



METADATA DICTIONARY – Course Number Crosswalk (State/District/SCED)

A crucial success factor of any enterprise education data system is the adoption of comprehensive metadata standards—and compliance with them. ESP Solutions Group can work with your education agency to standardize data elements, indicators, naming conventions, definitions, codes, and formats for improved data quality, comparability, and interoperability.

Metadata are literally "data about data." They are information that describe data.

ESP has provided nationally recognized expertise in the domain of education data and metadata since 1993. We specialize in working with education agencies to document data about data (metadata). Our combined knowledge and experience with metadata come together in **DataSpecs™**, a metadata inventory tool that facilitates sharing of data across all applications without the need to reformat them time and time again. More specifically, DataSpecs is used by ESP to manage the documentation of each data collection, repository, indicator/index, and output/report through our metadata services.

The data elements entered into DataSpecs can be linked (mapped) to data elements in existing education agency data dictionaries or to national standards, including Common Education Data Standards (CEDS), the National Center for Education Statistics (NCES) Handbooks, *EDFacts*, the Schools Interoperability Framework (SIF), and the National Education Data Model (NEDM).

BENEFITS

- DataSpecs maintains a comprehensive inventory of existing data elements in a custom metadata dictionary.
- DataSpecs locates data elements in collections from which they are being gathered, repositories where they are maintained, and outputs where they are used.
- DataSpecs allows certified users to enter, modify, and track elements as they appear throughout the agency in collections, repositories, and outputs (reports).
- DataSpecs assigns each collection/repository to a data steward – an office or person who is the owner of the metadata and responsible for its maintenance.

Output: A report, file, or other media for providing or accessing data.

WHY AN EDUCATION AGENCY MUST HAVE A METADATA DICTIONARY

Eleven Objectives Met Using ESP's DataSpecs™ Metadata Dictionary Tool

There are many reasons an education agency needs DataSpecs to manage its metadata. In general, metadata are data about data, not the data themselves. So, DataSpecs is a tool useful for describing how an organization, specifically an education agency, manages its data to ensure quality and usability.

Below are the major reasons an education agency uses DataSpecs. Typical problems that are heard around education agencies are stated. Then for each, an objective is presented and an example is provided along with a sample report to illustrate what DataSpecs produces.

Problem Heard	MY PEOPLE WASTE TIME REPORTING THE SAME DATA OVER AND OVER.
Objective 1	To reduce the burden imposed by the redundancy from collecting and storing the same data multiple times, often with conflicting definitions, formats, and codes. This is the major complaint heard from schools to districts, from districts to states, and from states to the U.S. Department of Education. Each level of the education enterprise believes that the next higher one asks for the same data to be reported multiple times, usually with different definitions or in different formats. Everyone is working on the problem, but without a single point of management such as DataSpecs, the progress is hit and miss.
Example	The data elements (represented as "items") on each data collection throughout the agency are entered into DataSpecs. Each item is mapped to a common element that represents all of the items across all collections that are similar. This creates a linkage across all the agency's collections and allows analysis of the redundancy. The office/steward may then determine which redundancies can be deduplicated and which are necessary.
DataSpecs Report	Element Profile (for a single element, a report showing all collections and repositories where the element is used as an item.)



Problem Heard	HOW CAN WE KNOW IF THESE COURSES ARE EQUIVALENT ACROSS DIFFERENT SCHOOLS AND DISTRICTS?
Objective 2	To align local courses to state and/or national (SCED) course classification systems to allow the crosswalking of course codes for electronic reporting across entities using state or national standards. Without common course identifiers, education agencies must create articulation agreements or informal crosswalks to align each others courses for acceptance of credit. DataSpecs provides for crosswalking from an uploaded local course code file to an uploaded state course code file; or from either to the NCES SCED course classification system.
Example	A district and SEA want to create a teacher-student data link but do not have local to state course mapping. To do so, they upload their course catalogue into DataSpecs and run the AutoMapping function. Local courses are mapped to SCED numbers, reviewed by local staff, accepted or edited as appropriate, and finalized. A file is downloaded into the local student information system and the SCED codes are imported. The state can use SCED now for all districts.
DataSpecs Report	Local to SCED Course Mapping

Problem Heard	EVERY TIME A NEW DEMAND FOR DATA ARISES WE INITIATE ANOTHER DATA COLLECTION.
Objective 3	To respond efficiently with existing data without imposing unnecessary increased burden when new mandates for data arise. The No Child Left Behind Act was the classic example of a mandate that imposed significant new requirements for reporting. States with individual student record level reporting systems were able to respond more easily than those with aggregate reporting from districts. When a new mandate arises, an agency must be able to inventory authoritative data sources to find the ones with the right elements and periodicities to meet the demand. Otherwise a new collection is required—or initiated unnecessarily.
Example	The Legislature passes a bill requiring the SEA to determine how many times the SEA asks an LEA each year to include their superintendent's email address on a report. There are three possibilities. Ask the LEAs, poll the SEA offices, or run a DataSpecs report.
DataSpecs Report	Element Profile (for a single element, a report showing all collections and repositories where the element appears as an item.)

