CENTER *for the* INTEGRATION *of* IDEA DATA

# The Use of Common Education Data Standards to Support Data Integration

**January 2016** • **Authors: Andrea Hall and Bill Huennekens**

## Data Integration

Quality data allow programs to provide context and meaning to research, evaluation, and accountability efforts and provides evidence to evaluate and demonstrate the effectiveness of programs. Data integration, the process of taking data from disparate systems, combining them, and transforming them into something meaningful, can increase data quality.[1] However, the process of integrating data systems is not without its challenges. By starting small and taking advantage of tools that are available, it is manageable. Data integration is a multi-step process with many tasks in each step.

This brief will provide state education agencies (SEAs) with a process for using the Common Education Data Standards (CEDS) tools to identify what data they collect and where the data are stored, conduct a map analysis, and ensure that elements appearing across the data systems are defined the same. This process is a critical step for integrating disparate data systems and realizing the data quality benefits of integrated data.

## Why CEDS

CEDS (https://ceds.ed.gov) is a shared vocabulary for a key set of education data elements. It was developed with educational stakeholders from across the P-20W spectrum (early learning through postsecondary and workforce). CEDS is not a data collection system; rather, CEDS establishes common definitions and option sets (or permitted values) for elements. While CEDS is widely known as a way for SEAs and other entities to compare and share information, it is also a powerful resource for SEAs that want to compare data across their own systems.

While CEDS is widely known as a way for SEAs and other entities to compare and share information, it is also a powerful resource for SEAs that want to compare data across their own systems.

CEDS includes various tools to assist SEAs with data integration that is crucial to maintaining and expanding their statewide longitudinal data systems (SLDS).

[1] *CIID Benefits of Integrated Data System* (https://ciidta.grads360.org/#communities/pdc/documents/7641s)

IDEAs that Work U.S. Office of Special Education Programs · aem · TA&D NETWORK · Department of Education

CEDS' tool suite includes Align[2], Connect[3], and myConnect[4]. Much of this brief will focus on Align as the reports used with that tool are an excellent resource for documenting, comparing and analyzing the various data in their SEA's systems, which is a necessary component of data integration.

