U.S. DEPARTMENT OF EDUCATION

ED*Facts* Workbook
SY 2007–08

Version 4.1

July 2008

U.S. Department of Education
Margaret Spellings
Secretary

Office of Planning, Evaluation and Policy Development
Williamson "Bill" M. Evers
Assistant Secretary

Performance Information Management Service
Ross Santy
Director

July 2008

This technical guide is also available on the Department's Web site at:
www.ed.gov/edfacts

On request, this publication is available in alternate formats, such as Braille, large print, or computer diskette. For more information, please contact the Department's Alternate Format Center at (202) 260-0852 or (202) 260-0818.

DOCUMENT CONTROL

DOCUMENT INFORMATION

Title:	ED*Facts* Workbook SY 2007-08
Revision:	Version 4.1
Issue Date:	July 2008
Security Level:	Unclassified
Filename:	SY 2007-08-edfacts-workbook-4-1.doc

DOCUMENT HISTORY

Version Number	Date	Summary of Change
2.0	10/17/05	Version 2.0 or subsequent updates (i.e., 2.1., etc.) contain information about the SY 2005–06 requirements.
3.0	12/06	Version 3.0 or subsequent updates (i.e., 3.1., etc.) contain information about the SY 2006–07 requirements.
4.0	1/08	Initial version for SY 2007-08.
4.1	7/08	Updates made to the following sections: • 1.3, corrected bullet formatting; • 4.3, corrected citation (section 1111, not 111); • 5.3.6, corrected rows 1 and 3 in the table; • 5.3.7, corrected description of files 121 and 121 in the table; • 5.3.14, removed data group 685 from table and added explanatory footnote; • 5.4.2, added 685 to list of approved data groups not included in the SY 2007–08 data collection; • 7.8, clarified that warnings will still appear on the error report even after an explanation is provided; • 7.9, corrected last sentence; • 11.1.1, clarified what states should do if data with an established due date will be late; • 13.1.2, relabeled data group 644 in table 13.1-3; and • 13.3, corrected the due date year for the 2007-08 assessment data.

RELEASE AUTHORIZATION

Name	Signature / Date
Author	Roberta Garrison-Mogren
QA	
Release Authority	

Preface

ED*Facts* is a collaborative effort of the U.S. Department of Education (ED), state education agencies (SEAs) and industry partners to centralize state-reported K through 12 educational performance data into one federally coordinated data repository located in ED. The purposes of ED*Facts* are to

- Place the use of robust, timely performance data at the core of educational decision making and policymaking

- Reduce state and district burden and streamline data practices

- Improve state data capabilities by providing resources and technical assistance

- Provide data for planning, policy and management and the federal, state and local levels

The Education Data Exchange Network (EDEN) is a centralized, coordinated repository of state K through 12 education and performance data residing at ED. EDEN includes: (1) the EDEN Submission System (ESS), an electronic data system capable of receiving data on more than 100 data groups at the state, district, and school levels and (2) the EDEN Survey Tool (EST), which collects data not currently collected through ESS.

The *EDFacts Workbook SY 2007–08* is a reference guide to using ESS to submit data for school year (SY) 2007–08. States report data to ED using ESS. States transmit data through ESS to meet data requirements of annual and final grant reporting, specific program mandates and the *Government Performance and Results Act*.

Table of Contents

1.0 Introduction

1.1 Purpose

This document provides information on how to submit files through the EDEN Submission System (ESS) into ED*Facts* for school year 2007–08. This document assumes a basic understanding of the ED*Facts* data framework. An explanation of the ED*Facts* data framework is found in appendix B.

The ESS shares an opening Web page with the Consolidated State Performance Report (CSPR) and other data submission tools; however, this document does not cover how to use these systems. Each of the submission systems has a Getting Started guide or a User Guide.

1.2 ED*Facts* Collection

The Office of Management and Budget (OMB) approved the "Annual Mandatory Collection of Elementary and Secondary Education Data for ED*Facts*" for SY 2007–08 on October 17, 2007. This data collection activity is approved for 3 years under OMB Control Number 1875-0240. The total data approved for collection by ED*Facts* includes:

- Data submitted by state education agencies (SEAs) through ESS to ED*Facts.*

- Data collected from other sources (i.e., U.S. Bureau of the Census).

- Civil Rights Data Collection (CRDC).

- Metadata from SEAs and other sources.

- Submission plans for data to be submitted by states to ESS.

ED*Facts* does not collect individual student or staff-level information. All information provided to ED*Facts* is aggregated—often by categories such as grade level. Although some of the data files may contain small numbers, none of the information is linked to specific students or staff members. In submitting data to ED*Facts*, states cannot suppress the data in small data cells except as specifically authorized by federal statute.

This document focuses on data submitted through ESS, metadata and the data provided by each state's annual state data submission plans.

1.3 ED*Facts* Coordinator

Each SEA identifies a staff member who serves as the ED*Facts* Coordinator. The ED*Facts* Coordinator is the official SEA contact for the ED*Facts* team. The Department of Education recommends the following knowledge, skills, abilities and responsibilities for each state's ED*Facts* Coordinator.

Knowledge, Skills, and Abilities

ED*Facts* Coordinators should be completely committed to improving education practices through improving the use of timely, quality data in education and

administration decisions. In addition, ED*Facts* Coordinators should possess the following knowledge, skills and abilities:

1. Knowledge about federal and state education programs, laws and regulations including the ability to:

   o Identify and work with education program data stewards.

   o Work with high-level state education policymakers.

   o Understand education data sources, work processes and users' needs.

   o Coordinate and manage data work flow processes with different programs.

   o Perform data quality checks for completeness, accuracy and validity.

2. Knowledge about information technology and data storage, transmission and use, including the ability to:

   o Translate technology terminology into education program language.

   o Understand the difference between good information and good technology.

   o Understand and use ED*Facts* and state technology tools.

   o Understand the federal and state data-sharing processes.

   o Understand the differences among multiple hardware tools.

   o Understand the differences among multiple software tools.

   o Compensate for the differences among hardware and software tools.

   o Plot hardware and software development trends and plan upgrades.

   o Understand and work closely with computer programmers.

   o Work closely with database designers, builders and managers.

3. Knowledge about statistics (the analysis, interpretation and presentation of data), including the ability to:

   o Translate statistical principles into useful education terminology.

   o Bring additional analytic power to the available education data.

   o Explain the analytical advantages of sharing quality data.

   o Identify logical and statistical flaws in a data presentation.

   o Present data in accurate and meaningful ways.

4. Skills in program management, diplomacy and negotiation, including:

o   Ability to plan, organize and manage the data collection and reporting process.

o   Ability to use project management tools.

o   Ability to understand and accommodate the many different programs' needs.

o   Ability to lead a data governance board and a data stewards group.

o   Ability to negotiate win-win data sharing solutions between/among programs.

5. Skills in teaching and motivating, including:

o   Ability to articulate the "big picture" (ideal future state) to skeptics.

o   Ability to help each "part" appreciate its place in the "whole" organization.

o   Ability to translate between professional languages.

o   Ability to keep communication open, active and professional.

o   Ability to motivate the unmotivated.

Responsibilities

ED*Facts* Coordinators are critical to the success of ED*Facts*. As reflected by the wide range of skills required for the assignment, the responsibilities of ED*Facts* Coordinators are wide ranging. These responsibilities include the following tasks.

1. Work with education program experts to identify data requirements.
2. Work with SEA leadership team to identify requirements.
3. Understand education data sources, work processes and users' needs.
4. Coordinate and manage acquisition of data from different state program offices.
5. Perform quality checks on data for completeness, accuracy and validity.
6. Translate technology requirements for education program leadership teams.
7. Promote policies and practices to secure good information and good technology.
8. Use the ED*Facts* and state technology tools to transmit data.
9. Document the federal and state data-sharing processes.
10. Compensate for the differences among multiple district and state hardware tools.
11. Compensate for the differences among multiple district and state software tools.
12. Forecast hardware and software development trends and plan upgrades.
13. Work closely with state computer programmers and contractors.
14. Work closely with database designers, builders and managers.
15. Translate statistical principles into useful education terminology.
16. Promote techniques for getting additional analysis from the available education data.
17. Promote the analytical advantages of sharing quality data.
18. Improve the logical and statistical presentation of data.

19. Present data in accurate and meaningful ways.
20. Plan, organize and manage the data collection and reporting process.
21. Use project management tools to manage the data collection process.
22. Accommodate the different education programs' (sometimes conflicting) needs.
23. Lead a data governance board and a data stewards group.
24. Negotiate win-win data-sharing solutions among programs.
25. Articulate the "big picture" (ideal future state) to skeptics.
26. Help each "part" appreciate its place in the "whole" organization.
27. Translate between professional languages.
28. Keep communication open, active and professional.
29. Motivate the unmotivated.

The ED*Facts* Coordinator may designate additional persons to serve as SEA submitters. These are usually SEA staff or contractors who have the knowledge needed to extract the data from existing state systems, put the data into the ED*Facts* file formats, submit the data and answer questions or handle programming issues that arise in the submission of the data.

1.4   Access to Systems

For SEAs, there are three types of ED*Facts* Users: ESS Submitters, ESS Reviewers and ED*Facts* Reporting System Users. Each SEA can have one or more users of the ESS and one user of the ED*Facts* Reporting System. The differences between these user types are described below.

ESS Submitter—This role of this ESS user is to submit files and review reports on the status of files submitted.

ESS Reviewer—This user provides access to state submission and transmission reports but does not allow access to the Transmittal Authorization sub-application. These personnel do not need the ability to submit files but only an ability to review and analyze error reports and data. These users may include the Common Core of Data (CCD) Coordinator and the *IDEA* Part B Data Manager.

ED*Facts* Reporting System User—One user per state can access the ED*Facts* Reporting System. The ED*Facts* Reporting System contains reports on the status of submissions through ESS, data submitted through the CSPR tool and data submitted through ESS.

The ED*Facts* Coordinator approves users for both systems. ESS may be accessed at https://eden.ed.gov/EDENPortal/.

1.5   Technical Assistance

In addition to this workbook, the following documents are available to assist SEAs in submitting data through ESS. All the documents listed below are available on the ED*Facts* Web site—http://www.ed.gov/edfacts.

EDEN Submission System (ESS) Users Guide—This guide addresses the basic mechanics of system access and file transmission. The figures in this document contain screen prints of the ESS to depict step-by-step instructions for ESS users. The guide is updated annually and when major system modifications affect user procedures.

EDEN Submission System's Release Notes—These documents provide a description of the technical enhancements to each version of ESS.

File/XML Specifications—These documents provide policy guidance and technical instructions for building the files that are submitted through ESS. File/XML specifications apply to a specific school year. XML validation schemas (.xsd) and style sheets (.xsl) are also available.

ED*Facts* Business Rules Guide—This document lists all the business rules that ESS uses to check the quality of data submitted to the system.

FAQs—This document lists answers to frequently asked questions (FAQs).

ED*Facts* Quarterly Newsletters—Newsletters contain information on relevant legislation, regulations, data collection information, training sessions, meetings and other events that occurred in the last quarter or are scheduled in the future.

To assist states with data submission, analysis and reporting, ED provides a dedicated Partner Support Center (PSC). The PSC can be reached via telephone (877-HLP-EDEN) or e-mail (eden_ss@ed.gov). PSC distributes e-mail announcements about system shut downs, reminders of due dates and technical hints. SEA staff members who use ESS automatically receive these e-mails.

In addition to the documents, the following tools are available upon request from the PSC:

ED*Facts* Data Set Viewer—The ED*Facts* Data Set Viewer is a customized Access Database that allows users to review data groups and categories for SY 2006–07 and SY 2007–08.

PSC File Format Checker (PSC Internal Tool)—This tool can be installed locally and states can use it to check for non-XML format errors prior to attempting to submit to ESS.

Other sources of information include the following:

Partner Support Center Broadcasting—Informative due date reminders, additional guidance support, etc.

EDICs—The Education Information Collection System provides access the information collection document sets of all OMB approved collection efforts. The Web site location is http://edicsweb.ed.gov/

ED encourages SEAs to share information among themselves. To facilitate this, ED maintains and distributes upon request the list of ED*Facts* Coordinators to staff at SEAs. ED also maintains a Listserv at edfacts@lists.ccsso.org. ED*Facts* Coordinators, submitters and other SEA staff members can initiate and participate in discussions on the email list server.

2.0 Data Submission Calendar

This section describes the calendar for submitting files to ESS for SY 2007–08.

2.1 ESS Files for Previous School Years

SEAs should submit files for SY 2005–06 and SY 2006–07[1] based on their ED*Facts* State Submission Plan. SEAs can monitor progress with the State Plan Execution Report (LEAD004) in the ED*Facts* Reporting System.

SEAs should be in lock status with their SY 2006-07 CCD **before** submitting data for SY 2007-08.

2.2 ESS Files for SY 2007–08

SEAs should begin submitting data for SY 2007–08 by submitting a complete directory of LEAs and schools. A directory file for the SEA should be submitted only if there were changes. The other files submitted through ESS are assigned to one of four collection periods:

- Early—January 2008 through April 2008
- Middle—February 2008 through June 2008
- End—June 2008 through September 2008
- Close-out—September 2008 through February 2009

The collection periods are based on ED's understanding of when data are available in most states. Through the EDEN Submission Plan, SEAs schedule the specific month in which they will submit each file.

Chapter 7.0 contains more information on the submission process.

2.3 Submission Plans

In January 2008, the PSC distributed the SY 2007–08 Submission Plan template. The Submission Plan template was due to ED by February 29, 2008. Chapter 6.0 contains more information on the Submission Plans.

---

[1] Data for SY 2003–04 and SY 2004–05 can no longer be submitted through ESS.

3.0 Department Regulation and Guidance

This section describes Department regulations and guidance affecting data submission through the ESS.

3.1 Regulation—Requiring the Electronic Reporting of Data to ED

On January 25, 2007, Final Regulations were published to amend 34 CFR part 76 governing state reporting requirements. The revised regulations allow the Secretary and the U.S. Department of Education, through the *Paperwork Reduction Act* clearance process, to require states to report certain information electronically. The regulations also establish that the Department may take administrative action against a state for failure to submit reports in the manner prescribed by the Secretary. This is intended to significantly reduce reporting burden and to streamline data collections currently required by the Department by bringing elementary and secondary education data in through the Annual Mandatory Collection of Elementary and Secondary Education Data for ED*Facts* (1875-0240). The failure to submit ED*Facts* data constitutes a failure to comply substantially with the law applicable to the funds provided by the federal government. The immediate remedy for this failure is the development of a State Transition Agreement Plan as required in the regulation. The Transition Agreement should address all three components of the annual ED*Facts* data submission requirements: the required data, the required format, and the time by which the required data must be submitted.

3.2 Guidance—Maintaining, Collecting and Reporting on Race and Ethnicity Data

The Secretary of Education published "Final Guidance on Maintaining, Collecting and Reporting Racial and Ethnic Data to the U.S. Department of Education" on October 19, 2007. This guidance implements the Office of Management and Budget's (OMB) *1997 Standards for Maintaining, Collecting and Presenting Federal Data on Race and Ethnicity*.

This guidance directly addresses three sets of issues:

1. How educational institutions and other recipients will collect and maintain race and ethnicity data from students and staff.

2. How educational institutions and other recipients will aggregate race and ethnicity data when reporting those data to the Department.

3. How data on multiple races will be reported and aggregated under the *Elementary and Secondary Education Act of 1965 (ESEA),* as reauthorized by the *No Child Left Behind Act of 2001 (NCLB).*

The guidance requires the collection of race and ethnicity data in a two-part question (or two-question format) which first asks whether the respondent is Hispanic/Latino and then allows the respondent to select one or more races from the following five racial groups:

1. American Indian or Alaska Native.

2. Asian.

3. Black or African American.

4. Native Hawaiian or Other Pacific Islander.

5. White.

The guidance requires educational institutions and other recipients to report the aggregated race and ethnicity data in the following seven categories:

1. Hispanic/Latino of any race; and for individuals who are non-Hispanic/Latino only.

2. American Indian or Alaska Native.

3. Asian.

4. Black or African American.

5. Native Hawaiian or Other Pacific Islander.

6. White.

7. Two or more races.

Educational institutions and other recipients are required to implement this guidance by fall of 2010 in order to report data for the 2010–2011 school year.

4.0 Standard Data Definitions

As part of the drive to streamline data collections across ED, ED*Facts* has standardized several definitions used through the ED*Facts* data groups. Although some of the previous definitions were similar, they were not identical, and states following the reporting requirements for various Program Offices to the letter often were required to classify students differently for different programs. This standardization of data definitions will reduce recordkeeping and reporting burden on the states, ultimately increasing the level of data quality.

4.1 Career and Technical Education (CTE) Concentrator

A secondary student who has earned 3 or more credits in a single CTE program area (e.g., health care or business services), or 2 credits in a single CTE program area, but only in those program areas where 2 credit sequences at the secondary level are recognized by the state and/or its local eligible recipients.

4.2 Children with Disabilities (*IDEA*)[2]

Children having mental retardation; hearing impairment, including deafness; speech or language impairment; visual impairment, including blindness; serious emotional disturbance (hereafter referred to as emotional disturbance); orthopedic impairment; autism; traumatic brain injury; developmental delay; other health impairment; specific learning disability; deaf-blindness; or multiple disabilities and who, by reason thereof, receive special education and related services under the *Individuals with Disabilities Education Act (IDEA)* according to an Individualized Education Program (IEP), Individualized Family Service Plan (IFSP), or service plan.

4.3 Limited English Proficient Students (Also Known as English Language Learners)[3]

In coordination with the state's definition based on Title 9 of *ESEA*, students:
    (A) who are ages 3 through 21;
    (B) who are enrolled or preparing to enroll in an elementary school or a secondary school;
    (C) (who are i, ii, or iii)
        (i) who were not born in the United States or whose native languages are languages other than English;
        (ii) (who are I and II)
                (I) who are a Native American or Alaska Native, or a native resident of the outlying areas; and
                (II) who come from an environment where languages other than English have a significant impact on their level of language proficiency; or

---

[2] Definition adapted from PL 108-446, Section 602(3), 34 C.F.R. Part 300.8(1)(a).
[3] Definition adapted from PL 107-110, Title IX, Part A(25).