Problem Heard	WE ARE BUYING A NEW SOFTWARE APPLICATION. HOW DO WE TELL THE BIDDERS WHAT THEY MUST BE SURE THEIR DATA MODEL WILL MATCH OUR DATA STANDARDS?
Objective 4	To create the comprehensive functional data requirements for a new software application, database, or data store to ensure interoperability with extant systems. A major mismatch with new software applications, especially major purchases such as data warehouses, can be their data models. Commercial products may come with an established database from another industry or level of the education enterprise that may not meet the needs of your agency. Customizing a commercial product may be difficult or impossible if that is not a requirement during procurement.
Example	The education agency is procuring a data warehouse.
DataSpecs Report	Repository Profiles (for each of the major repositories that will need to be incorporated into the data warehouse); Output Profiles (for each of the major outputs that will need to be produced from the data warehouse).

Problem Heard	WHO IS THE GO-TO PERSON FOR OFFICIAL DATA WHEN MORE THAN ONE OFFICE HAS THE DATA?
Objective 5	To supplement this calendar with data element level detail identifying the authoritative data source for each data element used for the organization's official statistics and for required reporting. The same data elements are shared in multiple locations throughout an organization; however, a single authoritative data source should be designated.
Example	The Deputy State Superintendent of Schools is meeting with the SEA staff to discuss grant applications, and everyone has been arguing over the number of low-income students in the state. What is the real number?
DataSpecs Report	Official Statistics: Source Data Elements



Problem Heard	HOW DO I KEEP UP WHEN ALL THE REPORTS ARE DUE? HOW DO I KNOW WHEN THE NEW OFFICIAL DROPOUT RATES WILL BE PUBLISHED? HOW DO I KNOW WHICH OFFICE IS THE PLACE TO GO FOR OFFICIAL STATISTICS?
Objective 6	To maintain and publish a comprehensive calendar and directory of the periodicity of the organization's official data collection and reporting activities. Anyone within the organization needs to be able to refer to a master calendar and know when a data submission is due or when an official statistic will be available. The data seeker also must know who is responsible and where to go to access the data.
Example	The State Board of Education Chair is addressing the Legislature and wants to know when the annual dropout statistics will be announced. The Superintendent of the largest urban district is cutting staff but first wants to know when the state data report is due for funding.
DataSpecs Report	Periodicity of Official Statistics

Problem Heard	DO WE REALLY HAVE ALL THE DATA REQUIRED FOR EDFACTS? WHAT ARE WE MISSING?
Objective 7	To analyze and identify the gaps between required data mandates, needs, and available data. This analysis provides confirmation that a mandate can be met, and if not, which data elements must be collected. Before consolidation of data systems, in the old days of stove pipes, a program office collected all the data it needed, stored them, and reported them on to their own federal office. No longer is that the model. The USED/IES/NCES and their program offices are moving toward consolidated collections. SEAs are moving toward consolidated reporting from a consolidated data store. This makes the challenge of searching for shared authoritative data sources and determining that all the mandated elements for a report are available with the required periodicity even greater. The ever-changing requirements from USED exacerbate the challenge for an SEA. These requirements are defined in DataSpecs as "outputs."
Example	EDFacts has become a continual series of submission files from an SEA to USED throughout an annual cycle. Currently there are 105 submissions, 205 files, 2,407 elements, and 253 distinct elements. An SEA needs to know at any time if there is a gap between these requirements and the available data in its repositories.
DataSpecs Report	EDFacts Map and Gap Analysis

