Implementation Guidelines for SDWARS/UCMR

Volume IV: XML Format

EP001S4
EPA 816-R-01-022C

December 2001

Contents

FIGURES

TABLES

Chapter 1
Introduction

This volume documents the specifications for submitting data for the Unregulated Contaminant Monitoring Rule (UCMR) program to the U.S. Environmental Protection Agency (EPA) using Extensible Markup Language (XML) standards. Only laboratories that are registered with EPA's Central Data Exchange (CDX) can submit data to the Safe Drinking Water Accession and Review System/UCMR (SDWARS/UCMR), the information system that supports the collection of data for the UCMR.

Preliminary analytical data are entered by laboratories on behalf of the Public Water Systems (PWSs) participating under the rule. The PWS can then review their data online. Data approved by the PWS are then available for state and EPA review before being transmitted to the National Contaminant Occurrence Database (NCOD). To learn more about CDX and how to register, see Volume I, or contact the CDX Technical Support by phone (1-888-890-1995) or by e-mail at <EPACDX@lmi.org> with "UCMR" in the subject line.

INTENDED USERS OF GUIDELINE

This implementation guideline (IG) describes the requirements and procedures necessary for UCMR participants to transmit XML through the EPA's CDX to the SDWARS. The scope of this IG is solely to transmit monitoring data from laboratories for PWSs to EPA's SDWARS/UCMR through the CDX.

ORGANIZATION OF THE GUIDE

This IG is the fourth volume of a 5-volume set. This volume gives details about the formatting requirements for XML to be used with the EPA-provided document type definition (DTD) to create a well-formed, valid XML document. The other volumes are listed below:

- *IG Volume I*—introduces the CDX and electronic reporting and options for submitting data for UCMR.
- *IG Volume II*—describes completing and submitting web forms in detail.
- *IG Volume III*—contains details about Electronic Data Interchange (EDI) in case EPA decides to exchange monitoring data using EDI (not applicable to UCMR).

- *IG Volume V*—documents in detail the formatting requirements for a delimited flat file.

RESPONSIBLE ENTITY

These guidelines are published under the authority of

Office of Environmental Information
Information Collection Division
Central Receiving Branch
1200 Pennsylvania Ave, NW, Mail Stop 2823
Washington, DC 20460

The Office of Environmental Information (OEI) helps ensure EPA collects high-quality environmental information and makes it available to the public. OEI provides guidance to assist EPA in the collection, management, analysis and dissemination of environmental information.

Within OEI, the Office of Information Collection (OIC) is the agency lead for information collection programs, including how EPA obtains and manages information. The OIC works closely with many partners, stakeholders, facilities, other federal agencies and states. CDX, which is managed by OIC, is the EPA's new infrastructure for supporting the exchange of environmental data between EPA and its external partners.

Over the next several years, CDX will expand to become the point of entry for nearly all environmental data submissions to the agency. It will also improve collection, management, and sharing of environmental information among states, tribes, and EPA so they can achieve their respective and shared environmental goals.

The CDX Technical Support staff will answer questions about XML for data monitoring or compliance reporting.

Support for electronic reporting is available via e-mail, fax, or telephone.

- *By telephone.* Person-to-person telephone support is available between 8:00 a.m. and 6:00 p.m. (EST/EDT) on our toll-free line at 1-888-890-1995.

- *By fax.* You may request assistance 24 hours a day; support personnel will return calls between 8:00 a.m. and 6:00 p.m. (EST/EDT). Our fax number is 703-917-7105.

- *By e-mail.* Send e-mail to <EPACDX@lmi.org> with "UCMR" in the subject line. Responses will be sent between 8:00 a.m. and 6:00 p.m. (EST/EDT).

HOW TO USE THIS VOLUME

This guideline should be used to supplement your knowledge of XML and document type definitions (DTDs). This guideline presents EPA's environmental monitoring data system and describes tools for implementing it.

You can find general information about the CDX and electronic reporting in Volume I. The remainder of this volume is structured as follows:

- *Chapter 2—Business Issues*. This chapter discusses business issues of standards and maintenance.

- *Chapter 3—Legal and Security Considerations.* This chapter discusses legal and security considerations.

- *Chapter 4—Environments*. This chapter explains the communications and software requirements for submitting data to CDX.

- *Chapter 5—Error Corrections and Resubmissions*. This chapter explains correcting errors and resubmitting data.

- *Appendix A—UCMR XML DTD, Version 2.1*. This appendix lists the UCMR DTD.

- *Appendix B—UCMR File Structure Tree Diagram*. This appendix provides a graphical presentation of the UCMR file structure.

- *Appendix C—UCMR Data Dictionary: XML DTD, Version 2.1.* This appendix is a data dictionary.

- *Appendix D—UCMR XML Instance*. This appendix provides an example XML file based on the UCMR XML DTD, Version 2.1.

- *Appendix E*—Abbreviations.

HOW TO GET MORE INFORMATION

EPA has websites that you may find useful in submitting a UCMR electronically:

- The CDX website is at <http://EPACDX.lmi.org>. Only registered users can access this site. To gain access, register by using the process described in Volume I.
- General public information about the EPA CDX is at <http://www.epa.gov/cdx>.
- The Office of Ground Water and Drinking Water (OGWDW) maintains a web page about UCMR at <http://www.epa.gov/safewater/ucmr.html>.

Chapter 2
Business Issues

This section describes when data may be submitted and the basis for using XML specifications. It also introduces the structure of the submission format which is further defined in Appendixes A through C.

THE XML DOCUMENT

This IG is not an XML tutorial, and it is assumed that you are reasonably knowledgeable about XML operations. If you want to learn more about XML, a number of commercial and nonprofit sources are available. A place to start would be the World Wide Web Consortium (W3C). The W3C develops specifications, guidelines, software, and tools that lead to recommendations that are widely adopted as the standard for web developers and users. *XML 1.0 Recommendation* (published in February 1998) is the recommendation employed for defining the UCMR XML DTD. For more information about current W3C recommendations, visit their website at <http://www.w3.org>.

Only laboratories will submit electronic files to the SDWARS/UCMR database. Files will contain data about batches (analytical or extraction) and sample results. After you submit a file, the CDX will return a message to your "MyCDX" inbox that the file has been successfully processed, or that errors have been detected and the file has been rejected. A copy of the message will also be sent to your e-mail address you registered with CDX. You should correct rejected files and resubmit with a different file name (Chapter 5 discusses error correction and resubmissions). Depending on specific codes placed in the file and the results of SDWARS/UCMR range check validations (see Chapter 5 for details of data range checks), you may need to log into SDWARS/UCMR and further process the data through web forms (see Volume II for more on web form processing).

File-Naming Convention

The file name must be unique from all other files submitted through CDX. The file name must follow the format: "UCM + (Laboratory ID) + (unique alphanumeric identifier assigned by your laboratory).xml" (The alphanumeric identifier may be an incremental counter, such as "UCMIL00028001.xml."). The length must not exceed 40 characters and may include an underscore "_", but not spaces or symbols (e.g., $, @, &, %).

Your laboratory ID begins with the two-letter postal code for your state followed by a five-digit numeric code assigned by the EPA during a past performance

evaluation (PE) study (e.g., FL12345). If you do not know your lab ID, please contact the Safe Drinking Water Hotline (see Chapter 1).

Timing of Transactions

The CDX system operates 24 hours a day, 365 days a year. See the CDX website for any specific changes in operating hours. Each file submitted will result in a processing confirmation record being placed in the your "MyCDX" inbox, as well as your e-mail address registered in your CDX account. You are encouraged to check your "MyCDX" inbox regularly, as well as follow the detailed instructions for verifying submissions in Volume II.

Submissions typically will be uploaded to SDWARS/UCMR within 30 minutes after they are received, but may take longer. In order to ease demands on system resources from peak periods (7:00 a.m.–10:00 p.m. EST), CDX has a file filter that sets aside relatively larger files (above 200 KB) for overnight processing. If after your file is uploaded to CDX the filter determines that the file exceeds the daytime processing maximum, CDX will send you a message to your "MyCDX" inbox indicating the file will be processed overnight.

Maintaining This Implementation Guideline

EPA will be responsible for maintaining, updating, and distributing the IG as needed. EPA will work with all CDX stakeholders to maintain a XML architecture that balances accommodating the bulk of users with incorporating new techniques and technologies. Submitters will be responsible for ensuring that their process remains consistent with this guideline. If a system is modernized, data formats and processes must remain consistent with what is defined in this IG and the UCMR program. EPA will coordinate changes with UCMR stakeholders before revising this IG and its content.

DOCUMENT STRUCTURE

Document Type Definition

The UCMR XML implementation uses a DTD developed specifically for submitting monitoring data. The DTD describes the format and the data elements required for submitting valid, well-formed monitoring data in XML to SDWARS/UCMR. The DTD is in Appendix A; a tree diagram of the hierarchical file structure is in Appendix B; and a data dictionary is in Appendix C.

UCMR File Structure

This section describes the logical structure of the UCMR submissions and the details of using the data for implementing electronic reporting. In addition, this section describes the business use, and explains the elements for mapping to

and from other application systems. A full explanation of the elements is in Appendix C.

The format of the UCMR hierarchical file structure is defined in Table 2-1. The requirement column indicates if the element is required or optional and the occurrence column indicates how many times an element may occur.

For example, 0-*n* means that the element can occur zero or more times (with no maximum upper limit) or that it is optional. An occurrence of 0-1 means the element may occur once or not at all for each occurrence of the parent element. A value of 1 means that the element must occur once for every occurrence of a parent element. An occurrence value of 1-n means the element must occur at least once for each occurrence of a parent element, but may occur more times. The condition column indicates any conditions that apply to the element.