(iii) who are migratory, whose native languages are languages other than English, and who come from an environment where languages other than English are dominant; and

(D) whose difficulties in speaking, reading, writing, or understanding the English language may be sufficient to deny the individuals (who are denied i or ii or iii)

(i) the ability to meet the state's proficient level of achievement on state assessments described in section 1111(b)(3);

(ii) the ability to successfully achieve in classrooms where the language of instruction is English; or

(iii) the opportunity to participate fully in society.

To be classified as limited English proficient, an individual must be A, B, C, and D. For C, an individual can be i, ii, or iii. If C-ii, the individual must be I and II. For D, an individual must be denied i or ii or iii.

4.4   Migrant

4.4.1 Eligible migrant children (also referred to as eligible migrant students)[4]

Children who are, or whose parents or spouses are, migratory agricultural workers, including migratory dairy workers, or migratory fishers, and who, in the preceding 36 months, in order to obtain, or accompany such parents or spouses, in order to obtain, temporary or seasonal employment in agricultural or fishing work (A) have moved from one local education agency (LEA) to another; (B) in a state that comprises a single LEA, have moved from one administrative area to another within such LEA; or (C) reside in an LEA of more than 15,000 square miles, and migrate a distance of 20 miles or more to a temporary residence to engage in a fishing activity.

4.4.2 Participating migrant children (also referred to as participating migrant students)

Children who participate in Migrant Education Programs (MEP) under Title I, Part C, including those served under continuation of services authority.

4.5   Private School Students

States should include private school students in only a small subset of ED*Facts* files related to *IDEA* Part B and Title I participation. The specific files are as follows:

***IDEA* Part B.** The following files that collect data on *IDEA* Part B include data on children with disabilities who have been parentally placed in private schools or facilities:

- N/X002 Children with Disabilities (*IDEA*) – School Age file {Data Group 74}—This file collects data on school-age children with disabilities for Table 1, Child Count, and Table 3, Educational Placement.

---

[4] Definition adapted from P.L 107-110, Title I, Part C and 34 C.F.R 200.31.

- N/X089 Children with Disabilities (*IDEA*) – Early Childhood file {Data Group 613}—This file collects data on children with disabilities (IDEA) who are ages 3 through 5 for Table 1, Child Count, and Table 3, Educational Placement.

- N/X009 Children with Disabilities (*IDEA*) – Exiting Special Education file {Data Group 85}—This file collects data on children with disabilities for Table 4, Exiting Special Education Programs.

Title I Participation. The following file that collects data on Title I can include data on students in private schools if those students receive services under Title I:

- N/X037 Title I – Participation file {Data Group 548}—This file collects data on students participating in Title I. These data have been collected through the Consolidated State Performance Report (CSPR) in section 2.1.3.

5.0 School Year 2007–08 Files

This section summarizes the files to be submitted for school year 2007–08. The list of files, by collection period, is in the EDEN Submission Plan Workbook. The files and collection periods are also available in the data framework and in the ED*Facts* Data Set Viewer.

5.1 Background

The files to be submitted are based on the data groups in the SY 2007-08 ED*Facts* Data Set. The data groups in that collection were organized into files and assigned to one of four collection periods based on when the data should be available in the SEAs.

Data groups are assigned IDs between 1 and 999. Files are also assigned IDs between 1 and 999. To differentiate, "ID#" is used without the leading zeros for data groups. "N" or "X" or both ("N/X") is added for files and include the leading zeros. "N" is used for fixed and delimited (non-XML) files, and "X" is used for XML files.

5.2 Directory

The first step in submitting files to ED*Facts* is to submit the directory.[5] Each SEA is required to submit a complete directory file for LEAs and schools. For SY 2007–08, ED plans to obtain Locale (ID#17), Congressional District Number (ID#13), County Name (ID#572), Geographic Location (ID#14), DUNS Number (ID#6) and Title 1 District Status (ID#582) from other federal sources instead of SEAs. The "Directory files" section (10.0) of this workbook provides more information on directory data and submitting directory files.

5.3 Relationships Among Files

In ED*Facts*, the data are organized differently from the organization used in the legacy collections. This subsection explains how certain files (data groups) relate to other files (data groups). This subsection also associates files with programs of ED. This subsection is not intended to be a detailed crosswalk between ED*Facts* and legacy collections. This subsection does not supplant the file specification guidance. Data submitters will need to carefully review the guidance in the individual file specifications before submitting files.

At the request of many SEAs, for SY 2007–08, the status files N/X101 and N/X102 were broken into smaller files.

5.3.1 Non-fiscal Common Core of Data

The CCD is collected for the National Center for Education Statistics (NCES). Section 13.1 contains a detailed crosswalk between CCD and ED*Facts.*

---

[5] The directory and grades offered files in ED*Facts* are the same as the "Universe" in CCD.

5.3.2 Consolidated State Performance Report

A detailed crosswalk between ED*Facts* and the CSPR for SY 2006–07 is included in the "Getting Started Guide for the CSPR tool." Once the CSPR for SY 2007–08 has been approved, a crosswalk between ED*Facts* and the CSPR for SY 2007–08 will be issued.

5.3.3 Accountability data

Under *NCLB*, states determine the Adequate Yearly Progress (AYP) status of each LEA and school. The matrix below illustrates the relationship between AYP status and the indicators used to calculate it and ED*Facts* data groups and files.

AYP Status (Alternative Approach Status) N/X103, ID#32 (N/X103, ID#617)					
Reading/Language Arts		Mathematics		Other indicators	
Participation	Proficiency	Participation	Proficiency	Elementary/ secondary	High school graduation rate
N/X110, ID#553	N/X111, ID#552	N/X108, ID#555	N/X109, ID#554	N/X106, ID#556	N/X107, ID#557

In addition, Improvement Status for schools (ID#34) and LEAs (ID#662) are reported in N/X130 *NCLB* file. Graduation Rate Tables (ID#563) are also reported in N/X041.

5.3.4 Assessment data

Under *NCLB*, states assess students annually. The matrix below illustrates the relationship between the data on the assessments and ED*Facts* data groups and files.

Student Participation		
Reading/Language Arts	Mathematics	Science
N/X081, ID#589	N/X081, ID#588	N/X081, ID#590
Student Performance		
Reading/Language Arts	Mathematics	Science
N/X078, ID#584	N/X075, ID#583	N/X079, ID#585

5.3.5 Highly qualified teachers (HQT)

Under *NCLB*, core academic classes must be taught by highly qualified teachers.

Elementary	N/X063, ID#381
Secondary	N/X064, ID#383

5.3.6 Title I, Part A of *ESEA*—improving basic programs operated by LEAs

As part of the CSPR, ED collects data on participants and staff in Title I, Part A of *ESEA*. Title I, Part A programs can be classified into four types: Public School Wide Programs (Public SWP), Public Targeted Assistance Programs (Public TAS), Neglected and Delinquent Programs (N and D), and Services to Children in Private Schools (Private). The matrix below illustrates the relationship between ED*Facts* data groups and files and the data collected on participants and staff in Title I, Part A programs.

File #	Data group ID	Data group name	Public SWP	Public TAS	N and D	Private
N/X134	670	Title I Participation Tables	X	X	X	X
N/X036	549	Title I TAS Services Table		X		
N/X037	548	Title I SWP/TAS Participation Tables	X	X		
N/X065	550	Title I TAS Staff Funded (FTE) Table		X		

In addition, the status of a school as Title I (ID#22) is reported in N/X129 CCD file. The status of LEA (school district) as a Title I District is derived from ED records.

5.3.7   Title I, Part C—migrant students (children) and the Migrant Education Program

Data are collected on students eligible for funding and for students served under the migrant student program. Unduplicated counts are also collected by timeframe: regular school year, summer session and 12 months. The best way to understand the data on migrant students is the following matrix.

Migrant student group	Reporting Period		
	12 month unduplicated	Regular school year	Summer / intersession
Eligible enrolled in public schools		N/X132, ID#110	
Eligible for funding purposes—from the program	N/X121, ID#634 (formerly Count #1 on the Migrant Child Count Report)		
Eligible and served			N/X122, ID#635 (formerly Count #2 on the Migrant Child Count Report)
Served (including after period of eligibility expired)	N/X054, ID#102	N/X123, ID#636	N/X124, ID#637
Services to students		N/X145, ID#684	

In addition, whether a school consolidated its MEP funds into a SWP (ID#514) is collected in N/X132 School file.

MEP staff is collected by headcount (ID#625) and FTE (ID#515) in N/X065.

5.3.8   Title I, Part D—prevention and intervention programs for children and youth who are neglected, delinquent or at risk

As part of the CSPR, ED collects data on Title I, Part D. The diagram below illustrates the data collected. The table below the diagram explains the relationship to ED*Facts* files and data groups.
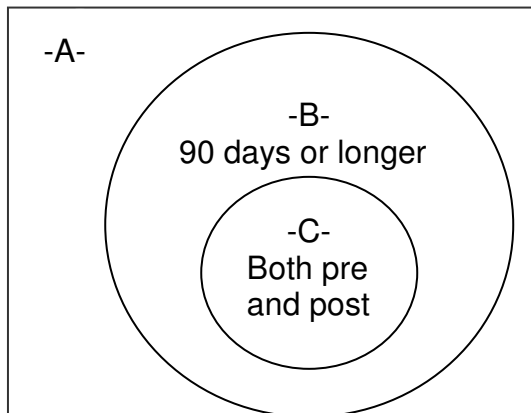
Explanation of diagram		Subpart 1		Subpart 2	
A	All N, D or at-risk youth	N/X119	ID#656	N/X127	ID#657
B	Youths who were in programs 90 days or longer.	N/X135	ID#672	N/X135	ID#671
C	Youths who were in programs 90 days or longer AND took both a pre- and post-test.	N/X113	ID#628	N/X125	ID#629

5.3.9   Title II, Part D—enhancing education through technology

ED is required to report on the performance of Title II, Part D. The following data are used in the performance indicators.

Integrated Technology Status	N/X131	ID#524
Personnel Skilled in Technology	N/X071	ID#526
Computers	N/X028	ID#525
8th-Grade Technology Literacy	N/X117	ID#650

5.3.10 Title III—English language proficiency

Title III of *ESEA* provides grants for the education of limited English proficient students. The best way to understand the data for students who are served by Title III and/or are limited English proficient is to focus on the universe included in each file.

Universe of students who are included	What data are collected about the universe of students included	Reporting period	File #	Data group ID
All students assessed and	Count of the students who were assessed and determined to be	October 1	N/X047	116

Universe of students who are included	What data are collected about the universe of students included	Reporting period	File #	Data group ID
determined to be LEP	LEP			
	Count of the students who were assessed and determined to be LEP	School year	N/X141	678
	The testing status for the annual assessment on English proficiency of the students in N/X141 who were present during the testing window	Testing window	N/X137	674
	The results of the annual assessment on English proficiency (This is the subset of students in N/X137 who were tested.)	Testing window	N/X139	676
Students who took native language assessments	The results of native language assessments, if state has native language assessments	Testing window	N/X049	272
Students enrolled in any program designed for LEP students	Count of students enrolled in a program designed for limited English proficient students	October 1	N/X046	123
Students enrolled in Title III programs	Count of students who are enrolled in a program designed for limited English proficient students that is funded by Title III	School year	N/X116	648
	The testing status for the annual assessment on English proficiency of the students in N/X116 who were present during the testing window	Testing window	N/X138	675
	The results of the annual assessment on English proficiency (This is the subset of student in N/X138 who were tested.)	Testing window	N/X050	151
Monitored Former Title III Students	Count of the monitored former Title III students who no longer receive services and the results of their content assessment	School year	N/X126	668
Immigrants	Count of the immigrant students whether LEP or not	School year	N/X045	519

In addition to data collected on students, N/X067 collects data on teachers in Title III programs (ID#422).

Title III data also includes the Annual Measurable Achievement Objectives (AMAOs):

Proficiency Attainment	N/X103 AYP Status	ID#518
Progress	N/X103 AYP Status	ID#569

5.3.11 Title IV, Part A—*Safe and Drug-Free Schools and Communities Act*

ED is required to prepare an annual report to Congress on the *Safe and Drug-Free Schools and Communities Act*. Some of the data for that report come from the following files:

Persistently Dangerous Status	N/X130 School	ID#36
Discipline – Incidents	N/X030	ID#523
Discipline – Counts of Students	N/X136	ID#673
Truants	N/X131 LEA	ID#664

5.3.12 *Gun Free Schools Act* report

ED is required to prepare an annual report to Congress on the *Gun-Free Schools Act*. Some of the data for that report come from the following files:

Filing of *GFSA* Report by LEAs and Schools (ID#603)	
N/X131 LEA	N/X132 School

Firearms Incidents	
Students involved	Number of incidents
N/X086, ID#596	N/X094, ID#601

5.3.13 Title X, Part C – *McKinney-Vento*—homeless children and youths

ED is required to prepare an annual report to Congress on the *McKinney-Vento* grants. Some of the data for that report comes from the following files:

Homeless Enrolled	N/X118	ID#655
Homeless Served (This is a subset of the students reported in N/X118)	N/X043	ID#560

5.3.14 Rural education

ED will use the following data to monitor and allocate rural education grants.[6]

REAP Alternative Funding Status	ID614 in N/X131 LEA

5.3.15 *Individuals with Disabilities Education Act (IDEA)*—Section 618 tables

In the past, *IDEA* Part B data have been collected through tables. Many states have already been approved to submit the *IDEA* Part B data through EDEN only. Below is a crosswalk between the ED*Facts* files and the tables. After the crosswalk is a written description of the relationship. Detailed crosswalks for each table are included in the relevant file specifications.

---

[6] Data Group 685 was included in the 2007-08 OMB package and approved for collection. It was also included in this section of Workbook version 4.0, released in January 2007. However, this data group will not be collected for SY 2007-08 or 2008-09 .

IDEA data collection	Legacy collection	EDFacts file	Data group
Child Count and Educational Environment	Table 1/ Table 3	N/X002	74
		N/X089	613
Staffing	Table 2	N/X070	486
		N/X099	609
		N/X112	647
Exiting	Table 4	N/X009	85
Discipline	Table 5	N/X005	512
		N/X006	475
		N/X007	476
		N/X088	598
		N/X143	682
		N/X144	683
Assessment	Table 6	N/X003	447
		N/X004	491
		N/X093	618

Table 1/Child Count
Report of Children with Disabilities Receiving Special Education under Part B of the *Individuals with Disabilities Education Act* (OMB # 1820-0043)

- N/X089 (ID#613) is used to submit early childhood data.

- N/X002 (ID#74) is used to submit school-age data.

Table 2/Personnel
Personnel Employed to Provide Special Education and Related Services for Children with Disabilities (OMB #1820-0518)

- N/X070 (ID# 486) is used to submit data on teachers employed or contracted to provide services to children with disabilities.

- N/X112 (ID#647) is used to submit data on paraprofessionals employed or contracted to provide services to children with disabilities.

- N/X099 (ID#609) is used to submit data on related-service personnel employed or contracted to provide services to children with disabilities.

Table 3/FAPE
Part B, *Individuals with Disabilities Education Act* Implementation of FAPE Requirements (OMB # 1820-0517)—The data on educational environments are submitted in the same files as the Table 1/Child Count data.

Table 4/Exiting
Report of Children with Disabilities Exiting Special Education during the School Year (OMB # 1820-0521)

- N/X009 (ID#85) is used to submit all the exiting data.

Table 5/Discipline

Report of Children with Disabilities Subject to Disciplinary Removal (OMB # 1820-0621)—The diagram below is intended to illustrate the relationship among the files submitted.

- N/X088 (ID# 598) is used to submit number of children with any type of removals by total removal length.

- N/X006 (ID#475) is used to submit number of children who had in- or out-of-school suspensions and expulsions.

- N/X005 (ID#512) is used to submit number of children who had removals to interim alternative education settings.

- N/X007 (ID# 476) is used to submit the number of removals of students to interim alternative education settings by school personnel by the reason for the removal.

- N/X143 (ID#682) is used to submit the total number of removals.

- N/X144 (ID#683) is used to submit data on provision of educational services to students who were expelled.
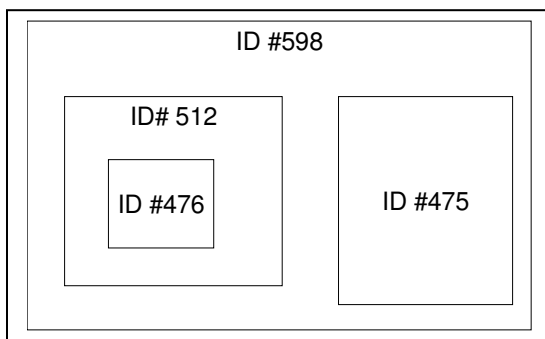
Table 6/Assessment
Report of the Participation and Performance of Students with Disabilities on State Assessments by Content Area, Grade and Type of Assessment (OMB # 1820-0659) This table collects three types of data on assessments: participation, no assessment and results.

- N/X093 (ID#618) is used to submit the participation by type of assessment.

- N/X004 (ID#491) is used to submit the data on students not assessed by the reasons not assessed and students with invalid results.

- N/X003 (ID# 447) is used to submit the results of the assessments.

5.3.16 *Perkins*—career and technical education

The Carl Perkins Program has performance indicators. The following data are used for those indicators.

Concentrators Leaving Secondary Education	N/X082	ID#521
Academic Attainment (Subset of N/X082 who took the *NCLB* state assessment)	N/X142	ID#681
Graduates (Subset of N/X082 who received a diploma)	N/X083	ID#320

5.3.17 Financial

GEPA requires a report on the allocation of funds. The following data are used for that report. In addition, financial data collected will be used to monitor flexibility in *NCLB* to transfer funds.

Offered Programs	N/X035	ID#470
Funding Allocations	N/X035	ID#547
Funds Transferred Under *NCLB*	N/X120	ID#663

5.3.18 Other

The following data groups are also collected and used by all Program Offices.

- School Poverty Percentage—N/X132, ID#31.

- School Totals—N/X131, ID#454.

5.4   Dropped ED*Facts* Data Groups

From one year to the next, data groups may be dropped from the ED*Facts* data collection for numerous reasons, primarily due to changes in program reporting requirements. The following sections describe data groups that have been removed from the ED*Facts* data collection effort for SY 2007–08.

