DATA QUALITY CAMPAIGN
Using Data To Improve Student Achievement

SUPPORTING EDUCATION POLICY AND PRACTICE THROUGH COMMON DATA STANDARDS

A Policymaker's Guide

April 2012

Consider the high-priority challenges facing education stakeholders today, such as measuring teacher effectiveness, implementing the Common Core State Standards, aligning K–12 and postsecondary education efforts with workforce demands, efficiently allocating resources, and ensuring that students stay on track to success. These efforts all rely on the efficient collection, management, and use of education data—and stakeholder confidence in those data, particularly if they will be used for high-stakes personnel, accountability, or other decisions.

Over the last decade, state, federal, and private investments of political will and resources have contributed to significant progress in building robust statewide longitudinal data systems. Attention is now shifting—as it should be—to the work of ensuring that data are used by stakeholders to answer critical questions and inform decisionmaking from the kitchen table to the classroom to the state capitol. In the absence of common data standards that ensure the quality, comparability, and efficient sharing of data, these efforts' impact will be limited and their implementation costly and ineffective.

Conditions for Useful Data: Quality, Comparability, and Efficiency in Sharing

Policymakers have a responsibility to ensure that education data meet three conditions to ensure their usefulness:

1. Quality
2. Comparability
3. The ability to be shared efficiently

1. Data are high quality regardless of where, how, and by whom they were collected, inputted, and stored.

Education data are collected, entered, and stored by a wide variety of individuals with different levels of technical expertise, from the school secretary to teachers to data systems

administrators. This is true in every sector, and an oft-cited information technology maxim speaks to this challenge—"garbage in, garbage out."

Education leaders and policymakers have a responsibility to the stakeholders using data to do all that they can to ensure the quality of education data.

2. Data from different systems can be compared, allowing users to draw valid conclusions.

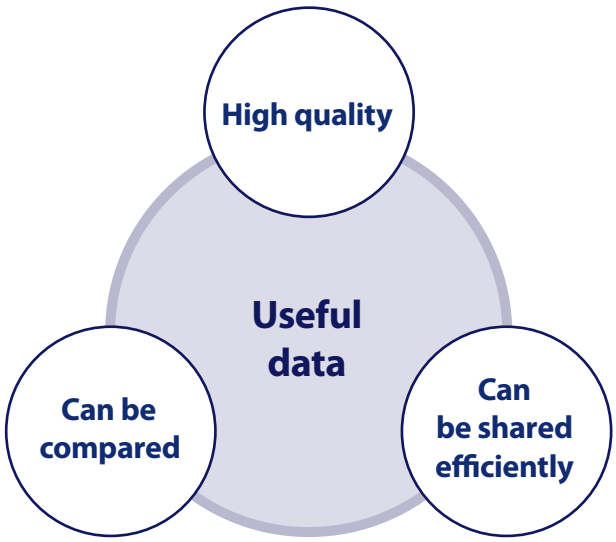
Schools, districts, programs, and states are constantly compared using performance indicators. Parents look at data about schools' test scores and graduation rates when deciding where to move and buy a home or at retention and job placement rates when helping their children select a higher education institution. Schools and districts are often viewed in relationship to their peers as part of state and federal accountability systems in decisions made about incentives, consequences, and support. And states' performance outcomes are frequently ranked on indicators that show up in national reports, the front pages of newspapers, and political debates about policymakers' performance.

Such comparisons will be more visible and more meaningful and will carry more weight due to the implementation of the Common Core State Standards and common assessments. The data used to produce these indicators often originate in different data systems. To draw valid conclusions, the indicators must be produced with data that are comparable. *Education leaders and policymakers have a responsibility to ensure data comparability.*

