COLLECTION AND DOCUMENTATION OF GENERAL DESCRIPTIVE GEOSPATIAL SITE DATA

Version 3.5

OFFICE OF SUPERFUND REMEDIATION AND TECHNOLOGY INNOVATION, AND
OFFICE OF RESOURCE CONSERVATION AND RECOVERY, AND
FEDERAL FACILITIES RESTORATION AND REUSE OFFICE

of the

OFFICE OF LAND AND EMERGENCY MANAGEMENT
UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

February 2019

Document Revision History

Date	Author	Version	Description
04/15/2016	Michael Alford	1.0	First Draft
05/27/2016	Michael Alford	1.1	Final Draft
06/14/2016	Michael Alford	1.2	Integrate Comments
08/17/2016	Michael Alford	1.3	Integrate EGAC Comment
09/12/2016	Michael Alford	1.4	Integrate Enforcement & R3 comments
03/16/2017	Michael Alford	1.5	Added Institutional and Engineering Controls as optional attributes of site features. Added comparison to RCRA CA in appendix copy of complete schema.
03/28/2017	Michael Alford	2.0	Created combined Superfund/RCRA document
4/25/2017	Michael Alford	2.1	Integrated OSRTI changes and modified data schema to accommodate Long Term Stewardship issues
5/25/2017	Michael Alford	2.2	Completed last edits for general review
6/12/17	Michael Alford	2.21	Added Lisa Messinger (R7) comments to guidance and John Wieber (R8) corrections to data schema field lengths
7/26/17	Michael Alford	2.22	Added discussion of R7 Parts. Moved Institutional and Engineering Control references from the Point, Line and Boundary tables to the Feature Identifier section of the master table.
8/7/17	Michael Alford	2.23	Added Appendix on Parts.
12/1/17	Michael Alford	3.0	Integrated R8 Data Dictionary with the combined Superfund/RCRA Data Schema
12/15/17	Michael Alford	3.1	Revised data schema descriptive diagrams and descriptive text for schema overview. Added cover and updated TOC.
01/30/18	Michael Alford	3.2	Aligned narratives with OLEM Directive 9200.2-191 Revised descriptions of collection methods to include regional input on multiscale online map interpolation and illustrations
2/16/18	Michael Alford	3.3	Incorporated Jennifer Sutton comments. Removed appendices and converted them to standalone Technical Supplement 1.
3/15/18	Michael Alford	3.3	Same version number with clarifications of Jennifer Sutton comments on institutional controls and records schedules. Modified schema to avoid using AREA as a field name, as this conflicts with ESRI reserved field names. Modified schema Block 5 to clarify equivalent importance of Controls in Superfund and RCRA CA.
4/3/18	Lisa Jenkins	3.3	Some clarification of scope, inclusion of federal facilities, addition of RCRA records schedules and web services, and final product manager review.
5/10/18	Michael Alford	3.4	Added Glossary of Acronyms

Date	Author	Version	Description
2/4/19	Mary Mazerik	3.5	Incorporated feedback, updated field types and provided additional acronyms.

Table of Contents

1 INTRODUCTION

1.1 Purpose

The purpose of this guidance is to provide a uniform method for collecting, documenting and managing geospatial information for sites subject to cleanup across all EPA Regional Offices under the following EPA's statutory authorities: It currently supports cleanups under Superfund (the Comprehensive Environmental Response, Compensation and Liability Act, CERCLA) and the Resource Conservation and Recovery Act (RCRA) Corrective Action (CA) program, including Federal Facilities under both programs. It defines mandatory and optional elements for documenting cleanup site locations, boundaries and internal features. It provides for the transfer of this information to the authoritative Superfund and RCRA CA headquarters information systems for distribution to the public and other users, and for the documentation of this information in the Environmental Dataset Gateway (EDG).

This guidance is provided for all EPA organizations, personnel or agents (including contractors and grantees) of EPA who design, develop (directly or indirectly), compile, operate or maintain general geospatial descriptive information on Superfund and RCRA CA sites and their internal features. It does not cover geospatial information developed to document site sampling programs and associated laboratory quality assurance programs. It assumes that users are familiar with basic Superfund or RCRA CA terminology, procedures and program operations.

The guidance does not dictate which geographic components of which sites are to be collected. That policy will be determined and issued by program and/or region. This guidance provides national standards for collecting the data and formats for sharing the data so as to simplify the creation of local collection practices and enable the collecting of comparable data.

This guidance will be updated from time to time as appropriate.

1.2 Governing Policies and Standards

This guidance implements the objectives of the National Geospatial Data Policy (NGDP, 2005)[1] and Latitude/Longitude Data Standard (2006)[2] consistent with Environmental Protection Agency (EPA) and Federal requirements, including the EPA Enterprise Information Management Policy (EIMP, 2015)[3], Open Data Executive Order 13642 (2013)[4], OLEM Directive 9200.2-191[5] and OMB Memorandum M-13-13 (2013)[6]. For Superfund, it addresses data collection on proposed, final and deleted National Priorities

[1] EPA. 2005. National Geospatial Data Policy. CIO Policy Transmittal 05-002. Classification No.: 2121. Available online at: https://www.epa.gov/sites/production/files/2014-08/documents/national_geospatial_data_policy_0.pdf.

[2] EPA. 2006. Latitude/Longitude Data Standard. Standard No.: EX000017.2. Available online at: https://iaspub.epa.gov/sor_internet/registry/datastds/findadatastandard/epaapproved/latitudelongitude/LatLongStandard-v2a_10022014.pdf.

3 EPA. 2015. Enterprise Information Management Policy (EIMP) Cataloguing Information Procedure. Available online at: https://www.epa.gov/sites/production/files/2015-05/documents/2135-p-01_0_eimp_catal_epa_info_proc_final_2015_03_25.pdf.

[4] Executive Office of the President. 2013. Executive Order 13642, Making Open and Machine Readable the New Default for Government Information. Available online at: https://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-.

[5] OLEM Directive 9200.2-191, *Geospatial Superfund Site Data Definitions and Recommended Practices,* Nov. 29, 2017. Available at: https://semspub.epa.gov/src/document/HQ/100000739.

[6] Executive Office of the President. 2013. Memorandum M-13-13, Open Data Policy-Managing Information as an Asset. Available online at: https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf.

List (NPL) and Superfund Alternative Agreement sites. For RCRA, it addresses all current and future Corrective Action sites. It addresses Federal Facilities Docket and other Federal Facilities sites. Collection of geospatial data consistent with this guidance promotes EPA's ability to share information with the public and to perform analyses to support program implementation and evaluation in a consistent fashion.

2 LIFE CYCLE OF CLEANUP SITE GEOSPATIAL DATA

The EIMP requires that EPA information be managed as a strategic asset, consistent with Executive and Legislative acts, orders, directives, strategies and guidance. It further specifies that information be planned and managed according to a defined information life cycle process that is appropriate to the information type and in accordance with enterprise systems and solutions.

OLEM therefore establishes the following life cycle workflow for the collection, management and use of cleanup site geospatial information, consistent with regional and headquarters information management systems and practices.

The following sections define the Target OLEM geospatial life cycle, which is not yet operationalized. In particular, processing, storage and synchronization of data between the regional offices and headquarters systems will take some time to complete. Interim data flows and data publishing guidance will be provided as Annexes to this guidance as steps toward the Target are available. Where possible, the sections below describe general features of the expected interim steps toward the Target.

2.1 Data Collection or Acquisition

Descriptive geospatial site[7] information is collected by Superfund site contractors, RCRA Corrective Action facility owners and their contractors, federal facilities, state hazardous waste agencies and/or EPA regional staff. It falls into three main categories:

- **Site Points** identifying the general location of the site or the location of specific internal point features, such as wells, sampling points and site geometry elements (such as site centroids or bounding box coordinates).
- **Site Lines** identifying such features as roads, fences, rights of way, utility corridors, streams, pipes or pipelines.
- **Site Boundaries or Perimeters** identifying the geographic extent of the site as a whole, internal sub-elements such as Superfund Operable Units or RCRA Solid Waste Management Units (SWMUs)/Areas of Contamination (AOCs), institutional or engineering controls, or the extents of areas of contamination.

Site geospatial data evolve and may be refined during a site's lifecycle. These data generally should be collected and maintained throughout the cleanup process. As site data are refined, updated data should be collected as they become available. Typically, geospatial data are collected throughout the lifecycle of a site.

2.1.1 General Site Boundaries

Superfund Primary Boundaries: In the early stages of a Superfund site's management, its overall geospatial extent may not be known or formally defined. While a single point may be the only available geospatial information early in the Superfund process, detailed site maps or geospatial data files, depicting polygons for a site's perimeter,[8] area of contamination and operable units (OUs), generally

[7] In practice, sites do not have to be contiguous areas of land: they may be divided by streams, have areas within the site perimeter that are not include in the site, or be composed of multiple separated contaminated areas considered as one site for program purposes. Examples of a variety of site types are provided in Addendum 3.
[8] If the Operable Units are discontiguous, the Site Boundary typically consist of multiple independent polygons.

should all be available by the time EPA issues a Record of Decision (ROD) for a remedial action or signs an Action Memorandum (AM) for a non-Time-Critical Removal action.

RCRA Corrective Action Boundaries: Per 40 CFR 260.10, a RCRA facility is defined as all contiguous land and structures, other appurtenances and improvements on the land used for treating, storing or disposing of hazardous waste. A facility may consist of several treatment, storage or disposal operational units (e.g., one or more landfills, surface impoundments or combinations). For the purpose of implementing corrective action under Sec. 264.101, all contiguous property under the control of the owner or operator seeking a permit under subtitle C of RCRA may also define a facility. This definition also applies to facilities implementing corrective action under RCRA Section 3008(h). Any non-contiguous properties, with the same owner/operator, would still be identified as separate facilities with differing EPA Handler IDs in RCRAInfo.

