

Electronic Data Deliverable (EDD) Specification Guidance Document



Civil & Environmental Consultants, Inc.
Pittsburgh · Chicago · Cincinnati · Cleveland · Columbus · Detroit
Export · Indianapolis · Nashville · Phoenix · St. Louis

Civil & Environmental Consultants, Inc.

**Electronic Data Deliverable
(EDD)
Specification**

For updates to the software, visit www.cecinc.com/edd

©2008-2010, Civil & Environmental Consultants, Inc.
333 Baldwin Road
Pittsburgh, PA 15205
(800) 365-2324

CEC EDD Specification

1.	Overview.....	1
2.	Description of the EDD Specification	2
3.	Using the Verifier Software	4
4.	Verifier Warnings and Common Resolutions.....	5
5.	Contact Information	8

Appendix: EDD Specification Detail

1. Overview

Civil & Environmental Consultants, Inc. (CEC) has developed a company-wide system for management of analytical data. This will allow CEC to manage all of its analytical data through one central database and system. To facilitate use of this system and control transcription errors, we have developed an electronic data delivery (EDD) format for transfer of analytical results. All analytical laboratories working with CEC are required to provide their data in this format, in addition to the hard copy or printable electronic formats provided otherwise.

To assist labs in proper preparation of the EDD files, we have also developed File Verifier software that can be used to check data files for correct format before they are submitted to CEC. A copy of this software is included with this guide. Instructions for its use are presented in Section 3.

2. Description of the EDD Specification

The detailed EDD specification is included in the Appendix to this booklet. Files are to be prepared in as tab-delimited ASCII (flat) file. Comma-delimited (CSV) and Excel formats are not acceptable. Quotation marks should NOT be used to delineate text fields.

CEC's specification relies on information and methods used in the Electronic Data Deliverable format developed by Region V of the USEPA, although *our specific format is not identical to the EPA format*. The EPA EDD website can be found at

<http://www.epa.gov/region5superfund/edman/index.html>. Our specification calls upon the Valid Values document associated with EPA's EDD. This document is available at <http://www.epa.gov/region5superfund/edman/download/MFEDappendix.pdf>.

The first line of the EDD must be the title line, reproduced exactly as shown in the specification. For convenience, a setup file containing this line is included on the enclosed CD as blankfile.txt. This file can also be downloaded from our website, at www.cecinc.com/edd/blankfile.txt.

Some guidance on specific items of the specification is presented below.

- Note that SampleID or LabID must be unique for each sample included in a data file. This requirement likewise applies to laboratory QA samples (matrix spikes, duplicates, method blanks).
- CEC's system uses CAS Numbers as the primary identifiers for chemical identity. This allows different chemical synonyms for chemicals to be applied to different projects, based on regulatory language, client preference, permit content, etc. In cases where no CAS Number applies to an analysis (such as indicators or combined parameters), a code from the Valid Values should be used, if available. If no Valid Value exists, check Attachment 2 of the CEC specification. In the absence of guidance in both of these documents, apply a unique code of up to 15 characters.
- The hyphens must be included in the CAS numbers; for example: 124-48-1 is right; 124481 is not.
- The parameter name is included for reference only, but may NOT be omitted. CEC uses this text to build a database of parameter synonyms.

- Data qualifiers are to follow the EPA Valid Values list.
- Basis is included to distinguish whether solids analysis is reported on a dry-weight or as-received basis. Since this column cannot be blank, include “N” for liquid samples.
- Total_or_dissolved is included to designate whether water samples were field-filtered (for “dissolved” analyses) or not (for “total” analyses). Like Basis, this field cannot be blank, so include “U” for solid samples or samples where such handling is otherwise not applicable.
- aMethod includes the analytical method. Where possible, this should use the designation provided in the Valid Values document. Consistency in presentation is important in this field so that the software can automatically identify whether analyses on a group of samples were performed in the same manner or not.
- Special is used to designate “special” or unusual handling of analyses. It is most commonly used to identify leach analyses of soils. Conventional totals analysis of solids or liquids should be designated by a blank in this column. Like aMethod, consistency in presentation in this column is very important. Insertion of non-critical information will result in difficulty in automatic tabulation and analysis of the data.

The most common errors that we have seen in preparation of the EDDs include:

1. *Improper modifiers attached to chemical names:* Water samples will often be reported with the field filtration indicated in the chemical name (“Calcium, total” and “Calcium, dissolved”). The parameter name should include ONLY the actual chemical name, and the total_or_dissolved field should be used to indicated filtration status.
2. *Improper file formats:* Note that Excel files are not acceptable. Also, only CEC’s EDD format can be used. Do not make any modifications to CEC’s format.
3. *Non-unique sample IDs:* If one data set includes several analyses from a single location, a distinct sampleID must be used for each individual sample. One way of accomplishing this is by appending a date or sequential identifier to the sample location. The unique sampleID is separate from the laboratory’s internal sample identification number.
4. *Excessive “assistance” from Excel:* If using Excel to clean up the data before exported to a flat file, be aware that it tends to try to convert CAS Numbers to dates and makes other

attempts to convert text into numbers. It may also add blank lines to the end of the data file.

