

IMPROVING PERFORMANCE DATA QUALITY A LIFECYCLE APPROACH

Office of the Chief Financial Officer
and the
Office of Environmental Information

What are performance data?

- Data collected, analyzed, or reported by EPA to characterize progress in achieving specific performance targets:
 - Environmental
 - Human health
 - Administrative
 - Quantitative and non-quantitative
 - Short and long-term
- Support performance measures/goals
X (amount) of Y (units) by Z (time/period of performance)
- Varied primary and secondary data sources, including federal, state, and program-specific data and systems
 - More than 50 national databases alone (see examples)
- Reported into EPA's Performance Management System (Measures Central: BAS, PERS, ACS)

Performance Data Sources and Systems: Water

- SDWIS
- DW-NIMS
- UIC database
- NHANES
- WBDO
- PRAWN
- STORET
- WATERS
- GLENDA
- Surf Your Watershed
- ATTAINS
- PCS
- WATA
- GRTS
- ICES-NPDES
- CW-NIMS
- CBR
- STARS
- USACoE/RAMS
- EMAP/NCA

How are performance data used?

- Public reporting and accountability (PAR)
- Agency priority setting (Strategic Plan)
- Annual planning and budgeting and justification (Annual Plan)
- Program management and evaluation (PART, EPAStat, QMR)
- Regional and Program annual commitments (NPM Guidance/ACS)
- Organizational assessments (SES)
- Individual performance standards (PARS)

Data of known quality and reliability are critical for credible programmatic and Agency-wide management and policy decisions, and for public accountability

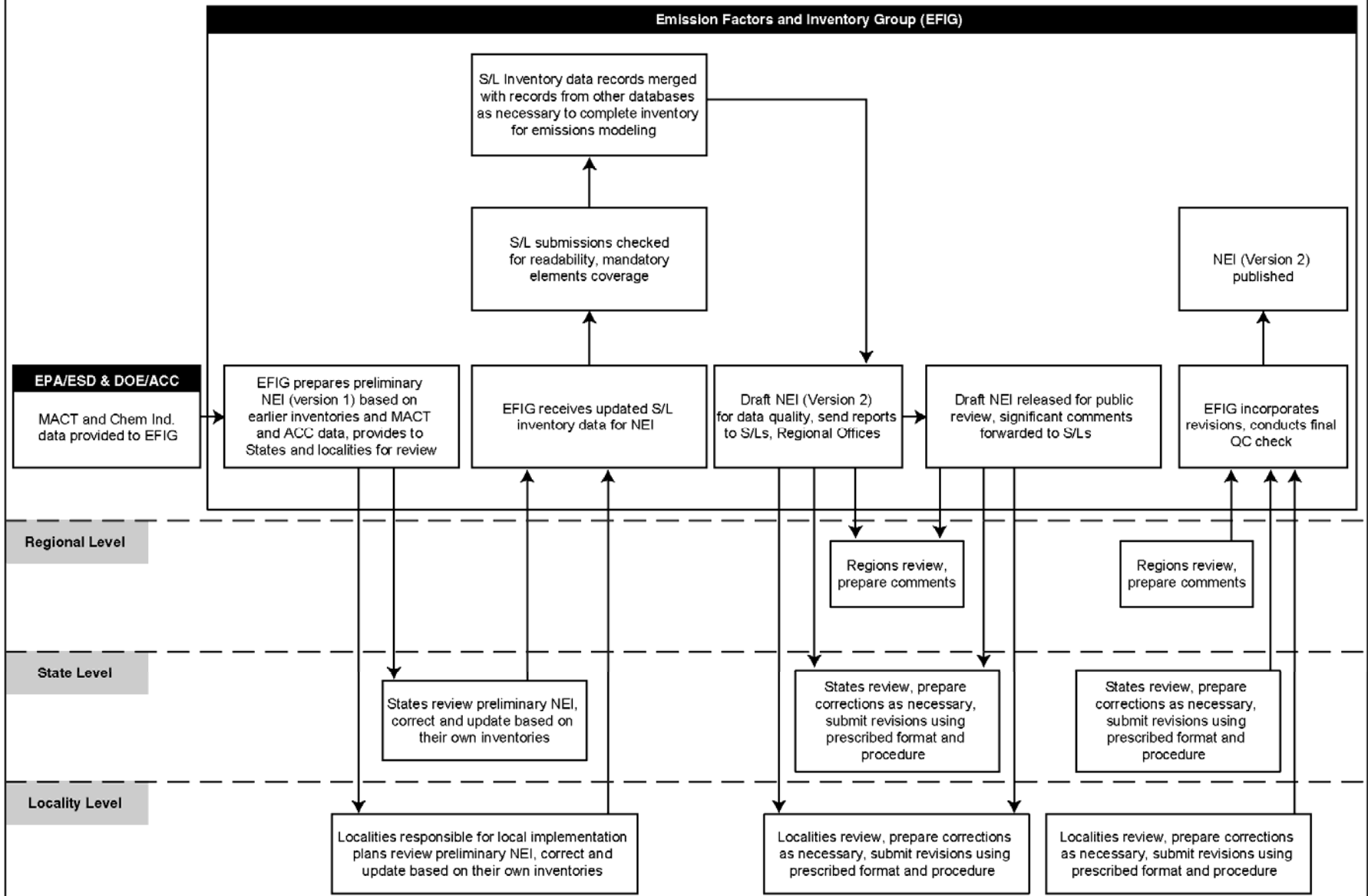
Challenge of performance data quality?

