Developers' Guide to the IUCLID i6z Format
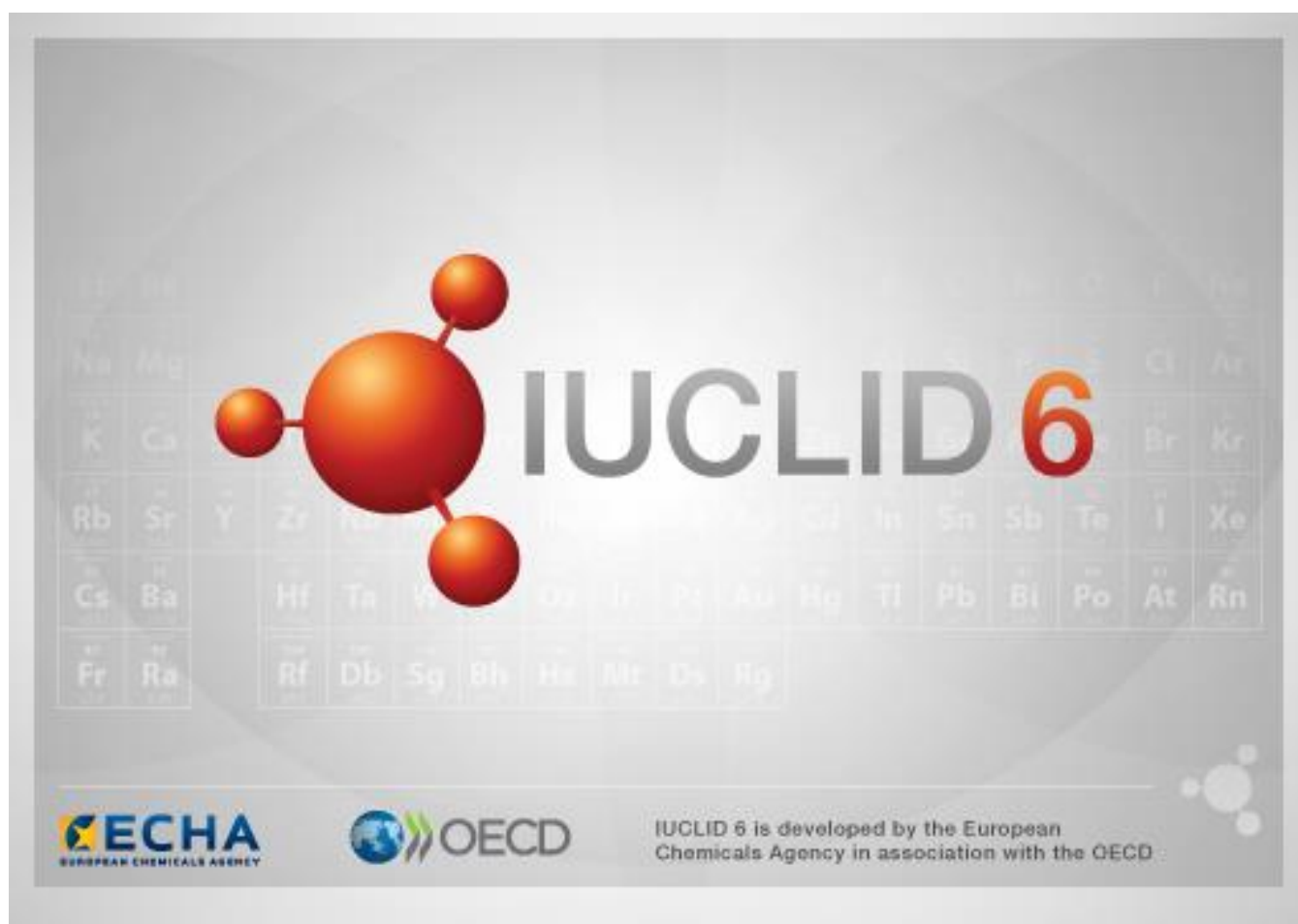
Changes to this document

Version	Changes	Date
1.0	1st version	06/02/2020
2.0	2nd version	25/11/2021
3.0	3rd version - Addition of an example for the PlatformMetada - Introduction of platform-container/v2	19/05/2023

Legal notice

The information in this document does not constitute legal advice. Usage of the information remains under the sole responsibility of the user. The European Chemicals Agency does not accept any liability with regard to the use that may be made of the information contained in this document.

IUCLID 6 is developed by the European Chemicals Agency in association with the OECD.

If you have questions or comments in relation to this document, please send them to ECHA via the information request form at the address below, quoting the reference and issue date given above:

https://echa.europa.eu/contact

IUCLID 6

European Chemicals Agency

Mailing address: P.O. Box 400, FI-00121 Helsinki, Finland

Visiting address: Telakkakatu 6, FI-00150 Helsinki, Finland

Table of Contents

Table of Figures

Table of Tables

1. Introduction

In IUCLID 6, the exchange of chemical information, from either datasets or dossiers, is facilitated via a zip/archive file that has the extension `i6z`, which stands for IUCLID 6 zip. Chemical information can be exported as an `i6z` file from an installation of IUCLID 6, and then imported into another. An `i6z` file has a well-defined and structured format that contains information on the IUCLID 6 entities, documents, and attachments it contains. The export feature of IUCLID 6 provides an advanced filtering mechanism that allows a user to select which of the interrelated entities are included in the archive.

2. Contents of an i6z file

This section contains a detailed description of the contents of an `i6z` archive file. An `i6z` file contains the following types of content:

- **manifest file (**`manifest.xml`**):** A single XML file providing top-level metadata of the archive and summarizing the archive's contents.
- **IUCLID document files (**`*.i6d`**):** IUCLID 6 XML files including information about the IUCLID 6 entities. By entities, we refer to:
  - **documents:** containing top-level information concerning the document along with the actual chemical information content;
  - **attachments:** containing metadata concerning the actual attachment content linked to the documents, either directly or via its enclosing fields.
- **IUCLID attachments:** The actual content in its original format (`docx`, `pdf`, etc.) that is attached to the documents and placed under the /attachments directory.
- **eXtensible Stylesheet Language files (**`*.xsl`**):** The files that describe how to display the XML content of IUCLID document into a human readable, HTML-based format via a web browser (optional).
- **Stylesheet (iuclid6_style.css):** A single `css` file to customize the presentation of the HTML elements (optional).

Each of the aforementioned types are further explained below. As a reference, please have a look at the following screenshot:

Figure 1: **Content of an i6z file**


```
<?xml version='1.0' encoding='UTF-8'?>
<?xml-stylesheet type="text/xsl" href="manifest.xsl"?>
<manifest xmlns="http://iuclid6.echa.europa.eu/namespaces/manifest/v1" xmlns:xlink="http://www.w3.org/1999/xlink">
    <general-information>
        <title>IUCLID 6 container manifest file</title>
        <created>Tue Feb 20 15:32:21 EET 2018</created>
        <author>User Super</author>
        <application>IUCLID6 (2.3.0, build of 09/02/2018 17:03)</application>
        <submission-type>R_INT_ONSITE</submission-type>
        <archive-type>RAW_DATA</archive-type>
        <legislations-info>
            <legislation>
                <id>core</id>
                <version>3.0</version>
            </legislation>
            <legislation>
                <id>domain</id>
                <version>3.0</version>
            </legislation>
        </legislations-info>
    </general-information>
    <comment>Remarks to be exported in the manifest file of the i6z</comment>
    <base-document-uuid>f5716c60-d07c-415c-b4c0-c1a796eaaf04/0</base-document-uuid>
    <contained-documents>
        <document id="f5716c60-d07c-415c-b4c0-c1a796eaaf04/0">
        <document id="4f88bc7f-395c-4d0b-997b-14e8c9aef605/0">
        <document id="8e257611-e925-3ae3-8ad9-63f29ba34a71/0">
        <document id="1e991422-239c-4b49-8a42-3f4730aa51a0/0">
        <document id="b61cc3b9-ca0e-35cc-8ed4-8284448a9716/0">
        <document id="ECB5-2898232d-99d1-4db6-8eb9-aeb53444b1df/0">
        <document id="6463936f-3354-371c-89d6-a1a1787159d4/0">
        <attachment id="11e18000-3093-4cb4-92cf-8dfc6dc81778/0">
    </contained-documents>
</manifest>
```
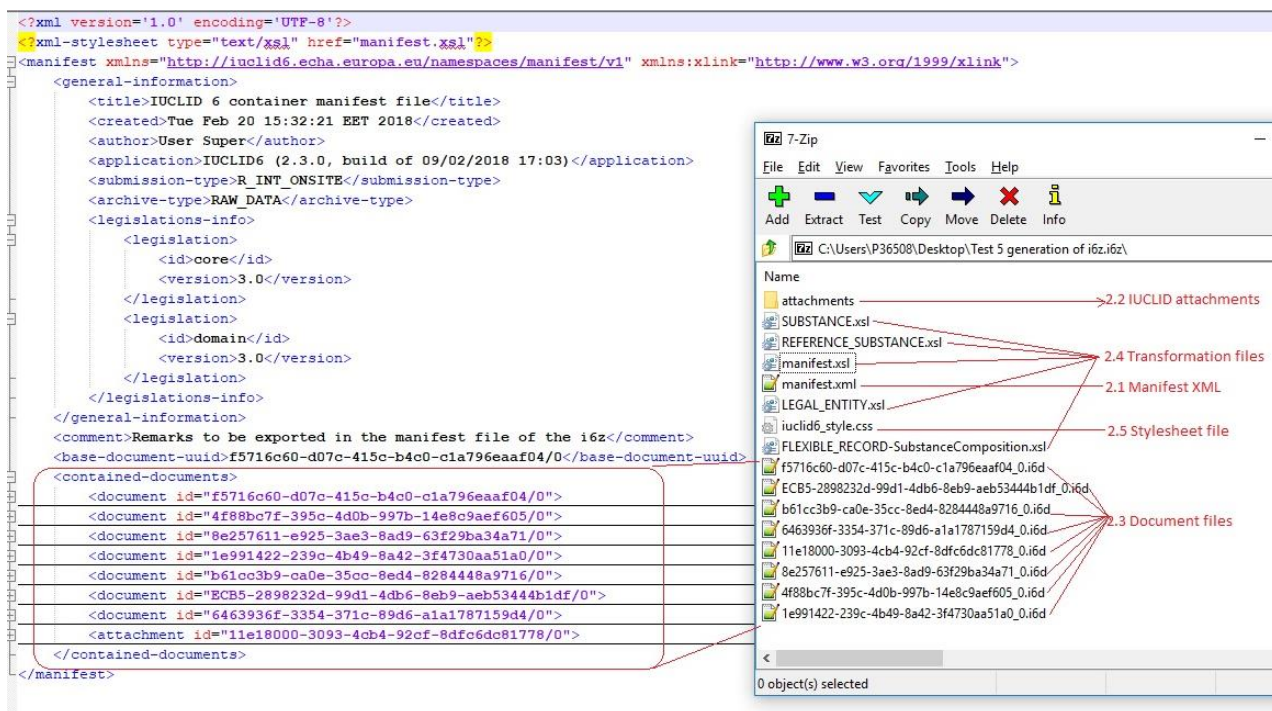
i6z: Mandatory contents

Exporting a dataset/dossier from IUCLID 6 results in an `i6z` archive that contains all the aforementioned types. However, the import operation does not require the existence of the `xsl` and `css` files. In other words, these files are optional, and can therefore be omitted from the `i6z` file.

