ensure that they enter data correctly. This not only decreases the number of individual errors, but also the repercussions of those errors when data are pulled into and used in different systems and reports.

Increasing Efficiency Also Raises Data Quality

This collaborative relationship between Metro Nashville Public Schools (TN)'s data quality team and staff entering data at different levels allows them to identify cases where changes in processes or technology could make improvements to the efficiency and effectiveness of data collection, and adjustments often are based on feedback they receive from staff. For example, when the SEA mandated a collection on the prevalence of dyslexia among students as part of a state focus on reading, the totals initially were difficult for the LEA to collect because the data were at the school level and not necessarily specified in reports. The information sometimes was in students' individualized education programs (IEPs), but not all students with dyslexia have IEPs. The data quality team made recommendations in collaboration with the district department (who are the subject matter experts on the use of the data) and the IT departments. These recommendations resulted in processes for consistent collection of dyslexia information and a data collection form within the SIS accessible to the people who need it. Automation was created so that when staff in schools entered information into the form, the state classification needed for state reporting would be assigned and reported accurately to the state data system. This new process allowed Metro Nashville Public Schools (TN) to provide accurate reporting on the prevalence of dyslexia among students to the state in an efficient and effective way.

Metro Nashville Public Schools (TN) was similarly successful in addressing how student academic plans could be improved. In the past, school counselors needed to review individual student transcripts to determine whether students were on track to graduate. To reduce burdens and increase efficiency, the Office of Data Quality and Integrity served as the liaison with schools and district offices to identify various stakeholder needs and then worked with the IT department to create a solution tailored to the needs of the LEA. Because the Data Quality and Integrity team knows the data and dependencies needed to configure the academic planner tool in the SIS, it was able to create a solution that allows schools to more quickly and proactively determine if students are on track for graduation. The academic planner tool also has helped schools address an SEA requirement that high school students have a 4-year plan. Prior to the development of the tool, each school approached the planning process differently. Now that the academic planner tool includes these plans, there is greater consistency across schools in the LEA, and district leaders are better able to monitor progress and ensure that all students have 4-year plans and are moving successfully toward achieving them.

In another example of system efficiency increasing data quality, Metro Nashville Public Schools (TN)'s Data Quality Dashboard provides data validation reports generated from the district's data warehouse. This dashboard makes it easier for school staff to monitor data quality by presenting on-demand, actionable information that is updated daily and includes what errors exist and information required to correct them. Three levels of errors are reported:

- First, the reports identify critical errors that affect accountability, funding, or student instruction/services.
- Next, there is a level of errors that need to be corrected but are not high priority.
- Finally, there are "warnings" for deviations from standardized business rules, which may have exceptions that eliminate them as errors. These are checked to be certain they are not exceptions.

Errors also are categorized by type—scheduling, attendance, or grading, for example.

The Data Quality Dashboard can be accessed by anyone at the school level who has access to the SIS. The system shows the user the number of errors of different types, and by clicking on the number, the user receives a more detailed report with additional information that assists in correcting the errors. This technology tool makes it easier for school leaders and staff to identify

data issues, correct them in an ongoing way, and have information that assists in identifying the root causes of the issues so the same errors are minimized in the future. The Dashboard also can be viewed at the district level, which allows for the identification of trends across schools. When the same data errors show up frequently in large numbers, the root cause is often the need for better communication of expectations or processes among district leaders or staff.

Identifying Root Causes of (and Solutions for) Data Quality Issues

The Metro Nashville Public Schools (TN) data quality team sees root causes of data quality issues falling into three areas: people, processes and policies, and technology. The team identifies the root causes of each data quality issue, then develops solutions appropriate to the specific area.

People: People need to know what needs to be done and how to do it. Clear communication of expectations is important, and monitoring of compliance is a basic requirement. However, because there is a lot of focus on compliance within data collection and reporting, staff may think of data entry as a collection of business rules and timelines rather than a crucial contribution to the ability of the district to make decisions, improve, and perform the reporting required to receive funding and ensure accurate accountability. Making staff members aware of the uses of the data they enter, as well as their importance, increases adherence to practices that ensure quality data. For example, accurate enrollment and scheduling data are critical to ensuring school funding and are used in multiple systems and reports.

Training is similarly essential. The Office of Data Quality and Integrity works closely with the training department to ensure that training on the SIS includes instructions for operating the SIS as well as information on the business rules that guide the data entry. Staff do not receive access to data systems until they have been trained, and the data quality team works closely with the training team to ensure that documentation is accurate and current. Through training, staff members learn about the technology and its related processes.

To provide additional support for staff in the schools, the Data Quality and Integrity team offers labs where school staff can receive additional guidance on complex data entry activities (such as master scheduling) that occur less frequently.

Processes and Policies: Many requirements for data collection come from policies on what must be reported established by the SEA. These requirements form the basis for the business rules that dictate what data must be collected. Then, it is time to establish processes that provide consistent ways for data collection to be operationalized. Ideally, these processes should establish the "who, what, when, where, and how" of the data collection.

For attendance, the state establishes policies regarding who is responsible and accountable for keeping accurate data, as well as defining the minimal code set that must be collected and reported. In the past, the district did not have standardized processes for many aspects of attendance data entry at the school level, and schools took many localized approaches that caused issues and errors when the data were used outside school operations. For example, if attendance data were not entered by a specific time, the callout system would not notify parents that their child was absent.