- Complexity of data sources, systems, analyses, users and uses, and presentations
- Potential bias in interpretation of result by competing interests: program, policy, management, and accountability
- Inconsistent application of quality system
- Unclear roles and responsibilities

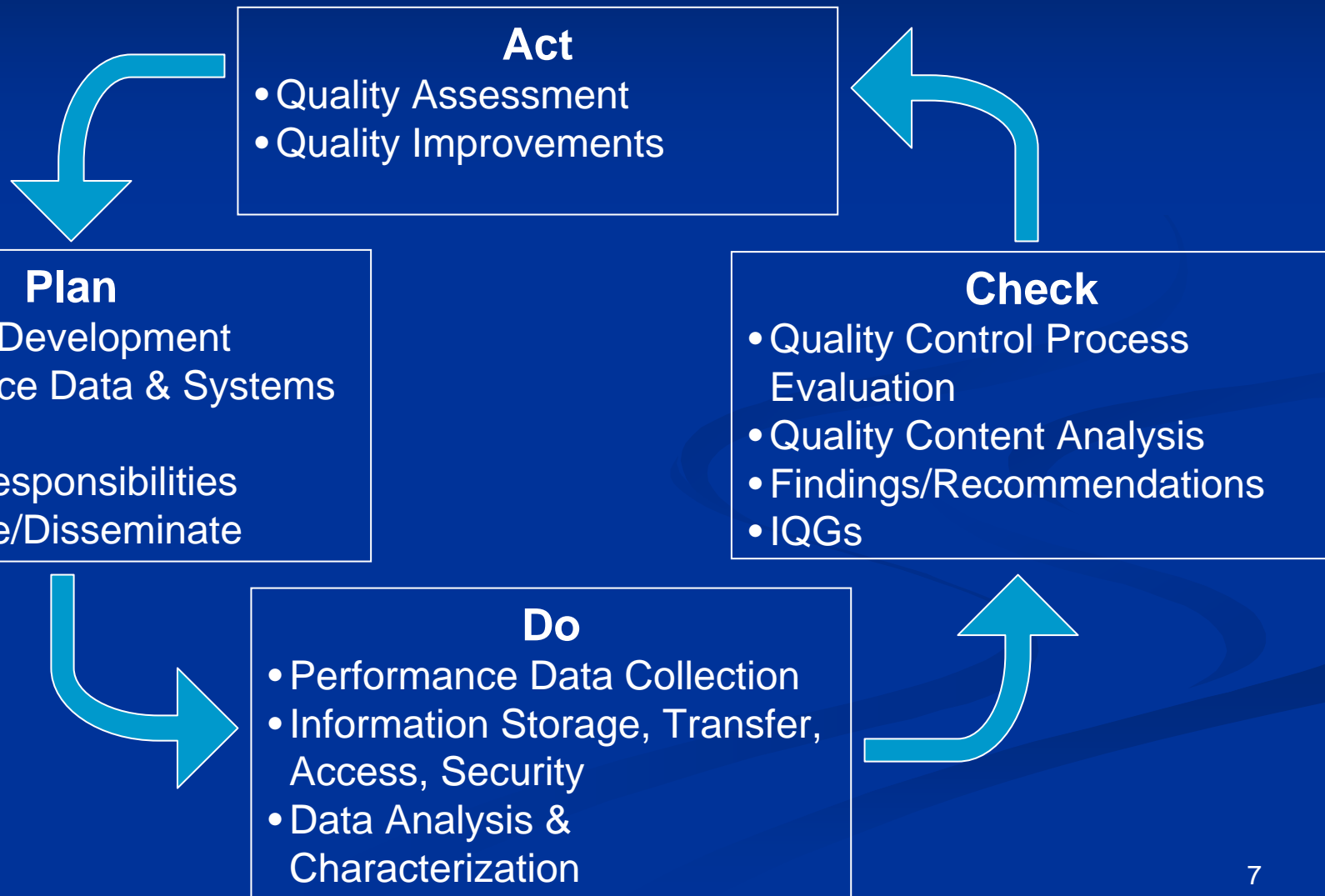
National Emissions Inventory (NEI) – Conceptual map of data flows and sources, responsible organizations, and relevant quality regimes

