|  |  |
| --- | --- |
|  | EUROPEAN COMMISSIONDIRECTORATE-GENERALEDUCATION AND CULTURE |

DG EAC Information Systems

Learning Opportunities and Qualifications in Europe

|  |  |
| --- | --- |
| Date: | 02/05/2014 |
| Authors: | Notis Pikros ,Vangelis Xanthakis |
| Revised by: | Karin van der Sanden |

# Overview

The portal Learning Opportunities and Qualifications in Europe allows searching for learning opportunities and qualifications using a faceted search approach; this approach allows users to browse results by iteratively narrowing the scope of their search using applicable criteria (utilizing metadata retrieved from diverse “**Data Sources**” like the internal database, National Data Sources and National Coordination Points).

To facilitate the faceted search interface, the portal Learning Opportunities and Qualifications in Europe has to locally index the data available at the Data Source local databases. This means that the portal Learning Opportunities and Qualifications in Europe stores the required data in the Learning Opportunities and Qualifications in Europe database instead of retrieving the data when queries are requested.

The architecture is a modular approach consisting of three (3) operations. The following figure illustrates the flow of operations:



Figure 1‑1: Learning Opportunities and Qualifications in Europe Interconnection interface overview

The staging area serves as an intermediate data storage mechanism, removing the burden from data providers to implement an interface for searching/filtering and exposing their data.

The staging area requires an XML file from the data providers, with a full or incremental data export of qualifications/learning opportunities residing in the Data Source database.

The XML file, placed inside a zip file, shall provide for a **specific data source** either:

* a **full data set**: that includes a list with all the entities that will replace everything in the staging area,
* a **partial data set**: that includes a list of entities to be updated/created,
* a **delete data set**: that includes a list of entities to be deleted

# Data Source Development

From the Data Source point of view in order to implement the proposed approach and provide information to the staging area, the following actions are required:

## Summary of information requested/provided

The following information is provided or requested:

|  |  |
| --- | --- |
| **Document** | **Comment** |
| **Requested documents:** | These documents should be completed by the Data Sources and provided to LO and Q in Europe Portal |
|  LO and Q in Europe Data-Source-Metadata.doc | Holds Data Source information (such as name) |
| LO and Q in Europe-LO-Mapping.xls | Learning Opportunities mapping document |
| LO and Q in Europe-Qualifications-Mapping.xls | Qualifications mapping document |
|  |  |
| **Provided documents:** |  |
| LO and Q in Europe Interconnection.doc | This document - Learning Opportunities and Qualifications in Europe Interconnection Interface & Guidelines |
| learningOpportunitiesSample.zip | Sample Learning Opportunities XML |
| qualificationsSample.zip | Sample Qualifications XML |
| learningOpportunities.xsd | Learning Opportunities XML Schema |
| qualifications.xsd | Qualifications XML Schema |
| delete.xsd | Delete records XML Schema |

Table 2‑1: Summary of information requested/provided

The following sections detail how this information is used or provided.

## Provide Data Source metadata

The metadata include:

* Data Source Name
* The URL of the Data Source web site
* A description of the Data Source in English, official language and any additional languages.
* The official logo of the Data Source in JPG or PNG format. The dimensions should not exceed 250x100 pixels.
* Indication of what data the Data Source provides, such as Qualifications and/or Learning Opportunities

The metadata are provided in the form of a word document:

* “LO and Q in Europe Data-Source-Metadata.doc”

## Map Data Source database elements to the Learning Opportunities and Qualifications in Europe elements

This step helps to document the mapping of data source elements to the Learning Opportunities and Qualifications in Europe elements. This mapping assists in the extraction of Data Source data in the XML format required by the Learning Opportunities and Qualifications in Europe portal.

The following excel files are provided to describe the mapping of the data source fields to the Learning Opportunities and Qualifications in Europe fields:

* **Learning Opportunities**: LO and Q in Europe-LO-Mapping.xls
* **Qualifications**: LO and Q in Europe-Qualifications-Mapping.xls

Mapping involves the following two (2) steps:

1. Mapping datasource **fields** to Learning Opportunities and Qualifications in Europe **fields**.

For example the datasource field for the Learning Opportunities and Qualifications in Europe "Local\_ID" may be "ID" in the local database, Learning Opportunities and Qualifications in Europe field "Requirements" may be "Prerequisites" in the local database, Learning Opportunities and Qualifications in Europe field "Start\_date" may be "Start\_date" in the local database.