5.4.1 Data groups dropped from SY 2006–07 to SY 2007–08

The following is a list of data groups that were included in the SY 2006-07 ED*Facts* data set and were removed from the ED*Facts* collection for SY 2007–08:

Data Groups Dropped for School Year 2007-08				
Data group #	Data group name	Steward	Reporting level	File spec number
33	Distinguished School Status	OESE	School	N101
351	AP Program Status Table	OCR	School	NA
410	Teachers New to District This Year	OESE	School, LEA	N061
455	AYP – Schools Meeting	OESE	LEA, State	N103
461	AYP – Districts Meeting	OESE	State	N103
482	Prekindergarten/Preschool Children Receiving Special Education and Related Services	OCR	LEA	NA
520	Promotion/Graduation Testing Status	OCR	School	NA
532	AYP – Schools Not Meeting	OESE	LEA, State	N103
534	AYP – Districts Not Meeting	OESE	State	N103
568	Children with Disabilities (OCR) Tables	OCR	School, LEA	NA
579	School Choice – Offered	OII/SES-Choice	LEA, State	N010
586	Student Performance in Writing	OESE	School, LEA,	N080

Data Groups Dropped for School Year 2007-08				
Data group #	**Data group name**	**Steward**	**Reporting level**	**File spec number**
	Tables		State	
591	Students Tested in Writing Tables	OESE	School, LEA, State	N081
595	Average Daily Attendance	OESE	School	N102
604	Operational Change	EDEN	School, LEA	N101
612	Supplemental Services – Offered	OII/SES-Choice	LEA, State	N102
615	RLIS Program Table	OESE/SST	State	NA
620	Vocational Concentrator Dropout Tables	OVAE	LEA, State	N100
621	LEP Placement Table	OELA	LEA, State	N085
623	LEP Accommodation Table	OELA	State	NA
640	GED Status	OCR	LEA	NA
665	Magnet Status (CCD)	NCES	School	N101

Note: In some cases, the data group listed in the table above was one of several data groups included in that file specification. In those cases, the file specification itself has not been dropped, just that component. Additionally, there are seven data groups shown here with NA in the file specification number column. These data groups either were never collected for SY 2006–07 or they were part of the Office for Civil Rights (OCR) data collection effort.

5.4.2 OMB-cleared data groups removed from the ED*Facts* data collection

The following data groups were approved as part of the SY 2007–08 data collection but will not be included in the SY 2007–08 data collection effort.

Data Group Number	File Specification Number	Data Group Name
677	N/X140	LEP Recent Arrivals
685	N/X133	Average Daily Attendance (Rural)

6.0   Prepare the SY 2007-08 Submission Plan

The ED*Facts* Coordinator completes an ED*Facts* SEA Submission Plan Workbook for SY 2007–08. A complete Submission Plan enables the ED*Facts* Coordinator to obtain accurate reports on the SEA's progress in submitting SY 2007–08 files. The SY 2007–08 plan should be completed for the <u>entire year, not for individual ED*Facts* collection periods</u>. When the Submission Plans are distributed, a due date will be provided. That date will be no earlier than February 29, 2008.

ED considers the Submission Plan submitted on the due date to be the state's <u>official</u> SY 2007-08 Submission Plan. Periodically, ED will run reports of actual file submissions against the data in the official Plan. However, we encourage ED*Facts* Coordinators to consider the Plan a 'living document' and to update it as more accurate information related to data submission becomes available. Updating the Submission Plan is done by contacting the PSC and submitting revised plans to PSC.

To complete the Workbook, the ED*Facts* Coordinator needs to identify the data source, estimate the number of schools and districts for which data will be reported and specify the date on which the file will be submitted for each level (i.e., SEA, LEA, school). The ED*Facts* Coordinator will work with the CCD Non-fiscal Coordinator, *IDEA* Part B Data Manager, CSPR Coordinator and other program managers to complete the Submission Plan. Section 5.3 contains information on the data to be submitted for SY 2007–08 that may be helpful in preparing the plan. Section 13.0 contains information on the transformation of legacy collections that may also be helpful in preparing the plan.

Additional workspace has been included in the Workbook template for SEA use.

7.0 Submission Process

This section summarizes the process for submitting files through ESS to ED*Facts*:

1. Submit the directory.
2. Gather the data.
3. Create the files.
4. Check the files.
5. Submit the files.
6. Correct format and validation errors and address reasonability warnings.
7. Monitor progress.

This section also provides information on what to do with large files and what happens to files that have high error rates.

7.1 Submitting the Directory

The directory files for LEAs and schools must be submitted before any other files can be submitted. Workbook section 10.0 "Directory Files" describes the directory.

7.2 Gather the Data

Based on the Submission Plan, the SEA submitter will be assigned one or more files to submit. The SEA submitter should read the file/XML specification for those files as outlined in workbook section 9.0 "File/XML specifications." The SEA submitter may need to review the "Guidance" and "Requirements for Submitting This File" sections of each file/XML specification with the Program Office owner or steward of the data in the state.

After confirming the source of the data, the submitter should decide if a staging database is necessary. In some database systems, creating views or queries is sufficient to pull the data together for a file. In other cases, the data would be pulled from the system and put into a staging database to organize the data in order to create the file. The best way to stage data is in a relational database. Excel or another spreadsheet application can be used to gather and stage the necessary data. However, the submitter must use caution when using Excel or other spreadsheet applications. If the cell formats are not set to "Text", Excel will remove the leading zeros when the files are saved. If the leading zeros are omitted in a field, the files may generate error messages when they are uploaded to ESS.

Each row of data in the staging database corresponds to a record in the data section of a non-XML file or one TABLETYPE object for XML files. For example, if student membership was submitted, a row of data in the staging database would be created for each grade level, race/ethnicity and sex combination. Consequently, there is a row for first-grade Asian males, a row for first-grade Asian females, a row for first-grade American Indian males, a row for first-grade American Indian females, etc. These rows

would convert to records in the data section of the non-XML file and to TABLETYPE objects for XML files.

The data in the staging database or staging view/query should be the same as what is laid out in the Data Record Definition section of the non-XML file specification document or objects for XML files.

While it is easier to group the data by education unit when reviewing them, subtotals and totals do not have to be contiguous to the detailed data. ESS sorts files by the identifiers to match the total records with the education units. If a file is split, all the data for a single education unit must be in the same file.

Some rows with counts of zero may need to be included because submitting the count of zero allows the data to be distinguished from missing or not collected data. Zeros provide a more accurate and complete picture. For example, in the membership file at the school level, if a school with seventh-grade enrollment has no seventh-grade American Indian male children, the school file must still include a row for this category with a count of zero. However, if the school does not have a seventh grade, then all the rows of data associated with the seventh grade can be omitted. Because the guidance varies between file specifications, it is important to read section 2.0 of the file specifications to determine how to handle zero and missing counts.

7.3   Create the File

After the SEA submitter gathers the data, the SEA submitter creates the file. Workbook section 8.0 "ED*Facts* Files" contains detailed information on creating files in the formats accepted by EDEN.

Files can be quite large. While ED has upgraded ESS to handle much larger files, local internet bandwidth bottlenecks can present problems when submitting very large files. There are two solutions:

1. Compress large files.
2. Split large files.

7.3.1 Compress large files

The ESS has been designed to recognize files compressed with the Zip data compression algorithm. The ESS will accept a Zip file and automatically extract (i.e., decompress) the file.

Plain text files, like ED*Facts* submission files, are very amenable to Zip compression. File size reductions of 50 to 95% are common with text files. There are a variety of commercial or free file compression utilities that create Zip files.

Zip-compressed files must have the .zip extension when uploaded to the ESS. The file name included in the Header Record (or FILETRANSMIT object if XML) should have the extension appropriate to the format of the uncompressed file.

7.3.2 Split large files

If compression is insufficient or not available, large files can be split into smaller files. A file **must** be split at an education unit boundary— the data for a single education unit must not be split between files. For example, an SEA may submit two files with school-level membership data, perhaps one with one school to test the format and a second file with all the other schools. The system will combine the data from the two files. However, the system cannot combine a file that contains some of the membership data for a school and a second file that contains the rest of the membership data for the school. The second file's data for that school will overwrite the first file's data for that school. For example, if the first file submitted contains membership data of a school for grades K through three and a second file contains membership data for that same school for grades four through six, the data in the second file will replace the data in the first file. In this example, only the data on grades four through six from the second file will move to the database for that school.

Each of the split files must still be properly formatted according to the file/XML specification. Each must have a unique File ID and File Name. Each non-XML file must have a Header Record with the correct record count for the individual file. Each XML file must be a valid XML document with valid FILETRANSMISSION start and end tags.

7.4 Check the File

Before the file is submitted through the ESS, the file should be reviewed. Workbook section 8.1 "ED*Facts* File Formats" contains specific rules for each of the file formats. The file should also be checked for the following:

Header Record—The file name in the Header Record should be the same as the external file name. The Transmittal Status Report displays the file name from the Header Record, not the external file name. Using the same name in both locations ensures that the file can be readily and correctly identified when technical support is needed.

FIPS/ZIP Codes—FIPS and ZIP codes are string fields, not numeric fields. Therefore, states with FIPS codes from 1 through 9 and ZIP codes that start with zero must include the leading zeros.

SEA, LEA and School IDs—SEA, LEA and School IDs, both NCES and state, are string fields, not numeric fields. Therefore, leading zeros should be used if they are part of the ID. An error will occur if IDs do not match exactly, i.e., an NCES ID number submitted as 749 but the NCES ID is 00749. Where leading zeros are meaningful, such as in NCES IDs, include the leading zeros and fill with blanks in string fields.

Percentages—Fields that contain percentages must be reported as decimals in the format 5,4; meaning there are five total digits with four digits after the decimal point. Ninety-five percent is then represented as "0.9500."

Subtotals and Grand Totals—Subtotals and grand totals must be equal to or greater than the sum of their parts. If the subtotal or grand total is less than the sum of its parts, an error is identified.

Completeness—Files do not need to contain all the schools and LEAs, but each file must contain all of the data for a specific education unit. **Subsequent submissions for a single education unit overwrite previous submissions of that file type for that education unit.** For example, submitting Grade Level Membership counts for an LEA in one transmittal file followed by the Third Grade Membership counts for the **same** LEA in another transmittal file will result in the Grade Level counts in transmittal #1 being over-written by the Grade Level counts in transmittal #2 for that LEA, resulting in only third-grade membership data in the database.

File type—The file type in the Header Record and in the File Transmit XML object must be in ALL CAPS. The ESS is set up for an exact match of file type that is case-sensitive. If a match isn't found, the file will receive a format error.

Field length—The field length cannot exceed the length specified in the non-XML file specification regardless of type of file (XML, comma-delimited, tab-delimited, or fixed). One method for checking the length of each field in a delimited file is to open the file in Excel and use the length function the application provides. Instructions on how to use this function are as follows:

1. Choose an empty field or create a new column to hold the result of the function.
2. Type "=LEN(CELL)" into the formula bar above the spreadsheet, where CELL is the target cell whose length you want to count. This cell should now contain the total number of characters in the target cell.
3. To have this formula work for every cell in the target cell's column, select the cell containing the LEN formula and copy it. Next, select the result cell's entire column and choose the paste command to have the formula apply to every cell in that column.
4. Each cell in the target cell's column should now have a corresponding cell that contains the target cell's number of characters. You can now sort the spreadsheet by this column to see if any of the counts exceed the maximum length of the field as defined in the file specification.

7.5 Submit the File

Once the SEA submitter has gathered the data and formatted the files to match the file/XML specifications, the SEA submitter logs in to the EDEN Portal to submit the data to ESS. SEA submitters will receive an acknowledgment for receipt of files and messages that request actions on the part of the SEA. The *EDEN Submission System User Guide* explains in detail how to submit files to ESS.

Performance testing on the ESS has shown that the ESS can successfully process a directory file of 300 megabytes. However, the EDEN Portal is a shared environment where processing resources are often used by more than one SEA. This can result in the system being slow. To help manage burden, SEAs should submit files as soon as they are created instead of creating many files and submitting them as a batch. Submitting outside the peak time of 8 AM to 5 PM Eastern Time is also helpful. Files can be submitted over the weekend, primarily Saturday, because the system is down for some period of time on most Sundays for routine maintenance.

ESS generates an e-mail acknowledging receipt of the file to the SEA submitter who sent the file. The SEA submitter receives messages via ESS when the file has processed successfully, is a duplicate of the most recent transmittal file or has errors. Specific error messages are not provided in the e-mail but are available in reports in ESS.

7.6 Correct Format Errors

Format edits take place when the file is first submitted to ESS. A format edit is a check that determines whether the data are in the correct file format. Format errors are the most serious error type and prevent any further processing of the file. The SEA must resubmit the file for reprocessing to correct the error.

Conditions that cause format errors include, but are not limited to:

- Absence of a Header Record.

- An improperly formatted Header Record.

- An invalid file type specified in the Header Record.

- Absence of a Carriage Return/Line Feed (CRLF) at the end of a record.

- Invalid record length.

- An invalid file format or extension (.txt, .csv or .tab) specified in the file name in the Header Record.

- A comma- or tab-delimited file that is improperly delimited (i.e., absence of a comma or tab between fields).

- Improperly formatted XML.

- Invalid XML tags.

- XML tags in a different order than specified in the schema.

Using End-of-File (EOF) characters will cause the system to look for a new record/file and result in a validation error. Instead use CRLF.

The file must be resubmitted if there are format errors.

7.7 Correct Validation Errors

Validation edits take place while the file is still in ESS. A validation edit is performed against the values in each field. A validation error occurs if the value is not included in the permitted values column in the file/XML specifications. A value that fails the edit check is always incorrect. The file will not be transferred to the staging database. The SEA must correct the data values and resubmit the entire file before processing can continue. When a file passes all validation edits, the file is transferred to the staging database.

Examples of conditions that will cause validation errors include, but are not limited to

- The file is missing one or more mandatory data groups (i.e., Agency Name or Address Line 1 in the Directory file).

- A post office box is specified in a Location Address line.

- A required count is blank or invalid.

- A numeric field contains non-numeric characters.

- An invalid Permitted Value is used.

- A required subtotal or total record is missing.

- A subtotal or total record is less than the sum of its parts.

Files must be resubmitted if there are validation errors. If, during the course of processing a file, 1,000 or more errors are identified, further processing of that file will end. For example, if an inappropriate string is included in every record in a file, the processing will stop after 1,000 errors are identified. The *Business Rules* document contains all of the validation edits in ESS.

7.8   Address Submission Errors

Submission edits take place when the file is in the staging database. These edits ensure that submitted data meet or exceed an acceptable level of reasonability by checking the values entered in a field against other similar values in the same file or across files. If a discrepancy is found (e.g., a value falls outside of the acceptable range), a submission error is issued. Any such errors that arise are reported to the SEA submitter via the Submission Error Report and the Match Error Report available for viewing and/or download from the ESS Portal. If, during the course of processing a file, 1,000 or more errors or warnings are identified, further processing of that file will end. The *Business Rules Guide* document contains all of the edits in ESS.

Submission discrepancies reported on the Submission Error Report are categorized as errors (urgent) or warnings (less urgent) depending on their level of significance and the level of urgency with which corrections need to be made.

Errors must be corrected with a resubmission. Once the error is corrected, it will no longer appear on the error report. If the data flagged as in error are actually correct, then the SEA submitter should contact the PSC to have a support representative bypass the edit for the state.

Warnings occur when data appear to be unusual or unexpected. States should review the data to determine if the data are accurate. If the data are incorrect, they should be corrected with a resubmission. If the data are correct, ED recommends that states use the explanation field of the data record to explain the data and what was done to verify that they are correct. Please note that even after explanations are provided, the warning will remain on the report.

Many of the submission edit checks in ED*Facts* are byproducts of the CCD–ED*Facts* merger project. These edits mimic match and edit checks formerly conducted by CCD. Submission errors that take the place of CCD's match edit checks are considered

critical in nature, and, therefore, errors must be corrected to ensure that data are accurate.

Other submission edits are from the transformation of the *IDEA* Section 618 tables into ED*Facts*. Additional submission edits will be added in the future.

7.9   Monitor Progress

ED*Facts* Coordinators can track their overall progress toward meeting their Submission Plans using the State Plan Execution Reports in ED*Facts*. ED*Facts* Coordinators should update the Submission Plan if delays or other problems are identified.

In addition to the reports in ED*Facts*, ESS displays four reports that can also be used to monitor progress:

- **Transmittal Status Report—**This report updates to reflect the "real time" status of every file submitted by the SEA.

- **Submission Error Report—**These reports reflect all submission discrepancies by level identified by ESS when comparing submitted data within and across files.

- **Match Error Report—**This report reflects all submission discrepancies identified by ESS when comparing submitted data to data from the previous school year.

- **Submission Progress Report—**This report reflects the last action taken by the SEA or ED with regard to a specific submission type .

Detailed information about these reports is available in the *EDEN Submission System User Guide.*

8.0   ED*Facts* Files

This section describes the four file formats that are accepted by ESS and how these files are created.

8.1   ED*Facts* File Formats

The four file formats accepted by ESS are:

- Fixed
- Comma-delimited
- Tab-delimited
- eXtensible Markup Language or XML[7]

8.1.1 Fixed format

In a fixed format file, each data element in both the Header Record and Data Record has a predefined length (for example, 15 characters) and a beginning position within the record. Thus, the software knows exactly where to find each individual field (including filler fields) within the record. When entering the specific data into the file, all fields must be filled, beginning in the Start position, even if most of the positions are blank. The suffix for fixed-format files is **.txt.**

The following rules apply to fixed-format files submitted to ESS.

- File extensions must be ".txt."

- Filler fields should be blank.

- Numeric fields must not contain commas.

- String fields should be left justified and filled with trailing blanks.

- Numeric fields (total records in file in the header record, file record number in the data record and the "count" in the data record) can be either right or left justified.

8.1.2 Comma-delimited file format

The comma-delimited file format is similar to the fixed format for order, record definitions and comments; however, rather than being a fixed length of characters, each field is separated by a comma. In instances where strings contain commas (,), there is the potential for causing incorrect processing, so use fixed or tab-delimited files when commas are included in the data. The suffix for the comma-separated file format is **.csv.**

The following rules apply to comma-delimited files submitted to ESS.

