

Reporting of establishment information under the Seveso Directive (eSPIRS)

Data model documentation



Cover design: EEA

Cover photo: Jorge Franganillo, Creative Commons Attribution 2.0 Generic (<https://goo.gl/rqHYk6>)

Layout: EEA

Acknowledgments

This data model documentation was prepared by the [European Environment Agency](#) (EEA) in cooperation with the [European Topic Center for Health and Environment](#) (ETC/HE). The data model was produced using Unified Modelling Language (UML). Please contact seveso.helpdesk@eea.europa.eu for any questions, clarifications or comments.

Version control

Version ID	Description	Date
0.1	Internal draft for discussion with the EEA	October 2023
0.2	Internal draft addressing comments from the EEA	December 2023
0.3	Version for sharing with the Member State Reporters	February 2023

Contents

Acknowledgments	2
Version control	2
Contents	3
1 Introduction	4
1.1 The legal basis	4
1.2 The submission procedure	4
2 Seveso eSPIRS Data Model	5
2.1 Structure of the data model	5
2.2 Logic of the data model.....	5
2.3 Access to the data model	9
3 Detailed description: feature types, data types and code lists	10
3.1 Feature types.....	10
3.1.1 ContextualInformation.....	10
3.1.2 SevesoEstablishment.....	10
3.1.3 SevesoEstablishmentConfidentiality.....	11
3.1.4 ProductionSite.....	12
3.1.5 ProductionSiteConfidentiality	13
3.1.6 CompetentAuthority	Error! Bookmark not defined.
3.1.7 EstablishmentSubstances	13
3.2 Data types.....	13
3.2.1 Identifier.....	13
3.2.2 ThematicIdentifier	14
3.2.3 AddressDetails	14
3.2.4 NACEIndustry.....	14
3.2.5 SPIRSIndustry.....	14
3.2.6 Confidentiality.....	15
3.3 Code lists	16
Glossary of Terms	17

1 Introduction

This document describes the new data model that has been developed in order to handle the necessary administrative information for industrial facilities that European Member States (MS) and other reporting countries are required to report on Seveso eSPIRS (e-Seveso Plant Information Retrieval System) under the [Seveso-III Directive](#) (Directive 2012/18/EU). The resulting dataflow is hereafter referred to as the 'Seveso eSPIRS reporting'.

This document defines the scope and logic of the Seveso eSPIRS reporting dataflow. In addition to this document, two further related documents will be prepared which address other aspects of the data flow, namely the following:

- The quality assurance of the data submissions: A Quality Assurance / Quality Control (QA/QC) logic and,
- A Manual for Reporters, which will guide reporters on the practicalities of the reporting.

1.1 The legal basis

This data model document tackles the data reporting on entities within scope of the Seveso-III Directive and implemented through [Commission Implementing Decision 2022/1979](#) (CID).

1.2 The submission procedure

Countries will submit the Seveso eSPIRS dataflow, including administrative data on sites and establishments under the Seveso-III Directive to the EEA [Reportnet 3 platform](#) following a procedure that will be described in detail in a separate Manual for Reporters. Reportnet 3 provides an interface for interacting with the data in tabular form and natively supports data import in csv format. The EEA will harvest country data submissions to aggregate a European-wide dataset and release that dataset, excluding any confidential data, for analysis, public access, and support of environmental programs.

2 Seveso eSPIRS Data Model

2.1 Structure of the data model

The data model presented in this document provides an overview of all the data elements specified by the CID, including the data types, the obligatory and confidential status and their inter-relationships, with the objective of preparing countries to establish their data collection and management systems for the new reporting obligation. Indications on the data format that will be used to interact with Reportnet 3 will be presented in the separate document (Manual for Reporters).

The Seveso eSPIRS data model is made up of data fields (i.e. a single pieces of information), called attributes and is composed of three main elements:

1. **Feature types:** contains multiple attributes to collect information about a concept (e.g. *SevesoEstablishment*)
2. **Data types:** used when an attribute has several items of information (e.g. an address)
3. **Code lists:** a series of pre-defined values to standardise the information gathered in certain attributes (e.g. a status)

A code list is typically indicated via the appending of the text 'value' at the end of its name. A data type in turn may contain multiple attributes.