For the purposes of this document, the boundary encompassing a facility is referred to as the Entire Facility. Practice may vary on the degree of detail recorded in the Entire Facility polygon. The Entire Facility boundary may, for example, have to include discontiguous polygons, such as where one or more parcels are separated by public roads. The boundary may also, as part of the graphic image, include the boundaries of internal parcels that together make up the Entire Facility.[9]

However, EPA has authority to require corrective action to be taken for releases beyond the facility boundary where necessary to protect human health and the environment. In these cases, the RCRA Corrective Action boundary may be different from the Entire Facility boundary and may be delineated by the area of contamination (AOC) included under a general RCRA facility boundary. There are times when the Entire Facility may not be the boundary defined in a Hazardous Waste Management Permit or under post-closure care, both of which may have a corrective action component. In these cases, the general RCRA Corrective Action boundary may also differ from the Entire Facility boundary and should be revised to incorporate the additional property where appropriate.

2.1.2 Additional Site Features

In addition to the general site boundary, regions may add additional point, line or boundary features to a site's record to enhance the public understanding of the site's environmental issues, institutional controls and engineering solutions. Such enhancements are particularly valuable for larger sites where the site boundary may encompass resident populations, dispersed areas of contamination and/or multiple distributed institutional controls or engineering solutions. They are also desirable where a site involves a so-called "antecedent" boundary – a larger area such as a military base within which the Superfund site, RCRA TSD or corrective action units are located.

Regional Offices should update a site's geospatial records over time to reflect such changes as the delisting of site sub-areas, the redrawing of internal features or the imposition of institutional or engineering controls.

[9] Including the boundaries of all parcels that together make up the Entire Facility, whether contiguous or discontiguous, is standard practice in Region 7. Region 7 refers to parcel boundaries as Parts, with each Part identified by its Part ID number. For the purposes of this guidance, therefore, Parts are considered to be an optional Boundary type. They are useful (1) to document the internal parcel components of the Entire Facility boundary and (2) to document particular parcels that are included within an Institutional or Engineering Control. See sections 4.5.2, 4.5.3 and Addendum 4.

2.2 Processing and Storage at Regional Offices and at Headquarters

All site cleanup geospatial information is compiled at the Regional Office for quality assurance and review prior to being synchronized with the appropriate authoritative headquarters program information database.

- **Superfund:** Because Superfund is a federal program, geospatial information is collected either by regional staff or by private contractors operating under regional direction. All site geospatial data is therefore under the direct management of the program. The authoritative system for this information is the Superfund Enterprise Management System (SEMS). Federal Facility data is also stored in SEMS, and the storage and flow of federal facilities data will generally follow Superfund data.
- **RCRA Corrective Action:** Because hazardous waste management is a delegated program in some states, geospatial information may be collected by the state, by the facility or by regional staff. Because of this variability, the scope and extent of RCRA CA geospatial data will vary from region to region and be collected at the discretion of regional management. The authoritative system for this information is Resource Conservation and Recovery Information (RCRAInfo).

Despite programmatic differences on where site data is initially collected, all geospatial data is to be compiled at the Regional Office using that region's geospatial information management system common data schema: the OLEM Cleanup Site Geospatial Data Schema (CSGDS). It will then be aggregated to create two comprehensive, national site cleanup data sets—one for Superfund and one for RCRA Corrective Action—with parallel, compatible data structures.[10]

General Geospatial Data Flow: The intended general target data flow, described below, is the basis for the structure, design and terminology of the CSGDS.

The goal is to keep regional and headquarters geodatabases up to date and consistent with each other using the following general workflow.

- All site geospatial information, including narrative descriptions and required metadata, are originated in the regions and synchronized with headquarters.
- All other site attribute data for which the headquarters is the authoritative source, such as site/facility program status indicators, the site/facility address of record and personal contact information, are synchronized from the headquarters system to the regional geodatabase.

The minimum set of site geospatial data elements, metadata and site attributes to be maintained for export from SEMS and RCRAInfo using the CSGDS is discussed in Section 3.0 below. These include all fields synchronized between SEMS and RCRAInfo and the regional geodatabases.

Under the target data flow, Regional Offices will aggregate all descriptive site records for their regions into a single ESRI Geodatabase that is accessible for synchronization with SEMS and RCRAInfo and maintained in Geographic Coordinate System Decimal Degrees, WGS 84 Datum.

Extensibility of the Cleanup Site Geospatial Data Schema: Regions are free to extend their site and facility information and regional geodatabase to include region-specific fields, but this information will not be captured by the headquarters systems and therefore will not be automatically disseminated to

[10] Creating aggregated data sets is the goal, but modification of SEMS and RCRAInfo to accomplish this automatically is not yet scheduled. In the interim, options include compiling the data sets manually and posting them to existing repositories or visualizing multiple regional web services on platforms such as CIMC.

the public and other users. For public enquiries concerning site geospatial attributes not maintained at headquarters, OLEM will refer questions to the appropriate region.

Interim Geospatial Data Flows: Until SEMS and RCRAInfo are modified to accept geospatial data in the CSDGS, all data will be maintained in the regions and transferred manually to Headquarters for compilation and distribution in whatever format is deemed most practical.

- For Superfund sites, the SEMS team has developed web services[11] to allow joining of attributes and data synchronization between the regional geodatabases and SEMS. The have also developed, for the regional geodatabases, data to be harvested and merged into national data sets residing on the Shared Enterprise Geodata and Services (SEGS) servers. Federal Facilities data can also be processed in this way.
- For RCRA Corrective Action Facilities, similar processes will be developed and hosted initially by Cleanups in My Community (CIMC) from data pulled from the RCRAInfo hosted flat files. Site attributes from the CIMC RCRA web service can then be joined with regional geodatabases. A process will also be developed for the regional geodatabases data to be harvested and merged into national data sets residing on the Shared Enterprise Geodata and Services (SEGS) servers.

The SEGS national geospatial web services will reside on an EPA enterprise geospatial framework that can be used by other EPA data systems and applications, such as the EPA GeoPlatform Online (GPO), the Environmental Dataset Gateway (EDG, see below), as well as the Facility Registry Service (FRS) and Cleanups in My Community (CIMC).

Metadata Records for the Environmental Dataset Gateway: The EDG is a web-based metadata (data about data) portal that supports the discovery of and access to EPA's environmental dataset resources. The EDG contains metadata records contributed by EPA offices and links to geospatial and non-geospatial resources (e.g., data, services, or applications) described by those metadata records. Unrestricted information that is contributed to the EDG is shared with interagency data sharing portals, including Data.gov and Geo.Data.gov.

Owners of all datasets created or owned by EPA must document these datasets in EDG using a standard template. The template used for geospatial datasets follows the Federal Geographic Data Committee (FGDC) and International Standards Organization (ISO) format.

Two types of EDG records are potentially required for the Superfund Site and RCRA Corrective Action Facility geospatial data sets that are the subject of this guidance:

- Regional layer datasets for both Superfund and RCRA
- Headquarters distributed national data sets aggregating all regional layer datasets into one geodatabase.

Each regional office that assembles a complete or partial data set of Superfund Site or RCRA Corrective Action geospatial data must complete a metadata record in EDG.

When a national aggregated data set is compiled from these regional layers, the owner of that compilation must complete a separate EDG metadata record.

Note that:

[11] Region 8 has developed web services and is sharing a python script that associates attributes in SEMS with an existing site record in the regional geodatabase.

- The EDG record supplies metadata for the dataset as a whole, not for the individual sites or facilities contained in the dataset
- Appropriate metadata for site- or facility-level records within the dataset is contained in the individual record. This is found in the CSGDS itself and is compliant with the EPA National Geospatial Data Policy and the EPA Enterprise Information Management Policy.

2.3 Dissemination and Use

Web services will be the ultimate method of distribution for these datasets, either via the SEGS web services developed by EPA or, eventually, via web services developed to extract directly from the authoritative systems SEMS and RCRAInfo. Data feeds will be limited to fields included in the CSGDS. Web services will be limited to the fields marked as Export (color coded green in this guidance).

Systems that use geospatial information for site cleanup currently include, but are not limited to:

- Geospatial Data
  - EPA GeoPlatform
  - Cleanups in My Community (CIMC)
  - EPA Facility Registry Services (FRS)
  - Envirofacts
- Metadata Records[12]
  - EPA Environmental Dataset Gateway (EDG)
  - Data.gov (indirectly via EDG)

2.4 Storage and Disposition

The CSGDS provides for versioning of all features that will be maintained in, and exported as web services by, the authoritative Headquarters program information management systems: SEMS and RCRAInfo. Features may be updated an unlimited number of times, with the effective date of each change carried with the feature on export. Ultimately any and all chronological versions of each site feature will be maintained within SEMS or RCRAInfo.

In the interim, regions maintain their own geodatabases and implement their own processes to preserve archive records of changed features, such as using the SEGS script to identify records that change when creating national feature classes and pushing old records to a separate geodatabase to archive historical records. Until SEMS and RCRAInfo are updated to implement the versioning of features and other fields, there will be no standard method for retaining historical geospatial information.

Regional databases will not maintain multiple measurements of the same point, line or polygon as archival records. The regional geodatabase will retain only those measurements determined by the Regional Offices to be authoritative.

[12] EDG geospatial metadata templates are based on the Federal Geographic Data Committee (FGDC) *Content Standard for Digital Geospatial Metadata*. This standard will be replaced by the North American Profile of ISO 19115/19139 (NAP), which is currently in development.

