Flat File Submission Method Instructions

Introduction

The Civil Rights Data Collection (CRDC) is a mandatory data collection of the U.S. Department of Education's Office for Civil Rights. For more information about the CRDC, visit Civil Rights Data | U.S. Department of Education.

Two methods will be available for local educational agencies (LEAs) to report data:

- **Online Data Entry**: The online data entry method using the Web interface during the survey open period requires users to hand-key responses into the CRDC data elements. Using only the Web-based submission method may not be practical for large LEAs that must enter data for many schools.
- **Flat File Submission (FFS)**: The FFS method requires users to report data by preparing and uploading one or more text-based flat files in comma-separated value (.csv) format. The FFS method allows an LEA to prepare one or more data files that contain data elements required in the CRDC. This method may be used to provide all the data for the CRDC or a partial set of CRDC data. These files must conform to the FFS formats and specifications included within this document. These files can be uploaded once the survey opens.

LEAs may also choose to combine these two methods. For example, an LEA may choose to use the FFS method for the majority of its CRDC submission, and then allow individual schools to enter data via the online survey tool for a subset of specific questions (e.g., interscholastic athletics).

Purpose

The purpose of this document is to provide instructions for LEAs on how to create flat files for submission.

Who Should Use the Flat File Submission Method?

Any LEA can choose to submit their data by using the FFS method. However, the FFS method is primarily intended for use by LEAs that have an existing central student database. The CRDC FFS method is designed to be flexible to meet the needs of school districts to organize and upload CRDC data elements. Mid-size and even small school districts may find that creating flat files for some or all their CRDC data would be more efficient than keying in data through the online interface.

Flat File Creation Process

Creating File Templates

There are two file templates for the CRDC—Excel Template LEA Form and Excel Template School Form. Each may be submitted once or in multiple, partial submission uploads, as desired.

The CRDC provides many options for school districts to create and organize their flat file submissions. In choosing the number and order of data elements to organize into flat files, consider how the data are stored in information systems within the LEA.

- For large school districts, grouping large chunks of the survey data into files similar to Part 1 and Part 2 of the survey may be most efficient, if their information systems already group the data in that manner.

- Other school districts may choose to group together similar types of data stored in unique data systems. For example, one option may be to create separate files for the student enrollment data, school-level expenditures, and school personnel, if these data are stored in different systems.

Required File Format

The CRDC will use a comma-delimited (.csv) file format for flat file submissions. Users may generate files in the comma-delimited format using many different programs, ranging from database programs to spreadsheet applications, such as Excel, to simple text editors. In a comma-delimited file, each field is only as long as it needs to be to hold the data (no trailing blanks) and is separated from the next field by a comma. Lines of data are separated by a carriage return/line feed.

The following rules apply to all comma-delimited files:

- File extensions must be ".csv."
- Carriage Return/Line Feed (CRLF) must immediately follow the last field in each record (line).
- All fields, except for the last field in a record, must be followed by a comma.
    - For fields with values, the comma should immediately follow the permitted value.
    - For fields with no value, only the comma is entered. When the last field in a data record layout has no value, the data will look like the record ends in a comma, but that comma represents the blank field.
- Numeric data values must not contain commas within the field.
- For all alpha/numeric data values, use double quotes as text qualifiers.

The following rules are particular to how data must be formatted and organized within the file for purposes of the CRDC:

- Each data element being submitted in the file must be defined in the first line, which then constitutes a data header line for the particular submission file.
- Additional lines of data provide values for the data elements defined in the first line.
- The first data header, and the first value reported in each additional line, must always be for either an LEA ID or School ID. This ensures every value reported is associated with an LEA or school. The CRDC survey uses the LEA IDs and School IDs assigned by NCES for purposes of the Common Core of Data (CCD) and other federal surveys. No unique State or LEA IDs can be used in submitting data for the CRDC.
    - A CRDC/NCES LEA ID is always a seven-digit code representing the specific school district or other educational agency. The first two digits represent the state ID.
    - A CRDC/NCES School ID is always a twelve-digit code comprising the combination of the LEA ID with an additional five digits or characters representing the specific school or other educational institution within that LEA.
- Fields cannot exceed the maximum length defined for the data element.
- Data element values NOT being reported in the file may be indicated in one of two ways (more detailed information about partial file uploads is provided in the Partial Files section in this document):

- Completely omit the data element by not having the data element header in the first line of the file at all.
- Enter a blank field in the form of a comma with no value between it and the preceding field.