2.1. Manifest

A IUCLID 6 archive, i.e. `i6z` file, has a manifest file which describes its contents. The file `manifest.xml` contains a table of contents of all the data files that are available in the `i6z` file. It is important to notice what is included in the manifest, and how the information is organised:

- **general-information:** section that describes the top-level information of the container file;
- **comment:** optional field that contains the user-defined text during the export operation of the specific dataset/dossier;
- **base-document-uuid:** The key of the document, the export operation was initiated from; this document is included in the archive and its metadata should appear under contained-documents section.
- **contained-documents:** section listing the metadata of all the documents and attachments included in the archive.

2.1.1. Sample of manifest.xml

Manifest xml file for a Substance dataset
<?xml version='1.0' encoding='UTF-8'?>

Manifest xml file for a Substance dataset

```xml
<?xml-stylesheet type="text/xsl" href="manifest.xsl"?>
<manifest xmlns="http://iuclid6.echa.europa.eu/namespaces/manifest/v1"
xmlns:xlink="http://www.w3.org/1999/xlink">
   <general-information>
      <title>IUCLID 6 container manifest file</title>
      <created>Tue Feb 20 15:32:21 EET 2018</created>
      <author>User Super</author>
      <application>IUCLID6 (2.3.0, build of 09/02/2018 17:03)</application>
      <submission-type>R_INT_ONSITE</submission-type>
      <archive-type>RAW_DATA</archive-type>
      <legislations-info>
         <legislation>
            <id>core</id>
            <version>3.0</version>
         </legislation>
         <legislation>
            <id>domain</id>
            <version>3.0</version>
         </legislation>
      </legislations-info>
   </general-information>
   <comment>Remarks to be exported in the manifest file of the i6z</comment>
   <base-document-uuid>f5716c60-d07c-415c-b4c0-c1a796eaaf04/0</base-document-uuid>
   <contained-documents>
      <document id="f5716c60-d07c-415c-b4c0-c1a796eaaf04/0">
         <type>SUBSTANCE</type>
         <name xlink:type="simple" xlink:href="f5716c60-d07c-415c-b4c0-c1a796eaaf04_0.i6d">Test substance for checking the i6z</name>
         <first-modification-date>2018-02-20T09:09:17Z</first-modification-date>
         <last-modification-date>2018-02-20T11:16:53Z</last-modification-date>
         <uuid>f5716c60-d07c-415c-b4c0-c1a796eaaf04/0</uuid>
         <links>
            <link>
               <ref-uuid>4f88bc7f-395c-4d0b-997b-14e8c9aef605/0</ref-uuid>
               <ref-type>REQUIRED_LEGAL_ENTITY</ref-type>
            </link>
            <link>
               <ref-uuid>4f88bc7f-395c-4d0b-997b-14e8c9aef605/0</ref-uuid>
               <ref-type>REFERENCE</ref-type>
            </link>
            <link>
               <ref-uuid>8e257611-e925-3ae3-8ad9-63f29ba34a71/0</ref-uuid>
               <ref-type>REFERENCE</ref-type>
            </link>
         </links>
         <representation>
            <reference-substance>
```

IUCLID 6

Manifest xml file for a Substance dataset

```xml
            <name>25167-70-8      </name>
            <IUPAC-name>25167-70-8      </IUPAC-name>
            <CAS-number/>
            <inventory-number/>
         </reference-substance>
         <legal-entity>
            <name>Predefined Legal entity</name>
            <town/>
            <country/>
         </legal-entity>
      </representation>
   </document>
   <document id="4f88bc7f-395c-4d0b-997b-14e8c9aef605/0">
      <type>LEGAL_ENTITY</type>
      <name xlink:type="simple" xlink:href="4f88bc7f-395c-4d0b-997b-
14e8c9aef605_0.i6d">Predefined Legal entity</name>
      <first-modification-date>2017-12-18T10:01:44Z</first-modification-date>
      <last-modification-date>2017-11-15T11:19:39Z</last-modification-date>
      <uuid>4f88bc7f-395c-4d0b-997b-14e8c9aef605/0</uuid>
      <links/>
      <representation>
         <name>Predefined Legal entity</name>
         <town/>
         <country/>
      </representation>
   </document>
   <document id="8e257611-e925-3ae3-8ad9-63f29ba34a71/0">
      <type>REFERENCE_SUBSTANCE</type>
      <name xlink:type="simple" xlink:href="8e257611-e925-3ae3-8ad9-
63f29ba34a71_0.i6d">25167-70-8      </name>
      <first-modification-date>2018-01-12T15:26:44Z</first-modification-date>
      <last-modification-date>2012-12-05T13:16:44Z</last-modification-date>
      <uuid>8e257611-e925-3ae3-8ad9-63f29ba34a71/0</uuid>
      <links/>
      <representation>
         <name>25167-70-8      </name>
         <IUPAC-name>25167-70-8      </IUPAC-name>
         <CAS-number/>
         <inventory-number/>
      </representation>
   </document>
   <document id="1e991422-239c-4b49-8a42-3f4730aa51a0/0">
      <type>FLEXIBLE_RECORD</type>
      <subtype>SubstanceComposition</subtype>
      <name xlink:type="simple" xlink:href="1e991422-239c-4b49-8a42-
3f4730aa51a0_0.i6d">Composition.001</name>
      <first-modification-date>2018-02-20T09:46:31Z</first-modification-date>
      <last-modification-date>2018-02-20T09:47:12Z</last-modification-date>
```

IUCLID 6

Manifest xml file for a Substance dataset

```xml
        <uuid>1e991422-239c-4b49-8a42-3f4730aa51a0/0</uuid>
        <links>
          <link>
            <ref-uuid>f5716c60-d07c-415c-b4c0-c1a796eaaf04/0</ref-uuid>
            <ref-type>CHILD</ref-type>
          </link>
          <link>
            <ref-uuid>b61cc3b9-ca0e-35cc-8ed4-8284448a9716/0</ref-uuid>
            <ref-type>REFERENCE</ref-type>
          </link>
          <link>
            <ref-uuid>ECB5-2898232d-99d1-4db6-8eb9-aeb53444b1df/0</ref-uuid>
            <ref-type>REFERENCE</ref-type>
          </link>
          <link>
            <ref-uuid>6463936f-3354-371c-89d6-a1a1787159d4/0</ref-uuid>
            <ref-type>REFERENCE</ref-type>
          </link>
        </links>
        <representation>
          <parent>
            <type>SUBSTANCE</type>
            <name>Test substance for checking the i6z</name>
          </parent>
          <reference-substance>
            <name>25167-70-8    </name>
            <IUPAC-name>25167-70-8    </IUPAC-name>
            <CAS-number/>
            <inventory-number/>
          </reference-substance>
          <legal-entity>
            <name>Predefined Legal entity</name>
            <town/>
            <country/>
          </legal-entity>
        </representation>
      </document>
      <document id="b61cc3b9-ca0e-35cc-8ed4-8284448a9716/0">
        <type>REFERENCE_SUBSTANCE</type>
        <name xlink:type="simple" xlink:href="b61cc3b9-ca0e-35cc-8ed4-
8284448a9716_0.i6d">Automatically generated during migration to IUCLID 6, no data
available</name>
        <first-modification-date>2018-01-12T15:26:44Z</first-modification-date>
        <last-modification-date>2009-12-09T08:02:39Z</last-modification-date>
        <uuid>b61cc3b9-ca0e-35cc-8ed4-8284448a9716/0</uuid>
        <links/>
        <representation>
          <name>Automatically generated during migration to IUCLID 6, no data
```

Manifest xml file for a Substance dataset

```xml
available</name>
          <IUPAC-name>Automatically generated during migration to IUCLID 6, no data
available</IUPAC-name>
          <CAS-number/>
          <inventory-number/>
        </representation>
      </document>
      <document id="ECB5-2898232d-99d1-4db6-8eb9-aeb53444b1df/0">
        <type>REFERENCE_SUBSTANCE</type>
        <name xlink:type="simple" xlink:href="ECB5-2898232d-99d1-4db6-8eb9-
aeb53444b1df_0.i6d">2,4,4-trimethylpentene</name>
        <first-modification-date>2018-01-12T15:26:44Z</first-modification-date>
        <last-modification-date>2016-05-26T06:43:41Z</last-modification-date>
        <uuid>ECB5-2898232d-99d1-4db6-8eb9-aeb53444b1df/0</uuid>
        <links/>
        <representation>
          <name>2,4,4-trimethylpentene</name>
          <IUPAC-name>2,4,4-trimethylpentene</IUPAC-name>
          <CAS-number>25167-70-8</CAS-number>
          <inventory-number>246-690-9</inventory-number>
        </representation>
      </document>
      <document id="6463936f-3354-371c-89d6-a1a1787159d4/0">
        <type>REFERENCE_SUBSTANCE</type>
        <name xlink:type="simple" xlink:href="6463936f-3354-371c-89d6-
a1a1787159d4_0.i6d">Mixed xylenes (in hydrocarbons)</name>
        <first-modification-date>2018-01-12T15:26:44Z</first-modification-date>
        <last-modification-date>2015-03-04T09:04:48Z</last-modification-date>
        <uuid>6463936f-3354-371c-89d6-a1a1787159d4/0</uuid>
        <links/>
        <representation>
          <name>Mixed xylenes (in hydrocarbons)</name>
          <IUPAC-name>Mixed xylenes (in hydrocarbons)</IUPAC-name>
          <CAS-number/>
          <inventory-number/>
        </representation>
      </document>
      <attachment id="11e18000-3093-4cb4-92cf-8dfc6dc81778/0">
        <name xlink:type="simple" xlink:href="11e18000-3093-4cb4-92cf-
8dfc6dc81778_0.i6d">Balsamiq Mockups 3.lnk</name>
        <first-modification-date>2018-02-20T11:16:48Z</first-modification-date>
        <last-modification-date>2018-02-20T11:16:48Z</last-modification-date>
        <uuid>11e18000-3093-4cb4-92cf-8dfc6dc81778/0</uuid>
        <container-uuid>f5716c60-d07c-415c-b4c0-c1a796eaaf04/0</container-uuid>
        <linked-attachments>
          <linked-doc xlink:type="simple"
xlink:href="attachments/070818f5e12266e6ac74bbf799bc15ae.lnk">Balsamiq Mockups
3.lnk</linked-doc>
```

Manifest xml file for a Substance dataset

</linked-attachments>
      </attachment>
    </contained-documents>
</manifest>

2.1.2. <general-information> section

As already stated, this section contains the top-level information of the container file.