The periodic update of datasets provided to the systems which use the geospatial information (listed above) will include only the most recent feature versions. Thus, for any given geospatial feature at any given date, OLEM will maintain and distribute only one geospatial measurement.

All geospatial information incorporated into SEMS or RCRAInfo is retained per the appropriate records schedule.

Per Draft EPA Records Schedule 1036, Superfund sites are likely to fall under either:

- **Item A:** *Historically significant site-specific records (See Guidance for explanation.)* Includes Superfund, Oil Spills, Tribal land (where EPA is the lead), Formerly Used Defense Sites on the National Priorities List (NPL), and Federal Facilities.

  NARA Disposal Authority: Pending
  - **Permanent**
    - Close remedial site-specific records when deleted from the NPL. Close landmark cases when case is completed. Close all others when archived (See Guidance for explanation of archived status.) in the Superfund Enterprise Management System (SEMS).
    - Transfer to the National Archives 30 years after file closure.

- **Item C:** Long-term site-specific records.

  NARA Disposal Authority: Pending
  - **Disposable**
    - Close when activity, project, or topic is completed.
    - Destroy 30 years after file closure.

Note that no data may be deleted until this schedule becomes final, and no records may be deleted until at least 30 years after the file is closed, defined as "when activity, project, or topic is completed." Furthermore, retention requirements for geospatial records will be the same as for any other records pertaining to that site or facility. Geospatial records pose no different or additional retention requirements than any other form of electronic records, so decisions regarding their disposition will be handled as part of any and all decisions affecting the file as a whole.

RCRA data records are likely to fall under one of the following records schedules:

- National RCRA Information System (RCRAInfo): http://intranet.epa.gov/records/schedule/final/0257.html
- Compliance and Enforcement (specifically RCRA Corrective Action files): http://intranet.epa.gov/records/schedule/final/1044.html
- Information and Technology Management (input – part e): http://intranet.epa.gov/records/schedule/final/1012.html

3 OVERVIEW OF THE CLEANUP SITE GEOSPATIAL DATA SCHEMA

Figure 1 provides a general structural view of the Cleanup Site Geospatial Data Schema. Details are covered in Section 4.0: Data Coding Guidance.

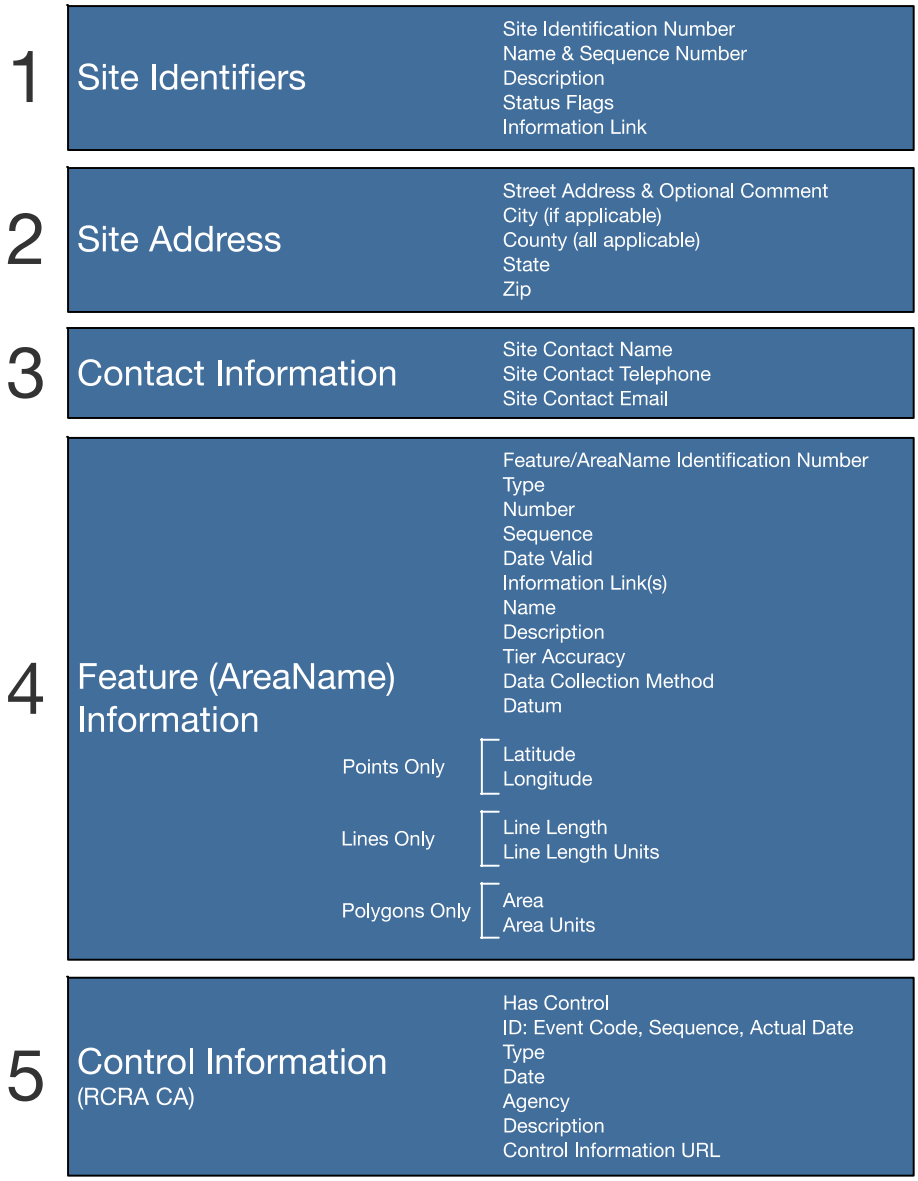


Figure 1: Structure of the Cleanup Site Geospatial Data Schema (CSDGS)

Site Identifiers (1): This data group includes the EPA Site/Hander Identification number; EPA Site/Hander name (and sequence number, if the name has changed); site/facility description; codes indicating the site's current NPL status and whether it is a federal facility and a URL link to additional site information. The EPA Site/Hander ID information is supplied by the region; all other information is supplied by the headquarters database.

Site Address of Record (2): This data group is supplied by the headquarters database. An address comment may be added by the region to clarify a non-standard address.

Site Contact (3): Headquarters databases will provide the name, phone and email of the most current contact person for a given site appropriate for public contacts. Where multiple contacts exist, the contact used for Superfund will typically be the Remedial Project Manager (RPM); for RCRA CA, it will typically be the facility's appointed Public Point of Contact (POC).

Feature Identifiers (4): This data block includes the type, number and sequence number of the feature described, as well as the geometry itself and all feature descriptions.

- Data Accuracy:
  - If the entry in question is for a site identifier of record extracted from either SEMS or RCRAInfo, then the user will have to provide the projected binary file to supplement the text coordinates received from the headquarters system. The recorded Tier Accuracy is derived from the collection method used and should be equal to or better than the Tier Accuracy required by the type of feature documented (see Section 4.3.1).
  - Where regions are using historical sources to digitize geospatial information, collection methods may either be unknown, or be known not to meet current accuracy targets. In such cases, they will simply provide the known or supposed historical collection method. On export, it will be clear to the user that the value in question—although the best available—is not up to current accuracy standards.
- The Datum[13] shows the projection or coordinate system in which the data are currently stored, which should be WGS84; it is either derived from the collection method or entered manually.
- POINTS: If the feature is a point, then the Latitude and Longitude are input in decimal degrees as a convenience to users.
- LINES: If the feature is a line, then the length of the feature and the unit of measure (e.g., feet, meters) is generated from the geometry for output to users.
- POLYGON: If the feature is a polygon, then the area and the unit of measure is generated from the geometry for output to users.

Controls (5):

Control information is collected for both Superfund and RCRA Info and can be coded using Block 5 for either program. This Block uniquely identifies each control, describes the control briefly and provides an optional URL for retrieving control documentation.

Certain fields in this Block are, however, unique to RCRA CA and are designed to allow compatibility with differences between how RCRAInfo and SEMS identify control information. SEMS uses a unique ID (CONTROL_ID); RCRA Info currently requires three separate fields:

- EVENT_SEQ: The sequence number of the EVENT of which the control is associated
- ACTUAL_DATE: The date at which the control became effective

[13] A datum is a reference system that approximates the irregularities of the Earth's surface, against which measurements of latitude, longitude and elevation can be made. The most commonly encountered North American datums are the North American Datum of 1927 (NAD27), North American Datum of 1983 (NAD83) and the World Geodetic System 1984 (WGS84). The NADs were designed to be most accurate for measurements in North America than WGS84. WGS84 varies only slightly from NAD83 and is more commonly used because it applies generally to the entire globe (hence its use in commonly accessible digital globes such as Google Earth).

- EVENT_CODE: a subset of RCRAInfo Event codes relating to the establishment and termination of ICs and ECs

Block 5 also provides the option for a Region to provide a REGIONAL_SUPP_CONTROL_ID equivalent to Superfund's CONTROL_ID. This is field could receive supplemental regional IDs that may be developed for the Long-Term Stewardship program or other functions. Or, over the long term, RCRAInfo might be updated to provide for a single ID for controls, in place of the current EVENT_SEQ, EVENT_CODE, ACTUAL_DATE. In that case, REGIONAL_SUPP_CONTROL_ID would assume equivalence to CONTROL_ID, enabling direct compatibility between the two programs.

4 DATA CODING GUIDANCE

The CSGDS forms the structure for exchanging geospatial site information between regional GIS databases and headquarters. In general, the regions are responsible for maintaining current geospatial information, including narrative descriptions and metadata. The appropriate Headquarters system provides all non-geospatial data for which that system is the authoritative source. This ensures that regional and Headquarters geospatial records contain complete, authoritative and consistent data at all times.

