FACILITY TEAM PHASE II SUMMARY

E-Enterprise Facility Team

Version 1.0

May 2018

Table of Contents

Revision Log

Date	Version No.	Description
April 12, 2018	0.01	Preliminary draft
April 16, 2018	0.02	Incorporated first round of feedback from Facility Team co-chairs
April 20, 2018	0.03	Incorporated second round of feedback from Facility Team co-chairs
April 26, 2018	0.04	Incorporated updates to SharePoint description, CAER content and final round of comments from Facility Team co-chairs
April 27, 2017	0.05	Incorporated additional comments from Facility Team co-chairs
May 7, 2018	0.06	Incorporated comments from Facility Team co-chairs
May 11, 2018	1.0	Final

1.0 Introduction

Facility data is at the core of federal and State, Local, and Tribal (SLT) environmental regulatory processes. Knowing a facility's name, ownership, location and characteristics are key to a comprehensive picture of past, current, future, and potential environmental impacts. Linked to other critical environmental data such as ambient air and water quality data, census figures, and other demographic information, facility data has the capacity to provide a comprehensive picture enabling co-regulators to better protect human health and the environment.

Sharing integrated facility identity information between SLTs and EPA is critical to achieving the vision of E-Enterprise for the Environment (E-Enterprise). EPA and each SLT's environmental agency—as well as each program office (such as air, water, and waste) within those agencies—separately collect, record, and maintain both the specific data needed for each set of regulations and core information to identify each regulated facility by name, address, geographic location, owner, etc. Facilities currently must report and update the same basic identification data to multiple programs, at multiple levels of government, and at multiple times. Oftentimes, there are discrepancies among these disparate sets of facility identification data (synonymous with facility data management) that prevent accurate correlation across programs and agencies. E-Enterprise offers an opportunity to develop approaches to integrating facility identification data across co-regulators and their programs and correcting facility data as it is being reported.

Phase I of this effort focused on a thorough analysis of the current state of facility data management. A summary of Phase I findings is available in the *Facility Data Integration Discovery and Analysis Document* which is available on the E-Enterprise SharePoint site at this link and on the Facility page of the publically accessible E-Enterprise web page.

1.1 Phase II Overview

Facility Team members organized Phase II topics into analysis work streams to discuss user stories, business rules, a shared facility data model and application programming interface (API), and to consider the integration between the Facility Team and the E-Enterprise Combined Air Emissions Reporting (CAER) team. The original Phase II scope initially considered three pilot efforts; however, due to resource constraints and partner availability, the team proceeded with one pilot related to State Master Data Management with Rhode Island.

This document summarizes the work completed during Phase II. Each section provides an overview of work completed in each work stream. Deliverables or artifacts that document the work completed by the work streams are available in the Facility Team work space in EPA's SharePoint site. SharePoint is the working space used by the Facility Team and other E-Enterprise teams as a way to collaborate in real time on project deliverables. It requires a login and password. If you are part of E-Enterprise governance or project teams and need SharePoint access, please contact one of the Facility Co-Chairs (see contacts at the end of this document). For other questions related to E-Enterprise, please visit the E-Enterprise 'Contact Us' page at: https://www.epa.gov/e-enterprise/forms/contact-us-about-e-enterprise-environment. Final versions of work stream deliverables will also be available on the Facility page of the E-Enterprise web page. The products of each of the work streams are not included here due to size, length

and complexity. However, the corresponding sections of this document provide an effective summary of the work products that are in SharePoint.

1.2 Phase II Results

During Phase II, the Facility Team successfully executed a variety of tasks across multiple work streams. These accomplishments are in line with the 5-year plan envisioned during Phase I. Phase II results included:

- Developing User Stories to document broad requirements and guide future pilot services development,
- Evaluating a shared data model to determine the best approach to share facility data,
- Examining the integration of work under the Facility Team with that of the Combined Air Emissions Reporting (CAER) E-Enterprise team,
- Documenting business rules that will be critical to sharing facility data across the enterprise, and
- Successfully piloting three web services in a test environment and developing a Shared Facility Services Implementation Guide.