Table 1:   **XML elements in the manifest file under <general information>**

Element	Type	Description	Required
- title	- xs:string	- The name that should be displayed as a first element, once opening the XML file via a web browser. For the time being, it is hardcoded to "IUCLID 6 container manifest file"	- Yes
- created	- xs:string	- The creation date-time of the i6z file. - Format: EEE MMM dd HH:mm:ss z yyyy	- Yes
- author	- xs:string	- The IUCLID user (i.e. concatenation of "First name " and "Last name") that created the i6z file. Alternatively, if the file is generated automatically, the name of the system can be used.	- Yes
- application	- xs:string	- It is by default "IUCLID6" and in a parenthesis the release version and build date/time are mentioned. However, generically, this is the name of the application or software module, including its version, used to generate the file.	- Yes
- submission-type	- xs:string	- The submission type of the dataset: For the dossier it is extremely useful; For raw datasets it is less useful; anyhow, choosing a submission type prior to exporting a raw dataset is mandatory, hence this information will exist.	- Yes
- archive-type	- (RAW_DATA | DOSSIER_DATA | CHEMICAL_INVENTORY)	- Indicates the type of the documents included in the archive. Poison Centre and SCIP notifications have to be submitted as dossier, so DOSSIER_DATA should be used.	- Yes
- legislations-	- complexType: list(legislation)	- Contains an optional list of legislation elements. This section mentions information relevant to the IUCLID legislations; in this section, the	- Yes

IUCLID 6

Element	Type	Description	Required
info		legislations that provide the documents (found within the i6z file) are listed. For instance, in the example mentioned above, the documents found within the respective i6z file are provided by two IUCLID legislations: CORE and DOMAIN.	
- legislation	- complexType: sequence(id, version)	- The name and version of the legislation provider whose at least one document is included in the archive	- No
- id	- xs:string	- The name of the legislation provider	- Yes
- version	- xs:string	- The version of the legislation provider	- Yes

Below, a sample of the section <general-information>:

Example of <general-information> section
<general-information>   <title>IUCLID 6 container manifest file</title>   <created>Tue Feb 20 15:32:21 EET 2018</created>   <author>User Super</author>   <application>IUCLID6 (2.3.0, build of 09/02/2018 17:03)</application>   <submission-type>R_INT_ONSITE</submission-type>   <archive-type>RAW_DATA</archive-type>   <legislations-info>   <legislation>   <id>core</id>   <version>3.0</version>   </legislation>   <legislation>   <id>domain</id>   <version>3.0</version>   </legislation>   </legislations-info> </general-information>

2.1.3. <comment> section

This is an optional field that contains the user-defined text during the export operation of the specific dataset or dossier.

Example of <comments> section
<comment>Remarks to be exported in the manifest file of the i6z</comment>

2.1.4. <base-document-uuid> section

This field keeps the key of the document the export operation was initiated from. For additional information concerning the document key/identifier, see section *3.1 Document identification*. The following cases should be mentioned:

1. **Dossier export:** the key of the dossier header `<document_uuid>/<snapshot_uuid>`;

2. **Entity export:** the key of the entity (*Substance*, *Mixture/Product*, *Template*, *Legal entity*, *Reference substance*, etc.) to be exported;

3. **Document export:** the key of the section document the export operation initiated from.

Example of <base-document> section
<base-document-uuid>f5716c60-d07c-415c-b4c0-c1a796eaaf04/0</base-document-uuid>

2.1.5. <contained-documents> section

As already stated, this section lists the metadata of all the documents and attachments included in the archive. It consists of a list of the following two subsection elements:

1. **document:** This element contains the metadata and top-level information of all IUCLID 6 documents included in the archive;

2. **attachment:** This element contains the metadata of all the attachments included in the archive.

2.1.5.1. <document> subsection

Below is an analysis of the attributes and elements included in the subsection `<document>`.

Table 2:   XML elements in the manifest file under subsection <document> in <contained-documents>

Element/ Attribute	Type	Description	Require d
- @id	- xs:string	- The unique identifier of the document. See section *3.1 Document identification*.	- Ye s
- type	- xs:string - Eligible values are: • ANNOTATION • ARTICLE • CATEGORY • DOSSIER • FIXED_RECORD • FLEXIBLE_RECORD • ENDPOINT_STUDY_RECOR D • FLEXIBLE_SUMMARY • ENDPOINT_SUMMARY • ASSESSMENT_ENTITY • LEGAL_ENTITY	- The type of the `document`.	- Ye s

IUCLID 6

Element/ Attribute	Type	Description	Required
	• MIXTURE • REFERENCE_SUBSTANCE • SITE • CONTACT • LITERATURE • SUBSTANCE • TEMPLATE • TEST_MATERIAL_INFORMA TION • INVENTORY • CUSTOM_ENTITY • CUSTOM_SECTION		
- subt ype	- xs:string	- The subtype in case of section document. This information is not applicable for entity documents. `<type>.<subtype>` uniquely identify the section document type and represent the document definition identifier.	- No
- nam e	- xs:string with two required attributes: • xlink:type • xlink:href	- It is the name of the document as specified by the user.	- No
-	- xlink:type	- The xlink:type attribute of the element specifies the type of the link. Its value is always "simple"	-
-	- xlink:href	- The `xlink:href` attribute of the element specifies the name of the XML file (`i6d`) that contains information regarding the actual content of this document. - The filename of the i6d file by convention must respect the following format. "*documentUuid_dossierUuid.i6d*" -	-
- first - mod ifica tion -	- xs:dateTime	- The date-time in ISO-8601 that the document was created	- Ye s

IUCLID 6

Element/ Attribute	Type	Description	Required
date			
- last-mod ifica tion - date	- xs:dateTime	- The last modification date-time in ISO-8601 format of the document.	- Yes
- uuid	- xs:string	- The unique identifier of the document. - See section *3.1 Document identification*.	- Yes
- links	- linksType: list(link)	- It contains a list of the documents that are referenced from/linked to the given document. It is mandatory that all documents contained in the list to be present in the i6z archive file.	- No
- link	- complexType: sequence(ref-uuid, ref-type)	- A single link between two documents	- No
- ref-uuid	- xs:string	- The unique identifier of the document that is referenced from/linked to the given document	- Yes
- ref-type	- xs:string - Eligible values are: • PARENT • CHILD • REFERENCE • USES_TEMPLATE • REQUIRED_LEGAL_ENTITY • CATEGORY • ATTACHMENT • DOSSIER_SUBJECT • ANNOTATION	- Specifies the type of the relationship between the given and the linked document	- Yes
- Represen-tation	- A complex element containing a sequence of the following optional elements: • subject-type • last-name • first-name • organisation • parent • reference-substance • name • town	- Additional metadata per document needed to produce the document's representation in the IUCLID 6 Swing UI. - Not all elements are applicable for all document types. Only a subset of them are filled per case.	- No

Element/ Attribute	Type	Description	Require d
	• country • legal-entity • IUPAC-name • CAS-number • inventory-number • title • author • reference-type		

Manifest: mandatory contents

The elements `links` and `representation` are used to facilitate the generation of an archive summary without the need to read/parse any document xml file (`i6d`).

Below, an example of the subsection `<document>` under the section `<contained-documents>`.

Example of <document> section under <contained-documents>
| ```
<document id="1e991422-239c-4b49-8a42-3f4730aa51a0/0">
  <type>FLEXIBLE_RECORD</type>
  <subtype>SubstanceComposition</subtype>
  <name xlink:type="simple" xlink:href="1e991422-239c-4b49-8a42-3f4730aa51a0_0.i6d">Composition.001</name>
  <first-modification-date>2018-02-20T09:46:31Z</first-modification-date>
  <last-modification-date>2018-02-20T09:47:12Z</last-modification-date>
  <uuid>1e991422-239c-4b49-8a42-3f4730aa51a0/0</uuid>
  <links>
    <link>
      <ref-uuid>f5716c60-d07c-415c-b4c0-c1a796eaaf04/0</ref-uuid>
      <ref-type>CHILD</ref-type>
    </link>
    <link>
      <ref-uuid>b61cc3b9-ca0e-35cc-8ed4-8284448a9716/0</ref-uuid>
      <ref-type>REFERENCE</ref-type>
    </link>
    <link>
      <ref-uuid>ECB5-2898232d-99d1-4db6-8eb9-aeb53444b1df/0</ref-uuid>
      <ref-type>REFERENCE</ref-type>
    </link>
    <link>
      <ref-uuid>6463936f-3354-371c-89d6-a1a1787159d4/0</ref-uuid>
      <ref-type>REFERENCE</ref-type>
    </link>
  </links>
  <representation>
``` |

Example of <document> section under <contained-documents>

