

Environmental Information



EN2017

FACILITY INTEGRATION: TOWARDS A ROSETTA STONE

Ben Way, Wyoming Department of Environmental Quality

Joshua Kalfas, Oklahoma Department of Environmental Quality

Ron Evans, U.S. EPA, Office of Air and Radiation

Susan Joan Smiley Baker, U.S. EPA, Office of Environmental Information

2017 Exchange Network National Meeting

Innovation and Partnership

May 16-18, 2017

**Sheraton Philadelphia Society Hill Hotel
Philadelphia, Pennsylvania**

<http://www.exchangenetwork.net/en2017>

ABSTRACT

This session presents the initial outcomes and lessons learned from the work of the Facility Team in Phase II to develop a common understanding of facility information for E-Enterprise. Reflections from Oklahoma's experience frame the challenges. User stories, business rules and the results of pilot work streams continue to develop a path forward for shared facility information.

Agenda

- Challenges and Lessons Learned from Oklahoma
Joshua Kalfas, *Oklahoma DEQ*
- E-Enterprise Facility Phase II Outcomes and Lessons Learned
Facility Team Co-Chairs
- Questions
- Session tomorrow @ 8:30 to get your input on ideas/plans for next steps and what you hear today



E-ENTERPRISE
for the environment

Modernizing the business of environmental protection

Presentation Section Structure

Topics Include:

- **General Background**
- **Oklahoma Background**
- **Integration Complexity**
- **Lessons Learned**
- **Looking Forward**

General Background

Q: What is ‘facility integration’?

A: ‘Facility integration’ is the practice or process of mapping and using ‘facility’ data produced by a primary source for a secondary purpose. Establishing relationships with a primary source is required in ‘facility integration.’

General Background

Q: What is ‘facility’?

A: Within the context of ‘facility integration,’ ‘facility’ is the most general term used to identify either something that is a place or something that has an ability to conduct activities that are of environmental interest.

facility_{GHGRP}

facility_{TRI}

facility_{EIS}

General Background

Q: What is ‘facility integration’ for?

A: It depends.

- **Use case – what the final (data) outcome is used for:**
 - Identify ‘nearby’ facilities
 - Support reporting air emissions to programs (CAER)
- **Use case – what a tool is used for (how it is used):**
 - Link programs’ facilities
 - Update facility attribution

Oklahoma Background

Driver:

- **Comprehensive Facility View**

Tools:

- **MilkCow**
- **Nexus**
- **Master Facility List**
- **Fido**
- **Facility Management System***
- **Facility Profiler**
- **nothing**

Oklahoma Background

Users:

- High level viewers, not data providers

Solutions:

- Required significant resources
- Integration product conflicted with Primary Data Sources
- Did not support use cases or goals

Oklahoma Background

Barriers:

- Fundamental disconnect between programs
 - [Organization : ‘facility’] relationship varies
 - Investments in existing systems varies
 - Freedom to change varies
- “My data is the best,” says everyone.
- Documentation

Oklahoma Background

Barriers (continued):

- **Data models did not support use case(s)**
- **Use case(s) did not support solution(s)**
- **Solution(s) did not meet goal(s)**
- **Lack of clearly defined use case(s) and goal(s)**
- **No use of integrated data by data submitters**

Integration Complexity

Linking

The process of connecting two or more records or datasets based on a common attribute or spatial location.

**Uniqueness is maintained.
a.k.a. associating,
connecting, relating.**

Vs.

Merging

The process of combining two or more records or datasets into a single record or dataset. A common attribute or spatial location is not required. Uniqueness is not maintained.

Integration Complexity

Link	Prog _A :Prog _B	Merge
n/a	1:0	n/a
appropriate	1:1	may/may not be appropriate
appropriate?	1:many	may/may not be appropriate
may/may not be appropriate	many:many	may/may not be appropriate
may/may not be appropriate	1:part	inappropriate

- What might be a ‘facility’ in one program may be subordinate to a ‘facility’ in another program.
- “Appropriate” is goal specific and assumes known cardinality between any two programs.