The Facility Team expects to continue advancing facility integration in Phase III by piloting expanded services with additional partners and deploying shared facility services to production. In Phase III, the Facility Team hopes to more fully explore governance, develop a micro-service to populate facility status, and perform a detailed mapping of facility data across the four air systems that are part of the E-Enterprise CAER project. Expansion of the E-Enterprise facility services will include development and pilot implementation of services including sub-facility and geospatial fields.

2.0 User Stories

The User Story work stream developed user stories based on real and hypothetical experiences of state, industry, EPA, and the public that document and corroborate the necessity of a shared facility approach. A user story is a simplified version of a requirement and describes who the user is, what they are trying to do, and why they are trying to do it.

The User Story work stream's goal was to capture approximately 80% of user stories related to facility data integration in the initial document so that the other work steams could be initiated. The work stream members identified the following five high-level requirements for facility integration:

1. Streamline Data Operations
2. Increase Data Accuracy
3. Support Program Analyses
4. Support Data System Interoperability
5. Improve Public Understanding

Using these high-level requirements as a foundation, the work stream members developed 43 user stories and linked each user story to potential users (e.g., public, state data steward, etc.). Four examples of User Stories are provided below:

- As a concerned citizen, I want easy access to environmental data for all media so that I can quickly identify any environmental concerns near to where I live.
- As a submitter for a regulated entity, I want facility data to be automatically populated so that I don't have to spend time entering data that has not changed since my last submission.
- As a staff member at a regulatory agency, I want to have access to the latest facility data (potentially supplied by another agency/data source) so that I can be certain that the information that I am working with is current.
- As an environmental program staffer, I want to use data from other programs so that I can verify the accuracy of data submitted under my program.

The work stream team members also identified governance issues and questions to explore in future phases of this effort. The user stories developed by this work stream are available in SharePoint at this link.

3.0 Data Model and API Work Stream

The Data Model and API work stream explored technical alternatives for conceptual facility data management as well as applied pilot work. Through a series of group discussions, presentations, and collaborative analysis of documents, the team compiled a collection of artifacts documenting key data model and API topics. The scope of work completed through this work stream will be expanded in later phases as shared facility services are expanded and tested more fully. A summary of the artifacts is available in SharePoint at this link.

EPA has recently extended the Facility Registry Service (FRS) data model to meet evolving needs for more granular facility information. This work resulted in the ability of FRS to manage sub-facility data. The initial work was completed to meet requirements for air facilities, but this work can be leveraged to support sub-facility data within other media such as wastewater, drinking water, or hazardous waste. The FRS data model will continue to expand in order to support the range of programs' facility data.

4.0 CAER Mapping

In this work stream, the Facility Team is collaborating with the E-Enterprise CAER team to complete a data gap analysis. This data gap analysis will inform changes or additions needed to the FRS data model and the FRS API in order to support states' data needs as well as the needs of federal emissions reporting programs under the scope of CAER. At the federal level, the four federal programs and systems within the scope of CAER are:

- Program: National Emissions Inventory (NEI) – System: Emissions Inventory System (EIS)
- Program: Toxics Release Inventory (TRI) – System: TRI Made Easy-Web (TRI-MEweb)
- Program: Risk and Technology Review (RTR) – System: Compliance and Emissions Reporting Interface (CEDRI)
- Program: Green House Gas Reporting Program (GHGRP) – System: Electronic Greenhouse Gas Reporting Tool (e-GGRT)

In this work stream, the Facility Team is leveraging the work of the CAER sub-teams that are developing the detailed CAER reporting requirements and data mapping from the air programs' perspectives.

The gap analysis is evaluating data element differences between the FRS data model, the FRS API, the four federal systems within the scope of CAER, and systems from four states (Iowa, North Carolina, South Carolina, and Massachusetts) that support state air programs. The result of this gap analysis is a spreadsheet available in SharePoint at this link.

This detailed analysis is laying the foundation for defining requirements in the next phase for the development of web services to support the combined emissions reporting form.