When the district went with an SIS that relied on period attendance entry to create a daily code, it became extremely important to put consistent processes in place that ensured accurate attendance data entry for every period so that daily attendance values used for a large amount of reporting were accurate and usable across schools and for state reporting.

The SEA sometimes adds new data collections and often adds new codes to many existing data collections. This requires the LEA to develop new policies and processes to ensure effective data collection, reporting, and use.

Technology: To limit technology issues as root causes for data quality problems, Metro Nashville Public Schools (TN) works to ensure that the technology used serves the staff members entering the data. Issues such as troublesome interfaces or confusing field names are identified and, if possible, adjusted to increase clarity and understanding. Because most data entry in the LEA is now at point of service, forms and interfaces need to be laid out in a way that data collection can occur at the time it is received and in conjunction with the work being done.

Additionally, the LEA tries to develop ways that technology can simplify processes for schools, such as use of the academic planner tool. The schools' technology systems already included the tool, but schools were not taking advantage of the planner because they still were identifying on-track graduation by reviewing individual transcripts. Metro Nashville Public Schools (TN)'s data quality team demonstrated the value of building out the technology, including its potential to reduce the amount of staff time spent reviewing transcripts. Once staff saw the benefit of the academic planner tool, they understood that using it could make their work more efficient and effective.

Moving Forward

Throughout the creation and expansion of the LEA's data quality efforts, members of the Metro Nashville data quality team have recognized and promoted the idea that consistency is key. Additionally, they have worked to increase the understanding of all those involved in the data process that all data entered throughout the LEA come together in myriad ways to answer questions and provide a larger picture of education in the LEA. By showing individual data staff members why their contributions are critical, they have helped staff see data as part of an integrated system, as well as a tool to improve teaching and learning. Moving forward, the data quality team will continue this focus on consistency and maintain efforts to advocate for updated technology that meets the needs of all who enter and work with data in the LEA.

West Virginia: Improving Data Quality in a Centralized Data System

In the West Virginia Department of Education (WVDE), achieving data quality requires collaboration; the responsibility for ensuring that data are complete, accurate, and valid lies with those who collect, review, and report data. WVDE has a centralized system called WVEIS (West Virginia Education Information System) that allows the SEA to implement consistent data standards and manage the SIS for LEAs.

Development of WVEIS

WVEIS was developed in the 1980s, partially in response to the need to decrease the time it took the SEA to provide data to policymakers. At that time, the process for the SEA to pull data that could answer policy questions was lengthy and included contacting schools, requesting data, and correcting any errors prior to responding to policymakers. The requirement for all LEAs to use WVEIS eventually was codified in state code.

The system was rolled out broadly in the 1990s, and over time, WVDE identified advancements to increase usability. The original WVEIS, which functioned on a basic DOS platform, eventually was replaced by a network-dependent web-based data entry interface called WVEIS on the Web (WOW) and then by a more advanced web-based platform colloquially known as "The Portal," which provided a more advanced interface accessible from anywhere. WOW included controls that improved the quality of data and improved the user experience, but over time, WVDE identified additional areas for improvement. New data management needs, combined with the need to update WVEIS to align with current technologies and user behaviors, led WVDE to develop WVEIS 2.0 and begin the transition to this updated platform.

Evolution of WVEIS 2.0

WVDE began the transition to WVEIS 2.0 prior to the 2022-23 school year. The new platform is a large-scale change, built upon careful evaluation and planning by SEA leaders and data experts. As they approached the transition, WVDE staff made a concerted effort to tighten data standards, which included data format, data definitions, data field requirements, and course codes. The data team also designed functions to ease the burdens of end users while improving data quality across the state. WVEIS 2.0 provides easy access to data standards and definitions and restricts the values that users can input to minimize inconsistencies. In the past, WVDE allowed more variation in data entry, which resulted in additional work for SEA staff who had to manage discrepancies after data were submitted. LEA staff are learning the new and updated codes that were introduced with the new system. Although the introduction of new standards has required some initial training, the transition is increasing the quality of the data because users must check code definitions instead of entering traditional codes that may be outdated or incorrect. These adjustments have ensured standardization across the state and made data processes much easier for the SEA. LEA staff also have seen how standardization improves the quality of the data they use to support students.

Challenges and Lessons Learned

WVDE has faced challenges as it works through the transition to WVEIS 2.0. Some challenges resulted from importing data from the old SIS. For example, past inconsistencies in course coding among LEAs meant that the SEA data team had to clean the data after they were submitted and ensure that the final data aligned with the required codes and standards. This history of the SEA making these adjustments on the back end meant that data transferred from the old system still could be incorrect. Moving forward, data controls that require the use of standardized codes, course names, and definitions will eliminate these inconsistencies and increase data quality.

Moving beyond these earlier variations in course code use also required new understanding on the part of the LEAs, which the SEA needed to facilitate. Courses and their descriptions have changed considerably since WVEIS was developed, and some LEAs had been entering older codes that were no longer valid. Users sometimes were surprised that old codes did not import into the WVEIS 2.0 and that they could not type their local name for the course. WVDE is working to build edit checks into the new system, making it easier for LEAs to see errors prior to submission. Users can see data edits via a dropdown list and are able to quickly export issues by category or overall.