## Getting Started With Data Integration

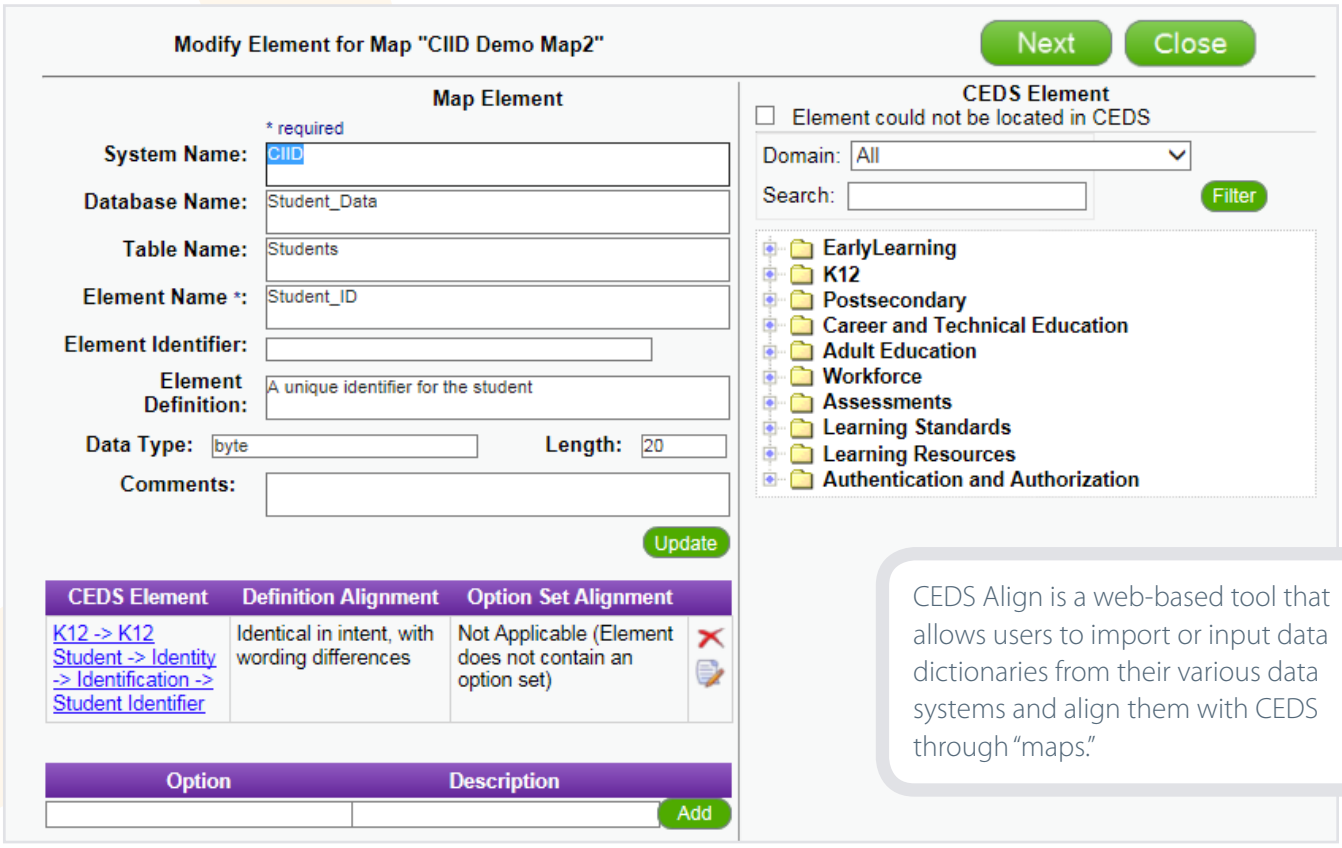
Identifying what data are collected and where they are stored is critical for data integration. SEAs must first develop an exhaustive list of data systems that will be integrated. They then need to review what data are collected in each system and determine if there are any commonalities.

Identifying what data are collected and where they are stored is critical for data integration.

CEDS Align is a web-based tool that allows users to import or input data dictionaries from their various data systems and align them with CEDS through "maps." This alignment serves as the bridge that allows the user to look across data systems to see commonalities and discrepancies.

Align is similar to a metadata[5] management or enterprise architecture[6] tool in that the map stores information such as data sources, data elements, definitions, data types, and option sets (see Figure 1). Maps are private and available only to assigned users until the Map Administrator (map's owner) releases it for public viewing.

Figure 1. Align (https://ceds.ed.gov/learnCedsAlignment.aspx)




CEDS Align is a web-based tool that allows users to import or input data dictionaries from their various data systems and align them with CEDS through "maps."

[2] Align is a tool that uses CEDS as a bridge for comparing across multiple data systems. See https://ceds.ed.gov/align.aspx.

[3] Connect is a tool to document reporting requirements such as EDFacts at the element level. See https://ceds.ed.gov/connect.aspx.

[4] myConnect is a tool that allows users to join an Aligned data system with Connections to explore how the data use may be applied in their system. See https://ceds.ed.gov/myconnect.aspx.

[5] Metadata is data about data, learn more at: https://en.wikipedia.org/wiki/Metadata and https://nces.ed.gov/forum/pub_2009805.asp

[6] Learn more about enterprise architecture at: https://en.wikipedia.org/wiki/Enterprise_architecture

Data dictionaries can be uploaded to CEDS as a single map or as separate maps, and there are benefits to both options. In choosing between options, SEAs should consider the purpose of the mapping process. When multiple data sources are going to be mapped, an SEA needs to determine if the desired result is (1) a consolidated list of all data elements or (2) a list of common elements across data sources. If the former is the goal, a single map should be created with naming conventions determining the data source. If the latter is the goal, multiple maps (one for each data source) should be created so reports across maps can be generated. The reporting tools in CEDS work with either approach. For example, if an SEA loads a single map with multiple data sources, the *Data Dictionary* report will display a consolidated list of data elements within that map by system, database and table. Alternatively, if an SEA loads multiple maps, the same *Data Dictionary* report can be run, but will display side-by-side comparisons of the elements across the maps, identifying those that are common in each data source.