Integration Complexity

My facility is not your facility.

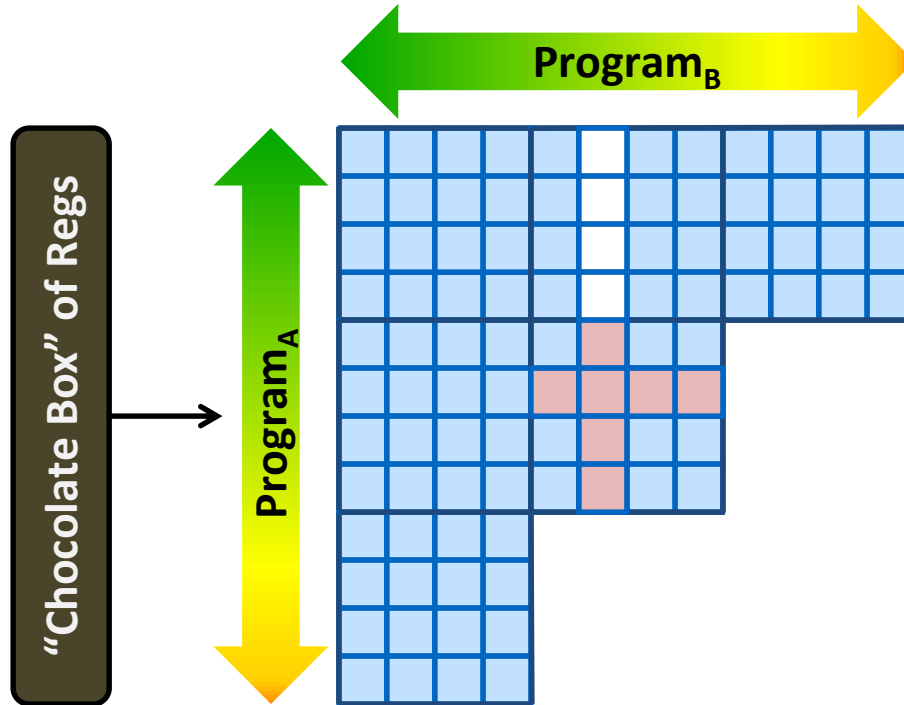
My 'facility' is not your 'facility.'

My 'facility' may/may not be your 'facility.'

My Facility_{EIS} may/may not be your Facility_{TRI}.

A:B and A:C does not mean B:C

Integration Complexity



Program Divergence

- Regulations
- Professional Definitions
- Program Business Rules
- Technology/Tools

Integration Complexity



“Blind monks examining an elephant”
Hanabusa Itchō, (1652 – 1724)

How can anyone describe the whole until he has learned the total of the parts?

<http://davidmengart.blogspot.com/2014/03/of-elephants-and-blind-men.html>

Lessons Learned

Contributors must be explicit and avoid using professional jargon

- **IT and program Subject Matter Experts (SMEs) are required to collaborate throughout the entire process**
- **Large % of time spent is used to ensure people are talking about the same thing (personal observation)**

Lessons Learned

Data providers must be stakeholders in integration outcomes

- **Out of site is out of mind...**
- **Garbage in → garbage out**
- **Vested interest precludes garbage**

Lessons Learned

‘Facility’ cardinality across programs is not defined and is not consistent

- **Cross-program cardinality is buried in regulations**
- **Defining cardinality is goal specific**
 - **Integration up (least common multiple)**
 - **Integration across (lowest common denominator)**

Lessons Learned

Technology is not a barrier to ‘facility integration’

- **Diverging program regulations, business rules, needs, and practices are significant obstacles**
- **Cross program/holistic/integrated subject matter expertise to guide integration is uncommon**