Table 2-1. UCMR PWSS Structure

UCMR file structure	Requirement	Occurrence	Condition
UCMR_PWSS	Required	1	
Header_Data	Required	1	
Base_Header_Data	Required	1	
CDX_Identification	Required	1	
Schema_Version	Required	1	
Environment	Required	1	
Report_Type	Required	1	
Customer_Header_Data	Required	1	
Transaction	Required	1	
Transaction_Purpose	Required	1	
Transaction_Date	Required	1	
Transaction_Time	Required	1	
Detail	Required	1	
Lab_Id	Required	1	
Batch	Optional	0-*n*	
Batch_Id	Required	1	
Extraction_Analysis_Date	Required	1	
Analytical_Method	Required	1	
Analyte	Required	1-*n*	
Analyte_Code	Required	1	C1
Spiking_Concentration	Required	1	
Analytical_Precision	Required	1	
Analytical_Accuracy	Required	1	

Table 2-1. UCMR PWSS Structure (Continued)

UCMR file structure	Requirement	Occurrence	Condition
PWS	Optional	0-*n*	
PWS_Id	Required	1	
Facility	Required	1-*n*	
Facility_Id	Required	1	C2
Facility_Sample_Point	Required	1-*n*	
Sample_Point_Id	Required	1	C3
Sample_Point_Sample	Required	1	
Sample_Id	Required	1	
Sample_Collection_Date	Required	1	
Analysis_Type	Required	1	
Lab_Sample_Comment	Optional	0-1	
Analysis	Required	1-*n*	
Analyte_Code	Required	1	
Batch_Id	Required	1	C4
Analytical_Method	Required	1	C4
Analysis_Result	Required	1	
Value	Optional	0-1	C5
Result_Sign	Required	1	C5
Presence	Optional	0-1	C6
Anaylsis_Status	Optional	0-1	
Reviewer_Status	Required	1	
Lab_Result_Comment	Optional	0-1	

Table 2-2 defines the codes in the Condition column of Table 2-1.

Table 2-2. Conditions Applicable to Elements

Code	Description
C1	Analyte_Code must represent an analyte that can be tested for by the preceding Analytical_Method.
C2	Facility_Id must represent a facility that is recorded in SDWARS/UCMR as being associated with the preceding PWS_Id.
C3	Sample_Point_Id must represent a sampling point that is recorded in SDWARS/UCMR as being associated with the preceding Facility_Id.
C4	As part of the analysis tag, the Batch_Id and Method_Id must relate to data in a preceding batch or that are already entered in SDWARS/UCMR.
C5	If value is null, then Result_Sign must be “lt.” If value is not null, the Result_Sign must be “eq.”
C6	Presence is not utilized at this time for UCMR reporting of List 1 analytes.

Maintaining the DTD Version or Release

The DTD in Appendix A is for a particular version of the XML standard. EPA will continue to operate in that version until business requirements indicate that the DTD needs to be revised and updated to a newer DTD or schema version. At a minimum, EPA will coordinate the changes through OGWDW. Updates are posted on the EPA CDX website under "What's New."

Submitters must use XML software that is fully compliant with the W3C recommendations for the version noted in the DTD, and that can support creating an XML document that is valid and well-formed.

Chapter 3
Legal and Security Considerations

Electronic reporting of UCMR data reduces the burden of reporting, collecting data, and record keeping for both reporting facilities and responsible environmental agencies by eliminating the labor, time, and other costs of submitting paper reports. Electronic reporting does not lessen or alter any of the responsibilities or liabilities under good business practices.

AUTHENTIC SUBMISSIONS

Drinking water regulations do not require you to electronically sign data submitted to SDWARS/UCMR. However, your laboratory should consider the submissions as your official copy of record.

UCMR business rules require laboratories to mark each sample "approved" before the responsible PWS can review it. In turn, each PWS must mark a sample "approved" before a state or EPA entity can review it, or before it can be moved onto other EPA data systems.

You can mark uploaded files as "approved." If successfully uploaded, these files will move directly to the PWS review process. If rejected, you must correct and resubmit them. However, there are also certain data range checks that the system will review. If data fall outside of the range limits, SDWARS/UCMR will alter the status from "lab approved" to "lab hold" (see Chapter 5 for details of data range checks). For such records, you must use the web form to manually approve the records.

ELECTRONIC REPORTING

CDX Registration

You must be registered with CDX before submitting an XML document. Refer to Volume I of these IGs to review the registration process.

XML Submission

The XML submission itself does not have to be authenticated. However, EPA expects you to maintain proper safeguards and security over the computer systems that will generate the XML document. At a minimum, an authorized official must use a secure means to release the UCMR XML document. The security can

include using password- or token-based entry into the system to release the completed monitoring data so the XML can be generated.

Approving Submissions

If you are recognized by CDX as having the Approver role for your lab, you may access the website and approve your lab's data. To have the Approver role you must select that role in CDX and be confirmed through your lab's sponsor letter. After the data are approved by the lab and PWS, they are considered to be the official data for SDWARS/UCMR. See Volume I of these IGs for more information about registering.

ELECTRONIC RECORD KEEPING

Regardless of how the facility keeps records, your monitoring application system and its XML system should at least adhere to procedures described below.

Backup and Recovery Procedures

The systems should back up all data and programs at least daily. Backup media should be stored off the site. XML systems should maintain archives of all transmissions, sent and received, which should be readily accessible for at least 30 days to ensure that they can be retransmitted if requested by EPA or other regulatory agencies. You should maintain logs of all transmitted data and verify that EPA received the transmission.

Alternative plans should be developed to accommodate unforeseen problems, such as loss of a data center or local phone system, or a catastrophic act of nature that prevents exchanging transmissions for an extended period. Alternatives include using remote backup systems temporarily or using different third-party service providers.

Audit Considerations

You should maintain an adequate audit trail to ensure that they can substantiate, when needed, information exchanged electronically. In an information systems environment, an audit trail typically focuses on the transactions in the system—the data processed, input/output devices accessed, and the date and time that activities occurred. Documents in paper form usually are available to validate information input or output from information systems. However, in an electronic environment, paper versions of data do not exist. Therefore, an audit mechanism for the XML environment should be more comprehensive to substantiate that the information was transmitted and received electronically. An XML audit trail is a full set of records (maintained in either electronic or paper format) documenting the data received, sent, retained, and stored. This set of records should accurately reflect the events as they occurred.

Procedure Documentation

You should maintain current and detailed documentation of your backup and disaster and recovery procedures. Also, you should document record-keeping procedures (either on paper or electronically).

Chapter 4
Environments

This chapter summarizes the XML data flow between you and the CDX. The chapter begins with a discussion of the communications path between the your computer and CDX.

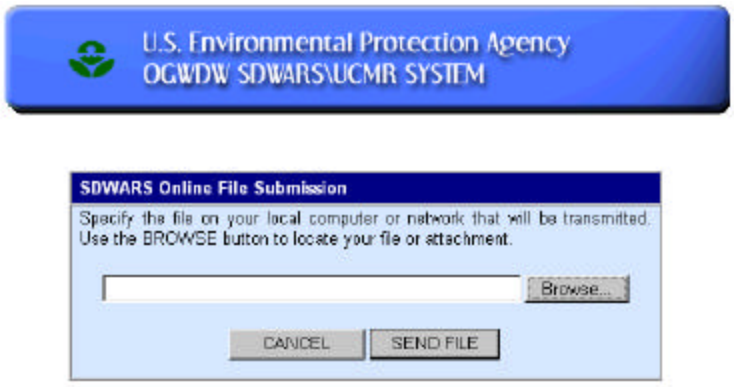
COMMUNICATIONS

Once a file is generated, it can be uploaded through the CDX website. Once you have established an account with CDX and your sponsor letter is processed at CDX, you will have access to the UCMR: File Upload link in the "Available Account Profiles" at the bottom of the "MyCDX" account web page (see Figure 4-1).

Figure 4-1. CDX Account Web Page

The UCMR: File Upload link will take you to the UCMR file upload web page (see Figure 4-2). On the UCMR File Upload screen you may enter the directory location of the flat file or click on the Browse button to locate the file. Once the file has been located, you may click on the Send File button.

Figure 4-2. File Upload

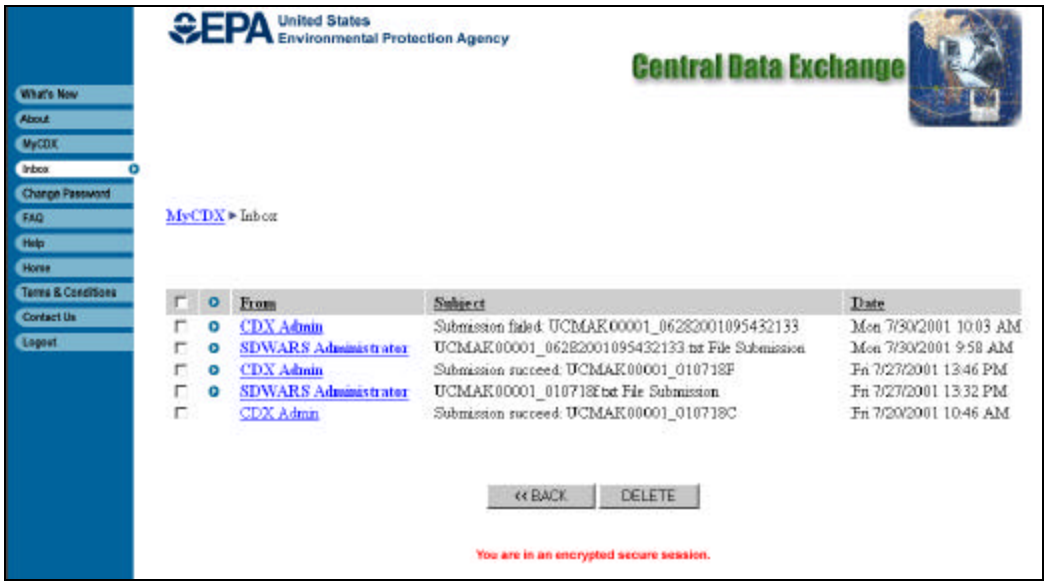


You will receive an acknowledgement that CDX received the file on the screen (see Figure 4-3) and in your "MyCDX" inbox, and to your registered e-mail address. If you encounter problems with uploading the file, contact CDX Technical Support.

Figure 4-3. File Submission Confirmation

Depending on file size and system activity it may take a few minutes for CDX to upload your file. Once uploaded, CDX will send you an upload confirmation message or an error notification to your "MyCDX" inbox (see Figure 4-4). A copy of the error notification will be sent to your registered e-mail address as well. Error messages are discussed in Chapter 5. If you receive a successful upload confirmation message (Figure 4-4), your data has been loaded to CDX and has been queued for processing into SDWARS/UCMR.