## Data Map Analysis

Once the data systems have been identified, and the data collected across those systems have been mapped in the CEDS Align tool, the SEA then moves on to the map analysis. The purpose of the map analysis is to identify data for consolidation and eliminate redundancies.

CEDS contains numerous map- and element-based reports to assist in the map analysis, each designed to assist users with understanding and managing their data (see Figure 2). The *Data Dictionary Only* report is the most instrumental in assisting SEAs with their data integration efforts. This report shows a list of all elements in the selected maps, side by side, and reveals how the elements line up across each map, regardless of whether the elements are from data systems in the same educational domain or not. This report can easily be modified to also show the CEDS elements to which the maps were aligned either by selecting the "Show CEDS Alignment" under Options in the *Data Dictionary Only* report or by selecting the *Data Dictionary + CEDS Info* report.

As an example, an SEA has two maps in CEDS (CIID Demo Map 1 and CIID Demo Map 2). Both maps are selected when the *Data Dictionary Only* report is run. Any element that exists in both of the selected maps will display side by side for the user to review and analyze (see Figure 3). This is an efficient way to identify where data are being collected and maintained in multiple systems. Additionally, CEDS allows users to drill down on each element to review additional information about the data, such as definitions, data types, and option sets.

Figure 2. CEDS Reports

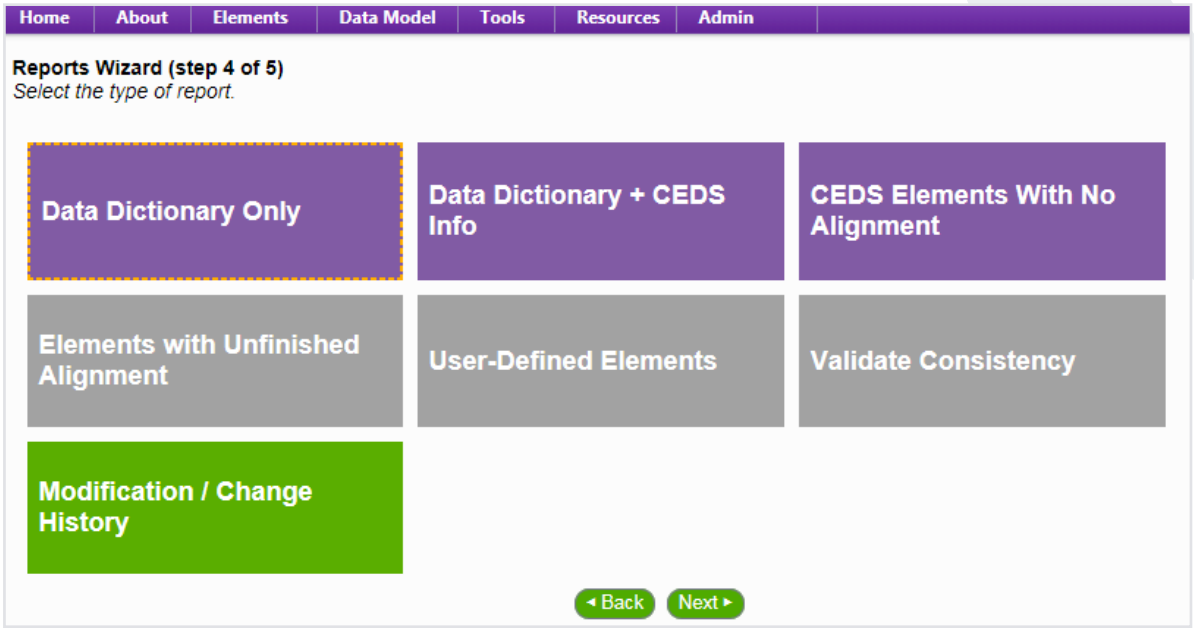
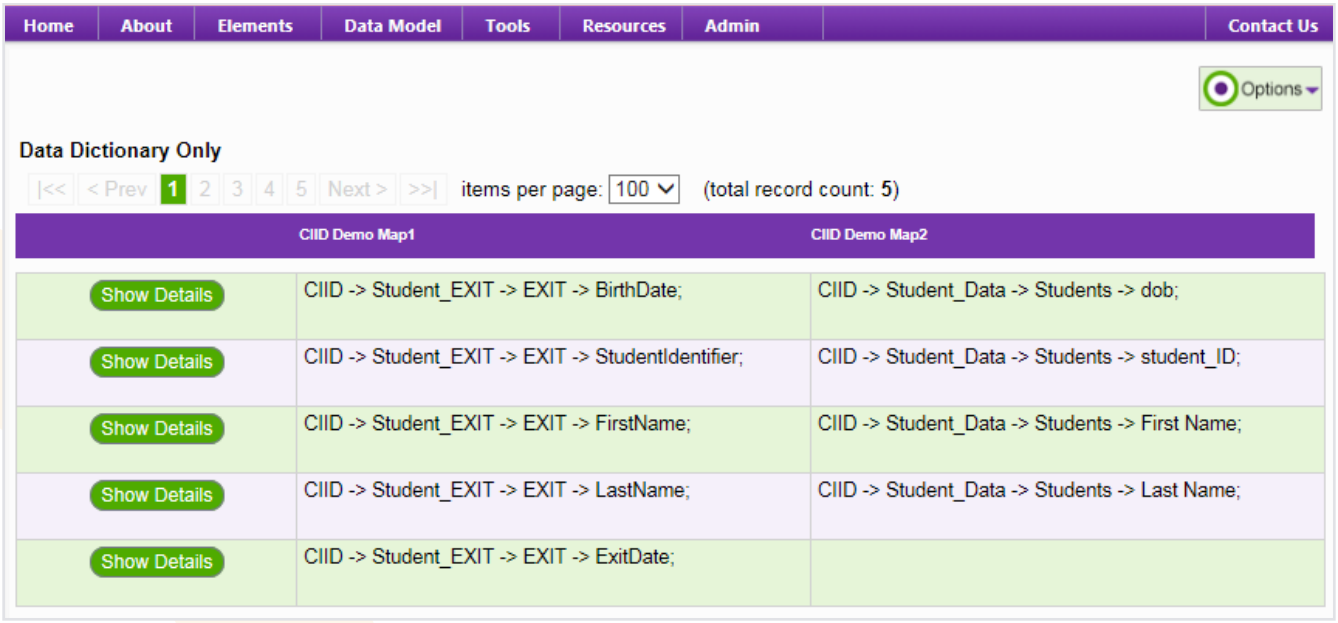


Figure 3. Data Dictionary Only report