5.0 Business Rules

The Facility Business Rules work stream members documented high level business rules to enable integration of facility information across programs, among partners, and across the enterprise. The team consisted of members from 11 States and several U.S. EPA programs. The members provided specific information on how their programs use facility information. Programs discussed by work stream members included state and federal systems for Air Pollution Control, Drinking Water, National Pollutant Discharge Elimination System (NPDES), and EPA's Facility Registry Service.

Using the current systems' frameworks of facility data management, the team reflected on the business rules necessary to integrate facility information across various existing systems. These business rules will continue to be refined in future phases and as facility integration shared services are developed. Within the business rules document, rules are grouped into ten categories and lettered A through J. Each letter contains a summary business rule and a sub-set of detailed rules to support it. A summary list of the high-level business rules is provided below. A summary of the business rules and the full business rules document are available in SharePoint at this link.

A. Facility Record Definition: The definition of facility should stay within the purview of the originating program.
B. Data Ownership, Data Stewardship, and Collaborative Use: Clear communication about authorship and ownership is required. For ease of application, duplicate datasets may exist as long as appropriate citation is used.
C. Data Completeness: Mandatory and optional data elements are needed to ensure consistency and enable data linkages and reconciliation across programs.
D. Metadata: Metadata for each facility record will allow facility integrators to identify the data source, assess the data quality submitted by the originating program, and enable integration. Some metadata will be entered manually, some will be derived. In all cases, metadata used in facility integration are required elements of a program record.
E. Data Standardization: Programs use various conventions for their facility records. Standardization will help facilitate data integration. These business rules do not create mandatory data standard requirements for programs, but the

standardization criteria used must be identified to enable integration. The integration standardization will be determined by the originating programs.

F. Data Validation: High quality data needs to be available for integration. Validation is needed to ensure that submitted facility data meets certain QA/QC criteria.
G. Creation of a New Facility Record: For integration purposes, it is important to know from where the data originates in the case of new program facility records, and determine cases when connections between new and existing records should apply.
H. Modification of a Facility Record/Hierarchy Decision Making: To establish a clear hierarchical decision process, it is essential to identify who can edit the data, who has primacy over the program record (not the master record), what is the relationship between primacy-overwrite-timing, and what is the link back to the primary source. The hierarchy decision modeling is still an area that needs to be further tested and might be program-specific.
I. Data Deletion/Removal History: For historical reasons, information should not be deleted if it is tied to other records, as existing linkages should be maintained. Development of other potential business rules related to data deletion/removal/history will depend on additional work needing to be done to define the use of "status" in FRS.
J. Upcoming Topics That Need to Be Addressed.

6.0 State Master Data Management Pilot

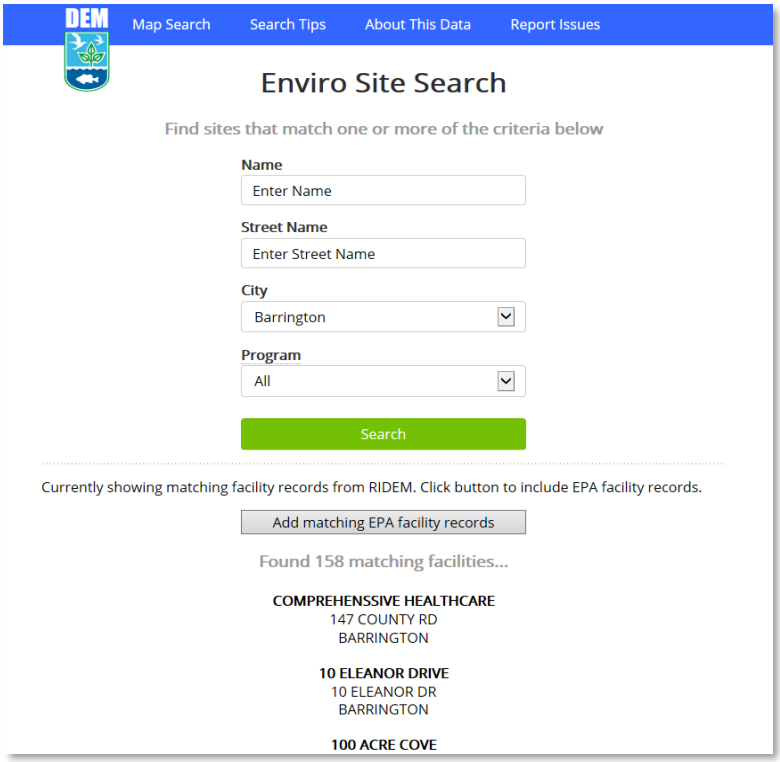
The purpose of the State Master Data Management (MDM) pilot during Phase II was to explore how facility data can be jointly managed by the EPA and a state with a mature MDM system and processes. Implementing a joint management framework provides the enterprise with comprehensive, up-to-date facility information. It allows the participating partners to incorporate additional facility information into their systems (including public-facing websites and services), improving transparency and data quality for all stakeholders.

