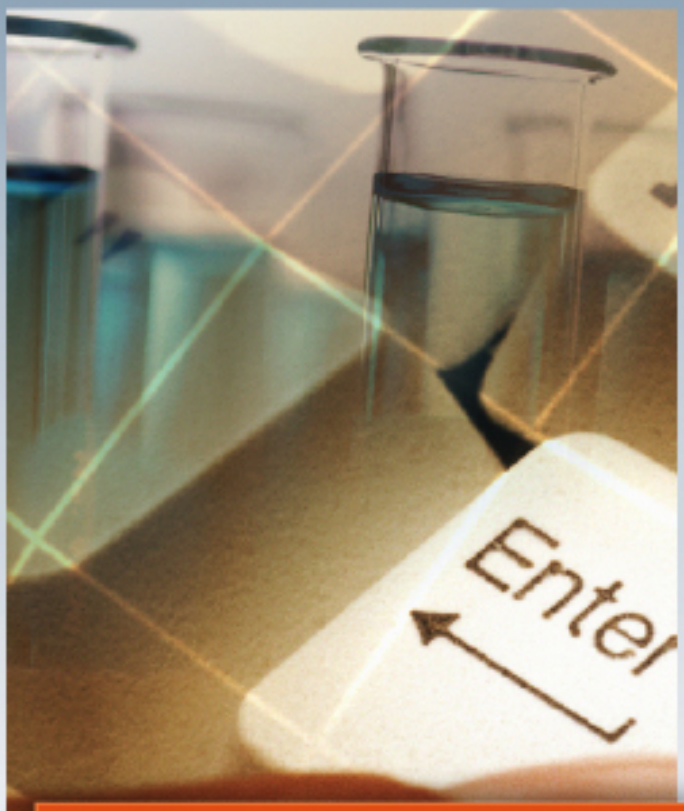
http://echa.europa.eu

Data Submission Manual 6

Submission of bulk pre-registrations (RELEASE 1)



For latest news and most up-to-date information please consult the ECHA website

Reference: ECHA-08-B-07-EN

Date: 22-07-2008

Language: English

If you have questions or comments in relation to this document please send those (quote the reference and issue date) using the information request form to ECHA helpdesk. The information request form can be accessed in the REACH helpdesk section of the ECHA website at:

http://echa.europa.eu/reach/helpdesk/echahelp_en.asp

Table of Contents

1. Introduction

REACH, the new EU chemicals regulation, requires that chemical substances on their own, in preparations and those which are intentionally released from articles have to be registered to the European Chemicals Agency (ECHA). The regulation applies to substances manufactured in, or imported to the EU in annual quantities of 1 tonne or more per company, unless the regulation indicates otherwise.

The obligation to register applies from 1 June 2008. The chemicals currently on the EU market which meet the definition of phase-in substances[1] should be pre-registered between 1 June and 1 December 2008. Companies who pre-register their substances can benefit from extended registration deadlines. The deadline depends on the tonnage band and the hazardous properties of the substance. The staggered deadlines are:

- 30 November 2010,
- 31 May 2013 or
- 31 May 2018

The information to be submitted to the European Chemicals Agency (ECHA) for pre-registration is relatively limited. It includes:

- The name of the substance as specified in section 2 of Annex VI of REACH, i.e.
  - EINECS number (if available and appropriate)
  - CAS number and CAS name (if available)
  - IUPAC[2] name or other international name(s)
  - Other names (usual name, abbreviation and trade name),
- Envisaged deadline and tonnage band for the registration
- Name and contact information of a Contact person or Third party Representative who will act as the contact point in data sharing.

A pre-registrant should also provide when applicable identifiers of related substances which may be relevant for deriving data for the substance pre-registered. This is a way to indicate which data can be shared by read-across, (quantitative) structure-activity relationships ((Q)SARs)) and grouping of substances.

---

[1] Phase-in substance means a substance which meets at least one of the following criteria:
(a) it is listed in the European Inventory of Existing Commercial Chemical Substances (EINECS)
(b) it was manufactured in the Community, or in the countries acceding to the European Union on 1 January 1995 or on 1 May 2004, but not placed on the market by the manufacturer or importer, at least once in the 15 years before the entry into force of this Regulation, provided the manufacturer or importer has documentary evidence of this
(c) it was placed on the market in the Community, or in the countries acceding to the European Union on 1 January 1995 or on 1 May 2004, before entry into force of REACH Regulation by the manufacturer or importer and was considered as having been notified in accordance with the first indent of Article 8(1) of Directive 67/548/EEC but does not meet the definition of a polymer as set out in REACH Regulation, provided the manufacturer or importer has documentary evidence of this.
Source: REACH Article 3 (20) definition

[2] International Union of Pure and Applied Chemistry (IUPAC)

There are two options to pre-register substances by using REACH-IT.

- By entering the pre-registration information directly online in the REACH-IT system: data for each phase-in substance are encoded manually following a step-by-step procedure (also called online wizard) which proceeds to a number of validations (e.g. verification of the existence of the substance in the EC inventory where EINECS numbers are stored).
- By preparing the pre-registration information outside of the REACH-IT system and storing this information in a "pre-registration file": when the file is considered as ready for submission, it can then be uploaded into the REACH-IT system so that the pre-registration process can be finalised.

The pre-registration file generated outside of the REACH-IT system may contain either:

- Information on one phase-in substance only (single pre-registration file).

  In this case, there is no constraint in the way the substance is identified: it can be by chemical name, EC number, CAS number and name, or other identifiers etc… When the pre-registration file is uploaded into REACH-IT, the online wizard is initialised so that the validation steps mentioned above can be done.

- Information on several phase-in substances (bulk pre-registration file).

  This file will be processed by the system without manual intervention by the pre-registrant. For this reason, it is only possible to pre-register in bulk substances that are identifiable by an EC number (EINECS or NLP number). In case of multi-constituent substances, this constraint does not apply to the substance itself but to each of its constituents which must be identified by an EC number (EINECS or NLP number).

The pre-registration file must be fully compliant with the REACH-IT pre-registration format (XML pre-registration format) to be successfully processed by REACH-IT. This XML format has been published in February 2008 on the IUCLID 5 web site.

The preparation of the pre-registration file (XML file) can be done either by using the IUCLID 5 pre-registration module or any company IT system that generates a XML file compliant with the REACH-IT XML pre-registration format.

The bulk pre-registration option will be available only until the end of the pre-registration period, i.e. 1st December 2008 while the other two will remain available to newcomers[3].

The purpose of this document is to describe how a legal entity can pre-register phase-in substances in bulk using the REACH-IT application.

