

EQulS EDD

Quick Reference Guide

for Laboratories

MARCH 2021

Table of Contents

Glossary and Acronym List	ii
1 Introduction	1
2 EQuIS EDD Format Files.....	2
2.1 Format File Structure	2
2.2 Standard EQuIS Formats	3
2.3 Proprietary EQuIS Formats	3
3 EQuIS EDD Specifications and Creating EDDS	4
3.1.1 EDD Conventions	4
3.2 EDD Tools in EDP	5
3.2.1 EDD Descriptions	5
3.2.2 Blank EDDs.....	6
3.3 Reference Value Files.....	7
4 EQuIS Data Processor Workflow	8
4.1 EQuIS EDD Package Naming Conventions	9
4.1.1 ZIP File Name.....	9
4.1.2 Flat Text File Name for EDDs within the ZIP Package	9
4.1.3 EQuIS EDD File Naming Convention Rules	10
4.2 Processing EQuIS EDD in EDP Standalone.....	11

Glossary and Acronym List

Electronic Data Deliverable – An Electronic Data Deliverable (EDD) is a tabular, flat file format for sharing, manipulating, and using data, and is commonly a text or Microsoft Excel file type.

EQuIS (Environmental Quality Information Systems) – Customizable Schemas (relational database table structures) for environmental and geotechnical data management. Platform runs on Microsoft Azure or Microsoft SQL Server. Each database may have one to thousands of projects (i.e., facilities).

EQuIS Data Processor – The standalone version of the EQuIS Data Processor (EDP) is a powerful desktop interface to check for data completion and referential integrity, identify and correct errors, and create compressed files containing multiple related EDDs in a single useable format for upload and storage in EQuIS. Enterprise EDP is a web application that performs the same checks as EDP Standalone, but the entire process may be automated.

Facilities – The physical (site) or logical (project) extent of data made available for reporting or modeling. Each facility is stored in EQuIS as a FACILITY_CODE.

Format File – Structured in XML, the EQuIS format file set contains the definitions and restrictions for each individual field in available data tables and is the essence of data checking in EDP and EQuIS. The format files control data checks, such as range checking, reference values, formatting and enumerations. Format files are typically encrypted.

Format Definition File – The EQuIS Format Definition (EFD) file defines the format sections (EDD tables) and the columns in each of those tables. In addition to table definition, the EFD file defines the data mapping to the EQuIS database and any associated business rules, including data type consistency and lookup values. The EFD is denoted with either an XSD (unencrypted) or XSE (encrypted) file extension.

FTP – File Transfer Protocol is a standard network protocol used for the transfer of computer files between a client and server on a computer network.

LIMS (Laboratory Information Management System) – A software-based laboratory and information management system with features that support a modern laboratory's operations.

Reference Value File – For each EQuIS database, reference values are defined and can be exported to a reference value file. This file is part of the EDD file format set and is denoted with an RVF file extension. These reference values are required to process an EDD for errors.

XML – Extensible Markup Language. Used to write format files.

1 Introduction

Electronic environmental analytical data can be directly submitted to an EQuIS™ database by the laboratory processing the samples. This document provides a brief overview of the process, data requirements, and use of the EQuIS Data Processor (EDP) to check for data completion and referential integrity prior to uploading the data into EQuIS.

Data are submitted to EQuIS as Electronic Data Deliverables (EDDs), which are tabular, flat file formats for sharing, manipulating, and using data and are commonly a text or Microsoft Excel file type. An EDD file follows a consistent design meant to organize information in a useful format and typically contains header row(s) that describe what information should be completed in each column and in what format that data should be entered. EQuIS EDDs are based on a specific EQuIS format and are packaged into an archive zip file along with an EQuIS user certificate applying a set naming convention.

Laboratories will need to setup their Laboratory Information Management System (LIMS) to generate the EDDs to match the EQuIS format selected by the client requesting the analyses. In addition, each client database will have specific reference values for their projects. Coordination with the client is critical to appropriately configure the LIMS.