- File extension must be ".csv."

---

[7] Appendix E provides a discussion of XML and the advantages of reporting using that format.

- Except for the last field in a record, each field, whether valued or not, must be followed by a comma, including filler fields. When the last field in a data record layout is filler, the data will look like the record ends in a comma, but it ends with a blank filler field.

- Filler fields should be blank.

- For valued fields, the comma should immediately follow the permitted value.

- Data values (text or numeric) must not contain commas within the field.

- Carriage Return/Line Feed (CRLF) must immediately follow the last field in each record.

- Fields cannot exceed the maximum length.

A comma indicating the end of the last data field in the Header Record or Data Record should not be used.

ED recommends, however, using tab-delimited over comma-delimited format as imbedded commas in string fields can be problematic in comma-delimited files. Using quotes/double quotes to imbed commas in a comma-delimited format does not work in MS SQLServer.

8.1.3 Tab-delimited file format

The tab-delimited file format is similar to the fixed format for order, record definitions and comments; however, rather than being a fixed length of characters, each field is separated by a tab character. The suffix for tab-delimited file formats is **.tab**.

The following rules apply to tab-delimited files submitted to ESS:

- File extension must be ".tab."

- Except for the last field in a record, each field, whether valued or not, must be designated by a tab character, including filler fields. When the last field in a data record layout is filler, the data will look like the record ends in a tab, but it ends with a blank filler field.

- Filler fields should be blank.

- For fields that contain a value, the tab should immediately follow the data value

- Data values must NOT be enclosed in double quotes ("). Double quotes within a field will cause a Format Error.

- Numeric fields must not contain commas.

- CRLF must immediately follow the last field in each record.

- Fields cannot exceed the maximum length.

A tab indicating the end of the last data field in the Header Record or Data Record should not be used.

8.1.4 XML Files

XML files are described in appendix E. In addition to the XML specifications, SEA submitters will need the XML schemas and the non-XML file specification for the allowable field. XML style sheets are also available. The suffix for the XML file format is **.xml**.

The following rules apply to XML files submitted to ESS:

- File extension must be ".xml."

- XML tags must comply with the XML schemas.

- XML tags must be in the order specified in the schema.

- Numeric data should not contain commas.

8.2 Creating Files

Each file has two file/XML specifications. One specification provides the specifics on submitting the fixed, comma-delimited and tab-delimited files (non-XML files) and the other XML files. This section includes a description of the processes for creating non-XML and XML files.

8.2.1 Creating non-XML files

Within each non-XML file (fixed, comma-delimited and tab-delimited formats), there is a header row and the data section. One of the data items in the header row is a count of the rows in the data section. For that reason, the header section is often added to the file after the data section is compiled. The number of positions for the total records in the file header field length is 10, and the file record number field length has been changed to 10.

The data section consists of lines of plain text. The only real difference among the three non-XML file formats used by ESS is how the border between the fields is defined. The delimited file formats (comma and tab) use a comma or tab to define the fields. The fixed format uses a fixed location to define the fields.

Copy or export the data from the staging area to a text file. The process used to copy or export is dependent on the application used for staging and is beyond the scope of this workbook. The copy or export process will also determine which non-XML format to use.

The fields included and their order and structure should match the description in the relevant data record definition section of the file specification. Make sure all the "filler" fields are present.

Fields that are defined as "Number" can contain values with leading zeros or without leading zeros. Additionally, when submitting fixed format files, "Number" fields can be either right-or left-justified. This applies to the following fields: "Total Records in File" in the Header, "File Record Number" in the Data Record, and the "Count" fields in the Data Record, e.g., Teacher Count, Student Count.

If the copy or export process included a row with the field names at the top of the file, delete that line.

Once the data section is created, then the Header Record can be added to the top of the file following the description in the header record definition of the relevant file specification.

8.2.2 Creating XML files

There are two approaches to making XML submissions. The original approach essentially duplicates the non-XML file format creating a TABLETYPE object to hold the data record. The modified approach takes advantage of the power of XML schemas for data validation. Currently, an additional file transformation step is required if using the validation schema approach. When using either approach to building an XML file, remember that XML is case sensitive.

8.2.2.1 Using the basic approach

The "original" approach is the file format defined in the XML specifications. Essentially all of the formats for files that are tables have a similar XML definition. There is an AGENCY element for each education unit—SEA, LEA or school. The table data for a row are contained in a <TABLETYPE> element. Each data category is defined by a <CATEGORY> element. The <TABLETYPE> and <CATEGORY> elements have attributes to clarify what they contain.

LEA	Grade level	Race/ethnicity	Sex (membership)	Amount
1234567	01	AS	M	26
1234567	01	AS	F	28

Example: If the table above contains a portion of the LEA membership data for a state, the comparable section of an original approach XML file looks like:

```
<AGENCY FIPSSTATECODE="99" STATEAGENCYIDNUMBER="01"
        STATELEAIDNUMBER="1234567">
  <TABLETYPE TYPEABBRV="MEMBER" TOTALINDICATOR="N">
      <CATEGORY TYPE="GRADELVMEM" VALUE="01"/>
      <CATEGORY TYPE="RACEETHNICITY" VALUE="AS"/>
      <CATEGORY TYPE="SEX" VALUE="M"/>
      <AMOUNT>26</AMOUNT>
  </TABLETYPE>
  <TABLETYPE TYPEABBRV="MEMBER" TOTALINDICATOR="N">
      <CATEGORY TYPE="GRADELVMEM" VALUE="01"/>
      <CATEGORY TYPE="RACEETHNICITY" VALUE="AS"/>
      <CATEGORY TYPE="SEX" VALUE="F  "/>
      <AMOUNT>28</AMOUNT>
  </TABLETYPE>
  …
</AGENCY>
```

XML files have the advantage of being both human and computer readable. It may be easier for a person to spot an error in the XML file than a non-XML file. Most modern database systems have tools for generating XML.

8.2.2.2   Using the validation schemas

With the validation schemas approach, a special file, called an XML schema, is used that defines the legal contents of an XML file. Each of the submission files has its own validation schema. A schema defines the valid structure for a file, the valid element names, order, format, legal code values, numeric value ranges, etc. Most errors and warnings can be identified by validating an XML file prior to submission.

Using the validation schema format, the same section of the membership file would look like:

```
<Agency>
    <StateAgencyIdNumber>01</StateAgencyIdNumber>
    <StateLeaIdNumber>1234567</StateLeaIdNumber>
    <Membership>
        <Grade>01</Grade>
        <RaceEthnicity>AS</RaceEthnicity>
        <Sex>M</Sex>
        <TotalIndicator>N</TotalIndicator>
        <Amount>26</Amount>
    </Membership>
    <Membership>
        <Grade>01</Grade>
        <RaceEthnicity>AS</RaceEthnicity>
        <Sex>F</Sex>
        <TotalIndicator>N</TotalIndicator>
        <Amount>28</Amount>
    </Membership>
</Agency>
```

The validation schema knows that a <Membership> element will contain <Grade>, <RaceEthnicity>, <Sex>, <TotalIndicator> and <Amount> elements. It validates that they are in that order. It validates that "01", "AS" and "M" are legal content for the respective elements. It validates that TotalIndicator must be "N" for this combination of elements. It ensures that Amount is an integer.

An XML schema for the original layout can verify that the <TABLETYPE> element contains some number of <CATEGORY> elements and an <AMOUNT> element. It cannot verify, however, that the right combination of <TYPE>s are used or validate that the VALUE="01" is valid when TYPE="GRADELVMEM".

Currently, XML transform technology must be used to convert the validated XML file to the "original" approach XML structure—the data remain the same, but the tagging and structure of the file are modified to the acceptable format. The XML transform technology takes a style sheet (a special type of XML file) to do the transformation.

The XSL Transform process moves data from one XML representation into another XML representation. The transformation processor only needs a special XML file, called an XSL Transform Style Sheet, containing the transformation rules.

Thus, a full-featured XML schema can be defined specifically for each file submission. States can create files to be validated by the "validation schemas." An XSL Transformation can convert the file into the format accepted by the ESS. This is shown diagrammatically below.

Figure 8.2.2.2-1: XSL Transformation
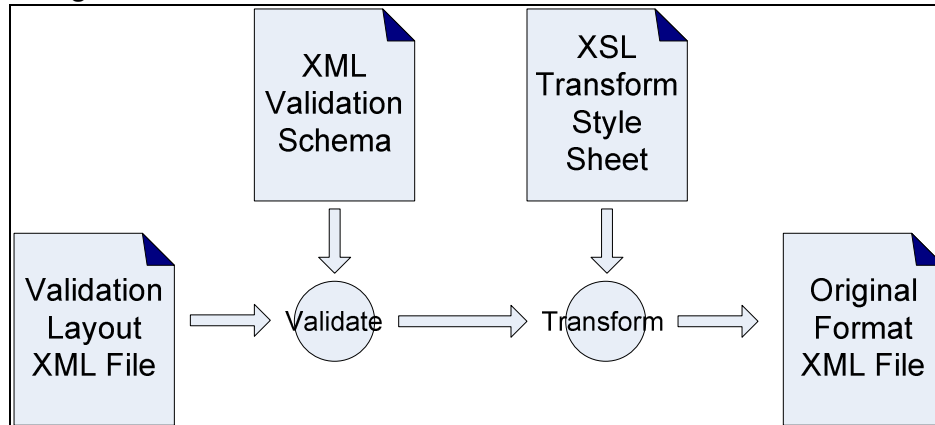
Validation Schemas and XSL Transform documents are available for each of the File Submissions. These are specific to both the file type and level. That is, there is a validation schema and XSL transform for the SEA-level membership file and a separate set for the LEA-level membership file.

Another advantage of the validation schema format is these were designed so the XML file is easier to generate from the common database applications.

Eventually, the ESS will be modified to perform the Transform. The "validation format" files will be accepted directly by the ESS.

A variety of commercial and free tools are available for editing, validating and transforming XML files. Both the validation schemas and the transformation style sheets are posted with the file/XML specifications.

A basic discussion of XML can be found in appendix E.

9.0   File/XML Specifications

The purpose of this section is to explain how the file/XML specifications are constructed so that the user can interpret the information correctly. This guidance applies specifically to the SY 2005–06, SY 2006–07 and SY 2007–08 file specifications.

For each file, two specification documents are available—one for XML files and the other for non-XML (fixed-format, comma-delimited and tab-delimited) files. File/XML-specification documents contain descriptions of the relevant data groups, permitted values and record layouts that must be used to submit the files. XML specifications begin with an "X" prefix and a three-digit number. Non-XML specifications begin with an "N" prefix and the same three-digit number as their XML counterpart. The only exception to this is N039 Grades Offered, which does not have an XML counterpart. The grades offered data are combined with the X029 Directory XML specifications.

9.1   Organization of the File/XML Specifications

Separate versions of each file specification are created for each version of ESS. In some cases, updated versions are also created to correct errors in the file specifications. Version 2.0 and subsequent updates (i.e., 2.1, 2.2, etc.) are used to build files for SY 2005–06.[8] Version 3.0 and any updates are used to build files for SY 2006–07. Version 4.0 and any updates are used to build files for SY 2007–08. The most recent updates for each version of the file specifications are on the ED*Facts* Web site, www.ed.gov/edfacts.

Each file/XML specification document for SY 2007–08 is organized as follows. Subsequent parts of this section will address each section of the specifications in greater detail.

Section	Title and Description
1.0	**Purpose**
	The purpose section gives the data group name(s) and identification number(s) included in this submission file. It also references the legal citation for collecting the data or the legacy collection the data group replaces.
	Any questions about the meaning of the data being collected can be resolved by referring to the data framework or ED*Facts* Data Set Viewer and looking up the data group name and number referenced in this section. See the *ESS User Guide* for guidance on using the data framework.

---

[8] Version 1.0 and subsequent updates (i.e., 1.1, 1.2, etc.) were used to build files for SY 2003–04 and SY 2004–05. SY 2003–04 and SY 2004–05 are disabled and are no longer collected by ESS.

1.1   Changes from the SY 2006–07 File/XML Specification

This section describes any changes that were made to this file/XML specification from the prior year.

1.2   Requirements for Submitting Data Groups

The requirements section indicates at what level (SEA, LEA, school) files are to be created and submitted. This section also contains information about whether the data apply to all schools and LEAs or only a subset. For example, graduation data are submitted only for schools and LEAs that have 12th grade.

It also contains the data requirements matrix that shows valid combinations of categories and total indicators for the file. This matrix is a very important reference for those staff creating the files.

2.0   Guidance for Submitting File[9]

This section provides guidance on any special considerations or clarification that may be necessary for the file to be submitted. The SEA submitter should review this section with the program office data steward.

2.1   Definitions

If applicable, a subsection of definitions is provided.

3.0   File Naming Convention[10]

This section describes the convention for naming the files to be uploaded to ESS. The file-naming-convention section is customized in each file/XML specification document for the portion of the name that varies according to the file being submitted.

3.1   Population Status/XML Specification Headings[11]

This section in the non-XML specifications Population Status explains how to read the column in sections 4.0, 5.0, and 6.0. The "Pop" column is used to clarify when a field is mandatory, optional or is required in accordance with 1.2-1.

This section in the XML Specification Headings section includes a headings and codes table. The last section details the Char (characteristics) column of the EDEN XML Format Codes described further in Table 9.6-1.

4.0, 5.0, 6.0 SEA, LEA and/or School File Descriptions

---

[9] File/XML specification documents for SYs 2005–06 and 2006–07 Guidance for submitting file are in section 3.
[10] File/XML specification documents for SYs 2005–06 and 2006–07 File naming convention are in section 2.
[11] This information is in section 3 for SY 2006-07 file specifications.

These section(s) each describe one level of a file—SEA, LEA or school—as appropriate for the file being described. The organization of these chapters is identical.

For the non-XML specifications, these chapters document the Header Record, then the Data Record and end with partial examples of fixed, comma and tab-delimited layouts.

XML files are composed of a hierarchy of XML objects. The sections of the File Description chapters in the XML specifications define the XML objects beginning with the innermost object and working out to the root FILETRANSMIT object.

Appendix

A few file/XML specification documents have an appendix if additional clarifying information is necessary (see for example the Directory File/XML specifications N029/X029).

9.2   Single vs. Multiple Data Group Files

The structure of the files can be generalized into two types based on whether they report multiple data groups or a single data group.

Many of the data groups consist of a single item, count, flag or characteristic per education unit. Examples of these are Address Mailing (data group 8), Title I School Status (data group 22), School Poverty Percentage (data group 31) and School Totals (data group 454). These are often gathered with similar data groups in a file that has a single record per education unit but multiple data groups per record. Files of this type are mostly confined to the directory and status files.

The majority of the files collect a single data group, but that data group consists of a table or tables of multiple records per education unit. These data groups may contain multiple levels of aggregation and categorical reporting. For example, the LEA level membership file collects for each LEA

- Membership counts by each grade level by race/ethnicity group by sex.
- Subtotal counts of membership by grade level by race/ethnicity group.
- Subtotal counts of membership by grade level by sex.
- Subtotal counts of membership by race/ethnicity group by sex.
- Subtotal counts of membership by grade level.
- A grand total membership count.

The LEA Membership file (N/X052) collects just this one data group but that consists of 18 records per grade level per LEA (10 detail and 8 subtotal records) plus an additional 11 records per LEA (10 subtotal and 1 grand total record).

Section 1.0 "Purpose" in the file/XML specification is the place to look to identify the number and type of data groups collected in each file.

9.3   Data Requirement Matrix

For the single data group files (the files that contain tables), Section 1.2 "Requirements for submitting this data group" of the file/XML specification contains the data requirements matrix. This matrix describes which data categories are being collected and how the counts are to be aggregated. It also indicates which records are considered subtotals of others in the submission and whether the submission should have a grand total for each education unit.

The data requirements matrix from the Membership file (file/XML specification 52) is shown below.

Example Table 1.2-1: Membership Table, Required Categories and Applicable Totals

Category set	Table name	Grade level (membership)	Race / ethnicity	Sex (membership)	Total indicator	Comments
Category Set A	**MEMBER**	X	X	X	N	Enrollment by Grade Level (Membership), Race/ Ethnicity and Sex (Membership)
Subtotal 1	**MEMBER**	X	X		Y	Subtotal by Grade Level (Membership) and Race/Ethnicity
Subtotal 2	**MEMBER**	X		X	Y	Subtotal by Grade Level (Membership) and Sex (Membership)
Subtotal 3	**MEMBER**		X	X	Y	Subtotal by Race/Ethnicity and Sex (Membership)
Subtotal 4	**MEMBER**	X			Y	Subtotal by Grade Level (Membership)
Grand Total	**MEMBER**				Y	Grand Total

Each row in the matrix defines a set of rows to be submitted in the file. The first row in the example indicates membership must be reported by grade level, race/ethnicity and sex (i.e., fifth-grade Asian males). The second row indicates a membership subtotal reported by grade level and race/ethnicity (i.e., fifth-grade Asians).

The "N" under Total Indicator on row one of the data requirements matrix indicates that this is the lowest level of detail ED*Facts* will collect. A state may have summarized individual records to get this, but for ED*Facts* purposes, this is neither a subtotal nor grand total but a detail record.

The "Y" on row 2 indicates that ED*Facts* considers this record to be a subtotal or grand total of the detail records which are reported in the line 1 data records in this example. While processing the file, ESS will sum the detail data by grade level and race/ethnicity. Any amount reported in the second set of data records must be greater than or equal to the calculated detail total. That is, the reported total third-grade Hispanics from the

second data set must be greater than or equal to the sum of all the third-grade Hispanics (Male and Female) reported in the detail data.

The data requirements matrix shows which records are considered detail (total indicator = N) and which records are considered subtotals or grand totals of those details (total indicator = Y). Subtotal records contain one or more categories. Grand total records never contain categories.

9.4   File Naming Convention

States will be submitting many files. Having a naming convention simplifies the process of identifying

- The contents of the file.
- The file type.
- The level of the file (i.e., SEA, LEA, or school).
- The submitting state.