Lessons Learned

Specific technology/tools are required

- Relational Database
- Geographic Information System or Service

Outputs from automated processes will always require integration SME review

Responsibility must be explicit

Looking Forward

‘Facilities’ are emergent

- Regulated ‘sub-facility’ parts cause emergence**
- The place or thing that has an ability to conduct activities that are of environmental interest “wags” the ‘facility’**
- ‘Facility’ is a simplified term indicating interface between regulatory abstraction (environmental interest) and physical reality**

Looking Forward

Clearly and explicitly identify goals

- Final (data) outcome
- Tools (and uses)
- Required for effective cost/benefit analysis

Failure to comply is inevitable

- Build in break points/crumple zones
- Control deformation and explicitly state limits

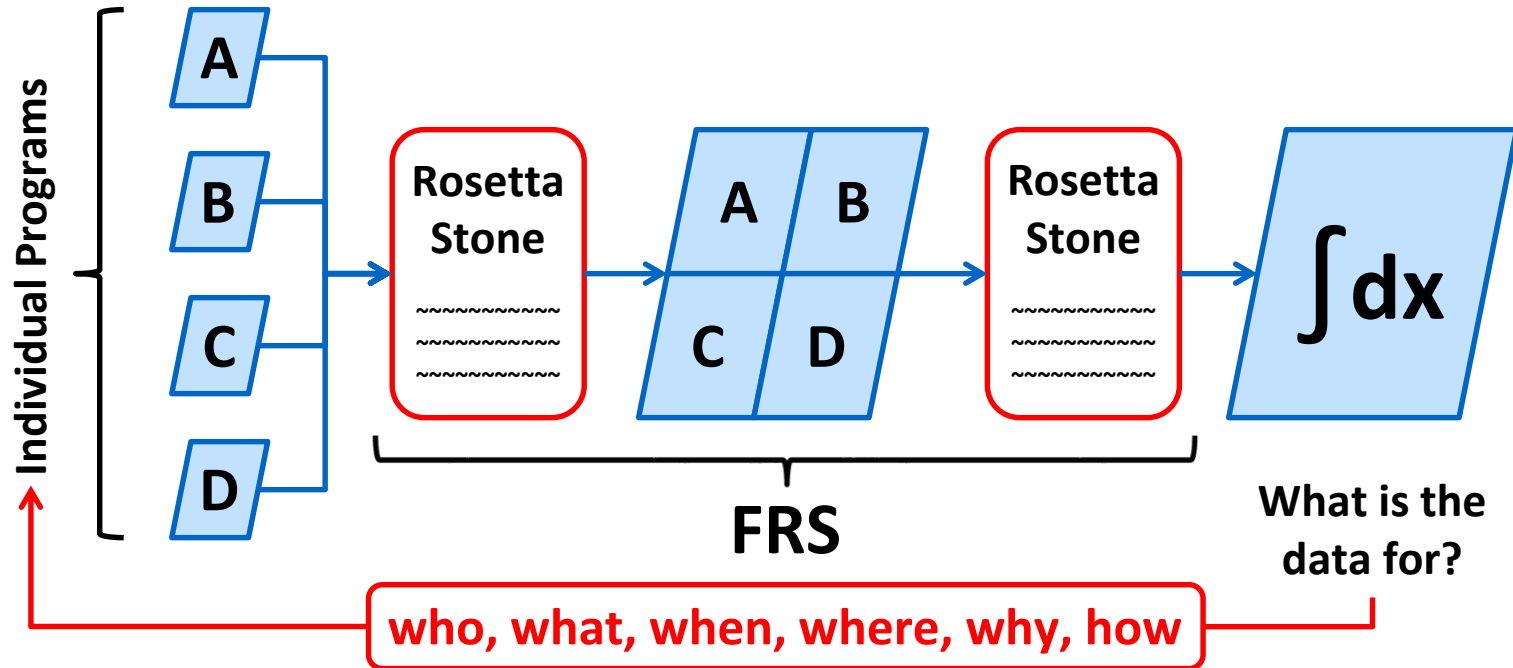
Looking Forward

Regulations diverge, so begin converging

- **Data Models**
- **Schema**
- **Elements**
- **Data Standards**
- **Business Rules**

