

```

Schema" xml version="1.0" encoding="UTF-8"
enetwork" _ <xsd:schema
nDefault="unqualified" targetNamespace="http://www.epa.gov/exchangenetwork"
mon_v3_0.xsd" , xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:nei="http://www.epa.gov/exchangenetwork"
elementFormDefault="qualified"
version="3.0">
<xsd:include schemaLocation="EN_NEI_Common_v3_0.xsd"
- <!--
XML 3.0 Start of Schema Header
-->
_ <xsd:annotation>
ngenetwork</xsd:documentation>
Point</xsd:documentation>
XML 3.0 Point data</xsd:documentation>
Available:http://www.epa.gov/exchangenetwork/
cumentation>
<xsd:documentation>
Environmental Protection Agency
input format</xsd:documentation>
<xsd:documentation>
encoding="UTF-8" ?>
user</xsd:documentation>
<xsd:documentation>
ce="http://www.epa.gov/exchangenetwork"
tp://www.w3.org/2001/XMLSchema"
tp://www.epa.gov/exchangenetwork"
default="qualified" attributeFormDefault="unqualified"
SchemaLocation="EN_NEI_Common_v3_0.xsd"
ader
Schema Name: NEI XML 3.0
d:documentation>
entation>Current Version
http://www.epa.gov/exchangenetwork/
on>
entation>Description: The NEI XML 3.0 Point data
nat</xsd:documentation>
entation>Application: Varies by
:documentation>
entation>Developed By: Environmental Information
ng="UTF-8" ?>
http://www.epa.gov/exchangenetwork/
www.w3.org/2001/XMLSchema"
www.epa.gov/exchangenetwork"
="qualified" attributeFormDefault="unqualified"
Location="EN_NEI_Common_v3_0.xsd"

```

Facility ID XML Schema User's Guide

Release Status: Draft
Revision Date: 5/21/2009

TABLE OF CONTENTS

SCHEMA OVERVIEW	5
LEGEND	5
OVERVIEW OF CHANGES FROM VERSION 2.3	6
COMPOSITE STRUCTURE	6
RECONCILIATION WITH FACILITY DATA STANDARDS	6
MULTIPLE SCHEMA VARIANTS	6
DATA ORIGATION AND UNIQUE FACILITY IDENTIFICATION	6
GEOGRAPHIC LOCATION	7
ALTERNATIVE NAMES AND IDENTIFIERS	7
ELECTRONIC ADDRESSES	7
AFFILIATIONS	7
FACILITY DETAILS OVERVIEW	8
FACILITY DETAILS MODULE	11
FACILITY MODULE	12
ENVIRONMENTAL INTEREST MODULE	14
AFFILIATION MODULE	16
AFFILIATE MODULE	17
FACILITY INTEREST OVERVIEW	18
FACILITY INDEX OVERVIEW	19
FACILITY COUNT OVERVIEW	20

Acknowledgements

Facility ID Integrated Project Team	
Tom Aten – WI DNR	Ken Blumberg – EPA, Region I
Dennis Burling – NE DEQ	Pat Garvey – EPA, OEI
Ana Greene – EPA, OEI	Alex Klaessig – EPA OEI
Noel Kohl – EPA, Region V	Wendy Nye – EPA, Region I
Andrew Putnam – CO DPHE	Gerald Schwandt – MN PCA
Bob Simpson – EPA, Region II	Kurt Rakouskas – ECOS
Mitch West – OR DEQ	
Contractor Support	
Bethel Helene – Lockheed Martin	Veeree Kadans – Lockheed Martin
Kevin Lyons – Windsor Solutions, Inc.	Guy Outred – Windsor Solutions, Inc.
Bill Rensmith – Windsor Solutions, Inc.	Clint Teng – Lockheed Martin
TK Conrad – Windsor Solutions, Inc.	

Prepared by:



4000 Kruse Way Place
Building 2, Suite 285
Lake Oswego, OR 97035
(503) 675 7833

Document Revisions

Release Date	Status	Changes
3/19/2007	DRAFT	Original Draft Release
4/13/2009	DRAFT	Revised to incorporate elements of the GeoRSS Application Profile.
5/21/2009	DRAFT	Revised to include section summarizing changes from version 2.3.

DRAFT

Schema Overview

The Facility ID XML Schema is used to exchange facility site information among partners of the exchange network. In order to better support the many uses of this schema, four root components are available for exchange. These four components are described as follows:

- **Facility Details** – This component contains a detailed list of facility information including environmental interests and affiliates. This element is optimized for exchanging detailed facility information for a specific user query or for database synchronization between exchange partners.
- **Facility Interest** – This component contains a list of abbreviated facility information including environmental interests. This element is optimized for exchanging a list of high-level facility site information based on an ad hoc user query.
- **Facility Index** – This component contains a list of abbreviated facility information. This element is optimized for exchanging a list of high-level facility site information based on an ad hoc user query.
- **Facility Count** – This component contains only one element, FacilityCountMeasure.