All fields indicated in green are Export fields for general public use. They reflect the most current state of the site. Historical information (information associated with earlier sequence numbers and effective dates) may be accessed from the system itself.

The CSGDS shows two columns of field names, one for Superfund and one for RCRA CA. Wherever possible the field names correspond to field names currently used in SEMS or RCRAInfo. In addition, the parallel format documents field correspondences between the two systems so that if a site migrates from one program to the other, geospatial information will carry over accurately and no geospatial information will be lost.

The following pages describe the entire CSDGS.

4.1 Site Identifiers

For this block of information, the headquarters database will be the source of authoritative data. Regions may, at their discretion, accept data contributions from states, regulated facilities or other providers as they compile their regional geospatial layers. They may elect to extend the schema presented here to include that original source information.

Cleanup Site Geospatial Data Schema V3.5 BLOCK 1 FEATURE IDENTIFIERS						
Req/Opt	Source	SUPERFUND Field Name	RCRA CA Field Name	Field Size	Description	Comments
R	R	EPA_PROGRAM	EPA_PROGRAM	VARCHAR(30)	Superfund, RCRA CA	
R	R	REGION_CODE	REGION	VARCHAR2(2)	The number of the EPA Regional Office, from 1 to 10.	
R	R / Hq	EPA_ID	HANDLER_ID	VARCHAR2(12)	Assigned EPA program site identification number.	Would be the CERLIS site ID for a Superfund site, Handler ID for a CA site.
R	Hq	SITE_NAME	HANDLER_NAME	VARCHAR2(100)	Assigned site name. Should be a duplicate of the long name used in the programmatic database.	For RCRA CA, could be a state assigned name.
S	Hq	NPL_STATUS_CODE	*No Corrective Action Equivalent*	VARCHAR2(1)	Supplied by system based on dropdown choice.	

Cleanup Site Geospatial Data Schema V3.5 BLOCK 1 FEATURE IDENTIFIERS						
Req/Opt	Source	SUPERFUND Field Name	RCRA CA Field Name	Field Size	Description	Comments
S	Hq	FEDERAL_FACILITY_DETER_CODE	FED_WASTE_GENERATOR	VARCHAR2(1)	Supplied by system based on dropdown choice.	
RA	Hq	URL_ALIAS_TXT	FACILITY_INFO_URL	VARCHAR(2000)	A URL to an address that documents the site in more detail. Most likely it would be to Envirofacts, from which most other information would also be available.	This would provide a single additional link to externally available full site data.
	R	*No Superfund Equivalent*	REGIONAL_PROFILE_URL	VARCHAR(2000)		

Req./Opt.: R=Required; **RU**=Required but may include "Unknown," such as for null values; **RA**=Required if applicable or if available; **O**=Optional; **S**=System generated, such as a feature ID number; **D**=Derived from another supplied value (e.g., area derived from a supplied boundary).
Source: R=Entered by Regional Office staff; **Hq**=Supplied by Headquarters by synchronization from SEMS or RCRAInfo; **V**=Varies Hq or R: only current example is NPL point lat/long; **S**=Entered by System.
Color: Light Blue=not exported (for internal EPA use); **Light Green=**exported, by any method (flat file, web service); **Dark Green=**exported, and can include multiple field entries (e.g., multiple URLs, multiple Handler Types, etc.)

4.2 Site Address

Cleanup Site Geospatial Data Schema V3.5 BLOCK 2 SITE ADDRESS						
Req/Opt	**Source**	**SUPERFUND Field Name**	**RCRA CA Field Name**	**Field Size**	**Description**	**Comments**
R	Hq	STREET_ADDR_TXT	LOCATION_STREET_NO	VARCHAR2(50)	Either the urban street number and name, a descriptive rural address or a facility references, such as a military base name.	This must be the authoritative address as maintained in the relevant headquarters system. For Superfund the field size is VARCHAR2(50), in RCRAInfo, the field size is VARCHAR2(12). In the CIMC webservice, the LOCATION_STREET_NO and LOCATION_STREET1 are concatenated and called LOCATION_STREET to ensure what the user sees is all on one line.
R	Hq	*No Superfund Equivalent*	LOCATION_STREET1	VARCHAR2(50)	Street Name	In the CIMC webservice, the LOCATION_STREET_NO and LOCATION_STREET1 are concatenated and called LOCATION_STREET to ensure what the user sees is all on one line.

O	R	ADDR_COMMENT	*Not used*	VARCHAR(1000)	Explanation of the site address, especially if it is rural style (e.g., 5 miles west of X)	This may already be available in SEMS. Otherwise it should be provided by the regional GIS data developer.
RA	Hq	CITY_NAME	LOCATION_CITY	VARCHAR 2(60)	City name the site is within or has as a mailing address.	City is not always a part of cleanup site addresses. For Superfund the field size is VARCHAR2(60), in RCRAInfo, the field size is VARCHAR2(25).
RA	Hq	COUNTY	LOCATION_COUNTY_NAME	VARCHAR2(50)	Names of all counties that are included within the site.	For RCRAInfo, imputed from COUNTY_CODE. For Superfund the field size is VARCHAR2(50), in RCRAInfo, the field size is VARCHAR2(27).
R	Hq	STATE_CODE	LOCATION_STATE	VARCHAR2(2)	Postal abbreviation	R8 uses State FIPS ID or code, else valid value text.
RA	Hq	ZIP_CODE	LOCATION_ZIP	VARCHAR(14)	Generally, the 5-digit code, extensible to 9	

Req./Opt.: R=Required; **RU**=Required but may include "Unknown," such as for null values; **RA**=Required if applicable or if available; **O**=Optional; **S**=System generated, such as a feature ID number; **D**=Derived from another supplied value (e.g., area derived from a supplied boundary).
Source: R=Entered by Regional Office staff; **Hq**=Supplied by Headquarters by synchronization from SEMS or RCRAInfo; **V**=Varies Hq or R: only current example is NPL point lat/long; **S**=Entered by System.
Color: Light Blue=not exported (for internal EPA use); **Light Green=**exported, by any method (flat file, web service); **Dark Green=**exported, and can include multiple field entries (e.g., multiple URLs, multiple Handler Types, etc.)

4.3 Contact Information

This block of information is supplied by SEMS or RCRAInfo. It contains the name and contact information of the most appropriate person available in the region (or, at the region's discretion, a contact at the state or facility) to answer public questions on the site's geospatial and basic descriptive information.

Cleanup Site Geospatial Data Schema V3.5 BLOCK 3 CONTACT INFORMATION						
Req/Opt	Source	SUPERFUND Field Name	RCRA CA Field Name	Field Size	Description	Comments
R	Hq/R	SITE_CONTACT_NAME	FACILITY_CONTACT_NAME	VARCHAR(255)	Contact name for general site inquiries, generally the RPM in Superfund and the facility's Public Point of Contact for RCRA Corrective Action. Source is the most recent available in SEMS or RCRAInfo.	
R	Hq / R	PRIMARY_TELEPHONE_NUM	FACILITY_CONTACT_TEL	VARCHAR2(20)	Telephone number for the above contact.	
R	Hq	SITE_CONTACT_EMAIL	*Not used*	VARCHAR2(100)		
R	Hq	*No Superfund Equivalent*	CONTACT_EMAIL_ADDRESS	VARCHAR2(80)	Email address for the above contact.	
R	Hq	*No Superfund Equivalent*	CONTACT_PHONE_AND_EXT	VARCHAR2(22)	Phone number for the above contact.	

Req./Opt.: R=Required; **RU**=Required but may include "Unknown," such as for null values; **RA**=Required if applicable or if available; **O**=Optional; **S**=System generated, such as a feature ID number; **D**=Derived from another supplied value (e.g., area derived from a supplied boundary).
Source: R=Entered by Regional Office staff; **Hq**=Supplied by Headquarters by synchronization from SEMS or RCRAInfo; **V**=Varies Hq or R: only current example is NPL point lat/long; **S**=Entered by System.

Color: **Light Blue**=not exported (for internal EPA use); **Light Green**=exported, by any method (flat file, web service); **Dark Green**=exported, and can include multiple field entries (e.g., multiple URLs, multiple Handler Types, etc.)

4.4 Site Feature/AreaName Information

This block of information identifies the geospatial feature in question and links to the appropriate geospatial information feature class that contains the required geometry and metadata.

The type of geographic feature (called Features in Superfund and AreaNames in RCRA) identified here differ between the two programs. The choices in the dropdowns provided are limited to the most common choices encountered in practice.[14] Because so many features are either rarely encountered or unique to individual sites, the "Other" option (Other Feature Point, Other Feature Line, Other Feature Boundary) should be used freely, with the appropriate name and description provided in the feature class table.

Cleanup Site Geospatial Data Schema V3.5 BLOCK 4 SITE FEATURE/AREANAME INFORMATION						
Req/Opt	**Source**	**SUPERFUND Field Name**	**RCRA CA Field Name**	**Field Size**	**Description**	**Comments**
S	S	SITE_FEATURE_ID	*Not used*	TBD	Unique feature identifier supplied by system.	Will vary depending on the format and system supporting the data.
O	R	*No Superfund Equivalent*	ENTIRE_FACILITY_IND	VARCHAR2(1)	A Y/N indicator for is the area the entire facility or is it not the entire facility	
R	S	SITE_FEATURE_TYPE_CODE	*Not used*	VARCHAR(2)	Select from dropdown. See Table 1 below.	"Other" will be the feature type available to describe all unique site-specific features. See table below.
R	R	SITE_FEATURE_CLASS	*No Corrective Action Equivalent*	VARCHAR2(100)		

[14] One of the options for RCRA CA is "Part," a term currently used only in Region 7. Parts are usually individual land parcels that make up part of a Facility. They are primarily used by Region 7 to document component parcels within a Facility; they are included because they are a generic term that may become more widespread across EPA regions in the future.