Implement best practices, modernize, and streamline

Food for Thought



Food for Thought

- **Scientific Management (of knowledge)**
- **Epistemology**
- **Ontology**
 - **Upper**
 - **Domain**
- **Horizontal Integration (business)**
- **Vertical Integration (business)**

E-Enterprise Facility Phase II Background

Facility Team Co-Chairs

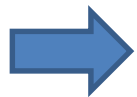
Facility Team: Component of E-Enterprise

E-Enterprise Leadership
Council (EELC)

Management Board
(MB)

Interoperability and
Operations Team

E-Enterprise Teams



Advanced Monitoring
Communications Team
Facility Team
Integrated Watershed Monitoring Networks
Local Government Portal
Pesticides Data Accessibility and Label Matching
Smart Tools for Inspectors
Tribal Roadmap Workgroup

Combined Air Emissions Reporting (CAER)
e-Permitting
Integrated Identity Solution Project (ISOL)
Leak Repair and Detection (LDAR)
Measures and Metrics Workgroup
Portal Development
Shared Services IPT

Solving a Problem

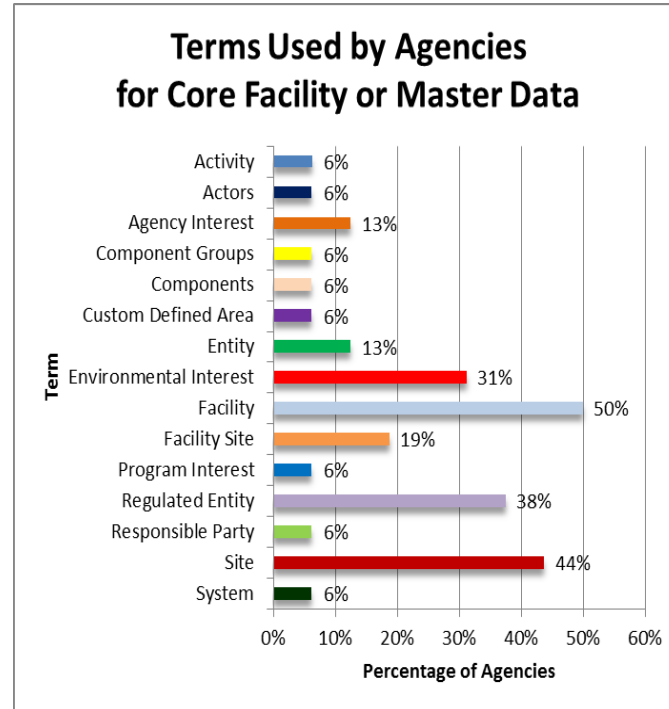
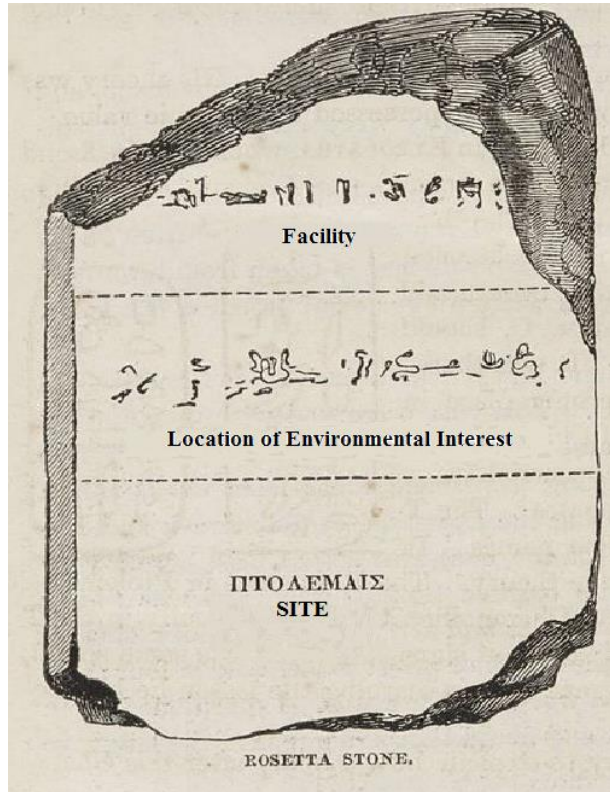
- Integrated, reconciled, facility information is a key to solving the problem of:
 - Reducing regulatory burden
 - Increasing transparency
 - Ensuring best data available to make decisions
 - Improving data quality