The software correlates values to the data elements by matching field order (number of comma-separated values) between the data header line and the additional lines, which is why it is important to check that data fields are properly lined up with the data header line prior to upload. Each data field in the first line must contain a data element name, and each data field in the additional lines must correspond to one of those data header fields.

Tips for Creating .CSV Files

Find menu options for **Export** or **Save As**: Many database systems (e.g., Oracle, SQL) and spreadsheet applications (e.g., Excel, OpenOffice) allow for the output of tabular data into a comma-delimited format in an automated fashion. Typically, users will find menu options for **Export** or **Save As** allow them to generate files in .csv from whatever data they select in those programs. Spreadsheet applications will also allow users to easily view comma-delimited files as spreadsheets, with comma-separated values displayed as columns and lines of data displayed as rows, allowing users to check that their data are properly formatted and 'lined up' without actually changing the format or file extension.

Making Changes or Corrections to the File: Users should be cautious when viewing or working with an already created .csv file in a spreadsheet application. Many versions of applications such as Excel will insert extraneous data into the file if the user saves into the spreadsheet format (e.g., .xls), which will persist even if the file is once again exported or saved as .csv. Best practice for once a properly formatted .csv file has been generated is to use spreadsheet applications for viewing the file only, not for making and saving changes to it. If corrections must be made, do so in the original source file and export/save to .csv again.

Check for leading zeros: Spreadsheet applications may interpret text fields as numerical fields and omit required data such as leading zeroes in ID numbers, both for viewing a .csv file and for saving as a spreadsheet. Best practice for once a properly formatted .csv file has been generated, therefore, is to use spreadsheet applications for viewing the file only, not for making and saving changes to it. If corrections must be made, do so in the original source file and export/save to .csv again.

Numeric Values: Numeric values must not contain commas within the field. For example, use **1876** and not **1,876** when preparing flat files for the CRDC.

Required Data Element Headers - REQUIRED for 2023-24

A new header has been added for each table that has an aggregated by sex category.

X - Nonbinary students

Table 1. Example of new header

CRDC Question/Table	CRDC Table Cell/Data Element	Element Name
Student Enrollment	Number of Hispanic Nonbinary Students	SCH_ENR_HI_X

The first line of each FFS file must contain the standardized data element names which serve as headers to clearly identify the data elements for which values are being provided. Each data element header is less than 32 characters and has a specific field length and set of permitted values. Generally, data elements follow a consistent naming convention. Each data element name begins with either **SCH** for school-level data elements or **LEA** for LEA-level data elements, followed by a short abbreviation for the type of data and disaggregation. There are no blank spaces within data element names. Instead, underscores "_" are used in the names to separate categories of data. For example, a school-level data element name for the CRDC question where schools report the total number of students enrolled is **SCH_ENR_HI_F**, the school-level enrollment of Hispanic female students. For all data elements, a consistent set of race/ethnicity abbreviations are used:

HI: Hispanic	**BL**: African American/Black
AM: American Indian/Alaska Native	**WH**: White
AS: Asian	**TR**: Two or more races
HP: Native Hawaiian/Pacific Islander	

Figure 1. Race/Ethnicity abbreviations

CRDC School and LEA flat files can organize the data element headers in any order, omit any of the data elements other than the School ID or LEA ID, and provide blank values, even for the data elements defined in the header line. However, each value submitted in the later lines must correspond to a data element header to be applied to the correct data element.

The FFS standardized data element names for each of the two file templates (LEA and School) can be found in the following documents:

- CRDC List of Elements (Flat File Specs): Excel spreadsheet with tabs for LEA and school level data elements. Each tab lists all the data elements required by the CRDC for that file/form.
- FFS Table layouts, at both the school and LEA level, that map the data element headers to individual cells in the CRDC table layouts:
  - CRDC School-Level Table Layout
  - CRDC LEA-Level Table Layout

Permitted Values

The CRDC has the following types of permitted values:

- Integer: Any whole number (no decimals or fractions)
- Yes/No: Each data element must be either Yes (3 characters) or No (2 characters)
- Name Format: Maximum 35 characters, including spaces (separate fields for first and last)
- Phone Format: Maximum 25, minimum 10 characters, including spaces, parentheses, hyphens, or periods, accommodating numerous different ways of formatting phone numbers such as (XXX) XXX-XXXX or XXX-XXX-XXXX
- Email Format: Maximum 254 characters in format of _____@_____._____
- URL Format: Maximum 2000 characters with no spaces, must contain a period (.); @ character not accepted
- Decimal: Maximum 10 characters for FTE values, 16 characters for finance values, with two decimal places

Partial Files

For the CRDC, LEAs continue to have the ability to submit partial data files. The following is an explanation of options for creating these partial files.