Cleanup Site Geospatial Data Schema V3.5 BLOCK 4 SITE FEATURE/AREANAME INFORMATION						
Req/Opt	**Source**	**SUPERFUND Field Name**	**RCRA CA Field Name**	**Field Size**	**Description**	**Comments**
R	R	SITE_FEATURE_TYPE	*Not used*		Supplied by GIS personnel. In this case the feature type is the site boundary. Other types are listed based on regional input.	
R	R	SITE_FEATURE_NUMBER	*Not used*	VARCHAR2(10)	If more than one feature of the same type (such as OUs) exist, increment this number for each feature. For RCRA CA, the governing factor is an Event rather than a feature. For this reason, all Handler features should be unique. This and the following field could be used if necessary, however.	
R	R	SITE_FEATURE_SEQUENCE	*Not used*	VARCHAR2(2)	If feature is updated over time, increment this number for each update	
R	R/Hq	SITE_FEATURE_NAME	AREA_NAME	VARCHAR2(50)	This would be the common name that a feature is referred to. Often OUs are referred to by name and not just number.	For RCRA CA, this is the common name that a feature is referred to. Area name can be anything – a point, line or polygon. For Superfund the field size is VARCHAR2(50), in RCRAInfo, the field size is VARCHAR2(40).

Cleanup Site Geospatial Data Schema V3.5 BLOCK 4 SITE FEATURE/AREANAME INFORMATION						
Req/Opt	**Source**	**SUPERFUND Field Name**	**RCRA CA Field Name**	**Field Size**	**Description**	**Comments**
R	R	SITE_FEATURE_DESCRIPTION	AREA_NAME_DESCRIPTION	VARCHAR2(1000)	Create a short narrative description of the feature. For Superfund, explain how the boundary may include property boundaries, extent of contamination, or observance of geographic boundaries such as streams.	For RCRA CA, the description will almost always be that it includes the property boundary of the Handler (Federal Facilities may be an exception, since the CA area may be a subsector of a larger government/military installation).
R	R	CLEARED_PUBLIC_RELEASE	CLEARED_PUBLIC_RELEASE	VARCHAR2(1)	(Y/N) flag indicating if the feature is cleared for public release via associated web services. Default would be Y. Features not cleared for public release would be available to internal EPA staff via the intranet GeoPlatform or other internal platforms.	Default would be TRUE. Features not cleared for public release would be available to internal EPA staff via the intranet GeoPlatform or other internal platforms
RU	R	LAST_CHANGE_DATE	GIS_FEATURE_LAST_CHANGE_DATE	Date	SF: Date from which the feature is valid RCRA CA: Represents the day when piece of information is created. It does not represent the date that the area name was defined.	Example: If an OU boundary is changed because of partial delisting, insert the date of delisting.
RU	R	ORIGINAL_CREATION_DATE	DATA_COLLECTION_DATE	Date	Date the feature was originally collected/created.	

Cleanup Site Geospatial Data Schema V3.5 BLOCK 4 SITE FEATURE/AREANAME INFORMATION						
Req/Opt	**Source**	**SUPERFUND Field Name**	**RCRA CA Field Name**	**Field Size**	**Description**	**Comments**
RU	R	SITE_FEATURE_SOURCE	*No Corrective Action Equivalent*	100	The original source of the feature. Could be a regional GIS team, contractor, PRP, state, other.	
RA	R	FEATURE_INFO_URL	*No Corrective Action Equivalent*		A URL to an address that documents the feature in more detail.	These fields will link to an external table to permit an indeterminate number of URLs to be associated with this feature.
O	R	FEATURE_INFO_URL_DESC	AREANAME_INFO_URL_DESC	VARCHAR(20)	A very short description of the URL, such as the name of the system (e.g., SEMS, RCRAInfo) or the content of the URL target (e.g., Institutional Control)	
RU	R	HORIZ_COLLECT_METH_CODE	HORIZONTAL_COLL_CODE	VARCHAR2(30)	SEMS code that best describes how the data were created/collected.	Values to be updated from existing EPA Latitude Longitude Standard.
R	R	HORIZ_COLL_DATUM	*Not used*	VARCHAR2(5)	Required if not automatically specified by POINT_COLL_METHOD. Values can include NAD83, WGS84. Any measurements in NAD27 should be converted to WGS84.	Can be imputed from collection method or supplied by user if necessary.

Cleanup Site Geospatial Data Schema V3.5 BLOCK 4 SITE FEATURE/AREANAME INFORMATION						
Req/Opt	**Source**	**SUPERFUND Field Name**	**RCRA CA Field Name**	**Field Size**	**Description**	**Comments**
R	R	*No Superfund Equivalent*	HORIZONTAL_COLL_DESC	VARCHAR2(100)	The text description of the method used to create the feature. The collection method implies the accuracy tier for the measurement.	
RU	R	*No Superfund Equivalent*	HORIZONTAL_ACC_MEASURE	VARCHAR(1)	Specific accuracy for this measurement in meters, if available.	Required, if available
R	R	TIER_ACCURACY_CODE	TIER_ACCURACY_CODE	VARCHAR(1)	This dropdown supplies the Tier number to the system (values 1 - 6 used for cleanup sites in both programs).	This may be imputed directly from the horizontal collection method using the
R	R	*Not Used*	TIER_ACCURACY_DESC	VARCHAR(25)	The text description of the tier accuracy code.	Should be determined by the collection method code. Code will now be NGDP Accuracy Tier number.
R	R/HQ	*No Superfund Equivalent*	AREA_ACREAGE	NUMER (13,2)		
R	V	LATITUDE_DECIMAL_VAL	LATITUDE_MEASURE	Double	Decimal: 6 significant digits provided. The value in this field must be reprojected in WGS84 if its original collection method is not WGS84.	For the NPL point vintage, must be the value published in the Federal Register.

Cleanup Site Geospatial Data Schema V3.5 BLOCK 4 SITE FEATURE/AREANAME INFORMATION						
Req/Opt	**Source**	**SUPERFUND Field Name**	**RCRA CA Field Name**	**Field Size**	**Description**	**Comments**
R	V	LONGITUDE_DECIMAL_VAL	LONGITUDE_MEASURE	Double	Decimal: 6 significant digits provided. The value in this field must be reprojected in WGS84 if its original collection method is not WGS84.	For the NPL point vintage, must be the value published in the Federal Register.
O	R	GIS_LENGTH	LENGTH	Double	The length as derived from GIS, with units. Units TBD: probably English, such as in feet if areas are shown in acres.	This value should be calculated by GIS personnel using an appropriate projection for their part of the country.
R	R	GIS_UNITS	LENGTH_UNITS	15	The length units as derived from GIS.	
R	R	GIS_AREA	*Not used*	Double	The area as derived from GIS, not the internal database area field.	This value should be calculated by GIS personnel using an appropriate projection for their part of the country.
R	R	GIS_AREA_UNITS	*Not used*	15	The area units as derived from GIS.	

Table 1: Most Frequently Used Horizontal Collection Methods		
HORIZONTAL COLLECTION METHOD CODE (Interim)	**COLLECTION METHODS**	**ESTIMATED TIER ACCURACES**
107-025	Classical Survey	Tier 1 (<1m)
103-030	Survey Grade GPS	Tier 1 (<1m)
107-026	Conversion from Metes and Bounds (with <1m accurate staring point)	Tier 1 (<1m)
104-022	Online Map Interpolation (1m or less image resolution)	Tier 2 (1 – 5 m)
104-023	Online Map Interpolation (3m+ image resolution)	Tier 3 (6 – 25m)
103-028	GPS Handheld Unit	Tier 3 (6 – 25m)
103-030	GPS Smartphone (<25m accuracy)	Tier 3 (6 – 25m)
104-022	Paper Map Interpolation (at least 1:24,000)	Tier 3 (6 – 25m)
107-027	Conversion from Metes and Bounds (w/ uncertain starting point or uncertain north bearing)	Tier 3 (6 – 25m)
103-031	GPS Smartphone (unknown accuracy)	Tier 4 (26 – 100m)
101-008	Urban Style Address Match	Tier 4 (26 – 100m)
101-009	Rural Style Address Match	Tier 5 (101 – 200m)
105-000	Unknown	Tier 5 (101 – 200m)

Table 2: Example Site Feature Types	
Superfund Site Feature Types	**RCRA Corrective Action Site Area Names**
NPL Point (as published in Federal Register)	Facility Address Point
Primary Point (as indicated by the primary flag in SEMS)	
Site Address Point	Facility Entrance Point
Site Entrance Point	Approximate Center Point of Facility
Approximate Center Point of Site	Facility Geometric Centroid (point)
Site Geometric Centroid (point)	Other Feature Point
Other Feature Point	----------
----------	Road (line)
Road (line)	Right of Way (line)
Right of Way (line)	Fence
Discharge Pipe (line)	Discharge Pipe (line)
Other Feature Line	Other Feature Line
----------	----------
Site Boundary	Entire Facility
OU Boundary	Part
Institutional Control Boundary (PC)	Solid Waste Management Unit (SWMU)
Institutional Control Boundary (GC)	Area of Contamination (AOC)
Institutional Control Boundary (ID)	Institutional Control Boundary (PR)
Institutional Control Boundary (EP)	Institutional Control Boundary (GC)
Engineering Control Boundary (GW)	Institutional Control Boundary (ID)
Engineering Control Boundary (NG)	Institutional Control Boundary (EP)
Contamination Boundary (undefined)	Engineering Control Boundary (GW)
Contamination Boundary (Soil)	Engineering Control Boundary (NG)
Contamination Boundary (Groundwater)	Other Feature Boundary
Contamination Boundary (Surface Water)	
Contamination Boundary (Sediment)	
Contamination Boundary (Air)	
Antecedent Boundary	
Deleted Area Boundary	
Other Feature Boundary	

4.5 Control Information

This block of information describes all institutional or engineering controls that apply to a facility. At this time, control information will be compiled only for RCRA Correction Action facilities, not for Superfund sites. No single unique control identifier is available from RCRAInfo. Instead, controls must be identified by supplying the Event sequence number, Event code and the "actual date" that applies to the control. Links to relevant control documents should be supplied in all cases.

