Understanding the New Facility Registry Service's Data Model

Matthew Kelly

U.S. EPA, Office of Environmental Information

- **2017 Exchange Network National Meeting**
- May 16-18, 2017
  Sheraton Philadelphia Society Hill Hotel
  Philadelphia, Pennsylvania
- http://www.exchangenetwork.net/en2017

Overview

- FRS overview
- New data model – the details
- Examples of new data model in action
- Looking ahead

FRS Overview

FRS' 3 main goals are to:

Integrate Data

Improve Data

Provide Data

FRS links facilities and geospatial data across Programs, States, Tribes and other Federal Agencies
- Based on a consistent set of data elements (address, facility name, location)

FRS Overview

Allows users to identify…

- The permits/regulations that pertain to a particular facility, and evaluate its compliance

- The location of a particular facility

- Regulated facilities within a particular sector

Used for enforcement, rulemaking, reporting, analysis and emergency response

FRS Overview

What do we integrate?

- 33 EPA program systems: http://www.epa.gov/enviro/frs-data-sources

- 54 state systems:
  http://www.epa.gov/enviro/frs-exchange-network

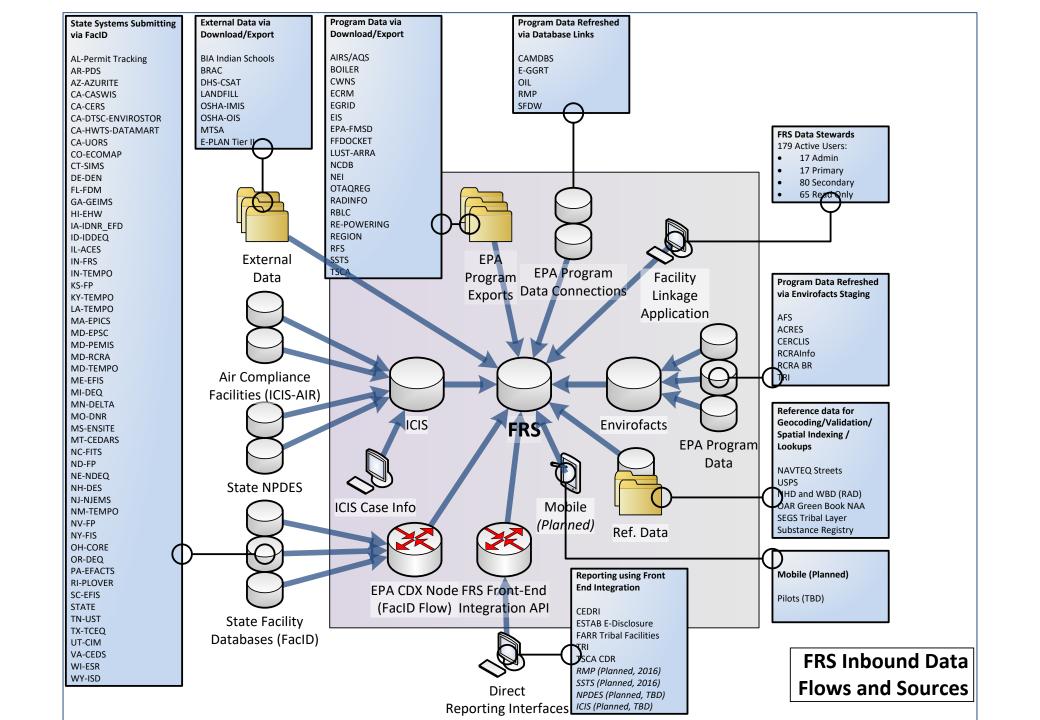- 4 tribal systems

- 8 systems from other federal agencies

www.epa.gov/frs

State Systems Submitting via FacID

AL-Permit Tracking
AR-PDS
AZ-AZURITE
CA-CASWIS
CA-CERS
CA-DTSC-ENVIROSTOR
CA-HWTS-DATAMART
CA-UORS
CO-ECOMAP
CT-SIMS
DE-DEN
FL-FDM
GA-GEIMS
HI-EHW
IA-IDNR_EFD
ID-IDDEQ
IL-ACES
IN-FRS
IN-TEMPO
KS-FP
KY-TEMPO
LA-TEMPO
MA-EPICS
MD-EPSC
MD-PEMIS
MD-RCRA
MD-TEMPO
ME-EFIS
MI-DEQ
MN-DELTA
MO-DNR
MS-ENSITE
MT-CEDARS
NC-FITS
ND-FP
NE-NDEQ
NH-DES
NJ-NJEMS
NM-TEMPO
NV-FP
NY-FIS
OH-CORE
OR-DEQ
PA-EFACTS
RI-PLOVER
SC-EFIS
STATE
TN-UST
TX-TCEQ
UT-CIM
VA-CEDS
WI-ESR
WY-ISD