Legend

Throughout this document, the presented figures will all use the same diagramming conventions. Below you will find a description for each of the utilized conventions:

Diagram Object/Text	Name	Description
Component Name	Component	This object represents a logical component in the schema.
SC: Component Name	Shared Component	Components that are prefaced with the text “SC:” represent Shared Schema Components.
<u>Element Name</u>	Unique Identifier	Elements that are underlined and bold are considered unique identifiers for the component that houses the element. A value must be provided for this element that is unique across all records that are exchanged for the parent component.
Element Name	Required Element	Elements that are bold are considered required and must be provided as a part of the exchange.
Element Name	Optional Element	Elements that are not underlined or bold are considered optional and are not required to be provided as a part of the exchange.
→	Required Relationship	This object represents a required relationship between one logical component and another. The arrow points from the parent component to the child component.
...▶	Optional Relationship	This object represents an optional relationship between one logical component and another. The arrow points from the parent component to the child component.

Overview of Changes from Version 2.3

The 3.0 version of the Facility Identification flow includes a number of improvements upon the previous 2.3 version. A complete listing of all changes can be found in the “change_log.txt” file in the schema folder. Here is a summary of the major changes:

Composite Structure

The 2.3 version of this flow allowed for the use of a payload variant comprised of nine separate XML files as an alternative to using a consolidated single XML file. This approach made it easier for some organizations to more readily move to XML from previous flat file type exchanges, but also negated some of the structural integrity that XML enforces in the data. The 3.0 version has removed this format option.

Reconciliation with Facility Data Standards

The Exchange Network's Core Reference Model (CRM) was developed to encourage the adoption of data standards within all new XML schema designs. The 3.0 schema makes extensive use of the CRM's shared schema components (SSCs) wherever possible in order to align with Exchange Network Facility Data Standards (FDS). For the most part, the same information is represented in 3.0 as in 2.3, but is organized in slightly different ways. This reorganization means that the data structures now follow the same organization as many other flow schemas that are based upon the CRM.

Although the vast majority of the data elements are the same, some 2.3 elements have been dropped as they were not explicitly included within the CRM (or FDS, upon which the SSCs were based). Specifically removed elements are: Organization DUNS #, Organization Type, Employer Identifier, Ultimate Parent Name, and Ultimate Parent DUNS #. However, many of these can still be accommodated in the new 3.0 schema using some of the more generic data structures it provides. For example, should an organization collect and wish to share any Ultimate Parent Name data, this can be represented by within the Affiliation schema structure using a separate Organization affiliation of type “Ultimate Parent Facility.”

Multiple Schema Variants

The 2.3 version supported only one root component to share detailed facility data providing comprehensive details for each exchanged facility. The 3.0 version provides four different root components with varying levels of detail: Facility Details, Facility Interest, Facility Index, and Facility Count. This allows data consumers to build more efficient data services that only retrieve the level of detail necessary for the operation. For example, if an operation only calls for a summarized list of facilities rather than full detail on each individual facility, the Facility Index schema can support this need with a lighter payload than the Facility Detail schema.

Data Origination and Unique Facility Identification

As partners have the ability to exchange and re-exchange facility data between partners, it has become apparent that there is a need to better expose the original source/steward of the exchanged facility data. Furthermore, there is a growing need to be able to uniquely identify each facility across the network, regardless of its origin. To achieve these two goals, three data elements are used to define where the data originated from, and how the originator uniquely identifies that facility. These elements are: Originating Partner Name, Information System Acronym Name, and Facility Site Identifier.

Geographic Location

Numerous enhancements have been made for the specification of geographic locations. The previous version allowed for specifying only one geographic coordinate for a facility, and the geometric type was limited to a single point (i.e., latitude/longitude). In the new version, locations can be specified as points, lines, “envelopes” (i.e., bounding box), and polygons, using the GeoRSS standard (<http://georss.org>). In addition, multiple geographic attributes can now be specified for a single facility. For example, a facility may have a series of points of interest as well as a pipeline specified as a line and a bounding area specified as a polygon. Finally, a facility can specify a primary facility location separately from the many other related geographic attributes.

As a result of moving to the new GeoRSS standard, certain elements are now expressed differently. For example, Latitude and Longitude are now specified in one element, POS (e.g., “45.523875 -122.670399”).

Alternative Names and Identifiers

In the previous version of the schema, a facility could be assigned only one alternative name (this was a flow in the schema) and no alternative identifiers. In the new version, a facility can have an unlimited number of alternative names and identifiers.

Electronic Addresses

The previous version supported electronic addresses in the form of email addresses for individuals and organization. In the 3.0 version, an electronic address of type Email, Internet, Intranet, HTTP, FTP, or Telnet can be assigned to individuals, organizations, facilities and/or environmental interests. This allows for facilities to share a web address pointing to additional information about a facility or environmental interest (e.g., hosted on the originating data source’s web site).