## Resolving Differences Among Common Elements

A key step in the data integration process is ensuring that elements common across the data systems are defined the same way so the data can be combined. Once all common elements have been identified via the *Data Dictionary Only* report, the SEA can then use the information in Align to determine if the element definitions and option sets are comparable. If the SEA's definitions and option sets are not comparable, the SEA can review the definition and option sets available in the related CEDS elements. Several SEAs have reported either adopting the CEDS definition and option sets for specific elements in their entirety or using the CEDS definition and option sets as a starting point for discussions within the SEA of suitable substitutes.

### CEDS in Action

Data integration is an ongoing and evolving process. Reporting requirements can change, elements can be discontinued, others may be updated, and new elements might be introduced. Even SEAs with integrated SLDS systems, such as Kansas, must continually monitor and update their data systems as things change. To address this, the Kansas State Department of Education integrated CEDS into its daily work by including a routine comparison of the SEA's data elements with the CEDS standard. If a new element was being introduced to a system within the SEA's SLDS, a review was done in CEDS to see if there was a matching element. If there was, and if the definition aligned with the intent of the element being introduced, then the CEDS version was adopted. Additionally, when new versions of the CEDS standards opened for public comment, Kansas held meetings with data stewards and program leadership to review the updates. Not only did this give key stakeholders an opportunity to provide input into the CEDS standard, it fostered a continued collaboration between programs—a necessary component of a successful integrated data system. These meetings also gave programs the chance to review existing data elements, option sets, and definitions for accuracy and relevancy.