Figure 4-4. CDX Inbox with Successful Upload Message



SOFTWARE REQUIREMENTS

You may implement XML parsing and validating according to your Laboratory Information Management System (LIMS) or other information system architecture. The XML products that are used should be compliant with the W3C recommendation for XML, the version noted in the DTD, and be capable of validating the XML document against the DTD. Information about W3C recommendations is available at <http://www.w3.org>.

UCMR XML DATA FLOW

The first step to creating an XML document is to prepare the extraction of the data from the LIMS. Some data transformation software may require you to extract a delimited flat file from the LIMS and parse (map) the flat file elements to the elements of the DTD. Other XML products have capabilities to communicate directly with databases and can parse to the DTD straight from the database tables. Your provider of XML software and LIMS should be consulted about the capabilities of the software. Another consideration is the ability of the software to validate XML documents. Validation ensures the XML document conforms to the specifications for XML and DTD.

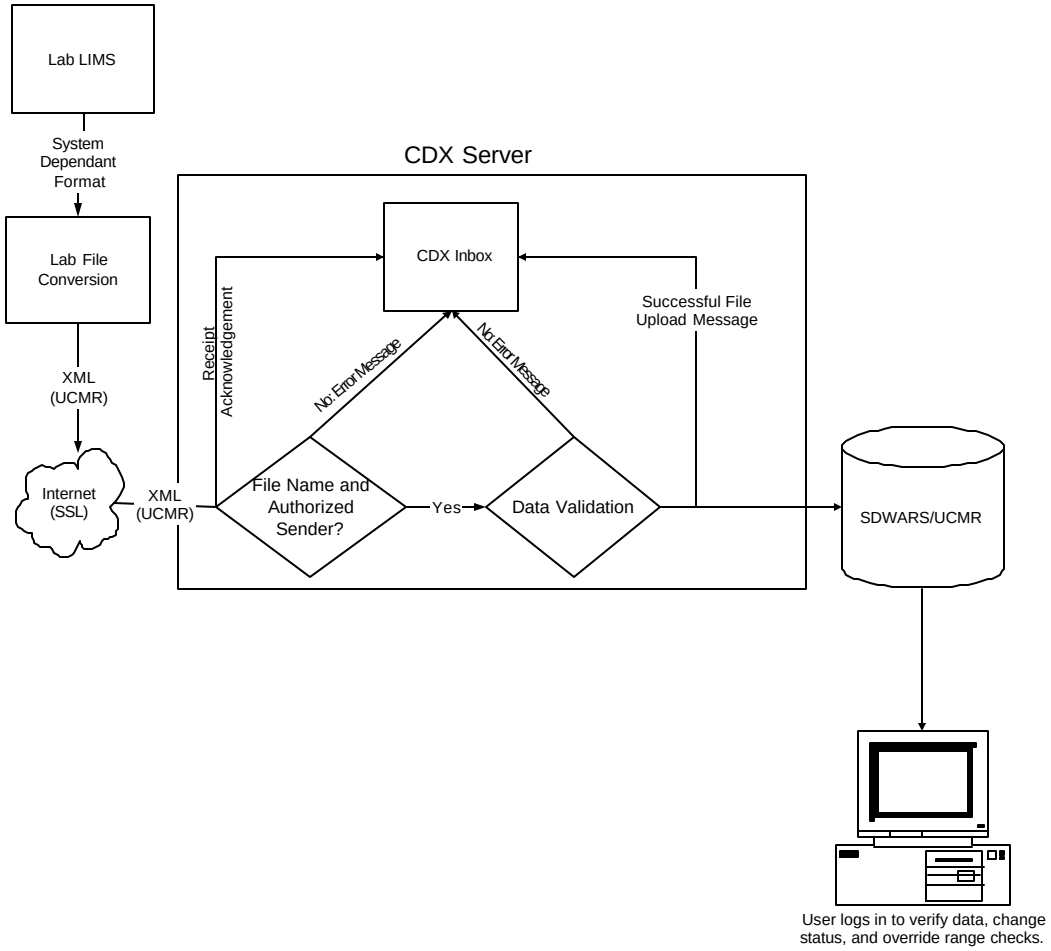
When an architecture for extracting the data has been determined, the data must be parsed into an XML document according to the UCMR DTD (in Appendix A). EPA will maintain the most recent UCMR DTD on the CDX website. The lab should validate the resultant XML document with the structure of the DTD to ensure it is valid and well-formed.

The validated XML document is uploaded to CDX as described above. CDX will poll the file submission directory on the CDX server and send XML files to the CDX parser for validation and inputting into SDWARS/UCMR. Figure 4-5 represents the processing of a UCMR XML submission.

If errors prevent the processing of the submission, CDX will send an error message to your "MyCDX" inbox and registered e-mail address (Figure 4-4). After making the corrections, the data can be resubmitted (refer to Chapter 5). You can call the CDX Technical Support for additional help.

Once the data are loaded into SDWARS/UCMR, you will be able to enter SDWARS/UCMR via your CDX account to view the data (see IG Volume II for a discussion of web access). If you applied an "approval" status in the Reviewer_Status element of the XML document, you will only be able to view the data online through the search function. If you applied a "hold" in the Reviewer_Status element or don't provide any status in the XML document, you will be able to view the data online and make edits. A PWS cannot review your laboratory data in SDWARS/UCMR until someone from your laboratory has approved the submitted data.

Figure 4-5. UCMR XML Process



Chapter 5
Error Corrections and Resubmissions

CORRECTING ERRORS

If a submission contains errors that prevent it from being parsed into SDWARS/UCMR, CDX will notify you of the error by sending a notification to your "MyCDX" inbox (see Figure 5-1) and your registered e-mail address.

Figure 5-1. CDX Inbox with Error Notification

Data checks occur at two stages during the file upload process—authentication and data validation. Authentication occurs as soon as CDX receives the file. CDX authenticates your account and the file name. These messages are sent to your "MyCDX" inbox and your e-mail address registered with CDX. Table 5-1 shows the error messages you could receive.

Table 5-1. Authentication Error Messages

Error message
Archive1 fail. File name larger than 75 characters.
File: gggggggggggggggg was not accepted. The user ID yyyyyyy is not authorized to submit for lab id: xxxxxxxxxxxxx. For questions, contact CDX technical support at 1-888-890-1995.
The file submitted for UCMR Lab ID xxxxxxxxxxxxxxxxxxx has some problems. Please access your “MyCDX” inbox for further information. Thank you for your cooperation.

The data validation function performs checks by the parser and additional checks in two parts—format and UCMR validations. There are two different error messages for XML—parsing errors and data validation errors. CDX will parse the file. If the file is not valid and well-formed, CDX will provide a text file listing the parsing errors (see Figure 5-2). This file is sent to your “MyCDX” inbox. While a copy of the error notification will be sent to your registered e-mail address, you can only view the list of errors through the error notice in your “MyCDX” inbox. The parsing error messages are generic messages. If 50 errors are found, CDX stops parsing the file and sends the error message, noting that the file may contain additional errors.

Figure 5-2. Error Message from Parser

An error occurred while processing a file submission to UCMR.

Submission information:

File name: UCMEP00005005-AMOFLYNN1.xml

Error date: Jul 11, 2001

Error time: 3:11:41 PM

For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.

4 Error(s) occurred during parser validation. Please verify you are using PWSS DTD verison 2.1 and the case of your tags. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.

Line Number, 785. Error Message, End tag does not match start tag 'Analysis'.

Line Number, 785. Error Message, Element Analysis not complete, expected elements '[Batch_Id]'.

Line Number, 785. Error Message, End tag does not match start tag 'Sample_Point_Sample'.

Line Number, 785. Error Message, Unexpected EOF.

If the file is valid and well-formed, CDX will conduct data validation and verifications according to UCMR. If the file does not pass the checks, CDX will provide an error message to your "MyCDX" inbox. Here too, a copy of the notification will e sent to your e-mail address directing you to view the error details online. A copy of the error notification is available to the CDX Technical Support, in case the submitter has questions about the error report.

The message (see Figure 5-3) is numbered and provides an explanation, the extract of the record that contains the error, and an administrator message. The administrator message is applied to assist technical support when typical solutions do not work. The technical support staff may request the associated administrator message to assist them and the programmers in identifying a solution to an error. The text file lists up to the first 50 errors in the submitted file. Message number one of Figure 5-3 does not include an administrator message while message number two does.

SDWARS/UCMR verifies the data by first checking general format and verifies IDs (i.e., correct format for date, batch ID, valid PWS ID, valid method, valid sampling point ID). Table 5-2 lists common error messages for data formatting and general validations. These messages identify the probable cause for the error and recommended corrective actions.

Figure 5-3. Text File with Details of Errors

An error occured while processing a file submission to UCMR.

Submission information:
File name: UCMAK00001_06292001113952750-JKELLOG1.txt
Error date: Jun 29, 2001
Error time: 11:52:6 AM

For further assistance, please contact the CDX Technical Support staff at 1 888 890-1995.

1. ERROR MESSAGE EXPLANATION:
SPK_CONCENTRATION must equal N/A or be greater than 0. For further assistance, please contact the CDX Technical Support staff at 1 888 890-1995.

LAB_IDENT_CD: AK00001
BATCH_IDENT_CD: AST2251887
METHOD_CD: ASTM D5790
ANALYTE_CD: 2254
SPK_CONCENTRATION: 0
RESULT_PRECISION: 71
ACCURACY: 19
EXTRACTION_ANALYSIS_DATE: 24-Jan-01
USER_LOGIN_ID:
LAST_CHANGE:

2. ERROR MESSAGE EXPLANATION:
The Batch associated with this Sample does not exist. Please add the Batch data or change the Batch ID and resubmit your data. See BATCH_ID below. For further assistance, please contact the CDX Technical Support staff at 1 888 890-1995.