These elements are shown in Figure 1, Figure 2 and Figure 3.

Multiplicity is used to indicate if an attribute need be reported at all, in addition to how many times this attribute can be populated. Multiplicity in the data model reflects the optionality detailed in the legal basis (i.e. the Seveso-III Directive and the CID). As an example, multiplicity of [0..*] means a value for a certain attribute need not be reported at all, but it can be reported once, or reported multiple times. Reporting of attributes where no multiplicity is specified in the data model diagrams (Figure 1 and Figure 2) will be mandatory (if the multiplicity were to be listed for such situations this would be [1..1]).

2.2 Logic of the data model

ProductionSite and *SevesoEstablishment* will always need to be reported. Single geographic points must be provided for both these feature types in the data model. If an establishment contains substances that are hazardous according to the Seveso-III directive, then the substances data can be reported through the *EstablishmentSubstances* feature type, which collects information on name, quantity, and physical properties on the substances within scope of the Seveso-III Directive. Confidentiality is collected for these feature types through two additional confidentiality feature types (*ProductionSiteConfidentiality* and *SevesoEstablishmentConfidentiality*). The confidentiality for *EstablishmentSubstances* is embedded within the feature type.

Figure 1 Diagram of Seveso Reporting eSPIRS dataflow feature types

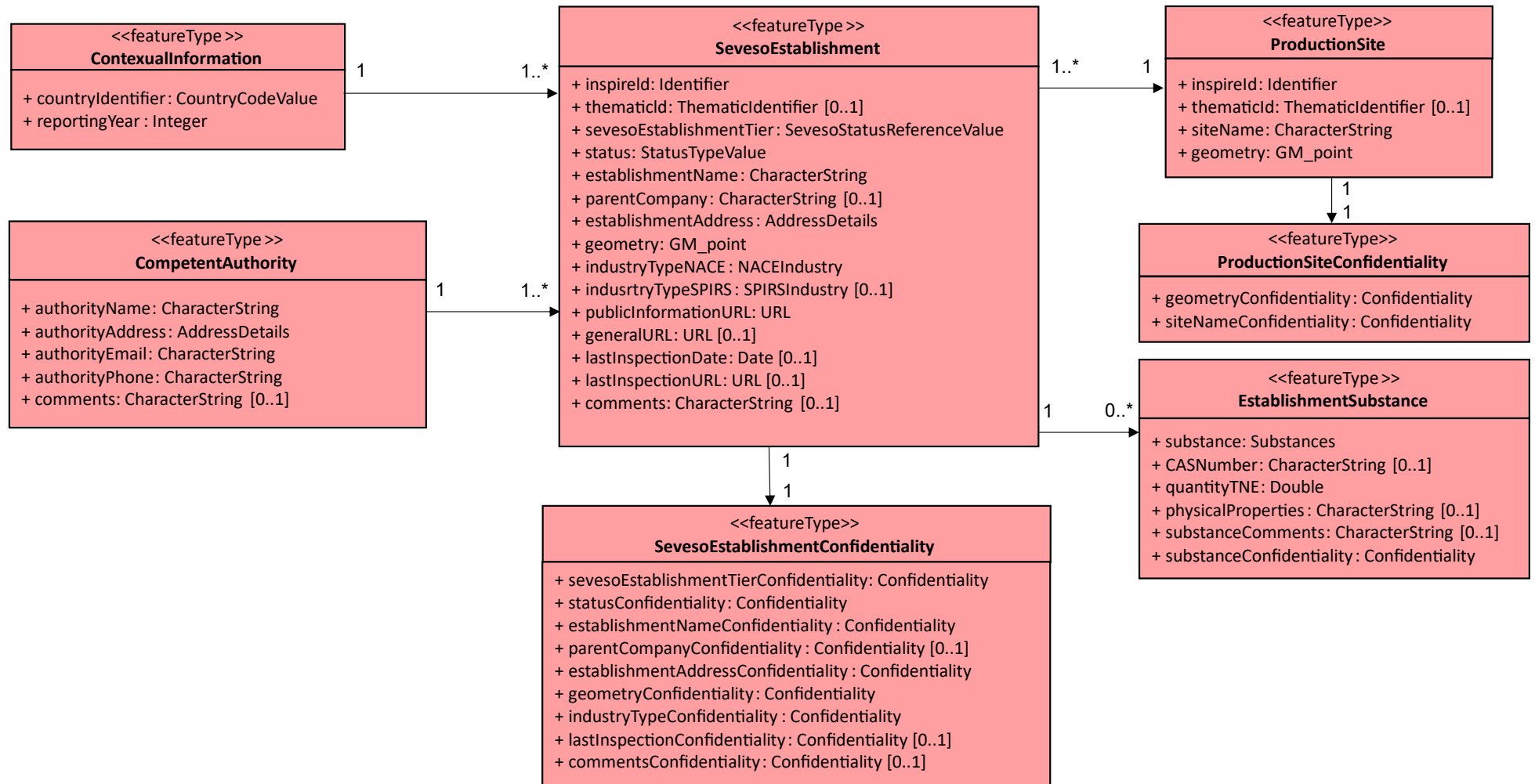


Figure 2 Diagram of Seveso Reporting eSPIRS dataflow data types

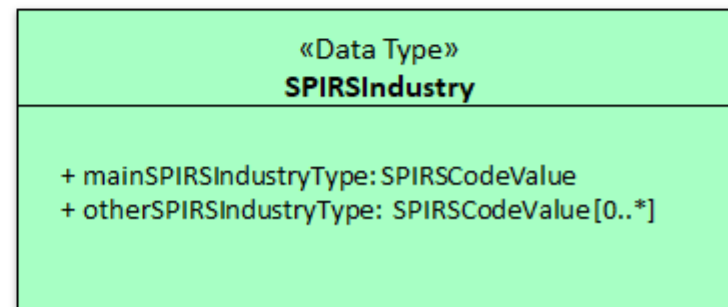
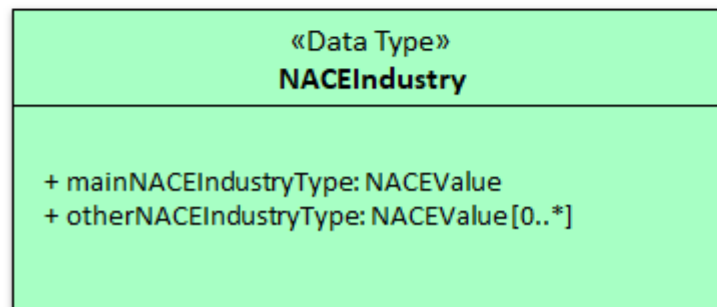
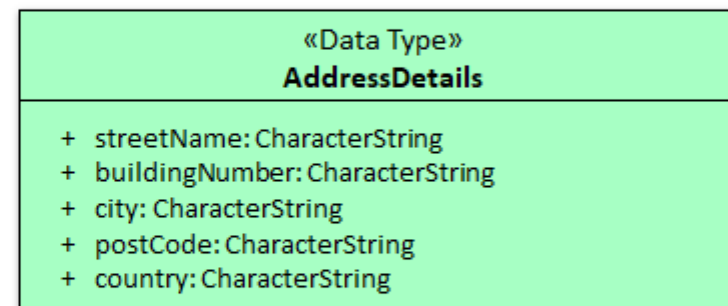
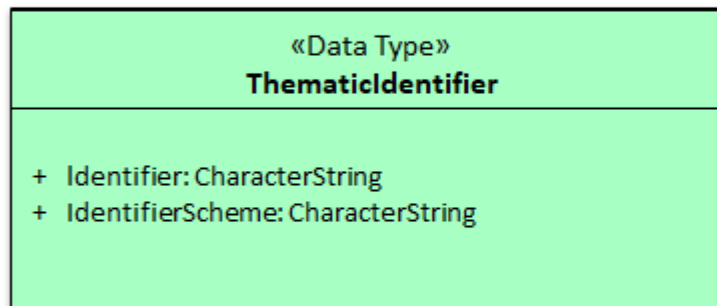
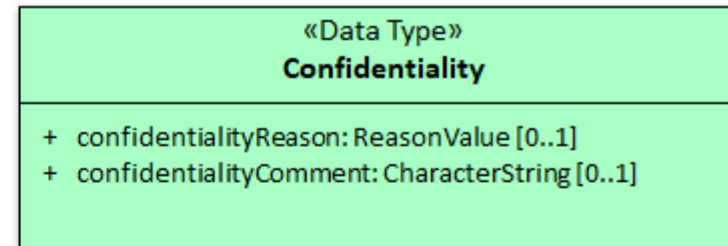
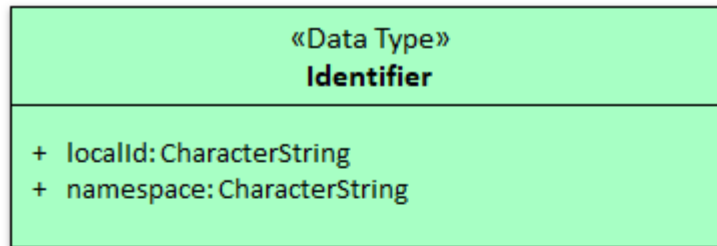
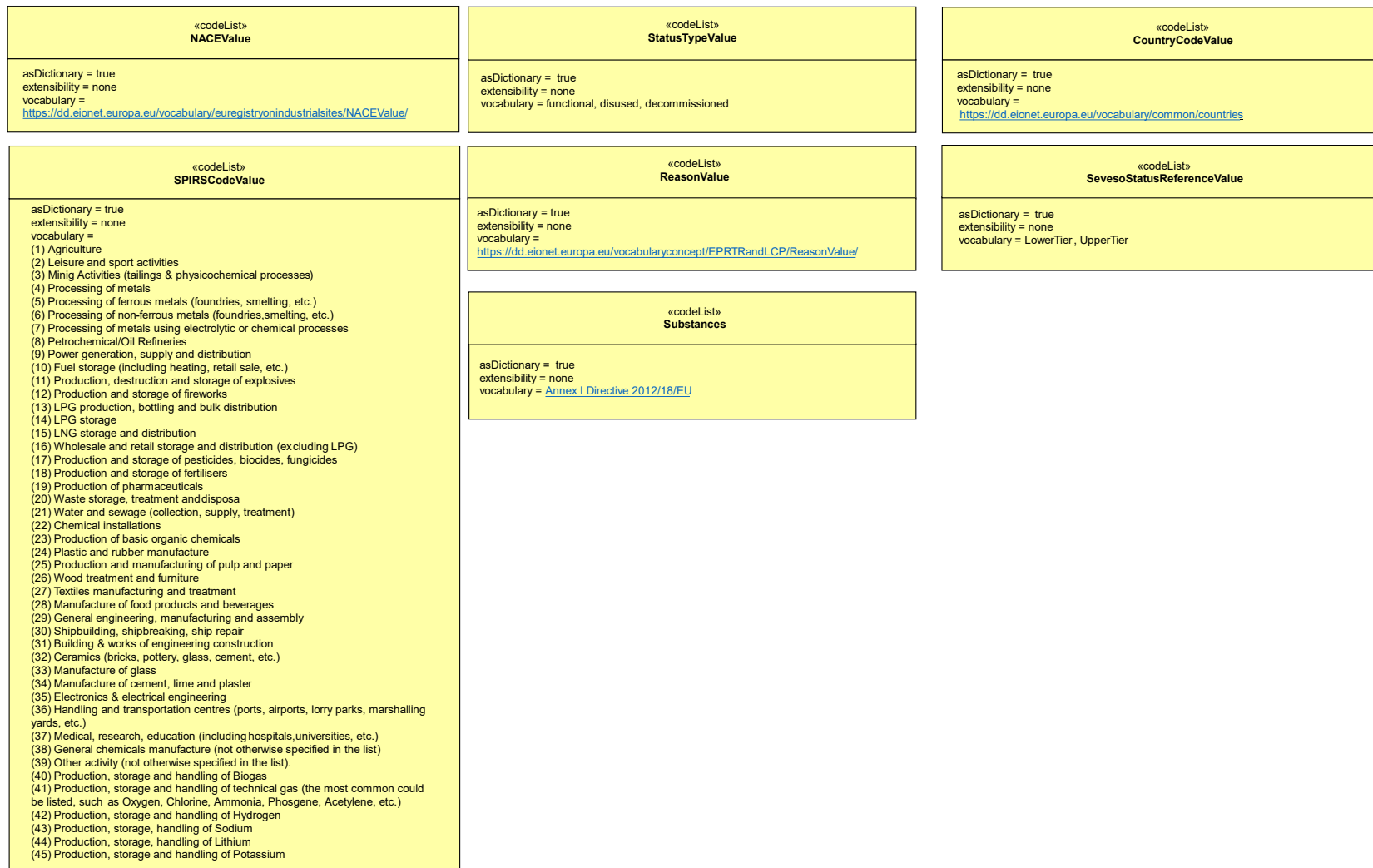