Partial Data for Individual LEA or School Records

As before, a specific LEA or school form record does not need to contain all data elements for the form. Subsequent submissions for a specific LEA or School form record need only contain additions or changes to data elements submitted previously. The reasons for submitting partial files, even for an individual LEA or school record are, as follows:

- Chronological – some reportable data become available sooner than other data, and the respondent chooses to submit these data in separate files, as they become available.
- Categorical – different users at the LEA or school are responsible for gathering and submitting different types of data (e.g., athletics, discipline, finance) and therefore choose to submit separate files.
- Skip logic – certain responses on guiding questions in the survey allow users to entirely skip other questions and corresponding data elements on the form, as defined in the CRDC Table Layouts.

The header line and/or data element name field included in the template allows partial data submissions to be performed, not only by using blank 'filler' fields, but also allows the option of completely omitting the fields for which data are not being submitted by taking out the entire column or row.

Table 2. Example 1 – Partial Files Using Blank "Filler" Fields

Data Element Headers	LEA_ID	LEA_ENR	LEA_SCHOOLS	LEA_ENR_NONLEAFAC	LEA_PS_IND
Initial Submission	1234567	12345		7890	Yes

Second Submission	1234567		23		No
Resulting data in database	1234567	12345	23	7890	No

Example 1: Partial Files Using Blank "Filler" Fields
In Example 1 above, the user is submitting two partial data submissions using the blank 'filler' method. The first line shows the headers present in the first line of the submission file. The second line shows data for an LEA record in the initial file submission. The third line shows the data submitted for the same LEA in a second file submission.

When the system reads the first record, it writes all fields with data to the database. When the system reads the second record, it only reads and updates fields containing data. Any fields in the second record containing blanks will not overwrite those same fields in the database from prior file submissions for that LEA, for reasons discussed in the Partial Files section above. The last row shows the data that have been written to the data base system from the two-file submission.

Table 3. Example 2 – Partial Files by Omitting Data Element Headers

Data Element Headers	LEA_ID	LEA_ENR	LEA_ENR_NONLEAFAC	LEA_PS_IND
Initial Submission	1234567	12345	7890	Yes

Data Element Headers	LEA_ID	LEA_SCHOOLS	LEA_PS_IND
Second Submission	1234567	23	No

Data Element Headers	LEA_ID	LEA_ENR	LEA_SCHOOLS	LEA_ENR_NONLEAFAC	LEA_PS_IND
Resulting data in database	1234567	12345	23	7890	No

Example 2: Partial Files by Omitting Data Element Headers
In Example 2, instead of using blank fields, the data element header line is edited to only show elements for which data will be submitted.

As before, only values entered in the file are written to the database. In this instance, there is no requirement to determine what to do with a blank value, although if one were provided it would still not overwrite anything. In fact, users may mix and match the two methods of creating and preparing a file for upload. They may delete whole sections of the data element names for those portions of the survey they are not prepared to report on at all, while filling in blank values for specific elements within sections that they are submitting data for at this time.

There are multiple benefits to this approach:

- Users do not need to submit all data for each record in a single LEA and School file submission. Data can be collected/submitted as it becomes available.
- Users can make corrections to the data for individual schools or LEAs without having to resubmit the data for all of them. If previously submitted data needs to be revised, simply re-submit the updated data for that individual school or LEA.

Partial Records for Files

A specific file does not need to contain all records for the submission (e.g., all school records for the LEA, all LEA records for the LEA group). An initial file can contain only some of the required records. Subsequent files need only contain records that were not submitted previously.

For example, the original file could contain records for the elementary schools, and the second file could contain records for the high schools. As another example, if after submitting the school file it is discovered that a school was not included in the file, submit another file that contains just the record for that one school.

The benefit of this approach is that all records do not need to be submitted for the LEA or LEA group in one file submission. Records can be collected/submitted as they become available.