5. *Overuse of comments:* The “comments” field should be used only for information that is specifically of interest to a specific analytical result. These should not be comments related to the sample as a whole, and the same comment should not be repeated for every analysis unless there is a specific issue. The comments should not repeat information indicated elsewhere in the EDD, such as analytical method, laboratory, sample description, etc.
6. *Improper QA qualifiers:* The list of qualifiers must conform to EPA’s qualifier list. Alternative values are not permitted.
- 7.

3. Using the Verifier Software

The CD included with this booklet includes installation files for CEC's Verifier software. The software can also be downloaded from www.cecinc.com/edd. Requirements for running the software include a PC equipped with Windows 2000/ XP or higher and Microsoft's .NET Framework 2.0.

To install the software, double-click the file, "setup.exe", located on the CD. You will then be guided through the setup process. Note that the setup program will verify that Microsoft's .NET Framework 2.0 is installed on the PC. If the setup program detects that the .NET Framework 2.0 is not installed, you will be prompted to download the .NET Framework from Microsoft. As an alternative to downloading the .NET Framework 2.0 setup program, you can install the .NET Framework from the CD. If you choose to install the .NET Framework from the CD, cancel out of the CEC Verifier setup process and install the .NET Framework 2.0 by double-clicking the file "dotnetfx.exe" located on the CD. You will then be guided through the process of setting up the .NET Framework 2.0 on your PC. Once the .NET Framework is installed, you may then double-click the file, "setup.exe" on the CD to install the CEC Verifier software.

4. Verifier Warnings and Common Resolutions

Below are the warning messages reported by the Verifier and some of the most common explanations and resolutions.

"Header line is empty. Data file verification aborted."

This message is displayed when the first line in the data file is empty. The first line in the data file must be the header row. And, the header row must be a tab delimited list of fields as specified in the data file specification.

"Header line is not tab delimited. Data file verification aborted."

This message is displayed when the first line in the data file is not tab delimited. The first line in the data file must be the header row. And, the header row must be a tab delimited list of fields as specified in the data file specification.

"Incorrect number of fields detected in header row. Data file verification aborted."

This message is displayed when the number of fields in the header row does not match the number of fields specified in the data file specification. Extra tabs in the header row can cause this message to be displayed.

"Header fields are invalid. Data file verification aborted."

This message is displayed when the header row does not match a particular data specification. The first line in the data file must be the header row. And, the header row must be a tab delimited list of fields as specified in the data file specification. Numerous problems can cause this verification failure including an empty header row, no tabs in the header row, incorrect number of fields in the header row, incorrect order of fields in the header row, field names are not spelled correctly, and/or the data file is not a tab-delimited ASCII text file.

"Line not tab delimited."

This message is displayed when the line contains no tabs. In a tab delimited data file, every line is expected to have tabs including the header row and all data lines.

"Incorrect Number of fields detected."

This message is displayed when the line contains too few or too many data values. Extra or missing tabs in the line can cause this message to be displayed. Null/empty values that are not tabbed properly can also cause this message to be displayed. The rule of thumb is that the header line and all data lines should always contain the same number of tabs and the number of tabs should correspond to the number of fields as specified in the data file specification.

"CAS Number/Parameter Name combination is not unique in the file."

This message is displayed when the program detects that there is more than one parameter name for a given CAS number. A CAS number can be associated with only one parameter name in the data file.

"Primary Key fields are not unique in the file."

This message is displayed when the combined values of the following fields are not unique in the file: SampleID, CASnumber, Basis, total_or_dissolved, Laboratory, aMethod, and Special.

"Null values are not permitted for field..."

This message is displayed when a null/empty value is detected for a field that does not permit null/empty values.

"Time format is invalid for field..."

This message is displayed when a time field contains a value that is not formatted correctly. A time field must be in the following format: hh:mm (1-2 digits followed by a colon then 2 digits).

"Date format is invalid for field..."

This message is displayed when a date field contains a value that is not formatted correctly. A date field must be in the following format: mm/dd/yyyy (1-2 digit month followed by a forward slash, then 1-2 digit day followed by a forward slash, then 4 digit year).

“Value for...” [field name] “...is invalid.”

This message is displayed when a field contains a value that is not permitted. Several fields in the data file specification are restricted to a particular set of values.

