INCLUDING QUALIFICATIONS IN THE LEARNING OPPORTUNITIES AND QUALIFICATIONS IN EUROPE PORTAL

Technical specifications and metadata schemata

Contents

1. INTRODUCTION

I. The need for interoperability

When publishing information on qualifications, the various stakeholders involved in the process (national public authorities, awarding bodies, citizens and the Commission) want efficient and effective interactions between them. Interoperable information exchanges are the solutions to meet the stakeholders' demands. The exchanges in question will ensure the connection and linking of national qualification databases to the Learning Opportunities and Qualifications in Europe (LOQ) portal[1].

According to the European Interoperability Framework (EIF) [2] "interoperability, within the context of European public service delivery, is the ability of disparate and diverse organisations to interact towards mutually beneficial and agreed common goals, involving the sharing of information and knowledge between the organisations, through the business processes they support, by means of the exchange of data between their respective ICT systems".

The EIF describes four levels of interoperability: legal, organizational, semantic and technical. This document focuses on semantic and technical interoperability:

- **Semantic interoperability** enables organizations to process information from external sources in a meaningful manner and to ensure that the precise meaning of exchanged information is understood and preserved throughout exchanges between parties.
- **Technical interoperability** ensures that any systems used by the parties are technically linked.

Concerning the publishing of qualifications, the combination of semantic and technical interoperability will allow the national qualification databases and the LOQ portal to be linked together and to exchange meaningful information.

II. Why do we aim for interoperability?

The Member States have different education and training systems. The Commission and the European Qualification Framework Advisory Group (EQF AG) aim for a better comparability and transparency of qualification systems and qualifications across Europe. Making it easier to access and reuse information on qualifications on the web, as well as on other search and delivery platforms, contributes to this transparency.

Information can only be comparable throughout Europe if all Member States and other actors have a common view on attributes that are specific for qualifications, such as: title, learning outcomes profile (knowledge, skills and competence), awarding body (AB), EQF level and relations to occupations. Everyone who publishes information on qualifications needs to make sure that users can easily find what they are searching for.

An important part of this common language is an agreed 'metadata schema'. Such a schema tells IT systems, search engines and web portals, such as the LOQ portal, how to search for information and how to use existing sources already available on the web.

---

1 https://ec.europa.eu/ploteus/
2 http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf

III.   What are metadata and what is a metadata schema?

Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource[3]. Metadata is often called data about data or information about information.

The National Information Standards Organization (NISO) defines three main types of metadata:

- **Descriptive:** describes a resource for purposes such as discovery and identification. It can include elements such as the title of the awarding body (AB);
- **Administrative:** provides information to help manage a resource, like the "modification date" field of a Qualification.
- **Structural metadata:** explains how the data are structured together. A list of Learning Outcomes structuring a Qualification for example.

A schema is a logical plan showing the relationships between metadata elements. This is normally done by establishing rules for the use and management of metadata, specifically as regards the semantics, the syntax and the optionality of values (content)[4].

A metadata schema, as a set of the three types of metadata in question and is designed to serve a specific purpose. Within the scope of this document the purpose is to increase transparency on qualifications across Europe.

IV.   How can Member states use the schema to publish information as Linked Open Data?

Linked Open Data refers to a style of publishing and interlinking structured data on the Web. This way of publishing depends on clearly defined meta-data structures and uses the web to create links between data from different sources. The basic idea behind it is that the value and usefulness of data increases the more they are interlinked with other data.

Two basic requirements for Linked Data are:

- Take advantage of the Resource Description Framework (RDF) which is the general framework for how to describe any Internet resource such as a Web site and its content[5] and use it to publish structured data on the web under an open license;
- Interlink data from different data sources by using RDF links.

Using the RDF data model, structured data about Qualifications on the web is published using the metadata schema described in section 2. The qualifications schema in this document is an evolution of the existing EQF data model (the XML Schema), which is the model on which the national qualifications databases already rely upon. The new schema will provide more elements and relations between the NQFs and cover metadata about qualifications in a more consistent and complete way.