WVDE also has learned from earlier transitions that it is critical for the SEA to consider the data process and how LEAs use the system to enter and maintain data before upgrading the technology. The WVEIS team worked closely with LEAs and considered their feedback as they coordinated the transition. Previously, the SEA required a specific data entry process without adequate testing and had to revert to its older process. This issue taught the SEA the importance of making sure that transitions are gradual, well explained, and carefully planned to minimize problems for users. In the WVEIS 2.0 transition, the data team is making the system user-friendly and providing opportunities for LEAs to practice new processes prior to the full rollout. Recognizing that former DOS green screen users like to be able to tab between fields, WVDE made sure the portal allows this feature. The SEA offered trainings on the new system and provided "sandboxes" with LEA data available that allowed users to access the portal and learn how to use it prior to the launch of WVEIS 2.0.

Moving Forward

WVDE expects to continue the transition process for WVEIS 2.0 throughout the 2022-23 and 2023-24 school years. Beyond the rollout of the system and the learning curve for LEAs, the

SEA will be testing the success of the system during each major data task over the year, such as the first set of report cards, the December child count, year-end assessments, and high school transcripts. The data team will need to watch for complications or issues, and promptly address them in the system to ensure that it runs effectively and improves data quality for West Virginia.

Maryland: New Data Collection Provides Opportunity to Improve Data Quality

The Maryland State Department of Education (MSDE) continues to enhance data quality and data governance while implementing new processes and data collections in response to state initiatives.

Updating Data Governance Processes to Improve Data Quality

The SEA's standard process requires LEAs to batch their data in accordance with the appropriate MSDE data collection manual and submit the data to MSDE via a secure network. The MSDE data collection manuals used by LEAs ensure the processes of data collection, transmission, and verification complies with all relevant federal and state statutory guidelines. Maryland is a strong local control state that does not currently use an SIS to capture data collections.

MSDE uses a web-based data collection system that enables the SEA to review error reports and LEA-related comments. This review process allows MSDE staff to work with the LEAs to discuss their data, resolve errors, and make data loading as seamless as possible. Using Statewide Longitudinal Data System (SLDS) Grant funds (https://nces.ed.gov/programs/slds/), the SEA has been systematically implementing initiatives put forward by the governor, many of which focus on expanding data collections and improving data governance.

MSDE works closely with the 24 LEAs in the state to ensure consistency and review data concerns. The SEA's Student Records System Manual (https://www.marylandpublicschools.org/about/Documents/OCP/Publications/StudentRecordsSystemManual2020.pdf) offers high-level guidance on data governance and the submission processes for LEAs. In addition, state regulations require each LEA to have a local accountability coordinator (LAC); these individuals meet biweekly to receive updates regarding assessments, accountability, and data collections. LACs are encouraged to share concerns and generate solutions within this forum for specific situations encountered by their LEAs that may not be covered explicitly by the guidance in the state's manuals.

Implementing a New Data Collection

Based on a 2019 law requiring all LEAs to convert student home addresses into corresponding census tract and block numbers, MSDE partnered with the Maryland Longitudinal Data System (MLDS) Center (https://mldscenter.maryland.gov/) to implement a new data collection in the 2022-2023 school year. The statute also required the MLDS Center to develop a protocol for converting the student home addresses for the collection. Prior to the collection, the SEA brought four of the state's 24 LEAs together for a workgroup. The workgroup discussed data included in LEA's school information systems that could be used to determine the needed information and they considered potential barriers LEAs may face. MSDE and MLDS Center staff met with the workgroup multiple times, and the MLDS Center used input from the group to develop a utility that converts student addresses into census tract and block numbers. One of the reasons this utility was developed was to protect student privacy by keeping individual student addresses from being submitted to the state. The SEA piloted the utility with the workgroup to test the process, provide feedback for protocol documentation, and allow MSDE to inform and improve its function prior to a larger rollout.

During the summer before the new collection, MSDE and MLDS Center staff provided technical assistance to each LEA. The technical assistance webinars explained the rationale for the collection, specified the requirements, and demonstrated how the utility would be implemented. This provided LEAs the opportunity to review their current data and ensure that

their student addresses were accurate. The design of the utility was to create a common file layout in the format needed for LEA submission.

In the first collection, 96.88 percent of student addresses could be converted into census tract and block numbers. In addition to following the protocol for the utility, LEAs referenced the MSDE manual that includes business rules and frequently asked questions. The utility also generated a detailed error list to help improve the quality of the LEA's data. MSDE worked with the LEAs to address their concerns (such as a student not having a set address) and kept superintendents aware of how the collection was progressing. Keeping varied stakeholders engaged with the process helped to encourage buy-in for the new collection.

Challenges Addressed

As the new collection was implemented, MSDE mitigated several challenges related to the utility and the nature of census tract and block data. Some LEAs had firewall issues when downloading earlier versions of the utility, while others found that their input file was not created correctly, and they needed assistance with coding. When LEAs were unsure why an error occurred, the SEA worked through the issues with them to find resolutions. Once MSDE received files from the LEAs, they verified that state assigned student IDs were valid in the state's system and the submitted geolocation IDs were valid 15-digit numbers.

LEAs were invested in improving their conversion rate percentages and worked with MSDE and the MLDS Center to address the challenge of student addresses that could not be converted to census tract and block numbers. For example, students living in mobile home parks often have the mobile home park address listed rather than specific addresses within the park. Some apartment addresses posed similar challenges. LEAs indicated that students living in these areas often were those who would qualify for higher funding if properly identified. MDSE is working with LEAs to identify more accurate addresses for students. In the case of mobile homes, one possibility is to use latitude and longitude coordinates for each mobile home. Additionally, the SEA is updating the utility to provide better information about census tracts and blocks and how LEAs can identify the best geolocation ID when multiple matched addresses are returned.