The **field** mapping should be described in the following sheets of the mapping excels:

* + **Learning Opportunities**: sheet "**LO**" in LO and Q in Europe-LO-Mapping.xls
	+ **Qualifications**: sheet "**Qualification**" in LO and Q in Europe-Qualifications-Mapping.xls
1. Mapping datasource **field** **values** to Learning Opportunities and Qualifications in Europe **field** **values**,

For example the local subject categories should be mapped to the Learning Opportunities and Qualifications in Europe subject values, the local study type values should be mapped to the Learning Opportunities and Qualifications in Europe study type values.

The field **value** mapping should be described in the corresponding sheets of the mapping excels, for example "EQF Level" values in sheet "EQF Level", "Languages" values in sheet "Languages, etc.

Please note the following mapping considerations:

* Text field lengths
	+ All text fields are limited to 4000 characters unless stated otherwise inside the mapping excel files.
	+ The “description” field allows larger texts.
* The URL/link fields must include the http:// or https:// prefix.
* It is intended to link Learning Opportunities records to Qualification records via the fields below, so an effort should be made to provide matching values for these fields:
	+ "QualificationAwarded" field of Learning Opportunity records
	+ "Title" field of Qualification records

The mapping process should involve both technical people that understand the local database fields and education people that understand how the values of the fields should be mapped.

## Create the XML file

This step requires the development of a mechanism to extract data from the Data Source database into an XML file. It should use as input the mapping of the previous step (section 2.3 Map Data Source database elements to the Learning Opportunities and Qualifications in Europe elements).

The XML files **must validate** against the corresponding XML schema (XSD):

* **Qualifications**: qualifications.xsd
* **Learning Opportunities**: learningOpportunities.xsd
* **Delete**: delete.xsd

### Full/Incremental update

A **Qualifications** XML must include a Qualifications root element and one Qualification element for each Qualification, as shown below:

<?xml version="1.0" encoding="UTF-8"?>

<tns:Qualifications

 **DatasourceComment**="This is the datasource comment"

 **Incremental**="true"

 **Key**="XXXXXXXXXXXXXXXXXX"

 **XsdType**="Qualification" **XsdVersion**="XXX"

 xmlns:tns="http://ec.europa.eu/Qualifications/vXXX"

 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

 xsi:schemaLocation="http://ec.europa.eu/Qualifications/vXXX
 qualifications\_v0.x.y.xsd ">

 <tns:Qualification>

…

 </tns:Qualification>

 <tns:Qualification>

…

 </tns:Qualification>

 <tns:Qualification>

…

 </tns:Qualification>

</ tns:Qualifications>

A **Learning Opportunities** XML must include a LearningOpportunities root element and one LearningOpportunity element for each Learning Opportunity, as shown below:

<?xml version="1.0" encoding="UTF-8"?>

<tns:LearningOpportunities

**DatasourceComment**="This is the datasource comment"

 **Incremental**="false"

 **Key**="XXXXXXXXXXXXXXXXXXXXXX"

 **XsdType**="Learning Opportunity" **XsdVersion**="XXX"

 xmlns:tns="http://ec.europa.eu/LearningOpportunities/vXXX"

 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

 xsi:schemaLocation="http://ec.europa.eu/LearningOpportunities/vXXX
 learningOpportunities\_v0.x.y.xsd ">

 <tns:LearningOpportunity>

…

 </tns:LearningOpportunity>

 <tns:LearningOpportunity>

…

 </tns:LearningOpportunity>

 <tns:LearningOpportunity>

…

 </tns:LearningOpportunity>

</ tns:LearningOpportunities>

The XML files share the following common metadata attributes on the root element:

* The **Key** attribute is mandatory and it is the identifier of the owner/provider of the data. The Key will be supplied confidentially to each Data Source.
* The **Incremental** attribute is optional and defines whether the update is an incremental (true) or a full one (false). If the attribute is missing the update is considered as full.
* The **XsdType** attribute is mandatory and defines the type of the XSD (XML schema) that the XML should comply with. Together with the XsdVersionserve as reference to the XSD.
* The **XsdVersion** attribute is mandatory and defines the XSD version of the XsdType that the XML should comply with. Together with the XsdTypeserve as reference to the XSD. The XXX are used instead of the real version (e.g. XsdVersion="0.3.4") to avoid referencing obsolete versions due to future version updates.
* The **DatasourceComment** attribute is an optional string field that allows the datasource to comment on the XML file, such as date generated, datasource versioning information, etc.

This field is optionally provided by the datasource and is currently ignored during the processing of the XML.

Since the XML file must support internationalized data the file encoding must be UTF-8.

### Translation of text fields

Providing a translation for a given text element (e.g. Title) can be done by adding a new node of the same element, but with the attribute “language” set to the language of the provided translation.