In that way and exactly as the traditional web can be crawled by following hypertext links, the so-called 'Web of Data' can be crawled by following RDF links. Working on these "crawled data", search engines and other online services can provide sophisticated query capabilities, similar to those

---

3 NISO (2004) Understanding Metadata. Bethesda, MD: NISO Press
4 ISO 23081-1:2006 Information and documentation - Records management processes - Metadata for records - Part 1: Principles
5 For more info see: https://en.wikipedia.org/wiki/Resource_Description_Framework

provided by conventional relational databases. Because the query results themselves are structured data – and not just links to HTML pages - they can be immediately processed. This enables a new class of applications based on the Web of Data.Regarding the implementation point of view, there are many **tools and techniques** which support the publishing of qualifications as Linked Data on the Web (T. Heath and C. Bizer[6]) and a detailed process as described in the work of the W3C Government Linked Data Working Group (GLDWG)[7].

Currently, the main publishing option Member States have chosen is a Relational Database Management System (RDBMS) producing XML (or some other files formats). Member States currently using this XML schema will need to update it in order to be compliant with the qualification metadata schema. The Commission will propose such an XML Schema to the Member States and will also support them in implementing it (cf. section III, "Finalizing the schema", Table: Summary of actions). The modified XML will be able to express, contain and capture all metadata (entities, properties and their relationships) specified in the qualification schema.

Depending on architecture and usage scenarios, stakeholders will be able **to choose the most appropriate publishing method** which will be further explained in a targeted and detailed "user training manual", foreseen for early 2016.

V.   Publishing qualifications data and connecting to the LOQ portal

The Member States' authorities have various options to publish linked open data under an open license and connect it to the LOQ portal:

- Techniques like RDF "injections" in HTML pages.
- Other formats such as XML, CSV, Microsoft Excel or PDF resources.
- Atoms feeds.
- Static RDF resources.
- Through a linked data frond end like Pubby[8].
- Using a relational database publishing tool such as D2R[9].
- Wrapper implemented around an API, for example RDF Book Mashup[10.]

There is no best publishing option; each authority will decide what solution fits its current publishing system. The experts' team can help the authority in this assessment.

Once the publishing option is determined by the authority, the web scraping (or harvesting) process starts[11]. Thanks to the process in question, qualification metadata can be extracted from web sites. Different web scraping techniques exist; the most efficient is the "semantic annotation recognition". This technique relies on the fact that the pages that are being scraped contain metadata or semantic mark-ups and annotations, which can be used to locate specific data chunks.

After the scraping process the collected data will be structured (based on the schema) and re-published in the LOQ portal.

---

6 Tom Heath and Christian Bizer (2011) *Linked Data: Evolving the Web into a Global Data Space* (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool.
7 http://www.w3.org/TR/ld-bp/
12 http://www.w3.org/TR/2013/REC-prov-o-20130430/
12 http://www.w3.org/TR/2013/REC-prov-o-20130430/
12 http://www.w3.org/TR/2013/REC-prov-o-20130430/
12 http://www.w3.org/TR/2013/REC-prov-o-20130430/

VI.   Who will use this metadata schema?

The metadata schema will benefit various user groups:

- **Stakeholders** who use it to publish their information as reusable and machine-readable data, such as awarding bodies and ministries.
- **The Commission** who uses it to combine the published information, exploit it and reuse it in web portals (such as the Learning Opportunities and Qualifications portal or EURES Drop'pin), in online services (such as job matching features of EURES or CV creation in Europass) and in semantic assets for republication as part of an interlinked data set.
- **Public or private online service providers**, that have an interest in information on qualifications, such as online job portals and career guidance services, training platforms or search engines, can reuse the published information.
- **End users**, such as jobseekers, learners, teaching staff, employment advisers and recruiters will benefit from the more transparent information on qualifications.

This document describes the schema for the qualifications. It will be refined based on contributions from the Member States, different workshops and a pilot implementation. "Open issues" this version are listed at the end of this document (cf. section III).

The schema is derived from the work of the EQF Advisory Group (EQF AG), the work done on the Learning Opportunities and Qualifications in Europe portal. The former publishes, amongst others, information on qualifications and the EQF. Table 1 below shows the evolution of the data model for qualifications from 2013 to 2015.

Qualifications from National Qualification Frameworks (NQFs) are published on the portal and integration is done via XMLs validated by a qualifications XML schema. The proposed RDF schema will be influenced by the model proposed in the XML schema.