Headquarters Level

Source: The 1999 National Emission Inventory Preparation Plan (http://www.epa.gov/ttn/chief/net/nei_plan_feb2001.pdf)



Performance Data/Measure Quality Life Cycle



Plan

- Measures Development
 - Program performance questions
 - Performance metrics
 - Data sources
 - Acceptance criteria/data quality objectives
- Performance Data & Systems QA Plan
 - Sampling and laboratory analysis plans/protocols/SOPs
 - Reporting and record keeping
 - Database/computer system design, operation, security
 - Data input and export protocols/verification
 - Mathematical/statistical analysis and presentation
 - Acceptance criteria/process for secondary data
- Roles & Responsibilities
 - QA Plan development and oversight
 - Quality reviews/assessments/audits/approvals
 - Training

Who?

Program and Systems Staff/Managers/QA Managers

Do

- Performance Data Collection
 - Field sampling, monitoring data
 - Laboratory/field analyses
 - Secondary data collections
 - Model input data
 - Documentation (field and laboratory logs, system updates/maintenance logs)
- Information Storage, Transfer, Access, Security
 - Databases/computer systems
 - Record keeping
 - Presentations
 - Information products (reports, internet)
- Data Analysis/Characterization/Performance Result
 - Aggregation/tabulation
 - Statistical analyses
 - Modeling

Who?

Program and Support Staff/Management

Check

- Quality Control Process Evaluation
 - Adherence to/problems with QA plans
 - Data, database, and computer systems integrity
- Quality Content Analysis
 - Validity of *a priori* assumptions about data/measure
 - Statistical characterization of variability and/or uncertainty
- Information Quality Guidelines
 - Pre-dissemination review
 - OCFO's IQG Compliance Form
- Findings/Recommendations
 - Implications for interpreting performance and trends
 - Options to improve performance data quality
 - Data Quality Section of PAR or similar documents

Who?

IM Officer/QA Manager/IQG Officer/External Reviewers

Act

- Opportunities Assessment
 - Consideration of external reviews
 - Options to address quality issues
 - Costs and benefits of improvements options
 - Recommendations for specific improvements

- Examples
 - QA plans, protocols, SOPs
 - Performance measures
 - Data sources/collection, analyses, systems
 - Dissemination/Training
 - Performance data presentations/products

Who?

Program and Systems Staff and Management/QA Officers

Goals for Lifecycle Approach to QA

- Minimize errors by defining quality principles throughout data lifecycle, roles and responsibilities
- Better alignment of measure and program goal
- Better interpretation of results using data of known quality
- Improve confidence in data for program management and public accountability

2 Sediment Case Studies

Great Lakes National Program Office (multistate + Canada)

- Annual – cubic yards of contaminated sediments remediated by anyone
- GLNPO Excel data base of questionnaire results

Superfund (HQ/Region 10)

- Puget Sound prioritized Superfund sediment sites – 2011 target of 200 acres remediated
- CERCLIS modified for this performance measure

Plan

- Program & partners choose measure and data sources, address responsibility for data collection, manipulation, supporting database, reporting and records
- Systematic process for data & information suitability:
 - Performance criteria are set for performance measure data and database
 - Acceptance criteria are set for existing data
- Oversight & review roles determined

“Plan” IM/IT Issues for Sediment Measures

- Completeness – specify response/data entry for x percentage of remediated sites
- Timeliness – “freeze” data input for reporting period
- Comparability – data definitions for diverse reporters & remediation types
- Consistency and comparability (year-to-year) – set & track these indicators
- Consistency – objective data entry decisions with SOPs & training
- Reproducibility – oversight of data base procedures

Do

- Performance data are collected
- Database is populated (using definitions)
- Error checks done, data reported, stored
- Data manipulations (if any) are made
- Measurement is reported with meta data
- QA oversight (if any) is documented

“Do” IT/IM Issues for Sediments

- Completeness, comparability and consistency – data source and entry procedures known
- Reproducibility and integrity – consistent retrieval, manipulation and reporting
- Responsibility to reject or follow up on data problems
- Oversight to verify the data are managed and stored correctly
- Data limitations report includes data management issues