2. Pre-registration information

General information on substance identification in REACH-IT

The REACH-IT pre-registration module follows the principles set up in the Guidance for identification and naming of substances under REACH. It enables to pre-register substances which are either:

- Well defined substances, i.e. substances with a defined qualitative and quantitative composition that can be sufficiently identified based on the identification parameters of REACH (section 2 of Annex VI). These substances may be either:
    - Mono-constituent substances, or
    - Multi-constituent substances
- "UVCB substances": Substances of Unknown or Variable composition, Complex reaction products and Biological materials.

To facilitate the description of a substance in REACH-IT during the pre-registration process, a general identification scheme is applied. This scheme has to be respected regardless of the substance type (mono- or multi-constituent substance or UVCB) and method (online or via a single XML file).

It defines a cascade of several identification levels from the preferred ones (i.e. EC number) to the less preferred ones (i.e. chemical name), such as follows:

1. All constituents of a substance should, whenever possible, be identified using an EC number (EINECS or NLP number).
2. If this is not applicable, because the substance has not yet been included in the EC Inventory, it should be identified by providing its CAS number and CAS name.
3. If the CAS information does not exist either, the substance constituent can be identified by providing one or more chemical names. In any case, REACH-IT will support the applicant in using stronger

[3] Potential registrants who manufacture or import for the first time a phase-in substance in quantities of one tonne or more per year or use for the first time a phase-in substance in the context of production of articles or import for the first time an article containing a phase-in substance that would require registration
Source: REACH Article 28 (6)

identification information if it is already available in the system by e.g. proposing similar chemical names.

Remarks:

- In case of multi-constituent substance, REACH-IT requires to pre-register each individual constituent (e.g. constituent A, constituent B, constituent C). At the end of the pre-registration process, the system will automatically assign the name “Reaction mass of A, B and C” to the newly pre-registered substance. At the end of the pre-registration period, new EC numbers will be assigned to these Reaction masses.
- UVCB substances should be pre-registered following the algorithm in place for mono-constituent substances.
- As mentioned in the introduction, only substances having all their constituents identified by an EC number can be pre-registered in bulk.

Contact information

Contact information has to be provided for each pre-registered substance. This information will be displayed in the so-called pre-SIEFs: i.e. REACH-IT space where pre-registrants of the same substance are initially put together for starting the data sharing negotiations. The following three options are available.

1. If the pre-registrant does not specify any contact information during pre-registration of a substance, the contact information provided during company sign-up will be displayed in the related pre-SIEF.
2. If the pre-registrant specifies contact information during pre-registration of a substance, this contact information will be displayed in the related pre-SIEF.
3. If the pre-registrant prefers to hand over SIEF formation and data sharing responsibility to a Third Party Representative (TPR), REACH-IT allows specifying contact information for this TPR at the level of each pre-registration.

Remarks:

- Before the TPR contact detail can be specified in a pre-registration, the TPR must have signed up in REACH-IT and received from the system a unique identifier (UUID). This UUID must be communicated to the pre-registrant before he starts pre-registering his/her substances.
- The TPR then takes over responsibility for the substance in the pre-SIEF and can coordinate with the other potential registrants.
- If such a representative is assigned to a pre-registration, his contact information, provided during his/her own sign-up, will be displayed in

the pre-SIEF, even if an additional contact person has been specified by the pre-registrant during pre-registration.

For more information on REACH and pre-registration, please consult the pre-registration section of the ECHA website at http://echa.europa.eu.

3. Bulk pre-registration

Bulk pre-registration of phase-in substances has been implemented within the REACH-IT application in order to accommodate the needs of legal entities that have to pre-register a high number of substances and where manual online creation is not a workable option.

As already mentioned, bulk pre-registration is restricted to phase-in substances identified by an EC Number (i.e. EINECS or NLP number). It means that:

- Mono-constituent or UVCB substances must be listed in the EC Inventory,
- Multi-constituent substances must have each of their constituents listed in the EC inventory.

The bulk pre-registration file containing the data to pre-register has to be generated outside REACH-IT and has to respect the REACH-IT XML pre-registration format.

3.1 Preparation of the bulk pre-registration file

The bulk pre-registration file in principle contains information on multiple pre-registrations. Each pre-registration within the file should include the following information:

- Identification of the substance to be pre-registered, in the form of EC numbers (EINECS or NLP numbers) listed for all constituents as explained in section 2 of this document,
- If appropriate: Identification of related substance(s) in the form of EC numbers (EINECS or NLP numbers) listed for all constituents of each related substance specified,
- Estimated tonnage band and envisaged registration deadline,
- Contact person information,
- Third party representative information, if appropriate,
- Remark information,
- Option to receive pre-SIEF related information in the REACH-IT internal mailbox system.

Limitation on the number of substances submitted in one file

In order to facilitate the process of the bulk pre-registration files, the amount of substances allowed in one file is limited to **500**.

This constraint was introduced for keeping the process time of a bulk file in REACH-IT at a reasonable level and to guarantee the workability of the system for the pre-registrants: if one entry in the bulk file contains invalid information, the whole bulk file will be rejected by the system. It is therefore easier for the pre-registrant to handle smaller files.

Remark

There is no limitation on the overall number of pre-registrations that a company can perform, since it is possible to submit several bulk pre-registration files.

Format of the file

The bulk pre-registration file must be fully compliant with the XML pre-registration format available on the IUCLID 5 website in order to be successfully uploaded into the REACH-IT system.

This XML format (also referred to as XML schema definition) has been published in advance of the start of the pre-registration phase, so that companies and third party vendors could prepare interfaces to existing systems and generate bulk pre-registration files in the appropriate XML format from data stored in these systems.

More information is available in the pre-registration subsection of the IUCLID 5 website at http://iuclid.eu.

Also available on the IUCLID 5 website:

- a pre-registration Toolkit containing examples of bulk pre-registration files,
- a pre-registration file checker, which enables to validate of the XML format of the pre-registration file. Note: this checker does not guarantee that the XML file can be submitted without failures.

Use of the IUCLID 5 pre-registration plug-in

The IUCLID 5 pre-registration plug-in enables companies to use information already stored in their local IUCLID 5 database. The plug-in extracts the data required for pre-registration and prepare bulk (and single) pre-registration files compliant with the pre-registration XML format.

For more information about download, installation and usage, please consult the pre-registration section of the IUCLID 5 website at http://iuclid.eu.

Use of IT systems other than IUCLID 5

Pre-registration files prepared with applications other than IUCLID 5 must fulfil both the XML validation constraints and the REACH-IT business rules to be successfully uploaded into REACH-IT.