LAB_IDENT_CD: AK00001
BATCH_IDENT_CD: ASTOUNDING
METHOD_CD: ASTM D5475
SAMPLE_IDENT_CD: AST2052644164
ANALYTE_CD: 2272
ANAL_RESULT_VALUE: 6
ERROR_EXIST:
LAB_REVIEW:
LAB_REVIEW_DATE:
PWS_REVIEW:
PWS_REVIEW_DATE:
STATE_REVIEW:
STATE_REVIEW_DATE:
EPA_REVIEW:
EPA_REVIEW_DATE:
SENT_TO_EPA_DATE:
PRESENCE_ABSENCE:
LESS_THAN_MRL: 0
STATE_REGION_ID: 99
PWS_ID: 990000001
FACILITY_ID: 00002
SAMPLING_POINT_ID: 4354
STATUS_CD: 10
USER_LOGIN_ID:
LAST_CHANGE:

Administrator Message:
ORA-02291: integrity constraint (UCMR.FK_FIELD_ANALYTES_BATCH_QC) violated - parent key not found

Table 5-2. Common Formatting Data/Verification Error Messages

Error message	Administrator message
The Batch ID and Method combination already exist in SDWARS/UCMR. Please change the Batch ID or Method and resubmit your data. See BATCH_IDENT_CD and METHOD_CD: below. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	ORA-0001 unique constraint (UCMR.PK_TableName) violated.
Your Lab ID is not known. Verify your Lab ID is correct and resubmit. See LAB_IDENT_CD below. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	ORA-02291 integrity constraint (UCMR.FK_TableName_TableName) violated—parent key not found.
This Method is unknown to SDWARS/UCMR. See METHOD_CD below. Verify the method and resubmit your data. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	UCMR.FK_BATCH_QC_METH0D_ ANALYTE_IDS
The analyte and method combination is unknown to SDWARS/UCMR. See ANALYTE_CD and METHOD_CD below. Verify the analyte and method, then resubmit your data. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	UCMR.FK_BATCH_QC_METH0D_ ANALYTE_IDS
The record below shows that this Method is not associated with this Analyte in SDWARS/UCMR. Please verify the METHOD_CD and ANALYTE_CD are correctly associated. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995	UCMR.FK_BATCH_QC_METH0D_ ANALYTE_IDS
The Sample ID and PWS ID combination already exists. Please change the Sample ID and resubmit your data. See SAMPLE_IDENT_CD below. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	ORA-00001 unique constraint (UCMR.PK_TableName) violated.
Your Lab ID could not be found in SDWARS/UCMR. Please verify your Lab ID and resubmit your data. See LAB_IDENT_CD below. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	ORA-02291 integrity constraint (UCMR.FK_TableName_TableName) violated—parent key not found.
The following PWS ID could not be found in SDWARS/UCMR. Please verify the PWS ID and resubmit your data. See PWS_ID below. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	ORA-02291 integrity constraint (UCMR.FK_TableName_TableName) violated—parent key not found. FK Violation (UCMR.FK.BATCH_QC_ METHOD_ANALYTE_IDS)
The following PWS ID and Facility ID combination could not be found in SDWARS/UCMR. Please verify the Facility ID belongs to this PWS and resubmit your data. See FACILTY_ID and PWS_ID below. If the Facility is correct, please contact the PWS and have them update their inventory in SDWARS/UCMR. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	ORA-02291 integrity constraint (UCMR.FK_TableName_TableName) violated—parent key not found. K Violation (UCMR.FK.SAMPLES_ PWS_FACILITY)
The PWS ID, Facility ID, and Sampling Point ID combination could not be found in SDWARS/UCMR. Please validate the Sampling Point ID belongs to this Facility and PWS, then resubmit your data. See FACILTY_ID, PWS_ID, and SAMPLING_POINT_ID below. If the PWS is registered on your client list and the Sampling Point is correct, contact the PWS and have them update their inventory in SDWARS/UCMR. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	ORA-02291 integrity constraint (UCMR.FK_TableName_TableName) violated—parent key not found. FK Violation (UCMR.FK.SAMPLES_PWS_ FAC_SAMPLING_POINT)

Table 5-2. Common Formatting Data/Verification Error Messages (Continued)

Error message	Administrator message
The Sample Type is invalid. Please change the Sample Type and resubmit your data. See SAMPLE_TYPE below. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	SAMPLES FK Violation (UCMR.FK.SAMPLES_VALID_ SAMPLE_TYPE)
The PWS ID, Sample ID, Batch ID, Method ID combination already exists for your Lab. Please validate this data or change the file to a "Replacement" and resubmit your data. See PWS_ID, SAMPLE_ID, BATCH_ID, and METHOD_ID below. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	FIELD ANALYTES PK Violation
The Batch associated with this Sample does not exist. Please add the Batch data or change the Batch ID and resubmit your data. See BATCH_ID below. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	FIELD ANALYTES FK Violation (UCMR.FK.FIELD_ANALYTES_ BATCH_QC)
Your Lab ID could not be found in SDWARS/UCMR. Please verify your Lab ID and resubmit your data. See LAB_ID below. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	FIELD ANALYTES FK Violation (UCMR.FK. FIELD_ANALYTES_LABS)
The PWS ID could not be found in SDWARS/UCMR. Please validate the PWS ID and resubmit your data. See PWS_ID below. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	FIELD ANALYTES FK Violation (UCMR.FK. FIELD_ANALYTES_PWS)
The PWS ID and Facility ID combination could not be found in SDWARS/UCMR. Please verify the Facility belongs to this PWS and resubmit your data and that the PWS is registered on your client list. See PWS_ID and FACILTIY_ID below. If the Facility is correct please contact the PWS and have them update their inventory in SDWARS/UCMR. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	FIELD ANALYTES FK Violation (UCMR.FK. FIELD_ANALYTES_ PWS_FACILITY)
The following PWS ID, Facility ID, and Sampling Point ID combination could not be found in SDWARS/UCMR. Please verify the Sampling Point belongs to this Facility in the PWS and resubmit your data. See PWS_ID, FACILTY_ID, and SAMPLING_POINT_ID below. If the PWS is registered on your client list and the Sampling Point is correct please, contact the PWS and have them update their inventory in SDWARS/UCMR. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	FIELD ANALYTES FK Violation (UCMR.FK. FIELD_ANALYTES _ PWS_FAC_SAMPLING_POINT)
Application terminated. No data was processed. Document must contain an Environment field and field value must be T or P. For further assistance, please contact the CDX Technical Support staff at 1-888-890-1995.	
Application terminated. No data was processed. Document must contain Header data and Batch QC or Sample and Result data	
Application terminated. No data was processed. Document must contain a Report Type field and field value must be UCMR to identify transaction as a UCMR report.	
Application terminated. No data was processed. Document must contain a Transaction Purpose field and field value must be o or r.	

The final validations that CDX conducts are UCMR data validations. Table 5-3 lists common error messages for UCMR data validations.

Table 5-3. UCMR Validation Errors

Error message	Administrator message
RESULT_PRECISION,ACCURACY,SPK_CONCENTRATION must equal N/A	ORA-20100
RESULT_PRECISION must equal N/A or be greater than or equal to 0	ORA-20101
New.ACCURACY||':ACCURACY must equal N/A or be greater than or equal to 0	ORA-20102
SPK_CONCENTRATION must equal N/A or be greater than 0	ORA-20103
EXTRACTION_ANALYSIS_DATE must be between January 1, 1985 and sysdate	ORA-20104
SPIKE CONCENTRATION may not equal MISSING	ORA-20105
ACCURACY may not equal MISSING	ORA-20106
RESULT_PRECISION must be a number or N/A	ORA-20107
ACCURACY must be a number or N/A	ORA-20108
SPK_CONCENTRATION must be a number or N/A	ORA-20109
The Collection Date DD-MON-YYYY must be before the Extraction Analysis Date DD-MON-YYYY	ORA-20200
For Method '||:new.method_cd||', the analytical result value must be less than MRL	ORA-20201
ANAL_RESULT_VALUE is not null, so LESS_THAN_MRL must be null or zero	ORA-20202
ANAL_RESULT_VALUE must equal to N/A or be greater than or equal to an MRL of MRL	ORA-20203
LESS_THAN_MRL is N/A, so ANAL_RESULT_VALUES must be null	ORA-20204
Either LESS_THAN_MRL or ANAL_RESULT_VALUES must be not null in order to be approved	ORA-20205
Your STATUS CODE is status_cd and STATUS CODE must be 10 or 50 in order to change the ANAL_RESULT_VALUE	ORA-20206
Your STATUS CODE is status_cd and STATUS CODE must be 10 or 50 in order to change the LESS_THAN_MRL	ORA-20207
ANAL RESULT VALUES must be a number or N/A	ORA-20208

RESUBMITTING DATA

A file containing only original data or including original and replacement data can be submitted. The analyte results data may be replaced only if the existing results have not been approved. When resubmitting replacement data, you *must* use the code "r" in the transaction_purpose element. The "r" indicates the transaction contains records that will replace existing records. **The laboratory cannot resubmit a file that has the same name as a previous submission. You must rename the file before resubmitting it.** The resubmission will overwrite the

existing data in the SDWARS/UCMR database. You may confirm that the data was overwritten by viewing the records in the SDWARS/UCMR web application.

Users from your laboratory are the only users that may edit your sampling and analytical data. You can only edit data that is in "lab hold" status in the SDWARS/UCMR database. If you need to edit data that it has already been approved, you can contact the PWS to return the control of the record back to your lab for corrections. If the record was also approved by the PWS, then the PWS must request that the SDWARS/UCMR database administrator return editing rights to your laboratory. The SDWARS/UCMR database administrator can be contacted through CDX Technical Support.

CHECKING THE RANGE OF DATA VALUES

In addition to the edit validations described in the preceding sections, SDWARS/UCMR reviews several fields for specific values or ranges. In some cases it may entirely reject the record for failing the range check, while in others it may create a warning that requires the laboratory to manually review and either approve or modify the results. This section describes each of these fields.

UCMR Data Range Checks

The UCMR has specific ranges of values for different analytes and methods. Information about MRL values is in the following regulations:

- 40 CFR, parts 9, 141, and 142, "Revisions to the Unregulated Contaminant Monitoring Regulation for Public Water Systems;" final rule, September 17, 1999
- 40 CFR, part 141, "Unregulated Contaminant Monitoring Regulation for Public Water Systems: Analytical Methods for Perchlorate and Acetochlor;" "Announcement of Laboratory Approval and Performance Testing (PT) Program for the Analysis of Perchlorate;" final rule and proposed rule, March 2, 2000
- 40 CFR, part 141, "Unregulated Contaminant Monitoring Regulation for Public Water Systems: Analytical Methods for List 2 Contaminants and Clarifications;" final rule, January 11, 2001.