External Data via Download/Export

BIA Indian Schools
BRAC
DHS-CSAT
LANDFILL
OSHA-IMIS
OSHA-OIS
MTSA
E-PLAN Tier II

Program Data via Download/Export

AIRS/AQS
BOILER
CWNS
ECRM
EGRID
EIS
EPA-FMSD
FFDOCKET
LUST-ARRA
NCDB
NEI
OTAQREG
RADINFO
RBLC
RE-POWERING
REGION
RFS
SSTS
TSCA

Program Data Refreshed via Database Links

CAMDBS
E-GGRT
OIL
RMP
SFDW

FRS Data Stewards

179 Active Users:
- 17 Admin
- 17 Primary
- 80 Secondary
- 65 Read Only

Program Data Refreshed via Envirofacts Staging

AFS
ACRES
CERCLIS
RCRAInfo
RCRA BR
TRI

Reference data for Geocoding/Validation/ Spatial Indexing / Lookups

NAVTEQ Streets
USPS
NHD and WBD (RAD)
OAR Green Book NAA
SEGS Tribal Layer
Substance Registry

Mobile (Planned)

Pilots (TBD)

Reporting using Front End Integration

CEDRI
ESTAB E-Disclosure
FARR Tribal Facilities
TRI
TSCA CDR
RMP (Planned, 2016)
SSTS (Planned, 2016)
NPDES (Planned, TBD)
ICIS (Planned, TBD)

External Data

EPA Program Exports

EPA Program Data Connections

Facility Linkage Application

ICIS

FRS

Envirofacts

EPA Program Data

Air Compliance Facilities (ICIS-AIR)

State NPDES

ICIS Case Info

State Facility Databases (FacID)

EPA CDX Node (FacID Flow)

FRS Front-End Integration API

Mobile *(Planned)*

Ref. Data

Direct Reporting Interfaces

FRS Inbound Data Flows and Sources

7

FRS Outbound Data Flows

Reporting Interfaces using FRS Lookup Services: CEDRI, TRI, E-GGRT

Web Queries: Facilities + Enforcement Compliance

Geodatabases: ESF-10 Infrastructure

Applications Consuming FRS Data via Envirofacts: Envirofacts Reports, Cleanups In My Community, NEPAssist, MyEnvironment, Enviromapper, EJScreen, MyProperty + *more...*

Custom Extract: .csv

Future Exchange (TBD)

Metadata Cataloging

Custom Extracts: Geospatial formats

State Systems

HSIP

Chem EO Partners

Data.gov

INTERNET

Direct Reporting Interfaces

ECHO

Exchange Network FacID Flow

EPA Applications

Internet FRS EZ-Query

Chem EO Exchange (Planned)

EDG

Clip N Ship

GeoPlatform Mapping Applications

FRS Lookup API

ECHO

EPA CDX Node (FacID Flow)

Geodata Downloads

Envirofacts Public EnviroP

FRS Public Queries

Mobile *(Planned)*

Geo Metadata

Internet Geodata Servers

INTRANET

Custom Extract: .csv

Intranet FRS EZ-Query

Data Stewardship

Facility Linkage Application

FRS

ER Flex Viewers

Internal GeoPlatform Mapping

Planned Expansion of Mobile: R1 SPCC Inspection, START Contractors, National Evidence Inspection System

Web Queries

Envirofacts Intranet Queries

Envirofacts Intranet EnviroI

ICIS

FRS Intranet Queries

EPA Regions Programs

Sidewinder FTP Server

Intranet Geodata Servers

Mobile *(Planned)*

Bidirectional Exchange for Enforcement and Compliance

Database Links / Extracts

ER GIS Geodatabases

ArcGIS Server: igeo.epa.gov, orator.epa.gov ...