During Phase II, EPA partnered with the Rhode Island Department of Environmental Management (RIDEM) to explore joint facility data management. RIDEM uses a state-developed customized system to manage facility data for their agency and they have a robust data governance framework in place to manage facility data. A crucial aspect of this pilot was the use of shared web services to enable this joint facility data management while acknowledging and minimizing disruption to a state's implemented MDM system and set of policies and practices.

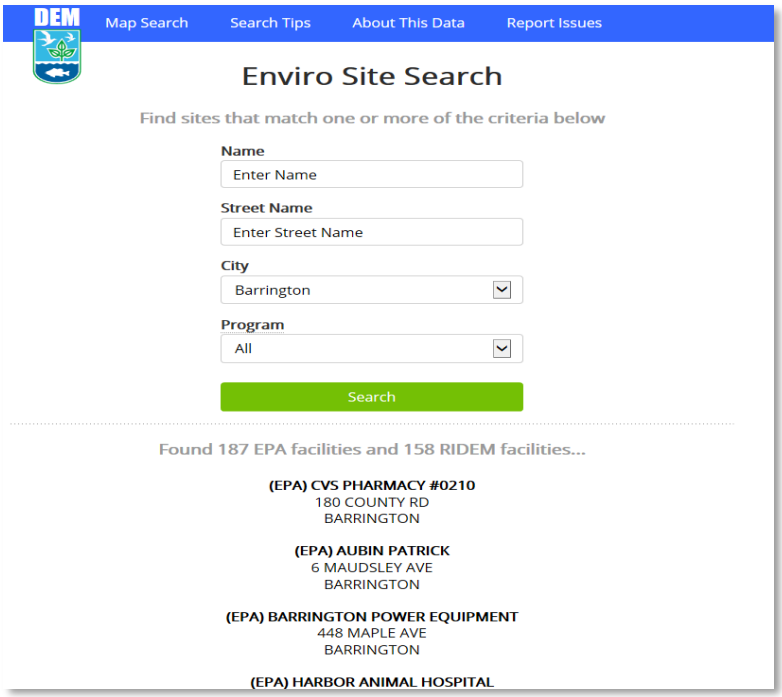
Together, EPA and RIDEM implemented the State MDM Pilot. The pilot included requirements gathering, service development and documentation, and an evaluation of impacts to "downstream" users of FRS facility data. As a result of the pilot work completed during Phase II, the State MDM Pilot team developed and implemented three shared facility services that enable real-time sharing of facility information. The suite of Shared Facility Services included two submit services and one query service.

- Submit to FRS – enables real-time sharing of facility information between a partner system and EPA's FRS. This enables an update in FRS upon the addition of a new record or edit of an existing record in a partner's system.
- Submit to Partner – enables real-time sharing of facility information between FRS and a partner's system when an edit is made to the record in FRS.
- Query FRS – enables a partner to perform a real-time call of FRS so a partner can incorporate FRS data directly into their systems or make FRS data available to the public via their system.