Lessons Learned

MSDE leaders identified several key lessons that emerged while the SEA brought the new collection forward. First, piloting is the best way to introduce new collections. In addition to identifying potential concerns, running a pilot provides the SEA with time to engage with LEAs to discuss the forthcoming collection, increasing both understanding and buy-in.

SEA leaders also note that greater transparency can increase LEA focus on data quality. Giving LEAs the timeline for the new collection and offering ongoing technical assistance elevated trust levels with LEA staff. LEAs also benefit from the SEA considering various state requirements and deadlines and putting strategies in place for data collection timelines.

Moving Forward

Maryland will use this data to support the implementation of a methodology to calculate compensatory education funding, but the new census tract and block data collection also provides new opportunities for LEAs to use data strategically and think creatively about data quality. Local agencies can use this new data for mapping purposes, combine it with national sources of data such as the Census Bureau's American Community Survey, and potentially identify areas where targeted programs may be needed. Moving forward, the SEA plans to work with LEAs to help them facilitate their data quality efforts.

Part Three:
Data Quality Tip Sheets

The various stakeholders involved in data collection, reporting, and use share many tasks and responsibilities related to ensuring data quality, but there are also actions and concerns specific to different roles. These tip sheets are designed to highlight responsibilities, considerations, and actions associated with stakeholders commonly involved in the data collection process.

In addition to focusing on the specific components of data quality, all who work with education data must consider data privacy in all steps of the data collection and reporting process.

The Student Privacy Policy Office (SPPO) leads the U.S. Department of Education's (ED) efforts to protect privacy. SPPO provides leadership, oversight, and coordination to ensure ED and field compliance with federal privacy laws, regulations, and policies. SPPO manages and maintains the ED's privacy program to include the enforcement of student privacy laws, the development and evaluation of privacy policy, and the management of privacy risks across ED. More information can be found at https://www2.ed.gov/about/offices/list/opepd/sppo/index.html.

The Role of Office Staff

Responsibility

You are responsible for entering important data accurately and completely and understanding how the data will be used. By entering the data correctly, you lay the groundwork for all further steps in the data quality process.

Things to Think About

- Who uses the data that you enter? Users might include the state education agency (SEA), school board members, principals, parents, teachers, students, and payroll staff.
- How are data used to make important instructional or operational decisions such as student placement?
- What is the effect of the data you enter on schools or programs?
- What can you do to ensure the accuracy of the data you enter?
- When do the data need to be available to the end users?
- What training documents or standards are available to guide data entry?

Things to Do

- Enter data accurately and promptly.
- Ask for and attend professional development programs.
- Ask for or create appropriate instructions and documentation to manage data.
- Ask clarifying questions to understand your role and your responsibilities.
- Identify barriers to effective data entry procedures and communicate these to the principal or other appropriate personnel.
- Develop and run regular quality assurance checks to monitor data quality.
- Correct errors as they arise.
- Provide feedback to leadership regarding ways to improve data quality.
- Share ideas and best practices with peers.

Outcomes

- Increased organizational efficiency
- Better information for instructional staff leads to better outcomes for students
- Increased accuracy of funding if the data collection is used in funding calculations: for example, attendance data
- The minimization of redundant work required to fix errors
- More accurate representation of the school when data are used for public reporting, such as attendance or behavior

The Role of the Teacher[13]

Responsibility

You are responsible for entering timely and accurate data about your students, as required. Because these data are used to make key decisions that affect your students, your adherence to accuracy and completeness is vital to the data quality process.

Things to Think About

- Who uses the data that you enter? Users might include parents, students, school board members, principals, other teachers, or payroll staff.
- How do you use data to make important individual and group instructional decisions (for example, progress toward the attainment of content standards, or the need for remediation or intervention)?
- What is the effect of the data you enter on students' educational experiences?
- What is the impact of incomplete or inaccurate data?
- What can you do to increase the accuracy of data?
- What expectations are there on timeliness of reporting?
- What training is available to guide proper reporting?

Things to Do

- Ask for and attend training on the use of data in the instructional program. Use online training resources available.
- Ask for appropriate instructions and documentation.
- Ask clarifying questions about your data responsibilities.
- Identify barriers to effective data entry and communicate these to the principal or other appropriate personnel.
- Use only district-approved software.
- Enter data accurately and promptly.
- Check your work for accuracy and completeness.
- Provide feedback to leadership regarding ways to improve data quality.
- Share good ideas and best practices about data entry with your peers.
- Encourage students to use the data you make available to highlight issues earlier when problems are brought to your attention and to improve data quality.

Outcomes

- Your ability to make sound educational decisions about your students will be based on high-quality, accurate data.
- Increased parent and student engagement with the availability of timely classroom data
- More accurate representation of the school when used for public reporting such as attendance or behavior
- Accurate, complete data sets help inform other staff who work with students
- Increased accuracy of funding if the data collection is used in funding calculations such as attendance

13 The suggestions in this tip sheet may apply to others in the school, such as a counselor or nurse.

The Role of the Principal

Responsibility

As the chief instructional leader, you play a key role in ensuring the quality of data collected, reported, and used in the school. Your leadership guides the staff in reporting and using the most accurate data possible.

Things to Think About

- How do you and your staff use data to measure student outcomes?
- How certain are you that the data you are reporting are correct?
- What data are collected in your school?
- Which stakeholders use the data collected at your school?
- How are data from your school incorporated into the larger reporting contexts that guide resources to your location?
- What can you and your staff do to produce quality data?
- Who are your data stewards responsible for validating accuracy of data?
- What training resources are available, and how are these promoted to ensure awareness?
- What are the expected timelines for reporting?
- Are you creating the time and opportunity for staff to manage data effectively?