8

Integrating State data – FacID flow

- Created in early 2000s as a result of an effort similar to the Facility IPT
- Facility Identification (FacID) flow allows Partners to share their integrated facility/site data with FRS
- Using the Exchange Network, Partners can share information on facilities, sites, monitoring stations, and other place-based areas subject to environmental regulation or environmental interest
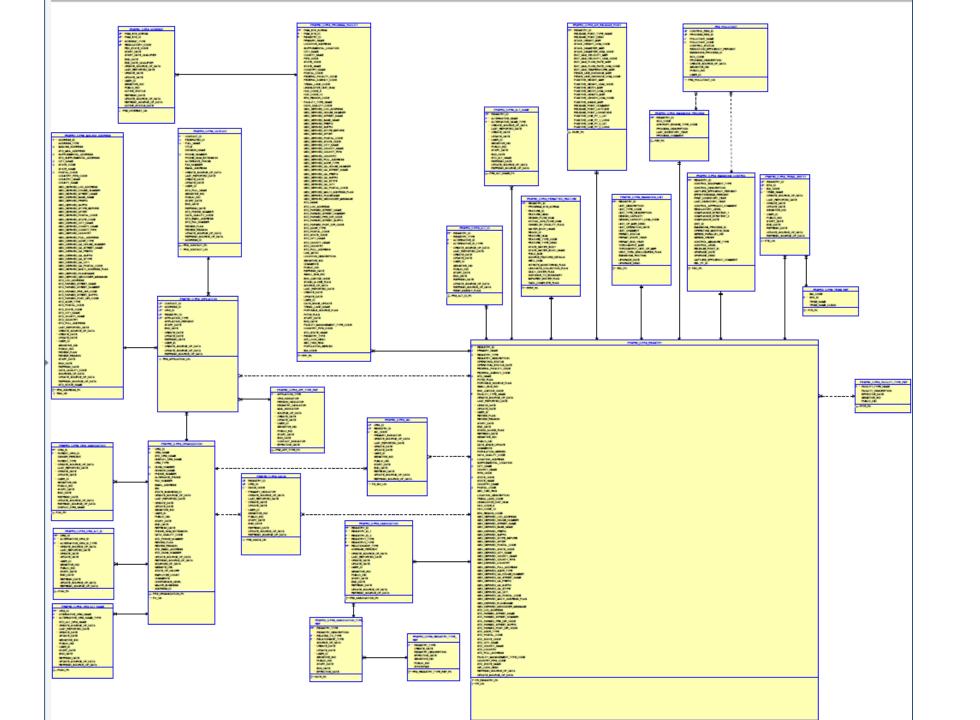
Integrating State data – FacID flow

State of current facility information in the Exchange Network.

Documentation for flowing facility data via the Exchange Network can be found at http://www.exchangenetwork.net/data-exchange/facility-identification

Drivers for New Facility Data Model

- FRS treats everything as a facility, whether facility, well, outfall pipe, or other feature of interest
  - Programs already report sub-facility features
  - FRS only does master data management for facility and geospatial coordinates not other features
  - Need a model that allows the capability to perform master data management on organizational and contact data for all features
- E-Enterprise Profile requirements necessitate a more nuanced FRS model to support more granular facility reporting and analysis needs being requested by EPA Programs (e.g., Combined Air Emissions Reporting, RTR)
- Need more robust history data for facilities and features
  - Provide facility configuration at specific time
- Improved capabilities between CDX and FRS

OAR-specific Facility Model

Program "Components"

Stores program-specific version of a "component"

Storing similar information as Core Attributes, but includes Unique ID for Program records

"Component" Attribute tables

- Emissons Release Point
- Emissions Process
- Emissions Pollutant
- Permitted Features
- Emissions Control
- Tribal Entity Attributes
- Emissions Unit

Affiliations

Stores the association of Organizations, Contacts, and "Components"

Organization
- Includes name, DUNS, Mailing Address

Contacts
- Name, Address, Phone Number, Source of data

Associations

Manage and store relationship of "components"

Core "Component" Attributes

Attributes that could apply to "Components"
- ID, Type, Name, Operating Status, Address, City, State, Zip, etc.
"Component" covers broad range of data
- Capacity to store complex facility data down to most granular levels

Flattened structure allows model to easily adjust to new relationship of "component" to "component"

Adding "components", program-by-program, only requires the addition of new attribute tables above

13

Facility Model Questions

Program "Components"

Does your program currently integrate data with FRS?

What is your version of a "component" and its attributes?

"Component" Attribute tables

How do you describe your program's "components"?

What is the overlap of core "component" attributes and program-specific attributes?

Can we see your database schema? Application interfaces?

Affiliations

Does your program have unique organizational data requirements?

Need for web services to provide organization data?

Associations

What are the relationships of your program's components?

Relationship of your program's components to other programs components?

Core "Component" Attributes

Are your program's core "component" needs met with this data schema?

Does your program have a data collection component that could benefit from Facility Widget?

To what degree will FRS services and applications need to be updated?