Figure 3 Diagram of Seveso Reporting eSPIRS dataflow code lists



2.3 Access to the data model

During the development phase, all materials related to the Seveso Reporting eSPIRS dataflow are published at the project website: <https://eionet.europa.eu/reportnet/seveso>.

This website will contain the latest version of the following documents as soon as they are ready:

- The Seveso Reporting eSPIRS data model (this document)
- The Manual for Reporters
- The QA/QC manual

When the Seveso Reporting eSPIRS dataflow progresses to an operational phase, all materials will be moved to a dedicated location on the reporting platform.

3 Detailed description: feature types, data types and code lists

This section systematically progresses through all attributes included in the feature types, the characteristics of data types where these are referred to in the feature types, and finally it refers to the code lists which are used across the data model.

3.1 Feature types

A feature type is the core entity in a Unified Modelling Language (UML) data model. It represents a class of data together with relevant attributes. The attributes may refer to data types or be populated with values from code lists or with numeric or text data.

3.1.1 ContextualInformation

The *ContextualInformation* feature type is designed to provide contextual information about both the reporting country and the reporting year for which a submission is made. These data are required in order to build a consistent database over time. The feature type consists of two attributes:

- **countryIdentifier**: This attribute is populated with values from the *CountryCodeValue* code list which contains a list of countries that will report eSPIRS data. This is automatically available in the reporting system.
- **reportingYear**: Defines the year that the submission covers (the year the data refers to, e.g. if the information is correct as of 2023, but reported in 2024, reportingYear should be 2023).

3.1.2 CompetentAuthority

This feature type is used to provide the name of the competent authority, a body responsible for reporting data under the Seveso Directive, as designated by the Member State. This contains five attributes to collect identification and contact information about this authority:

- **authorityName**: The name of the competent authority.
- **authorityAddress**: This attribute links to the *AddressDetails* data type, which contains all the necessary attributes to detail and address.
- **authorityEmail**: This attribute is a character string of the email address operated by the name detailed in the *authorityName* attribute. EEA encourages the use of functional mailboxes instead of individual's email accounts. This can be used in a generic sense and can refer to mailboxes for teams and departments within the authority. The email address provided should adhere to the common format and therefore indicate a functioning mailbox.
- **authorityPhone**: This attribute is a character string of the phone number for the authority referenced in the *authorityName* attribute. This should include all necessary area codes in order to be contacted from outside the reporting country.
- **comments**: A character string describing information regarding the competent authority. This attribute has a multiplicity of [0..1].

3.1.3 SevesoEstablishment

The *SevesoEstablishment* feature type is linked to the *ContextualInformation* and *CompetentAuthority* feature types with multiplicity of [1..*] which enables both the reporting of either one or multiple Seveso Establishments. It contains the following attributes (unless otherwise specified, they are mandatory with multiplicity [1..1]):

