

Instructions for using the CRDC Flat File Submission Method

Table of Contents

1.	Introduction	2
1.1	Purpose.....	2
1.2	Who should use the FFS method?.....	2
1.3	Major Changes for SY 2011-12 CRDC Submitters.....	2
2	Flat File Creation Process	3
2.1	Creating File Templates.....	3
2.2	Required File Format.....	3
2.3	Required Data Element Headers.....	5
2.3.1	Permitted Values.....	6
2.4	Partial Files.....	6
2.4.1	Partial Data for Individual LEA or School Records.....	6
2.4.2	Partial Records for Files.....	8
2.4.3	Treatment of Blank Fields in Partial Files.....	8
2.5	Special Values.....	8
2.5.1	Not Submitted (NS).....	8
2.6	Record Layouts.....	9
2.7	File Size.....	9
2.8	File Naming Convention.....	9
3	Other CRDC Resources	10
3.1	Current Resources.....	10
3.2	Future Resources.....	10
4	Acronyms	10

1 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 <http://www.ed.gov/ocr/data.html>.

Two methods will be available for Local Education Agencies (LEAs) to report data.

- Online Data Entry: Online data entry using the Web interface during the survey open period will require users to hand-key in responses to the CRDC data elements. Using only the Web-based submission method may not be practical for large LEAs that must enter data for a large number of schools.
- Flat File Submission (FFS): The second method is to report data by preparing and uploading one or more text-based flat files. The FFS method allows an LEA to prepare one or more data files that contain data elements required in the SY 2013-14 CRDC. The FFS method may be used to provide all of the data for the CRDC or a partial set of CRDC data. These files must conform to the FFS formats and specifications included with 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, then allow individual schools to enter data via the online survey tool for a subset of specific questions (e.g. interscholastic athletics).

1.1 Purpose

The purpose of this document is to provide instructions for school districts on how to create flat files for submission in Fall 2014.

1.2 Who should use the FFS method?

Any LEA can choose to submit using the FFS method. However, the FFS method is primarily intended for use by LEAs that have an existing central student database. The 2013-14 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 of their CRDC data would be more efficient than keying in data through the online interface.

1.3 Major Changes for SY 2011-12 CRDC Submitters

Based on feedback from school districts, the 2013-14 CRDC FFS will allow school districts to:

- Submit files with a subset of fields (i.e. a field does not need to be present for every data element); and
- Organize data elements in any order.

If you participated in the SY 2011-12 CRDC, there are a number of changes to the SY 2013-14 CRDC FFS from the prior collection that you should be aware of:

- SY 2013-14 CRDC FFS requirements for developing files have changed:
 - School-level and LEA-level FFS files can be created for as few as one data element or for any combination of data elements in Part 1 or Part 2, provided that each file contains the required data element headers.
 - FFS files **must** contain data element headers that clearly identify the data elements for which any values will be provided and should not contain extraneous blank columns or rows.
 - Part 1 and Part 2 data can now be submitted within the same files, using the same file templates.
 - The FFS forms now only require two separate file templates, one for LEA and one for School(s).
 - Some of the new questions are optional for SY 2013-2014 but will be required in the SY 2015-2016 collection. Please review the SY 2013-14 CRDC Table Layouts and the List of Proposed Continuing, New, and Dropped CRDC Data Elements for School Years 2013–14 and 2015–16 for full details on these changes.
 - Comma-separated value (.csv) files will be required.
- The 2013-14 submission system will also have important changes to how FFS files are submitted and processed.
 - LEAs will submit the FFS files via upload to the SY 2013-14 CRDC survey application, which will then validate and run data checks on the submitted data. Respondents will receive feedback reports in response to their data file uploads showing the submission status and results of these checks. Data corrections, if needed, may be made via the online forms or by submitting a new flat file.

2 Flat File Creation Process

2.1 Creating File Templates

There are two file templates for the 2013-14 CRDC, LEA and School. Each may be submitted once or in multiple, partial submission uploads as desired, as further described in section 2.4.

The 2013-14 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 is stored in information systems within your LEA.

- For large school districts familiar with the 2011-12 FFS, concatenating all the Part 1 data, similar to the 2011-12, field layout, may be most efficient.
- Other school districts may choose to group similar types of data stored in unique data systems together. For example, one option would be to create separate files for the

student enrollment data, school-level expenditures, and school personnel if these data are stored in different systems.

2.2 Required File Format

The SY 2013-14 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.
 - o For fields with values, the comma should immediately follow the permitted value.
 - o 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 SY 2013-14 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 a School ID. This ensures every value reported is associated with an LEA or school.
- Fields cannot exceed the maximum length defined for the data element.
- Data element values NOT being reported in the particular file may be indicated in one of two ways (more detailed information about partial file uploads is provided section 2.4):
 - o Completely omit the data element by not having the data element header in the first line of the file at all.
 - o 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:* As mentioned at the beginning of the section, 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 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.
- *Check for leading zero's:* Spreadsheet applications may interpret text fields as numerical ones 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 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 2013-14 CRDC.