Affiliations

Previously, details about an organization or individual were replicated for each relationship (i.e., affiliation) an organization or individual had to a facility or environmental interest. This resulted in significant duplication of information when, as is often the case, one person served many roles at the same facility and was affiliated with many environmental interests.

In the new version, details about an organization or individual are stored in one area of the schema, while its relationship between the facility and its environmental interests are stored in a separate area. This normalization of affiliation data reduces the requirement for data replication and, as a result, reduces the overall payload of the FACID XML file while retaining all of the same information as in the 2.3 version.

Facility Details Overview

The Facility Details component provides the ability to exchange detailed information about facility sites as well as any associated environmental interests and affiliates. This component is optimized for exchanging detailed facility information for a specific user query or for database synchronization between exchange partners.

The following diagram describes the logical components contained in the Facility Details component of the Facility ID XML Schema:



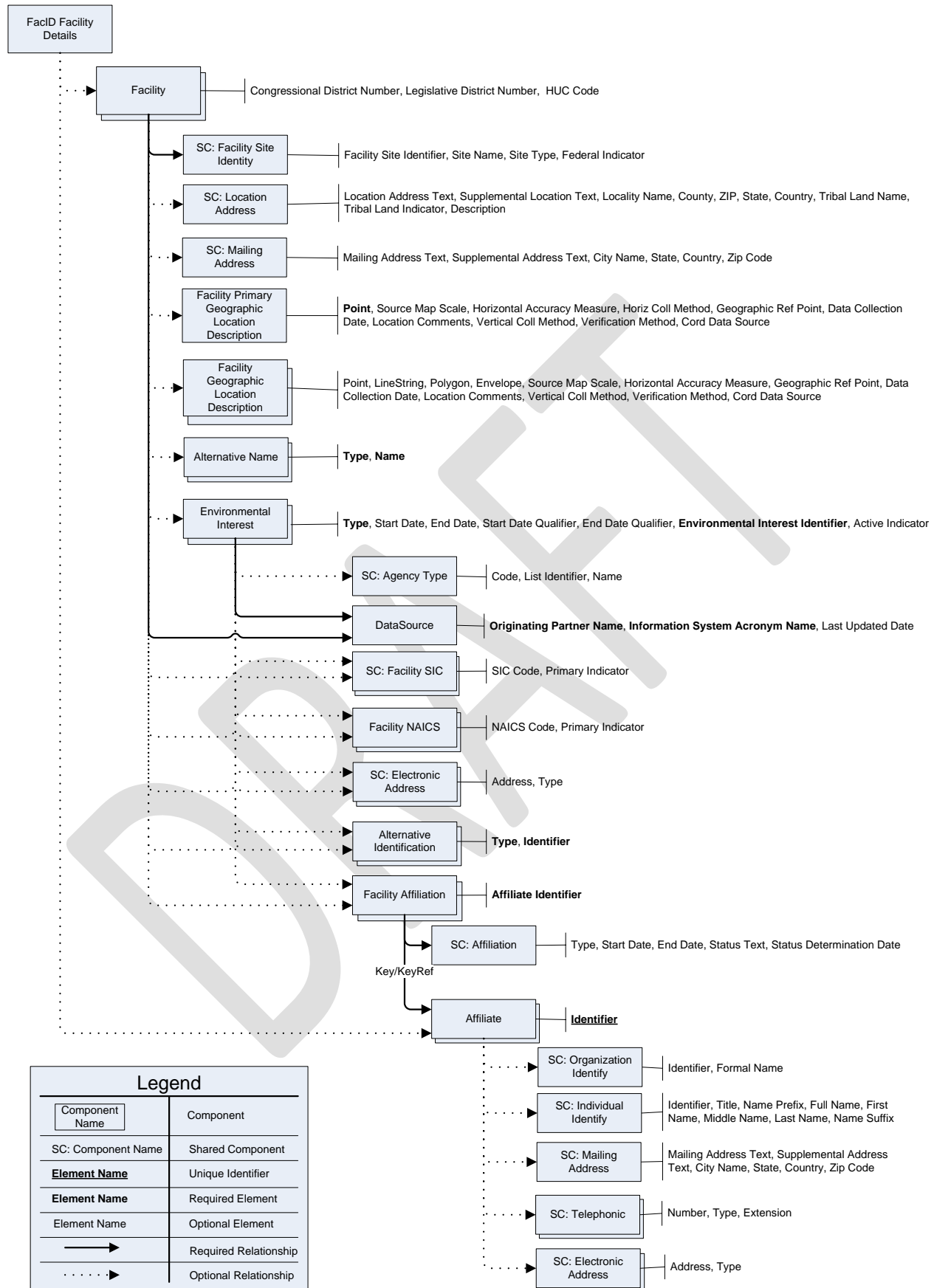


Figure 1: Logical Representation of the Facility Details Component

It is worth noting that the Facility ID schema leverages Exchange Network Shared Components (SCs) that are defined in the NEIEN Core Reference Model (CRM), as well as components from the GeoRSS Application Profile (<http://georss.org>). Where available, and appropriate, the SCs were implemented in this exchange.