Benefits of the New Facility Data Model

- Management of sub-facility information
  - Users can track the relationship between stacks and pieces of equipment to facility
- Historical tracking for facility information
  - Configuration, facility name, ownership, contacts
- Ability to view a "facility profile" from multiple perspectives
  - Program-specific interests, "conceptual groups"
- Providing a context for regulatory requirements
  - Clarification on what the environmental interest is at a facility, program-by-program
- Capacity to better associate facility/components with organizations, contacts

New Facility Data Model – An Example

PROGRAM SYSTEM	INTEREST TYPE
BIENNIAL REPORTERS (BR)	HAZARDOUS WASTE BIENNIAL REPORTER
COMPLIANCE AND EMISSIONS DATA REPORTING INTERFACE (CEDRI)	COMPLIANCE AND EMISSIONS REPORTING *
ELECTRONIC GREENHOUSE GAS REPORTING TOOL (E-GGRT)	GREENHOUSE GAS REPORTER
EMISSION INVENTORY SYSTEM (EIS)	HAZARDOUS AND CRITERIA AIR POLLUTANT MAJOR *
FLORIDA - FIESTA DATA MAINTENANCE	STATE MASTER
ICIS-AIR (AIR)	AIR MAJOR *
INTEGRATED COMPLIANCE INFORMATION SYSTEM (ICIS)	FORMAL ENFORCEMENT ACTION
NATIONAL POLLUTANT DISCHARGE ELIMINATION SYSTEM (ICIS-NPDES)	ICIS-NPDES MAJOR
	STORM WATER INDUSTRIAL *
	ICIS-NPDES
	NON-MAJOR *
RACT/BACT/LAER CLEARINGHOUSE (RBLC)	AIR MAJOR *
RESOURCE CONSERVATION AND RECOVERY ACT INFORMATION SYSTEM (RCRAInfo)	SQG (Y)
RISK MANAGEMENT PLAN (RMP)	RMP REPORTER *
TOXIC RELEASE INVENTORY SYSTEM (TRI)	TRI REPORTER *
BOILER	BOILER
EMERGENCY RESPONSE INFORMATION SYSTEM (E-PLAN)	TIER 2 REPORTER
Occupational Safety and Health Administration (OSHA) Inspection System (OIS)	OSHA ESTABLISHMENT

*have associated enforcement/compliance activities

New Facility Data Model – An Example

27 Units from Emission Inventory System (EIS)

- ➢ 6 Boilers
- ➢ 1 Chemical Reactor
- ➢ 1 Kiln
- ➢ 1 Storage Unit

- ➢ 2 Incinerators
- ➢ 14 Unclassified
- ➢ 1 Reciprocating IC Engine
- ➢ 1 Distillation column/Stripper

9 units from U.S. Energy Information Administration (EIA-860)

- ➢ 4 Generators (STEAM TURBINE)
- ➢ 5 Boilers
  - 1 STOKER (SPREADER, VIBRATING GATE, SLINGER)
  - 2 WALL FIRED (OPPOSED, REAR, FRONT, AND SIDE WALL)
  - 2 OTHER

56 units from Residual Risk and Technology Review (RTR)

- 17 units (match FRS EIS Unit IDs)
- 39 units (listed as unit ID model)

New Facility Data Model – An Example

- <u>EIS</u>
  - **Facility Site** <Maintains> **Emissions Units** (27)
  - **Facility Site** <Releases Emissions Via> **Fugitive Release** (13)
  - **Facility Site** <Releases Emissions Via> **Stack** (15)
  - **Emissions Unit** <Is Operated By> **Emissions Process** (64)
  - **Emissions Process** <Vents Emission Thru> **Fugitive Release** (21)
  - **Emissions Process** <Vents Emission Thru> **Stack** (43)
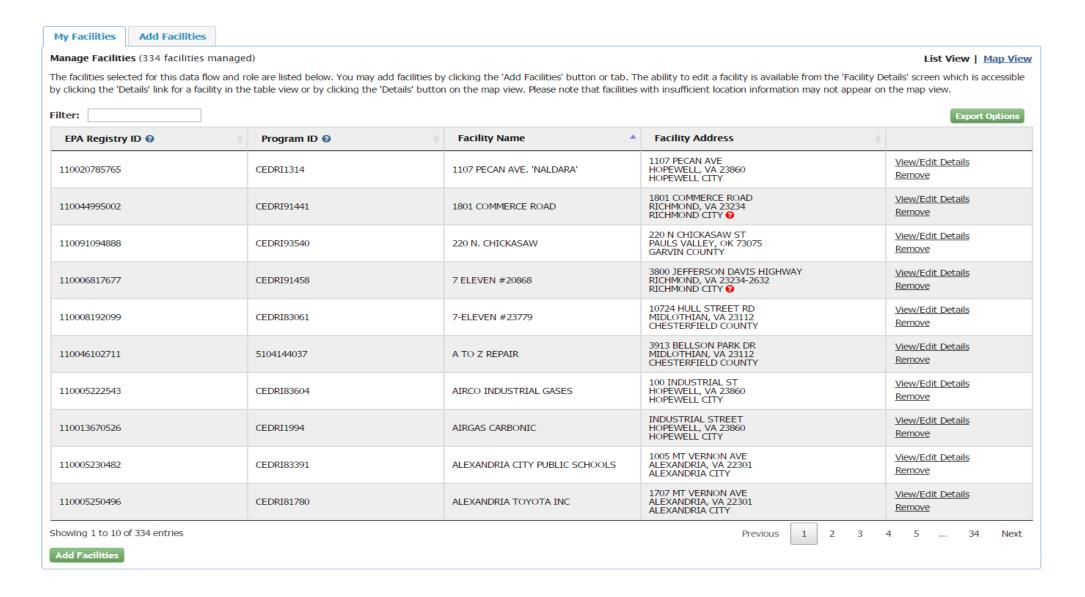- <u>EIA-860</u>
  - **Facility Site** <Maintains> **Emissions Unit** (9)
  - **Emissions Unit** <Is Associated with> **Emissions Unit** (12)
  - **Emissions Unit** <Has Particulates Controlled By> **Control Method** (2)