Treatment of Blank Fields in Partial Files

As discussed, sometimes a flat file submission will contain a particular data header, but one or more of the data fields in that column will be a blank 'filler' field. This could be because the submitter chose that method of making a partial submission for a single school or LEA, or because the file contains data for multiple schools or LEAs and values for that data element were available for some but not others.

Whenever a file contains such blank fields, they will be treated in a particular manner by the system. Each blank field will be read as a null (missing) value for that data element by the system, but such null values will not overwrite previously submitted values. This is necessary to allow partial files to be submitted without wiping out data for all of the data elements that are not included as part of a particular file.

Special Values

Not Submitted (NS)

Since the system will disregard blank values where a previous value has been submitted, users cannot 'blank out' or delete previously submitted data element values by submitting a file with such blanks. If it is discovered that data were submitted for fields that should not have been submitted at all, upload a file in which the fields to be deleted as *not yet reportable* or *not to be submitted at this time* are indicated using the special code **NS** for **Not Submitted**.

Record Layouts

The record layouts for the CRDC form records must follow the CRDC specifications that require the first row of data to include applicable data element names as headers.

The comma-delimited text file formats for the CRDC are presented in the following documents:

- CRDC LEA Form Excel Template
- Excel Data Collection Template School Form

Each document provides three tabs for use in creating data files for FFS:

- **Form Instructions** include the relevant instructions for all the items in the given survey form.
- **Form Data Entry** is a highly formatted, interactive Excel sheet providing users

with an easy way to enter data to the questions and fields in the appropriate formats. Data entered through this tab are also reflected in the **Form Consolidated** tab.

- **Form Consolidated** is a simple Excel sheet with all the data element headers populated in the top row, suitable for export to a .csv format text file once all the desired elements/fields have been populated. It can be populated using the formatted **Form Data Entry** tab.

Recommended File Size

The recommended maximum file size is 30 MB to optimize upload and processing times. This recommended maximum file size should be large enough to accommodate all school districts uploading all CRDC data elements for all schools. For example, a single data file with complete data from 1600 schools would be approximately 25 MB.

Special Note for state educational agencies (SEAs)*:* For files that exceed 30MB, processing times may be longer. Users will be notified in real time of the system's progress in uploading and processing the submitted data. The submission system will accept larger files, but users submitting data for multiple LEAs may expect longer processing times. A single data file with data from 10,000 schools in a very large SEA having 1,000 districts would be approximately 200 MB. Processing time for file uploads will vary depending on file size and system load at the time of submission.

File Naming Convention

Since the CRDC tool will accept multiple file uploads regardless of file name, it is possible for respondents taking advantage of this feature to generate large numbers of files. The file naming convention below is suggested to help school districts differentiate and track all versions of their flat file submissions. This is a suggested naming convention, but ultimately the file can be named anything that the submitter wishes to name it, and there is not a restriction on the number of characters that can be used to name the file.

Acronyms

CRDC	Civil Rights Data Collection
FFS	Flat File Submission
CRLF	Carriage Return/Line Feed
LEA	Local Educational Agency
CSV	Comma Separated Values
SCH	School
ED	U.S. Department of Education
SY	School Year
EOF	End-of-File

Table 4: Suggested File Naming Convention for Files Before Uploading to the Site

Where	Means	Limit in Characters
1234567	LEA ID (include leading zero, if applicable)	7
lev	Abbreviation for form level: • LEA for a file containing Local Education Agency form record(s) • SCH for a file containing a school form record(s)	3

vvvvvvv	Alphanumeric string designated by the SEA to uniquely identify the individual submission (e.g., ver0001, v120921, athlete, discipline, etc.). Can include version identification, if resubmitting.	7
.ext	Extension identifying the file format: .csv	4

Other CRDC Resources

In addition to this instruction document, the following documents are available to assist LEAs in submitting data through flat files.

Creating File Templates

- CRDC Table Layouts – These documents describe the data that are being collected in the CRDC survey. They also provide definitions for all terms related to the data being collected.
  - LEA Form Table Layout
  - School Form Table Layout
- CRDC General Overview, Changes and List of Data Elements
- File Specifications – These documents provide formatting instructions for building the .csv flat files.
  - List of Elements (Flat File Specs)
  - CRDC Excel Data Collection Template LEA Form
  - CRDC Excel Data Collection Template School Form

Other Resources

Several helpful guides, instructions, training videos, and other resources are available at the CRDC Resource Center.

Contact the Partner Support Center

(855) 255-6901 | PSC Contact Us