For the purposes of describing the Facility Details component, this section is divided into modules, based on logical groupings of the information. Each Module is described in detail in the following sections.

For detailed data element information, please refer to the Facility ID Data Exchange Template (DET) spreadsheet.

Note: The root of the Facility Details Component is FACID_FacilityDetails_v3.0.xsd .

Facility Details Module

The Facility Details element is the root element of the Facility Details component of the Facility ID schema. The Facility Details data block serves as a container for zero or more Facilities and Affiliates.

Many partners who exchange information manage Affiliates using a consolidated warehouse system. This results in unique affiliates being shared across facility sites and/or environmental interests. For this reason, Affiliates have been normalized in the Facility ID exchange and no longer require the same Affiliate to be replicated for each relationship to a Facility or Environmental Interest. Using this approach helps to minimize the amount of information that is exchanged using this schema. In order to support this functionality, the Affiliate component can now be found as an element within this root element.

The following diagram describes the Facility Details data block contained in the Facility ID XML Schema:

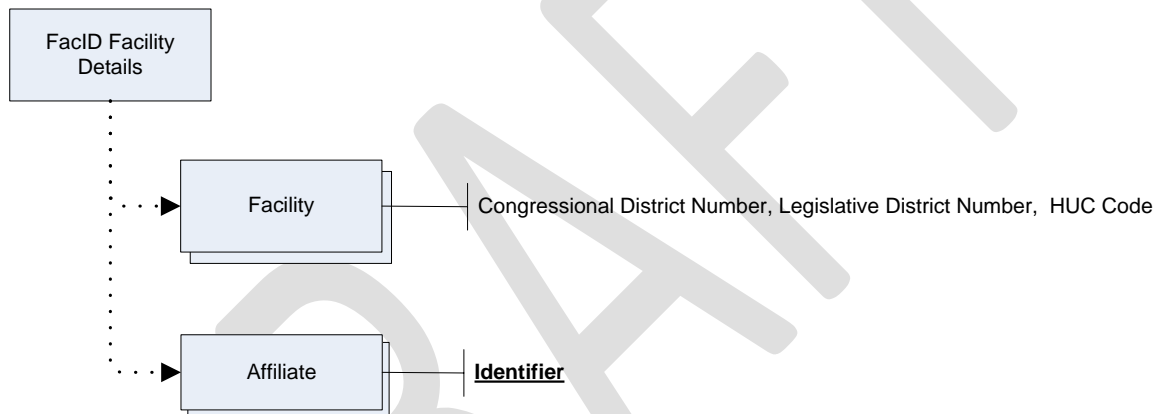


Figure 2: The Facility Details Data Block

Note: The Facility element of this exchange is considered optional in order to enable providers to return pertinent error messages, if an issue were to arise during the exchange of information.

Facility Module

The Facility data block contains data for a single facility site. This data block can contain information about the facility site including; the source of the facility site information, facility site identification information, a primary geographic location for the facility, many related geographic locations, a location address, a mailing address, many alternative names, many environmental interests, many SIC codes, many NAICS codes, many electronic addresses (email address, web address, etc.), many alternative ID's and many affiliations.

The following diagram describes the Facility data block contained in the Facility ID XML Schema:

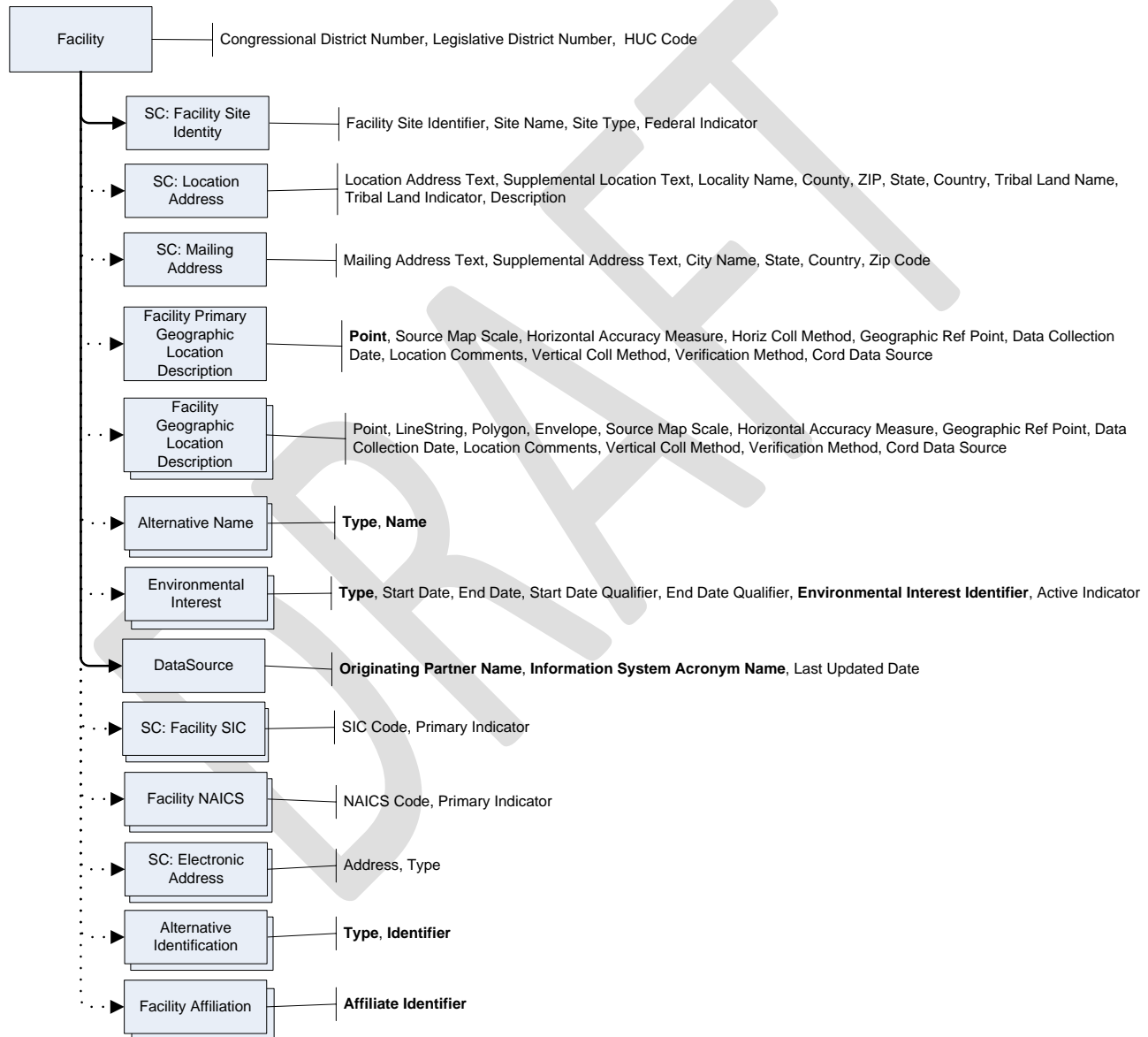


Figure 3: The Facility Data Block

Many of the components that are available for the Facility module are also available in the Environmental Interest module. The schema is structured in this manner due to the fact that some partners manage facility site information at the facility site level, while others manage facility site information at the environmental interest level. By replicating these components in both modules,

partners who use either facility site management approach can exchange information utilizing the schema.

There are now two components for describing geographic location. Facility Primary Geographic Location Description is used to specify a location intended to represent the entire facility. It must contain a Point (i.e., a latitude/longitude) and can additionally contain locational metadata. Facility Geographic Location Description is used to specify one or more geographic locations related to the facility. Each location can be defined for a Point, LineString, Polygon, or Envelope (i.e., box).

The DataSource component in this module is used to specify the partner who originally provided the facility site information that is being exchanged. In addition, the information system that managed the data as well as the date the information was last updated will be specified in this component. These values are an important component in deciphering where information was originally attained, following an exchange. This is becoming a critical component of the exchange as partners are becoming surrogate providers for information that was originally provided by another partner. As a hypothetical example, the Kansas Department of Health and Environment (KDHE) provides facility site information to EPA on a regular basis. If KDHE were to subsequently solicit a list of all facility sites located in Kansas from EPA, some returned facility sites will represent the facility sites that KDHE provided to EPA in a previous exchange and some will represent facility sites that originated from EPA managed systems. For this reason, it is useful for KDHE to have exposure to the originating partner and information system that manages the information.

In order to effectively exchange facility site information, the combination of the Originating Partner Name and Information System Acronym Name element values in the DataSource component and the Facility Site Identifier element in the Facility Site Identity component must be unique. This uniqueness will allow partners to exchange and identify a distinct set of facility sites, which is a critical component of utilizing the exchanged information..

Partners can now specify one or more Alternative IDs for a facility site. This provides the ability to exchange alternative IDs that may help identify a facility site, such as an FRS ID.

The schema now supports the exchange multiple Electronic Addresses (web addresses, etc.). This functionality was added in order to allow partners to specify a web URL which can provide additional information pertaining to a facility site, if one is available.

Note: Partners can now specify multiple Alternative Names and Geographic Locations for a Facility.

Environmental Interest Module

The Environmental Interest data block contains data for a single Environmental Interest. This data block can contain information about the environmental interest including; the source of the environmental interest information, a environmental interest start and end dates, environmental interest identification information, many SIC codes, many NAICS codes, many electronic addresses (email address, web address, etc.), many alternative ID's and many affiliations.