- **inspireId**: This crucial attribute links to the *Identifier* data type and is the fundamental building block that enables the effective sharing of geospatial environmental information and is the key to enabling the crosslinking of eSPIRS and eMARS reporting.
- **thematicId**: This attribute is to be used by reporting countries to provide an additional identifier specific to their own national reporting system. This attribute links to the *ThematicIdentifier* data type. This attribute is subject to [0..1] multiplicity so it does not need be populated.
- **sevesoEstablishmentTier**: This attribute is populated with the *SevesoStatusReferenceValue* code list and details whether the establishment is a lower or upper tier establishment according to Annex I of the Seveso-III Directive.
- **status**: This attribute is populated with the *StatusTypeValue* code list and details the operational status of the establishment.
- **establishmentName**: This attribute is a character string to provide the name of the Seveso establishment.
- **parentCompany**: This attribute is a character string to provide the name of the parent company. The attribute is subject to a [0..1] multiplicity.
- **establishmentAddress**: This attribute links to the *AddressDetails* data type which includes attributes to detail the address of the establishment.
- **geometry**: This attribute is populated with the point coordinates of the entity. The point coordinates should indicate the approximate centre of an entity. They are required using ETRS89 (2D) -EPSG:4258 coordinate reference system with a 10-meter accuracy (i.e., to five decimal places).
- **industryTypeNACE**: This attribute links to the *NACEIndustry* data type which includes attributes to detail the main and any other economic activities of an establishment, using the Eurostat NACE classification. The attribute is subject to [1..*] multiplicity, allowing for the reporting of more than one economic activity.
- **industryTypeSPIRS**: This attribute links to the *SPIRSIndustry* data type which includes attributes to detail the main and any other SPIRS industry types. The attribute is subject to [0..*] multiplicity, allowing for the reporting of no industry types or many.
- **publicInformationURL**: This attribute is to be completed with the URL where information referred to in Annex V of the Seveso-III Directive is made available to the public.
- **generalURL**: This attribute is to be completed with the URL to the parent company website. This attribute has a multiplicity of [0..1].
- **lastInspectionDate**: This attribute is to be completed with the date on which the establishment was last inspected. Actions such as site visits, checks of internal measures, systems and reports and follow-up documents, and any necessary follow-up, undertaken by or on behalf of the competent authority to check and promote compliance of establishments with the requirements of the Seveso-III Directive are considered as an inspection. This attribute has a multiplicity of [0..1].
- **lastInspectionURL**: This attribute is to be completed with the URL to the last inspection conclusions. This attribute has a multiplicity of [0..1].
- **comments**: A character string describing information regarding the reported Seveso establishment. This attribute has a multiplicity of [0..1].

3.1.4 SevesoEstablishmentConfidentiality

The *SevesoEstablishmentConfidentiality* feature type is designed to describe attribute confidentiality within the *SevesoEstablishment* feature type. Reporting countries decide if these attributes will be published by EEA in public data products. It contains the following attributes:

- **sevesoEstablishmentTierConfidentiality**: This attribute links to the **Confidentiality** data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the **sevesoEstablishmentTier** attribute confidential.
- **statusConfidentiality**: This attribute links to the **Confidentiality** data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the **status** attribute confidential.
- **establishmentNameConfidentiality**: This attribute links to the **Confidentiality** data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the **establishmentName** attribute confidential.
- **parentCompanyConfidentiality**: This attribute links to the **Confidentiality** data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the **parentCompany** attribute confidential. This attribute has multiplicity of [0..1].
- **establishmentAddressConfidentiality**: This attribute links to the **Confidentiality** data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the **establishmentAddress** attribute confidential.
- **geometryConfidentiality**: This attribute links to the **Confidentiality** data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the **geometry** attribute confidential.
- **industryTypeConfidentiality**: This attribute links to the **Confidentiality** data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the **industryTypeNACE** and **industryTypeSPIRS** attributes confidential. This attribute has multiplicity of [1..*].
- **lastInspectionConfidentiality**: This attribute links to the **Confidentiality** data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the **lastInspectionDate** and **lastInspectionURL** attribute confidential. This attribute has multiplicity of [0..1].
- **commentsConfidentiality**: This attribute links to the **Confidentiality** data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the **comments** attribute confidential. This attribute has multiplicity of [0..1].

3.1.5 ProductionSite

This feature type is linked to the parent **SevesoEstablishment** feature type. It is mandatory to report a Production Site for each Seveso Establishment. It has multiplicity of *..1 which enables more than one Seveso Establishment to link to the same Production Site. It contains the following attributes:

- **geometry**: This attribute is populated with the point coordinates of the entity. The point coordinates should indicate the approximate centre of an entity. They are required using ETRS89 (2D) -EPSG:4258 coordinate reference system with a 10-meter accuracy (i.e. to five decimal places).
- **inspireId**: This crucial attribute links to the **Identifier** data type and is the fundamental building block that enables the effective sharing of geospatial environmental information and is the key to enabling the crosslinking of the eSPIRS and eMARS reporting.
- **thematicId**: This attribute is to be used by reporting countries to provide an additional identifier specific to their own national reporting system. This attribute links to the **ThematicIdentifier** data type. This attribute is subject to [0..1] multiplicity so it does not need be populated.