```
    <parent>
      <type>SUBSTANCE</type>
      <name>Test substance for checking the i6z</name>
    </parent>
    <reference-substance>
      <name>25167-70-8</name>
      <IUPAC-name>25167-70-8</IUPAC-name>
      <CAS-number/>
      <inventory-number/>
    </reference-substance>
    <legal-entity>
      <name>Predefined Legal entity</name>
      <town/>
      <country/>
    </legal-entity>
  </representation>
</document>
```

2.1.5.2. <attachment> subsection

Below is an analysis of the attributes and elements included in the attachment subsection.

Table 3:   XML elements in the manifest file under subsection <attachment> in <contained-documents>

Element/Attribute	Type	Description	Required
- @id	- xs:string	- The unique identifier of the attachment. See section *3.1 Document identification*.	- Yes
- name	- xs:string with two required attributes: • xlink:type • xlink:href	- The value of this element is the original filename of the attachment.	- No
	- xlink:type	- The xlink:type attribute of the element specifies the type of the link. It's value is always "simple"	
	- xlink:href	- The `xlink:href` attribute of the element specifies the name of the XML file (`i6d`) that contains information regarding the actual content of this attachment.	

				- The filename of the i6d file by convention must respect the following format "*attachmentUuid_dossierUuid.i6d*"		
- first-modification-date		- xs:dateTime		- The date-time in ISO-8601 format that the attachment was created.		- Yes
- last-modification-date		- xs:dateTime		- The last modification date-time in ISO-8601 format of the attachment.		- Yes
- uuid		- xs:string		- The unique identifier of the attachment. See section *3.1 Document identification*.		- Yes
- container-uuid		- xs:string		- The unique identifier of the document to which the attachment is linked.		- Yes
- linked-attachments		- list(linked-doc)		-		- No
- linked-doc		- xs:string with two required attributes: • xlink:type • xlink:href		- The value of this element is the name of the attachment. -		- No
		- xlink:type		- The xlink:type attribute of the element specifies the type of the link. It's value is always "simple"		
		- xlink:href		- The `xlink:href` attribute specifies the actual attachment content file under the attachments directory. The file name of the attachment must be the MD5 hash of the attachment file, with the same extension as the file. The file name is case insensitive.		

Below is an example of the subsection `<attachment>` under the section `<contained-documents>`.

Example of <attachment> section under <contained-documents>

```
<attachment id="ECB5-d3a3b15d-d83f-4b21-acc9-001c31505c6d/0">
  <name xlink:type="simple" xlink:href="ECB5-d3a3b15d-d83f-4b21-acc9-001c31505c6d_0.i6d">110-54-3-
```

```
V2.jpeg</name>
   <first-modification-date>2017-10-06T13:08:00Z</first-modification-date>
   <last-modification-date>2017-10-06T13:08:00Z</last-modification-date>
   <uuid>ECB5-d3a3b15d-d83f-4b21-acc9-001c31505c6d/0</uuid>
   <container-uuid>ECB5-27a8cb0b-5cb5-4ab8-ac18-4eab59415f7f/0</container-uuid>
   <linked-attachments>
      <linked-doc xlink:type="simple" xlink:href="attachments/e3cde6b1d8fba4ac4ce414ebe9610ccf.jpeg">110-54-
3-V2.jpeg</linked-doc>
   </linked-attachments>
</attachment>
```

2.2. IUCLID Attachment files

An i6z file archive contains a separate XML (i6d) file for each attachment. The file name contains the ID of the attachment, as referred to in the manifest file. An example is given below:

Reference in `manifest.xml`

```
<attachment id="IUC5-cd3a12ce-8b98-4932-9c9b-7cbdc79c93bf/0">
```

Filename:

`IUC5-cd3a12ce-8b98-4932-9c9b-7cbdc79c93bf_0.i6d`

Below is a description of the attributes and elements included in the element `<Attachment>` inside an `i6d` file for an attachment.

Table 4:   XML elements in the i6d file for an attachment

Element/Attribute	Type	Description	Required
- documentKey	- xs:string	- The unique identifier of the attachment. See section *3.1 Document identification*.	- Yes
- name	- xs:string	- It is the name of the uploaded attachment	- Yes
- creationDate	- xs:dateTime	- The date-time in ISO-8601 format that the attachment was created	- Yes
- lastModificationDate	- xs:dateTime	- The last modification date-time in ISO-8601 format of the attachment	- Yes
- remarks	- xs:string	- The remarks provided by the user during the attachment uploading	- No
- md5	- xs:string	- The MD5 hash (either in lowercase or	- Yes

IUCLID 6

Element/Attribute	Type	Description	Required
		uppercase) of the uploaded attachment content	
- mimetype	- xs:string	- The media type of the attachment content	- No
- content	- xs:string with two required attributes: • xlink:type • xlink:href	- The value of this element is the name of the attachment. -	- No
	- xlink:type	- The xlink:type attribute of the element specifies the type of the link. It's value is always "simple"	
	- xlink:href	- The xlink:href attribute specifies the actual attachment content file under the attachments directory. The file name is the MD5 hash (either in lowercase or uppercase) of the uploaded attachment content with the extension of the actual file.	

Below is an example of the subsection `<attachment>` under the section `<contained-documents>`.

Example of <Attachment> in the i6d file for an attachment

```xml
<?xml version='1.0' encoding='UTF-8'?>
<Attachment xmlns="http://iuclid6.echa.europa.eu/namespaces/platform-attachment/v1"
xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <documentKey>ECB5-65104712-25bd-4366-bbd6-2d085a0c4eb3/0</documentKey>
  <name>108-88-3-V2.jpeg</name>
  <creationDate>2017-10-06T21:08:00.675+08:00</creationDate>
  <lastModificationDate>2017-10-06T21:08:00.675+08:00</lastModificationDate>
  <md5>8c930da8a8967b53dc73b919a2309c1b</md5>
  <mimetype>image/jpeg</mimetype>
  <content xlink:href="attachments/8c930da8a8967b53dc73b919a2309c1b.jpeg" xlink:type="simple"/>
</Attachment>
```

IUCLID 6

Please note that every IUCLID attachment i6d file must be referenced by at least one IUCLID document file, described below.

2.3. IUCLID Document files

Besides attachments, there are also XML-based documents (`i6d`) that contain top-level information concerning the IUCLID 6 document along with the document's actual chemical information content. The name of the `i6d` file is based on the document identifier that is `<document_uuid>_<snapshot_uuid>.i6d`. If the document is a raw document, and therefore not part of a dossier, the snapshot_uuid is 0.

Below is an analysis of the elements included in the element `<document>` within the document-related `i6d` files.

- **PlatformMetadata (cf.** `platform-metadata.xsd`**):** section that contains the top-level information of a IUCLID 6 document such as document identifier, name, type and subtype, etc;
- **Content (e.g.** `LEGAL_ENTITY-4.0.xsd`**):** section that contains the chemical information of the specific document;
- **Attachments (cf.** `platform-attachment.xsd`**):** section that lists the attachments that are directly linked to the document;
- **ModificationHistory (cf.** `platform-modification-history.xsd`**):** section that lists the entries of the document's modification history.

2.3.1. <PlatformMetadata> section

Table 5:   **XML elements under the section <PlatformMetadata> in an i6d file for a document**

Element	Type	Description	Require d
- iuclidVersi on	- xs:string	- The current IUCLID version used for exporting the i6z archive. - If another system is used to create the IUCLID file, it should indicate the version of IUCLID 6 that is compatible with the file generated: • For the first version of the PCN format, this value is '3.1.1'. • For the first version of the SCIP format and the latest version of the PCN format, this value is '4.2.1'.	- Y es
- document Key	- xs:string	- The unique identifier of the document. See section *3.1 Document identification*.	- Y es
- documentT ype	- xs:string - Eligible values are: - ANNOTATION - ARTICLE	- The type of the document	- Y es

IUCLID 6

Element	Type	Description	Required
	- CATEGORY - DOSSIER - FIXED_RECORD - FLEXIBLE_RECORD - ENDPOINT_STUDY_RECORD - FLEXIBLE_SUMMARY - ENDPOINT_SUMMARY - ASSESSMENT_ENTITY - LEGAL_ENTITY - MIXTURE - REFERENCE_SUBSTANCE - SITE - CONTACT - LITERATURE - SUBSTANCE - TEMPLATE - TEST_MATERIAL_INFORMATION - INVENTORY - CUSTOM_ENTITY - CUSTOM_SECTION		
- definitionVersion	- xs:string	- The definition version of the exported document. This value: • indicates that the content section follows the document format of the specified version' • during import operation, this value drives the resolution of the proper document's `xsd` to run the validation with.	- Yes
- creationDate	- xs:dateTime	- The date-time in ISO-8601 format that the document was created.	- Yes
- lastModific	- xs:dateTime	- The last modification date-	- Y

Element	Type	Description	Required
ationDate		time in ISO-8601 format of the document.	es
- name	- xs:string	- It is the name of the document as specified by the user. Usually the document name is a representation of the document content but it can also be another identifier chosen by the user.	- Yes
- document SubType	- xs:string	- The subtype in case of section document. This information is not applicable for entity documents. <type>.<subtype> uniquely identify the section document type and represent the document definition identifier.	- Yes
- parentDoc umentKey	- xs:string	- In case this document is a section document, this element keeps the unique identifier of its parent document. See section *3.1 Document identification*.	- Yes
- orderInSec tionNo	- xs:nonNegativ eInteger	- In case this is a section document, the order of the document with the specific definition identifier (type, subtype) under the provided parent entity	- Yes
- submission Type	- xs:string	- Applicable only for dossier archives. Indicates the submission type used during dossier generation. The value is specified in case the XML concerns: • the dossier document • the documents (ARTICLE/SUBSTANCE/MIXTURE) under the dossier with a submission type different than the one of the dossier document	- Yes
- submission TypeVersio n	- xs:string	- The version of the submission type used to generate the dossier	- Yes
- submitting LegalEntity	- xs:string	- Only for the dossier document: the legal entity document identifier that originated from the dossier creation	- Yes

Element	Type	Description	Required
- dossierSubject	- xs:string	- Only for the dossier document, it contains the document key (unique identifier) of the dossier subject document (ARTICLE, SUBSTANCE, MIXTURE, CATEGORY) which is the document from which the dossier was created	- Yes
- i5Origin	- xs:boolean	- Flag indicating whether this document originated in IUCLID 5 format, and has been migrated to the current IUCLID 6 format.	- Yes
- creationTool	- xs:string	- Element that specifies the application this document was first created with. The default value "IUC6" is provided for documents created with IUCLID 6.	- Yes
- snapshotCreationTool	- xs:string	- In case of dossier archive, element that specifies the application this dossier was created from. Upon dossier creation by IUCLID this is filled in with "IUC6".	- Yes

Below is an example of the subsection <PlatformMetadata> for documents before the 7.0.0 IUCLID6 version

---

XML example of <platform metadata> section of a DOSSIER HEADER document

```
<i6c:PlatformMetadata xmlns:i6m="http://iuclid6.echa.europa.eu/namespaces/platform-metadata/v1"
xsi:schemaLocation="http://iuclid6.echa.europa.eu/namespaces/platform-container/v1 platform-container.xsd">

  <i6m:iuclidVersion>2.5.0</i6m:iuclidVersion>

  <i6m:documentKey>98051ac9-53d0-45b0-acae-42c28665de54/98051ac9-53d0-45b0-acae-
42c28665de54</i6m:documentKey>