Problem Heard	THE DATA WE GET ARE WRONG, BUT NO ONE KNOWS HOW TO FIX THEM. WE WONDER WHAT DECISIONS PEOPLE MADE WHEN THEY SUBMITTED THE DATA WE SEE REPORTED.
Objective 8	To document the business rules that apply to the data in a collection, a repository, or an output to ensure appropriate interpretation and use. Business rules describe not only the characteristics that ensure the quality of a data element but also the proper relationship of data elements to each other. Level 1 business rules can simply be determined within a single record. Level 2 business rules can be determined across records within a set of records. Level 3 business rules can be determined across sets of records, such as across years, or across different districts. In ESP's implementations of our State Report Manager™ product, that applies business rules for districts to clean their data, there can be dozens to hundreds of rules applied when a set of data are submitted. DataSpecs provides a tool for formulating and documenting these business rules, and passing them on to an application that will apply them to the data.
Example	An LEA submits a "Teacher/Class/Student" data file to its SEA. It goes first to the SEA's State Report Manager application for the checking of business rules. From DataSpecs, these rules have been captured and are being enforced before the data can be certified for use by the SEA. Examples: <ul style="list-style-type: none"> — Class size must be fewer than 45. — Teacher identifier must be valid (>5000000000, <6000000000). — Student cannot be in more than 8 classes.
DataSpecs Report	Business Rule Profile



Problem Heard	WE HAVE TO RE-ENTER DATA INTO MULTIPLE SOFTWARE APPLICATIONS EVERY TIME A STUDENT'S INFORMATION CHANGES.
Objective 9	To increase the interoperability among information systems to ensure efficient and accurate exchange of data. Data exchanges among databases, software applications, and program offices don't work, or worse are misleading, if the element definitions and standards are not identical. Interoperability is essential today because no one has time to manually re-key data and to correct entry errors. SIF is one option for interoperability. LEAs use SIF for horizontal interoperability among local software applications. LEAs and SEAs use vertical interoperability for local-to-state reporting.
Example	When a student enrolls in a school, that student's demographics and family characteristics need to be entered into multiple local and state applications. All of those applications must use the same element dictionary attributes for efficient exchange (without crosswalking or translation). DataSpecs will hold the item/element name, definition, codes for each application (considered a repository) and align them with the standard element definition as adopted by the agency to ensure interoperability. If SIF is adopted, DataSpecs maps to SIF and the objects in which the elements appear.
DataSpecs Report	Repository Profile; SIF Alignment Report

Problem Heard	I NEED TO COLLECT ALL NEW DATA FOR MY RESEARCH PROJECT.
Objective 10	To provide data for research from existing sources without burdening schools with new collections. Agencies need a process for providing researchers an overview and analysis of available data that may satisfy their needs as an alternative to collecting new data.
Example	A graduate student who is designing a dissertation study of student mobility patterns and their effects on graduation rates may not know that the SEA has a common student identifier across multiple years and all schools in the state. With a full understanding of which data files contain enrollment, withdrawal, dropout, and graduation data, the researcher may be able to design a study that uses only extant data.
DataSpecs Report	Repository Profiles

Problem Heard	CAN WE TRUST THESE DATA?
Objective 11	To increase the quality of the data throughout the agency. Often, data quality is challenged whenever someone wants to challenge the data themselves. A compelling response is excellent documentation, comprehensive definitions, and solid business rules. This provides an argument that the data providers, the data stewards, and the reporters of the data are compliant with those definitions and rules.
Example	A fully loaded and accessible metadata dictionary is the best documentation to show that an education agency has established and is following standards for quality data.
DataSpecs Report	All DataSpecs reports support this objective.

PRICING

The cost of the project depends on the desired customization and integration of the resulting DataSpecs tool. This project is most often combined with *ISInsight*™, Data Governance and Enterprise Architecture Services, CourseWalk™, or as part of a larger systems integration project. Smaller, more tailored projects are also possible – ESP will work with your agency at no charge to determine the best possible solution that fits within your budget.

WHY ESP?

ESP Solutions Group specializes in improving the quality of education data. Our staff has worked directly with the development of all NCES data handbooks, Handbooks On-line, the U.S. Department of Education's metadata dictionary, and many other national efforts to define best practices for education data collection and reporting. ESP personnel have advised many local school systems, all 52 state-level education agencies, and the U.S. Department of Education on the practice of P-20 data management. We are nationally recognized as leading experts in understanding the data and technology implications of state reporting, the No Child Left Behind Act (NCLB), EDEN/EDFacts, the Schools Interoperability Framework (SIF), Common Education Data Standards (CEDs), and EDFACTS Shared State Solution (ES3).

CourseWalk™

CourseWalk™ + DataSpecs™ = Teacher/Student Data Link
Standardized course codes are essential to tracking student performance and improving instruction in the long run. They are also extremely useful for standardizing information in student transcripts. ESP's CourseWalk tool facilitates the matching of local course codes to state or national course code standards. Adding CourseWalk to your DataSpecs project completes the teacher/student data link. See Objective 2.