UIC Data Quality Tools & Process

Trang Le
OGWDW, U.S. EPA
2010 GWPC Annual Forum
Pittsburgh, PA -- September 26-29, 2010
Objective:
To explain how to use the UIC database tools to facilitate data quality improvement

Agenda:
1. EPA’s Nat’l UIC Data System Framework
2. UIC Data Validation Process --
   -- Ensuring Appropriate Data Mapping
   -- Validating Incoming Data
   -- Verifying Received data
National UIC Database System Framework

OGWDW Management:
- 7520 summary info
- UIC PAM Measures
- UIC Inventory Reporting Sys
- SWP/UIC Integration
- Class I Special Needs
- Class V Invent. Initiative
- Congr/GAO/OMB Questions

ICIS Reports (ICDS)

70+ Regional and state databases

Data Source’s site

XML Doc

Schematron

UIC Operational Database (Pre-Prod Database)

70+ Regional and state databases

XML Doc

State/DI XML Doc

Schematron

Parse and Load XML Doc

Data ready for reports

XML Conversion Tool

Staging Access Database

(1)

CDX (XML Doc Repository)

Standard reports and ad-hoc queries

UIC Reporting Database (Prod Database)

XML Conversion Tool

Backend Reports:
1. Backend invalid data
2. Changes from last data submission

3. Backend invalid data
4. Changes from last data submission
UIC Data Validation Process

Validate input data

Ensure proper data mapping

Verify correctness of received data
Ensuring Appropriate Data Mapping

Logical Mapping form:

- Identify matched data elements
- Recognize default value uses
- Document data mapping for evaluation & references
Validating Incoming Data

Validation Software:

- Execute UIC business rules
- Check if incoming data is complete, accurate, consistent
- Flag out data errors for correction

Validation Reports
Verifying Received Data with States

- **Adhoc Query tool:**
  - Verify quality of submitted data
  - Enable to drill down supporting data
  - Provide UIC draft reports: 7520s, PAMs, Inventory

- **Collaborate with States to ensure appropriate data use**
Questions?