  <i6m:parentDocumentKey/>

  <i6m:name/>

  <i6m:documentType>DOSSIER</i6m:documentType>

  <i6m:documentSubType>R_1-10_PC</i6m:documentSubType>

  <i6m:orderInSectionNo/>

  <i6m:definitionVersion>3.0</i6m:definitionVersion>
```

XML example of <platform metadata> section of a DOSSIER HEADER document

<i6m:creationDate>2018-03-09T12:22:55Z</i6m:creationDate>

<i6m:lastModificationDate>2018-03-09T12:22:55Z</i6m:lastModificationDate>

<i6m:submissionType>R_1-10_PC</i6m:submissionType>

<i6m:submissionTypeVersion>reach 3.0</i6m:submissionTypeVersion>

<i6m:submittingLegalEntity/>

<i6m:dossierSubject>50672294-fceb-4ff4-9321-bc19c1bc9224/98051ac9-53d0-45b0-acae-42c28665de54</i6m:dossierSubject>

<i6m:i5Origin>false</i6m:i5Origin>

<i6m:creationTool>IUC6</i6m:creationTool>

<i6m:snapshotCreationTool>IUC6</i6m:snapshotCreationTool>

</i6c:PlatformMetadata>

Below is an example of the subsection <PlaformMetadata> for documents of the 7.0.0 IUCLID6 version

XML example of <platform metadata> section of a DOSSIER HEADER document

<i6c:PlatformMetadata xmlns:i6m="http://iuclid6.echa.europa.eu/namespaces/platform-metadata/v1" xsi:schemaLocation="http://iuclid6.echa.europa.eu/namespaces/platform-container/v2 platform-container-v2.xsd">

<i6m:iuclidVersion>7.0.0</i6m:iuclidVersion>

<i6m:documentKey>98051ac9-53d0-45b0-acae-42c28665de54/98051ac9-53d0-45b0-acae-42c28665de54</i6m:documentKey>

<i6m:parentDocumentKey/>

<i6m:name/>

<i6m:documentType>DOSSIER</i6m:documentType>

<i6m:documentSubType>R_1-10_PC</i6m:documentSubType>

<i6m:orderInSectionNo/>

<i6m:definitionVersion>8.0</i6m:definitionVersion>

<i6m:creationDate>2023-03-22T10:17:55Z</i6m:creationDate>

<i6m:lastModificationDate> 2023-03-22T10:17:55Z</i6m:lastModificationDate>

<i6m:submissionType>R_1-10_PC</i6m:submissionType>

<i6m:submissionTypeVersion>reach 8.0</i6m:submissionTypeVersion>

<i6m:submittingLegalEntity/>

<i6m:dossierSubject>50672294-fceb-4ff4-9321-bc19c1bc9224/98051ac9-53d0-45b0-acae-42c28665de54</i6m:dossierSubject>

<i6m:i5Origin>false</i6m:i5Origin>

<i6m:creationTool>IUC6</i6m:creationTool>

<i6m:snapshotCreationTool>IUC6</i6m:snapshotCreationTool>

</i6c:PlatformMetadata>

IUCLID 6

2.3.2. <Content> section

This section contains the chemical information of the specific document. At this point, we should describe in short the key concepts of IUCLID 6 document types and structure. In IUCLID 6 every structure that keeps chemical information is called document. There are the following type of documents:

- **Entity:** This is a top-level document without temporal or other dependencies to other documents. Examples are: ANNOTATION, ARTICLE, CATEGORY, LEGAL_ENTITY, MIXTURE, REFERENCE_SUBSTANCE, SITE, CONTACT, LITERATURE, SUBSTANCE, TEMPLATE, TEST_MATERIAL_INFORMATION. A sub-category of the entities is the composite entities that refer to documents, which hold additional children documents underneath. Examples of composite entities are: SUBSTANCE, MIXTURE, TEMPLATE. ARTICLE has been defined as a simple entity because no other documents are included under it.
- **Section:** This document lives only under a composite entity and can be any of the following types: FIXED_RECORD, FLEXIBLE_RECORD, ENDPOINT_STUDY_RECORD, FLEXIBLE_SUMMARY, ENDPOINT_SUMMARY. The section documents are further categorized based on the kind of information they hold. The extra categorization is described by the document subtype.

Every document, entity or section, has a specific and well-defined structure. This hierarchical, tree-based data structure consists of the following nodes:

- **root:** the document itself
- **leaf:** the document fields of the various types. IUCLID 6 supports numerous data types, primitives and composite. The data types are:
  - Text, including rich-text field which content should be provided as HTML
  - Picklist (single/multiple)
  - Physical quantity
  - Physical quantity range
  - Data protection / confidentiality
  - Attachments (single/multiple)
  - Address
  - Boolean
  - Date
  - Number
  - Document reference (single/multiple)
  - Section types
- **internal nodes:** the blocks that specify a set of fields or nested blocks (blocks inside blocks) under a specific header. This header acts as an aggregator of related chemical information. There are two types of blocks:
  - **single blocks:** a block with a single occurrence per document. This block can be uniquely identified by its path
  - **repeatable blocks:** a block with an unbounded number of occurrences. These blocks have the same document path and exist multiple times so to uniquely identify them, there are assigned with block identifiers. See section *3.1 Document identification*.

Every node has a type/identifier. The identifier of the root node is the concatenation of its document type/subtype information. By node path we mean the location of the node inside the tree. Its value is the concatenation of all the identifiers from the root node up to the node itself.

The Content section contains the XML-based representation of the hierarchical document structure where leaf elements provide the values of the document fields and internal nodes are blocks consisting of other fields or nested blocks. There is a standard XML representation of

IUCLID 6

the composite field elements. In case of repeatable blocks, the UUID attribute is utilized to uniquely identify the block inside the document.

An example that illustrates the XML format for a document definition is given below. Let's consider the following sample endpoint study record, which contains a single and a repeatable block, each containing various types of fields:

ENDPOINT_STUDY_RECORD.Density

AdministrativeData: single block

DataProtection: data protection field

Endpoint: picklist field

AttachedJustification: repeatable block

AttachedJustification: attachment field

ReasonPurpose: picklist field

XML example for ENDPOINT_STUDY_RECORD.Density

```
<i6c:Content>
  <ENDPOINT_STUDY_RECORD.Density
xmlns="http://iuclid6.echa.europa.eu/namespaces/ENDPOINT_STUDY_RECORD-Density/2.0"
xmlns:i6="http://iuclid6.echa.europa.eu/namespaces/platform-fields/v1">
    <AdministrativeData>
      <DataProtection>
        <i6:confidentiality/>
        <i6:justification>Justification_text</i6:justification>
      </DataProtection>
      <Endpoint>
        <i6:value>2311</i6:value>
        <i6:remarks>REMARKS_ENDPOINT_STUDY_RECORD.Density.AdministrativeData.Endpoint</i6:remarks>
      </Endpoint>
      …
      <AttachedJustification>
        <entry i6:uuid="66687d92-2220-4b95-9b21-61cc692a9a8c">
          <AttachedJustification>599777e4-6619-4863-bb1f-02cd2dc913fc/3cfc96de-e097-4b1d-a7ac-
39b1d108593d</AttachedJustification>
          <ReasonPurpose>
            <i6:value>60006</i6:value>

<i6:remarks>REMARKS_ENDPOINT_STUDY_RECORD.Density.AdministrativeData.AttachedJustification.ReasonPurp
ose</i6:remarks>
          </ReasonPurpose>
        </entry>
        <entry i6:uuid="74f0fafb-aaee-49ab-853c-3de1da94ab5d">
          <AttachedJustification>7424406b-eb1a-4ccf-a625-ea272a106d2d/3cfc96de-e097-4b1d-a7ac-
39b1d108593d</AttachedJustification>
          <ReasonPurpose>
```

IUCLID 6

XML example for ENDPOINT_STUDY_RECORD.Density

```xml
                <i6:value>60010</i6:value>

<i6:remarks>REMARKS_ENDPOINT_STUDY_RECORD.Density.AdministrativeData.AttachedJustification.ReasonPurp
ose</i6:remarks>
                </ReasonPurpose>
            </entry>
            <entry i6:uuid="9c8425c5-7630-494d-9f63-99d98ad6ae75">
                <AttachedJustification>396555fa-795b-461c-b83d-a297830f3726/3cfc96de-e097-4b1d-a7ac-
39b1d108593d</AttachedJustification>
                <ReasonPurpose>
                    <i6:value>60009</i6:value>

<i6:remarks>REMARKS_ENDPOINT_STUDY_RECORD.Density.AdministrativeData.AttachedJustification.ReasonPurp
ose</i6:remarks>
                </ReasonPurpose>
            </entry>
        </AttachedJustification>
        …
      …
    </ENDPOINT_STUDY_RECORD.Density>
</i6c:Content>
```

Important notes

- The content of this section depends on the definition format of the document under the specific version
- During import, an XSD-based validation verifies that the structure of the i6d document of the specific version is the correct one. That means that the sequence of elements is the expected one and all field values match the XSD-defined types.

2.3.3. <Attachments> section

Section that lists the attachments that are directly linked to the document. The content of this section is an unbounded list of references to attachment identifiers to which this document is linked.

The following example indicates that there is an attachment linked directly to this document that has the identifier:

b9800746-5007-3ad7-b2c8-cd8a544bbdf9/0

By directly, we mean that this content is directly linked to the document itself, and not included in any of the document's fields.

XML example of <Attachments> section

```xml
<i6c:Attachments xmlns:i6a="http://iuclid6.echa.europa.eu/namespaces/platform-attachment/v1"
xsi:schemaLocation="http://iuclid6.echa.europa.eu/namespaces/platform-attachment/v1 platform-
attachment.xsd">
```

IUCLID 6

```
  <i6a:AttachmentRef>b9800746-5007-3ad7-b2c8-cd8a544bbdf9/0</i6a:AttachmentRef>
</i6c:Attachments>
```

2.3.4. <ModificationHistory> section

This section lists the entries of the document's modification history. Every entry is a single operation that took place on the specific document and specifies the date of the action, the user that run the action, the submitting legal entity of the user and the modification remarks if any.

Table 6: XML elements under the section <ModificationHistory> in an i6d file for a document

Element	Type	Description	Required
- Date	- xs:dateTime	- The date-time in ISO-8601 format the action was performed on the document	- Yes
- Author	- xs:string	- The userName of the user that performed the modification	- Yes
- LegalEntity	- xs:string	- The description of the submitting legal entity of the user. This information contains the concatenated value of the LE name, city and localized country information	- Yes
- Remarks	- xs:string	- The modification comment	- Yes

Below is a sample of the modification history section.

XML example of < ModificationHistory > section