SDWARS/UCMR Data Range Checks

SDWARS/UCMR will check the range of some data values. Table 5-4 shows the list of elements that SDWARS/UCMR checks. SDWARS/UCMR assigns a range or value that the element must satisfy, and the range or value it should satisfy. The "must" checks will prevent the data from loading and will result in an error notification to your "MyCDX" inbox and registered e-mail address. However, the

"should" checks will not result in preventing the data from loading. Instead, the "should" checks will flag results as having potential errors for your laboratory to confirm. **If data are submitted that violate the "should" criteria, then the status of those records will be "lab hold" and you must use the web forms to override the range check (or correct the value through web forms or a replacement file).** You cannot approve data where a range check notification has not been addressed. Thus, the data cannot be sent to the PWS to approve and will remain in the "lab hold" status.

You have two options to edit a range check. In order to override the warning message and approve the data, you must use the web forms. If, in reviewing the data, you believe the data should be corrected, you may do so either by using the web form or by correcting the data in your system and issuing a replacement transaction.

Table 5-4. Range Checks Performed by SDWARS

Element	Criteria	
	Must be	Should be
Accuracy	≥ 0 or N/A < 32000	≤ 200% ≥ 10%
Collection date	Not later than the current date Not earlier than 1/1/85	
Extraction/analysis date	Not later than the current date Not earlier than the collection date Not earlier than 1/1/85	≤ 60 days from the sample collection date
Precision	≥ 0, "MISSING"[a], or N/A < 32000	≤ 99%
Result (value)[b]	≥ MRL or N/A < 32000	< 10 x MRL
Spiking concentration	> 0 or N/A < 32000	≤ 200

Note: N/A indicates "not analyzed." If an analyte is not analyzed for a batch, then precision, accuracy and spiking concentration must be "N/A". If an analyte is not analyzed for the sample then the result value is "N/A".

[a] "MISSING" is allowed when a duplicate result was not available.

[b] If the result_sign is "eq," then the result value field must be numeric and comply with the above criteria. Otherwise the result sign is "lt" and the result value is "null." Method EPA 515.3 will not have a result value, it must always be reported as less than MRL.

There is no mechanism in this version to allow SDWARS/UCMR to issue a notice back to your CDX or registered e-mail account when a "should" condition is encountered. The only way to check for the "should" case range checks is through the online application. (See the UCMR IG Volume II for details on addressing range checks online.)

Appendix A
UCMR XML Document Type Definition, Version 2.1

```
<?xml version='1.0' encoding='UTF-8' ?>
<!-- This XML DTD was developed on behalf of EPA for reporting data related to the Unregulated Contaminant
     Monitoring Regulation (UCMR)
     Author:              Paul Macias
     Company:             Logistics Management Institute
     Contact Information: pmacias@lmi.org
     Schema Version:      2.1
     Date:                04/13/01
     -->
<!-- Acronyms used in this DTD:
     UCMR = Unregulated Contaminant Monitoring Regulation,
     CDX = Central Data Exchange,
     PWSS = Public Water System Supervision,
     PWS = Public Water System
     -->

<!ELEMENT UCMR_PWSS (Header_Data , Detail)>

<!ELEMENT Header_Data (Base_Header_Data , Customer_Header_Data)>

<!ELEMENT Detail (Lab_Id , Batch* , PWS*)>

<!ELEMENT Base_Header_Data (CDX_Identification , Schema_Version , Environment , Report_Type)>

<!ELEMENT Customer_Header_Data (Transaction)>

<!ELEMENT CDX_Identification (#PCDATA)>
<!-- "CDX_Identification" is the sender's CDX user ID.
     -->

<!ELEMENT Schema_Version (#PCDATA)>
<!-- The version of this schema is "2.1"
     -->

<!ELEMENT Environment (#PCDATA)>
<!-- Valid environment codes are:
     T:  Test
     P:  Production
     -->

<!ELEMENT Transaction (Transaction_Purpose , Transaction_Date , Transaction_Time)>

<!ELEMENT Report_Type (#PCDATA)>
<!-- The Report_Type code must be:
     UCM:  Unregulated Contaminant Monitoring
     -->
```

```
<!ELEMENT Transaction_Purpose (#PCDATA)>
<!--  Valid Transaction_Purpose values are
      o: Original submission
      r: Replacement submission
      -->

<!ELEMENT Transaction_Date (#PCDATA)>
<!ATTLIST Transaction_Date  e-dtype NMTOKEN  #FIXED 'date' >

<!ELEMENT Transaction_Time (#PCDATA)>
<!ATTLIST Transaction_Time  e-dtype NMTOKEN  #FIXED 'time' >

<!ELEMENT Lab_Id (#PCDATA)>
<!-- For Lab_Id use the federal laboratory number -->

<!ELEMENT Batch (Batch_Id , Extraction_Analysis_Date , Analytical_Method , Analyte+)>

<!ELEMENT PWS (PWS_Id , Facility+)>

<!ELEMENT Batch_Id (#PCDATA)>
<!-- For Batch_Id provide an extraction batch number. If the analytical method does not involve
      extraction in the process, provide the analysis batch number.
      -->

<!ELEMENT Extraction_Analysis_Date (#PCDATA)>
<!ATTLIST Extraction_Analysis_Date  e-dtype NMTOKEN  #FIXED 'date' >

<!ELEMENT Analytical_Method (#PCDATA)>
<!-- For Analytical_Method use the SDWIS code list for methods.
      -->

<!ELEMENT Analyte (Analyte_Code , Spiking_Concentration , Analytical_Precision , Analytical_Accuracy)>

<!ELEMENT Analyte_Code (#PCDATA)>

<!ELEMENT Spiking_Concentration (#PCDATA)>

<!ELEMENT Analytical_Precision (#PCDATA)>

<!ELEMENT Analytical_Accuracy (#PCDATA)>

<!ELEMENT PWS_Id (#PCDATA)>

<!ELEMENT Facility (Facility_Id , Facility_Sample_Point+)>

<!ELEMENT Facility_Id (#PCDATA)>

<!ELEMENT Facility_Sample_Point (Sample_Point_Id , Sample_Point_Sample+)>

<!ELEMENT Sample_Point_Id (#PCDATA)>

<!ELEMENT Sample_Point_Sample (Sample_Id , Sample_Collection_Date , Analysis_Type ,
      Lab_Sample_Comment? , Analysis+)>

<!ELEMENT Sample_Id (#PCDATA)>
```

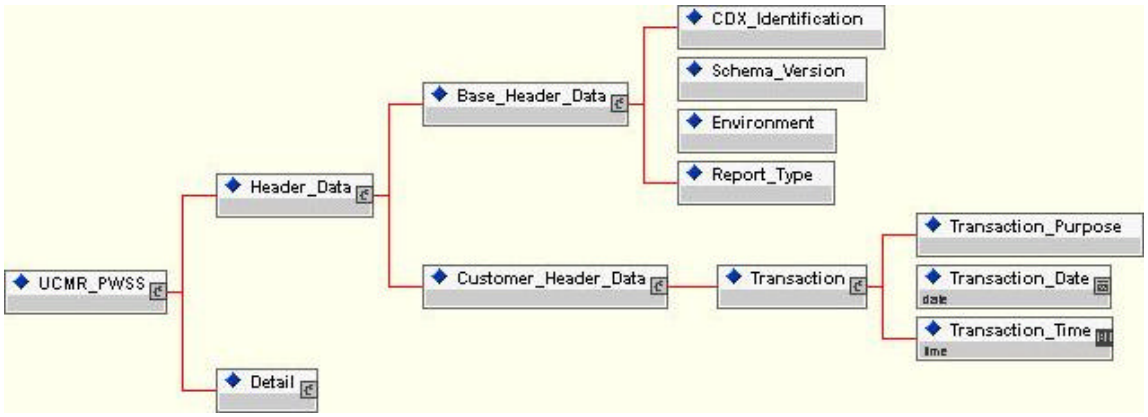
```
<!ELEMENT Sample_Collection_Date (#PCDATA)>
<!ATTLIST Sample_Collection_Date  e-dtype NMTOKEN  #FIXED 'date' >

<!ELEMENT Analysis_Type (#PCDATA)>
<!-- Valid Analysis_Type values are:
     rfs:  Raw field sample—untreated sample collected and submitted for analysis under this rule.
     rds:  Raw duplicate field sample—untreated field sample duplicate collected at the same time and place
          as the raw field sample and submitted for analysis under this rule.
     tfs:  Treated field sample—treated sample collected and submitted for analysis under this rule.
     tds:  Treated duplicate field sample—Treated duplicate field sample
          collected at the same time and place as the treated field sample and submitted for analysis under this rule.
     -->

<!ELEMENT Lab_Sample_Comment (#PCDATA)>

<!ELEMENT Analysis (Analyte_Code , Batch_Id , Analytical_Method, Analysis_Result , Analysis_Status?)>

<!ELEMENT Analysis_Result (Value? , Result_Sign , Presence?)>

<!ELEMENT Analysis_Status (Reviewer_Status , Lab_Result_Comment?)>

<!ELEMENT Value (#PCDATA)>
<!-- Values must be in the appropriate unit of measure for analyte as stated in the UCMR. Though the units of
     measure are not transmitted, the list of units of measure is:
     Degrees Celsius
     Micrograms per liter
     Nephelometric turbidity units
     pH units
     Micrograms per liter
     -->

<!ELEMENT Result_Sign (#PCDATA)>
<!-- Valid Result_Sign values are:
     lt: Less than
     eq: Equals
     -->

<!ELEMENT Presence (#PCDATA)>
<!-- Valid Presence values are:
     p: Present
     a: Absent
     -->

<!ELEMENT Reviewer_Status (#PCDATA)>
<!-- Valid Reviewer_Status values are:
     h: Hold result record
     a: Approve result
     -->

<!ELEMENT Lab_Result_Comment (#PCDATA)>
```

Appendix B
UCMR File Structure Tree Diagram

This appendix provides a graphical depiction of the UCMR file structure. It is not possible to provide a picture of the entire tree diagram at once due to size limitations. In order to present the tree diagram, we broke the file structure into three sections. Each section shows the detail for a specific portion of the file structure.