Check

- Clients (program manager, GAO, public) determine whether the performance measure and result as documented are representative of program performance
- External and internal quality assessments of data and database are evaluated
- Stakeholder feedback is considered

“Check” IT/IM Issues for Sediments

- Comparison to criteria
 - documented data limitations show whether the tolerance for error is reasonable
- Transparency – the performance measure process as documented is adequate for external reviewers/auditors
- Independent oversight/audits (e.g., IMO/QA manager) were documented
- The relative impact of the uncertainty of the existing data vs. performance data & management is assessed

Act

- Relative benefits/costs assessed for quality-related improvements
- Performance measures may be modified
- Data sources, collection methods may be changed and improved
- Documentation may be updated
- Reporting may be modified

“Act” IT/IM Issues for Sediments

- Transparent – verification that the data base documentation is updated and complete
- Accountability for future plans – all modifications aren’t accomplished at once, so requires monitoring & tracking
- Bottom line – quality criteria, procedures and documentation for data and database management need to support Agency performance measures

The new Quality Policy calls for an integrated perspective on quality assurance

Drinking Water Case Study

- Focus of case study is how well DW Program follows the PDCA quality mode
 - Plan: Establish the objectives and processes necessary to deliver results in accordance with customer requirements and the organization's policies
 - Do: Implement the process on program specific activities
 - Check: Monitor and measure processes, products and services against policies, objectives and requirements
 - Act: Take action to continually improve process performance (if appropriate)

