Standard Operating Procedure

Number: **012**

Name: **Data standardisation**

Title: Procedure for the standardisation of information assets on the EU Pollinator Hub

Category: Data management

Status: Approved

Version: 1.0

Date: 2024-02-07

Authors: Michael Rubinigg

Table of contents

1   Document history

1.1   Current Release

1.1.1   Drafted

Submitted on 2023-10-26 by Michael Rubinigg, quality manager.

1.1.2   Reviewed

Reviewed on 2023-11-14 by Noa Simòn Delso, general manager.

1.1.3   Approved

Approved on 2024-02-07 by Michael Rubinigg, quality manager.

1.2   Change history

n.a.

2   Abbreviations

EU   European Union

EUPH   EU Pollinator Hub

n.a.   not available

SOP   Standard Operating Procedure

UID   Unique Identifier

URL   Uniform Resource Locator

WI   Work Instructions

3   Introduction

The aim of the present SOP is to provide a procedure for the definition of classes and their translations into different languages (referred to as expressions) through a collaborative, software-based process. The software allows registered users to participate in a fully traceable process, including the possibility to submit and edit proposals, the possibility to comment on proposals and the approval of proposals by higher-level-entities (supervisors).

3.1   Definitions

Class is a data element, which describes objects, organisms, abstract concepts or processes that are sufficiently similar to be considered different from any other class in terms of their purpose.

Expression is defined as a single word, a group of words or a phrase in human language used to describe a class. Different expressions (synonyms) may be used to describe the same class in any given language.

Unit Catalog is defined as an application in the EUPH platform management tool used to manage dataset *Units*.

4   Validity

4.1   Area

The of validity of this document is limited to the process of standardisation (definition, translation) of classes in dataset *Dictionary* and units in dataset *Units*.

4.2   Time

The period of validity of this SOP begins on 2024-02-07 and ends with its revocation.

4.3   Related Documents

Document type	Number	Name
SOP	001	SOP instructions
SOP	004	Organisation chart
SOP	023	Method validation
WI	007	Citation management
WI	014	Dictionary management
WI	015	Translation management
WI	016	Dictionary application
WI	022	Discussion management

4.4   Related Tasks

Persons who have been assigned one of the following tasks, as defined in SOP-004, assume a responsibility within this SOP:

- Data standardisation

5 Procedure

5.1 Responsibilities

- The person responsible for data standardisation on the EUPH, as defined by SOP-004, is responsible for the management of the processes defined in this SOP.
- The task of carrying out certain parts of this process can be transferred to another person.

5.2 Classes

5.2.1 Definition

5.2.1.1 Application

- The process of standardisation of classes is carried out with the application *Dictionary*, accessed via the EUPH Application Module.

5.2.1.2 Scope

- WI-007, WI-014, WI-016 and WI-022 are used to execute all processes defined in section 5.2.1.
- The standardisation process includes the creation, update and deletion of a class and the process of its definition.

5.2.1.3 Assignment of a description to a class

- Unique descriptions of a class, fulfilling the following criteria, mut be provided:
  - A class must be described by single words (*e.g.* nouns, verbs, adjectives/adverbs) or, if not possible, by groups of words (*e.g.* compound nouns, compound verbs) or, if not possible, by a phrase.
  - The English language must be used to describe a class.
  - The description of a class must be as brief as possible, linguistically correct, and contain all essential information to distinguish it from other classes.

5.2.1.4 Assignment of draft status

- The status of a class is set to *Draft*.

5.2.1.5 Association to a specific dataset

- A class may be associated with an existing dataset hosted on the EUPH if the provider of this dataset has an understandable interest in defining the term according to his/her own criteria.
- If an equivalent class exists in the dataset, which is not associated to an existing dataset, the class that is associated to an existing dataset is defined as analogous to the class that is not associated to an existing dataset.

5.2.1.6 Association of parent entities

- A class is associated with another class, referred to as parent class, if the class is a subset of the parent class.

- A class can only be associated with a specific dataset if the parent class is associated with the same dataset, with the exception of classes acting as root classes for a given set of classes related to a dataset.