To help understand how to create an EQuIS EDD and/or setup a LIMS to generate an EQuIS EDD, the following topics are covered in this guide:

- EQuIS EDD Format File(s)
- EQuIS EDD Specifications and Creating EDDs
- EQuIS Data Processor Workflow

This guide is not intended to provide comprehensive, step-by-step instructions. More detailed documentation is available at <https://help.earthsoft.com/>.

2 EQuIS EDD Format Files

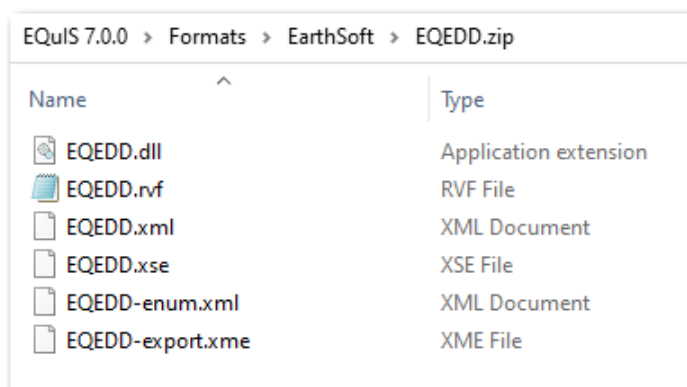
Typically, clients will request that EDDs are produced to meet the requirements of a specific EQuIS format. The format file is the essence of data checking in EDP and EQuIS. Structured in XML, the EQuIS format file set contains the definitions and restrictions for each individual field in available data tables. The format files control data checks, such as range checking, reference values, formatting, and enumerations.

Laboratories may submit data to EQuIS using one of the standard format files (e.g., EFWEDD, EQEDD, EZEDD, etc.) or a custom format file, depending on the needs of the client.

2.1 Format File Structure

An EQuIS EDD Format is comprised of several files and is usually provided as a .zip file. The EQuIS format file always includes the EQuIS Format Definition (EFD) file and the format .dll file but may also include other files.

Format Definition File – The EFD file defines the format sections (EDD tables) and the columns in each of those tables. In addition to table definition, the EFD file defines the data mapping to the EQuIS database and any associated business rules, including data type consistency and lookup values. The EFD is denoted with either an XSD (unencrypted) or XSE (encrypted) file extension.



EQuIS 7.0.0 > Formats > EarthSoft > EQEDD.zip	
Name	Type
EQEDD.dll	Application extension
EQEDD.rvf	RVF File
EQEDD.xml	XML Document
EQEDD.xse	XSE File
EQEDD-enum.xml	XML Document
EQEDD-export.xme	XME File

Format DLL File – Custom business rules that apply to the format are compiled in the format .dll file.

Reference Value File – For each EQuIS database, reference values are defined and can be exported to a reference value file (RVF). Reference values are required to process an EDD for errors in the EQuIS Data Processor. When an EDD is loaded into EDP for error checking, the desired reference values must also be loaded. Clients will provide the RVF to use in EDP Standalone or the LIMS.

Enumeration File – Enumerations are a way to enforce vocabulary on fields that have no reference tables. Enumerations can also restrict the vocabulary of fields that do have existing reference tables. This file is denoted as *-enum.xml.

Export File – The Export file is used to generate EDD exports from EQuIS and is generally not used by the laboratory. This file is denoted as *-export.xml.

2.2 Standard EQulS Formats

EarthSoft provides several standard EQulS formats for importing data. The formats range from simple single section EDDs to more complex, multi-section EDDs. Standard EQulS formats are available in the EarthSoft Formats folder (e.g., \EarthSoft\EQulS\Formats\EarthSoft). Clients may also directly provide an EQulS format or request contacting EarthSoft support for assistance.

The table below details several of the standard EQulS EDD formats that EarthSoft provides.