```
<i6c:ModificationHistory xmlns:i6h="http://iuclid6.echa.europa.eu/namespaces/platform-modification-history/v1"
xsi:schemaLocation="http://iuclid6.echa.europa.eu/namespaces/platform-modification-history/v1 platform-
modification-history.xsd">
  <i6h:Modification>
    <i6h:Date>2017-12-15T10:59:18Z</i6h:Date>
    <i6h:Author>MuylaertJoris</i6h:Author>
    <i6h:LegalEntity>Folat B.V. / Haarlem / Netherlands</i6h:LegalEntity>
    <i6h:Remarks/>
  </i6h:Modification>
  <i6h:Modification>
    <i6h:Date>2017-12-15T10:59:09Z</i6h:Date>
    <i6h:Author>MuylaertJoris</i6h:Author>
    <i6h:LegalEntity>Folat B.V. / Haarlem / Netherlands</i6h:LegalEntity>
    <i6h:Remarks/>
  </i6h:Modification>
  <i6h:Modification>
    <i6h:Date>2017-12-15T10:58:38Z</i6h:Date>
```

XML example of < ModificationHistory > section

```xml
    <i6h:Author>MuylaertJoris</i6h:Author>
    <i6h:LegalEntity>Folat B.V. / Haarlem / Netherlands</i6h:LegalEntity>
    <i6h:Remarks>Created</i6h:Remarks>
  </i6h:Modification>
</i6c:ModificationHistory>
```

3. Hints for developers

3.1. Document identification

A textual key is utilized to uniquely identify IUCLID 6 documents and attachments. In the IUCLID 6 domain, this structure is called DocumentKey and its format is <document_uuid>/<snapshot_uuid>

- **document_uuid:** It is the unique identifier of the document or attachment
- **snapshot_uuid:** Indicates the UUID of the dossier that the document or attachment belongs to. If the document is not yet part of the dossier, which means it is a raw and editable document, then the snapshot_uuid = 0.

The content of every subpart is a universally unique identifier (UUID) being a 128-bit number. In its canonical textual representation, the UUID is represented as 32 hexadecimal (base 16) digits, displayed in five groups separated by hyphens, in the form 8-4-4-4-12 for a total of 36 characters (32 alphanumeric characters and four hyphens).

Please note that for Legal Entity documents, it is allowed a prefix of the following format:

- ECHA-<document_uuid>
- IUC5-<document_uuid>
- ECB5-<document_uuid>

For additional information, please check:

Universally unique identifier - Wikipedia

RFC 4122: A Universally Unique IDentifier (UUID) URN Namespace

Finally, the information concerning the application this document originated from is stored in the corresponding metadata elements *creationTool* and *snapshotCreationTool*.

Below, a code snippet for UUID generation.

UUID generation

```java
java.util.UUID.randomUUID().toString();
```

3.2. Block identification

The IUCLID 6 domain contains blocks. A block is a collection of elements (fields or blocks) defined under the path of the block. There are two types of block:

1. **Single block:** A block with maximum one occurrence in the document context

2. **Repeatable block:** A block with multiple occurrences within the same document

IUCLID 6

Single blocks are uniquely identified per document via its path. Repeatable blocks however have the same path so they are assigned with UUID identifiers. The identifier does not follow the document identification pattern so there is no snapshot_uuid kept in the block entries.

In the sample of the exported document below, one can see that every entry of a repeatable block contains the UUID identifier.

Repeatable block identification

```
<i6c:Content>
  <ENDPOINT_STUDY_RECORD.Density
xmlns="http://iuclid6.echa.europa.eu/namespaces/ENDPOINT_STUDY_RECORD-Density/3.0"
xmlns:i6="http://iuclid6.echa.europa.eu/namespaces/platform-fields/v1">
    <!-- Single block -->
    <AdministrativeData>
      …
      <!-- Repeatable block -->
      <AttachedJustification>
        <entry i6:uuid="66687d92-2220-4b95-9b21-61cc692a9a8c">
          …
        </entry>
        <entry i6:uuid="74f0fafb-aaee-49ab-853c-3de1da94ab5d">
          …
        </entry>
      </AttachedJustification>
```

3.3. Cross-referencing XML documents

As already specified, an `i6z` archive includes a single file `manifest.xml`. This file provides some top-level metadata of the archive along with metadata of the documents and attachments included in the archive.

One of the main metadata information is the reference to the XML document (i6d) that contains the actual document content. Check below:

Referencing XML documents

```
<document id="fb199594-8b87-426b-a21d-eac5f1c930a9/3cfc96de-e097-4b1d-a7ac-39b1d108593d">
  …
  <name xlink:type="simple" xlink:href="fb199594-8b87-426b-a21d-eac5f1c930a9_3cfc96de-e097-4b1d-a7ac-
39b1d108593d.i6d">DSD - DPD.001</name>
  …
```

Similarly for the reference to the XML attachment information (i6d):

Referencing XML attachments

```
<attachment id="96072c4b-bb63-4d24-81f5-53747ddb7e3c/3cfc96de-e097-4b1d-a7ac-39b1d108593d">
  <name xlink:type="simple" xlink:href="96072c4b-bb63-4d24-81f5-53747ddb7e3c_3cfc96de-e097-4b1d-a7ac-
```

IUCLID 6

```
39b1d108593d.i6d">Excel-icon.jpg</name>
   …
</attachment>
```

As already indicated, XLink *(XML Linking Language)* is used to create hyperlinks within XML documents. This is the official W3C solution for XML linking.

XLink defines a set of attributes that may be added to elements of other XML namespaces. To get access to the XLink features, the XLink namespace must be declared at the top of the manifest document:

```
xmlns:xlink="http://www.w3.org/1999/xlink"
```

The attribute `xlink:type="simple"` creates a simple "HTML-like" link.

The attribute `xlink:href` specifies the URL to link to. In our case, this is the `i6d` XML document that contains the document or attachment content.Null or empty-value fields.

An `i6z` archive should contain documents specifying only the fields with actual content; null or empty-value fields should be omitted.

If a required field is missing from the XML document, the proper XSD validation error will be thrown.

3.4. Multilingual support

IUCLID 6 supports multilingual content (since version 6.3). All fields that include any kind of textual information have been modified to accept content in multiple languages. The affected fields are:

1. Text fields

2. Picklist fields (with and without remarks)

3. Data protection fields

4. Physical quantity fields

5. Physical quantity range fields

Except of the Text field that is a single-value field, all other fields are composite and consist of multiple subfields. For example the picklist field contains textual information on the "other" and "remarks" subfields. The physical quantity and quantity range fields consist, among others, of a picklist field so the "other" textual subfield is also present.

From the XML point of view, the implemented approach is to allow multiple occurrences of the same textual field/subfield. In addition to that, any multilingual field, could optionally contain the @xml:lang attribute that specifies the language and (optional) locale of the element content. The @xml:lang attribute is described in the XML Recommendation. Every occurrence of the @xml:lang attribute value should be unique in a field; as a result, it is not allowed providing content for the same language multiple times for the same field. Note that the value for the @xml:lang attribute must be the 2-letter ISO language code, for example, "en" or "de" or "fr".

From a Poison Centres Notification dossier point of view (the only submission type currently

IUCLID 6

using multilingual fields), the relevant specific languages should be provided for multilingual fields. The default values, without the language attribute, will not be taken into account and can be ignored.

Below are some examples of the content of the multilingual fields.

Multilingual XML content

```xml
<!-- Text field -->
<JustificationForTypeOfInformation>this is the default text</JustificationForTypeOfInformation>
<JustificationForTypeOfInformation xml:lang="en">This is the english text</JustificationForTypeOfInformation>
<JustificationForTypeOfInformation xml:lang="it">This is the italian text</JustificationForTypeOfInformation>


<!-- Single picklist field -->
<StudyResultType>
   <i6:value>1342</i6:value>
   <i6:other>this is the default other text</i6:other>
   <i6:other xml:lang="en">This is the english other text</i6:other>
   <i6:other xml:lang="it">This is the italian other text</i6:other>
   <i6:remarks>this is the default remarks</i6:remarks>
   <i6:remarks xml:lang="en">This is the english remarks text</i6:remarks>
   <i6:remarks xml:lang="it">This is the italian remarks text</i6:remarks>
</StudyResultType>


<!-- Multiple picklist field -->
<SpecificSubmissions>
   <RegulatoryProgramme>
      <i6:value>209</i6:value>
      <i6:remarks>this is the default remarks</i6:remarks>
   </RegulatoryProgramme>
   <RegulatoryProgramme>
      <i6:value>210</i6:value>
      <i6:remarks>this is the default remarks</i6:remarks>
      <i6:remarks xml:lang="en">This is the english remarks text</i6:remarks>
      <i6:remarks xml:lang="it">This is the italian remarks text</i6:remarks>
   </RegulatoryProgramme>
</SpecificSubmissions>


<!-- Physical quantity field -->
<Conc>
   <unitCode>1342</unitCode>
   <unitOther>this is the default text</unitOther>
```

Multilingual XML content

```xml
    <unitOther xml:lang="en">This is the english text</unitOther>
    <unitOther xml:lang="it">This is the italian text</unitOther>
    <value>1</value>
  </Conc>


  <!-- Physical quantity range field -->
  <Tension>
    <i6:unitCode>1342</i6:unitCode>
    <i6:unitOther>this is the default text</i6:unitOther>
    <i6:unitOther xml:lang="en">This is the english other text</i6:unitOther>
    <i6:unitOther xml:lang="it">This is the italian other text</i6:unitOther>
    <i6:lowerQualifier>&gt;</i6:lowerQualifier>
    <i6:upperQualifier>&lt;</i6:upperQualifier>
    <i6:lowerValue>1</i6:lowerValue>
    <i6:upperValue>2</i6:upperValue>
  </Tension>


  <!-- Data protection field -->
  <DataProtection>
    <i6:confidentiality>3441</i6:confidentiality>
    <i6:justification>this is the default justification text</i6:justification>
    <i6:justification xml:lang="en">This is the english justification text</i6:justification>
    <i6:justification xml:lang="it">This is the italian justification text</i6:justification>
    <i6:legislation>
      <i6:value>1342</i6:value>
      <i6:other>other legislation text</i6:other>
      <i6:other xml:lang="en">This is the english other text</i6:other>
      <i6:other xml:lang="it">This is the italian other text</i6:other>
    </i6:legislation>
  </DataProtection>
```

Below some examples of the XSDs concerning multilingual fields. Please note that all the multilingual complexTypes assigned to the document fields are specified inside `platform-fields.xsd` and declared under the http://iuclid6.echa.europa.eu/namespaces/platform-fields/v1 namespace. That is the reason for the prefix i6 when specifying these elements at the XML document as shown above.

Multilingual XSD content

Multilingual XSD content