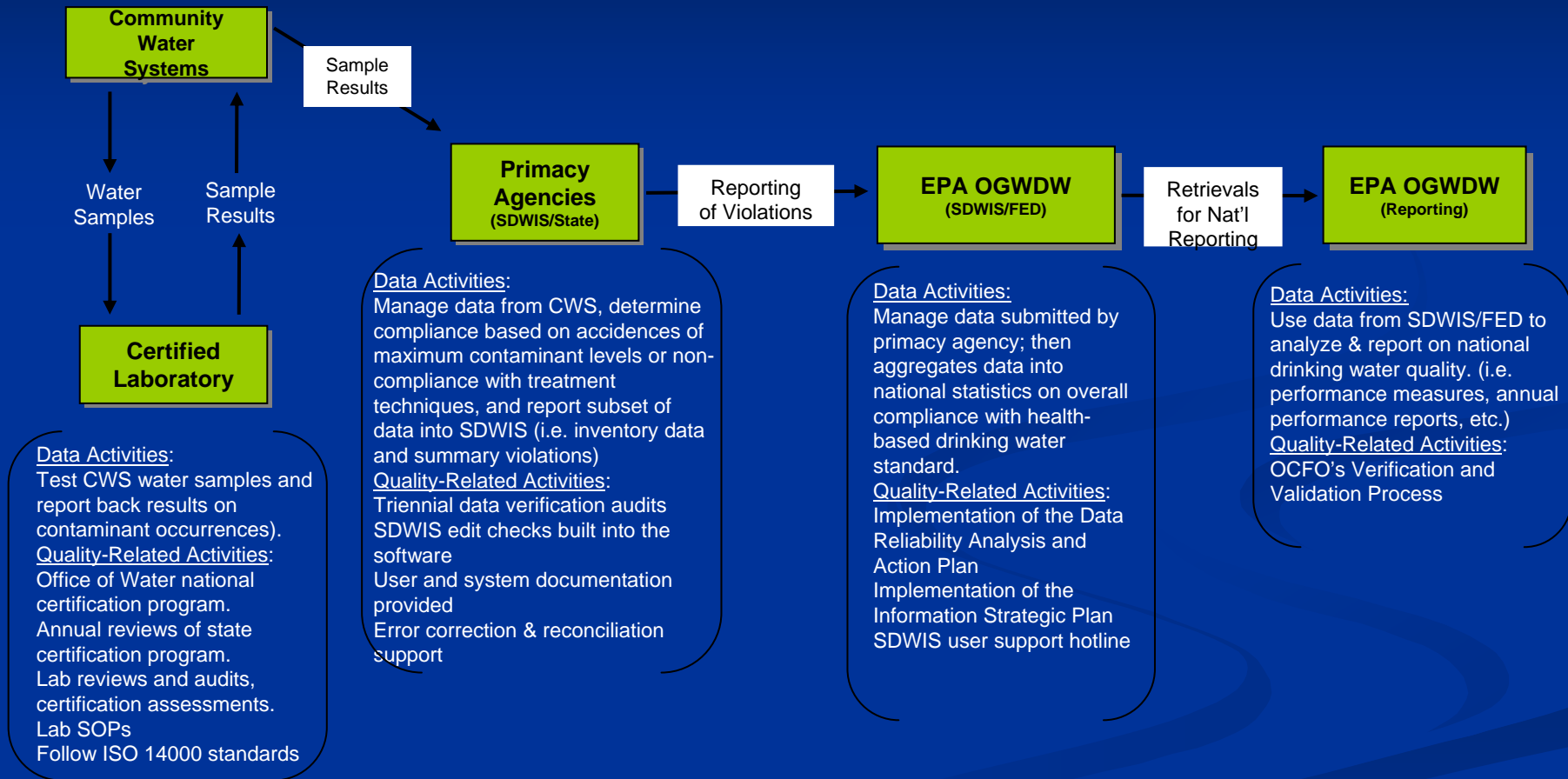
Performance Metric

- *Percent of the population served by community water systems that meet all applicable health-based drinking water standards*
= $1 - (\text{PopInViol} / \text{TotPop})$, whereby:
 - **PopInViol:** Population served by systems reporting violations. This number is based on TotPop and SysInViol
 - **TotPop:** The total population served by current, active water systems
 - **SysInViol:** The number of CWS systems reporting violations

Plan

- Systematic QA project planning to establish:
 - Objectives
 - Data requirements and criteria
 - Processes
 - Roles and responsibilities
- Project Planning, but QAPP was not developed for DW metric:
 - Project planning included objectives, data sources, processes and responsibilities
- SDWIS/FED Data Reliability Action Plan:
 - Includes processes and data standards
 - SDWIS considers it to be comparable to a QA project plan

Do – Process for Reporting SysInViol



Check - Oversight

- Primacy agencies' information systems and compliance determinations are audited about every 3 years
- SDWIS/FED edit checks built into the software to reject erroneous data
- OGWDW is working with the states to complete Data Quality Review and Plan
- State data verification and other quality assurance analyses
- SDWIS data quality was identified as an Agency weakness in 1999
- In a 2006 report the IG maintained that drinking water data are vulnerable to fraud and inappropriate procedures

ACT – Continual improvement

- ISP was implemented to improve tools and processes for creating and transferring data to EPA and allow the flow of data via CDX
- OGWDW is working with states to implement DRAP and ISP to improve the completeness, accuracy, timeliness, and consistency of the data in SDWIS/FED

Findings

- Complexity of Data Sources
- Difficulties Interpreting Results
- Inconsistent Implementation of Agency Quality Requirements
- Quality Gaps in the Life Cycle
- Unclear Roles and Responsibilities

Discussion Issues

- How do we improve QA planning for performance metrics?
- How do we fill in gaps in the life cycle?
- How do we get all parties to the table?
- Where should new guidance focus?
- How do we involve the QA community in promoting performance data quality?

Contacts

- Joe Greenblott, OCFO 202.564.4250
Greenblott.joseph@epa.gov
- Pat Mundy, OEI 732.906.6988
Mundy.patricia@epa.gov
- Ron Shafer, OEI 202.564.5173
Shafer.ronald@epa.gov
- Judy Lieberman, OCFO 202.564.8638
Lieberman.judy@epa.gov

Performance Data Sources and Systems

- Following slides provide example of major data sources and systems for performance data.
- Information compiled from FY2009 V/V
- These slides will not be presented, but provide useful background.

Performance Data Sources and Systems: Water

- SDWIS
- DW-NIMS
- UIC database
- NHANES
- WBDO
- PRAWN
- STORET
- WATERS
- GLENDA
- Surf Your Watershed
- ATTAINS
- PCS
- WATA
- GRTS
- ICES-NPDES
- CW-NIMS
- CBR
- STARS
- USACoE/RAMS
- EMAP/NCA

Performance Data Sources and Systems: Air

- ETS
- CASTNET
- NADP
- TIME/LTM
- AQS
- AIRNow DMC
- FREDS
- NEI Databases
- ATC
- RADNet
- CPPTS

Performance Data Sources and Systems: Land/Waste/Recycling

- DoC
- DoE/ACAA
- RCRIS
- WSTARS
- CERCLIS
- ACRES
- GSN CRM
- Performance Track on-line

Performance Data Sources and Systems: Toxics/Pesticides

- US Census
- Surveys & Publications
- USDA
- USGS NAWQA
- PCC/TESS
- Doane Marketing Research
- NHANES
- OPPIN/PRISM
- SAI
- FLPP
- RSEI
- CBITS
- HPVIS
- UNEP Partner Clearinghouse