EFWEDD.xsd	The EFWEDD format, also known as the 4-file format, is a standard format that is suitable for most labs. It offers two different sections for sample information (choose either Field or Lab), a test section, a results section, and a batch section.
EQEDD.xse	The EQEDD is one of the most extensive/detailed standard EQulS formats. It contains sections for location, lab, and field data. It is a good generic format to begin importing data.
ESBasic.xsd	ESBasic is a simple, single section lab EDD format. Use to import single sample tests.
EZEDD.xsd	EZEDD is another single section lab EDD format. Note that a SYS_LOC_CODE field has been added to EZEDD, creating the new format EZEDD_LOCATION .
Historical.xsd	The Historical format contains four sections for importing historical data. The four sections are: 1. Water Levels 2. Geology 3. Chemistry 4. Location Data
EuroEQEDD.xse	The EuroEQEDD is a multiple section format that includes sections for Locations, Wells, Water Levels, Field Samples, Field Results, Soil, Gas, and Lab Data.

2.3 Proprietary EQulS Formats

Proprietary EQulS EDD formats are formats that were requested and designed for a third party. An example of a proprietary format is the EPA Region 4 format. Proprietary EDD formats will generally be provided directly by the client or via a download page.

3 EQuIS EDD Specifications and Creating EDDS

The “specifications” of the requested EQuIS EDD format need to be understood to ensure that the LIMS correctly generates EDDs. Laboratories may request the EQuIS format specifications from the client or generate an EDD description file for the EQuIS format in EDP (see Section 3.X). This section details EQuIS EDD conventions and the structure of EQuIS EDD files.

3.1.1 EDD Conventions

An EDD file follows a consistent design meant to organize information in a useful format and typically contains header row(s) that define what each column (field) in the EDD represents and in what format that data should be entered. The column headers in EQuIS EDD sections contain the data field names from the associated EQuIS database tables.

The **Header Row 1** defines the data in each column.

The **Header Row 2** indicates the data type and field length.

	Column 1	Column 2	Column 3	Column 4
Header Row 1	<u>#sys_loc_code</u>	<u>x_coord</u>	<u>y_coord</u>	<u>coord_type_code</u>
Header Row 2	#Text(20)	Numeric	Numeric	Text(20)
Data Row 1	MW-1S	-82.317493	28.057509	LAT LONG
Data Row 2	MW-2I	-82.317659	28.057151	LAT LONG

Both rows start with the “#” sign, which indicates that these are comment rows and do not contain data. The comment header rows will not be uploaded into the EQuIS database and are only in the file to provide context to the user.

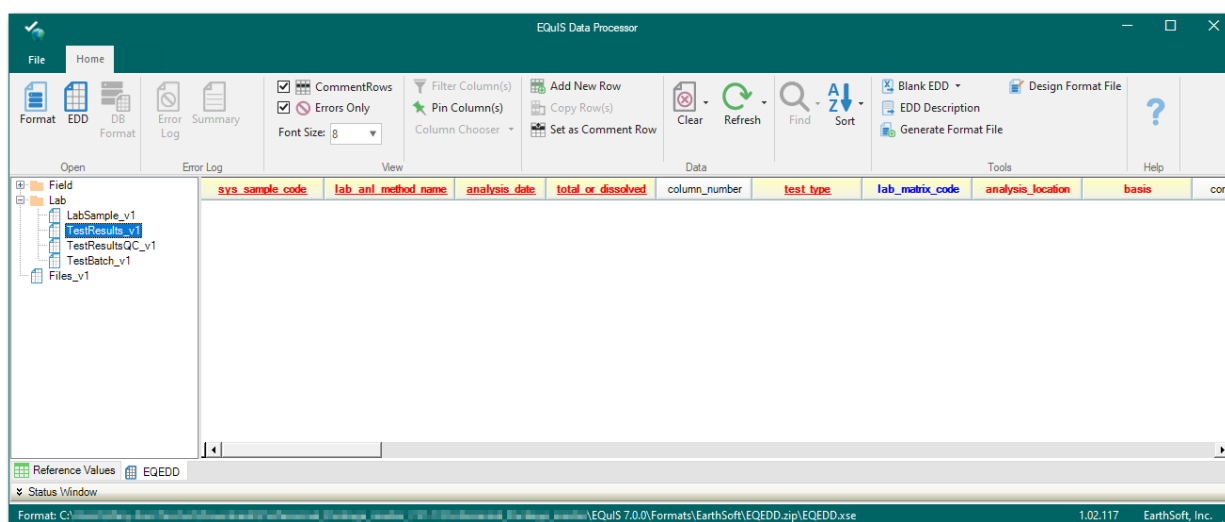
Note: Some EDDs are created without any header rows (analytical data from laboratories typically do not have header information).