For more information, please refer to the pre-registration section of the IUCLID 5 website at http://iuclid.eu and to Annex 1 of this document on Technical specification of bulk pre-registration files.

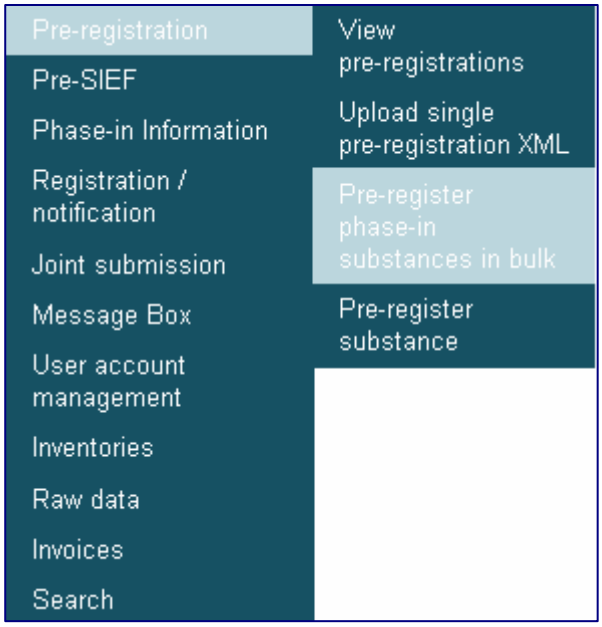
3.2 Submission of the bulk pre-registration file in REACH-IT

Pre-conditions

The user submitting the bulk pre-registration file must belong to a company already signed-up in REACH-IT and have a user role that allows him to perform the operation. In addition, the user must be logged in REACH-IT and the bulk pre-registration file must be stored in the user's personal computer.

Submission process

Step 1. From the *Pre-registration* menu select the option "Pre-register phase-in substances in bulk".



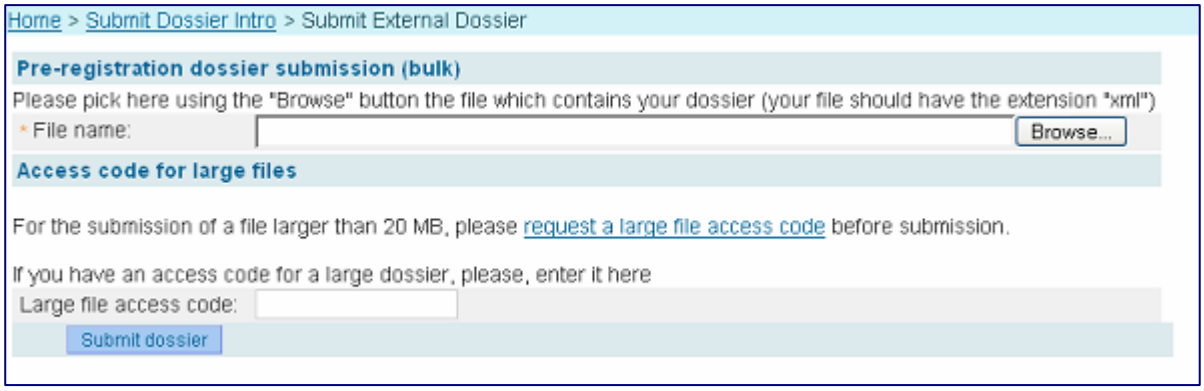
Step 2. Read the declaration which appears on the page.

If the substances in the bulk pre-registration file are all phase-in substances, tick the check-box to accept the declaration.

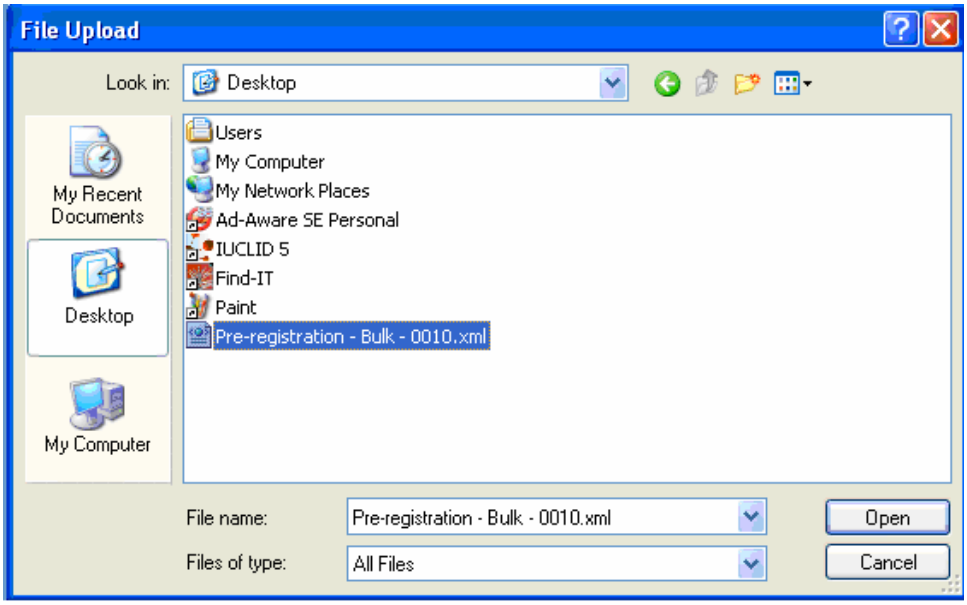
Click **Continue** to start the submission of a bulk pre-registration.

Step 3. Select the bulk pre-registration file to be submitted.

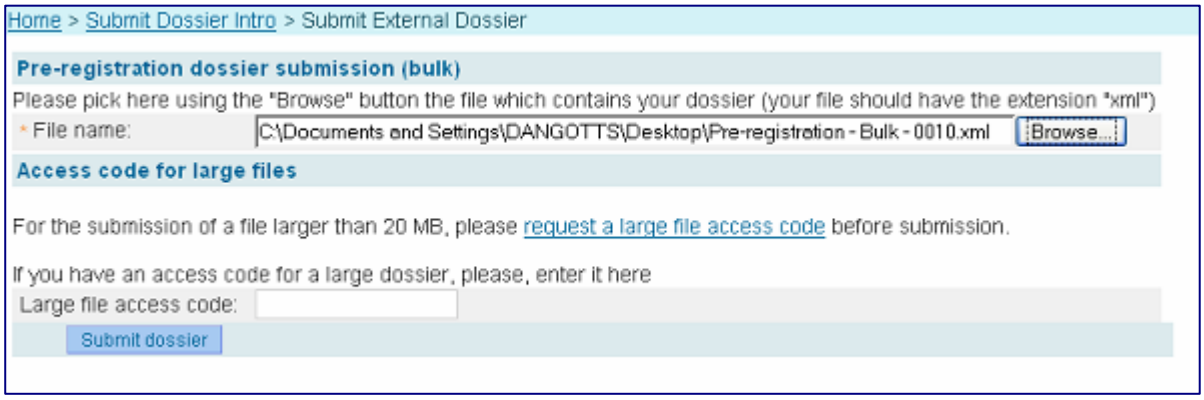
Click **Browse** to select and retrieve your bulk pre-registration file from your local PC.