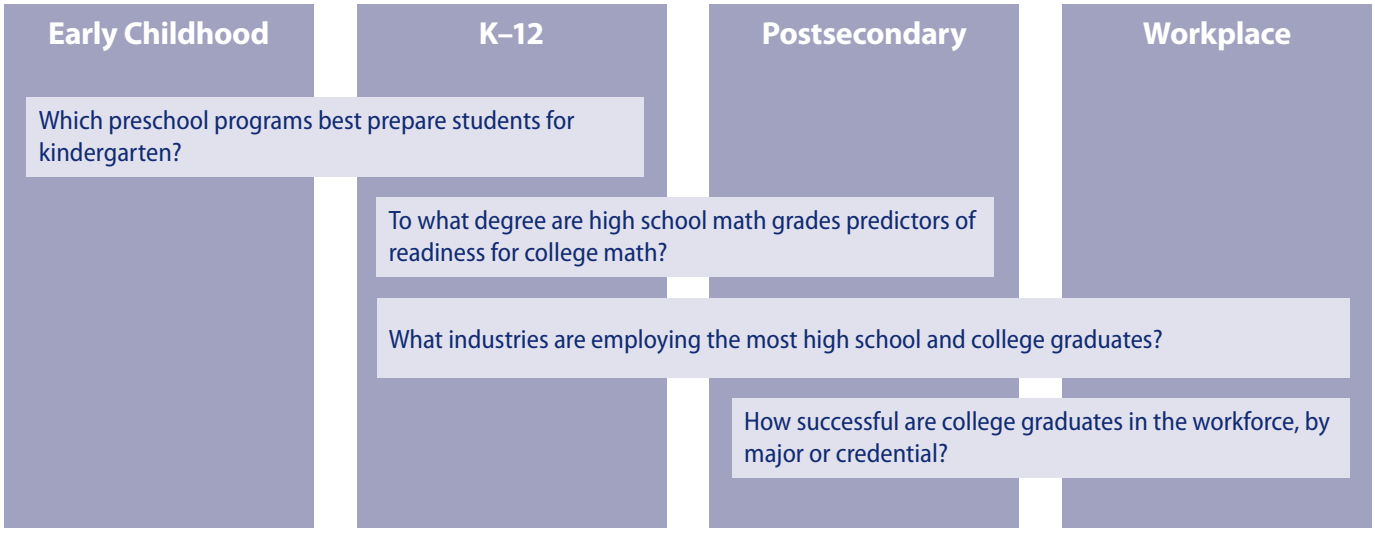
3. Appropriate data can be shared efficiently when necessary.

Increasingly, answering critical questions requires appropriate data to flow efficiently and effectively across traditional silos, including agency sectors, and states. Data must be exchanged within multiple levels of the education sector, such as for state data collections from districts and schools. There is also growing demand for coordinating integrated services across agencies, such as through the Promise Neighborhoods initiatives, which requires sharing data across agencies and with nontraditional partners. And the renewed national focus on the goal of graduating all students from high school ready for college and careers means that—more than ever before—the important questions facing the education sector are ones that require data from multiple sources that often span traditional boundaries. See the graphic on the next page for examples.

Creating the Conditions for Useful Data



Important Questions Often Require Data from Multiple Data Systems



Moreover, students are mobile—moving between grades and schools and across district and state lines. Some of those transitions are expected, as students move through their education career and enroll in various schools and participate in various programs. However, some of this mobility is unexpected: the U.S. Census found that, between 2010 and 2011, 1.8 million U.S. families with school-aged children moved, including 187,000 families that moved across state lines.[1] Schools, districts, and states need to be able to efficiently transfer student records to ensure a seamless transition for students and limit interruptions to the delivery of instruction and services. Moreover, the data from the various data systems must be able to be aggregated to produce indicators like those described above.

Education leaders and policymakers have a responsibility, particularly in this era of reduced resources, to ensure that data systems are positioned to exchange data as efficiently as possible, reducing the burden on stakeholders at all levels.

Standardization Supports Data Quality, Comparability, and Efficiency in Sharing

Unfortunately, these basic conditions for data use—quality, comparability, and the ability to be shared efficiently—are at risk when data are not standardized.

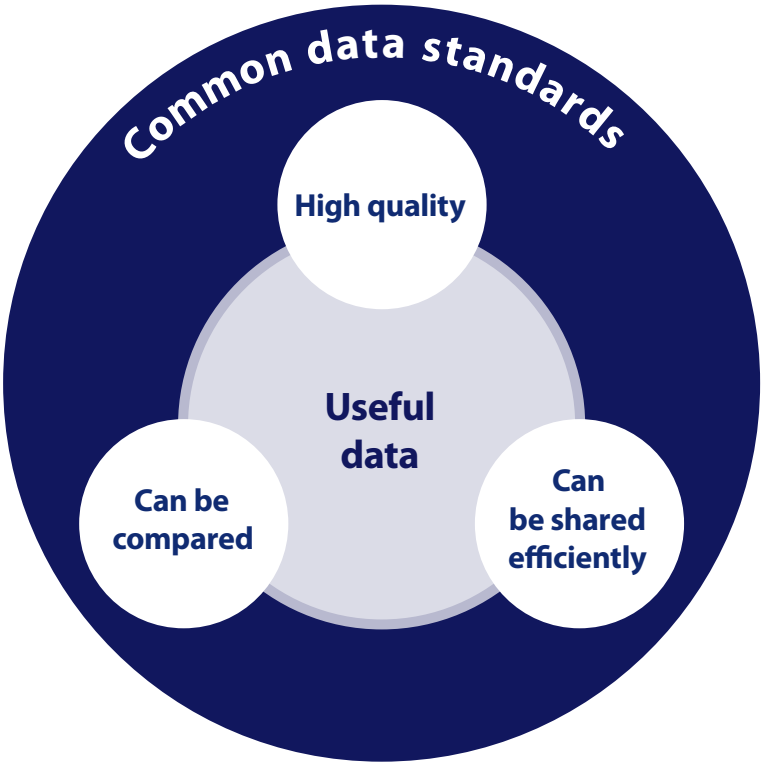
Consider how humans communicate. Our languages make use of sounds that represent concepts; symbols that represent sounds; and rules that govern the meaning and structure of alphabets, words, sentences, and paragraphs. When two people from different countries try to communicate, the lack of common language can serve as a barrier to quality and efficient communication. They can exchange words, but until they can find a way to