Residual Risk and Technology Review (RTR)

- CAER/OAR leveraging the new FRS data model for collecting "component" data from direct reporters

- Matching between Residual Risk and Technology Review (RTR) and NEI has been a problem in the past

- RTR facility data collection
  - Past - Via EPA ICRs, used spreadsheets and Microsoft® Access® files
  - Future -
    - Use EPA portal/CDX and facility widget (and beyond)
    - Same approach as CEDRI so test data can be provided linked to same facility attributes

- Through CAER, linking and use of FRS component data by other air programs will allow for improved matching and use of RTR data

RTR – 'My Facilities' List

My Facilities | **Add Facilities**

Manage Facilities (334 facilities managed)

List View | Map View

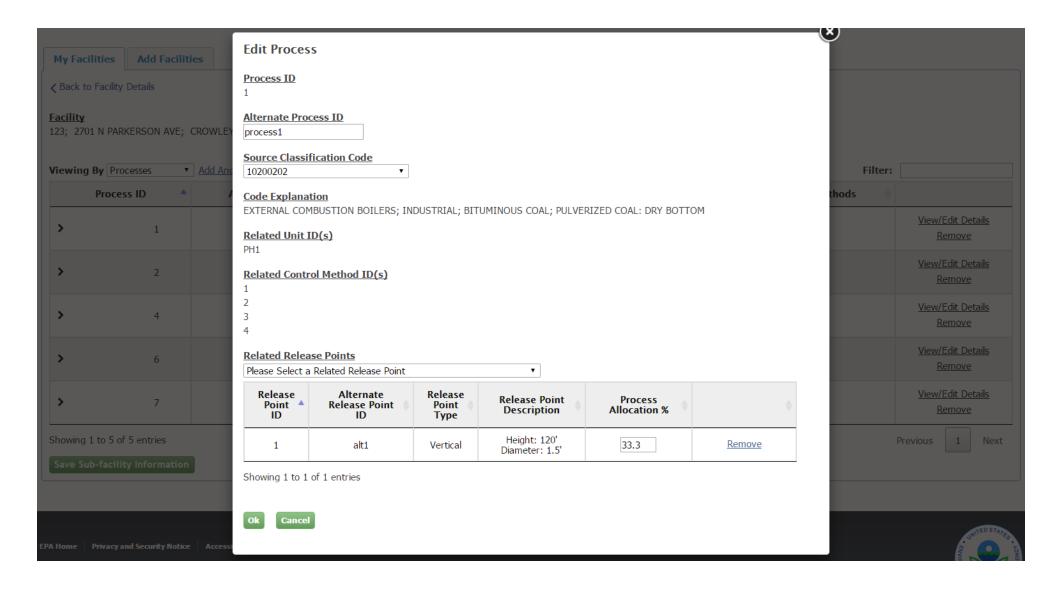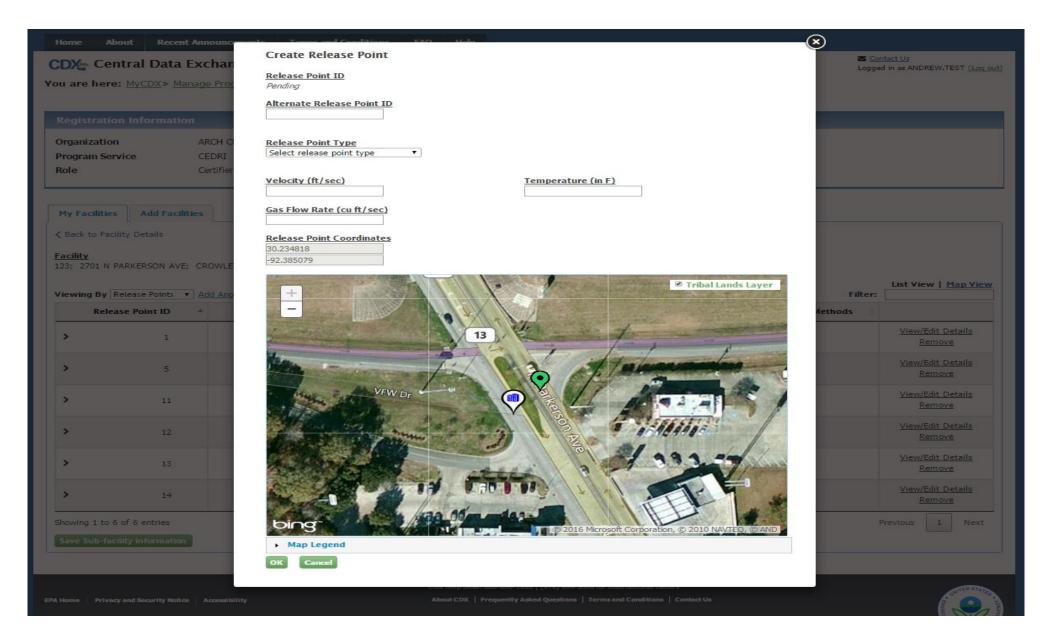The facilities selected for this data flow and role are listed below. You may add facilities by clicking the 'Add Facilities' button or tab. The ability to edit a facility is available from the 'Facility Details' screen which is accessible by clicking the 'Details' link for a facility in the table view or by clicking the 'Details' button on the map view. Please note that facilities with insufficient location information may not appear on the map view.