Your browser opens up a “File upload” dialogue box. Select your bulk pre-registration file and click **Open**.

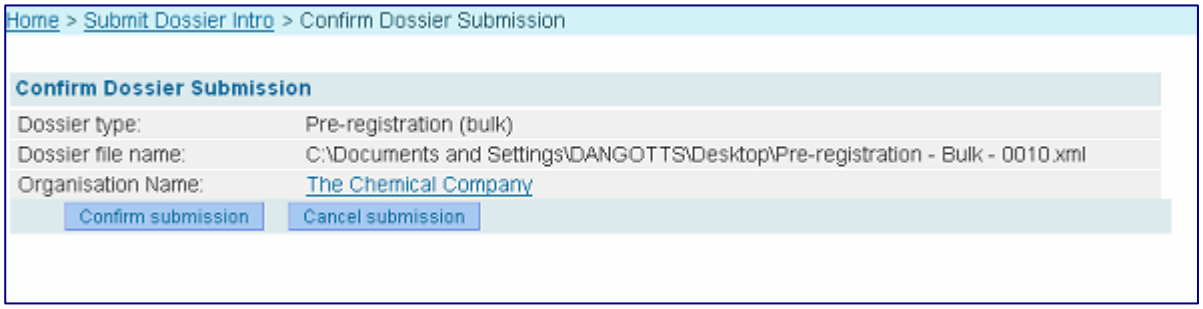


The “File name” entry field gets populated with the path to your file. Click **Submit dossier**.

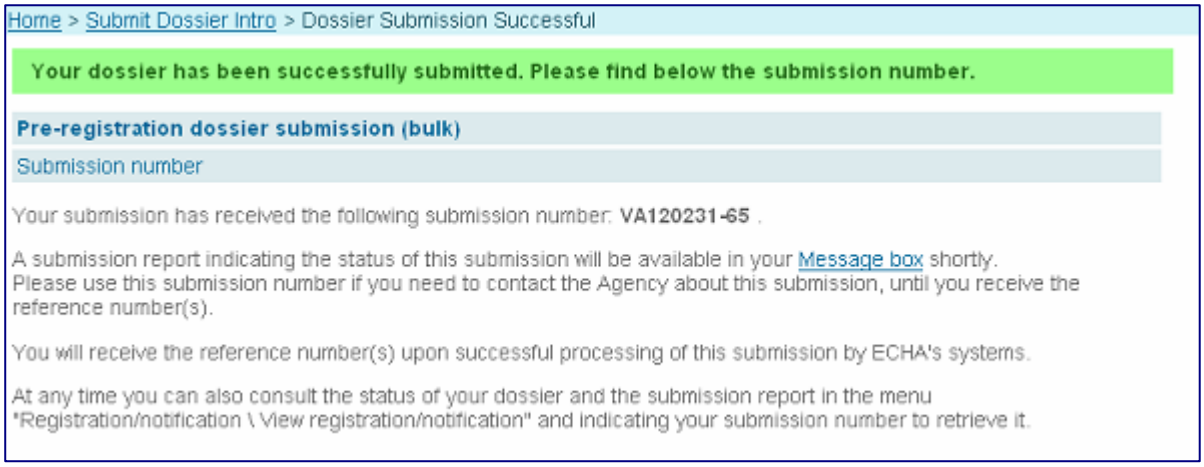


Step 4. The system will confirm your dossier type/file name and your organisation name.

If you do not wish to proceed with the submission, click **Cancel**. Otherwise, click **Confirm submission**.



Step 5. The system will confirm the successful submission and provide a submission number.



This submission number acknowledges the reception of the complete pre-registration file by the system. When the file is successfully processed, REACH-IT will then assign an individual pre-registration number to each individual substance contained in the file.

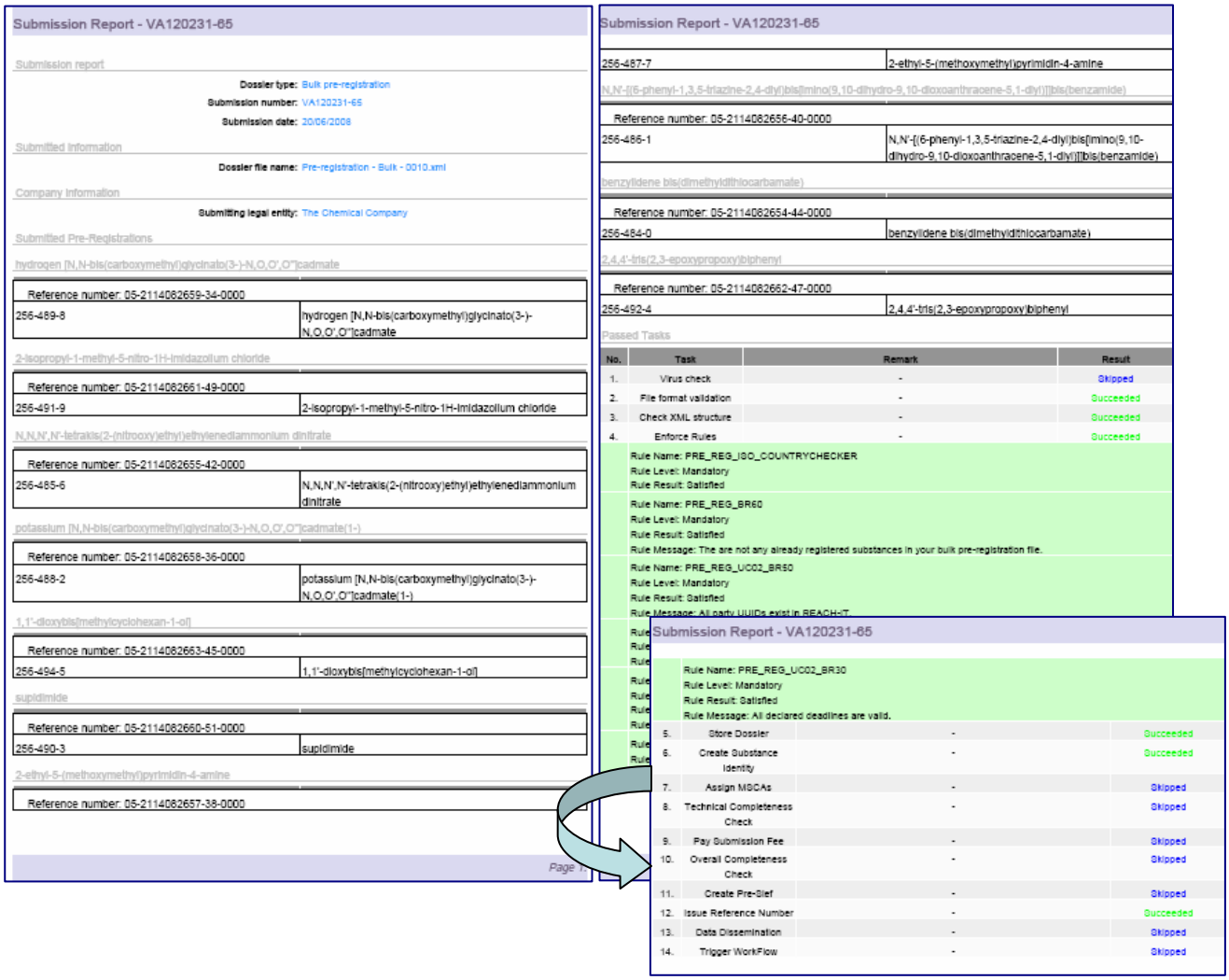
Remark:

For successful submission of one bulk pre-registration file, the system will send three internal messages:

- The first one at the time of submission, confirming that the file has been received by the system.
- The second one will indicate that reference numbers have been assigned to your pre-registrations (this message will not be sent if the file only contained updates).

- The third one indicating that the submission has reached the end of the pipeline, containing a link to the submission report.

Click Download submission report to open the report in PDF format.



3.3 Updating a pre-registration

All information within a pre-registration, apart from the substance identification, can be updated at a later time.

This can either be done either directly in REACH-IT by using the online encoding functionality or by submitting a new bulk pre-registration file containing updated information.

If the same substance identity is submitted twice with different information (e.g. different related substances), the second entry will be considered as an entry of the first one.

Remark:

This principle also applies if both entries for the substance are contained within the same file.

3.4 What happens after submission of the bulk pre-registration file?

The description below gives an overview of the steps involved to process the bulk pre-registration files.

The bulk pre-registration submission is an asynchronous process, i.e. the answer of the REACH-IT system to the user is not immediate. After submission of his/her file, the pre-registrant will receive a submission number as proof of successful submission. After this first step, the processing of the file is done asynchronously by the system without any user interaction.

The submitted files will be placed in a “queue” and processed according to the order of arrival in the queue (First in / First out principle).

Virus check

The bulk file is stored in a staging area for **6 hours**, to ensure that the virus scanner is updated and the file can be processed safely. If a virus is found, the submission will fail and the pre-registrant will be notified accordingly in the submission report.

File format validation

The format of the file is validated:

- The extension of the file should be **.xml**.
- The file should be an XML file.

If the format validation fails, the submission cannot be completed and the pre-registrant will be notified accordingly in the submission report.

Remark:

XML files can easily be altered, in particular when they are opened and edited with applications not meant to handle those files. In this case, they will not pass the format validation step.

Check of the XML structure

The system checks that the XML structure is correct, i.e. it proceeds to a number of verifications:

- the tonnage bands and registration deadlines must be written correctly,
- all mandatory fields are present in the file.

Please refer to the REACH-IT pre-registration format in the pre-registration section of the IUCLID 5 website at http://iuclid.eu for more information on the XML structure.

If the XML structure is not correct, the submission will fail and the pre-registrant will be notified accordingly in the submission report.

REACH-IT business rules

See section 4 of this document for more information on the REACH-IT internal business rules.

If these rules are not respected, the submission will fail and the pre-registrant will get a listing of all errors in the submission report.

Processing of the data contained in the bulk file

When the bulk pre-registration file reaches this step, all the data contained in the file are considered correct and the actual processing can start. It consists in:

- Checking if a pre-registration for each substance in the file already exists for the legal entity submitting the file. Based on the outcome of this check, the systems either updates the previous pre-registration or inserts the new one.
- Verifying if the contact person identified in the file already exists in the company contact list in REACH-IT. Based on this, the system either updates the previous entry or inserts a new one.
- Introducing the legal entity in the relevant pre-SIEF, identified by the contact information chosen by the pre-registrant: company details, contact person details or Third party representative details.
- Updating the pre-SIEF with information on related substances if provided in the bulk file.

If any of these tasks fails for some unidentified reason, the system will reintroduce the file into the queue. It will be processed again in due time. If the failure occurs several times, the file will be removed from the queue by an operator and the pre-registrant will be contacted.

Assignment of a Reference Number (i.e. pre-registration number)

A pre-registration number will be issued for every individual substance contained in the bulk pre-registration file if the substance does not have a pre-registration number assigned before (in case of updates).

The structure of the pre-registration number is the following:

<TYPE>-<BASE-NUMBER>-<CHECKSUM>-<INDEX-NUMBER>

Example: 05 - 1234567890 - 49 – 0000, where:

- 05 is the pre-registration submission type
- 1234567890 is the random unique 10-digit number
- 49 is the calculated checksum (changeable 2-digit number)
- 0000 is the index number

When this step is completed, the pre-registrant will get an internal message stating that reference numbers have been assigned to his pre-registration substances.

End of submission

The pre-registrant will be notified that all tasks have been completed successfully by receiving an internal message in his/her REACH-IT account. This message will contain a link to the submission report where all newly assigned reference numbers can be viewed.

Remark:

Only reference numbers assigned to initial pre-registrations are indicated in the submission report.

4. List of pre-registration Business rules

In addition to the constraints that apply to the pre-registration XML format, the REACH-IT system applies additional business rule validation on the submitted bulk pre-registration file.

If the information contained in the file does not meet the requirements of these rules, the submission will fail and the pre-registrant will be notified by an internal message explaining the reasons of the failure.

Business rule validation continues even after the first error has been found to so that the submission report contains all detected errors. It ensures that the pre-registrant can correct all errors at the same time.

The business rules enforced on the bulk pre-registration file are as follows.