The Rosetta Stone



- Inscribed with three versions of a 196 BC decree; rediscovered in 1799
- Top text in Ancient Egyptian hieroglyphic script
- Middle text in Ancient Egyptian Demotic script
- Bottom text in Ancient Greek
- The key to deciphering previously untranslated hieroglyphic language

Developing a common understanding of facility information



From Phase 1 Discovery and Analysis

Phase I Accomplishments

- **Information gathering** about a diverse set of partner experiences related to facility master data management (MDM)
- **Cataloging existing partner systems** and summarizing their features and characteristics
- Gathering information about **lessons learned** on system implementation, business process changes, and governance
- Conducting detailed 'one on one' **discovery sessions** with 3 state members to allow for a deeper dive into business processes, data models, technical approaches, lessons learned and challenges
- **Compared state business rules** with existing EPA FRS services and other EPA program business rules
- Authored, reviewed, and revised an **Discovery and Analysis Document**

Facility Team Goals

- Help programs and agencies manage responsibilities more efficiently
- Reduce industry reporting burden of redundant facility information
- Assemble more quickly the multi-media environmental data needed for consolidated reports, permits, and inspections
- Provide the public more complete understanding of regulatory obligations and environmental impacts at each facility
- Increase facility data accuracy

Common Vision for Facility Coordination and Collaboration

- Integration and correction of data in as near to real-time as possible
- Common facility profile model that allows for varying levels of granularity
- Shared business rules and mapping to common-enough terminology
- APIs flexible enough to work with EPA, state and other systems
- Shared good practices and tools

Focus of Work Streams

- The deliverables of the work streams progressively refined based on findings of other work streams
 - User Stories
 - Business Rules
 - Facility Model/APIs
 - EPA adoption of a state's Master Data Management (MDM) model pilot with Rhode Island
- Proactive communication for stakeholder input and status reporting

Comprehensive Use Cases (*User Stories*)

- Articulates full range of perspectives
- Stories associated with one or more key objectives:
 - Streamline data operations
 - Increase data accuracy
 - Support program analysis
 - Support data systems interoperability
 - Improve public understanding

Business Rules Approach

- Focus on the high level, cross cutting issues
- Ensure that rules address both system and data structure especially regarding data stewardship and hosting/maintenance of shared services
- As a prototype, work with the air media first
 - CAER as a live test bed
 - Has extensive EPA/State interest and responsibilities down to fine granularity
- Plan for governance for the short and long term
 - Resolve conflicts/issues to the extent possible; elevate to E-Enterprise Management Board and Leadership Council as needed

Facility Profile Model/API Approach

- Initial development of potential re-usable solutions to pilot that would accomplish these functions:
 - Correct facility data as it's reported
 - Curate facility data in a central location
 - Leverage FRS data quality tool, the Facility Linkage Application
 - States can currently have access; interface mods needed for state-specific requirements
- Integrate facility data system-to-system
 - Enable partners to share their integrated facility/site data with EPA's FRS
- ***Develop a data curation shared service***