- **siteName**: This attribute is a character string to provide the name of the production site.

3.1.6 ProductionSiteConfidentiality

The *ProductionSiteConfidentiality* feature type is designed to describe attribute confidentiality within the *ProductionSite* feature type. Reporting countries decide if these attributes will be published by EEA in public data products. It contains following attributes:

- **geometryConfidentiality**: This attribute links to the *Confidentiality* data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the *geometry* attribute confidential.
- **siteNameConfidentiality**: This attribute links to the *Confidentiality* data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the *siteName* attribute confidential.

3.1.7 EstablishmentSubstances

This feature type is linked to the parent *SevesoEstablishment* feature type. It is subject to [0..*] multiplicity which enables the reporting of multiple substances or none at all. It contains the following five attributes:

- **substance**: An attribute which is populated with values in the *Substances* code list.
- **CASNumber**: This attribute is a character string containing CAS Registry Number, a unique numeric identifier for substances. This attribute has multiplicity of [0..1].
- **quantityTNE**: An attribute containing a positive real number in a double-precision floating-point format representing the quantity of the substances, specified in tonnes triggering the Seveso status. This attribute is mandatory.
- **physicalProperties**: This attribute is a character string describing information regarding the physical properties of the substance. The attribute has a multiplicity of [0..1].
- **substanceComments**: This attribute is a character string describing information regarding the reporting establishment substances. This attribute has a multiplicity of [0..1].
- **substanceConfidentiality**: This attribute links to the *Confidentiality* data type, which contains all the necessary attributes to detail confidentiality and the associated reasons. This attribute enables reporting country to mark the information within the *substance* feature type confidential.

3.2 Data types

A data type is a UML data model element that defines both characteristics of data and which operations can be performed on the data. It will typically contain more than one attribute. The attributes can refer to other data types or be populated with code list values or with numeric or text data. Unless otherwise indicated, the data types listed below were created specifically for the Seveso Reporting eSPIRS data model.

3.2.1 Identifier

This is the standard INSPIRE data type designed to describe the Inspire ID for the entity being reported. Reporting countries will have the flexibility to use their existing national level identifiers provided they comply with the EU INSPIRE requirements for such IDs. When an INSPIRE Identifier is already available in the country system to identify the site or the establishment, reporters are expected to use the same identifier to designate the data in eSPIRS. This data type should be populated with two attributes:

- **localId**: A character string representing an identifier that is unique amongst all such entities in the reporting country.

- **namespace**: An attribute populated with a character string that defines the source of the data. It is recommended that the namespace starts with the two letter ISO country code, followed by the acronym of the reporting authority.

3.2.2 ThematicIdentifier

This is a data type used to provide an alternative identifier to the *InspireId* attribute (linked to the *Identifier* data type). The data type consists of two attributes:

- **Identifier**: This attribute is a character string of the thematic identifier. It has a multiplicity of [0..1].
- **IdentifierScheme**: This attribute is a character string denoting the scheme or system to which the identifier has relevance. It has a multiplicity of [0..1].

3.2.3 AddressDetails

A data type designed to collect address information. In the Seveso Reporting eSPIRS dataflow, this data type occurs in the *address* and *authorityAddress* attributes. It contains the following attributes:

- **streetName**: A character string populated with name of the street or road in which the building/property is located.
- **buildingNumber**: A character string populated with the building/property number or name. The character string format accounts for scenarios where a letter may be used in conjunction with a number (e.g. 67A).
- **city**: A character string populated with the name of the city where the building/property is located.
- **postcode**: A character string populated with the postcode of the building/property.
- **country**: A character string populated with the name of the country where the building/property is located.

3.2.4 NACEIndustry

A data type designed to collect information on the economic activity undertaken at the establishment. It contains two attributes:

- **mainNACEIndustryType**: This attribute is populated with values in the *NACEValue* code list. This should refer to the main economic activity occurring within the establishment.
- **otherNACEIndustryType**: This attribute is populated with the *NACEValue* code list. This should refer to other economic activities occurring within the establishment. The attribute is subject to [0..*] multiplicity, allowing for the reporting of multiple other activities or none for an establishment.