Reading from left to right and bottom to top, Figure B-1 shows the root element of UCMR_PWSS and its sub-elements, Header_Data and Detail. The Header_Data is expanded to reveal the detail of the Header_Data portion of the file structure. The Detail section of the file structure can be seen in Figures B-2 and B-3.

Figure B-1. UCMR_PWSS Tree Diagram with Header_Data



The Detail section of the file structure is, like the overall tree diagram, too large to depict as a single tree diagram. In order to best present the tree diagram, the Detail section is split into two sections. The first diagram illustrates the file structure four levels out with the details of PWS in Figure B-3.

Figure B-2. UCMR_PWSS Tree Diagram with Detail

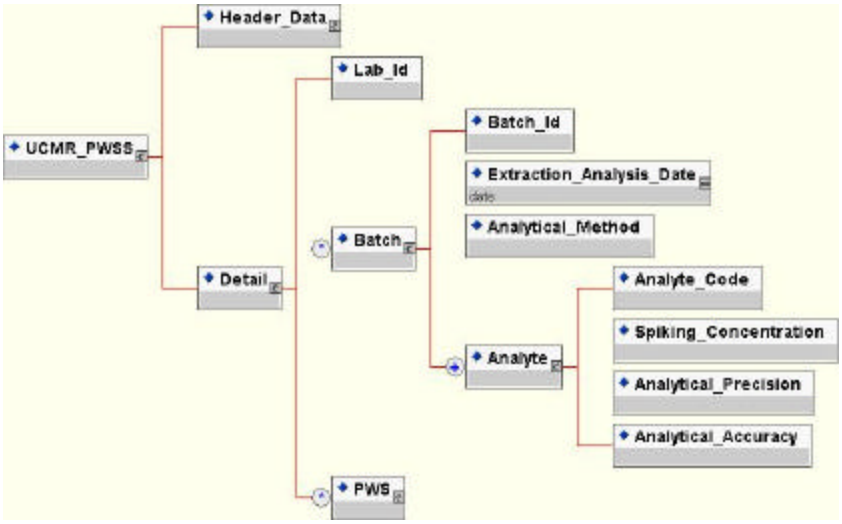
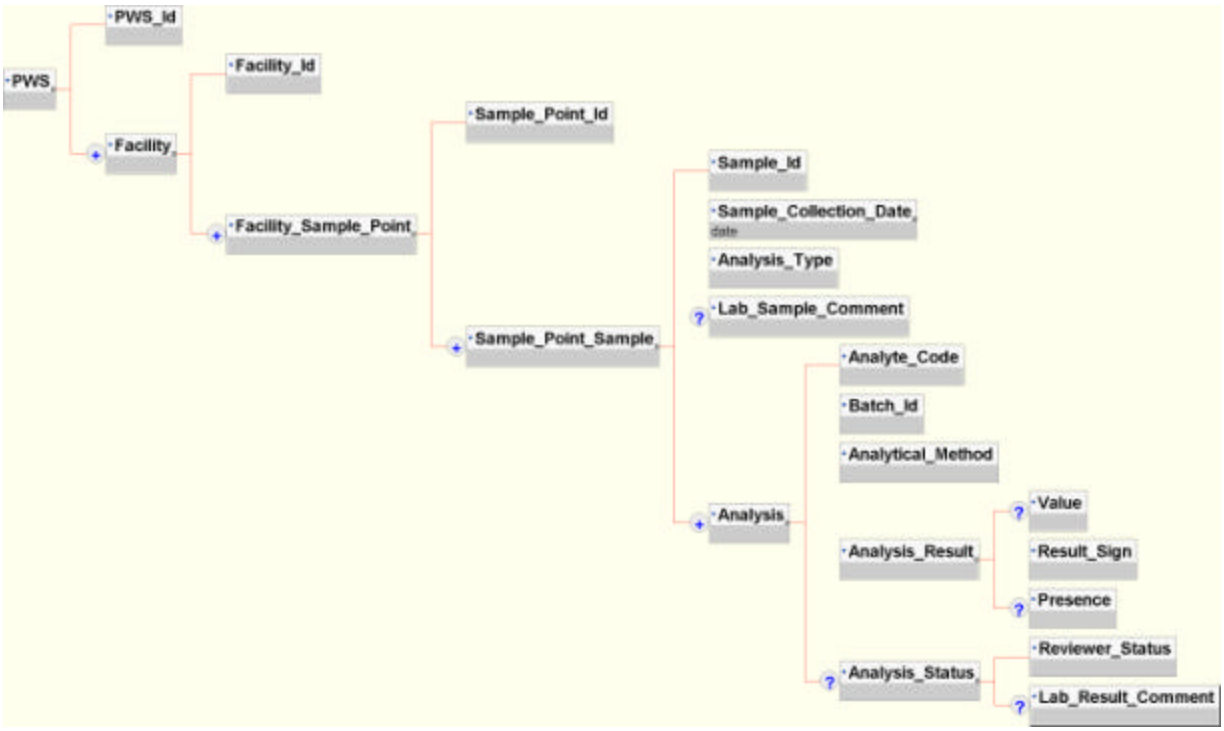


Figure B-3 shows the fully expanded PWS path exposing all of its children elements; whereas in Figure B-2, the full PWS path is not expanded due to size limitations of the diagram.

Figure B-3. PWS Path Tree Diagram



LOGICAL STRUCTURE

This section contains tables defining the UCMR DTD file content models. Each section is titled with the parent or container element and is followed by a table that lists the children or the contents of the container. The sub-element column defines the path. The description column defines the last element in the path in the sub-element column of the table. A more detailed definition with element attributes and associated code lists (e.g., Analyte Code List and Method Code List) is in Appendix C.

The UCMR_PWSS Element

The root element for the UCMR XML document is UCMR _PWSS, as shown in Table B-1. There are two elements below the root element. They are Header_Data and Detail. Table B-2 and Table B-3 describe the sub-elements respectively.

Table B-1. Top-Level Sub-Elements of UCMR_PWSS

Sub-element	Description
Header_Data	This sub-element occurs once at the beginning of a transaction file to provide transaction processing information. The sub-element is composed of standard EPA header elements and UCMR customer-specific data.
Detail	This sub-element contains the lab identification number, batch data-quality records, and PWS sampling results records. This sub-element occurs once for each transaction of UCMR data from a laboratory.

HEADER_DATA SUB-ELEMENT

This container sub-element contains secondary sub-elements that hold all the data for the transaction type and CDX routing.

Each document has one instance of Header_Data (see Table B-2).

Table B-2. Header_Data Sub-Elements

Sub-element	Description
Base_Header_Data	This sub-element defines the type of transaction and version. The sub-element contains information that informs the parser how to process the data properly.
Base_Header_Data/CDX_Identification	This sub-element contains the CDX user identification for the submitter of the file. The user ID will assist CDX if notifications are necessary for processing files.
Base_Header_Data/Schema_Version	This sub-element defines which version of the DTD schema was the basis for generating the XML transaction.
Base_Header_Data/Environment	This sub-element can define a transaction as test or production data. The use of "test" definition is not accepted by the UCMR file upload process. A file-testing capability is being developed for CDX.
Base_Header_Data/Report_Type	This sub-element defines the report as the UCMR transaction.
Customer_Header_Data	This sub-element enables the laboratory to generate transaction-specific information.
Customer_Header_Data/Transaction	This sub-element defines the entire transaction by the version, purpose, date, and time.
Customer_Header_Data/Transaction/ Transaction_Purpose	This sub-element defines the purpose of the transaction.
Customer_Header_Data/Transaction/ Transaction_Date	This sub-element identifies the date the transaction was prepared.
Customer_Header_Data/Transaction/ Transaction_Time	This sub-element identifies the time the transaction was prepared.

Detail Sub-Element

This container sub-element contains secondary sub-elements that hold all the laboratory, batch, and PWS data for recording analytical sample results.

Table B-3. Detail Sub-Elements

Sub-element	Description
Lab_Id	This sub-element identifies the laboratory that did the analysis submitted in the transaction.
Batch	This sub-element contains the batch ID, date, analytical method, and analytes. This sub-element occurs one or more times to provide a series of batches reported by the same lab.
Batch/Batch_Id	This sub-element identifies the extraction batch assigned by the lab (if extraction was not part of the method, use the analysis batch identification).
Batch/Extraction_Analysis_Date	This sub-element identifies the date that the processing of samples contained in each extraction batch (or an analysis batch if there is no extraction) was completed.
Batch/Analytical_Method	This sub-element defines the analytical method used to analyze the samples in the batch.
Batch/Analyte	This sub-element contains the analyte code and the data quality that are part of a batch. This element occurs one or more times for a batch.
Batch/Analyte/Analyte_Code	This sub-element identifies the applicable analyte for a batch quality control (QC) record. Analytes are identified by their SDWIS code value.
Batch/Analyte/Spiking_Concentration	This sub-element defines the concentration added to a sample to be analyzed for calculating analytical precision and accuracy if the value reported is assumed to be in the same unit of measure as that reported for analytical results.
Batch/Analyte/Analytical_Precision	This sub-element identifies the analytical precision used for measuring the analyte for the batch according to the observed variability of results for duplicate spike samples. Values are percentages (e.g., 95.5) or if not measured, “missing.” The decimal point is not included in the element size limit.