Goals for Conducting Facility Pilot

- Develop accurate, cross-functional view of facility
 - Compliance history, operating status, others
- Provide holistic view of facility information for analysis and impacts
- Improve data quality and accuracy
- Create awareness and insight into partner approaches and solutions

Goals for Conducting Facility Pilot

- Demonstrate an API to deliver data validation lookup using EPA's FRS for a set of facility attributes using a state facility/permit system
- Demonstrate a Master Data Management solution sharing facility information between a state and EPA's FRS
- Test tools and approaches for compatibility with state systems
- Document analysis of pilot results include technical barriers, possible solutions, data accuracy metrics, concepts for governance

EPA Adoption of a State's Master Data Management (MDM) Model Pilot

- Rhode Island partner
- Goal: Demonstrate EPA adopting appropriate linkages, allowing state data to be accepted as the master record

Focus on Product-based Outcomes

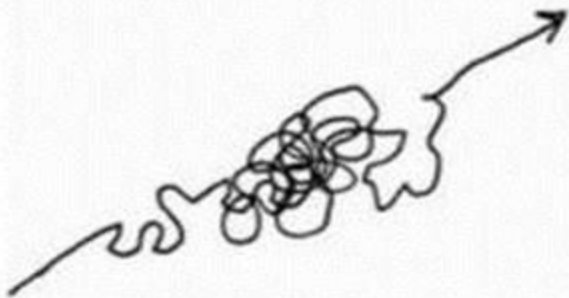
- Outcomes should be broadly applicable to the Enterprise
- Avoid point solutions even though it may solve a particular problem
- Use shared resources/services
- Develop repeatable processes

Success



what people think
it looks like

Success



what it really
looks like

E-Enterprise Facility Phase II Outcomes

Facility Team Co-Chairs and State
Pilot Participants

Primary Outcomes

- Expanded documentation of detailed requirements
- Initiated development of shared services (APIs)
- Developed framework for executing repeatable analysis of state/tribal requirements
- Continued progress toward next phase

Expanded User Stories

- Documented more use cases for future development
 - More than 50 user stories for backlog
 - Actors: states, federal, tribes, regulated facility, public
 - Actions: reporting, data corrections/updates, inspections, permitting, etc.
 - Outcomes: data quality, program value, reduce redundant data entry, etc.
- First cut completed and passed to Business Rules
 - User Stories available via SharePoint
- Always seeking comment/refinement and additional stories

Addressing Broad Business Rules

- Business rules are the heart of making facility coordination and collaboration work
- Business rules team tackled several high level, cross cutting areas which will have broad applicability
 - e.g., how do you create/modify a facility ID which will be recognized by disparate State and Federal programs

Business Rules Process

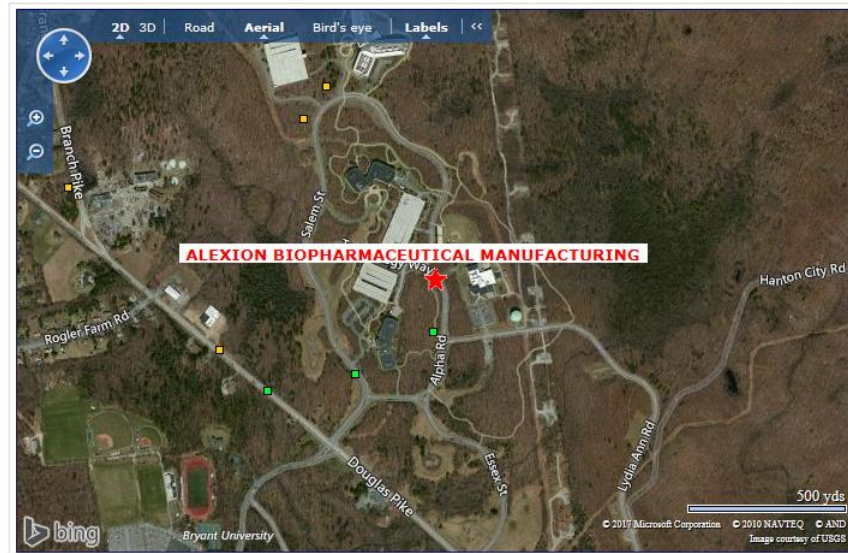
- Representatives from 6 EPA program areas and 6 states are part of the team
 - Each provided background information on their approach to handling these high level questions
- Extensive discussions on each have developed into draft business rules
- As part of the vetting process for these, the business rules team is working with the Rhode Island MDM pilot team
 - Focused evaluation of business rules developed by team
 - Business rules team is evaluating business rules developed for the RI pilot for broader application