Services were implemented in a pre-production environment to facilitate thorough testing without impacting production (live) data. The Query FRS Service developed as part of the pilot is planned to be available in production for use by partners by the summer of 2018. The two submit services require further testing with other co-regulators' facility systems to ensure broad usability. RIDEM integrated the Query FRS service within their ePLOVER web page. ePLOVER is RIDEM's publicly-available site search. The screenshots below illustrate the test implementation of the Query FRS Service.



Upon completing a search within ePLOVER, users are given the option to add matching EPA facility records.



If the option to add EPA facility records is selected, results are displayed with "(EPA)" appended to the front of the name. It should be noted that the search results and page configuration shown here is as it exists in RIDEM's test environment as of the creation of this document and does not necessarily reflect the final implementation that will be available in production in the near-future.

The State MDM Pilot Team developed a draft Implementation Guide to provide information to partners who wish to implement shared facility services and includes estimated levels of effort that potential partners can use when considering implementation of these services. This document is available on SharePoint at this link. This draft Implementation Guide will be a document that continues to evolve in continued facility work in subsequent phases as more partners test and implement the services.

The team also produced a summary document with detailed information completed during the Phase II State MDM Pilot. This document contains the detailed analysis of the requirements for the services, the results of the downstream evaluation testing, and the lessons learned through pilot services implementation. It is available in SharePoint at this link.

7.0 Phase III Planning

The Facility Team anticipates that Phase III will include a number of tasks to build on the results of Phase II and to continue on the current trajectory to realize the Team's goal of facility integration across the enterprise. Anticipated Phase III tasks are in line with the 5-year plan for the Facility Team as envisioned during Phase I. Completion of potential Phase III tasks will depend on availability of partners for pilots and other sub-teams and resource availability.

During Phase III, the Facility Team anticipates expanding the services developed during the State MDM Pilot with Rhode Island to incorporate additional functionality. The team expects to collaborate with an additional partner or partners to test expanded services in a non-production environment, complete a comprehensive review of downstream user

impacts, and, once sufficient testing and review is complete, move services to production for use and adoption by partners. The Facility Team is also planning to pilot the use of EPA's Facility Registry Service (FRS) tools to support MDM with a partner that does not have that capability in-house. This will demonstrate that a partner that does not have a facility MDM system can take advantage of shared facility services.

Additional work planned for Phase III includes completing a comprehensive gap analysis with the entire CAER community to pilot and test sub-facility web services with the four air systems that are part of CAER (NEI, TRI, GHG, and CEDRI). The Facility Team would also like to scope requirements for a unified facility status micro-service to support all environmental programs and to scope a facility governance framework. A facility governance framework will leverage and iterate upon the Business Rules, User Stories, and Shared Facility Services Implementation Guide developed during Phase II based on further pilot outcomes and the needs for governance in enterprise facility integration.

8.0 Facility Phase II Participants

The Facility Team wishes to acknowledge the following work stream and Pilot Team participants. The Facility Team would like to extend special thanks to the tremendous effort, time, attentiveness, and dedication show by the team members from the Rhode Island Department of Environmental Management who participated in the State MDM Pilot. Their contribution to this effort will help usher in a new era of joint collaborative information management, and their help in advancing the mission of E-Enterprise is greatly appreciated.

Name	Organization	Work Stream/Pilot Team Participation
Ron Evans	EPA, OEI Facility Team Co-Chair	Business Rules, State MDM Pilot, CAER
Joshua Kalfas	Oklahoma DEQ Facility Team Co-Chair	Business Rules, State MDM Pilot, Data Model and API, User Stories
Susan Joan Smiley	EPA, OEI Facility Team Co-Chair	Business Rules, State MDM Pilot, User Stories, Data Model and API
Ben Way	Wyoming DEQ Facility Team Co-Chair	Business Rules, Data Model and API, CAER
Sam Alves	Rhode Island DEM	State MDM Pilot
Warren Angell	Rhode Island DEM	State MDM Pilot
Gary Arbuckle	Texas TCEQ	Data Model and API