Each file name is limited to 25 characters (including the file extension). The format is:

ss*lev*FILETYPES*vvvvvvv*.ext

The elements of the file name are as follows:

ss	=	the 2-character USPS State Code value (See appendix E)
lev	=	the 3-character file level as follows: • For SEA files, specify **SEA** • For LEA files, specify **LEA** • For school files, specify **SCH**
FILETYPES	=	the 9-character file name abbreviation. (For example, the file name abbreviation for directory Files is DIRECTORY.) See the *EDEN Submission System User Guide* for a complete list of file name abbreviations. The file name abbreviation is also noted in each file/XML specification in section 3.0
vvvvvvv	=	any combination of up to 7 characters as specified by the SEA (when a file is split into multiple files as discussed in workbook section 7.3.2, this portion of the file name should differentiate among the multiple files)
.ext	=	the file extension as follows: • For fixed-format files, specify **.txt** • For comma-delimited files, specify **.csv** • For tab-delimited files, specify **.tab** • For eXtensible Markup Language (XML) files, specify **.xml**

Examples

Example of fixed-format SEA directory information file name for the state of Euphoria (EU) using version numbering as the final seven characters:

<div align="center">EUSEADIRECTORYVer0001.txt</div>

Example of comma-delimited LEA directory information file name for the state of Euphoria (EU) using a full date/version combination as the final seven characters:

<div align="center">EULEADIRECTORY110305A.csv</div>

Example of tab-delimited school directory information file name for the state of Euphoria (EU) using a month/day/version combination as the final seven characters:

<div align="center">EUSCHDIRECTORY1213V01.tab</div>

Example of an XML school directory information file name for the state of Euphoria (EU) using the file creator's initials, date and version as the final seven characters:

<div align="center">EUSCHDIRECTORYsk0214a.xml</div>

9.5   Non-XML File Description

Sections 4, 5 and/or 6 of the file/XML specifications each describe the layout for a specific file reporting level—SEA, LEA and/or school. All three sections are present when all three file levels are to be reported. Only sections 4 and 5 are present if data are only to be reported for SEA and LEA or for only LEA and school.

For non-XML files, sections 4.1, 5.1 and/or 6.1 define the Header Record layout. Sections 4.2, 5.2 and/or 6.2 define the contents and layout of the data records. Sections 4.3, 5.3 and/or 6.3 show partial example files for their respective file levels.

9.5.1 Header Record description

The Header Record is the first record in every non-XML file. The Header Record provides information as to the file type, number of data records in the file, file name, file identifier and file reporting period.

The structure of the Header Record is essentially the same for all files. It comprises

- **File Type**—Defines the type of file being submitted. This is a text string up to 50 characters long and must match the permitted value specified in the file specification.

- **Record Count**—The total number of records in the data section. Do not count the header record. Because the number of data records must be known to complete this portion of the header record, the header record is often created and added after the data section is created.

- **File Name**—The name of the file uploaded to the ESS, including the file extension. If the file was zipped prior to submission, this should be the

uncompressed file name and extension. The file name should follow the file-naming convention defined above. The Transmittal Status Report displays the name contained in the header record, not the external file name. For this reason, it is important that the name stored in the header record match the actual name of the file. Using the same file name in both locations ensures that the file can be readily and correctly identified when technical support is needed. Multiple transmittal files of the same type of file should each have a unique file name.

- **File Identifier—**This is a text string up to 32 characters long that uniquely identifies or describes the file. It can be simply a repeat of the file name, or it can further identify the file for SEA purposes. Additional information might include the creation date, file creator, version number, a brief description of the data on the file, etc. The file identifier is also shown on the Transmittal Status Report page of the ESS. Since this column can be sorted, starting the file identifier with a creator User ID can make it easy for multiple submitters to locate and track the status of their submissions.

- **File Reporting Period—**This is the academic school year for which the data are being reported. The required format is either "CCYY-CCYY" or CCYY CCYY." The proper format for the 2007–08 school year would be either "2007–2008" or "2007 2008."

- **Filler—**The filler field is blank. The purpose of the filler field at the end of the header record is to make it the same overall length as the detail record.

9.5.2 Data records description

The data records are submitted immediately following the header records and provide detail data as well as subtotals and grand totals. Each record appears on its own line in the file, and each ends with a CRLF.

The data section of the file is described in a table. An example is shown below.

Data element name	Start position	Length	Type	Pop	Definition/comments	Permitted values
File record number	1	6	Number	M	A sequential number assigned by the state that is unique to each row entry within the file.	
ID 559 FIPS state code	7	2	String	M	The two-digit Federal Information Processing Standards (FIPS) for the state, District of Columbia and the possessions and freely associated areas of the United States (i.e., Puerto Rico).	For a list of valid FIPS state codes, refer to the ED*Facts* Workbook.
… (rows omitted)	…	…	…	…	…	…

Data element name	Start position	Length	Type	Pop	Definition/comments	Permitted values
Table Name	45	20	String	M	The official student enrollment, including students both present and absent, excluding duplicate counts of students within a specific school or students whose membership is reported by another school or LEA.	**MEMBER**
… (rows omitted)	…	…	…	…	…	…
Sex (membership)	95	15	String	A	The concept describing the biological traits that distinguish the males and females of a species.	**M** = Male **F** = Female **MISSING**
… (rows omitted)	…	…	…	…	…	…
Explanation	111	200	String	O	Submitted files will be processed….	
Student count	311	10	Number	M	If the count provided is either a subtotal or grand total, it must be equal to or greater than the sum of its parts.	
Carriage return/line feed (CRLF)	321	1		M		

The columns in this table are:

- **Data Element Name—**The name of the data in each field

- **Start Position—**For a fixed-format file, the text character position in which the data should begin. In this sample, the file record number starts in character position 1 and ends in character position 6. The FIPS state code must be in character positions 7 and 8. This information is not used for comma- or tab-delimited files.

- **Length—**The maximum number of characters allowed for this field. Many fields will be shorter than the maximum, but the maximum length of a field applies to all files regardless of the type (.txt, .csv, .tab, or .xml).

- **Type—**The type of data, text string or numbers, that will be entered in the field

- **Pop[12] (Population status)—**The population status indicates when the field must be populated. The following population statuses are used:
  - o   M - Mandatory, this field must always be populated.
  - o   A - This field is populated in accordance with table 1.2-1.
  - o   O - Optional, data in this field are optional.

- **Definition/Comments—**The definition of the data group in each field. This information should be used in concert with section 2.0 (Guidance) in the file specification.

- **Permitted values—**The range of values that are allowed for the field. In the example, for this file the Table Name must be the text string "MEMBER." The category Sex (Membership) can have one of three values: "M", "F", "MISSING."

For every non-XML file except the directory files, the first five fields in the data records are the same. They are:

- **File Record Number—**A number that uniquely identifies each data row in the file submission. Typically these are simply the sequential row number.

- **FIPS State Code—**The two-digit Federal Information Processing Standards (FIPS) for the state, District of Columbia and the possessions and freely associated areas of the United States (i.e., Puerto Rico)

- **State Education Agency Number—**The state agency number for all SEAs is "01."

- **State LEA Identifier—**The identifier assigned to an LEA by the SEA. Also known as State LEA ID. This ID is a required field in LEA- and school-level files. For SEA-level files, a blank filler field is located in this position. An LEA directory record must exist with a matching state LEA ID.

- **State School Identifier—**The identifier assigned to a school by the SEA. Also known as state school identification number. This ID is a required field in school-level files. For SEA- and LEA-level files, a blank filler field is located in this position. A school directory record must exist with a matching SEA, LEA and school ID.

The remaining fields of the data record contain the data specific to the file.

9.5.3 Example layouts

Following the data record definition section for each level, the file specifications include partial sample layouts for that level. The fixed-format is shown first followed by sample comma-delimited and tab-delimited files.

---

[12] Previously, this column was "Mandatory/Optional." For SY 2005–06, and the initial versions of the SY 2006–07 file specifications, this column was "(M)andatory/(O)ptional." The column had either an M or an O. The M indicated that the field must always be populated. The O indicated either that the field was populated in accordance with table 1.2-1 or was optional.

9.6   XML File Descriptions

Sections 4, 5 and/or 6 of the XML specifications each describe the layout for a specific file reporting level—SEA, LEA and/or school.

Each XML document has a FILETRANSMISSION at its root. The FILETRANSMISSION object contains an AGENCY object for each of the education units in the file. For the single data group files, multiple TABLETYPE objects hold each of the data records within the AGENCY object. A CATEGORY object defines each of the category sets for the data record.

Sections 4, 5 and/or 6 have subsections that describe these XML objects starting with the most internal object and ending with the FILETRANSMISSION object.

Included for each XML object are the following XML characteristics:

- Element—name of the XML element tag

- Attribute—name of the XML attribute tag

- Category value—name of the category

- Char—the XML element or attribute characteristic

- Definition/comments—definition and additional comments related to formats or other business rules

- Permitted values—the range of values allowed for the field

The Char (characteristics) column in the XML format matrices accepts the following codes:

Table 9.6-1: EDEN XML Format Codes

Code	Characteristic
M	Mandatory Element/Attribute
O	Optional
C	Conditionally Required
MR	Mandatory and Repeatable Element
OR	Optional and Repeatable Element
CR	Conditional and Repeatable Element

9.6.1   File transmission XML object

Each XML file has a FILETRANSMISSION object at its root. The FILETRANSMISSION object has attributes holding the file type, file ID and school year. These are the same as those in the non-XML header record.

The FILETRANSMISSION object contains AGENCY elements that hold the data for each of the education units in the submission.

9.6.2 Agency XML object

The AGENCY objects contain the data for all the education units in the submission. An AGENCY object has attributes for the FIPS state code and state agency number ('01' for all SEAs). For an LEA-level file, the AGENCY object has an attribute for the state LEA ID. For a school-level file, the AGENCY object has an attribute for the state LEA ID and one for the state school ID.

All of the data for an education unit are in elements within the AGENCY object. There is one AGENCY object for each education unit in the file.

If the file is a single data group file, the individual records are stored in TABLETYPE objects. If the file is a multiple data group file, the data are contained in a series of elements and attributes defined specifically for the data groups.

9.6.3 Table type XML object

The TABLETYPE object stores the data from the data requirements matrix (table 1.2-1 in section 1.2) for single data group files. It has attributes for the table type and the total indicator flag. The category information is stored in a series of CATEGORY tags with the respective count in an AMOUNT element.

9.6.4 Category XML object

The CATEGORY object has two attributes: a TYPE attribute that names the category set and a VALUE attribute to store the category set code value.

10.0 Directory Files

This section describes the directory used in ED*Facts* and the file used to submit directory records. It also discusses changes to the directory. This section is intended to provide a basic outline of the directory in ED*Facts* and does not substitute for the directory file/XML specification when building and submitting directory files.

10.1 Hierarchy

ED*Facts* maintains a three-tier directory of education units in each of the 50 states, the District of Columbia, Puerto Rico, the outlying areas (i.e., American Samoa, Guam, Northern Marianas and the Virgin Islands), Bureau of Indian Education and Department of Defense schools. The three tiers are (1) the SEAs) (2) the LEAs, referred to as "agencies" in NCES' CCD and (3) the schools.

In some states, there are schools outside of the traditional three-tier system such as state schools, youth facilities and correctional institutions. States have been able to accommodate this additional tier by creating a virtual LEA that includes some or all of these additional schools. States should contact the PSC to determine if this solution is appropriate. All uses of these virtual LEAs should be documented in the metadata.

10.1.1 State education agency

An SEA is the agency of the state charged with primary responsibility for coordinating and supervising public elementary and secondary instruction, including the setting of standards for instructional programs. The SEA is the state agency that administers federal grant programs under the *Elementary and Secondary Education Act*.

10.1.2 Local education agency

An LEA is a governmental administrative unit at the local level that exists primarily to operate schools or to contract for educational services. These units may or may not be coterminous with county, city or town boundaries. In addition to traditional school districts, LEAs include

a) State-operated agencies that provide instruction or education support services for students in any of grades prekindergarten through 12 or ungraded. This may include, but is not limited to, the SEA; corrections, health and human services, juvenile justice and other state agencies; the governor's office; and the State Board of Education. SEAs should not report agencies operated by the Department of Defense Education Authority or the Bureau of Indian Education.

b) County or city governments that operate schools when these agencies perform the function of an LEA.

c) Agencies operating schools that provide instruction to grades in addition to prekindergarten through 12 or ungraded. For example, a local authority administering both community colleges and vocational-technical schools that

serve students in grades 10 through 12 or a board responsible for laboratory schools in universities.

d) Agencies administering charter schools—Charter schools under the authority of the school district in which they are situated are reported with that school district. Charter schools administered by some agency other than a school district are reported with that administering agency; for example, State Board of Charter Schools, State Charter School Licensing Authority, Board of Regents, etc.

e) Separate charter schools—Charter schools that are not under the authority of any agency are reported as their own education agencies because every school in the CCD universe must have an associated education agency. In these instances, the charter school is included in the hierarchy twice: once as an LEA and once as a school.

f) Agencies that do not operate schools—Report supervisory unions and education service agencies that provide services but do not operate schools. For example, supervisory unions (typically found in New England) provide administrative services for multiple school districts. Education service agencies usually provide administrative, professional development, research and evaluation or data processing services to multiple districts or may employ instructional staff that provides services in multiple districts. An example of this type of agency would be an education service agency that contracts to provide special education teachers and student support staff to several districts. In addition to supervisory unions, Regional Education Service Agencies (RESA) and Boards of Cooperative Education Services (BOCES) fall into this category.

g) LEAs that "tuition out" their students—Report agencies that are responsible for providing education to students residing within their boundaries even if the agency sends all of its students to other agencies or to private schools for their education; for example, a small township that finds it more feasible to send its students to a neighboring LEA than to operate schools itself.

The inclusion of non-traditional school districts results in some LEAs that do not report students.

10.1.3 School

A school (for the purpose of ED*Facts*) is an institution that provides educational services and:

- Has one or more grade groups (prekindergarten through 12) or is ungraded;

- Has one or more teachers;

- Is located in one or more buildings;

- Has assigned administrator(s);

- Receives public funds as its primary support and;

- Is operated by an education agency.

For purposes of this definition, "public funds" include federal, state and local public funds. "Located in a building" does not preclude virtual schools since the administrators and teachers are located in a building somewhere. An "education agency" is not limited to the SEA or LEA, but can include other agencies (example, corrections or health and human services) charged with providing public education services. In addition to what is traditionally considered a public school, schools include

a) Schools that have grades in addition to prekindergarten through 12 and ungraded, if one or more of grades prekindergarten through 12 are included. *Example:* Report a Vocational-Technical Center that offers grades 11, 12 and a post-secondary Associate's Degree (sometimes considered grades 13 and 14). *Exception:* Do not report postsecondary schools that offer prekindergarten services only; for example, a community college that includes a daycare center.

b) Schools that do not have students and staff because they are temporarily not in operation (but are expected to be in operation within 2 years). *Example*: Report schools that are closed for renovations and future schools. Report schools that have no students simply because there are no students that year (example, a "one-room schoolhouse" without students that year).

c) Schools that do not report students and/or staff because these are included in the data for some other school of record. *Example:* Report a Vocational-Technical Center that serves multiple high schools even if students attending the Vocational-Technical Center are included in the membership counts for these other high schools.

d) Charter schools. These operate under a public charter, with public funds and are public schools. Report charter schools regardless of what agency administers them.

e) Schools operated by agencies other than an LEA. Schools that fall within the definition of a school should be reported regardless of the agency responsible for them. This includes, but is not limited to, state-operated schools such as those operated by the SEA, corrections, health and human services, juvenile justice, other state agencies, the governor's office, or the State Board of Education. *Exception:* Schools under the administration of the Bureau of Indian Education and the Department of Defense Education Authority are reported by these agencies. These schools are not reported by the SEA of the state in which the school is located.

f) Schools that share the same campus or building, if they have separate administrators. *Example:* a single building with a principal for grades prekindergarten through 7 and a principal for grades 8 through 12 would be reported as two schools. Note that cases in which programs are spread across

several buildings or campuses are reported as a single school if they are under the direction of a single administrator (principal or equivalent).

g) Virtual schools. Although students are not physically present, there is a physical facility associated with the transmission of courses.

Inactive, closed, future and shared-time schools would not be expected to report students or AYP status.

10.2 Directory Data

For each education unit, a directory record is maintained that includes unique identifiers (i.e., name, identification numbers, location), contact information (i.e., addresses, phone number, Web site) and descriptive information (i.e., education type, operational status). The chart in appendix D of this workbook contains all the directory data by the level in the hierarchy that reports it. While not a part of the directory, the grades-offered data should be maintained in conjunction with the directory data.

10.2.1 Unique identifiers

Each education unit has two IDs: NCES and state identifier. These IDs are used as the primary identifiers in ED*Facts*. IDs are strings and not numbers. ED*Facts* considers "01" and "1" to be different IDs. Therefore, SEA submitters must be careful to maintain leading zeros in IDs.

In addition, the identification numbers are used to link the education units to each other in the hierarchy. For example, the record for a school includes the identification number of the LEA to which the school belongs.

The primary identifiers are used in coordination with the education unit name and addresses to identify duplicates. It is imperative that each education unit be included only once in the database.

For SY 2006–07, additional edits on directory data were added to prevent duplications and assist with issuing NCES IDs to new LEAs and schools.

The capability to report changes to state identifiers was new for SY 2006–07. The SY 2006–07 directory allows for the identification of a new state identifier and a prior state identifier. ED*Facts* will associate the institution to the new state identifier as of the supplied effective date.

New for SY 2007-08 is the school-year-specific directory. In the past, the state was required to submit only updates to its directory. The directory information rolled over from one school year to the next. As of SY 2007–08, states must submit a complete LEA and school directory each school year. State agency information does not need to be submitted each school year. Updates to the state agency directory file can be processed on an as-needed bases.

Each version of the directory specification identifies the SY to be included in the header. The SY specified in the header of the file will be used to determine what SY the changes and additions are processed against. Regardless of effective dates specified in

the submitted file, a directory file will only affect data for the SY specified in the header. For example, submitting a change using a SY 2006–07 (i.e., version 3.x) directory file will only affect data for SY 2006–07. To make the same change in a different SY requires that the change be submitted in a file formatted for that SY.