To understand data connections and assist with data population, the criteria for certain fields are denoted by the following conventions:

- Red and Underlined text indicates a *Primary Key* field.
- **Red** text indicates a *Required* field.
- **Blue** text indicates a field linked to a *Reference* table (e.g., look-up or valid value).

Images on the next page are an example of an EQuIS EDD in Microsoft Excel and the standard EQEDD format opened in EDP. These images illustrate EQuIS fields with the data field conventions denoted in the columns.

	A	B	C	D	E	F	G	H	I
1	#data_provider	sys_loc_code	x_coord	y_coord	surf_elev	elev_unit	coord_type_code	observation_date	coord_identification
2	#Text(20)	Text(20)	Numeric	Numeric	Numeric	Text(15)	Text(2147483647)	DateTime	Text(20)
3	TUTORLAB	B-30	-78.79201	39.581796	678.34 ft		LAT LONG	3/12/1997 12:00	1
4	TUTORLAB	B-31	-78.79198	39.582028	660.86 ft		LAT LONG	3/12/1997 12:00	1
5	TUTORLAB	B-32	-78.79188	39.58219	668.92 ft		LAT LONG	3/12/1997 12:00	1
6	TUTORLAB	B-33	-78.79184	39.585622	679.31 ft		LAT LONG	3/12/1997 12:00	1
7	TUTORLAB	B-34	-78.79176	39.586472	678.45 ft		LAT LONG	3/12/1997 12:00	1
8	TUTORLAB	B-38	-78.79201	39.588706	677.37 ft		LAT LONG	3/12/1997 12:00	1
9	TUTORLAB	B-4	-78.79198	39.591731	699.12 ft		LAT LONG	3/12/1997 12:00	1
10	TUTORLAB	B-42	-78.79188	39.592103	678.84 ft		LAT LONG	3/12/1997 12:00	1
11	TUTORLAB	B-44	-78.79184	39.59227	677.02 ft		LAT LONG	3/12/1997 12:00	1
12	TUTORLAB	B-45	-78.79176	39.592984	673.21 ft		LAT LONG	3/12/1997 12:00	1
13	TUTORLAB	B-46	-78.79201	39.593479	677.11 ft		LAT LONG	3/12/1997 12:00	1
14	TUTORLAB	B-47	-78.79198	39.595398	674.84 ft		LAT LONG	3/12/1997 12:00	1
15	TUTORLAB	B-48	-78.79188	39.596233	674.77 ft		LAT LONG	3/12/1997 12:00	1



3.2 EDD Tools in EDP

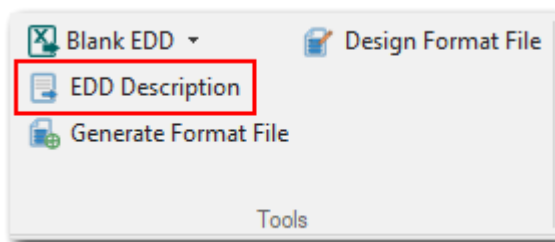
The Blank EDD and EDD Description tools within EDP facilitate data collection and processing. The Blank EDD tool creates an empty Microsoft Excel template that follows the EQuIS format file's specific layout for data entry. The EDD Description tool defines the specifications of the EQuIS format. These tools are useful references and may be incorporated into broader Standard Operating Procedures (SOPs). See this [article](#) for additional information on both tools.

3.2.1 EDD Descriptions

The EDD Description tool defines the specifications of the selected EQuIS format. Information about the selected format file is exported to a Microsoft Excel spreadsheet, including whether a field is a primary key (i.e., required), data type, related look-up tables, and EQuIS database mapping. The description file includes a separate worksheet detailing each section of the EDD format file.

To generate an EDD description file for a particular EQuIS format:

1. Open **EDP**.
2. Ensure that the desired format is loaded.
3. Click on the **EDD Description** button in the Tools group on the Home ribbon.