“Value for...” [field name] “...exceeds the permitted...” [number of characters allowed]

This message is displayed when the number of characters in a value exceeds the permitted character count.

“Value...” [field name] “...is not numeric.”

This message is displayed when a numeric field contains a value that is not a number.

5. Contact Information

If you need additional assistance in preparing the CEC EDD format, contact CEC's headquarters in Pittsburgh, Pennsylvania, via telephone on normal business days between 8:30 AM and 5:00 PM. The phone number is 800-365-2324. Ask the receptionist for assistance with preparing a laboratory upload.

You may also contact Deb Thompson by email at dthompson@cecinc.com.

Appendix

**ELECTRONIC FILE SPECIFICATION
CEC Site Data Management System**

File type: Tab-delimited text file, ASCII format

File Format:

Line 1: Header Row, tab delimited, as follows:

SampleID SampleDate SampleTime CASnumber ParamName Result Qualifier Units Basis total_or_dissolved Comments Laboratory aMethod Special MDL error RL LabID

Remaining Lines: One line per analytical result, format as follows

Content	Type	Maximum Length	Accepts Nulls	Explanation
SampleID	text	30	No	Client-assigned sample number
SampleDate	mm/dd/yyyy	n/a	No	Date of sample collection
SampleTime	hh:mm	n/a	Yes	Time of sample collection, in military time
CASNumber	text	15	No	CAS number for the analyte. If no CAS number exists, assign a unique code, using the conventions in Tables A-13 and A-15 of USEPA Region V EDD Valid Values*. A list of additional codes used by CEC is presented in Attachment 1.
ParamName	text	150	No	Name of the analytical parameter.
Result	numeric	n/a	No	Numeric result of the analysis. If the compound was not detected, insert the Practical Quantitation Limit (PQL).
Qualifier	text	6	Yes	Standard qualifier codes (from Table A-10 of USEPA Region V EDD Valid Values*). May use “=” for an unqualified result above detection limit.
Units	text	10	No	Standard units abbreviation. See Attachment 1 for valid units.
Basis	text	1	No	Basis of analysis. Use: D for dry weight W for as-received N for not applicable (such as water samples).
Total_or_dissolved	text	1	No	Whether the analysis represents a dissolved (field filtered) or total constituent analysis. Use: T for total D for dissolved U for not applicable (solid samples, etc.)
Comments	text	240	Yes	Brief information necessary for understanding this specific analysis. Do not duplicate information provided elsewhere in the file. Do not include laboratory identification.
Laboratory	text	50	No	Name of the laboratory performing the analysis
aMethod	text	25	Yes	Analytical method. Please use full citation, based on the Lab_AnI_Method_Name in Table A-16 of the USEPA EDD Valid Values*. Consistency in format is important.
Special	text	25	Yes	Special handling or preparation method. Common values include TCLP, SPLP, EP-Tox. Left blank for conventional analyses.
MDL	numeric	n/a	Yes	Minimum detection limit for analysis. Required only for radiological parameters.
error	numeric	n/a	Yes	Error value for the analysis. Required only for radiological parameters.
RL	numeric	n/a	Yes	Reporting limit – normally the PQL
LabID	Text	30	No	Laboratory sample identification number

Example Data:

SampleID	SampleDate	SampleTime	CASnumber	ParamName	Result	Qualifier	Units	Basis	t_or_d	Comments	Laboratory	aMethod	Special	MDL	error	RL	LabID
S-1	6/5/2003	8:20	7439-97-6	Mercury	0.0024		mg/L	N	U		Ace Labs	SW846 7470A	TCLP	0.000025		.00002	234X23345
S-1	6/5/2003	8:20	7439-97-6	Mercury	0.67		mg/kg	N	U		Ace Labs	SW846 7471A		0.0042		.001	234X23346
S-1	6/5/2003	8:20	297-97-2	Thionazin	33	U	ug/kg	N	U		Ace Labs	SW846 8141A		5.6		.01	234X23347
S-1	6/5/2003	8:20	298-04-4	Disulfoton	33	U	ug/kg	N	U		Ace Labs	SW846 8141A		6.6		.01	234X23348
S-1	6/5/2003	8:20	52-85-7	Famphur	33	U	ug/kg	N	U		Ace Labs	SW846 8141A		6.3		.01	234X23349
S-1	6/5/2003	8:20	56-38-2	Parathion	33	U	ug/kg	N	U		Ace Labs	SW846 8141A		7.2		.01	234X23350

Notes:

1. Provide information in title case, not all caps.
2. The chemical name should include the proper name of the chemical only. Do not include a chemical synonym as parenthetical information.
3. Some codes for non-CAS parameters are shown in Attachment 2.