10.2.2 Contact information

Basic contact information (address, telephone number and Web site) is also included in the directory for each education unit. Both mailing and location addresses are requested. As noted above, addresses are used to identify duplicates. Addresses are also used to obtain locale codes from the U.S. Bureau of the Census.

10.2.3 Descriptive data

The directory also includes descriptive data. The most important descriptive data are operational statuses.

Operational status
ED*Facts* provides data for the CCD, including the directory universe for the pertinent school year. A critical field in the CCD directory universe is the operational status of the schools and districts. There are different permitted values for schools and LEAs.

When states add a school or district, they may report the operational status as "NEW," "ADDED" or "FUTURE" as appropriate.

10.2.4 Grades offered

Grades offered is part of the directory. For XML files, grades-offered data are submitted with directory data. For non-XML files, grades-offered data are submitted as a separate file. Grades-offered records must exist for any LEA or school that has students enrolled. While it is expected that membership counts will be provided for any grade offered, it is understood that an exact match may not occur in all instances. In some instances, a grade offered may not have membership counts. However, if membership data are submitted for a grade, that grade must exist in grades offered for that education unit. If an LEA has no students in any grade, only the applicable directory data and a single record for grades offered that indicates there are no grades offered need to be submitted. Indicate no grades offered using the permitted value "NoGrades." No membership or other student-related data should be submitted for an education unit that has no students.

10.3 Directory Files

The directory file/XML specification contains detailed information on submitting directory records. This section summarizes the overall policies for the directory files.

10.3.1 School-year specific

Because the directory data are now school-year specific, the directory data from SY 2006–07 will not roll over to SY 2007–08. SEAs need to submit all LEAs and schools at the beginning of each school year.

10.3.2 Requirement for directory records

Directory files for LEAs and schools must be submitted before other file groups are submitted.

- A school (SCH) directory record will not be accepted until the LEA directory record associated with that school is in the system.

- A membership record (or any other non-directory file) for a school will not be accepted unless a record for that school (SCH) exists in the directory file.

- When supervisory unions are reported, a new supervisory union must be submitted in one file, and the institutions that are part of that supervisory union must be submitted in a subsequent file.

Since acceptance of all other files is dependent upon having complete directory data for every education unit, the SEA should confirm that all directory data are complete and accurate before any other type of file is submitted.

10.3.3 Closing of LEAs and schools

LEAs and schools are never deleted from the directory when they close. Instead, the operational status changes from OPEN to CLOSED through the directory file. LEAs and schools that close need only be submitted in the directory file once to indicate their change in operational status.

10.3.4 Changing the directory for a prior year

Each year NCES "locks" the directory for the CCD collection. While SEAs can make changes to the directory after the "locking," those changes could have unanticipated consequences. As a result, SEAs need to inform PSC in advance of making changes to a prior year directory.

10.3.5 Number one cause for duplicate LEAs and schools

Adding or removing leading zeros in state IDs and NCES IDs is the primary cause of duplicate LEAs and schools. Please be careful when you submit IDs with leading zeros.

10.4 Migrant Student Exchange System (MSIX)

The directory data in ED*Facts* is the Department's master or reference data for LEAs and schools. As a consequence, the directory data on LEAs and schools used in MSIX is supplied by ED*Facts*. The directory data cannot be updated through MSIX. Instead, the directory data on LEAs and schools must be updated in ED*Facts* and then transferred to MSIX.

11.0 Missing, Not Applicable, Not Collected, Zero

This section describes how to use Missing, Not Applicable, Not Collected, or Zero so that the data submitted accurately reflect the situation that exists in the SEA. The methodology below is an attempt to balance clear reporting against burden. The file specifications contain additional information about how this methodology applies to a specific set of data.

11.1 Missing

Missing means the SEA collects the data, but the data are not currently available. For example, the SEA may collect a particular set of data on a paper survey, but the data are not converted to an electronic form yet and thus are not available to submit. In this case, these data are currently missing for the entire state. As another example, a particular LEA may not have returned a survey to the SEA. In this case, the data are currently missing for that LEA.

- When Missing is a valid option, it is included in the Permitted Values field in the File/XML specifications.

- For string fields, "MISSING" is the Permitted Value.

- For numeric fields, the value for Missing is –1 (minus 1).

11.1.1 Data Group is missing for the entire state

If the data group is in its own file, then the file should be submitted when the data are available. For example, if an SEA does not have the SY 2007–08 data ready for data group #74 Children with Disabilities (*IDEA*) School Age (N/X002) by the February due date, then the SEA should not submit the file with prior year data or preliminary data. Instead, the SEA should submit the file when the data are finalized. The State Submission Plan should reflect the date that finalized data can be submitted. When the data have established due dates (i.e., *IDEA* Part B), the SEA should notify ED via PSC if finalized data will not be not available by the due date.

If the data group is not in its own file, then the file can be submitted using Missing or the approved numeric value for Missing. When the data become available, the SEA resubmits the entire file.

11.1.2 Data Group is missing for a particular education unit

The record for that particular education unit is not submitted. When the data become available, the SEA submits a file with the records for only the education units with the missing data.

11.1.3 Missing data or missing category label

A category of data can be collected by itself or it can be part of a broader category set group. Depending on what data are missing, the situations are handled differently. For example, some data groups request data by sex while others report sex by grade by race/ethnicity group. Reporting missing sex data in these two situations is different.

In the first case, the data are missing but the data labels are known. Consequently, the file would have a record for male with an amount of -1 and a record for female with an amount of -1.

If a data group requests data by grade by sex by race/ethnicity and the state only has grade by race/ethnicity, then the data are reported in the grade by race/ethnicity category set and the grade by sex by race/ethnicity category set is not submitted. If there is not a grade by race/ethnicity category set, the state reports grade by sex by race data using MISSING as the sex value.

11.2 Not Applicable

Not Applicable means the situation does not occur in the SEA, LEA or school. This is generally not a Permitted Value for most data groups or fields.

For non-table data groups (e.g., status data groups), when not applicable is an option it is included in the Permitted Values column of the file/XML specifications. In most cases, the Not Applicable situations have been specifically designed for a particular data group. For example, data group 18 (Grades Offered) includes the permitted "NOGRADES = No grades are offered at this educational unit" instead of not applicable.

For table data groups, Not Applicable is invalid. This is because the table data groups are counts of students or staff, and the result of a count is usually zero, one or more rather than Not Applicable. If an SEA identifies a situation where Not Applicable appears to be needed, the SEA should contact PSC.

11.3 Not Collected

Not Collected means the SEA does not collect the data group at all, does not collect the data group at a particular education unit level or does not collect a particular category for that data group.

For non-table data groups (e.g., status data groups), when Not Collected is an option, it is included in the Permitted Values column of the file/XML specifications.

For table data groups, if the data group is Not Collected at all by the SEA or is Not Collected at a particular education unit level, the table is not submitted. Since table data groups are generally in a file by themselves, the SEA does not send a file in for the data groups that are Not Collected. The data group should be reported as Not Collected through the State Submission Plan. See Section 11.5, When Submissions Are Not Required, below.

For table data groups where a particular category is Not Collected, the SEA should use the Not Collected Permitted Value for that category. For example, if race/ethnicity is not collected for student performance, it is expected that race/ethnicity is not collected for all grades within a school or LEA, and race/ethnicity should have a value of NOTCOLLECT. Submitting student performance data with race/ethnicity in some grades and not in others within the same category set results in an error.

If the data group contains more than one table (category set), Missing or Not Collected can be used for a category in one table and have other values in the second table. For

example, if children with disabilities data are collected by race/ethnicity and disability category, but are not collected by race/ethnicity and sex, race/ethnicity can be supplied in the first set of data while not collected is used for race/ethnicity in the second set.

11.4 Zeros

In numeric fields, zero (0) should be used when the education unit has no instances of a requested count, i.e., no student or staff to report within the scope of a specific data group. For instance, if a school has no Limited English Proficient (LEP) students, the correct response in the student performance files for those values is "0".

11.5 When Submissions Are Not Required

If the lack of students in a particular group would result in a submission with ALL zeros, the submission does not need to be created. For example, if there are no LEP students in a school, no submission for N/X049 LEP–Assessed in Native Languages (data group #272) needs to be created for that school. If you have just one student who needs to be reported for an education unit, then the submission needs to be created for that education unit, and all values need to be valid permitted values (a number, missing or zero).

If an entire file for a particular file/XML specification is missing, not applicable, or not collected, then the file should *not* be sent in. For example, if an SEA does not have native language assessments, the SEA would not submit files for the file/XML specification N/X049 LEP–Assessed in Native Languages (data group #272).

12.0 Metadata

Metadata are data collected to explain other data. For ED*Facts*, metadata include data that are needed to accurately interpret other data. For example, SEAs submit student performance data from state assessments. The crosswalk of the state performance levels to federal proficiency is metadata necessary for ED*Facts* to report state data accurately. ED plans to collect metadata in such a way that states need only to update them when they change. Below are examples of the types of metadata that will be collected to explain the data.

State Definitions. Some of the definitions of ED*Facts* Data Groups include references to state definitions. To understand the data, ED*Facts* will collect the state definitions.

General SEA Information. There are a number of basic components of SEAs that differ from one state to the next, for example

- *School Year/Start of School Year*—SEAs will be asked what 12-month period of time defines a school year.
- *Charter School Legislation*—SEAs will be asked when the state approved charter school legislation. ED will pre-populate based on information already collected.
- *Special Education Count Date*—SEAs will be asked on what specific date the child count required by *IDEA* was conducted.
- *Directory*—SEAs will be asked to update a general description of the education units in the state that was developed during previous state site visits.

Assessment. Each state has in place a state assessment system that is unique to that state. In order to properly interpret the data submitted on this subject, users of the data need information beyond the counts of students by performance level, including:

- *Performance levels*—SEAs will be asked how each performance level maps to "proficient" in *NCLB.*

- *Instrument*—SEAs will be asked for a description of the assessment instruments such as academic subjects and grades covered as well as the basic test administrative procedures and processes for handling the results.

- *Accommodations*—SEAs will be asked what accommodations are available/provided to students and under what circumstances.

- *Timing*—SEAs will be asked when the assessment was conducted, e.g., November or April.

AYP Reporting. As is the case with the previous examples, AYP reporting includes components that are state specific and, when known, help to inform interpretation of state data submissions, including

- ***Full Academic Year***—SEAs will be asked what period of time a student is at school for inclusion in AYP reporting.

- ***Targets***—SEAs will be asked for the target percentage of students proficient under *NCLB* by school year.

Graduation Rates. States calculate graduation rates differently. In order to properly interpret the data submitted, users of the data need to know how the graduation rate was calculated, including whether the average freshman graduation rate was used.

Safety and Discipline. To interpret data for the firearms incidents, the Department needs to know how each state implemented requirements for expulsion for 1 year and referrals to the justice system because of firearms violations. The Department also needs to know how the state law handles alternative placement, whether it is required or encouraged, and whether it is supported with state funds. Each year, SEAs will indicate whether changes occurred and, if changes were made, what those changes were. The Department also needs the state definition of violence (with and without injury), weapons possession, alcohol related, and illicit drug related.

Migrant Child Count. The migrant child count will be used to allocate funds of the Migrant Education Program, Title I Part C. The Department needs information about the states' calculation and validation procedures, including information systems used, the data collection and management procedures, the methods used to count the children and the quality control procedures.

Public School Choice and Supplement Services. Data are collected on the students who applied for public school choice and supplemental education services (SES). The definition of applied is needed to interpret the data. Data are also collected on who receives SES. The threshold for inclusion as receiving SES is needed to interpret the data.

13.0 Transformations of Legacy Collections

This section describes some of the transformations of legacy collections.

13.1 Non-Fiscal Common Core of Data (CCD)

To streamline and simplify data submission processes for the states, the NCES CCD non-fiscal survey is being incorporated into ED*Facts*. School year 2006–07 was the first year of a 2-year transformation. During SY 2006–07, data were submitted through ED*Facts* while ID generation and edit checks were conducted by the CCD system. Match and edit reports generated by the CCD system were available through ED*Facts*. During SY 2007–08, the data will again be submitted through ED*Facts*. In addition, ED*Facts* will assign new NCES IDs and conduct the edit checks. During the SY 2007–08 collection, the CCD system connectivity with ED*Facts* will remain to allow SEAs to finish submitting their SY 2006–07 data, including downloading edits reports.

For SY 2008–09 and beyond, ED*Facts* will collect the CCD non-fiscal data, generate the NCES IDs, and conduct the edit checks. NCES will take the data submitted through EDFacts and create the public CCD files and other publications.

In the sections below are more details on the process for SY 2007–08 and future school years and a list of files to be submitted.

13.1.1 Process for Submitting CCD Non-Fiscal Files

The following is the process for SY 2007–08 and future school years.

- SEAs submit the files listed in the crosswalk in section 13.1.2 according to the submission calendar in section 2.2.[13]

- SEAs begin by submitting the directory file. ED*Facts* will assign new NCES IDs. ED*Facts* will also apply the edits related to the directory data. The results of these edits will be displayed on a report referred to as the match report. SEAs must fix all errors on the match report before ED*Facts* will produce the reports with the results of the edits on non-directory data referred to as edit reports.

- Once the match reports are cleared, ED*Facts* will produce the edit report. Edits that rely on prior year data will use the SY 2006–07 data in ED*Facts*. If the final data for SY 2006–07 were changed by NCES instead of the SEA resubmitting the data through ED*Facts*, the data used in the comparison will not be the CCD's final data. Data are not "pushed" back from NCES into ED*Facts*.

- Match reports and edits for the early files must be finished by April 2008. Edits for the remaining files must be finished by February 2009.

---

[13] Prior to integration with ED*Facts*, the CCD non-fiscal files were submitted annually. ED*Facts* files are submitted in collection periods as explained in section 2.2.

- SEAs (both ED*Facts* and CCD coordinators) should contact the ED*Facts* PSC with any questions, concerns or suggestions for improvement.

13.1.2 Crosswalk of CCD data items to ED*Facts* data groups

Below is a list of the specific early, middle and closeout files that need to be submitted to ESS for CCD.

For the SEA level the following files need to be submitted:

Early collection period:
- N/X029—Directory

- N/X052—Membership

- N/X059—Staff FTE

Closeout collection period:
- N/X040—Graduates/Completers

Table 13.1-1 ED*Facts* Files for CCD State Non-Fiscal File

CCD State Non-Fiscal File					
CCD field name	CCD field position	CCD field description	ED*Facts* data group	ED*Facts* file specification	ED*Facts* file schedule
SEANAME	001 - 060	Name of State Education Agency	7 - State Agency Name	N/X029– Directory	Early
STREET	061 - 090	Mailing Address	8 - State Education Agency Mailing Address (Street portion)	N/X029 – Directory	Early
CITY	091 - 120	City	8 - State Education Agency Mailing Address (City portion)	N/X029 – Directory	Early
ZIP	121 - 125	ZIP Code (5 Digit)	8 - State Education Agency Mailing Address (Zip Code portion)	N/X029 – Directory	Early
ZIP4	126 - 129	ZIP + 4 Code	8 - State Education Agency Mailing Address (Zip Code Plus 4 portion)	N/X029 – Directory	Early
PHONE	130 - 139	Telephone Number	10 - State Education Agency Phone Number	N/X029 – Directory	Early
PKTCH - UGTCH	140 - 169	Teachers	528 - Staff FTE Tables	N/X059 – Staff FTE	Early
TOTTCH	170 - 175	Total FTE Teachers	Derived from 528 - Staff FTE Tables	N/X059 – Staff FTE	Early
AIDES - OTHSUP	176 - 253	Staff	528 - Staff FTE Tables	N/X059 – Staff FTE	Early
TOTGUI	(200 - 205)	Total Guidance Counselors	Derived from 528 - Staff FTE Tables	N/X059 – Staff FTE	Early

CCD State Non-Fiscal File					
CCD field name	**CCD field position**	**CCD field description**	**ED*Facts* data group**	**ED*Facts* file specification**	**ED*Facts* file schedule**
PK - WHUG	254 - 973	Students by Grade, Race/Ethnicity, and by Gender	39 - Membership Tables (Category Set A)	N/X052 – Membership	Early
MEMBER	974 - 981	Total Students	39 - Membership Tables	N/X052 – Membership	Early
AMTOT-WHTOT	982 - 1021	Total students by Race/Ethnicity	Derived from 39 - Membership Tables	N/X052 – Membership	Early
REGDIP-WHDIP	1022 - 1069	Diploma Recipients (Total and by Race/Ethnicity)	Derived from 306 – Graduates/Completer Tables	N/X040 – Graduates/Completers	Close-out
EQUIV-WHEQUIV	1070 - 1117	High School Equivalency Recipients (Total and by Race/Ethnicity)	Derived from 306 – Graduates/Completer Tables	N/X040 – Graduates/Completers	Close-out
OTHCOM - WHOHC	1118 - 1165	Other High School Completers (Total and by Race/Ethnicity)	Derived from 306 – Graduates/Completer Tables	N/X040 – Graduates/Completers	Close-out

For the LEA level, the following files need to be submitted:
Early collection period:

- N/X029 – Directory

- N039 – Grades Offered (these data are in X029 for XML submitters)

- N/X002 – Children with Disabilities (*IDEA*) - School Age

- N/X089 – Children with Disabilities (*IDEA*) - Early Childhood

- N/X052 – Membership

- N/X059 – Staff FTE

Middle collection period:
- N/X046 – LEP – Program

Close-out collection period:
- N/X032 – Dropout

- N/X040 – Graduates/Completers

- N/X122 – MEP Students - Eligible and Served Summer/Intersession

Table 13.1-2 ED*Facts* Files for CCD Agency Non-Fiscal File

CCD Agency File					
CCD field name	**CCD field position**	**CCD field description**	**ED*Facts* data group**	**ED*Facts* file specification**	**ED*Facts* file schedule**