4. In the Save dialog box, select the desired folder and name the EDD Description file as desired. By default, the EDD Description file will be named after the EDD format (e.g., *EQEDD.xls*) and saved to the same directory as the EDD format. Click **Save**.
5. Excel will open the file, which has separate worksheets detailing each section of the EDD format file.

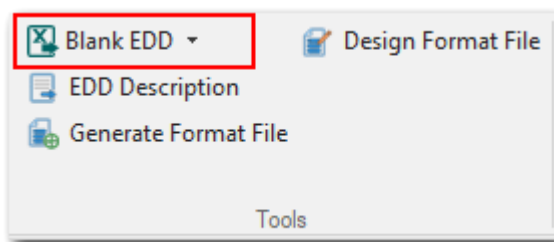
	A	B	C	D	E	F	G	H	I
	Field Name	Data Type	Key	Required	Default	Parent	Lookup	Database Mapping(s)	Comment
1	sys_sample_code	Text(40)	PK	Y		LabSample_v1.sys_sample_code			Unique sample identifier. Each sample at a facility must have a unique value, including spikes and duplicates. You have considerable flexibility in the methods used to derive and assign unique sample identifiers, but uniqueness throughout the database is the only restriction enforced by EQuIS.
2	lab_anl_method_name	Text(20)	PK	Y			rt_analytic_method.analytic_method		Laboratory analytical method name or description. A controlled vocabulary column, valid values can be found in the appendix in table lab_anl_method_name.
3	analysis_date	DateTime	PK	Y					Date and time of sample analysis in MM/DD/YYYY HH:MM:SS format. May refer to either beginning or end of the analysis as required by EPA.
4	total_or_dissolved	Text(10)	PK	Y			rt_fraction.fraction		Must be either 'D' for dissolved or filtered [metal] concentration, 'T' for total or undissolved, or 'N' for anything else
5	column_number	Text(2)							Values include either '1C' for first column analyses, '2C' for second column analyses or 'NA' for tests for which this distinction is not applicable.
6	test_type	Text(10)	PK	Y	INITIAL		rt_test_type.test_type		Type of test. Valid values include 'INITIAL', 'REEXTRACT1', 'REEXTRACT2', 'REEXTRACT3', 'REANALYSIS', 'DILUTION1', 'DILUTION2', and 'DILUTIONS3'
7	lab_matrix_code	Text(10)					rt_matrix.matrix_code		Code which distinguishes between different type of sample matrix. For example, soil samples must be distinguished from ground water samples, etc. The matrix of the sample as analyzed may be different from the matrix of the sample as retrieved (e.g. leachates), so this field is available at both the sample and test level.
8	analysis_location	Text(2)		Y			(Enumeration: analysis_location)		Must be either 'FT' for field instrument or probe, 'FL' for mobile field laboratory analysis, or 'LB' for fixed-based laboratory analysis
9	basis	Text(10)		Y	NA		(Enumeration: basis)		Must be either 'Wet' for wet_weight basis reporting, 'Dry' for dry_weight basis reporting, or 'NA' for tests for which this distinction is not applicable. The EPA prefers that results are reported on the basis of dry weight where applicable.
10	container_id	Text(30)							Report as null
11	dilution_factor	Numeric			1.0				Effective test dilution factor
12	prep_method	Text(20)					rt_prep_method.prep_method		Laboratory sample preparation method name or description
13	prep_date	DateTime							Beginning date and time of sample preparation in MM/DD/YYYY HH:MM:SS format
14	leachate_method	Text(15)							Laboratory leachate generation method name or description. The method name should be sufficient to reflect operation of the laboratory.
15	leachate_date	DateTime							Beginning date and time of leachate preparation in MM/DD/YYYY HH:MM:SS format
16	lab_name_code	Text(40)					rt_company.company_code		Unique identifier of the laboratory
17	qc_level	Text(10)					(Enumeration: qc_level)		May be either 'screen' or 'quant'
18	lab_results_id	Text(40)							Laboratory test results identifier. For screening a field sample number must be provided.