Things to Do

- Define points of contact and create processes responsible for verifying data reporting in systems periodically before key reporting dates.
- Ensure that all reports are checked for accuracy and reasonableness before "signing off" and sending to the district.
- Periodically spot-check with staff source documents against data entered to ensure that required data (for example, medical information) are being entered.
- Talk with school staff about the usefulness of the data for their needs.
- Ensure that your staff have access to appropriate technology tools.
- Provide staff with training on access to and navigation of technology tools.
- Attend meetings and training about data requirements to ensure up-to-date knowledge.
- Learn and follow all district policies, procedures, and guidelines as they relate to data requirements.
- Provide district data personnel with recommendations for improvements in data collection procedures.
- Allow and encourage staff to attend training in their areas of responsibility.
- Work with your staff and the district to develop and use standard procedures for data entry and reporting.
- Provide trained staff to back up data entry personnel during peak periods such as enrollment or scheduling.
- Ensure that staff use only district-approved software.
- Make sure any data used within your school has been properly vetted by appropriate personnel before use.
- Encourage the use of appropriate data to make daily instructional decisions.
- Provide an environment conducive to accurate data entry.
- Disseminate a calendar for data reporting deadlines.
- Make recommendations to the LEA for additional data that could help staff better inform their practice.

Outcomes

- The creation of a culture of data quality
- Increased quality of instructional and operational decisions
- More accurate representation of the school when used for public reporting such as attendance or behavior
- Increased accuracy of funding if the data collection is used in funding calculations such as attendance

The Role of Technology Support Personnel

Responsibility

You are responsible for managing, maintaining, and securing the hardware, software, and networks that allow staff to enter, store, secure, safeguard, and transfer data. Your role helps to uphold the districts' responsibilities to students and families and to direct data safely to the appropriate places for use in a larger context.

Things to Think About

- Do you have the appropriate hardware and software to allow efficient data entry, storage, and retrieval?
- Are standard data definitions used? Do you adhere to privacy and confidentiality laws and standards?
- Are you using national standards for collection and interoperability?
- How interoperable are the computer applications (that is, able to interact with each other using a minimum amount of programming resources)?
- Do you collaborate with other data professionals within the district?
- Do you have an effective data support process or service such as a helpdesk in place?
- Are the sufficient security infrastructure and tools in place to safeguard education data?
- Who needs access to the data systems, at what levels, and for what reasons?
- When will systems be under the most pressure for data entry or use, and what is your plan for mitigating any potential delays or downtime that users may experience?

Things to Do

- Work with the data coordinator and other staff involved with data entry to develop efficient editing and data verification procedures.
- Work with the data coordinator to provide adequate technical assistance and professional development. Provide a helpdesk and/or online help for data entry staff.
- When selecting computer applications, ensure that applications comply with district data standards and applications can be used in tandem with other district data systems.
- Develop an electronic audit trail so that people can identify potential flaws at each stage of data collection.
- Create and promote validation reports to help data review.

Outcomes

- Improved instructional programs and services
- Optimization of tools to support data collection and reporting
- Improved data security and privacy

The Role of the Data Steward[14] or Coordinator

Responsibility

You ensure that the statistical information reviewed by senior staff is valid and reliable. Furthermore, you enhance the information reporting process through staff development and collaboration with the various offices and programs responsible for producing data and information. Your role includes communicating how the data are ultimately used to help students and helping varied stakeholders understand the purpose of data collections and processes.

Things to Think About

- Do you have sufficient tools and reporting available to validate data quality?
- Does the information provided for review by the superintendent and senior staff represent facts based on accurate data that are thoroughly vetted?
- How well do stakeholders in your school district understand how data are used to benefit the instructional program and provide funds for services?
- How well do staff responsible for specific data collections understand the purpose of each collection?
- Are the staff responsible for entering data provided with ongoing professional development?
- Is there a process in place that allows "end users" to request or modify reports?
- How collaborative can you be with other offices, units, or divisions within the district?
- How do you deal with redundant requests for data?
- What process is in place to resolve discrepancies in information?

Things to Do

- Coordinate the data collection process.
- Provide professional development for staff members to encourage a culture of quality data in the school. The sessions might include
    - demonstrations that incorporate hands-on training;
    - "train-the-trainer" sessions;
    - examples that reflect real-life situations;
    - online resources, handbooks, or guidebooks, with copies of data entry screens, systematic instructions, and the rationale for entering the data;
    - descriptions of the procedures for requesting help (for example, a helpdesk phone number, or an online or e-mail query process); and
    - copies of the reports created from the data, enabling trainees who enter information to have a sense of how their work affects the operation of the school.
- Resolve discrepancies in information before reports are forwarded to senior staff.
- Develop a process that allows staff to request new reports or modifications of existing reports.
- Follow and promote state standards for data collection.
- Before reporting, use existing tools or create new validation tools to help ensure data quality.
- Collaborate with other data professionals within the district to establish authoritative data sources, business rules for data collections, and error checking procedures.

14 **Data stewards** are those individuals within an organization who are responsible for implementing data governance policies and standards and maintaining data quality and security. These staff members do much of the work related to anaging data, perform data checks to ensure accuracy, and compile the appropriate data elements to answer questions and fulfill reporting requirements. (*Forum Guide to Data Governance*; 2020; p. 10).