For example adding a translation for title and description is done as follows:

<?xml version="1.0" encoding="UTF-8"?>

<tns:LearningOpportunities

**DatasourceComment**="This is the datasource comment"

 **Incremental**="true"

 **Key**="XXXXXXXXXXXXXXXXXXXXXX"

 **XsdType**="Learning Opportunity" **XsdVersion**="XXX"

 xmlns:tns="http://ec.europa.eu/LearningOpportunities/vXXX"

 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

 xsi:schemaLocation="http://ec.europa.eu/LearningOpportunities/vXXX
 learningOpportunities\_v0.x.y.xsd ">

<tns:LearningOpportunity>

 <tns:LearningOpportunityId>123</tns:LearningOpportunityId>

 <tns:Title>Title</tns:Title>

 <tns:Title **language="de"**>Titel</tns:Title>

 <tns:Description>Description</tns:Description>

 <tns:Description **language="de"**>Beschreibung</tns:Description>

 …

</tns:LearningOpportunity>

…

</ tns:LearningOpportunities>

Elements with missing “language” attribute are considered as using the default language (e.g. language="en"). An English translation is **mandatory**; either by explicitly setting language = "en" or omitting the language attribute for an element.

Translation is provided for most text fields. Fields that are not translatable are classified data (e.g. Subject, Education Level, etc.) and non-translatable data (e.g. phone, email, postal code, etc.). More information can be found in the mapping documents.

### Delete records

A **Delete** XML must include a Delete root element and an Id element for each entity that will be deleted, as shown below:

<?xml version="1.0" encoding="UTF-8"?>

<tns:Delete xmlns:tns="http://ec.europa.eu/delete/vXXX"

 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

 **Key**="XXXXXXXXXXXXXXXXXXXXXXXXXXX"

 **XsdVersion**="XXX" **XsdType**="Delete">

 <tns:**EntityType**>Learning Opportunity</tns:**EntityType**>

 <tns:Id>Id0</tns:Id>

 <tns:Id>Id1</tns:Id>

</tns:Delete>

The **Key** attribute is mandatory and it is the identifier of the owner/provider of the data. The Key will be supplied confidentially to each Data Source. The **XsdType** and **XsdVersion** are mandatory attributes that define the type and the version of the XSD (XML schema) that the XML should comply with.

The **EntityType** element is mandatory and defines whether the entities to be deleted are Learning Opportunities (Learning Opportunity) or Qualifications (Qualification).

### Tips

The following are some suggestions for creating the XML files:

* For fields with sub-levels such as subjects, it is necessary to provide a value only at the lowest child level; there is no need to provide a value both at a child level and parent level.
* HTML input shall be inside CDATA section (<![CDATA[\*\*\*\*\*]]>).
* All fields with text values can have HTML input but with the following limitations in the allowed tags (and their attributes): "a", "b", "blockquote", "br", "cite", "code", "dd", "dl", "dt", "em","i", "li", "ol", "p", "pre", "q", "small", "strike", "strong", "sub","sup", "u", "ul"

### Sample XML files

The following sample XML files are provided:

Learning Opportunities:

* lo\_full\_sample.xml: Example for full LO update
* lo\_incremental\_1\_sample.xml: Example for incremental update based on full sample(first entity only)
* lo\_incremental\_2\_sample.xml: Example for incremental update based on full sample(second entity only)
* lo\_translation\_sample.xml: Example with translations of text fields.

Qualifications:

* qual\_full\_sample.xml: Example for full qualification update
* qual\_incremental\_1\_sample.xml: Example for incremental update based on full sample(first entity only)
* qual\_incremental\_2\_sample.xml: Example for incremental update based on full sample(second entity only)
* qual\_translation\_sample.xml: Example with translations of text fields.

Delete records:

* lo\_delete\_sample.xml: example on how to delete qualifications
* qual\_delete\_sample.xml: example on how to delete learning opportunities

## Upload XML

This step includes the packaging of the XML file into a zip file and uploading via the portal's graphical user interface

For performance reasons it is required that the XML is compressed before submitted; for compression the DEFLATE algorithm should be used and the compressed file should be in ZIP file format.

The Staging Area can be populated with the following ways:

* Bulk data upload Interface of the Learning Opportunities and Qualifications in Europe back office
* Bulk data upload Client: Automated tool (java based) that monitors a defined folder and connects Learning Opportunities and Qualifications in Europe back office to upload XML data files.

Details about the Learning Opportunities and Qualifications in Europe back office are provided in the "Data Manager User Guide".