```xml
<!-- Text field -->

<xs:element name="JustificationForTypeOfInformation" minOccurs="1" maxOccurs="unbounded"
type="i6:multiLingualTextFieldLarge"/>


<!-- Single picklist field -->

<xs:element name="StudyResultType" type="i6:multiLingualPicklistFieldWithSmallTextRemarks"/>

<xs:complexType name="multiLingualPicklistFieldWithSmallTextRemarks">

  <xs:complexContent>

    <xs:restriction base="picklistFieldWithSmallTextRemarks">

      <xs:sequence>

        <xs:element name="value" minOccurs="0" type="textFieldSmall" />

        <xs:element name="other" minOccurs="0" maxOccurs="unbounded" type="multiLingualTextFieldSmall" />

        <xs:element name="remarks" minOccurs="0" maxOccurs="unbounded"
type="multiLingualTextFieldSmall"/>

      </xs:sequence>

    </xs:restriction>

  </xs:complexContent>

</xs:complexType>


<!-- Multiple picklist field -->

<xs:element name="RegulatoryProgramme" minOccurs="0" maxOccurs="unbounded"
type="i6:multiLingualPicklistFieldWithMultiLineTextRemarks"/>

<xs:complexType name="multiLingualPicklistFieldWithMultiLineTextRemarks">

  <xs:complexContent>

    <xs:restriction base="picklistFieldWithMultiLineTextRemarks">

      <xs:sequence>

        <xs:element name="value" minOccurs="0" type="textFieldSmall" />

        <xs:element name="other" minOccurs="0" maxOccurs="unbounded" type="multiLingualTextFieldSmall" />

        <xs:element name="remarks" minOccurs="0" maxOccurs="unbounded"
type="multiLingualTextFieldMultiline"/>

      </xs:sequence>

    </xs:restriction>

  </xs:complexContent>

</xs:complexType>


<!-- Physical quantity field -->

<xs:element name="Conc" type="i6:multiLingualPhysicalQuantityField"/>

<xs:complexType name="multiLingualPhysicalQuantityField">

  <xs:complexContent>

    <xs:restriction base="physicalQuantityField">
```

Multilingual XSD content

```xml
        <xs:sequence>
            <xs:element name="unitCode" minOccurs="0" type="textFieldSmall" />
            <xs:element name="unitOther" type="multiLingualTextFieldSmall" minOccurs="0"
maxOccurs="unbounded"/>
            <xs:element name="value" type="xs:decimal" minOccurs="0"/>
        </xs:sequence>
      </xs:restriction>
   </xs:complexContent>
</xs:complexType>


<!-- Physical quantity range field -->
<xs:element name="Tension" type="i6:multiLingualPhysicalQuantityRangeField"/>
<xs:complexType name="multiLingualPhysicalQuantityRangeField">
  <xs:complexContent>
    <xs:restriction base="physicalQuantityRangeField">
      <xs:sequence>
        <xs:element name="unitCode" minOccurs="0" type="textFieldSmall" />
        <xs:element name="unitOther" minOccurs="0" maxOccurs="unbounded"
type="multiLingualTextFieldSmall"/>
        <xs:group ref="physicalQuantityRangeGroup"/>
      </xs:sequence>
    </xs:restriction>
  </xs:complexContent>
</xs:complexType>


<!-- Data protection field -->
<xs:element name="DataProtection" type="i6:multiLingualDataProtectionField"/>
<xs:complexType name="multiLingualDataProtectionField">
  <xs:complexContent>
    <xs:restriction base="dataProtectionField">
      <xs:sequence>
        <xs:element name="confidentiality" type="xs:string"/>
        <xs:element name="justification" maxOccurs="unbounded" type="multiLingualTextFieldLarge"/>
        <xs:element name="legislation" minOccurs="0" maxOccurs="unbounded">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="value" type="xs:string"/>
              <xs:element name="other" minOccurs="0" maxOccurs="unbounded"
type="multiLingualTextFieldSmall"/>
            </xs:sequence>
```

Multilingual XSD content

```
            </xs:complexType>

        </xs:element>

      </xs:sequence>

    </xs:restriction>

  </xs:complexContent>

</xs:complexType>
```

It should be noted that the requirement for the multilingual support created new platform-specific types that are extensions of the existing ones in order to increase the element's occurrence and account for the optional @xml:lang attribute.

Examples of the new types added in the platform-fields.xsd are:

Table 7:   Examples of the new types added in the platform-fields.xsd

Field	Type
multilingualTextField	- textual value with no size restriction
multilingualTextFieldSmall	- textual value with 255 characters length restriction
multilingualTextFieldMultiLine	- textual value with 2000 characters length restriction
multilingualTextFieldLarge	- textual value with 32768 characters length restriction
multilingualPicklistField	- complex type specifying value and other elements
multilingualPicklistFieldWithSmallTextRemarks	- complex type specifying value, other and remarks elements restricting the latter to 255 characters
multilingualPicklistFieldWithLargeTextRemarks	- complex type specifying value, other and remarks elements restricting the latter to 32768 characters
multilingualPicklistFieldWithMultiLineTextRemarks	- complex type specifying value, other and remarks elements restricting the latter to 2000 characters
multilingualDataProtectionField	- complex type specifying sub-elements (confidentiality, justification and legislation) of data protection field
multilingualPhysicalQuantityRangeField	- complex type specifying sub-elements (unitCode, unitOther, lowerQualifier, upperQualifier, lowerValue, upperValue) of physical quantity range fields with decimal lower and upper values
multilingualPhysicalQuantityIntegerRangeFi	- complex type specifying sub-elements

Field	Type
eld	(unitCode, unitOther, lowerQualifier, upperQualifier, lowerValue, upperValue) of physical quantity range fields with integer lower and upper values
multilingualPhysicalQuantityHalfBoundedField	- complex type specifying sub-elements (unitCode, unitOther, qualifier, value) of physical quantity half bounded fields with decimal value
multilingualPhysicalQuantityIntegerHalfBoundedField	- complex type specifying sub-elements (unitCode, unitOther, qualifier, value) of physical quantity half bounded fields with decimal value

3.5. Encoding picklist phrases

Part of the IUCLID 6 definition is the picklist-related fields along with their applicable phrase codes. These fields include single/multiple picklist, physical quantity, physical quantity range and data protection fields.

IUCLID 6 includes an XSD generation mechanism so that:

- Every element holding a picklist phrase code value is linked to a custom, text-based type listing the eligible phrase codes for the given picklist field. The eligible values are identified based on the phrase group code that is linked to the specific picklist field.

Encoding picklist phrases in XSD of OECD v3 legislation

```xml
<!-- XSD for v3 document with type ENDPOINT_STUDY_RECORD.BioaccumulationAquaticSediment -->
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns="http://iuclid6.echa.europa.eu/namespaces/ENDPOINT_STUDY_RECORD-
BioaccumulationAquaticSediment/3.0" xmlns:ct="http://iuclid6.echa.europa.eu/namespaces/oecd/v3"
xmlns:i6="http://iuclid6.echa.europa.eu/namespaces/platform-fields/v1" attributeFormDefault="qualified"
elementFormDefault="qualified"
targetNamespace="http://iuclid6.echa.europa.eu/namespaces/ENDPOINT_STUDY_RECORD-
BioaccumulationAquaticSediment/3.0">

<!-- Import the commonTypes.xsd that contains the phrasegroup-specific types that are included in the
commonTypes.xsd and are part of the specific namespace for OECD v3 legislation -->
<xs:import namespace="http://iuclid6.echa.europa.eu/namespaces/oecd/v3"
schemaLocation="commonTypes.xsd"/>

<xs:element name="ENDPOINT_STUDY_RECORD.BioaccumulationAquaticSediment">
  …
  <!-- Confidentiality value of data protection field -->
  <xs:element minOccurs="0" name="confidentiality" type="ct:N64"/>

  <!-- Regulatory programme value of data protection field -->
  <xs:element minOccurs="0" name="value" type="ct:N78"/>

  <!-- Phrase code of a picklist field -->
```

```xml
   <xs:element minOccurs="0" name="value" type="ct:PG6_60496"/>

   <!-- unitCode value of a physical quantity (range) field -->
   <xs:element minOccurs="0" name="unitCode" type="ct:E04"/>
   …
 </xs:element>
</xs:schema>
```

Below a part of the commonTypes.xsd listing the various phrase group-specific types per legislation provider.

OECD v3 legislation commonTypes.xsd

```xml
<!-- Every type defined is part of the OECD v3 name space (http://iuclid6.echa.europa.eu/namespaces/oecd/v3) -->
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns="http://iuclid6.echa.europa.eu/namespaces/oecd/v3"
xmlns:i6="http://iuclid6.echa.europa.eu/namespaces/platform-fields/v1" attributeFormDefault="qualified"
elementFormDefault="qualified" targetNamespace="http://iuclid6.echa.europa.eu/namespaces/oecd/v3">

   <xs:import namespace="http://iuclid6.echa.europa.eu/namespaces/platform-fields/v1"
schemaLocation="platform-fields.xsd"/>

   <xs:simpleType name="N64">
     <xs:restriction base="i6:textFieldSmall">
       <xs:enumeration value=""/>
       <xs:enumeration value="2732"/>
       <xs:enumeration value="2859"/>
       <xs:enumeration value="3441"/>
     </xs:restriction>
   </xs:simpleType>

   <xs:simpleType name="N78">
     <xs:restriction base="i6:textFieldSmall">
       <xs:enumeration value=""/>
       <xs:enumeration value="733"/>
       <xs:enumeration value="9000"/>
       <xs:enumeration value="5793"/>
       <xs:enumeration value="735"/>
       <xs:enumeration value="209"/>
       <xs:enumeration value="210"/>
       <xs:enumeration value="919"/>
       <xs:enumeration value="1313"/>
       <xs:enumeration value="1646"/>
```

OECD v3 legislation commonTypes.xsd

```xml
        <xs:enumeration value="1647"/>
        <xs:enumeration value="1648"/>
        <xs:enumeration value="1342"/>
      </xs:restriction>
  </xs:simpleType>


  <xs:simpleType name="PG6_60496">
    <xs:restriction base="i6:textFieldSmall">
      <xs:enumeration value=""/>
      <xs:enumeration value="62504"/>
      <xs:enumeration value="62505"/>
      <xs:enumeration value="62507"/>
      <xs:enumeration value="3437"/>
      <xs:enumeration value="58256"/>
      <xs:enumeration value="2282"/>
      <xs:enumeration value="62470"/>
    </xs:restriction>
  </xs:simpleType>