Pilot with Rhode Island

- Using RI's facility master data management system and EPA's FRS
 - PLOVER: Permits, Licenses & Other Vital Environmental Records
- Weekly collaboration between EPA and RI Department of Environmental Management for detailed requirements analysis
- Agile sprints for development

PLOVER Integration into FRS (Prod)

- 11,020 PLOVER program facilities integrated
 - 8,893 are not linked to another program record
 - 2,127 were linked to existing FRS master records
 - 643 were linked to another PLOVER record via the same master record



Examples – PLOVER facilities linked to existing FRS records:

https://iaspub.epa.gov/enviro/fii_query_detail.disp_program_facility?p_registry_id=110016712912
https://iaspub.epa.gov/enviro/fii_query_detail.disp_program_facility?p_registry_id=110012520600

Initiate Development of APIs

- Application Program Interfaces/Shared Services for:
 - Query services
 - Submit services
 - Admin services for joint data curation
 - Correcting data real-time as it is being reported
 - Correcting data after it has been reported

CAER Gap Analysis

- Identify gaps in FRS (new) data model and API for Air programs' reporting
 - Sub-facility components in scope of analysis
- Better understand a state's framework for managing Air data within the scope of CAER
- Develop repeatable process for more states to add to the gap analysis

E-Enterprise Facility Phase II

Lessons Learned

Facility Team Co-Chairs

Flexibility Required

- Difficult to define the outcome before the process starts
- Use guiding principles to keep objectives in sight
- Determine priorities collaboratively
- Align benefits and interest of involved parties

Resource Challenges

- Identifying all necessary resources up front difficult
- Program and IT support needed from all parties
- This “other duty as assigned” competes with program operations priorities
- Not enough time or money to do everything needed

Pilots Have Long Lead Times

- It takes many conversations to develop a shared understanding of
 - Systems
 - Processes
 - Partnerships
 - Technical requirements
- Analysis process must be thorough

Spectrum of Integration Needed

- A wide variation of integration solutions exist today across partner systems
 - Most built for high level tasks
- For facility information integration, flexible solutions needed
 - Leveraging shared services

We're Headed in a Good Direction

- Identifying commonalities
- Building shared services
- Engaging broadly
- Moving forward

Success Obstacles

- Working in silos - includes program management, program development, rules
- Inventing isolated solutions
- Believing problems are unique
- Waiting to engage

Steps For Integration Development

- Clearly identify specific goal
- Identify information/data needed for success
- Ensure data model(s) accommodate data needs
- Map data from source to source, acknowledging program data stipulations
- Build decision logic to support goal
- Leverage existing shared services, note shortcomings
- Verify data, mapping and logic meet goals

Facility Team Co-Chairs

Ben Way

Wyoming Department of Environmental
Quality

ben.way@wyo.gov

307.777.7017

Joshua Kalfas

Oklahoma Department of Environmental
Quality

joshua.kalfas@deq.ok.gov

405.702.4210

Ron Evans

U.S. EPA, Office of Air and Radiation

evans.ron@epa.gov

919.541.5488

Susan Joan Smiley

U.S. EPA, Office of Environmental Information

smiley.susan@epa.gov

919.541.3993

Questions and Discussion

Join us at 8:30 am on Thursday
for Facility Phase III Listening Session