			CCD Agency File		
CCD field name	**CCD field position**	**CCD field description**	**ED*Facts* data group**	**ED*Facts* file specification**	**ED*Facts* file schedule**
LEAID	001 - 007	NCES Education Agency ID	1 - NCES LEA Identifier	N/X029 – Directory	Early
STID	008 - 021	State Education Agency ID	4 - State LEA Identifier	N/X029 – Directory	Early
NAME	022 - 081	Name of Education Agency	7 - LEA Name	N/X029 – Directory	Early
PHONE	082 - 091	Telephone Number (Area Code and Phone Number)	10 - LEA Phone Number	N/X029 – Directory	Early
MSTREE	092 - 121	Mailing Address - Street	8 - LEA Address Mailing (Street portion)	N/X029 – Directory	Early
MCITY	122 - 151	Mailing Address – City	8 - LEA Address Mailing (City portion)	N/X029 – Directory	Early
MSTATE	152 - 153	Mailing Address - State (P.O. Abbreviation)	8 - LEA Address Mailing (USPS State Abbreviation portion)	N/X029 – Directory	Early
MZIP	154 - 158	Mailing Address - ZIP (digits 1 thru 5)	8 - LEA Address Mailing (Zip Code portion)	N/X029 – Directory	Early
MZIP4	159 - 162	Mailing Address – ZIP4 (digits 6 thru 9)	8 - LEA Address Mailing (Zip Plus 4 portion)	N/X029 – Directory	Early
LSTREE	163 - 192	Physical Location Address - Street	9 - LEA Address Location (Street Portion)	N/X029 – Directory	Early
LCITY	193 - 222	Physical Location Address - City	9 - LEA Address Location (City Portion)	N/X029 – Directory	Early
LSTATE	223 - 224	Physical Location Address - State (P.O. Abbreviation)	9 - LEA Address Location (USPS State Abbreviation portion)	N/X029 – Directory	Early
LZIP	225 - 229	Physical Location Address - ZIP (digits 1 thru 5)	9 - LEA Address Location (ZIP code portion)	N/X029 – Directory	Early
LZIP4	230 - 233	Physical Location Address - ZIP4 (digits 6 thru 9)	9 - LEA Address Location (Zip Plus 4 portion)	N/X029 – Directory	Early
AGTYP	234 - 234	Education Agency Type Code	453 - LEA Educational Agency Type	N/X029 – Directory	Early
UNION	235 - 237	Supervisory Union ID Number	551 - LEA Supervisory Union Identification Number	N/X029 – Directory	Early
CONAME	238 - 267	County Name	572 - County Name	Derived	Early
CONUM	268 - 272	FIPS County Code	12 - FIPS County Code	Derived	Early
BOUND	273 - 273	Operational Status Code	16 - Operational Status	N/X029 – Directory	Early
GSLO	274 - 275	Grade Span Offered -	Derived from 18 - Grades	X029 – Directory	Early

CCD Agency File					
CCD field name	**CCD field position**	**CCD field description**	**ED*Facts* data group**	**ED*Facts* file specification**	**ED*Facts* file schedule**
		Lowest grade	Offered	N039 – Grades Offered	
GSHI	276 - 277	Grade Span Offered - Highest Grade	Derived from 18 - Grades Offered	X029 –Directory N039 – Grades Offered	Early
UG	278 - 284	Ungraded Students	39 - Student Membership Tables	N/X052 – Membership	Early
PK12	285 - 291	PK-12 Students	Derived from 39 - Student Membership Tables	N/X052 – Membership	Early
MIGRNT	292 - 298	Migrant Students Served in a Summer Program	635 - MEP Students Eligible and Served Summer/Intersession Tables	N/X122 – MEP Students – Eligible and Served Summer / Intersession	Close-out
SPECED	299 - 305	Special Education - Individual Education Program (IEP)	Derived from 74 - Children with Disabilities (*IDEA*) School-Age Tables and 613 - Children with Disabilities (*IDEA*) Early Childhood Tables	N/X002 – Children with Disabilities (*IDEA*) –School Age N/X089 – Children with Disabilities (*IDEA*) – Early Childhood	Early
ELL	306 - 312	English Language Learner	123 - LEP Demographic Tables	N/X046 – LEP Program	Middle
PKTCH-UGTCH	313 - 347	Teachers	528 - Staff FTE Tables	N/X059 – Staff FTE	Early
TOTTCH	348 - 354	Total FTE Teachers	Derived from 528 - Staff FTE Tables	N/X059 – Staff FTE	Early
AIDES-OTHSUP	355 - 445	Staff	528 - Staff FTE Tables	N/X059 – Staff FTE	Early
TOTGUI	(383 – 389)	Total Guidance Counselors	Derived from 528 - Staff FTE Tables	N/X059 – Staff FTE	Early
TOTDPL	446 - 451	Total Diploma Recipients	Derived from 306 - Graduates/ Completer Tables	N/X040 – Graduates/ Completers	Close-out
AMDPLM-WHDPLU	452 - 541	Diploma Recipients by Race/Ethnicity and Gender	306 - Graduates/ Completer Tables	N/X040 – Graduates/ Completers	Close-out
TOTOHC	542 - 547	Total Other High School Completers	Derived from 306 - Graduates/ Completer Tables	N/X040 – Graduates/ Completers	Close-out
AMOHCM-WHOHCU	548 - 637	Other High School Completers by Race/ Ethnicity and Gender	306 - Graduates/ Completer Tables	N/X040 – Graduates/ Completers	Close-out
TOTD07-WHD12U	638 - 1213	Dropouts by Grade, Race/Ethnicity, and Gender	326 - Dropouts Tables	N/X032 – Dropout	Close-out

For the school level, the following files need to be submitted:

Early collection period:
- N/X029 – Directory

- N039 – Grades Offered (these data are in X029 for XML submitters)

- N/X129 – CCD

- N/X052 – Membership

- N/X033 – Free and Reduced Price Lunch

End collection period:
- N/X132 – School

Table 13.1-3 ED*Facts* **Files for CCD School Non-Fiscal File**

CCD School File					
CCD field name	**CCD field position**	**CCD field description**	**ED***Facts* **data group**	**ED***Facts* **file specification**	**ED***Facts* **file schedule**
LEAID	0001 - 0007	NCES Education Agency ID	1 - NCES LEA Identifier	N/X029 – Directory	Early
STID	0008 - 0021	State Education Agency ID	4 - State LEA Identifier	N/X029 – Directory	Early
SCHNO	0022 - 0026	NCES School ID	529 - NCES School Identifier	N/X029 – Directory	Early
SEASCH	0027 - 0046	State School ID	5 - State School Identifier	N/X029 – Directory	Early
LEANM	0047 - 0106	Name of Education Agency	7 - LEA Name	N/X029 – Directory (LEA file)	Early
SCHNAM	0107 - 0156	Name of School	7 - School Name	N/X029 – Directory	Early
PHONE	0157 - 0166	Telephone Number (Area Code and Phone Number)	10 - School Phone Number	N/X029 – Directory	Early
MSTREE	0167 - 0196	Mailing Address - Street	8 - School Address Mailing (Street portion)	N/X029 – Directory	Early
MCITY	0197 - 0226	Mailing Address - City	8 - School Address Mailing (City portion)	N/X029 – Directory	Early
MSTATE	0227 - 0228	Mailing Address - State (P.O. Abbreviation)	8 - School Address Mailing (USPS State Abbreviation portion)	N/X029 – Directory	Early
MZIP	0229 - 0233	Mailing Address - ZIP (digits 1 thru 5)	8 - School Address Mailing (Zip code portion)	N/X029 – Directory	Early
MZIP4	0234 - 0237	Mailing Address - ZIP4 (digits 6 thru 9)	8 - School Address Mailing (Zip Plus 4 portion)	N/X029 – Directory	Early

CCD School File					
CCD field name	**CCD field position**	**CCD field description**	**ED*Facts* data group**	**ED*Facts* file specification**	**ED*Facts* file schedule**
LSTREE	0238 - 0267	Physical Location Address - Street	9 - School Address Location (Street portion)	N/X029 – Directory	Early
LCITY	0268 - 0297	Physical Location Address - City	9 - School Address Location (City portion)	N/X029 – Directory	Early
LSTATE	0298 - 0299	Physical Location Address - State (P.O. Abbreviation)	9 - School Address Location (USPS State Abbreviation portion)	N/X029 – Directory	Early
LZIP	0300 - 0304	Physical Location Address - ZIP (digits 1 thru 5)	9 - School Address Location (ZIP code portion)	N/X029 – Directory	Early
LZIP4	0305 - 0308	Physical Location Address - ZIP4 (digits 6 thru 9)	9 - School Address Location (Zip Plus 4 portion)	N/X029 – Directory	Early
SCHTYP	0309 – 0309	School Type Code	21 - School Type	N/X029 – Directory	Early
STATUS	0310 - 0310	Operational Status Code	531 - School Operational Status	N/X029 – Directory	Early
GSLO	0311 - 0312	Grade Span Offered - Lowest grade	Derived from 18 - Grades Offered	X029 –Directory N039 – Grades Offered	Early
GSHI	0313 - 0314	Grade Span Offered - Highest grade	Derived from 18 - Grades Offered	X029 –Directory N039 – Grades Offered	Early
TITLEI	0315 - 0315	Title I Eligible School	22 - Title I School Status	N/X129 – CCD	Early
STITLI	0316 - 0316	School-wide Title I	22 - Title I School Status	N/X129 – CCD	Early
MAGNET	0317 - 0317	Magnet School	24 - Magnet Status	N/X129 – CCD	Early
CHARTR	0318 - 0318	Charter School	27 - Charter Status	N/X029 – Directory	Early
SHARED	0319 - 0319	Shared Time	573 - Shared Time Status	N/X129 – CCD	Early
FTE	0920 - 0324	Classroom Teachers	644 - Teachers FTE	N/X129 – CCD	Early
FRELCH	0325 - 0328	Free Lunch Eligible Students	565 - Free and Reduced Price Lunch Table	N/X033 – Free and Reduced Price Lunch	Early
REDLCH	0329 - 0332	Reduced-Price Lunch Eligible Students	565 - Free and Reduced Price Lunch Table	N/X033 – Free and Reduced Price Lunch	Early
FLE	0333 - 0336	Total of Free and Reduced-Price Lunch Eligible	565 - Free and Reduced Price Lunch Table	N/X033 – Free and Reduced Price Lunch	Early

CCD School File					
CCD field name	**CCD field position**	**CCD field description**	**ED*Facts* data group**	**ED*Facts* file specification**	**ED*Facts* file schedule**
MIGRNT	0337 - 0340	Migrant Students Enrolled in Previous Year	110 - MEP Students Eligible Regular School Year Tables	N/X132 – School	End
TOTPK-WHALU	0341 – 1364	Students by Grade, Race / Ethnicity, and by Gender	39 - Membership	N/X052 – Membership	Early

13.2 Consolidated State Performance Report (CSPR)

ED is again collecting the SY 2006–07 CSPR through a Web-based submission tool that shares an opening Web page with ESS. ED has included two levels of the CSPR pre-populated with data from ESS: (1) pre-population by ESS exclusively (i.e., participation in *NCLB* state assessments, school and districts accountability and participation in Title I, Part A programs) and (2) initial pre-population by ESS, with opportunity for manual entry by the SEA (i.e., academic achievement, school choice and supplemental services, graduation rates, homeless children and youths and migrant child count). For ESS data initially pre-populating the state's CSPR, the SEA has two options: (1) to accept data pre-populated from ESS or (2) overwrite the pre-populated data.

The SY 2006–07 CSPR is collected in two parts. Part I has been certified. Part II was collected according to the following schedule:

- January 22, 2008—Submission tool opened for CSPR Part II

- February 22, 2008—Submission tool closeed for CSPR Part II. All SEA submissions must be entered and certified.

After each part is closed, the Office of Elementary and Secondary Education (OESE) and other Program Offices in ED will review the submissions for data quality. SEAs should expect questions about their Part I data in mid-January and about their Part II data in mid-March. For each review period, SEAs will respond to questions and submit a newly certified CSPR, as needed. Prompt attention to the questions will enable ED to publish the CSPRs in a timely manner.

ED expects to collect data for the SY 2007–08 CSPR according to a similar schedule. ED plans to increase the amount of pre-population of data from ESS in the SY 2007–08 CSPR.

13.3 *Individuals with Disabilities Education Act* (*IDEA*) Section 618 Reporting

ED is continuing the transformation of Section 618 data collection from the previous legacy system (DANS) to ED*Facts.* A majority of states are already approved to submit Exiting (Table 4; N/X009), Child Count and Educational Environments (Tables 1 and 3;

N/X002 & N/X089) data exclusively through ED*Facts*. Therefore, three additional SY 2007–08 collections are being included for congruency analyses. These collections are Discipline (Table 5; N/X005, 006, 007, 088, 143, 144), Staffing (Table 2; N/X070, 099, 112) and Assessment (Table 6; N/X003, 004, 093).

The SY 2007–08 SEA data reporting dates for these collections are

> Discipline—November 1, 2008
> Staffing—November 1, 2008
> Assessment—February 1, 2009

ED is currently working with the Office of Special Education Programs to establish an alternative deadline for the submission of the SY 2007–08 LEA- and school-level files.

IDEA data edits have been included in the ESS, as of version 3.5.

13.4 *Gun Free Schools Act* Report

ED is also transitioning the *Gun Free Schools Act (GFSA)* data collection to ED*Facts*. Nine states opted to submit their SY 2005–06 *Gun Free Schools Act (GFSA)* report through ED*Facts*. The Department anticipates additional states will do so for their SY 2006–07 *GFSA* report. In order to complete most of the numeric sections of the report, the following file specifications and data groups must be submitted:

• N/X086 (Students Involved with Firearms); Data Group 596.

• N/X094 (Firearm Incidents); Data Group 601.

• N/X131 and N/X132 (GFSA Reporting Status); Data Group 603.

In addition, there are questions that require text responses. States will be instructed on how to provide these text responses.

13.5 Charter School Program Reporting

SEAs can receive grants from the Department to establish charter schools. SEAs report to the Department the name, state and NCES identification numbers and addresses of charter schools that receive subgrants under this program. ED*Facts* is then used by the Program Office to obtain the demographic and reporting data for those schools for reports to Congress and others.

Appendix A: ED*Facts* Glossary

Term	Description
AMAO	Annual Measurable Achievement Objective
AYP	Adequate Yearly Progress
Category	In the Data Framework and file/XML specifications, the breakouts or dimensions in data group tables (i.e., Sex, Grade Level). Formerly Disaggregation Categories.
Category set	A valid combination of Categories that can be reported for a Table. The Category Set defines the lowest level value of a table (a cell). For example, in the Membership Table, the category set is race/ethnicity, sex and grade level. The lowest level value for the Membership Table is the number of students for an education unit belonging to one racial/ethnic group, of male or female sex and in a specific grade (i.e., the number of third-grade Asian females).
Comment	In the Data Framework, specific business rules and other meaningful information related to the data group.
CCD	Common Core of Data collected by the National Center for Education Statistics (NCES)
DANS	Office of Special Education Program's Data Analysis System
Data collection	Process used by the U.S. Department of Education for collecting data from state education agencies (SEAs).
Data element	One specific piece of data to be included in the submission (otherwise known as a variable or field in other types of data collections).
Data group	A specific aggregation (i.e., a group) of related data that are stored in ED*Facts* to satisfy the specific information need of one or more ED program offices. A data group can be a single data element such as AYP status or a table of data such as the children with disabilities exiting special education tables. Both represent a conceptual piece of information.
Data record	Detailed information, in a specified order and format, for an education unit.
Data repository	Location in the EDEN Submission System where cleaned and edited files are stored. Data become the responsibility of ED and become available for review by Subject Matter Experts, i.e., ED Program Office staff.
Data topic	In the Data Framework, a categorization of data groups that describes what type of data are represented (i.e., student, staff, program, directory).
Data staging database	Located in the EDEN Submission System where files are received for Reasonability Edits and explanation review by ED.
ED*Facts*	A collaborative effort among the U.S. Department of Education, state education agencies and industry partners to centralize state-reported K-12 educational performance data into one federally coordinated, data repository located in the U.S. Department of Education.
ED*Facts* Coordinator	Official SEA contact for ED*Facts* and EDEN Submission System. This person is responsible for ensuring that the files are submitted, that errors are corrected and that files are approved for submission to the Data Repository; serves as the SEA contact person for completing the state Submission Plan.
EDEN	Education Data Exchange Network, a centralized, coordinated repository

Term	Description
	of state-reported, K-12, educational data residing at the U.S. Department of Education
Edit	Often referred to as a business rule. An indication of a problem in the format or contents of a file. ED*Facts* includes Format, Validation and Reasonability edits.
Education unit	Education entity such as SEA, LEA, or school.
Education unit level	Tier in the education hierarchy—school, LEA, or SEA level.
Error type	Category of error (format or validation error) in the data file that prevents complete transmission of the file.
File	See Transmittal File.
File extension	Ending portion of a file name that denotes the format used for this file (.txt, .tab, .csv, or .xml). Although .zip files may be submitted, when the file is unzipped it must create a file with one of the four permitted file extensions.
File format	One of four acceptable formats for submission of files. They include fixed length, comma-delimited, tab-delimited and XML.
File identifier	A 32-character string, including any combination of standard characters that is used to further identify a transmittal file.
File level	state = SEA, local education agency (including school districts and supervisory unions) = LEA, school = SCH
File/XML specification	A document that outlines the files, record layout and data elements in the record layouts.
FIPS	Federal Information Processing Standards; standards and guidelines developed by the National Institute of Standards and Technology (NIST) for federal computer systems. The list of FIPS State Codes is provided in appendix C.
Format error	A significant error in the format of the file requiring correction and resubmission by the SEA before the data are processed.
GFSA	*Gun-Free Schools Act*
Header record	Required for every file, the first record in every file. A header record provides file type, file name, file identifier and file reporting period.
IDEA	*Individuals with Disabilities Education Act*
LEA	Local education agency
LEP	Limited English proficient, also known as English language learner
MEP	Migrant Education Program
NCES	National Center for Education Statistics
NCLB	*No Child Left Behind Act of 2001*
NIST	National Institute of Standards and Technology, a non-regulatory federal agency within the U.S. Commerce Department's Technology Administration. NIST's mission is to develop and promote measurement, standards and technology to enhance productivity, facilitate trade and improve the quality of life.
OMB	U.S. Office of Management and Budget—the approval authority for all data collections; administers and enforces the *Paperwork Reduction Act of 1995*.