Table B-3. Detail Sub-Elements (Continued)

Sub-element	Description
Batch/Analyte/Analytical_Accuracy	This sub-element identifies the analytical accuracy used for measuring the analyte for the batch according to the percent recovered from spiked samples. Values are percentages (e.g., 95.5). The decimal point is not included in the element size limit.
PWS	This sub-element contains the facility and PWS_Id sub-elements. This sub-element occurs one or more times to provide a series of sample results.
PWS/PWS_Id	This sub-element identifies a PWS.
PWS/Facility	This sub-element contains the facility identification number and sample points. This sub-element occurs one or more times for each PWS.
PWS/Facility/Facility_Id	This sub-element identifies a facility within the PWS using an identification number that is unique within the PWS.
PWS/Facility/Facility_Sample_Point	This sub-element contains the sampling point identification. The identification must be either static or traceable to previous numbers and type identifications throughout the period of unregulated contaminant monitoring. This sub-element occurs one or more times for each facility.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Id	This sub-element defines the identification number, established by the state or at the state's discretion; the sampling point that is unique to the PWS for an intake for each source of water: a treatment plant, a distribution system, or other location of water treatment or delivery; and describes where samples may be taken.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample	This sub-element defines the element's number, type, sample comment, and analysis. This sub-element occurs one or more times for each sample point.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Sample_Id	This sub-element identifies the identification number of the sample.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/ Sample_Collection_Date	This sub-element identifies the date the sample is collected.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/ Analysis_Type	This sub-element identifies the type of sample collected.

Table B-3. Detail Sub-Elements (Continued)

Sub-element	Description
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/ Lab_Sample_Comment	This sub-element lists comments about the sample as it was collected. This sub-element is optional.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis	This sub-element contains the analytical result for a sample's single analyte and laboratory result status. The analysis data are associated with a batch reported in the same transaction or already recorded in SDWARS. This sub-element occurs one or more times for each sample.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis/ Analyte_Code	This sub-element identifies the applicable analyte for a sample analysis record. Analytes are identified by their SDWIS code value.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis/ Batch_Id	This sub-element identifies the batch (assigned by the lab) that the sample was a part of in producing the analytical result(s).
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis/ Analytical_Method	This sub-element identifies the analytical method that governed the analysis of the sample as part of the batch.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis/ Analysis_Result	This sub-element contains the analytical result for the analyte identified in the same analysis group.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis/ Analysis_Result/Value	This sub-element gives the result as a measured value based on the EPA defined unit of measure for an analyte.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis/ Analysis_Result/Result_Sign	This sub-element defines an analytical result as less than the EPA defined MRL or as being equal to the measured value.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis/ Analysis_Result/Presence	This sub-element defines whether or not an analyte is present. *(This element is currently not used, but is reserved for future use.)*
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis/ Analysis_Status	This sub-element contains the laboratory's status and comments for an analyte's analysis.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis/ Analysis_Status/Reviewer_Status	This sub-element gives the laboratory's status for an analyte's analysis.
PWS/Facility/Facility_Sample_Point/ Sample_Point_Sample/Analysis/ Analysis_Status/ Lab_Result_Comment	This sub-element lists the laboratory's comments for an analyte's analysis.

Appendix C
UCMR Data Dictionary: XML DTD, Version 2.1

CONTENTS

INTRODUCTION

This appendix contains the data dictionary for the UCMR DTD. This appendix also contains introductory materials about how to read the data dictionary. Only the data content elements are defined here. Group defining elements have a description in the logical structure of Appendix B.

HOW TO READ THIS DATA DICTIONARY

The data dictionary contains the element name, definition of the element, characteristics for the element, its relationship in the hierarchy, and an example of its use. Below you will find further explanation for each segment of the definition in the dictionary structure.

Element Name

Definition:

This section provides a textual definition of the element and how it is used in the DTD.

Characteristics:

The characteristics of an element provide information on the use of the element. Characteristics include data type (alpha-numeric or numeric), size or string length, requirement of optional or mandatory, and occurrences.

Parent Hierarchy:

The parent hierarchy section gives the relationship of the element to the root element.

Tagged Example:

The tagged example shows the element tagged with data as an XML document would appear.

A ELEMENTS

Analysis_Type

Definition:

The Analysis_Type element identifies the type of sample collected.

Characteristics:

Data Type: AN

Size: 3

Requirement: Mandatory (for <Sample_Point_Sample> loops)

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS

Detail

PWS

Facility

Facility_Sample_Point

Sample_Point_Sample

Tagged Example:

<Analysis_Type>rfs</Analysis_Type>

Code List:

Sample Type	Definition of Sample Type
rfs	Raw field sample—untreated sample collected and submitted for analysis.
rds	Raw duplicate field sample—untreated field sample duplicate collected at the same time and from the same place as the raw field sample and submitted for analysis. (only small systems)
tfs	Treated field sample—treated sample collected and submitted for analysis. (most UCMR systems)
tds	Treated duplicate field sample—treated field sample duplicate collected at the same time and from the same place as the treated field sample and submitted for analysis. (only small systems)

Analyte_Code

Definition:

Identifies the applicable analyte for a batch quality control (QC) record or an analysis result record. Analytes are identified by their SDWIS code value.

Characteristics:

Data Type: N

Size: 4

Requirement: Mandatory (for <Analyte> loops and <Analysis> loops)

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS
    Detail
        Batch
            Analyte
~and~
UCMR_PWSS
    Detail
        PWS
            Facility
                Facility_Sample_Point
                    Sample_Point_Sample
                        Analysis

Tagged Example:

<Analyte_Code>2009</Analyte_Code>

Code List:

SDWIS Code	UCMR Analyte Name
List 1–Assessment Monitoring Analytes	
2009	4,4'-DDE
1039	Perchlorate
2108	DCPA mono/di-acid degradates [a]
2027	Acetochlor
2052	EPTC
2251	MTBE
2254	Nitrobenzene
2266	2,6-dinitrotoluene
2270	2,4-dinitrotoluene
2272	Terbacil
2626	Molinate
List 2–Screening Survey Analytes	
3201	Aeromonas
2029	Prometon
2056	Diazinon
2102	Disulfoton
2103	Diuron
2104	Fonofos
2233	2-methyl-phenol
2254	Low-level Nitrobenzene
2268	1,2-diphenylhydrazine
2283	Linuron
2328	2,4-dinitrophenol
2332	2,4,6-trichlorophenol
2334	2,4-dichlorophenol
2545	Terbufos

[a]DCPA mono-acid degradate and di-acid degradate are not reported to SDWARS individually.

Analytical_Accuracy

Definition:

Identifies the Analytical_Accuracy for measuring the analyte for the batch according to the percent recovered from spiked samples. Values are to be percentages (e.g. 95.5). The use of a decimal point is *not* included in the element size limitation.

Characteristics:

Data Type:	N
Size:	1-5
Requirement:	Mandatory (for <Data_Quality> loops)
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - Batch
    - Analyte
      - Data_Quality

Tagged Example:

<Analytical_Accuracy>95.5</Analytical_Accuracy>

Analytical_Method

Definition:

The element Analytical_Method defines the analytical method used to analyze the samples in the batch.

Characteristics:

Data Type: AN

Size: 6-15

Requirement: Mandatory (for <Batch> loops)

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS

Detail

Batch

~and~

UCMR_PWSS

Detail

PWS

Facility

Facility_Sample_Point

Sample_Point_Sample

Analysis

Tagged Example:

<Analytical_Method>AOAC 990.06</Analytical_Method>

Code List:

Analytical Method	Abbreviated Method Name (40 character maximum)
AOAC 990.06	Organochlorine pesticides in water
AOAC 991.07	Nitrogen/phosphorus pesticides
AOAC 992.32	Chlorinated acid pesticides
ASTM D5317	Chlorinated organic acids, GC/ECD

Analytical Method	Abbreviated Method Name (40 character maximum)
ASTM D5475	Pesticides, nitrogen/phosphorus, LLE, GC, NPD
ASTM D5790	VOC, GC/MS, P&T, capcolumn
ASTM D5812	Pesticides, chlorinated, GC
EPA 1605	Membrane filter, aeromonas
EPA 314.0	Perchlorate, ion chromatography
EPA 502.2	VOC, GC, PID/ECD, P&T
EPA 507	Pesticides, nitrogen/phosphorus, LLE, GC, NPD
EPA 508	Pesticides, chlorinated, LLE, GC/ECD
EPA 508.1	Pesticides, chlorinated, SPE, GC/ECD
EPA 515.1	Herbicides, acids, LLE, GC/ECD
EPA 515.2	Herbicides, acids, SPE, GC/ECD
EPA 515.3	Herbicides, acids, LLE, GC/ECD
EPA 515.4	Herbicides, acids, micro LLE, GC/ECD
EPA 524.2	VOC, GC/MS, P&T, capcolumn
EPA 525.2	Organics, SPE, GC/MS
EPA 526	Organics, SPE, GC/MS
EPA 528	Phenols, SPE, GC/MS
EPA 532	Phenylurea, SPE, HPLC/UV
SM 6200 B	VOC, GC/MS, P&T, capcolumn
SM 6200 C	VOC, GC/MS, P&T, capcolumn
SM 6210 D	VOC, GC/MS, P&T, capcolumn

Analytical_Precision

Definition:

Identifies the Analytical_Precision for measuring the analyte for the batch according to the observed variability of results for duplicate spike samples. Values are to be percentages (e.g., 8.25) or "missing" if not measured. The use of a decimal point is *not* included in the element size limitation.

Characteristics:

Data Type:	N
Size:	1-5
Requirement:	Mandatory (for <Data_Quality> loops)
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - Batch
    - Analyte
      - Data_Quality

Tagged Example:

Example 1:

<Analytical_Precision>7.5</Analytical_Precision>

Example 2:

<Analytical_Precision>missing</Analytical_Precision>

B ELEMENTS

Batch_Id

Definition:

The element Batch_Id identifies the extraction batch assigned by the lab (if extraction was not part of the method, use the analysis batch identification).

Characteristics:

Data Type:	AN
Size:	1-15
Requirement:	Mandatory (for <Batch> loops)
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - Batch

~and~

UCMR_PWSS
- Detail
  - PWS
    - Facility
      - Facility_Sample_Point
        - Sample_Point_Sample
          - Analysis

Tagged Example:

<Batch_Id>2001032811</Batch_Id>

C ELEMENTS

CDX_Identification

Definition:

The CDX_Identification element contains the CDX user identification for the submitter of the file. The user ID provided will assist CDX if notifications are necessary for processing files.

Characteristics:

Data Type:	AN
Size:	8-30
Requirement:	Mandatory
Occurrences:	Once

Parent Hierarchy:

PWSS_UCMR

Header_Data

Base_Header_Data

Tagged Example:

<CDX_Identification>DJohnson</CDX_Identification>

E ELEMENTS

Environment

Definition:

The Environment element can define a transaction as test or production data. The use of "test" is not accepted by the UCMR file upload process.