Cleanup Site Geospatial Data Schema V3.5 BLOCK 5 CONTROL INFORMATION						
Req/Opt	Source	SUPERFUND Field Name	RCRA CA Field Name	Field Size	Description	Comments
O	R	HAS_CONTROL	*Not used*	1	Flag for whether the feature includes at least one Boundary Control, either Institutional or Engineering.	
R	R	CONTROL_ID	REGIONAL_SUPP_CONTROL_ID	VARCHAR(20)	Either a unique ID for a control under Superfund, or for RCRA CA, an optional field to receive supplemental regional IDs that may be developed for the Long Term Stewardship program or other functions.	A feature or AreaName may have any number of Controls. For Superfund, a unique ID may be available. For RCRA CA, this field will only be used in a Region develops a unique ID, or if RCRAInfo adopts such an ID in the future Currently, RCRAInfo identifies controls using 4 fields: HANDLER_ID, EVENT_SEQ, EVENT_CODE, ACTUAL_DATE. See below.

Cleanup Site Geospatial Data Schema V3.5 BLOCK 5 CONTROL INFORMATION						
Req/Opt	**Source**	**SUPERFUND Field Name**	**RCRA CA Field Name**	**Field Size**	**Description**	**Comments**
R	R	CONTROL_TYPE	EVENT_CODE	VARCHAR(30)	For Superfund, type of Control from dropdown, including both ICs and ECs. Values are Enforcement & Permitting Tools (EP); Government Control (GC); Proprietary Control (PC); Information Device (ID); Groundwater Control (GW); Non-Groundwater Control (NG). For RCRA CA, a subset of RCRAInfo Event codes relating to the establishment and termination of ICs and ECs: CA770GW, CA770NG, CA772EP, CA772GC CA772ID, CA772PR, CA780GW, CA780NG, CA782EP, CA782GC, CA782ID, CA782PR	Superfund and RCRA CA identification of specific ICs and ECs is not fully equivalent at this time. In the event that a CA site moves to Superfund, the scripts for identifying unique IDs for ICs and ECs will have to be customized based on the coding standards applying to SEMS and RCRAInfo at the time of the transfer.
RA	R	*No Superfund Equivalent*	ACTUAL_DATE	DATE	This is the date at which the control becomes effective, such as the filing date of a proprietary control.	Unique to RCRA CA
RA	R	*No Superfund Equivalent*	EVENT_SEQ	20	The sequence number of the EVENT of which the control is associated.	Unique to RCRA CA
R	R	*No Superfund Equivalent*	FEDERAL_FACILITY	VARCHAR2(3)	"Yes" or "No"	This will be readable on map by public. RCRAInfo does not distinguish between FF and FFDocket.

Cleanup Site Geospatial Data Schema V3.5 BLOCK 5 CONTROL INFORMATION						
Req/Opt	Source	SUPERFUND Field Name	RCRA CA Field Name	Field Size	Description	Comments
RA	R	CONTROL_DESC	CONTROL_DESC	VARCHAR(1000)	Briefly describe the Control.	The source would likely be the existing RCRAInfo IC_NOTE or EC_NOTE field, copied and pasted.
O	R	CONTROL_URL	CONTROL_URL	VARCHAR(255)	URL of the document(s) defining the Point Control. This may be a pointer to SEMS RM, a Drupal page of RCRA CA documents or some other repository.	Multiple URLs will be supported. Links may need to be updated over time, especially where the target is a Drupal page.

Req./Opt.: R=Required; **RU**=Required but may include "Unknown," such as for null values; **RA**=Required if applicable or if available; **O**=Optional; **S**=System generated, such as a feature ID number; **D**=Derived from another supplied value (e.g., area derived from a supplied boundary).
Source: R=Entered by Regional Office staff; **Hq**=Supplied by Headquarters by synchronization from SEMS or RCRAInfo; **V**=Varies Hq or R: only current example is NPL point lat/long; **S**=Entered by System.
Color: Light Blue=not exported (for internal EPA use); **Light Green=**exported, by any method (flat file, web service); **Dark Green=**exported, and can include multiple field entries (e.g., multiple URLs, multiple Handler Types, etc.)

4.5.1 Institutional Controls

Site features, or AreaNames (RCRA CA), may be subject to one or more Institutional Controls (IC)— non-engineered or legal controls that minimize the potential for human exposure by limiting land or resource use. ICs are established to supplement the effectiveness and integrity of interim or final remedies where residual contamination remains in place above unrestricted use action levels. Often, ICs are a critical component of the cleanup process and are used by the site manager to ensure both the short- and long-term protection of human health and the environment.

Institutional controls can govern physical site features, such as a no-dig order for a landfill cap. They can also govern activities related to the entire site or to a sub-area within the site, such as a formally designated groundwater restriction area.

Superfund and RCRA jointly recognize four types of Institutional Controls:[15]

- **Informational Devices:** Informational tools provide information or notification that residual or capped contamination may remain on site. Common examples include state registries of contaminated properties, deed notices, and advisories.
- **Enforcement and Permit Tools:** Under sections 104 and 106(a) of CERCLA, Unilateral Administrative Orders (UAO) and Administrative Orders on Consent (AOC) can be issued or negotiated to compel the land owner (usually a PRP) to limit certain site activities at both Federal and private sites; Consent Decrees (CD) can also be negotiated at private sites under 122(d). Under RCRA, institutional controls may be established through the use of enforcement and permit tools including Hazardous Waste Management Permits, administrative orders or other enforceable agreements.
- **Proprietary Controls:** These controls, such as easements and covenants, have their basis in real property law and are unique in that they generally create legal property interests. In other words, proprietary controls involve legal instruments placed in the chain of title of the site or property.
- **Governmental Controls:** Governmental controls are usually implemented and enforced by a state or local government and can include zoning restrictions, ordinances, statutes, building permits, or other provisions that restrict land or resource use at a site.

4.5.2 Engineering Controls

Engineering controls (EC) consist of engineering measures designed to minimize the potential for human exposure to contamination by either limiting direct contact with contaminated areas, reducing contamination levels, or controlling migration of contaminants through environmental media. Some examples of mappable engineering controls are capping, containment, slurry walls, extraction wells, or treatment methods that minimize contamination. Often ECs and ICs are used together as part of the remedy.

[15] See *Institutional Controls: A Site Manager's Guide to Identifying, Evaluating and Selecting Institutional Controls at Superfund and RCRA Corrective Action Cleanups*, USEPA, Office of Solid Waste and Emergency Response, September 2000. The above institutional control descriptions are quoted from this document.

It is possible to code the existence of engineering controls in the geospatial data schema below as an attribute to any feature, just as it is for institutional controls. Engineering controls coded here fall into the two categories used by RCRA CA:

- **Groundwater Controls:** These controls include any EC pertaining to groundwater, including in situ and ex situ treatment like bioremediation, in situ permeable reactor barriers, monitored natural attenuation, long-term monitoring, etc.
- **Non-groundwater Controls:** These controls include ECs not pertaining to groundwater, such as barriers or caps.

4.6 Geospatial Information Collection

The Regional GIS coordinators or their designees will be responsible for managing the input of disparate sources of geospatial data into the formats and projections of the Regional geodatabases. All geodatabases will be projected in Geographic Coordinate System, decimal degrees, WGS84 datum (GCS_WGS84) for consistency and compatibility, when exported for secondary users, with the most commonly used visualization systems (e.g., ESRI ArcGIS, AGOL, Google Earth, Bing, etc.).

4.6.1 Collection Methods and Accuracy Targets

Under the EPA *Enterprise Information Management Policy* (EIMP) and its supporting procedures, the collection of geospatial information is governed by three general principles:

1. The program requesting the geospatial data must specify, for each feature type, the minimum accuracy necessary for program business purposes.
2. Accuracy targets are stated as ranges ("tiers") rather than point values. There are ten tiers, ranging in value from Tier 1: < 1 meter to Tier 10: > 5000 m. Cleanup site measurements typically make use of accuracy Tier 2 (1 – 5m) through Tier 6 (201 – 999m).
3. Accuracy may be documented, at a minimum, by listing the technical method used to collect the measurement, with the accuracy tier achieved imputed from the method. Additional documentation is encouraged and should be entered if available. The Office of Environmental Information is responsible for maintaining the Agency's working list of correction method accuracies, including current and emerging methods as well as methods used in historical documents.[16]

Most collection methods can be associated with a specific accuracy tier. For these, an accuracy target is satisfied if the user employs a collection method associated with the target tier or a lower-numbered tier. This policy assumes, of course, that collection methods are used correctly. (See Addendum 1 for guidance on using GPS units in the field. Additional technical guidance and best practices for other common collection methods may be added to this document in the future.) Accuracy Tiers are a classification and characterization scheme based on a range of values. Their main purpose is to inform the user on suitability of use for various secondary purposes.