Term	Description
Permitted values	The valid values for a data element.
Pre-populate	To preload information into a survey request based on the data in the repository.
REAP	Rural Education Achievement Program
Record	Row within a fixed, comma-delimited, or tab-delimited file. This term does not relate to XML files.
Record layout	Description of the specific data elements and the order of elements per row within a flat file.
Reporting period	The specific date or period of time for which data are reported.
SEA	State education agency
SEA submitter(s)	Person(s) the SEA has approved to submit files and access the secure part of the EDEN Portal. Each submitter receives a personal USER ID and PASSWORD from ED.
Schema	A format specification for an XML file. Some schemas may be used for multiple transmittal files.
Start position	For the fixed-file format only, the starting position within the row of the element.
Steward	The Principal Office within ED that has the primary responsibility for the data group.
Submission	A set of related data belonging to only one education unit (i.e., one school or one LEA), for one school year and for one file. For example, Membership data for a single school for the school year 2003–04, or current year status data for an LEA for the school year 2004–05.
SWP	School-wide program for Title I
Table type	Table data groups such as Membership or Students Tested, often used to qualify a count or category.
TAS	Targeted Assistance Schools for Title I
Transmission	The process of transferring a file from the SEA to the EDEN System.
Transmittal (file)	A physical file of data to be transferred. A transmittal file may consist of one or many submissions for a given Education Unit Level (i.e., more than one school). Large files may be broken into subparts for transmittal, if preferred by the submitter.
Transmittal status report	"Real time" display of the status of every file sent to the EDEN Submission System from the Data Receipt and Processing System through to the Data Repository.
Validation error	Significant error in the data quality of the file requiring correction and resubmission by the SEA before the data are processed.
Warning	Possible error in data quality requiring an explanation and/or correction by the SEA before further processing.
XML	Extensible Markup Language, a universal format for preparing and posting documents and data to the Web. Allows users to be more flexible and accurate in describing information. Also used for passing information between computing systems, which are otherwise unable to communicate.

Appendix B: ED*Facts* Data Framework

This appendix provides an overview of the ED*Facts* data framework.

The data framework is a conceptual model that is used to define the data requirements for ED*Facts*. It is the means for describing the data to be collected through ED*Facts* during the paperwork clearance process. The data framework is also the means to describe data available in ED*Facts* for analysis and reporting. To help you understand the data framework, this document illustrates the structure, categorization and data exchange of the data group "Membership Table."

Structure

A data group is a specific aggregation (i.e., group) of related data that is stored in ED*Facts* to satisfy the specific information need of one or more ED Program Offices. Information on the structures of data groups is available in the data framework in ESS and in the ED*Facts* Data Set Viewer.

Every data group has a name, identification number, definition, code set or format, reporting period and reporting levels. The basic components of an example data group are

- **Name—**Membership Table

- **ID—**#39

- **Definition—**The official unduplicated student enrollment, including students both present and absent, excluding duplicate counts of students within a specific school or LEA or students whose membership is reported by another school or LEA

- **Code Set or Format—**Number

- **Reporting Period—**October 1

- **Levels—**State, LEA and school

Some data groups, like the example above, also have category sets. A category set is the grouping that an SEA or other data supplier uses to aggregate data before the SEA or other data supplier sends the data to ED. The example data group includes a category set of grade level (membership), race/ethnicity, and sex. The category set can be illustrated by the following table:

Membership by race/ethnicity, by sex, and by grade levels										
Grade Levels	American Indian/ Alaskan Native - Male	American Indian/ Alaskan Native – Female	Asian or Pacific Islander - Male	Asian or Pacific Islander – Female	Black (not Hispanic) – Male	Black (not Hispanic) – Female	Hispanic – Male	Hispanic – Female	White (not Hispanic) - Male	White (not Hispanic) - Female
K	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										

The category set concept can also be illustrated by the following diagram.

Grade Level (Membership) (K through 12) ◄──► Race/Ethnicity (OMB 5) ◄──► Sex (M or F)
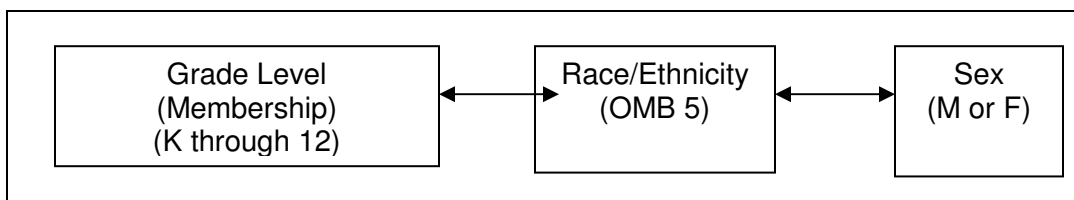
Each grade level (K through 12) is reported by each race/ethnicity (five in total) at the same time by each sex (male and female). The result if all grade levels are used is 130 numbers (13 grade levels permitted values * 5 race/ethnicity permitted values * 2 sex permitted values).

The Membership Table also includes the following subtotals:
- Subtotal 1—Grade Level (Membership), Race/Ethnicity

- Subtotal 2—Grade Level (Membership), Sex (Membership)

- Subtotal 3—Sex (Membership), Race/Ethnicity

- Subtotal 4—Grade Level (Membership)

The following tables illustrate the concept behind subtotal 1 by "grade level, race/ethnicity." Subtotal 1 combines the columns indicated in the diagram below:

Subtotal 1 by grade level, race/ethnicity										
Grade Levels	American Indian/Alaskan Native - Male	American Indian/Alaskan Native – Female	Asian or Pacific Islander - Male	Asian or Pacific Islander – Female	Black (not Hispanic) – Male	Black (not Hispanic) – Female	Hispanic – Male	Hispanic – Female	White (not Hispanic) - Male	White (not Hispanic) - Female
K	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										

The combined columns create the gray columns below. The gray columns below are subtotal 1. (AM = American Indian/Alaskan Native, AS = Asian or Pacific Islander, M = Male, and F = Female).

Subtotal 1 – Grade Level (Membership), Race/Ethnicity							
Grade Level	AM M	AM F	Total AM	AS M	AS F	Total AS	
K							
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							

The Membership Table Data Group also includes a grand total. The grand total is the gray box illustrated by the following diagram.

Grand Total - Membership									
Grade Level	Race/	Ethnicity	By	Gender					Total
Total									Grand Total

Categorization

The data categorization is done by sorting or grouping the data groups by similarities in content. Each data group is classified in three different ways: subject, scope and type. The data group Membership Table is categorized as

- Subject–Student Groups
- Scope–All Students
- Type–Background

The data framework includes all three different categorizations. The ED*Facts* Data Set Viewer includes only the subject categorization. The subject categorizations are directory, education unit, finance, program, staff, student, survey and online collection.

Data Exchange

Data are transmitted from the data supplier to ED using either the EDEN Submission System (ESS) or the EDEN Survey Tool (EST). The ESS is used by SEAs to submit data. The data for the Membership Table is available in all SEAs; therefore, this data group is collected using the ESS. Since the Membership Table is a table structure data group, it is its own file: Membership, X/N 052. The file specifications are available through the ED*Facts* Data Set Viewer or through the ED*Facts* Web site—
www.ed.gov/edfacts

Summary

Below is how the Membership Table appears in the OMB package for clearance. This presentation is similar to how it appears in the data framework and the ED*Facts* Data Set Viewer.

Data Group
Membership Table

ID: 39

Section: Student   **Reporting period:** October 1

Definition The official unduplicated student enrollment, including students both present and absent, excluding duplicate counts of students within a specific school or LEA or students whose membership is reported by another school or LEA.

Code set or format: Integer

Category A Grade Level (Membership), Race/Ethnicity, Sex (Membership)
Category B
Category C
Category D
Category E
Category F
Category G

Subtotal 1 Grade Level (Membership), Race/Ethnicity
Subtotal 2 Grade Level (Membership), Sex (Membership)
Subtotal 3 Sex (Membership), Race/Ethnicity
Subtotal 4 Grade Level (Membership)

Steward: NCES

Comment:

Grand Total: ☑   **Privacy:** ☑   **State Defined:**

Reporting Level
   School: ☑   **District/LEA:** ☑   **State/SEA:** ☑

The box "privacy" is checked when data are collected about students or staff that could result in small cell sizes, requiring suppression when ED*Facts* reports the data to maintain individual privacy. The box "state defined" is checked when the data definition is dependent on a state definition.

Appendix C: ED*Facts* Data Reporting

This appendix provides an introduction to ED*Facts* data reporting.

Use of Data

The primary customers for ED*Facts* education data are the program managers and analysts at ED. ED*Facts* data are used to evaluate the effectiveness and efficiency of federal education programs with the intent to improve program management and budgetary focus on those federal education programs that provide the best education outcomes for the nation's students and their families. State and local education agencies can use this information to evaluate their education status and progress. Eventually the public, including parents and students, will also have access to this information and be able to analyze and display information about the condition of education in their neighborhoods. The public data file will be modified to ensure that all of the necessary privacy and security requirements are completely enforced and meet the requirements of the *Family Educational Rights and Privacy Act (FERPA)*[14].

Sources of ED*Facts* Data

Although the primary source of ED*Facts* data is the data that SEAs submit through the EDEN Submission System (ESS), some data are obtained from other sources based on data provided by the SEA[15]. Data provided by other sources are listed in the following table along with their source and the SEA data on which they are based:

Table C-1: Data Sources for ED*Facts*

Data group ID	Data group	Potential source	Data group based on
17	Locale (Code)	NCES/Census	Address Location
14	Geographic Location	Census	Address Location
13	Congressional District Number	Census	Address Location
6	DUNS Number	Dunn & Bradstreet	Education Entity Name
572	County Name	NIST FIPS Publications	Address Location
12	FIPS County Code	NIST FIPS Publications	Address Location
582	Title I District Status	U.S Department of Education's records on Title I allocations	NCES IDs

[Source: File/XML specifications]

---

[14] *The Family Educational Rights and Privacy Act (FERPA)* (20 U.S.C. § 1232g; 34 CFR Part 99) is a federal law that protects the privacy of student education records. The law applies to all schools that receive funds under an applicable program of the U.S. Department of Education. For additional information see: http://www.ed.gov/policy/gen/guid/fpco/ferpa/index.html.

[15] ED recommends that ED*Facts* Coordinators maintain records of their SEA data sources for the data they submit to EDEN. This additional information will be critically important not only for future EDEN data collection efforts but also if ED has questions regarding a specific data file. This information is particularly useful when the data source is outside the ED*Facts* Coordinator's office (such as a specific SEA program office) and should include the contact person and office where the data came from, the type of data received (i.e., format, etc), the date they provided the information, and any other relevant notes.

Partnership with State Education Agencies
ED has worked with SEA representatives to implement and continuously improve ED*Facts*.

Although the early stages of ED*Facts* implementation require some duplicative reporting, the existing data collections will be retired as soon as SEAs demonstrate that they are submitting comparable data through the ESS. ED*Facts* produces three important benefits: a reduction in reporting burden on the SEAs and others, better quality and more complete data available for use by ED and the SEAs and more timely utilization of data.

Data Quality
Data quality is measured not only in its accuracy, but also in its completeness and timeliness. It is therefore critical that data be submitted to ED as soon as they are available for reporting and, as accurately as possible, for all education units available to the SEA. ED will perform validity, consistency and reasonability checks on all data submitted through the ESS. Additionally, ED conducts periodic comparative analyses of data submitted both through the ESS and through the existing legacy collection.

Appendix D: ED*Facts* Directory Data Groups

	SEA		LEA		School	
	ID#	Name	ID#	Name	ID#	Name
Unique Identifiers	559	FIPS State Code	559	FIPS State Code	559	FIPS State Code
	7	Education Entity Name	7	Education Entity Name	7	Education Entity Name
	6	DUNS Number*	6	DUNS Number*		
	570	State Agency Number				
			551	Supervisory Union Identifier		
			4	LEA Identifier (State)	4	LEA Identifier (State)
					5	School Identifier (State)
			1	LEA Identifier (NCES)	1	LEA Identifier (NCES)
					529	School Identifier (NCES)
Contact information	9	Address Location	9	Address Location	9	Address Location
	8	Address Mailing	8	Address Mailing	8	Address Mailing
	10	Telephone - Education Entity	10	Telephone - Education Entity	10	Telephone - Education Entity
	11	Web Site Address	11	Web Site Address	11	Web Site Address
	458	Chief State School Officer Contact Information				
Descriptive information			12	FIPS County Code*	12	FIPS County Code*
			572	County Name*	572	County Name*
			16	LEA Operational Status	531	School Operational Status
			453	Education Agency Type	21	School Type
			17	Locale*	17	Locale*
			18	Grades Offered	18	Grades Offered
			14	Geographic Location*	14	Geographic Location*
			13	Congressional District Number*	13	Congressional District Number*
			27	Charter Status	27	Charter Status
			582	Title 1 District Status*		
			669	Out of State Indicator	669	Out of State Indicator

* ED plans to obtain these data groups from sources other than the SEA.

Appendix E: Introduction to XML

XML stands for eXtensible Markup Language and is becoming the preferred format for exchanging data between disparate data systems.

XML documents look similar to HTML documents—the HyperText Markup Language upon which the World Wide Web is constructed. Both documents consist of data that have been "tagged."

Unlike HTML, XML is designed to describe the structure and semantics of the data, not its formatting. While HTML has predefined tags, XML does not.

Listed below is a sample of XML:

```
<?xml version="1.0" ?>
<note>
 <to>Jane</to>
 <from>Doug</from>
 <topic status="hot">Reminder</topic>
 <message>Don't forget the Membership files are
due next week</message>
</note>
```

The first line is the **XML declaration**. It is an optional line stating what version of XML is in use (normally version 1.0) and may also contain information about character encoding and external dependencies.

The remainder of the sample consists of nested *element*s, some of which have *attribute*s and *content*. An **element** typically consists of two tags, a *start tag* and an *end tag*, possibly surrounding text and other elements. The **start tag** consists of a name surrounded by angle brackets, like "<from>"; the **end tag** consists of the same name surrounded by angle brackets, but with a forward slash preceding the name, like "</from>". Element names are case-sensitive.

In addition to content, an element can contain **attributes**—name-value pairs included in the start tag after the element name. Attribute values must always be quoted, using single or double quotes, and each attribute name should appear only once in any element. In the example, the *topic* element has one attribute *status* which has a value of "hot."

The element's **content** is everything that appears between the start tag and the end tag. The content of the *to* element is just the text "Jane" while the contents of the *note* element are the four elements *to*, *from*, *topic* and *message*.

XML requires that elements be properly nested—both the start and ending tags for an element must be between the start and ending tags of its containing element. It is an error to start one element, start a second element, and then end the first before ending the second.

XML provides special syntax for representing an element with empty content. Such an element may only have attributes. Instead of writing a start tag followed immediately by an end tag, a document may contain the **empty element tag** where a slash *follows* the element name. The following two examples are functionally equivalent:

```
<CATEGORY TYPE="SEX" VALUE="M"></CATEGORY>
<CATEGORY TYPE="SEX" VALUE="M" />
```

Every XML document must have exactly one top-level root element. In the example, *note* is the root element.

There are two levels of "Correctness" for XML documents. An XML document is **well-formed** if it follows the rules listed above.

An XML document can, in addition, be **valid** if it has data that conform to a particular set of user-defined content rules. These rules are defined in an XML Schema document.

An XML schema for our example could specify that a *note* element must contain a *to* element, a *from* element, a *topic* element and a *message* element in that order. Rearranging the order or deleting one of the elements would be an error.

The XML Schema can define what the options are for the status attribute, maybe only "hot," "warm," "cool," or "cold." A *status* value of anything else would be an error.

Appendix F: State Abbreviations and FIPS Codes

This appendix defines the state abbreviations and the Federal Information Processing Standard (FIPS) codes (01-78).

Table F-1: State Abbreviations and FIPS Codes

State name	FIPS	State abbreviation
ALABAMA	01	AL
ALASKA	02	AK
ARIZONA	04	AZ
ARKANSAS	05	AR
CALIFORNIA	06	CA
COLORADO	08	CO
CONNECTICUT	09	CT
DELAWARE	10	DE
DISTRICT OF COLUMBIA	11	DC
FLORIDA	12	FL
GEORGIA	13	GA
HAWAII	15	HI
IDAHO	16	ID
ILLINOIS	17	IL
INDIANA	18	IN
IOWA	19	IA
KANSAS	20	KS
KENTUCKY	21	KY
LOUISIANA	22	LA
MAINE	23	ME
MARYLAND	24	MD
MASSACHUSETTS	25	MA
MICHIGAN	26	MI
MINNESOTA	27	MN
MISSISSIPPI	28	MS
MISSOURI	29	MO
MONTANA	30	MT
NEBRASKA	31	NE
NEVADA	32	NV
NEW HAMPSHIRE	33	NH
NEW JERSEY	34	NJ
NEW MEXICO	35	NM

State name	FIPS	State abbreviation
NEW YORK	36	NY
NORTH CAROLINA	37	NC
NORTH DAKOTA	38	ND
OHIO	39	OH
OKLAHOMA	40	OK
OREGON	41	OR
PENNSYLVANIA	42	PA
RHODE ISLAND	44	RI
SOUTH CAROLINA	45	SC
SOUTH DAKOTA	46	SD
TENNESSEE	47	TN
TEXAS	48	TX
UTAH	49	UT
VERMONT	50	VT
VIRGINIA	51	VA
WASHINGTON	53	WA
WEST VIRGINIA	54	WV
WISCONSIN	55	WI
WYOMING	56	WY
AMERICAN SAMOA	60	AS
BUREAU OF INDIAN AFFAIRS	59	BI[16]
DOD DOMESTIC	61	DD[16]
DOD OVERSEAS	58	DO[16]
GUAM	66	GU
MARSHALL ISLANDS	68	MH
MICRONESIA	64	FM
NORTHERN MARIANAS	69	MP
PUERTO RICO	72	PR
REPUBLIC OF PALAU	70	PW
VIRGIN ISLANDS	78	VI

[16] Not official U.S. Postal State Abbreviations. The state abbreviations for the Department of Defense (overseas) schools are AA, AE, and AP to indicate schools located in Asia, Europe, and the Pacific, respectively. For Department of Defense (domestic) schools and Bureau of Indian Education schools, state abbreviations correspond to the state in which the school is located.

To promote student achievement and preparation for global competitiveness by fostering educational excellence and ensuring equal access

www.ed.gov