The following diagram describes the Environmental Interest data block contained in the Facility ID XML Schema:

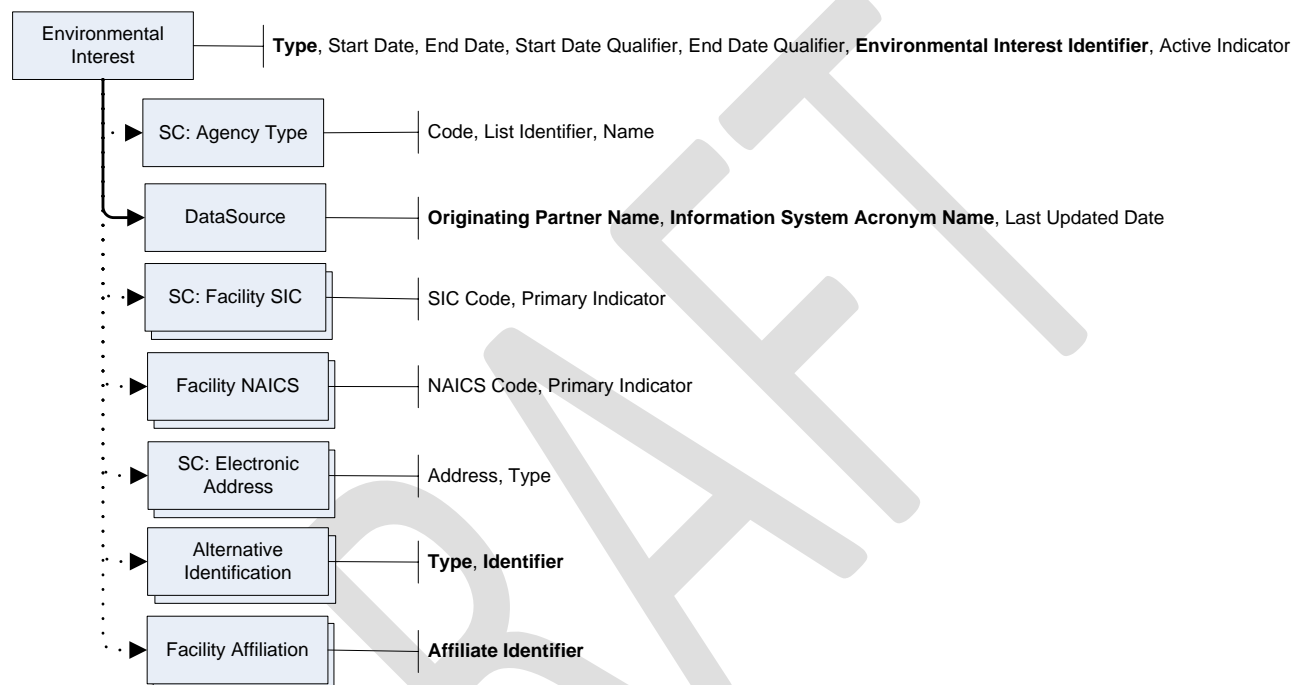


Figure 4: The Environmental Interest Data Block

As noted in the Facility Module section, many of the components that are available for the Facility module are also available in the Environmental Interest module. The schema is structured in this manner due to the fact that some partners manage facility site information at the facility level, while others manage facility site information at the environmental interest level. By replicating these components in both modules, partners who use either facility site management approach can exchange information utilizing the schema.

The Agency Type Shared Schema Component is available here to allow an information provider to specify the type of agency that manages the provided the information. This element has the flexibility to support the current types of agencies (Federal, State, Tribal, etc.) as well as supporting any future providers.

In order to effectively exchange environmental interest information, the combination of Originating Partner Name and Information System Acronym Name element values in the DataSource component and the Environmental Interest Identifier element in the Environmental Interest component must be unique. This uniqueness will allow partners to exchange and identify a distinct set of environmental interests, which is a critical component of utilizing the exchanged information.

The DataSource values (Originating Partner Name and Information System Acronym Name) must be provided to document the partner and partner system that provided the exchanged environmental information. It should also be noted that the DataSource value that is provided at the Environmental Interest level can differ from the value provided at the Facility level. This functionality is utilized when multiple source systems are utilized in the exchange. For example, if EPA were to provide facility site information to a partner, the Facility module information may come from the FRS system, while the Environmental Interest module information may come from auxiliary systems.

Partners can now specify one or more Alternative IDs for an environmental interest. This provides the ability to exchange alternative IDs that may help identify an environmental interest, such as a historical permit number.

The schema now supports the exchange multiple Electronic Addresses (web addresses, etc.). This functionality was added in order to allow partners to specify a web URL which can provide additional information pertaining to an environmental interest, if one is available.

Affiliation Module

The Affiliation data block serves as the container for an affiliation between a facility site and an affiliate or the affiliations between an environmental interest and an affiliate. This data block can contain information about the affiliation including; the unique identifier for the associated affiliate, the start and end date for the affiliation, the affiliation type, the affiliation status and information pertaining to the associated affiliate.