Two collection methods that cannot be uniquely associated with a particular accuracy tier are (1) online map interpolation (widely used) and (2) metes and bounds (frequently used in the RCRA program but rarely in Superfund). For these methods, collection accuracy may vary even when they are used

[16] See the *Latitude/Longitude Data Standard*, the *National Geospatial Data Policy* and subsequent updates and additional guidance issued from time to time under the EIMP.

correctly, and experienced professional judgment is necessary to estimate the correct tier value associated with an observation.

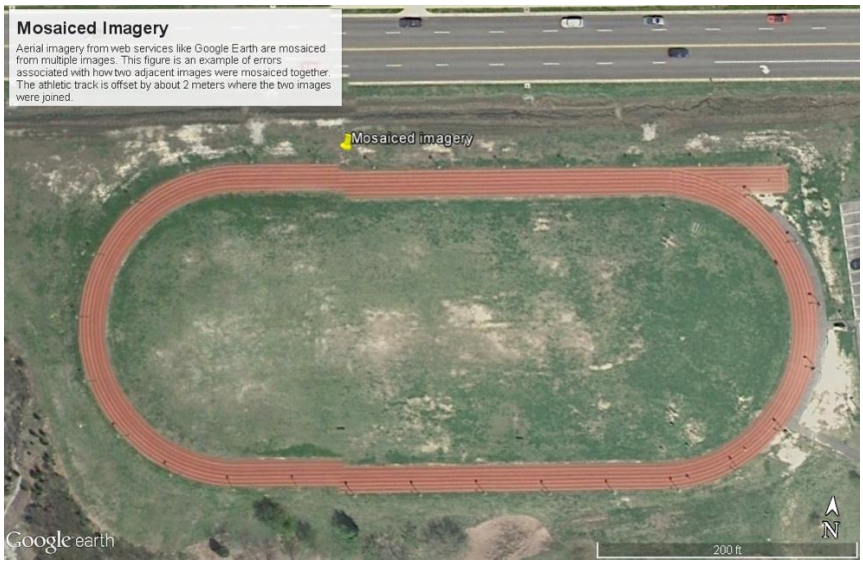
Regions may use any the collection methods specified in the 2007 Latitude Longitude Standard but note that many of these are out of date and do not correctly describe online map interpolation, currently the most common method of digitizing data. This guidance assumes that users will employ one of the six most common methods currently in use in the regions:

- **Professional Grade GPS (Tier 2: 1 – 5m):** These employ differential correction, usually through Space Based (or Satellite Based) Augmentation Systems (SBAS), for enhanced accuracy in the 2 – 5m range. They are more expensive than recreational grade GPS systems but are preferred for CSDGS purposes wherever available. In the process of performing professional grade GPS work, the data processing will generate a number of accuracy values, such as geometric dilution of precision (GDOP) among other calculated accuracy values based on parametric information gathered during observation. Professional GPS software will typically provide metadata with these calculated accuracy values. Where available, they should be entered into the CSDGS (HORIZ_ACCURACY_MEASURE or HORIZ_ACC_MEASURE).
- **Recreation grade GPS units and smartphones (Tier 3: 6 – 25m or Tier 4: 26 – 100m):** These can include recreational GPS handhelds or smartphones with auxiliary apps that can document the number of GPS satellites visible at the time of the measurement. Such units can reach about 10 m accuracy (within Tier 3), but in some situations (especially in more remote rural areas) are conservatively rated at Tier 4.
- **Web-based Map Interpolation (Tier 2, 3 or 4, based on professional judgment):** Variable-scale web map interpolation, using systems such as ESRI ArcGIS or Google Earth) has replaced paper maps for most cleanup site data collection where on-site measurements are impractical. Accuracy values in web mapping applications can be estimated via a combination of aerial imagery metadata and zoom level. Where possible, users should calculate and provide accuracy values. Further guidance may be forthcoming in the future, but as of this writing, tier levels for data collected using online maps should be determined by the professional judgment of an experienced GIS analyst based on the following:
  - Consider the metadata of the imagery, if available.
  - Prefer the use of image files, if available, before using a web service or use a service such as NAIP from the USDA for which metadata are available.
  - The most current imagery may not always be the most appropriate, especially if an historic image better represents the features that are to be derived from the image.
  - Consider image resolution: size of pixels relative to the size of the feature or to the details of the feature.
  - Consider topographic relief: is the area flat or hilly? Flat areas are more accurately georeferenced/orthorectified than hilly areas due to parallax.
  - Consider the ability to see the feature of interest: is the feature distinct or does it blend into other features or to the background?
  - Lighting: time of day when image was taken, e.g., is the feature in shadow and difficult to see in early morning with long shadows?
  - Age of photo: is the feature the same size today as when photo was taken, or has it been modified in the interim? This is not an issue if the purpose of tracing is to establish the location or size of a former feature as of the date of the aerial photo.

- Quality of georeferencing/orthorectification: it may be possible to quantify this variable by comparing a visible feature of known coordinates with apparent coordinates from the image. It may also be possible to determine qualitatively by comparing a single feature location across multiple aerial images, rejecting those that are inconsistent with most images.
- Height of feature relative to ground level: for example, tracing a building roof rather than the building outline at ground level can introduce error depending on where the building is relative to the flight path due to parallax; the higher above ground level the feature is, the greater the potential error due to parallax.

Example of parallax distortion error

- Scale at which image is viewed when tracing feature: when zoomed out to view a very large feature, small changes in where image is traced may result in larger errors; the amount of error can be estimated by zooming in on random points after tracing to measure the precision of the tracing.
- Image mosaic errors: adjoining images may misalign (as in the image below), making it difficult to establish absolute positioning of a feature.



Example of image mosaic error

- Number of points used to trace a feature: tracing curved features results in approximating the curve with multiple straight-line segments; the fewer points used on a curved feature, the greater the potential error between points.

- **Urban Style Address Matching (Tier 5: 101 – 200m):** Acceptable in urban or suburban areas for collecting a general site location point, address matching (also called geocoding) is the most rapid and convenient way to generate an acceptable site point, such as for Pre-CERCLA screens. Results from geocoding algorithms can vary by geography and conditions, such as where datasets combine actual coordinates associated with addresses with street-centerline interpolation routines. In addition, street-centerline interpolation results can vary as a function of street segment length. Geocoding algorithms, such as Oracle or ESRI geocoders, should be able to provide an accuracy estimate for each geocoded value. These values should be provided wherever possible.
- **Classical Surveying Techniques (Tier 1: <1m):** These may be used in concert with detailed engineering studies or site sampling programs and provide the highest level of precision (Tier 1: <1m) achievable. While not required for any general Superfund site data collection covered here, any data collected using classical surveying is more than adequate for all general purposes. Classical surveying methods typically employ error-of-closure checks, which should be available in submitted contract deliverables. Where available, these should be provided. Classical surveying describes boundaries as distance and bearing from point to point, beginning from a permanent benchmark, but may not include geographic coordinates of each point. Although the polygon described will be very precise, the survey accuracy in the real world will depend on how well the reference benchmark is defined relative to a real-world coordinate system, and how well bearings are defined (e.g., relative to true north, magnetic north at the time of the survey, or "plant" north.
- **Metes and Bounds (Tier 1, 2, or 3, based on professional judgment):** As for online maps, the tier level associated with a feature collected using metes and bounds must be determined using professional judgment. Some RCRA CA boundaries, such as Entire Facility areas, may correspond to parcel boundaries. Proprietary institutional control boundaries also require an area description through what is called a legal description written into the control itself. These parcel boundaries and legal descriptions are typically defined geospatially by a system commonly referred to as "metes and bounds," a narrative rather than graphic representation that describes the boundary as a series of points and vectors between the points, and which are created by surveyors using classical surveying techniques. As such, they can achieve Tier 1 accuracy if real-world coordinates of the starting point are known, and the bearings are relative to true north, or the real-world coordinates of at least 2 of the property corners are known. If unknown, the polygon created will need to be aligned with other features in the base map or air photo, so the resulting accuracy will depend on the alignment method and the accuracy of the air photo.

4.6.2 Common Feature/AreaName Types and Associated Accuracy Tiers

Table 3 below summarizes typical point, line and polygon features/AreaNames and their associated tier accuracy targets.

Table 3 Geospatial Feature/AreaName Types, Accuracy Targets and Acceptable Collection Methods		
General Site Points		
Accuracy Tier Value	**Description**	**Acceptable Methods**
Tier 2 (1 – 5 m)	Sampling Point (NPL)	GPS with real-time SBAS integration (professional or survey grade units); Online Map Interpolation if supported by best professional judgment.
Tier 3 (6 – 25m)	Bounding box (NW) < 50 acres Bounding box (SE) < 50 acres Monitoring station (air) Monitoring station (surface water) Monitoring station (well/groundwater) Sampling point (pre-CERCLA) Site entrance Site geographic centroid	Online map interpolation (ESRI ArcGIS, ESRI AGOL, Google Maps, Google Earth, Bing Maps, MapQuest, TIGER) if supported by best professional judgment. GPS Handheld Unit (recreational grade) GPS Smartphone, w/ accuracy within 25m (e.g., wi-fi assist and/or satellite documentation > 10 satellites)
Tier 4 (26 – 100m)	Site address urban Bounding box (NW) >50 acres Bounding box (SE) >50 acres	Geocoding GPS Unspecified (including Smartphone)
Tier 5 (101 – 200m)	Site address (geocode) rural Site center (representative point, approximate)	Urban Style Address Matching – Block Face Map Interpolation (online or paper < 1:20,000 scale)
Tier 6 (201 – 999m)	Site address (geocode) rural	Rural Style Address Match
General Line Features		
Tier 2 (1 – 5 m)	Fence (if EC) Other Line Feature (If EC)	GPS with real-time SBAS integration (professional or survey grade units); Online Map Interpolation if supported by best professional judgment.
Tier 3 (6 – 25m)	Road Right of Way Discharge Pipe Other	Online map interpolation (ESRI ArcGIS, ESRI AGOL, Google Maps, Google Earth, Bing Maps, MapQuest, TIGER) if supported by best professional judgment. GPS Handheld Unit (recreational grade)