3.2.5 SPIRSIndustry

A data type designed to collect information regarding the SPIRS industry type reporting under the *SevesoEstablishment*. In the Seveso Reporting eSPIRS dataflow, this data type occurs in the *industryTypeSPIRS* attribute. It contains two attributes:

- **mainSPIRSIndustryType**: This attribute is populated with the *SPIRSCodeValue* code list. This attribute identifies the main SPIRS industry type of the establishment.
- **otherSPIRSIndustryType**: This attribute is populated with the *SPIRSCodeValue* code list. This attribute identifies the other SPIRS industry types of the establishment. The attribute is subject to [0..*] multiplicity, allowing for the reporting of multiple activities or none for an establishment.

3.2.6 Confidentiality

This is a data type designed to describe the confidentiality of attributes within feature type being reported. Reporting countries will have the flexibility to decide if the information reported for the attributes will be published by EEA in public data products. This data type contains three attributes:

- **confidentialityReason:** If an element should be confidential, this attribute should be populated with the values from the *ReasonValue* code list, which contains a list of reasons to protect environmental information from public disclosure, originating from Directive 2003/4/EC on public access to environmental information. Otherwise, it should be left empty. If this attribute is populated, then the data contained in the specific attribute within feature type will not be made public.
- **confidentialityComment:** A character string describing information regarding the confidentiality of the reporting field. This attribute has a multiplicity of [0..1].

3.3 Code lists

Unless otherwise indicated, all code lists were created specifically for the Seveso Reporting eSPIRS data model.

The location of the codelists within the EEA infrastructure in relation to the Reportnet 3 platform is under construction and will be specified in a future version of this document. Some of the generic code lists are stored and maintained at the so-called [EEA Data Dictionary](#).

- **NACEValue:** This is a list of current NACE (Statistical Classification of Economic Activities in the European Community) codes which allows compliance with the legislation that requires this field to be reported. <https://dd.eionet.europa.eu/vocabulary/euregistryonindustrialsites/NACEValue/>
- **SPIRSCodeValue:** This is a list of current SPIRS (Seveso Plants Information Retrieval System) code to indicate the industry type. The list of SPIRS codes is covered in the Annex of Commission Implementing Decision (EU) 2022/1979.
- **CountryCodeValue:** This code list includes Strict ISO-3166 two-character country codes of the countries where the reported establishments are located.
- **ReasonValue:** This list covers the different subsections of Article 4 Directive 2003/4/EC on public access to environmental Information. <https://dd.eionet.europa.eu/vocabulary/EPTRRandLCP/ReasonValue/>
- **Substances:** A list containing name and classifications of substances to be reported under [Annex I of Directive 2012/18/EU](#).
- **SevesoStatusReferenceValue:** This code list relates to the Seveso Establishment Tier Indication. It defines whether the reporting establishment belongs to a lower tier or upper tier according to Annex I to Directive 2012/18/EU.
- **StatusTypeValue:** This code list relates to the operational status of the establishment, which can either be functional, disused or decommissioned.

Glossary of Terms

This section contains explanations of the key terms used in this document.

Boolean attribute – These attributes are evaluated as either true or false.

Data type – A data model element that defines both characteristics of data and which operations can be performed on the data.

Feature type – Represents a class of data together with relevant attributes.

Geospatial information – Data about a physical object that can be represented by numeric values in a geographic coordinate system.

INSPIRE – Infrastructure for Spatial Information in the European Community. The INSPIRE Directive (2007/2/EC) aims to establish an infrastructure for the sharing of environmental spatial data within the European Union. This will enable sharing among public sector organisations, facilitate public access to spatial data across Europe, and will aid in cross-boundary policy making.

Multiplicity – A definition of cardinality - i.e. the permitted number of elements - of some collection of elements.

Production Site – Represents the geographical location of the Seveso Establishment or a piece of land where the facility was, is, or is intended to be located.

Seveso Establishment – Represents the whole location under the control of an operator where dangerous substances are present in one or more installations, including common or related infrastructures or activities according to DIRECTIVE 2012/18/EU.

UML – Unified Modelling Language, a modelling language intended to provide a standard way to visualise the design of a system.

Voidable – In data modelling, voidable means that whenever information does not exist then it does not have to be provided.