interpret what each other is saying, they are unable to use the information or gain anything from their interaction. Even the right words, in the wrong order or with some extra words, might cause a communication problem. For example, if an American tourist orders coffee in a non-English-speaking country, he might be able to successfully request his order through significant energy, time, and frustration, or he might end up with a glass of iced tea.

Data systems trying to communicate are no different. Unfortunately, the education data systems in place across the country today too often speak different languages and dialects and lack the ability to efficiently and effectively translate. These various data systems were often built in silos, in response to local or state needs or to comply with state or federal data collections. The data standards used by these systems are drawn from a variety of sources, including definitions found in state or federal legislation, standards used by the mandated data collections, or the design of a piece of software purchased from a vendor. As a result, the education data sector is characterized by innumerable sets of data standards that overlap and are inconsistent.

These technical challenges undermine the ability to get the right data to the right people at the right time. And in today's education and economic environment, that is no longer acceptable. Tremendous amounts of money are invested each year in designing and implementing sophisticated data systems that do not communicate with each other, wasting resources and missing the opportunity to maximize the potential of these investments to improve decisionmaking and student success.

Common Data Standards Create the Conditions for Useful Data



Without common data standards ...

- Data quality is reduced because of the risk of error in entering, translating, and migrating data.
- Resources are spent duplicating data entries, translating and migrating data across systems, and producing data in multiple formats for various recipients that use different standards.
- Necessary data sharing across systems is delayed because of the extra time needed to translate and migrate data.
- Systems are unable to link data because they cannot spend the resources to conduct the necessary translations or migrations.
- Vendors are forced to tailor products to each system or state's specifications, increasing time and costs and inhibiting the development of new tools and services.

With widespread use of common data standards ...

- Data quality is improved because the individuals inputting data have clear guidelines around data formats, and error is reduced due to the limited need for translating, reentering, or recoding data.
- Sharing data across systems is efficient, seamless, and timely.
- Data collection and reporting burden is reduced.
- Limited resources are spent on innovative solutions, tools, and services.
- Users begin to rely on timely, accurate, easy-to-understand data.
- Useful research and analysis can be updated and replicated more efficiently.

The Way Forward: Common Data Standards

It is time for the education sector to standardize education data. When data are standardized, their quality is improved. When standardized data are shared, they have the same meaning on both ends of the exchange. Then, indicators produced with those data—such as high school graduation rates—would be reliable, meaningful, and comparable. Standardized data make transferring student data from one school or level of education to another easier, and they permit policymakers to learn how students fare as they move across institutions, state lines, and school levels—making the education sector more effective and efficient. In turn, this will facilitate smarter educational decisionmaking for all stakeholders—policymakers, administrators, even parents—and fuel an information-driven education sector focused on improving student outcomes.

So what are standardized data? They are data that conform to detailed standards for data element names, definitions, and formats. To ensure accuracy, these standards are as specific as possible, providing, for example, the exact name of each data element; its exact allowable length; whether it can be letters only, numbers only, or alphanumeric; whether or not special characters can be included; etc.

The Common Education Data Standards Initiative

As is the case in many industries and sectors, efforts to standardize education data have evolved over several decades through a variety of initiatives and entities, including the Schools Interoperability Framework (SIF) Association,[2] Postsecondary Electronic Standards Council (PESC),[3] National Center for Education Statistics (NCES) Handbooks,[4] National Education Data Model (NEDM),[5] and Ed-Fi™. (See box on next page for more about Ed-Fi.) The time has come for alignment across these efforts to create a central set of commonly agreed-upon fundamental standards so that states; districts; vendors; higher education entities; and other entities that collect, store, and share data use common standards that result in truly standardized education data.