Table 3 Geospatial Feature/AreaName Types, Accuracy Targets and Acceptable Collection Methods		
		GPS Smartphone, w/ accuracy within 25m (e.g., wi-fi assist and/or satellite documentation > 10 satellites)
General Site Polygons		
Tier 1 (<1m)	Institutional Control (PC)	Classical Survey Metes and Bounds using <1m accuracy starting point Survey-grade GPS
Tier 2 (1 – 5 m)	Engineering Controls (NG)	GPS with real-time SBAS integration (professional or survey grade units); Online Map Interpolation if supported by best professional judgment.
Tier 3 (6 – 25m)	Site Boundary OU Boundary Antecedent Boundary (<50 acres) Deleted Area Boundary Institutional Control (ID,GC,EP) Engineering Control (GW)	Online map interpolation (ESRI ArcGIS, ESRI AGOL, Google Maps, Google Earth, Bing Maps, MapQuest, TIGER) if supported by best professional judgment. GPS Handheld Unit (recreational grade) GPS Smartphone, w/ accuracy within 25m (e.g., wi-fi assist and/or satellite documentation > 10 satellites) Metes and Bounds using uncertain starting point or uncertain north orientation
Tier 4 (26 – 100m)	Site Boundary >50 acres Contamination Boundary (undefined) Contamination Boundary (Soil) Contamination Boundary (Ground Water) Contamination Boundary (Surface Water) Contamination Boundary (Sediment)	Online Map Interpolation if supported by best professional judgment. GPS Unspecified (including Smartphone)
Tier 5 (101 – 200m)	Contamination Boundary (Air)	Map Interpolation (online or paper < 1:20,000 scale)

Best practices for Superfund site boundary delineation: EPA Superfund site boundaries have historically been based on some combination of land surface ownership (cadastral or land parcel data) and/or extent of contamination. Making all or part of a boundary congruent with existing cadastral or land parcel data is considered a best practice when appropriate for site, operable unit and institutional control boundary delineation to ensure that private and public land parcels are explicitly identified as belonging to a site, operable unit or for specific institutional control implementation. Cadastral or land

parcel data is clearly not a best practice where sites border on bodies of water or where air deposition of substances are an issue. In these cases, extent of contamination is likely to be the critical concern.

Cadastral data is generally created and maintained as part of the city or county government tax assessment. See Addendum 2 for technical best practices for accessing and using government sources of this data.

For sites that are not well represented as one or more land parcels, direct field capture of a site perimeter using Global Positioning Systems (GPS) is also considered a best practice. Well planned and executed GPS capture can easily lead to high accuracy and precision for boundaries, with accuracies measured in the 3-meter or less range.

In cases where cadastral data is not available, creating boundaries via on-screen digitizing may be the best remaining option. Use of high-resolution imagery as a base map or background is highly recommended and will help the boundary developer to interpret boundaries relative to visible features such as fence lines, streets, building footprints, and so forth. Accuracy using this method is difficult to calculate, as the boundary will inherit all locational error associated with the base image as well as any error introduced in the interpretation and delineation phase. Accuracies can be in the 10-meter range for this method, assuming diligent effort is taken to minimize input error sources, similar to many map or photo interpretation methods.

Best practices for RCRA Corrective Action facility boundary delineation: Regional branch chiefs have discretion over what boundaries are to be collected for any given facility. Generally, the focus is on mapping institutional and engineering controls. Facility boundaries may or may not be included. If they are included, boundaries of the Entire Facility are virtually always the boundaries of the property parcel or parcels that compose the site. These may be available from state or local cadastral databases. Regions may wish to ground-truth these secondary sources using references such as ESRI ArcGIS or Google Earth. See Addendum 2 for technical best practices for accessing and using government sources of this data.

Privacy considerations in site or facility boundary delineation: Cadastral or land parcel information is used extensively by Superfund and almost exclusively by RCRA CA in defining site and facility boundaries. Since this information will become public, to what extent does the use of parcel data amount to personal identifiable information (PII) and therefore raise distribution concerns? OLEM believes that PII issues are not involved in the drawing of boundaries using parcel information for the following reasons:

- The boundaries in question are only geospatial shapes. They do not carry attributes related to the ownership, tax assessment or any other PII of the associated plot of land. Individual owners are not identified in the distributed geospatial data. These names may, of course, be included in the Superfund Site name (SEMS) or the RCRA Facility name (RCRAInfo), but in such cases the names are already legally and publicly associated with the regulatory action.
- Many states, counties and cities have created public web sites that contain complete parcel data, including ownership, property values, addresses and other details. Where these are available, it is a simple matter for a user to look up all details of property ownership for the site or facility in question.

4.6.1 Other Feature/AreaName Types and Associated Accuracy Tiers

Table 4 Superfund Geospatial Feature Types, Accuracy Targets and Acceptable Collection Methods	
Other Feature Types (point, line or polygon as appropriate) **Includes:** Adit, Berm, Cap, Clean-up area, Cultural resource, Ditch, Excavation, Facility, Fence Impoundment, Lagoon, Pile, Pit, Pond, Road, Right-of-way, Storage tank, Surface mine, Utility corridor, Waste in place, Waste treatment unit, Well, OTHER	
Tier 3 (6 – 25m) If relevant for engineering purposes	Online map interpolation (ESRI ArcGIS, ESRI AGOL, Google Maps, Google Earth, Bing Maps, MapQuest, TIGER), if supported by best professional judgment. GPS Handheld Unit (recreational grade) GPS Smartphone, w/ accuracy within 25m (e.g., wi-fi assist and/or satellite documentation > 10 satellites)
Tier 4 (26 – 100m) If not relevant for engineering purposes	Online map interpolation (ESRI ArcGIS, ESRI AGOL, Google Maps, Google Earth, Bing Maps, MapQuest, TIGER), if supported by best professional judgment. GPS Unspecified (including Smartphone)

Table 4 above lists several optional feature types that may be appropriate to include for some sites. The highly variable nature of site cleanups makes it impossible to create a definitive list. This list was compiled by reviewing a large number of features found across multiple sites in multiple regions and simplifying it to include only the most commonly encountered categories. Features can be recorded as either points, lines or polygons as appropriate.

Target accuracy tiers for these optional features are similarly indeterminate. Table 3 recommends that if a feature is relevant to the engineering of a site solution, the feature should be measured at target accuracy Tier 2 (1 to 5 m). If it is not part of an engineering solution, then Tier 3 (6 – 25m) is adequate. Do not bother to record a feature if it cannot be measured to within accuracy Tier 4.

5 RECOMMENDED DATA DISTRIBUTION DISCLAIMER

Site boundaries are subject to change for a variety of reasons. Common reasons include increased knowledge regarding the nature and extent of contamination, discovery of additional contaminants at the site, changes to current risk-based standards or to increase accuracy and precision of existing boundaries.

The following disclaimer, developed previously with OGC and OECA, will be included in all distribution of cleanup site geospatial data to notify users of the status of the data and that its distribution does not support or represent official Agency viewpoints.

> *The Agency is providing this geospatial information as a public service and does not vouch for the accuracy, completeness or currency of data. Data provided by external parties is not independently verified by EPA. These data are made available to the public strictly for informational purposes. Data do not represent EPA's official position, viewpoint, or opinion, express or implied. This information is not intended for use in establishing liability or calculating Cost Recovery Statutes of Limitations and cannot be relied upon to create any rights, substantive or procedural, enforceable by any party in litigation with the United States or third parties. EPA reserves the right to change these data at any time without public notice.*

6 GLOSSARY OF ACRONYMS

AOC	Area of Contamination
CD	Consent Decrees
CERCLA	Comprehensive Environmental Response, Compensation and Liability Act [Superfund]
CERCLIS	Comprehensive Environmental Response, Compensation and Liability Information System [predecessor to SEMS]
CIMC	Cleanups in My Community
CSGDS	Cleanup Site Geospatial Data Schema
EC	Engineering Control
EDG	Environmental Dataset Gateway
EIMP	Enterprise Information Management Policy
EP	Enforcement & Permit Tools [type of Institutional Control boundary]
EPA	Environmental Protection Agency
FF	Federal Facilities
FGDC	Federal Geographic Data Committee
FRS	Facility Registry System
GC	Government Control [type of Institutional Control boundary]
GPO	GeoPlatform Online
GW	Groundwater [type of Engineering Control boundary]
IC	Institutional Control
ID	Information Device [type of Institutional Control boundary]
ISO	International Organization for Standardization
NARA	National Archives and Records Administration
NG	Non-groundwater Control [type of Engineering Control boundary]
NPL	National Priorities List
OGC	Office of General Counsel
OECA	Office of Enforcement and Compliance Assurance
OLEM	Office of Land and Emergency Management
OU	Operable Unit
PC	Proprietary Control [type of Institutional Control boundary]
POC	Point of Contact
RCRA	Resource Conservation and Recovery Act
RCRA CA	RCRA Corrective Action

RCRAInfo	Resource Conservation and Recovery Act Information [database]
ROD	Record of Decision
RPM	Remedial Program Manager
SEMS	Superfund Enterprise Management System
SWMU	Solid Waste Management Unit
TSD	Transportation, Storage and Disposal
UAO	Unilateral Administrative Order
URL	Uniform Resource Locator