- PRE_REG_ISO_COUNTRYCHECKER

  Validate that the country specified for the contact person is identified with an ISO 3166-1 - alpha-2 code.

  The full list of countries and their alpha-2 codes is available from: http://www.iso.org/iso/country_codes/iso_3166_code_lists.htm

- PRE_REG_UC02_BR50

  Validate that the Third Party Representative UUID(s) exists in the REACH-IT system. The system will only accept existing UUID(s), belonging to signed-up representatives.

- PRE_REG_BR60

  Validate that the file does not contain pre-registrations of substances for which the submitting legal entity has already submitted a registration dossier.

- PRE_REG_EC_EXISTENCE

  Validate that all substances (incl. constituents and similar substances) within the file are identified by an existing EC number.

  More information on the EC inventory can be found from the IUCLID 5 website at http://iuclid.eu.

- PRE_REG_LIMITATION

  Validate that the file does not contain more than the restricted amount of **500** substances.

- PRE_REG_UC02_BR30

  Validate tonnage band and registration deadline:
  If the tonnage band is >= 1000 tonnes the deadline must be 30/11/2010.

- PRE_REG_UC02_BR40

  Validate tonnage band and registration deadline:
  If the tonnage band is >= 100 tonnes the deadline must be either 30/11/2010 or 31/05/2013

Annex 1 - Technical example of a bulk pre-registration file

The aim of this section is to give the developers generating bulk pre-registration XML files from IT systems other than IUCLID 5 a better overview on how to create bulk pre-registration files that can be successfully submitted to the REACH-IT application.

It is recommended to create examples with the IUCLID 5 pre-registration plug-in to get an understanding of the expected XML structure. Please also consult the documentation made available with the schema definition files published on the pre-registration section of the IUCLID 5 website at http://iuclid.eu.

Structure of the bulk pre-registration XML file

For readability the XML syntax is specified (**highlighted in bold red**) for each section/field.

Examples of correctly encoded field content (**highlighted in bold green**) are also provided.

Non-mandatory fields are highlighted in ***bold italics.***
Non-mandatory fields may be omitted if no content is reported.

UTF-8 encoding

As indicated later for many fields in the XML format, all text must be encoded following the UTF-8 format. In addition, some characters have to be "escaped". The IUCLID 5 pre-registration plug-in takes care of this task automatically, but also all other software might need to encode text strings.

Field lengths

Please consult the pre-registration schema definition file to obtain the field length restrictions.

Workflow to build the XML file

1. File header (mandatory)
2. File start (mandatory)
3. Contact Persons (CPs) section start (mandatory if CP data are reported; otherwise omit sections 3. through 4.)
  3.1. CP data start (mandatory) (repeat 3.1. through 3.2. for each CP)
    3.1.1. CP Title (optional)
    3.1.2. CP First name (mandatory)
    3.1.3. CP Last name (mandatory)
    3.1.4. CP Organisation start (mandatory if Organisation data are reported; otherwise omit sections 3.1.4. through 3.1.5.)
      3.1.4.1. CP Organisation name (optional)

3.1.4.2. CP Department name (optional)

3.1.5. CP Organisation data end (mandatory if Organisation data are reported; otherwise omit sections 3.1.4. through 3.1.5.)

3.1.6. CP Contact information data start (mandatory)

3.1.6.1. CP Phone number (mandatory)

3.1.6.2. CP Fax number (optional)

3.1.6.3. CP E-mail address (mandatory)

3.1.6.4. CP Postal address data start (mandatory)

3.1.6.4.1. CP Street (mandatory)

3.1.6.4.2. CP Street 2 (optional)

3.1.6.4.3. CP Postal code (mandatory)

3.1.6.4.4. CP City (mandatory)

3.1.6.4.5. CP Region (optional)

3.1.6.4.6. CP Country (mandatory)

3.1.6.4.7. CP Envelope address (optional)

3.1.6.5. CP Postal address data end (mandatory)

3.1.7. CP Contact information data end (mandatory)

3.2. CP data end (mandatory) (repeat 3.1. through 3.2. for each different contact person)

4. CPs section end (mandatory if CP are reported; otherwise omit sections 3.1 through 3.2)

5. Pre-Registrations (PRs) section start (mandatory)

5.1. PR data start (mandatory) (repeat 5.1. through 5.2. for each PR)

5.1.1. PR Substance identity data start (mandatory)

5.1.1.1. PR Substance component data start (mandatory) (repeat 5.1.1.1. through 5.1.1.2. for each component of the Substance)

5.1.1.1.1. PR Substance component EC number (mandatory)

5.1.1.2. PR Substance component data end (mandatory) (repeat 5.1.1.1. through 5.1.1.2. for each component of the Substance)

5.1.2. PR Substance identity data end (mandatory)

5.1.3. PR contact person identifier (optional)

5.1.4. PR Third party representative UUID (optional)

5.1.5. PR Registration deadline (mandatory)

5.1.6. PR Tonnage band (mandatory)

5.1.7. PR Similar substances data start (mandatory if Similar sustances data are reported; otherwise omit sections 5.1.7. through 5.1.8.)

5.1.7.1. PR Similar substance identity data start (mandatory) (repeat 5.1.7.1. through 5.1.7.2. for each Similar substance)

5.1.7.1.1. PR Similar substance component data start (mandatory) (repeat 5.1.7.1.1. through 5.1.7.1.2. for each component of the Similar substance)

5.1.7.1.1.1. PR Similar substance component EC number (mandatory)

5.1.7.1.2. PR Similar substance component data end (mandatory) (repeat 5.1.7.1.1. through 5.1.7.1.2. for each component of the Similar substance)

5.1.7.2. PR Similar substance identity data end (mandatory) (repeat 5.1.7.1. through 5.1.7.2. for each Similar substance)

5.1.8. PR Similar substances data end (mandatory if Similar sustances data are reported; otherwise omit sections 5.1.7. through 5.1.8.)

5.1.9. PR Information (optional)

5.2. PR data end (mandatory) (repeat 5.1. through 5.2. for each PR)

6. PRs section end (mandatory)

7. File end (mandatory)

1. File header (mandatory)

```
<?XML version="1.0" encoding="UTF-8" standalone="yes"?>
```

2. File start (mandatory)

The following XML lines are mandatory for the start of the file; they also include the specification of the correct XML namespaces.

```
<preReg:BulkPreRegistration XMLns:cti="http://echa.europa.eu/schemas/contactInfo"
XMLns:preReg="http://echa.europa.eu/schemas/preRegistration"
XMLns:sid="http://echa.europa.eu/schemas/substanceIdentity"
 preReg:schemaVersion="4.0">
```

3. Contact Persons (CPs) section start

It is not mandatory to provide data on CPs. The file should contain only the CPs referred to by at least one substance.