3.2.2 Blank EDDs

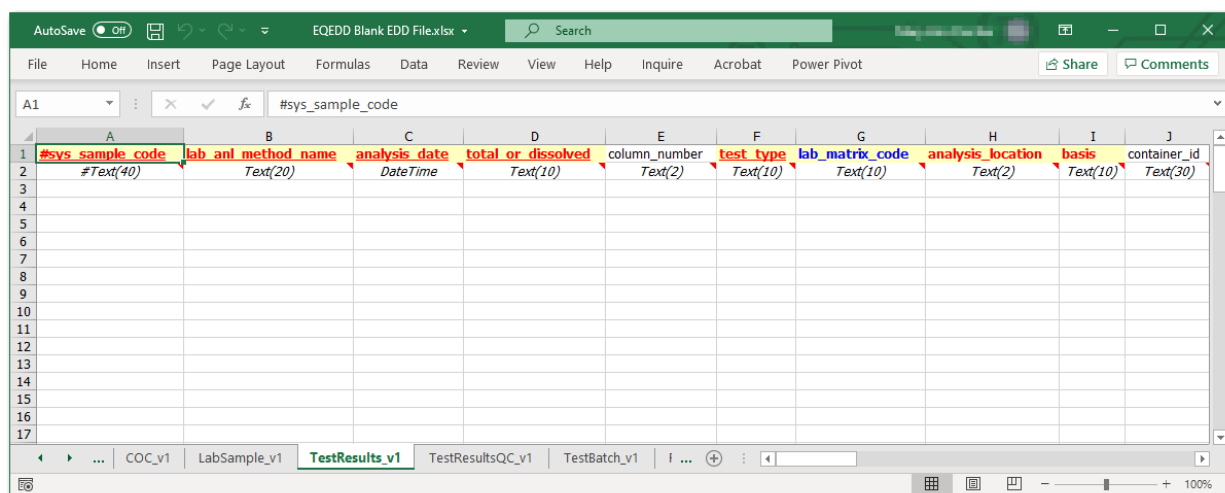
The Blank EDD tool creates an empty Microsoft Excel template that follows the EQuIS format file's specific layout for data entry. The Excel file contains the format columns and field types, along with the color-coding as displayed in EDP to indicate required fields, primary keys, and look-up values. Separate worksheets are named for each section of the format file. The Blank EDD file may be loaded into EDP for data checking after data have been entered.

To generate a Blank EDD file for a particular EQuIS format:

1. Open **EDP**.
2. Ensure that the desired format is loaded.
3. Click on the **Blank EDD** button in the Tools group on the Home ribbon.



4. In the Save dialog box, select the desired folder and name the Blank EDD file as desired. By default, the EDD Description file will be named after the EDD format (e.g., *EQEDD.xls*) and saved to the same directory as the EDD format. Click **Save**.
5. Excel will open the file, which has separate worksheets for each section of the EDD format file.



3.3 Reference Value Files

For each EQuIS database, reference values are defined and can be exported to a reference value file (RVF). Reference values are required to process an EDD for errors in EDP. When an EDD is loaded into EDP for error checking, the desired reference values must also be available.