- Collaborate with the SEA by providing feedback on data collection and reporting timelines, documentation, and supports.

Outcomes

- Improved use of data leading to improvements in student achievement
- Increased services provided to the district and schools
- More accurate representation of the school in public reporting such as such as attendance or behavior
- Increased accuracy of funding if the data collection is used in funding calculations such as attendance

The Role of the Superintendent

Responsibility

As a district superintendent, you are responsible for enhancing the education program, improving student achievement, and seeing that district policies are implemented to promote best practices, including those for data quality.

Things to Think About

- How can you promote a culture that emphasizes data quality and data use?
- How does your district use data to demonstrate achievement in education programs and drive action toward improvements where needed?
- How do you know that the data you review are accurate?
- What data are school staff responsible for entering into the student information system (SIS)?
- Is there an inventory of data collected in your district?
- Do personnel in your district understand the use of data in instructional programs?
- Do personnel in your district understand the use of data in funding programs?
- Are staff responsible for data entry receiving appropriate professional development and support?

Things to Do

- Establish a culture of maximizing the quality and the use of data.
- Set education benchmarks that use data to measure student outcomes.
- Support the development of a culture of quality data in your district through an effective professional development program.
- Encourage principals to make data-driven, building-level decisions.
- Support your information technology director, data stewards, and other staff in the promotion of more efficient data collection procedures, the use of technology to decrease data entry errors, and the movement toward applications that are "interoperable" (that is, applications that interact with each other using a minimum amount of programming resources).
- Support the allocation of funding to provide schools and department with the appropriate resources to collect and manage data.
- Establish a data governance team and assign staff to be data stewards.[15]
- Support the awareness and use of data standards and other data quality best practices.

Outcomes

- Reliable data that are useful for evaluating instructional programs and student achievement
- Increased engagement of students, parents, and staff using data to drive actions
- Effective decision-making
- Creation of a culture of data quality that can ultimately result in a more accurate representation of the district and schools when data are used for public reporting (for example, attendance, behavior)
- Increased accuracy of funding if the data collection is used in funding calculations such as attendance

15 **Data stewards** are those individuals within an organization who are responsible for implementing data governance policies and standards and maintaining data quality and security. These staff members do much of the work related to managing data, perform data checks to ensure accuracy, and compile the appropriate data elements to answer questions and fulfill reporting requirements. (*Forum Guide to Data Governance*; 2020; p. 10).

The Role of the School Board Member

Responsibility

You are responsible for setting policy and using data to inform decisions for and about your schools.

Things to Think About

- What information do you need?
- How does your district use data to demonstrate achievement in educational programs?
- How do you know that the data are accurate?
- What do personnel in a school do to collect and enter data?
- How burdensome are data requests for teachers, school leaders, and other district personnel?

Things to Do

- Ask questions to determine the maturing of the data culture. Is the district maximizing the quality and use of data?
- Set education benchmarks that use data to measure student achievement. Review data to see progress of work.
- Learn about the impact data have on program funding.
- Allocate appropriate resources to enable schools to meet their data collection and reporting mandates.
- Invest in data collection and management software as a routine cost of doing business.
- Request information or workshops when you need a deeper understanding of a topic.
- Review data to determine how they can be used to drive action toward improvements.

Outcomes

- By becoming knowledgeable about the data entry process and the district's procedures for ensuring data quality, you can rely on the information you use with more confidence.
- Investing resources to produce accurate, timely data can result in information that justifies programs or secures additional needed funds.

The Role of State Education Agency (SEA) Staff

Responsibility

You are responsible for creating a foundation of data standards and expectations and establishing ground rules for data collection and reporting.

Things to Think About

- What are the sources of data used in reporting?
- How can you help district staff understand how data are used at the state and federal levels?
- What tools can you provide to help districts adhere to standards?
- How can you minimize data collection and management burdens on district staff?
- What validation tools can be put in place to improve the quality of data?
- How will the SEA use the data?

Things to Do

- Establish communication channels between the SEA and districts to discuss data needs and requirements.
- Announce new collections and changes to existing collections with enough lead time for districts and student information system (SIS) vendors to make plans to collect the data.
- Clearly define business rules used in data reports.
- Provide data collection professional development and documentation for LEA staff members that will help promote data quality, such as
    - examples that reflect real-life situations they will encounter;
    - information on the reason for the data collection and how the data will be used to benefit schools and students, including reports generated from the data;
    - handbooks or user guides with depictions of data entry screens' systematic instructions; and
    - procedures to follow when questions arise, such as a helpdesk phone number or e-mail query process.
- Identify discrepancies in information and provide feedback and correction windows for districts before data and reports are published. Provide audits to data submitters that highlight unusual values (for example, a teacher with an assignment equivalent to 2FTE) or changes over time (such as an unexpected change in enrollment) to aid in the identification and correction of data.
- Collaborate with districts to set data collection and reporting timelines.
- Use data governance to promote cross-agency awareness and agreement before adding a new data collection or making a change to an existing data collection.
- Create tools or edit processes that help improve the quality of data.
- Use data in meaningful ways that show the value of the collection and drive action based on the data.
- Develop and promote use of a calendar for reporting timelines.

Outcomes

- Increased analytic capacity
- Reliable data that are useful for evaluating policies and programs, allocating resources and support, and identifying areas of success and areas where improvements are needed
- Improved data quality for state, federal, and public reporting

- Increased confidence in the information that enables data-informed decision-making
- Strong relationships between the SEA and LEAs with a combined focus on using data to drive action
- Increased stakeholder support for data collections

The Role of the Parent or Guardian

Responsibility

You are responsible for providing current and accurate information about your child to your child's school, including any changes or updates throughout the year.