Lynn Barnes	South Carolina DHEC	Business Rules
Michael Beatty	Texas TCEQ	Business Rules
Carla Bedenbaugh	South Carolina DHEC	Business Rules
Eric Brown	Colorado DPHE	Users Stories
Lynn Capuano	Exchange Network	State MDM Pilot
Joe Carioti	EPA, OECA	Business Rules
Scott Christian	EPA, OLEM	User Stories
Eric Cleckler	Alabama DEM	Business Rules
Regina Crolley	South Carolina DHEC	Facility Team State Co-Chair
Courtney Cswercko	Iowa DNR	Business Rules, Data Model and API, User Stories
Mary Curtis	EPA, OCFO	Business Rules
Brittany Decker	South Carolina DHEC	Business Rules
Sally Dombrowski	EPA, OAR	Business Rules, User Stories, CAER
Christopher Dunbar	New Hampshire DES	Business Rules
Laurie Fleet	Texas TCEQ	Business Rules
Scott Fontenot	EPA, OECA	State MDM Pilot
Pam Galli	Rhode Island DEM	State MDM Pilot, Data Model and API
Julia Gamas	EPA, OAR	Data Model and API
Beth Graves	ECOS	Business Rules, State MDM Pilot, Data Model and API
Carolyn Greenough	New Hampshire DES	Business Rules
Shana Harbour	EPA, OCFO	Business Rules
Kien Harris	Rhode Island DEM	State MDM Pilot
Ashley Inserillo	New Hampshire DES	Business Rules

David Jacobson	EPA, OECA	Business Rules, Data Model and API, User Stories
Lisa Jenkins	EPA, OLEM	User Stories
Matthew Kelly	EPA, OEI	Business Rules, State MDM Pilot, Data Model and API
Veronica Kenkel	Colorado DPHE	Business Rules
Won Kim	Oregon DEQ	Data Model and API
Mary Kimlinger	Minnesota MPCA	Business Rules
Arno Laud	Maryland DOE	Business Rules
Josie Lopez	EPA, Region 8	Business Rules
Anthony McClard	South Carolina DHEC	Business Rules
Hamilton McClean	New Hampshire DES	User Stories
Carol Miller	South Carolina DHEC	Business Rules
Christy Monk	Alabama DEM	Business Rules
Kaitlin Murphy	New Hampshire DES	Business Rules
Kelly Poole	ECOS	State MDM Pilot
Theresa Sabbia	New Hampshire DES	Business Rules
Tobias Schroeder	EPA, OCFO	Business Rules
Leslie Simpson	Colorado DHPE	Business Rules
Robert Simpson	EPA, Region 2	Data Model and API
Jill Slain	Wisconsin DNR	Business Rules, User Stories
Mark Wert	Massachusetts DEP	Business Rules
Chad Wilbanks	South Carolina DHEC	Business Rules
Nathan Wilkes	EPA, OEI	Data Model and API
Dana Stefan	Ross Strategic	Business Rules
Rob Willis	Ross Strategic	Business Rules
Tim Bizal	CGI Federal	State MDM Pilot

Chris Chafin	CGI Federal	State MDM Pilot
Jim Chilton	CGI Federal	State MDM Pilot, Data Model and API
Alison Ferner	CGI Federal	State MDM Pilot, Data Model and API
Luke Gentry	CGI Federal	State MDM Pilot, Data Model and API
Kimberly Hoke	CGI Federal	State MDM Pilot, Data Model and API
Lydia Lyshevski	CGI Federal	State MDM Pilot, Data Model and API
Nikki Spaeth	CGI Federal	State MDM Pilot

9.0 Contact Information

For additional information on the E-Enterprise Facility Team, to participate in future phases, or for information on shared facility services, please contact FRS_support@epa.gov or the Facility Team Co-Chairs:

Joshua Kalfas

Oklahoma Department of Environmental Quality

joshua.kalfas@deq.ok.gov

Ron Evans

Office of Air and Radiation, EPA

ron.evans@epa.gov

Susan Joan Smiley

Office of Environmental Information, EPA

smiley.susan@epa.gov

Ben Way

Wyoming Department of Environmental Quality

ben.way@wyo.gov