  <xs:simpleType name="E04">
    <xs:restriction base="i6:textFieldSmall">
      <xs:enumeration value=""/>
      <xs:enumeration value="2113"/>
      <xs:enumeration value="1976"/>
      <xs:enumeration value="1839"/>
      <xs:enumeration value="2468"/>
    </xs:restriction>
  </xs:simpleType>
</xs:schema>
```

3.6. XML namespaces

IUCLID 6 makes heavy use of namespaces in order to break up the various XML schemas (platform- and document-specific). In that way, we are able to define re-usable definitions and avoid conflicts of element and attribute names. Defining a default namespace for an element also saves us from using prefixes in all its child elements.

Below an analysis of:

- The namespaces supported by IUCLID 6;
- The alias for each namespace in order to fully qualify elements and types that are defined

IUCLID 6

in the corresponding XSD and are part of this namespace;
- The XSD file defining the entities that are part of this namespace;
- The XML file inside the i6z archive where this namespace is used.

Table 8: **XML namespaces**

Namespaces	Name space alias	XSD file	XML document usage
- http://iuclid6.echa.europa.eu/namespaces/manifest/v1	-	- manifest.xsd - Specifies the elements included in the manifest.xml along with their types	- manifest.xml
- http://iuclid6.echa.europa.eu/namespaces/platform-container/v1	- i 6 c	- platform-container.xsd - Specifies the sub-elements (PlatformMetadata, Content, Attachments, ModificationHistory) of the Document element. - It should be used for documents that were introduced before the 7.0.0 IUCLID 6 version	- Document-specific XML files (.i6d)
- http://iuclid6.echa.europa.eu/namespaces/platform-container/v2	- i 6 c	- platform-container-v2.xsd - Specifies the sub-elements (PlatformMetadata, Content, Attachments, ModificationHistory) of the Document element. - It should be used for documents that were introduced in the 7.0,0 IUCLID 6 version or greater. -	- Document-specific XML files (.i6d)
- http://iuclid6.echa.europa.eu/namespaces/platform-metadata/v1	- i 6 n	- platform-metadata.xsd - Specifies the sub-elements of the PlatformMetadata element under the Document element	- Document-specific XML files (.i6d)
- http://iuclid6.echa.europa.eu/namespaces/platform-attachment/v1	- i 6 a	- platform-attachment.xsd - Specifies the sub-elements of the Attachment element along with their types. This is also a sub-element of the Attachments under the Document element	- Attachment-specific XML files (.i6d) - Document-

Namespaces	Name space alias	XSD file	XML document usage
			specific XML files (.i6d)
- http://iuclid6.echa.europa.eu/namespaces/platform-modification-history/v1	- i 6 h	- platform-modification-history.xsd - Specifies the ModificationEntry sub-elements and types under the ModificationHistory which lies under the Document element	- Document-specific XML files (.i6d)
- http://iuclid6.echa.europa.eu/namespaces/platform-fields/v1	- i 6	- platform-fields.xsd - Specifies all the common types of the IUCLID fields - Imported by the auto-generated XSDs to gain access to its definitions/types	- Document-specific XML files (.i6d)
- http://iuclid6.echa.europa.eu/namespaces/<ENTITY_TYPE>/<DEFINITION_VERSION> - e.g. http://iuclid6.echa.europa.eu/namespaces/SUBSTANCE/3.0	- -	- Auto-generated XSD files <ENTITY_TYPE>-<DEFINITION_VERSION>.xsd - e.g. SUBSTANCE-3.0.xsd - Specifies the element with name <ENTITY_TYPE> under the Content element	- Document-specific XML files (.i6d) -
- http://iuclid6.echa.europa.eu/namespaces/<DOCUMENT_TYPE>-<DOCUMENT_SUBTYPE>/<DEFINITION_VERSION> - e.g. http://iuclid6.echa.europa.eu/namespaces/ENDPOINT_STUDY_RECORD-WaterSolubility/3.0	- -	- Auto-generated XSD files <DOCUMENT_TYPE>-<DOCUMENT_SUBTYPE>-<DEFINITION_VERSION>.xsd - e.g. ENDPOINT_STUDY_RECORD-WaterSolubility-3.0.xsd - Specifies the element with name <DOCUMENT_TYPE>.<DOCUMENT_SUBTYPE> under the Content element	- Document-specific XML files (.i6d)

3.7. IUCLID 6 document references

IUCLID 6 provides many data types to assign to every document field. One of these types allows a field to reference another IUCLID 6 document. These types are the DocumentReferenceField and DocumentReferenceMultipleField based on whether the field keeps a reference to a single or multiple documents respectively. The value of this field is the

IUCLID 6

identifier of the referenced document.

Below, the XML presentation of the document references which makes clear that a weak reference exists.

IUCLID 6 document references

```xml
<?xml version='1.0' encoding='UTF-8'?>

<!-- Field SUBSTANCE.OwnerLegalEntity referencing a single LEGAL ENTITY document -->
<OwnerLegalEntity>4f88bc7f-395c-4d0b-997b-14e8c9aef605/0</OwnerLegalEntity>

<!-- Field ENDPOINT_STUDY_RECORD.DermalAbsorption.DataSource.Reference referencing multiple LITERATURE
documents -->
<Reference>
  <i6:key>d247ab3f-2393-36f2-a603-48ec1a84d694/0</i6:key>
  <i6:key>f574f4b0-af68-300b-9aa2-85c6071280f7/0</i6:key>
</Reference>
```

3.8. Attachment content

The Attachment XML document (`i6d`) has the following structure.

Attachment document (.i6d)

```xml
<?xml version='1.0' encoding='UTF-8'?>
<Attachment xmlns="http://iuclid6.echa.europa.eu/namespaces/platform-attachment/v1"
xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <documentKey>ee038911-4d8c-4532-ba71-ca95716a6464/3cfc96de-e097-4b1d-a7ac-
39b1d108593d</documentKey>
  <name>Excel-icon.jpg</name>
  <creationDate>2016-09-13T10:15:47.427+02:00</creationDate>
  <lastModificationDate>2016-09-13T10:15:47.427+02:00</lastModificationDate>
  <remarks>Remarks value</remarks>
  <md5>6001a8c2832408420073f53543d7e2fe</md5>
  <mimetype>image/jpeg</mimetype>
  <content xlink:href="attachments/6001a8c2832408420073f53543d7e2fe.jpg" xlink:type="simple"/>
</Attachment>
```

3.8.1. Cross-referencing attachment content

The Attachment XML document contains only the metadata information of the actual attachment file. Instead of embedding the actual content inside the document, the `content` element provides the link to the actual attachment content. During import operation, the parser resolves the link and loads the attachment from the `i6z` archive.

3.8.2. MD5 hash calculation

The `Attachment` document contains a required `md5` element. The element's value is the MD5 hash of the actual attachment content linked to the specific Attachment document and

included in the `i6z` archive under the `attachments` directory. During import operation, the attached content is parsed and its MD5 hash value is calculated. The calculated value is checked against the one specified in the document. In case of any discrepancy, the import operation fails with the corresponding error message.

Below, a code snippet to generate the MD5 hash value in Java. Exception handling is omitted.

MD5 generation

```java
import java.io.InputStream;
import java.security.DigestInputStream;
import java.security.MessageDigest;

import javax.xml.bind.DatatypeConverter;

import org.apache.tika.io.TemporaryResources;
import org.apache.tika.io.TikaInputStream;

// The loaded attachment content
InputStream attachmentContent = null;

// Object implementing the MD5 digest algorithm in charge of running the hash computation of the attachment
// content
MessageDigest md5Digest = java.security.MessageDigest.getInstance("MD5");

// Creates a digest input stream, using the specified input stream and message digest
DigestInputStream digestInputStream = new DigestInputStream(attachmentContent, md5Digest);

// Read the digest input stream to properly update the md5Digest...

// The array of hex bytes for the resulting hash value
byte[] bytes = md5Digest.digest();

// The textual representation of the MD5 hash value
String md5 = javax.xml.bind.DatatypeConverter.printHexBinary(bytes).toLowerCase();
```

3.8.3. Mime type calculation

The `Attachment` document contains an optional `mimetype` element. The value indicates the media type of the actual attachment content linked to the specific Attachment document and included in the `i6z` archive under the `attachments` directory. During import operation, the media type of the attached content is recalculated. The latter value is checked against the one specified in the document. In case of any discrepancy, the import operation proceeds normally and the corresponding `warning` message is displayed to the user. The database eventually keeps the media type calculated from the application itself thus overriding the one included in the i6d document.

Below, a code snippet to generate the media type of an attachment in Java. Exception handling is omitted.

IUCLID 6

Media type generation

```java
import org.apache.tika.config.TikaConfig;
import org.apache.tika.detect.Detector;
import org.apache.tika.metadata.Metadata;
import org.apache.tika.mime.MediaType;

// …

// The loaded attachment content
InputStream attachmentContent = null;

// Provides a default configuration
TikaConfig config = TikaConfig.getDefaultConfig();

// Returns the configured detector instance
Detector detector = config.getDetector();

// Constructs a new, empty metadata
Metadata metadata = new Metadata();

// Detects the content type of the given input document. Returns <code>application/octet-stream</code>
// if the type of the document cannot be detected.
MediaType mediaType = detector.detect(attachmentContent, metadata);

// The textual representation of the media type object
String mimeType = mediaType.toString();
```

3.9. Date-Time

Iuclid6 supports two formats to express date-time:

1) EEE MMM dd HH:mm:ss z yyyy format: This date-time format is used only for the `created` element in the `<general-information>` section of `manifest.xml`. Below there is an example of a date using this format

EEE MMM dd HH:mm:ss z yyyy format date example

Thu Nov 04 13:17:07 EET 2021

2) YYYY-MM-DDThh:mm:ss[.mmm]TZD format: For all remaining date-time elements in the IUCLID format, the ISO-8601 format is used to express a date-time in the contents of an `i6z` archive file.

Both of the following formats are acceptable

- yyyy-MM-dd'T'hh:mm:ss[.mmm]'Z'

- yyyy-MM-dd'T'hh:mm:ss [.mmm](+/-)hh:mm

Below there are example of dates using ISO-8601 format

YYYY-MM-DDThh:mm:ss[.mmm]TZD format date examples
2015-04-16T13:35:19.957+03:00, 2012-03-29T10:05:45-06:00, 2014-03-12T16:10:19Z, 2014-03-12T16:10:19.957Z