Whenever the CP data in the file do not comply with the data formats and validations described below, normally the file will not be rejected in REACH-IT. However, an erroneous CP will be created in REACH-IT, which is not desirable since it would require further manual editing directly in REACH-IT.

The following XML line is mandatory for the start of the CPs section only if CP data are provided; otherwise, the XML lines in 3. through 4. must be omitted.

```
<preReg:ContactPersons>
```

3.1. CP data start

It is mandatory to provide a unique identifier for every CP. This identifier will be used to link the CP to the substance (**cp1** is used in the example below).

If the CP already exists in REACH-IT (based on the First name, Last name, e-mail comparison), the data will be updated. If the CP does not exist previously, it will be inserted.

The XML lines in 3.1. through 3.2. must be repeated for each CP.

The following XML line is mandatory for the start of the CP data, encode content as UTF-8.

```
<preReg:ContactPerson id="cp1">
```

3.1.1. CP Title (optional)

```
<cti:title>Mr</cti:title>
```

3.1.2. CP First name (mandatory)

```
<cti:firstName>John</cti:firstName>
```

3.1.3. CP Last name (mandatory)

```
<cti:lastName>Smith</cti:lastName>
```

3.1.4. CP Organisation data start

The following XML line is mandatory for the start of the CP organisation data only if at least one of the two fields described in 3.1.4.1. and 3.1.4.2. is reported.

```
<cti:organisation>
```

3.1.4.1. CP Organisation name (optional)

<cti:organisationName>*Company X*</cti:organisationName>

3.1.4.2. CP Department name (optional)

<cti:departmentName>*Chemical Department*</cti:departmentName>

3.1.5. CP Organisation data end

The following XML line is mandatory for the end of the CP organisation data only if at least one of the two fields described in 3.1.4.1. and 3.1.4.2. is reported.

</cti:organisation>

3.1.6. CP Contact information data start

CP Contact information data include several mandatory and optional fields, as described below.
The following XML line is mandatory for the start of the CP contact information data.

<cti:contactInformation>

3.1.6.1. CP Phone number (mandatory)

<cti:phone>
<cti:formattedNumber>+358123456789</cti:formattedNumber>
</cti:phone>

3.1.6.2. CP Fax number (optional)

<cti:fax>
<cti:formattedNumber>+358123456789</cti:formattedNumber>
</cti:fax>

3.1.6.3. CP E-mail address (mandatory)

The schema definition validates that the e-mail specific format is entered.

<cti:email>reach@companyx.com</cti:email>

3.1.6.4. CP Postal address data start (mandatory)

.
<cti:postalAddress>

3.1.6.4.1. CP Street (mandatory)

<cti:street1>Home street</cti:street1>

3.1.6.4.2. CP Street 2 (optional)

<cti:street2>*Second street*</cti:street2>

3.1.6.4.3. CP Postal code (mandatory)

<cti:zipCode>FIN-12345</cti:zipCode>

3.1.6.4.4. CP City (mandatory)

<cti:city>Kokkola</cti:city>

3.1.6.4.5. CP Region (optional)

<*cti:state*>*Ostrobothnia*</*cti:state*>

3.1.6.4.6. CP Country (mandatory)

The country in the contact person must be identified with an ISO 3166-1 - alpha-2 code. The full list of countries and their alpha-2 codes is available from:
http://www.iso.org/iso/country_codes/iso_3166_code_lists.htm

<cti:country>FI</cti:country>

3.1.6.4.7. CP Envelope address (optional)

<*cti:envelopeAddress*>*my envelope address*</*cti:envelopeAddress*>

3.1.6.5. CP Postal address data end (mandatory)

</cti:postalAddress>

3.1.7. CP Contact information data end (mandatory)

</cti:contactInformation>

3.2. CP data end (mandatory)

The XML lines in 3.1. through 3.2. must be repeated for each CP.

</preReg:ContactPerson>

4. CPs section end (mandatory)

The following XML line is mandatory for the end of the CPs section only if CP data are provided; otherwise, the XML lines in 3. through 4. must be omitted.

</preReg:ContactPersons>

5. Pre-registrations (PRs) section start (mandatory)

<preReg:PreRegistrations>

5.1. PR data start (mandatory)

Please note! The amount of substances allowed in a pre-registration file is in REACH-IT restricted to **500**.

This element contains the enableMessages attribute that defines if the company wishes to receive internal messages related to changes in the pre-SIEF for this substance. The attribute is either **true** or **false.**

The XML lines in 5.1. through 5.2. must be repeated for each PR.

One of the following alternative XML lines is mandatory for the start of the PR data.

<preReg:PreRegistration enableMessages="true">
or
<preReg:PreRegistration enableMessages="false">

5.1.1. PR Substance identity data start (mandatory)

The schema definition constraints duplicates in mono-constituent Substances (i.e. same EC number) and the third party software's should make sure that similar multi-constituents are not entered into the same file. In case this happens REACH-IT will consider the second entry as an update of the first.

Please note! Take caution that substance Identity can not be updated in REACH-IT.

<preReg:substanceIdentity>

5.1.1.1. PR Substance component data start (mandatory)

A mono-constituent is identified with one constituent and multi-constituents can have up to 10 contistuents. The XML lines in 5.1.1.1. through 5.1.1.2. must be repeated for each component of a Substance.

<sid:SubstanceConstituentIdentity>

5.1.1.1.1. PR Substance component EC number (mandatory)

This element must contain the EC number formatted as shown in the example below.

<sid:ECIdentity ecNumber="200-006-5" />

5.1.1.2. PR Substance component data end

The XML lines in 5.1.1.1. through 5.1.1.2. must be repeated for each component of a Substance.

</sid:SubstanceConstituentIdentity>

5.1.2. PR Substance identity data end

</preReg:substanceIdentity>

5.1.3. PR Contact person (CP) identifier (optional)

This element is used to link this specific substance to one of the contact persons defined earlier in the file.

<preReg:contactPerson idref="cp1" />

5.1.4. PR Third party representative (TPR) UUID (optional)

If used this element must contain a Company UUID of a Third party representative signed-up to REACH-IT.
The file must contain the UUID formatted as shown in the example below.

<preReg:thirdPartyRepresentative partyUUID="IUC5-c41494a4-b78c-4c14-84e7-c3b529fff842" />

5.1.5. PR Registration deadline (mandatory)

The file must contain one of the values **2010-11-30** or **2013-05-31** or **2018-05-31** as shown below.