Characteristics:

Data Type: AN

Size: 1

Requirement: Mandatory

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS

Header_Data

Base_Header_Data

Tagged Example:

<Environment>T</Environment>

Code List:

Environment Code Value	Environment
P	Production
T	Test

Extraction_Analysis_Date

Definition:

The element Extraction_Analysis_Date defines the date that the processing of samples contained in each extraction batch (or an analysis batch if there is no extraction) was completed.

Attribute:

<!ATTLIST Extraction_Analysis_Date e-dtype NMTOKEN #FIXED ‘date’ >

Characteristics:

Data Type:	N
Size:	8
Requirement:	Mandatory (for <Batch> loops)
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS

Detail

Batch

Tagged Example:

<Extraction_Analysis_Date>20010701</Extraction_Analysis_Date>

F ELEMENTS

Facility_Id

Definition:

The Facility_Id element defines the identification number of the facility.

Characteristics:

Data Type:	AN
Size:	1-6
Requirement:	Mandatory (for <Facility> loops)
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - PWS
    - Facility

Tagged Example:

<Facility_Id>00010</Facility_Id>

L ELEMENTS

Lab_Id

Definition:

The Lab_Id element identifies the laboratory that did the analysis submitted in the transaction.

Characteristics:

Data Type:	AN
Size:	7
Requirement:	Mandatory
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS

Detail

Tagged Example:

<Lab_Id>VA12345 </Lab_Id>

Lab_Result_Comment

Definition:

The Lab_Result_Comment defines comments about the sample and its analytical results.

Characteristics:

Data Type:	AN
Size:	1-250
Requirement:	Optional
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - PWS
    - Facility
      - Facility_Sample_Point
        - Sample_Point_Sample
          - Analysis
            - Analysis_Status

Tagged Example:

<Lab_Result_Comment>sample rerun due to machine failure</Lab_Result_Comment>

Lab_Sample_Comment

Definition:

The Lab_Sample_Comment element lists comments about the sample as it was collected. This element is optional.

Characteristics:

Data Type:	AN
Size:	1-250
Requirement:	Optional
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - PWS
    - Facility
      - Facility_Sample_Point
        - Sample_Point_Sample

Tagged Example:

<Lab_Sample_Comment>heavy rains </Lab_Sample_Comment>

P ELEMENTS

Presence *(reserved for future use)*

Definition:

The Presence element determines if an analyte is present or absent in the sample.

Characteristics:

Data Type: AN

Size: 1

Requirement: Optional

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS
Detail
PWS
Facility
Facility_Sample_Point
Sample_Point_Sample
Analysis
Analysis_Result

Tagged Example:

<Presence>a</Presence>

Code List:

Presence Indicator	Presence Indicator Definition
p	present
a	absent

PWS_Id

.Definition:

The PWS_Id element is the code for identifying each PWS.

Characteristics:

Data Type: AN

Size: 9

Requirement: Mandatory (for <PWS> loops)

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS

Detail

PWS

Tagged Example:

<PWS_Id>PA1234567 </PWS_Id>

R ELEMENTS

Report_Type

Definition:

The Report_Type element defines this report as the UCMR transaction.

Attribute:

<!ATTLIST Report_Type report_type CDATA #FIXED “ucmr”>

Characteristics:

Data Type: AN

Size: 4

Requirement: Mandatory

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS

Header_Data

Base_Header_Data

Tagged Example:

<Report_Type>ucmr</Report_Type>

Code List:

Report Type	Report Type Definition
ucmr	Unregulated-contaminant monitoring report

Result_Sign

Definition:

The element Result_Sign defines if the lowest concentration of an analyte is at the minimum reporting level, indicated by a value of less than MRL or equal to reported value (lt or eq). This element is required for all analytes except those for water quality.

Characteristics:

Data Type: AN

Size: 2

Requirement: Mandatory (for <Analysis_Result> loops)

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - PWS
    - Facility
      - Facility_Sample_Point
        - Sample_Point_Sample
          - Analysis
            - Analysis_Result

Tagged Example:

<Result_Sign>eq</Result_Sign>

Code List:

Result Sign	Result Sign Definition
lt	Less than MRL
eq	Equals (a value = or > MRL)

Reviewer_Status

Definition:

The Reviewer_Status defines the laboratory's status for the result. The value of this element is one of the following: h (hold) or a (approve).

Characteristics:

Data Type: AN

Size: 1

Requirement: Mandatory (for <Analysis_Status> loops)

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS
Detail
PWS
Facility
Facility_Sample_Point
Sample_Point_Sample
Analysis
Analysis_Status

Tagged Example:

<Reviewer_Status>a</Reviewer_Status>

Code List:

Status Options	Status Definition
a	Approve
h	Hold

S ELEMENTS

Sample_Collection_Date

Definition:

The Sample_Collection_Date element defines the date the sample is collected.

Attribute:

<!ATTLIST Sample_Collection_Date e-dtype NMTOKEN #FIXED ‘date’ >

Characteristics:

Data Type: N

Size: 8

Requirement: Mandatory (for <Sample_Point_Sample> loops)

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - PWS
    - Facility
      - Facility_Sample_Point
        - Sample_Point_Sample

Tagged Example:

<Sample_Collection_Date>20010701</Sample_Collection_Date>

Sample_Id

Definition:

The Sample_Id element defines the identification number of the sample.

Characteristics:

Data Type:	AN
Size:	1-15
Requirement:	Mandatory (for <Sample_Point_Sample> loops)
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
Detail
PWS
Facility
Facility_Sample_Point
Sample_Point_Sample

Tagged Example:

<Sample_Id>200105858 </Sample_Id>

Sample_Point_Id

Definition:

The Sample_Point_Id element defines the identification number, established by the state or at the state's discretion, the PWS, that is unique to the PWS for an intake for each source of water: a treatment plant, a distribution system, or other location of water treatment or delivery and describes where samples may be taken.

Characteristics:

Data Type:	AN
Size:	1-20
Requirement:	Mandatory (for <Facility_Sample_Point> loops)
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - PWS
    - Facility
      - Facility_Sample_Point

Tagged Example:

<Sample_Point_Id>0002E </Sample_Point_Id>

Schema_Version

Definition:

The Schema_Version element defines which DTD schema version was the basis for generating the XML transaction.

Characteristics:

Data Type: AN

Size: 1-4

Requirement: Mandatory

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS

Header_Data

Base_Header_Data

Tagged Example:

<Schema_Version>2.1</Schema_Version>

Spiking_Concentration

Definition:

The Spiking_Concentration element defines the concentration added to a sample to be analyzed for calculating analytical precision and accuracy where the value reported is assumed to be in the same unit of measure that is reported for analytical results.

Characteristics:

Data Type:	N
Size:	1-5
Requirement:	Mandatory
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - Batch
    - Analyte
      - Data_Quality

Tagged Example:

```
<Spiking_Concentration>20</Spiking_Concentration>
```

T ELEMENTS

Transaction_Date

Definition:

The Transaction_Date defines the date the transaction was prepared.

Attribute:

<!ATTLIST Transaction_Date e-dtype NMTOKEN #FIXED ‘date’ >

Characteristics:

Data Type: N

Size: 8

Requirement: Mandatory

Occurrences: Once

Parents:

UCMR_PWSS

Header_Data

Customer_Header_Data

Transaction

Tagged Example:

<Transaction_Date>20010701</Transaction_Date>

Transaction_Purpose

Definition:

The Transaction_Purpose element defines the purpose of the transaction.

Characteristics:

Data Type: AN

Size: 1

Requirement: Mandatory

Occurrences: Once

Parent Hierarchy:

UCMR_PWSS
Header_Data
Customer_Header_data
Transaction

Tagged Example:

<Transaction_Purpose>o</Transaction_Purpose>

Code List:

Transaction Purpose	Transaction Purpose Definition
o	Original (additions only)
r	Replace (updates and additions)

Transaction_Time

Definition:

The Transaction_Time element defines the time the transaction was prepared.

Attribute:

<!ATTLIST Transaction_Time e-dtype NMTOKEN #FIXED ‘time’>

Characteristics:

Data Type:	N
Size:	6
Requirement:	Mandatory
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
    Header_Data
        Customer_Header_Data
            Transaction

Tagged Example:

<Transaction_Time>183000</Transaction_Time>

V ELEMENTS

Value

Definition:

The Value element defines the numeric value of the analysis for chemical and microbiological results.

Characteristics:

Data Type:	N
Size:	1-15
Requirement:	Mandatory (for <Analysis_Result> loops)
Occurrences:	Once

Parent Hierarchy:

UCMR_PWSS
- Detail
  - PWS
    - Facility
      - Facility_Sample_Point
        - Sample_Point_Sample
          - Analysis
            - Analysis_Result

Tagged Example:

```
<Value>3</Value>
```

Appendix E
Abbreviations

ACES	Access Certificates for Electronic Services
B2B	business-to-business
CA	certifying agent
CDX	Central Data Exchange
CRK	customer retrieval key
CSI	Common Sense Initiative
DTD	document-type definition
EC	electronic commerce
EDI	Electronic Data Exchange
EPA	U.S. Environmental Protection Agency
FCC	Federal Communications Commission
GPEA	Government Paperwork Elimination Act
GPRA	Government Performance and Results Act
GSA	General Services Administration
I-3	Information Integration Initiative
ID	identification
IG	implementation guide
IT	information technology
LIMS	Laboratory Information Management Guide
NCOD	National Contaminant Occurance Database
OEI	Office of Environmental Information
OGWDW	Office of Ground Water and Drinking Water
OIC	Office of Information Collection
OLAP	online analytical processing
PE	performance evaluation
PIN	personal identification number
PC	personal computer
PWS	public water systems

REI	Reinventing Environmental Information
SDWARS	Safe Drinking Water Accession and Review System
SSL	secure socket layer
UCMR	Unregulated Contaminant Monitoring Rule
URL	uniform resource locator
XML	Extensible Markup Language
Y2K	Year 2000

Office of Water (4601)
Washington, DC 20460
EPA 816-R-01-022C
www.epa.gov/safewater
December 2001