## Summary

Data integration is crucial in managing and understanding data and in helping people make informed and effective decisions regarding education. To learn more about the value of an integrated data system, see the CIID topical brief *Benefits of an Integrated Data System* (*https://ciidta.grads360.org/#communities/pdc/documents/7641*). The tools available in CEDS provide a way for SEAs to identify what data they collect and where they are stored and helps them to compare the elements in their data dictionaries across data systems in the form of CEDS maps. This facilitates elimination of redundant data, sets standards for the data, and identifies a system of record for which various data entities can then be used by other systems.

It is important to note that the tools within CEDS do not have to be adopted in their totality to be useful. The tools are valuable whether they are used singularly, or in combination. For example, although the Connect tool builds off of the Align tool by linking SEA data elements and option sets to common data standards, it can also be used as a standalone tool to document and share uses of data or to create the framework and analysis recommendations for various reports. The important point is to use CEDS and the tools in a way that makes sense for your situation, adds value to the work you are doing, and provides a solution to a data management issue or challenge.

# Contacts and Resources

The CEDS site has numerous tutorials and transcripts (https://ceds.ed.gov/learnCedsAlignment.aspx).

For data integration technical assistance, contact CIIDTA@aemcorp.com.

Visit the CIID website for more information (https://ciidta.grads360.org/#program).

# Glossary of Terms

### CEDS

Common Education Data Standards (CEDS) is a national collaborative effort to develop voluntary, common data standards for a key set of K-12 and postsecondary variables (e.g., demographics, program participation, transition and course information). *Source: CEDS*

### CIID

The Center for the Integration of IDEA Data (CIID) provides technical assistance to SEAs to increase the capacity to report high quality data required under the Individuals with Disabilities Education Act (IDEA) Part B Sections 616 and 618. CIID supports the integration of IDEA data systems and processes with the Statewide Longitudinal Data System (SLDS). *Source: CIID*

### Data Dictionary

An agreed-upon set of clearly and consistently defined elements, definitions, and attributes. It helps an organization maintain consistency in its information systems. Database users and managers refer to a data dictionary to find out where specific data are located, whether they were reported correctly, how to use them appropriately, and what their values mean. *Source: CEDS*

### Data Element

An atomic unit of data that has precise meaning or precise semantics that can be defined and measured. *Source: CEDS*

## Data Source

A data source, in the context of computer science and computer applications, is the location where data that is being used come from. In a database management system, the primary data source is the database, which can be located in a disk or a remote server. The data source for a computer program can be a file, a data sheet, a spreadsheet, an XML file or even hard-coded data within the program. *Source: techopedia.com*

## Data Type

A description of the form or qualities (i.e., the "type") that the data element contains. Data element "types" include (1) Alpha/Numeric (AN): A data element for which any letter or number (or combination of letters and numbers) is appropriate and (2) Date (DT): A data element type that is specifically defined as a date. Generally, the format is specified for the standard, e.g. MMDDYY. *Source: CEDS*

## Map

A CEDS map is a data dictionary aligned to the Common Education Data Standards. *Source: CEDS*

## Option Set

The list of machine-readable codes and human-readable labels and/or definitions that specify the limited set of value options allowed for a data element. Code Set, Permitted Values, and Option Set are used interchangeably by some education agency IT professionals. *Source: CEDS*

## State Education Agency (SEA)

The state-level entity primarily responsible for the supervision of the state's public elementary and secondary schools. *Source: CEDS*

## Statewide Longitudinal Data System (SLDS)

In education, the SLDS is the unit-level data system designed for the collection, management, analysis, and reporting of statewide education data over time and across programs. *Source: CEDS*

## Tools

In CEDS, the Align, Connect and myConnect tools are designed to help users organize, analyze, understand, and compare data. *Source: CEDS*

Contact Us:
By email at CIIDTA@aemcorp.com
Visit the CIID website for more information at CIIDTA.org

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