2.3 Required Data Element Headers

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 element names 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 Question 7 where schools report the total number of 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
- AM: American Indian/Alaska Native
- AS: Asian
- HP: Native Hawaiian/Pacific Islander
- BL: African American/Black
- WH: White
- TR: Two or more races

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 or LEA ID, and provide blank values even for the data elements defined in the header line, but each value submitted in the later

lines must correspond to a data element header in order 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 SY 2013-14 List of Elements: Excel spreadsheet with tabs for LEA and school level data elements.
- FFS Table layouts, at both the school and LEA level, that map the data element headers to individual cells in the CRDC table layouts

Each document lists all of the data elements required by the CRDC for that file/form.

2.3.1 Permitted Values

The 2013-14 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 without spaces (separate fields for first and last)
- Phone Format: Maximum 20 characters, with 10 allotted for a 10-digit phone number and 10 allotted for extension information or to accommodate 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
- Decimal: Maximum 10 characters for FTE values, 16 characters for finance values, with two decimal places

2.4 Partial Files

For SY 2013-14 CRDC, LEAs continue to have the ability to submit partial data files. Following is an explanation of new options for creating these partial files.

2.4.1 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 several:

- Chronological – some reportable data becomes 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 you to entirely skip other questions and corresponding data elements on the form as not required, as defined in the SY 2013-2014 Table Layouts and File Specifications.

With the 2013-14 collection, the addition of the header line and/or data element name field to the template allows partial data submissions to be performed not only using blank ‘filler’ fields as

previously, but also allows the option of completely omitting the fields for which data is not being submitted by taking out the entire column or row.

Example 1: Partial Files Using Blank “Filler” Fields

Data Element Headers	LEAID	LEA_ENR	LEA_SCHOOLS	LEA_ENR_NONLEAFAC	LEA_PS_IND
Initial Submission	xxxxxxxxxxxxx	12345		7890	Yes
Second Submission	xxxxxxxxxxxxx		23		No
Resulting data in database	xxxxxxxxxxxxx	12345	23	7890	No

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 data base from prior file submissions for that LEA, for reasons discussed in section 2.4.3. The last row shows the data that has been written to the data base system from the two file submissions.

Example 2: Partial Files By Omitting Data Element Headers

Data Element Headers	LEAID	LEA_ENR	LEA_ENR_NONLEAFAC	LEA_PS_IND
Initial Submission	xxxxxxxxxxxxx	12345	7890	Yes

Data Element Headers	LEAID	LEA_SCHOOLS	LEA_PS_IND
Second Submission	xxxxxxxxxxxxx	23	No

Data Element Headers	LEAID	LEA_ENR	LEA_SCHOOLS	LEA_ENR_NONLEAFAC	LEA_PS_IND
Resulting data in database	xxxxxxxxxxxxx	12345	23	7890	No

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:

- You 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.
- You can make corrections to the data of individual schools or LEAs without having to resubmit the data for all of them. If previously submitted data needs to be revised, you can simply re-submit the updated data for that individual school or LEA.

2.4.2 Partial Records for Files

A specific file does not need to contain all records for your 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, your original file could contain records for your elementary schools, and your second file could contain records for your high schools. As another example, if after submitting your school file you discover that you forgot to include a school in your file, all you need to do is submit a file that contains just the record for that one school.

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

2.4.3 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 in order 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.

2.5 Special Values

2.5.1 Not Submitted (NS)

Since the system will disregard blank values where a previous value has been submitted, you cannot 'blank out' or delete previously submitted data element values by submitting a file with such blanks. If you discover that you submitted data for fields that should not have been

submitted at all, you will need to 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”.

2.6 Record Layouts

The record layouts for the SY 2013-14 CRDC form records have changed from the ones used in the SY 20011-12 survey. **Data files built using the SY 20011-12 record layouts will not be accepted unless they are modified to include the applicable data element names and fields no longer collected are removed.** Submissions must follow the SY 2013-14 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 SY 2013-14 CRDC are presented in the following documents:

- CRDC SY 2013-14 List of Elements: Excel spreadsheet with tabs for LEA and school level data elements.
- FFS Table layouts, at both the school and LEA level, that map the data element headers to individual cells in the CRDC table layouts

Each document provides a subset of the applicable standardized data element names populated as headers in the first data row in the file.

2.7 File Size

The recommended maximum file size is 30MB 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 Education Agencies: 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 200MB. Processing time for file uploads will vary depending on file size and system load at the time of submission.

2.8 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 to differentiate and track all versions of their flat file submissions.

Where	Means	Limit in characters
xxxxxxx	LEA ID (include leading zero if applicable)	7
lev	Abbreviation for form level: <ul style="list-style-type: none"> • LEA for a file containing Local Education Agency form record(s) • SCH for a file containing school form record(s) 	3

Where	Means	Limit in characters
vvvvvvv	Alphanumeric string designated by the LEA 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 – comma delimited (columnar)	4

3 Other CRDC Resources

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

3.1 Current Resources

- SY 2013-14 [CRDC Table Layouts](#) – This document describes the data that is being collected in the CRDC survey. This document also provides definitions for all terms related to the data being collected.
- [List](#) of Proposed Continuing, New, and Dropped CRDC Data Elements for School Years 2013–14 and 2015–16
- File Specifications – These documents provide formatting instructions for building your .csv flat files.

3.2 Future Resources

- CRDC Business Rules Guide – A spreadsheet listing all of the business rules that the CRDC online survey tool uses to check the quality of data submitted to the system.
- SY 2013-14 CRDC Users Guide – Information regarding the overall navigation and use of the SY2013-14 CRDC online survey tool.

4 Acronyms

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