Filter: [   ]

Export Options

EPA Registry ID	Program ID	Facility Name	Facility Address	
110020785765	CEDRI1314	1107 PECAN AVE. 'NALDARA'	1107 PECAN AVE HOPEWELL, VA 23860 HOPEWELL CITY	View/Edit Details Remove
110044995002	CEDRI91441	1801 COMMERCE ROAD	1801 COMMERCE ROAD RICHMOND, VA 23234 RICHMOND CITY ❓	View/Edit Details Remove
110091094888	CEDRI93540	220 N. CHICKASAW	220 N CHICKASAW ST PAULS VALLEY, OK 73075 GARVIN COUNTY	View/Edit Details Remove
110006817677	CEDRI91458	7 ELEVEN #20868	3800 JEFFERSON DAVIS HIGHWAY RICHMOND, VA 23234-2632 RICHMOND CITY ❓	View/Edit Details Remove
110008192099	CEDRI83061	7-ELEVEN #23779	10724 HULL STREET RD MIDLOTHIAN, VA 23112 CHESTERFIELD COUNTY	View/Edit Details Remove
110046102711	5104144037	A TO Z REPAIR	3913 BELLSON PARK DR MIDLOTHIAN, VA 23112 CHESTERFIELD COUNTY	View/Edit Details Remove
110005222543	CEDRI83604	AIRCO INDUSTRIAL GASES	100 INDUSTRIAL ST HOPEWELL, VA 23860 HOPEWELL CITY	View/Edit Details Remove
110013670526	CEDRI1994	AIRGAS CARBONIC	INDUSTRIAL STREET HOPEWELL, VA 23860 HOPEWELL CITY	View/Edit Details Remove
110005230482	CEDRI83391	ALEXANDRIA CITY PUBLIC SCHOOLS	1005 MT VERNON AVE ALEXANDRIA, VA 22301 ALEXANDRIA CITY	View/Edit Details Remove
110005250496	CEDRI81780	ALEXANDRIA TOYOTA INC	1707 MT VERNON AVE ALEXANDRIA, VA 22301 ALEXANDRIA CITY	View/Edit Details Remove