State, national, and federal leaders have recently come together through the Common Education Data Standards (CEDS) Initiative[6] to collaborate on that very goal: accelerating the widespread adoption and use of common data standards and shifting attention to the use of data by stakeholders. For a core set of data elements, CEDS has identified agreed-upon standard definitions, code sets, business rules, and technical specifications.

Policymakers should know the following five things about CEDS:

- **Aligned with existing sources:** CEDS is not being developed from scratch. CEDS draws on existing sources, including those mentioned above, to offer a central set of common education data standards. Existing sources are working to align with CEDS as it moves forward. For example, both the SIF Association and PESC have established a process for completely adopting, or aligning future versions of their standards with, future versions of CEDS.[7] Also, the U.S. Department of Education will use CEDS to inform future federal data collections and is moving to consolidate other standards-related projects, such as the NCES Handbooks and NEDM, with CEDS.[8]
- **Voluntary:** CEDS is voluntary. It is the responsibility of entities that collect and store data or that offer data standards—including state and local agencies, higher education entities, and vendors—to determine if and how to adopt and implement CEDS.
- **Iterative:** CEDS is an iterative process that continues to expand to include new elements. In January 2012, version 2.0 was released, which included a standard for elements across the K–12 and postsecondary systems.
- **Includes implementation tools:** CEDS provides an increasing number of tools that help data system leaders implement the standards, including a searchable element database, a logical data model, and the web-based CEDS Alignment Tool (http://ceds.ed.gov/alignmenttool) that allows stakeholders to import or input their organizations' data dictionaries and compare them to CEDS or other participating entities.
- **Not a data collection:** No data are actually collected through the initiative, nor does it require or mandate any data to be collected. Think of it as a voluntary dictionary that can be used if policymakers decide to speak about data issues.

For more information, see the CEDS website at www.ceds.ed.gov.

CEDS Provides the Data Foundation for Marketplace Tools for Educators

As the education marketplace develops new education tools designed to help educators and policymakers put data into action, CEDS can provide a foundation for their efforts to help ensure that the data behind the tools are high quality and able to be managed and shared efficiently. For example, the following efforts, which are led by prominent education foundations, have aligned with CEDS:

- **Ed-Fi™** is an open and XML-based solution to integrate information from a broad range of existing sources so it can be sifted, analyzed, and put to use by educators every day. It was developed with funding from the Michael & Susan Dell Foundation, and future technical developments will be guided by the foundation and the Ed-Fi Advisory Council, a group of state education agency representatives from states that are implementing Ed-Fi. Learn more at www.ed-fi.org.
- **The Shared Learning Collaborative (SLC)** is designing a shared technology infrastructure that will support the implementation of the Common Core State Standards and help states and districts provide teachers with the instructional data and tools they need. Instructional data will be linked to high-quality and diverse sets of curricular resources, so each student gets what he or she needs most at that moment in time. The Council of Chief State School Officers is coordinating the multistate alliance, which has initial funding from the Bill & Melinda Gates Foundation and Carnegie Corporation of New York. Learn more at www.slcedu.org.

Policymaker Actions to Support the Move to Standardize Education Data

When the tourist ordering coffee in a foreign land has trouble communicating, using hand signals, getting help from a friendly person in line, or thumbing through the translation book might be good alternatives. However, when the stakes are high—say, if the tourist lands in a hospital during his trip—the inability to communicate effectively and efficiently could be catastrophic.

The time is now to standardize education data. The results of *Data for Action 2011: DQC's State Analysis* demonstrate that states are poised to embrace a culture of using data for continuous improvement. States are collaborating on many high-stakes initiatives—such as measuring teacher effectiveness, implementing the Common Core State Standards, aligning K–12 and postsecondary education efforts with workforce demands, efficiently allocating resources, and ensuring that students stay on track to success—that require the use of shared or comparable data. The same political will, energy, and resources that coalesced to build robust longitudinal data systems over the past five years must now be harnessed to put into place practices and policies that will ensure that these data are high quality, comparable, and able to be shared efficiently.

To ensure that statewide data systems are interoperable and useful, state policymakers should facilitate conversations and action within their states to:

- **Learn** about the status of standardization in the state or system, current costs, challenges, and opportunities for moving forward. Examples:
  - Request that the state education agency submit its data dictionaries to the CEDS Alignment Tool and report back to policymakers on existing gaps.
  - Request information from state education agencies about the current extent of standardization; the costs of not standardizing data; and the costs, challenges, and opportunities for moving forward.