Things to Think About

- What data about your child are available to you?
- How can data from schools help you make educational decisions for or with your child?
- What information does your school ask for about your child? When does the school need that information?
- Do you know how the school staff use your child's information?
- What reports from the school, district, and state are available for you to review?
- Who is the best person to contact at your child's school or district if you have any questions or need to provide updates to your child's information?
- What is the best way to provide and verify information about your children? Is there an online registration process or another way?
- What is the best way to engage with your child's teacher and school to be aware of available data?
- Where can you find out more information about data use at your child's school or district?

Things to Do

- Ensure you review and update your child's information annually and whenever any changes occur.
- Inform the school if you notice inaccurate information about your child on a school document or report.
- Use data available to engage with teachers or school officials.
- Learn about and make use of tools available to parents to monitor students' attendance, behavior, grades, and more.
- Follow-up with teachers and district staff on incomplete or confusing data reporting to help with quality of data.
- Ask for additional information if you are unclear about what data the school needs about your child.
- Review the school's annual report card and other reports to see data about the school and how data are used.
- Understand your rights under the Family Educational Rights and Privacy Act (FERPA) and make an informed decision about sharing your child's directory information.

Outcomes

- Increased knowledge of how your child's information is used to improve services, outcomes, and instruction
- Increased knowledge of the school overall, populations of students served, education outcomes, and other highlights
- Correct student reporting
- The improved ability for schools to serve your child's unique needs
- Improved student outcomes

Reference List

Federal Committee on Statistical Methodology. (2020). *A Framework for Data Quality*. FCSM 20-04. Federal Committee on Statistical Methodology. September 2020.

National Forum on Education Statistics. (2020). *Forum Guide to Data Governance* (NFES 2020-083). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

National Forum on Education Statistics. (2016). *Forum Guide to Data Visualization: A Resource for Education Agencies*. (NFES 2017-016). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

National Forum on Education Statistics, Exit Codes Working Group. (2020). *Forum Guide to Exit Codes*. U.S. Department of Education. Washington, DC: National Center for Education Statistics.

National Forum on Education Statistics. (2021). *Forum Guide to Metadata* (NFES 2021110). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

National Forum on Education Statistics. (2018). *Forum Guide to Reporting Civil Rights Data* (NFES 2017-168). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

National Forum on Education Statistics. (2023). *Forum Guide to State Education Agency Support for Local Education Agencies in Civil Rights Data Reporting* (NFES 2023-026). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

National Forum on Education Statistics. (2021). *Forum Guide to Strategies for Education Data Collection and Reporting* (NFES 2021013). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Spiegelman, M., & Merlin, J. (2023, August 30). *Using Federal Education Data to Inform Policymaking: Part 1-Benefits and Advantages*. NCES Blog. U.S. Department of Education, National Center for Education Statistics. Retrieved from https://nces.ed.gov/blogs/nces/.

Spiegelman, M., & Merlin, J. (2023, August 30). *Using Federal Education Data to Inform Policymaking: Part 2-Challenges and Opportunities*. NCES Blog. U.S. Department of Education, National Center for Education Statistics. Retrieved from https://nces.ed.gov/blogs/nces/.

U.S. Department of Education. (2020). *Data Strategy*. Retrieved September 1, 2023, from https://www.ed.gov/sites/default/files/cdo/ed-data-strategy.pdf.

U.S. Department of Education. (2023). *Data Strategy*. Retrieved November 8, 2023, from https://www2.ed.gov/about/offices/list/opepd/ocdo/ed-data-strategy.pdf.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2021) Statewide Longitudinal Data System Webinar Summary: *Improving Data Quality: State Strategies to Enable Data Use*. Retrieved September 1, 2023, from https://slds.ed.gov/#communities/pdc/documents/20386.

U.S. Department of Education, National Center for Education Statistics, Common Education Data Standards. *CEDS Elements v. 10*. Retrieved August 28, 2023, from https://ceds.ed.gov/elements.aspx?v=10.

Related Resources

National Forum on Education Statistics Resources

Forum Guide to State Education Agency Support for Local Education Agencies in Civil Rights Data Reporting (2023)

https://nces.ed.gov/forum/pub_2023026.asp

This resource presents a variety of effective methods through which state education agencies (SEAs) can support their local education agencies (LEAs) in reporting civil rights data to the U.S. Department of Education's (ED's) Office for Civil Rights (OCR). In addition, the resource provides several detailed case studies from states that currently support their LEA reporting.

Forum Guide to Metadata (2021)

https://nces.ed.gov/forum/pub_2021110.asp

This resource presents and examines the ways in which metadata can be used by education agencies to improve data quality and promote a better understanding of education data. The resource highlights the uses of metadata, discusses how to plan and successfully implement a metadata system in an education setting, and provides examples of standard metadata items and definitions to assist agencies with standardization.

Forum Guide to Strategies for Education Data Collection and Reporting (SEDCAR) (2021)

https://nces.ed.gov/forum/pub_2021013.asp

This resource was created to provide timely and useful best practices for education agencies that are interested in designing and implementing a strategy for data collection and reporting, focusing on these as key elements of the larger data process. This new resource is designed to be relevant to the SEAs and LEAs of today, in which data are regularly collected for multiple purposes, and data collection and recording may be conducted by many different individuals within an agency.