The following diagram describes the Affiliation data block contained in the Facility ID XML Schema:

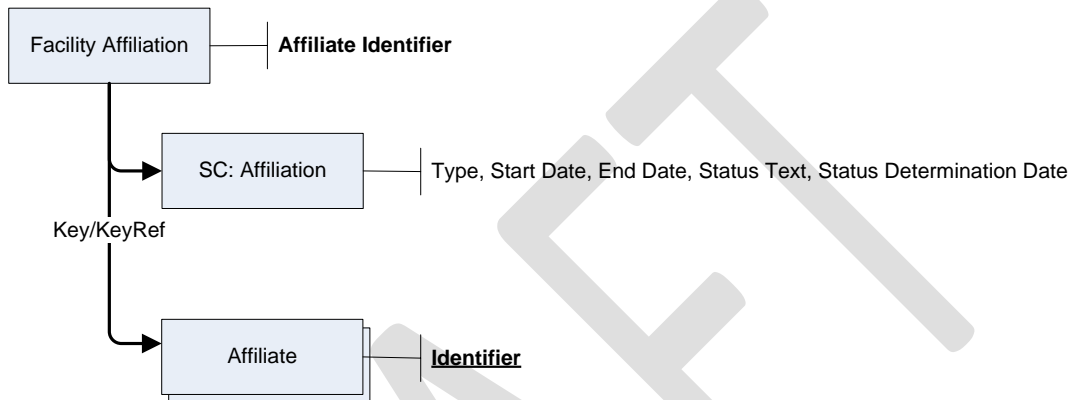


Figure 5: The Affiliation Data Block

As more partners manage affiliates using a consolidated warehouse system where unique affiliates are being shared across facility sites and/or environmental interests, the ability to exchange a consolidated set of affiliates has become more important. For this reason, Affiliates have been normalized in the Facility ID exchange and no longer require the same Affiliate to be replicated for each relationship to a Facility or Environmental Interest. The Affiliate component resides at the root level within this schema, but the diagram above shows the relationship from a Facility Affiliation to an Affiliate.

Note: In order to normalize the exchanged Affiliate information, this module utilizes the Key/KeyRef concept. This concept works similarly to the functionality provided by a Primary Key and Foreign key in a database environment, where the Key is the Primary Key and the KeyRef is the Foreign Key in the XML Schema environment. In this schema, the Affiliate Identifier element in the Facility Affiliation component is considered the KeyRef value and the Identifier element in the Affiliate component is considered the Key value.

Affiliate Module

The Affiliate data block serves as the container for information about an organization or an individual. This data block can contain information about the affiliate including; the unique identifier for the affiliate, organization identify information, individual identify information, a mailing address, many telephone numbers (phone, fax, etc.) and many electronic addresses (email address, web address, etc.).

The following diagram describes the Affiliate data block contained in the Facility ID XML Schema.

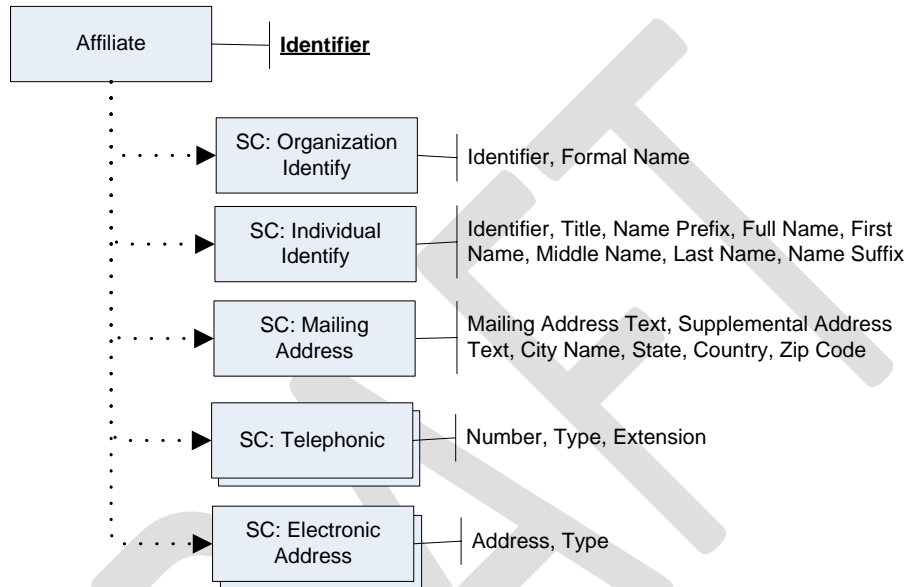


Figure 6: Affiliate Data Block

Note: The Identifier element in the Affiliate component is the Key value that will be used in the Key/KeyRef relationship to the Affiliation component.

Note: The provided Identifier value in the Affiliate component must be unique across all other Affiliates as this is the unique Key element for this component.

Note: An Affiliate must be considered either an organization or an individual. For this reason, only an Organization Identity or an Individual Identity can be provided, not both.