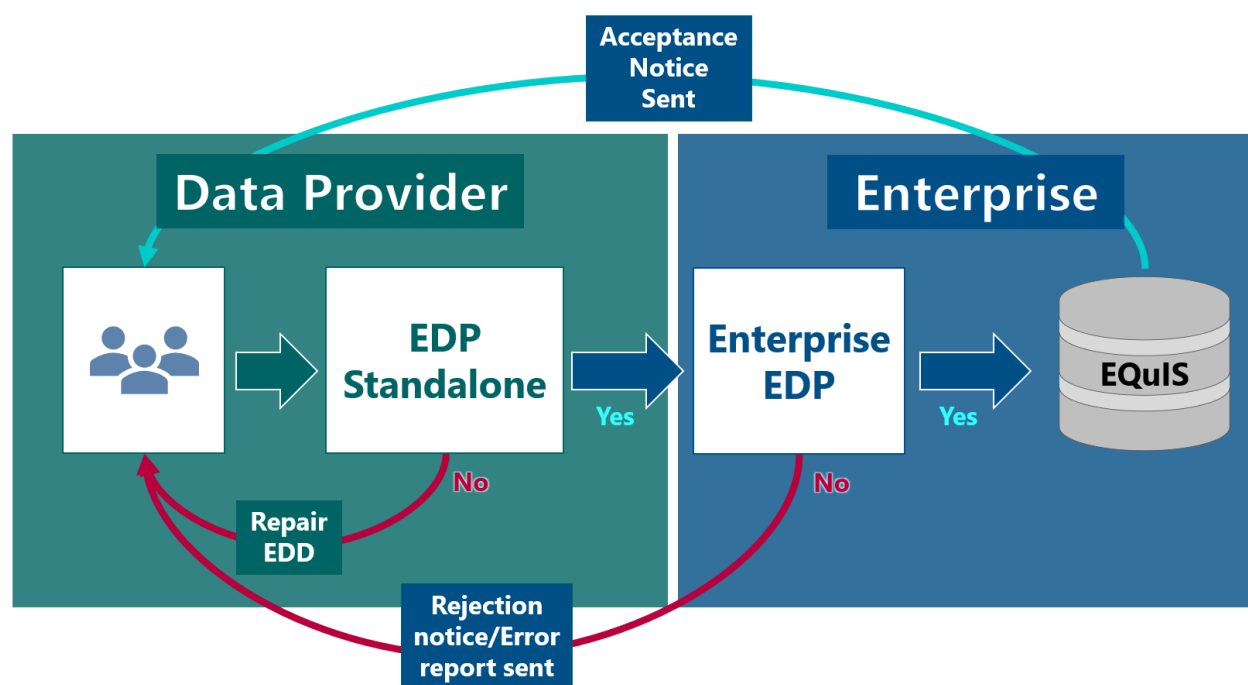
Since EDP Standalone (see Section 4) is not linked to a specific EQuIS database, the desired RVF must be loaded into EDP. The reference values used by each EQuIS format are different and vary by client. Thus, a separate RVF may be needed for each client and their selected EQuIS format. Clients should provide the most recent RVF to use in EDP Standalone or the LIMS.

4 EQuIS Data Processor Workflow

The EQuIS Data Processor, known as EDP, is a desktop or web interface for checking and importing data into the EQuIS database. There are different ways to interact with EDP:

EDP Standalone – Data providers (such as laboratories) check EDDs and resolve errors prior to submitting those EDDs to their client or an EQuIS database. EDP standalone is a desktop application not connected to an EQuIS database. Typically, EDDs that pass EDP Standalone checks are submitted to EQuIS via Enterprise EDP.

Enterprise EDP – Web application that performs the same checks as EDP Standalone, but the entire process may be automated. The purpose of Enterprise EDP is for the automated workflow to receive (via FTP, email, or web) and process EDDs. Acceptance and/or rejection notifications are automatically sent. Enterprise EDP expects incoming EDDs to be compressed (zipped) files containing the EDD data files and the data submitter's user certificate (*.usr).



Once the laboratory EDD has passed checks and been packaged in EDP Standalone, the client will determine how the laboratory EDD is submitted to EQuIS. Options include:

- Email the .zip file directly to the client (or specified representative) who will manually handle the upload process.
- Email the .zip file to the client specified email address that is part of the automated Enterprise workflow.

- Transfer the .zip file to the client specified FTP site that is part of the automated Enterprise workflow.
- Log onto the client Enterprise site and upload the .zip file using the EDP EDD Upload widget.

Although the laboratory EDD may have passed EDP Standalone, the data package may still be rejected by Enterprise EDP if database conflicts exist or needed parent records (e.g., sampling locations) are not yet in the database. Thus, the EDD submittal process may be iterative. The data error resolution process should be determined with each client.

4.1 EQuIS EDD Package Naming Conventions

As previously stated, Enterprise EDP expects incoming EDDs to be compressed (zipped) files containing the EDD data files and the data submitter's user certificate (*.usr). Email and FTP submittals must conform to the requirements outlined below. Failure to conform to these rules will result in the EDD being rejected by the Enterprise EDP workflow service.