5.2.1.7   Approval of a class

- A class is released when the manager of the discussion process, as defined in section 5.2.1.1, in collaboration with those participating in the process, has decided that the discussion process on its definition and its association with other classes has produced a satisfactory result.
- The status of a class is set to *Approved*.
- The validity of a class starts at the date and time of its release.

5.2.1.8   Update of a class

- If any change in the parameters of an approved class has to be applied, the manager starts an update process of a class.
- The validity of an existing class ends at the date and time at which an update with the same UID is released.
- A class is updated and released as described in section 5.2.1.4 through 5.2.1.7.

5.2.1.9   Deletion

- The manager of the process may delete a given class if keeping the class in the dictionary is no longer considered useful, if it is not a parent of another class and if it does not act as analogue of another class.
- The status of a class is set to *Expired*.
- The validity of a class ends at the date and time of its deletion.

5.2.2   Translation

5.2.2.1   Scope

- WI-007, WI-015, WI-016 and WI-022 are used to execute all processes defined in section 5.2.2.
- The translation process includes the creation, update and deletion of expressions for classes in human languages, as defined in section 5.2.2.

5.2.2.2   Assignment of draft status

- The status of an expression is set to *Draft*.

5.2.2.3   Approval of an expression

- An expression is released when the manager of the standardisation process, as defined in section 5.2.1.1, in collaboration with those participating in the process, has decided that the discussion process on its definition and its association with other expressions has produced a satisfactory result to which a majority of the parties agreed upon.
- The status of an expression is set to *Approved*.
- The validity of an expression starts at the date and time of its release.

5.2.2.4   Update of an expression

- If any change in the parameters of an approved expression has to be applied, the manager starts an update process of an expression.
- An expression is updated and released as described in section 5.2.2.2 through 5.2.2.3.

5.2.2.5   Deletion of an expression

- The manager of the process may delete a given expression, if keeping the class in the dictionary is no longer considered useful, and if it is not assigned as synonym by any other expression.

5.2.3   Public participation

5.2.3.1   Inclusion into the process

- WI-016 is used to execute all processes defined in section 5.2.3.
- Users with an approved account on the EUPH may participate in the process of definition and translations of classes.

5.2.3.2   Exclusion from the process

- Users who severely interfere with the standardisation process or violate legal regulations must be denied access to the EUPH.

5.3   Descriptors

5.3.1   Global descriptors

5.3.1.1   Application

- The process of standardisation of global descriptors is carried out with the application *Descriptors*, accessed via the EUPH Platform Management tool.

5.3.1.2   Initiation of the process

- The person responsible for the management of this processes may at any time initiate the creation, update or deletion of a global descriptor in the Descriptor Catalog.
- The person responsible for the management of this process creates a global descriptor not contained in the Descriptors Catalog if a formal request has been made.

5.3.1.3   Creation

- A global descriptor shall be linked to a class.
- A global descriptor shall receive the description of the class to which it is linked.
- The validity of a descriptor starts at the date and time of its release.

5.3.1.4   Update

- The validity of an existing descriptor ends at the date and time at which an update with the same UID is released.
- A descriptor is updated as described in section 5.3.1.1.

5.3.1.5   Deletion

- The status of a descriptor is set to *Deprecated*.
- The validity of a descriptor ends at the date and time of its deletion.

5.5   Units

5.5.1   Management of the Units Catalog

5.5.1.1   Application

- The process of standardisation of units is carried out with the application *Units*, accessed via the EUPH Platform Management tool.

5.5.1.2   Initiation of the process

- The person responsible for the management of this processes may at any time initiate the creation, update or deletion of a unit in the Units Catalog.
- The person responsible for the management of this process creates a unit not contained in the Units Catalog if a formal request has been made.

5.5.1.3   Creation

- The validity of a unit starts at the date and time of its release.
- Algorithms for the conversion of a given unit to a base unit are validated according to SOP-023 (Method validation).

5.5.1.4   Deletion

- The status of a unit is set to *Deleted*.
- The validity of a unit ends at the date and time of its deletion.

6   References

There are no sources in the current document.

7   Annexes

There are no annexes in the current document.