Showing 1 to 10 of 334 entries

Previous | 1 | 2 | 3 | 4 | 5 | ... | 34 | Next
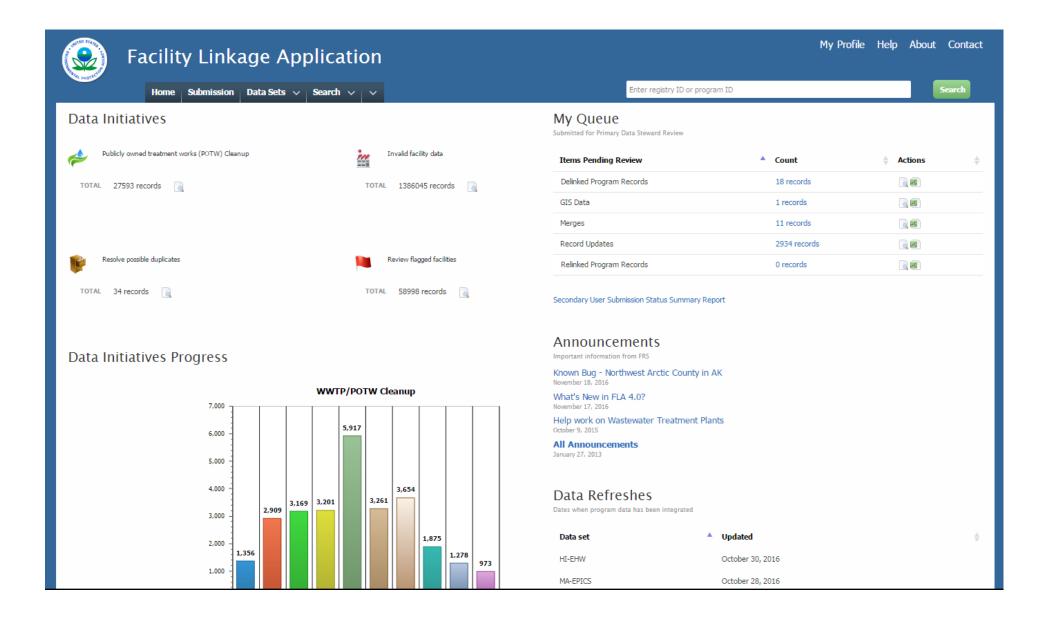
Add Facilities

RTR – Facility Details

RTR – Process details

RTR – Release Point details
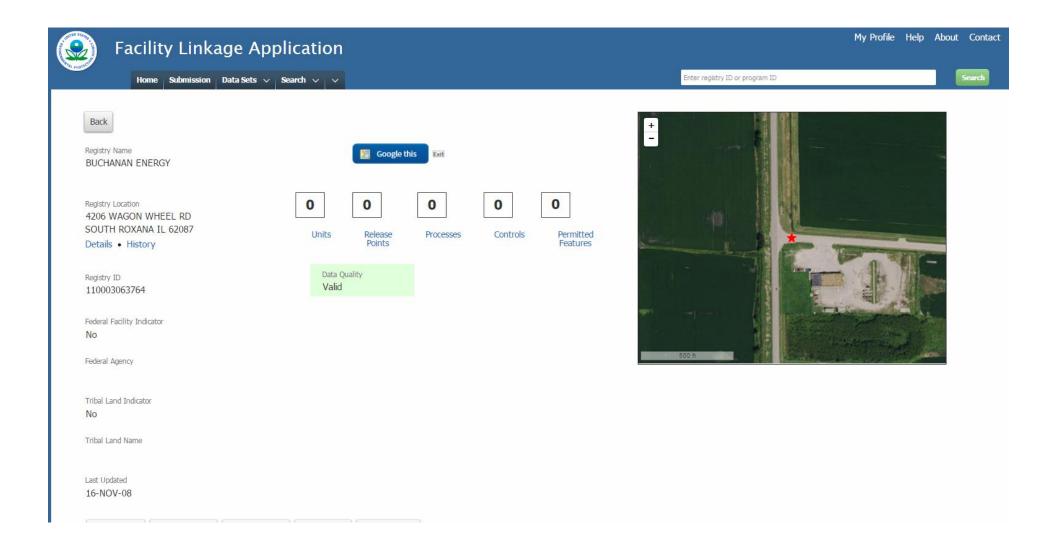
FLA – Home screen

Facility Linkage Application (FLA)

- Application used to correct facility data in FRS
- Influential in how data is viewed through FRS's data products and Agency application such as Envirofacts
- Fill in data gaps and discover data errors by linking programs records
- Correct locations of facilities to aid emergency responders, inspectors, rule makers, data modelers and others
- Reduce duplicates to provide a cleaner, easier to use dataset
- Network of about 160 data stewards making data quality improvements
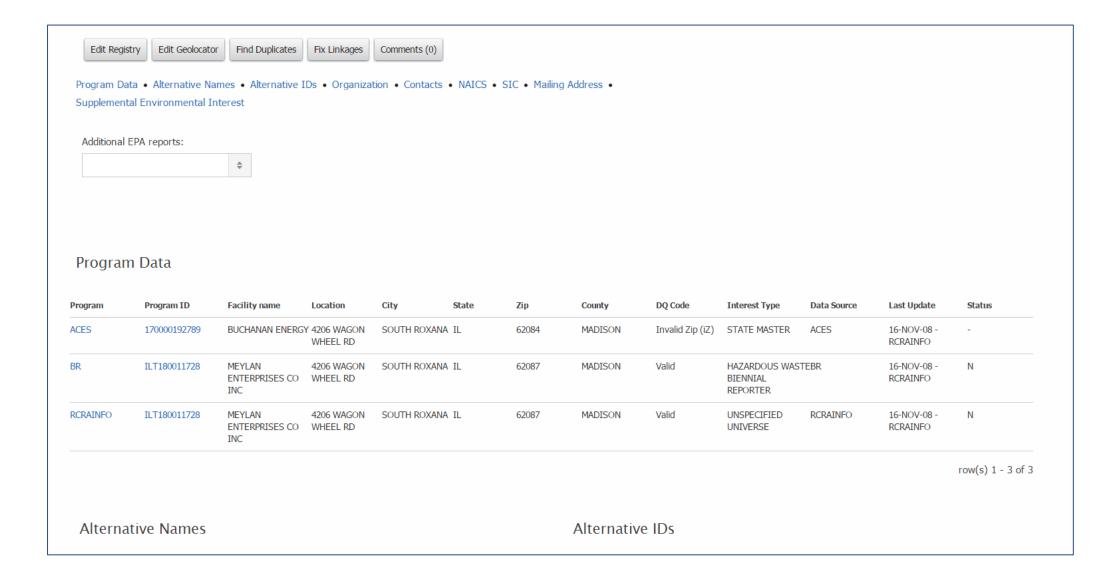 – Regional, Program, and State stewards