* Information on the USEPA Region V Electronic Data Deliverable is available on the Internet at <http://www.epa.gov/region5superfund/edman/index.html>. The Valid Values document can be downloaded from <http://www.epa.gov/region5superfund/edman/download/MFEDappendix.pdf>.

**Attachment 1
Units Abbreviations**

Unit	Description
%	Percent
%V	Percent by Volume
°C	degrees Celcius
°F	degrees Farenheit
cfs	Cubic Feet per Second
cfu/100ml	Colony Forming Units per 100 Milliliters
cfu/g	Colony Forming Units per Gram
cfu/ml	Colony Forming Units per Milliliter
colf/100ml	Coliform Bacteria per 100 Milliliters
colf/g	Coliform Bacteria per Gram
fibers/g	Fibers per Gram
fibers/kg	Fibers per Kilogram
fibers/l	Fibers per Liter
g/cc	Grams per Cubic Centimeter
g/g	Grams per Gram
g/kg	Grams per Kilogram
g/l	Grams per Liter
g/m3	Gram per Cubic Meter
g/ml	Grams per Milliliter
gpm	Gallons per Minute
kg/m3	Kilogram per Meter Cubed
mg/g	Milligrams per Gram
mg/kg	Milligrams per Kilogram
mg/l	Milligrams per Liter
mg/m3	Milligrams per Cubic Meter (PPBV)
mg/ml	Milligrams per Milliliter

Unit	Description
mgd	Million Gallons per Day
mL	Milliliter
mmhos/cm	Milliohms (MMHOS) per Centimeter
mS/cm	Microsiemens per Centimeter
nm	nanometers
ntu	Nephelometric Turbidity Units
pcf	Pounds per Cubic Foot
pCi/g	PicoCuries per Gram
pCi/kg	Picocuries per Kilogram
pCi/l	Picocuries per Liter
pCi/m3	Picocuries per Cubic Meter
pCi/ml	Picocuries per Milliliter
pg/g	Picogram per Gram
pg/kg	Picrograms per Kilogram
pg/l	Picogram per Liter
pg/m3	Picograms per Cubic Meter
pg/ul	Picograms per Microliter
pH	pH Units
SU	Standard pH Units
ug	Microgram
ug/g	Micrograms per Gram
ug/kg	Micrograms per Kilogram
ug/l	Micrograms per Liter
ug/m3	Micrograms per Cubic Meter
umhos/cm	Microhoms (UMHOS) per Centimeter

Attachment 2
CAS Equivalents for Non-CAS Parameters

Acidity, Total	ACID
Total Alkalinity	ALK
Bicarbonate Alkalinity	ALKB
Carbonate Alkalinity	ALKC
Hydroxide Alkalinity	ALKH
Phenolphthalein Alkalinity	ALKP
Biochemical Oxygen Demand	BOD
Biochemical Oxygen Demand, 20 Day	BOD20
Biologic Oxygen Demand, Five Day	BOD5
Neutralization potential as CaCO3 equivalence	CCE
Chemical Oxygen Demand	COD
Coliform	COLIF
Color	COLOR
Corrosivity	CORROS
Cresol (Total)	Cresol-T
Free Cyanide	FREE CN
p-Cymene	CYMP
Diesel Components	DIESELCOMP
Fecal Coliform	FECCOLIFOR
Ferric Iron	FE-FC
Ferrous Iron	FE-FS
Field pH	FieldpH
Flashpoint	FLASHPT
Gasoline-Range Organics	GASCOMP
Hardness	HARD
Ignitability	IGNITB
Lipids	LIPIDS
Percent Lipids	LIPIDS
Total Volatile Solids(TVS)	LOI
Motor Oils	MOIL
Total Kjeldahl Nitrogen	KN

Moisture Content	MOIST
m,p-Cresol	MPCRESOL
m/p-Xylene	MPXYLENE
Other Hydrocarbons	NA
Net Acidity	NetAcidity
Oil & Grease	OILGREASE
PCB Total-TCL	PCB
% Organic Matter	Pct Org Mat
pH	pH
Phenolics	Phenolics
Reactivity	REACTIVITY
Residue, Total	RESTOT
Reactivity - Cyanide	RX-Cyn
Reactivity - Sulfide	RX-Sulf
Specific Conductance	SC
Surfactants	SURFACT
Total Bacteria	TB
Total Dissolved Solids	TDS
Temperature	TEMP
Tetrachlorobenzenes	TETRAXBENZENE
Total Trihalomethanes	THM
Total Inorganic Carbon	TIC
Total Organic Carbon	TOC
Total Organic Halides	TOX
Total Organic Halides - Brominated	TOX_BR
Total Organic Halides - Chlorinated	TOX_CL
Total Organic Halides - Iodinated	TOX_I
Solids	TS
Total Suspended Solids	TSS
Turbidity	TURB