Please note! REACH-IT has the following constraints on the deadlines:

- if the tonnage band is **1000+,** the deadline must be **2010-11-30**
- if the tonnage band is **100-1000,** the deadline must be either **2010-11-30** or **2013-05-31**

One of the following lines must be entered

<preReg:registrationDeadline>2010-11-30</preReg:registrationDeadline>
or
<preReg:registrationDeadline>2013-05-31</preReg:registrationDeadline>
or
<preReg:registrationDeadline>2018-05-31</preReg:registrationDeadline>

5.1.6. PR Tonnage band (mandatory)

The file must contain one of the values **1000+** or **100-1000** or **10-100** or **1-10** as shown below.

Please note constraints specified in 5.1.5

<preReg:tonnageBand>1000+</preReg:tonnageBand>
or
<preReg:tonnageBand>100-1000</preReg:tonnageBand>
or
<preReg:tonnageBand>10-100</preReg:tonnageBand>
or
<preReg:tonnageBand>1-10</preReg:tonnageBand>

5.1.7. PR Similar substances data start (optional)

It is not mandatory to provide data on Similar substances for a pre-registered Substance.

The following XML line is mandatory for the start of the PR Similar substances data only if Similar substances data are provided; otherwise, the XML lines in 5.1.7. through 5.1.8. must be omitted.

<preReg:similarSubstanceIdentities>

5.1.7.1. PR Similar substance identity data start

Make sure to avoid indicating a similar substance identical to the pre-registered substance.

In REACH-IT a link will be created from the PR to the Similar substance. The link from a PR to a Similar substance can be deleted in REACH-IT.

The XML lines in 5.1.7.1. through 5.1.7.2. must be repeated for each Similar substance.

<preReg:SubstanceIdentity>

5.1.7.1.1. PR Similar substance component data start (mandatory)

A mono-constituent similar substance is identified with one constituent and multi-constituents can have up to 10 constituents.

The XML lines in 5.1.7.1.1. through 5.1.7.1.2. must be repeated for each component of a Similar substance.

```
<sid:SubstanceConstituentIdentity>
```

5.1.7.1.1.1. PR Similar substance component EC number

This element must contain the EC number formatted as shown in the example below.

```
<sid:ECIdentity ecNumber="200-007-0" />
```

5.1.7.1.2. PR Similar substance component data end

The XML lines in 5.1.7.1.1. through 5.1.7.1.2. must be repeated for each component of a Similar substance.

```
</sid:SubstanceConstituentIdentity>
```

5.1.7.2. PR Similar substance identity data end

The XML lines in 5.1.7.1. through 5.1.7.2. must be repeated for each Similar substance.

```
</preReg:SubstanceIdentity>
```

5.1.8. PR Similar substances data end

The following XML line is mandatory for the end of the PR Similar substances data only if Similar substances data are provided; otherwise, the XML lines in 5.1.7. through 5.1.8. must be omitted.

```
</preReg:similarSubstanceIdentities>
```

5.1.9. PR Information (optional)

Please note! In the IUCLID 5 pre-registration plug-in you can add up to 32,768 characters. The schema definition only allows 4000.

```
<preReg:information>my info</preReg:information>
```

5.2. PR data end

The XML lines in 5.1. through 5.2. must be repeated for each PR.

```
</preReg:PreRegistration>
```

6. PRs section end (mandatory)

```
</preReg:PreRegistrations>
```

7. File end (mandatory)

```
</preReg:BulkPreRegistration>
```

Final example

```
<?XML version="1.0" encoding="UTF-8" standalone="yes"?>
<preReg:BulkPreRegistration
XMLns:cti="http://echa.europa.eu/schemas/contactInfo"
XMLns:preReg="http://echa.europa.eu/schemas/preRegistration"
XMLns:sid="http://echa.europa.eu/schemas/substanceIdentity"
preReg:schemaVersion="4.0">
   <preReg:ContactPersons>
      <preReg:ContactPerson id="cp1">
         <cti:title>Mr</cti:title>
         <cti:firstName>John</cti:firstName>
         <cti:lastName>Smith</cti:lastName>
         <cti:organisation>
            <cti:organisationName>Company X</cti:organisationName>
            <cti:departmentName>Chemical Department</cti:departmentName>
         </cti:organisation>
         <cti:contactInformation>
            <cti:phone>
               <cti:formattedNumber>+358123456789</cti:formattedNumber>
            </cti:phone>
            <cti:fax>
               <cti:formattedNumber>+358123456789</cti:formattedNumber>
            </cti:fax>
            <cti:email>reach@companyx.com</cti:email>
            <cti:postalAddress>
               <cti:street1>Home street</cti:street1>
               <cti:street2>Second street</cti:street2>
               <cti:zipCode>FIN-12345</cti:zipCode>
               <cti:city>Kokkola</cti:city>
               <cti:state>Ostrobothnia</cti:state>
               <cti:country>FI</cti:country>
               <cti:envelopeAddress>my enveloaddress</cti:envelopeAddress>
            </cti:postalAddress>
         </cti:contactInformation>
      </preReg:ContactPerson>
   </preReg:ContactPersons>
   <preReg:PreRegistrations>
      <preReg:PreRegistration enableMessages="true">
         <preReg:substanceIdentity>
            <sid:SubstanceConstituentIdentity>
               <sid:ECIdentity ecNumber="200-006-5" />
            </sid:SubstanceConstituentIdentity>
         </preReg:substanceIdentity>
         <preReg:contactPerson idref="cp1" />
         <preReg:thirdPartyRepresentative partyUUID="xxxxxxxx" />
         <preReg:registrationDeadline>2010-11-30</preReg:registrationDeadline>
         <preReg:tonnageBand>1000+</preReg:tonnageBand>
         <preReg:similarSubstanceIdentities>
            <preReg:SubstanceIdentity>
               <sid:SubstanceConstituentIdentity>
                  <sid:ECIdentity ecNumber="200-007-0" />
               </sid:SubstanceConstituentIdentity>
            </preReg:SubstanceIdentity>
         </preReg:similarSubstanceIdentities>
         <preReg:information>my info</preReg:information>
      </preReg:PreRegistration>
   </preReg:PreRegistrations
</preReg:BulkPreRegistration>
```

European Chemicals Agency Annankatu 18, P.O. Box 400, FI-00121 Helsinki, Finland
Tel.: +358 9 6861 80 | Fax +358 9 6861 8210 | http://echa.europa.eu | info@echa.europa.eu