Forum Guide to Exit Codes (2020)

https://nces.ed.gov/forum/pub_2020132.asp

This resource provides best practice information for tracking data about when students transferred, completed high school, dropped out, or otherwise exited an education agency. The resource defines exit codes and reviews their use in an education agency; provides an updated, voluntary, common taxonomy for exit codes; discusses best practices and methods for addressing specific challenges in exit codes data collection; and features case studies that highlight different education agencies' approaches to and experiences with exit coding.

Forum Guide to Data Governance (2020)

https://nces.ed.gov/forum/pub_2020083.asp

This resource provides timely and useful best practices, examples, and resources for agencies implementing or updating their data governance programs. It provides an overview of data governance; discusses effective data governance practices, structures, and essential elements; describes how to meet privacy and security requirements while also meeting data accessibility and sharing needs; and includes detailed case studies from education agencies in their data governance efforts.

Forum Guide to Planning for, Collecting, and Managing Data About Students Displaced by a Crisis (2019)

https://nces.ed.gov/forum/pub_2019163.asp

This resource provides timely and useful best practice information for collecting and managing data about students who have enrolled in another district or school because of a crisis. It highlights best practices that education agencies can adopt before, during, and after a crisis and features contributions from agencies that have either experienced a crisis or received students who were displaced by a crisis.

Forum Guide to Reporting Civil Rights Data (2018)

https://nces.ed.gov/forum/pub_2017168.asp

This resource presents a variety of effective methods through which LEAs report civil rights data to the U.S. Department of Education's Office for Civil Rights. In addition, the guide provides examples of how SEAs can voluntarily help their LEAs with Civil Rights Data Collection (CRDC) reporting. The guide includes an overview of the CRDC, a discussion of the challenges and opportunities in reporting civil rights data, an explanation of the CRDC reporting process, and six case studies that examine how specific education agencies report civil rights data.

Forum Guide to Data Visualization: A Resource for Education Agencies (2016)

https://nces.ed.gov/forum/pub_2017016.asp

This resource recommends data visualization practices that will help education agencies communicate data meaning in visual formats that are accessible, accurate, and actionable for a wide range of education stakeholders. Although this resource is designed for staff in education agencies, many of the visualization principles apply to other fields as well.

National Center for Education Statistics (NCES) Resources

A Framework for Data Quality

https://nces.ed.gov/fcsm/pdf/FCSM.20.04_A_Framework_for_Data_Quality.pdf

This document provides policy and procedural guidance to federal agencies for ensuring and maximizing the quality, objectivity, utility, and integrity of information (including statistical information) disseminated by federal agencies.

ED*Facts* Business Rules in an SLDS Ecosystem

https://slds.ed.gov/#communities/pdc/documents/19986

Improving Data Quality: State Strategies to Enable Data Use

https://slds.ed.gov/#communities/pdc/documents/20386

Processes for Handling Multiple IDs to Ensure Data Quality

https://slds.ed.gov/#communities/pdc/documents/21203

Data Quality: Striking a Balance

https://slds.ed.gov/#communities/pdc/documents/20199

This Statewide Longitudinal Data Systems (SLDS) Grant Program issue brief provides strategies to improve data quality and data completeness.

Using Federal Education Data to Inform Policymaking: Part 1–Benefits and Advantages

https://nces.ed.gov/blogs/nces/post/using-federal-education-data-to-inform-policymaking-part-1-benefits-and-advantages

Using Federal Education Data to Inform Policymaking: Part 2–Challenges and Opportunities

https://nces.ed.gov/blogs/nces/post/using-federal-education-data-to-inform-policymaking-part-2-challenges-and-opportunities

Other Federal Resources

Data Quality Dimensions

https://ies.ed.gov/ncee/edlabs/regions/central/resources/pemtoolkit/pdf/module-5/CE5.3.2-Data-Quality-Dimensions.pdf

This table provides definitions and examples of the important components of quality data.

Data Quality Improvements from States Using Generate

https://eric.ed.gov/?q=data+quality&id=ED612037

This brief describes 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.

ED*Facts* Partner Support Center https://edfacts.communities.ed.gov/

Civil Rights Data Collection (CRDC) Partner Support Center https://crdc.communities.ed.gov

These resources provide technical assistance to SEAs around their respective federal data collections.

Regional Educational Laboratory (REL) Program Evaluation Toolkit

https://ies.ed.gov/ncee/edlabs/regions/central/resources/pemtoolkit/module5-ch1.asp

Module 5 provides an overview of data quality considerations. The module also covers the alignment of data to evaluation questions. The module contains three chapters that will help evaluators do the following:

- Chapter 1: Identify the two major types of data and describe how to use them in an evaluation.
- Chapter 2: Evaluate the quality of data, using six key criteria.
- Chapter 3: Connect data to your evaluation questions.

School Pulse Panel

https://ies.ed.gov/schoolsurvey/spp/

The School Pulse Panel is a study collecting information on the impact of the COVID-19 pandemic from a national sample of elementary, middle, high, and combined-grade public schools. Some survey questions are asked repeatedly to observe trends over time while others examine unique topics in a single month.

States Realizing Data Quality Improvements Through Data Integration

https://files.eric.ed.gov/fulltext/ED612036.pdf

This brief describes how 11 states worked with the Center for the Integration of IDEA Data (CIID) to integrate their Individuals with Disabilities Education Act (IDEA) data and to implement Generate and realized notable data quality improvements.