Example FRS Data Quality actions

- <u>Updating</u> – making changes to Facility Name, Address, Location
- <u>Merging</u> – reconciling duplicate master records in FRS. Choosing one record to serve as "master record", other becomes "program record"
- <u>Delinking</u> – reconciling program records that were erroneously linked with a "master record"
- <u>Finding duplicates</u> – search to find potential duplicate records
- <u>Edit geolocator</u> – method for changing latitude and longitude for a record. Also includes method and accuracy descriptions

FLA – Facility details screen

FLA – Facility details screen

Edit Registry Edit Geolocator Find Duplicates Fix Linkages Comments (0)

Program Data • Alternative Names • Alternative IDs • Organization • Contacts • NAICS • SIC • Mailing Address •
Supplemental Environmental Interest

Additional EPA reports:

Program Data

Program	Program ID	Facility name	Location	City	State	Zip	County	DQ Code	Interest Type	Data Source	Last Update	Status
ACES	170000192789	BUCHANAN ENERGY 4206 WAGON WHEEL RD		SOUTH ROXANA	IL	62084	MADISON	Invalid Zip (iZ)	STATE MASTER	ACES	16-NOV-08 - RCRAINFO	-
BR	ILT180011728	MEYLAN ENTERPRISES CO INC	4206 WAGON WHEEL RD	SOUTH ROXANA	IL	62087	MADISON	Valid	HAZARDOUS WASTEBR BIENNIAL REPORTER		16-NOV-08 - RCRAINFO	N
RCRAINFO	ILT180011728	MEYLAN ENTERPRISES CO INC	4206 WAGON WHEEL RD	SOUTH ROXANA	IL	62087	MADISON	Valid	UNSPECIFIED UNIVERSE	RCRAINFO	16-NOV-08 - RCRAINFO	N

row(s) 1 - 3 of 3

Alternative Names   Alternative IDs

FLA – Facility details screen

Alternative Names

Alternative Name	Data Source
BUCHANAN ENERGY	ACES
STAHLY CARTAGE CO	NOTIFICATION (RCRA)
STAHLY CARTAGE CO	RCRAINFO

row(s) 1 - 3 of 5   Next >

Alternative IDs

No Alternative IDs found.

Organization

Organization	Roles
MEYLAN ENTERPRISES	OPERATOR OWNER

row(s) 1 - 1 of 1

Contacts

No Contacts found.

NAICS

NAICS Code	Indicator	Description
339999	PRIMARY	ALL OTHER MISCELLANEOUS MANUFACTURING.

row(s) 1 - 1 of 1

SIC

No SIC found.

Looking ahead

- **E-Enterprise Portal**
  - Currently leveraging Facility look-up widget
  - Goal: to house facility, contact, organization "profiles"
- **Facility IPT with States**
  - Support workstreams, pilot potential facility data solutions with RI
  - CAER Gap Analysis to identify shortcomings in FRS data model, services
  - Opportunities to better leverage State data and enable states to leverage EPA data and web services
- **Combined Air Emissions (CAER) Workgroup**
  - "Shared Facility Attributes" database – future state of OAR data needs
  - Residual Risk Technology Review (RTR) project
    - Planning CDX flow, leveraging Facility look-up widget
- **Expanding data model to meet EPA Program, SLT needs**

Looking ahead – EN 2017 Presentations

- **Tuesday, May 16**
  - Session 1: Combined Air Emissions Reporting (Ballroom B)
- **Wednesday, May 17**
  - Session 7: Facility Data – Facility IPT (Ballroom B)
    - Facility Integration – Challenges and Lessons Learned from One State's Perspective (Oklahoma DEQ)
    - E-Enterprise Facility Phase 2 Effort