4.1.1 ZIP File Name

Critical information is encoded in the .zip archive file name. In addition to the .zip file name extension, Enterprise EDP is looking for three values: EQuIS Format, Facility Code, and a unique identifier. The zip file name should have the following nomenclature:

<Unique ID>.<FACILITY_CODE>.<Format Name>.zip

where,

<Unique ID> = A unique identifier, such as the date or Sample Delivery Group name. The nomenclature for the unique identifier is frequently specified the client.

<FACILITY_CODE> = The facility code for the project (i.e., facility) to which this EDD will be loaded. This is specified by the client.

<Format Name> = The EQuIS registered format definition (e.g., EFWEDD, EQEDD, EZEDD, or a custom format file).

4.1.2 Flat Text File Name for EDDs within the ZIP Package

When individual text files are included within the zipped EDD package, each file should have the following nomenclature:

<Unique ID>.<EDD Section Name>.txt

where,

<Unique ID> = A unique identifier, such as the date or Sample Delivery Group name. The nomenclature for the unique identifier is frequently specified the client.

<EDD Section Name> = The name of the section within the EDD (e.g., EFW2FSample, EFW2LabTST, etc.).

For example, a zip file may be received/named "20060425.SPRINGFIELD.EFWEDD.zip" that contains:

```
"20060425A.EFW2FSample.txt"
"20060425A.EFW2LabTST.txt"
"20060425A.EFW2LabRES.txt"
"20060425A.EFW2LabBCH.txt"
"user.usr"
```

where

20060425A = The unique identified representing April 25, 2006 first submittal (A)

Springfield = Facility Code

EFWEED = EQuIS Format

4.1.3 EQuIS EDD File Naming Convention Rules

The following .zip file naming rules must be adhered to for the EDD to properly load into the EQuIS database:

- Between each of the file name elements is a "." (period). It is very important that it is a period and not a "-" (dash), "_" (underscore), or any other character.
- The order of the ZIP File Name elements specified above is required: Unique ID, followed by FACILITY_CODE, followed by Format Name.
- If the FACILITY_CODE is not entered exactly, or if the period is not included between the elements, the EDDs will be rejected. This is because Enterprise will not recognize which facility to load the data in the EQuIS database.
- If the FACILITY_CODE contains any of the following characters, the EDD package cannot be processed by Enterprise EDP:

. : * ? " < > | \ /

- The files contained within the zipped file must include the EDD Section name. There may be more information in the names, as long as the EDD Section name is separated from the other information with a period.
- The order of the Flat Text File EDDs Name elements specified above is required: Unique ID followed by EDD Section Name. The EDD Section Name must be the last entry prior to the ".txt" file extension.
- Do not use the following characters in the EDD package name or file path, as the system will not appropriately interpret the file path:

. : * ? " < > | \ /

4.2 Processing EQuIS EDD in EDP Standalone

The following steps provide a brief overview of processing the laboratory EDD in EDP Standalone. For detailed information on EDP, please refer to the online documentation [here](#).

1. Open EDP Standalone.
2. Load the EQuIS Format and RVF specified by the client.
3. Load an EDD generated from the LIMS that matches the selected EQuIS Format and RVF specified by the client.
4. Resolve errors. Problematic EDDs are rejected and flagged for specific errors. EDD tables with errors are denoted in red, and individual fields within the EDD tables that have errors are highlighted with different colors, which signify the error type and facilitates correcting errors. All errors must be resolved prior to uploading the data to EQuIS. When errors in data files must be resolved by someone other than the person checking the data, the error log is a useful tool to communicate the errors occurring in the file.
5. Sign and Submit. Use the built in **Sign and Submit** function within EDP Standalone to create an EDD package automatically and prepare it for submission to EQuIS Enterprise. The EDD package consists of a .zip file with individual .txt files for each section (i.e., table) in the EDD as well as a user certificate. The .zip file name reflects the date and time, facility name, and format file used to create the file. Adjust the file name if needed to meet client specifications for the unique identifier.
6. Send the .zip file to the client based on the client workflow.