Many organizations, including DQC, have voiced their support of the CEDS Initiative to foster the collection and use of clear, consistent data through the use of common data standards. To view the list of supporters and add your voice, visit the CEDS website: http://ceds.ed.gov/voicesOfSupport.aspx.

- **Articulate** data standardization as an important goal for the state's education and related data systems. Make clear to stakeholders within the state that policymakers believe CEDS is vital to optimizing the quality, comparability, and efficiency of data. Examples:
  - Secure public commitment from state education leaders—including state board members; state education agency leadership; and postsecondary institution, system, and coordinating body leaders—to adopt and implement CEDS.
  - Communicate policymakers' commitment to standardization through a press release, executive order, legislation, federal grant application, statement, or hearing.
  - Ensure that vendors contracted through state-supported contracts use common data standards where applicable and appropriate.
- **Encourage collaboration** among critical players to find appropriate, effective, and efficient solutions. Examples:
  - Convene relevant stakeholders, such as K–12 and postsecondary education leaders, to discuss their progress in implementing CEDS, including engagement of local stakeholders, and to report on the extent of alignment between their data models and CEDS.
  - Ensure that work for K–12 includes engaging local school districts and regional entities and that postsecondary institutions, systems, and coordinating

bodies engage community colleges, P–20 councils, and public and private colleges and universities.

- Task a working group, such as the state's P–20 entity or P–20 data governance body, with the responsibility of ensuring that relevant state entities, such as department of education and higher education, are effectively using CEDS.
- Encourage state entities to use CEDS tools, such as the Alignment Tool, to help implement CEDS and inform P–20 data efforts.

Endnotes

[1] U.S. Bureau of the Census. *Current Population Survey, 2011 Annual Social and Economic Supplement. Internet Release, State-to-State Migration Flows (ACS), Table 9. General Mobility of Family Householders, by Type of Household, Race and Hispanic Origin of Householder, and Presence and Age of Own Children Under 18: 2010 to 2011.* Washington, DC, November 2011.

[2] Schools Interoperability Framework (SIF) Association: This community of school, regional, and federal education agencies; marketplace providers; and other organizations across three countries has developed the SIF Implementation Specifications—a set of platform-independent, vendor-neutral rules and definitions that enable software programs from different providers to share P–12 information.

[3] Postsecondary Electronic Standards Council (PESC): This association of colleges and universities; college and university systems; professional and commercial organizations; data, software, and service providers; nonprofit organizations and associations; and state and federal government agencies focuses on interoperability by enabling cost-effective connectivity between data systems to accelerate performance and service, simplify data access and research, and improve data quality along the education lifecycle.

[4] National Center for Education Statistics (NCES) Handbooks Online: The NCES Handbooks Online offer K–12 data definitions for a broad set of data elements. They are intended to serve as a reference guide for districts and states to adopt, adapt, or ignore but are in no way authoritative or directive.

[5] National Education Data Model (NEDM): NEDM is a searchable online catalogue of the data that already exist in the PK–12 education space. NEDM describes how the data in these various initiatives relate to each other as a reference tool to facilitate the identification, merging, and matching of data across different systems. It is similar to a taxonomy of the landscape.

[6] The National Center for Education Statistics works with a group of stakeholders to develop the standards. The CEDS consortium is responsible for advocacy, communications, adoption, and implementation of the standards. It is led by the Council of Chief State School Officers and State Higher Education Executive Officers and includes the Data Quality Campaign, the SIF Association, and PESC.

[7] SIF Association. *CEDS and SIF: From Architecture to Solution, SIF 3.0 to Support the Common Education Data Standards.* www.sifassociation.org/us/upload/news/61AF76_CEDS_and_SIF_From_Architecture_to_Solution.pdf

[8] http://ceds.ed.gov/FAQ.aspx

DATA QUALITY CAMPAIGN
Using Data To Improve Student Achievement
www.DataQualityCampaign.org

To download DQC resources, visit www.DataQualityCampaign.org and follow us on Twitter (@EdDataCampaign).

The **Data Quality Campaign (DQC)** is a national, collaborative initiative to encourage and support state policymakers' efforts to improve the availability and use of high-quality education data to improve student achievement. The campaign provides tools and resources that will help states implement and use longitudinal data systems, while providing a national forum for reducing duplication of effort and promoting greater coordination and consensus among the organizations focused on improving data quality, access, and use.