Facility Interest Overview

The Facility Interest component provides the ability to exchange an abbreviated set of information about facility sites as well as any associated environmental interests. This component is optimized for exchanging an abbreviated set of facility information for a specific user query or for database synchronization between exchange partners.

This component is structured in such a manner that the component has a minimal number of elements. This structure will support a simplified and condensed exchange of basic facility site and environmental interest information. This exchange can be utilized when comprehensive details for a facility site and the associated environmental interest are not required.

The following diagram describes the logical components contained in the Facility Interest component of the Facility ID XML Schema.

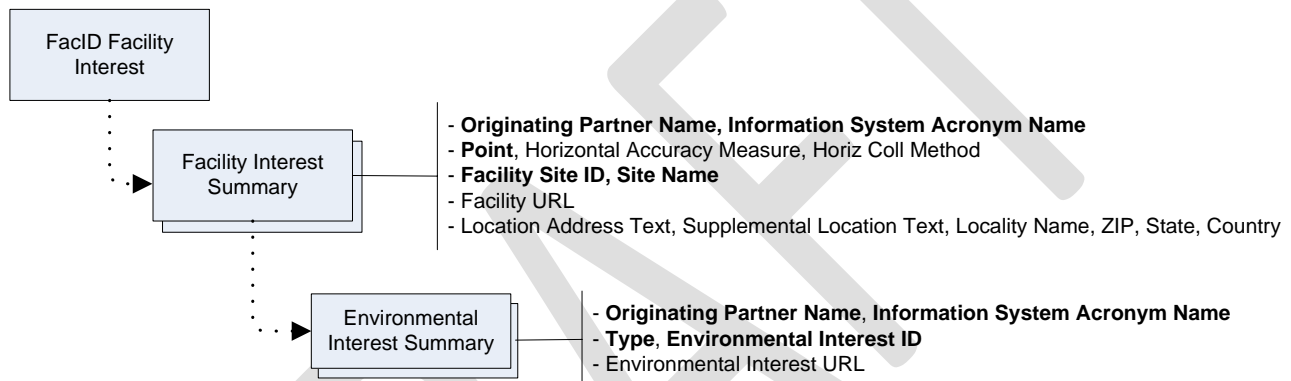


Figure 7: Logical Representation of the Facility Interest Component

For detailed data element information, please refer to the Facility ID Data Exchange Template (DET) spreadsheet.

Note: The root of the Facility Interest Component is **FACID_FacilityInterest_v3.0.xsd**.

Facility Index Overview

The Facility Index component provides the ability to exchange an abbreviated set of information about facility sites. This component is optimized for exchanging an abbreviated set of facility information for a specific user query or for database synchronization between exchange partners.

This component is structured in such a manner that the component has a minimal number of elements. This structure will support a simplified and condensed exchange of basic facility site information. This exchange can be utilized when comprehensive details for a facility site are not required.

The following diagram describes the logical components contained in the Facility Index component of the Facility ID XML Schema.

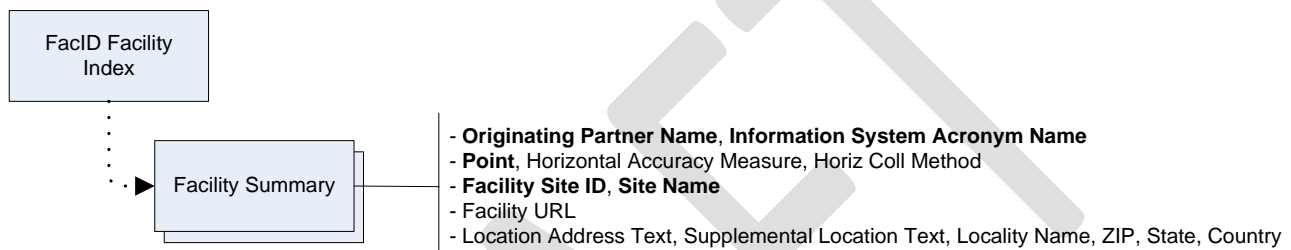


Figure 8: Logical Representation of the Facility Index Component

For detailed data element information, please refer to the Facility ID Data Exchange Template (DET) spreadsheet.

Note: The root of the Facility Index Component is **FACID_FacilityIndex_v3.0.xsd**.

Facility Count Overview

The Facility Count component contains only one element, FacilityCountMeasure, an integer value. This is used to support the GetFacilityCount_v3.0 data service.

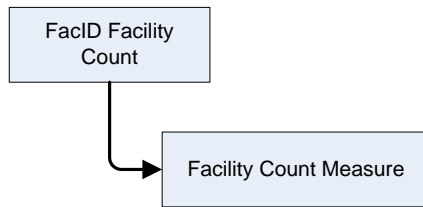


Figure 9: Logical Representation of the Facility Count Component

For detailed data element information, please refer to the Facility ID Data Exchange Template (DET) spreadsheet.

Note: The root of the Facility Count Component is **FACID_FacilityCount_v3.0.xsd**.