In the RDF schema, standardized vocabularies and ontologies are reused as much as possible. For example the Provenance ontology[12] is used for modelling ownership and place and time of origin; Organization[13] and Registered Organization[14] ontologies are used to model organizations such as awarding bodies (ABs) and accreditation agencies.

Section 2 describes the metadata schema classes, properties and vocabularies. Section 3 explains the open issues of the metadata schema and how they will be solved. Finally, the annex contains a list of abbreviations.

---

12 http://www.w3.org/TR/2013/REC-prov-o-20130430/
13 https://www.w3.org/TR/vocab-org/
14 https://www.w3.org/TR/vocab-regorg/

Table 1: Evolution table EQF DATA MODEL for QUALIFICATIONS

DATA			National qualifications data bases-2013	National qualifications data bases-2015
Title of the qualification			Required	Required
Thematic Area			Required	Required
Country/Region (code)			Required	Required
EQF Level			Required	Required
NQF Level			Required	Required
Description of the qualification	Either	Knowledge	Required	Required
		Skills		
		Competences		
	Or	Open text field describing what the learner is expected to know, understand and able to do		
Awarding body			Required	Required
Further information on the qualification			Optional	Optional
Source of information			Optional	Optional
Link to supplement			Optional	Optional
URL of the qualification			Optional	Optional
Information language (code)			Optional	Optional
Entry requirements			Optional	Optional
Expiry date (if relevant)			Optional	Optional
Ways to acquire qualification			Optional	Optional
Relationship to occupations			Optional	Optional
Credit points			*X*	*NEW Optional*
Volume of learning			*X*	*NEW Optional*
Accreditation and other quality assurance processes			*X*	*NEW Optional*

2. THE METADATA SCHEMA

The schema is flexible, so that it can be adapted to the situation in different Member States as well as to an international context.

For example, different actors can have different roles. An organization could be both the publisher and the awarding body of a qualification. But the roles of publisher and awarding body could be assigned to different organizations too, depending on the situation.
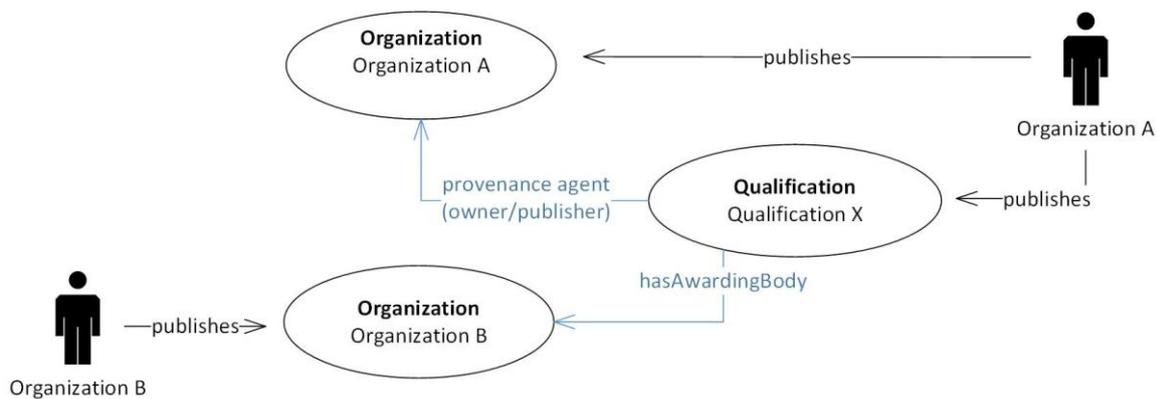

Figure 1- A conceptual view of the metadata schema

I.   Namespaces used in the schema

The Qualification Schema reuses terms from various existing specifications. Classes and properties specified in the next sections have been taken from the following namespaces:

- rdfs: http://www.w3.org/2000/01/rdf-schema#
- skos: http://www.w3.org/2004/skos/core#
- dct: http://purl.org/dc/terms/
- xsd: http://www.w3.org/2001/XMLSchema#
- iso-thes: http://purl.org/iso25964/skos-thes
- prov: http://www.w3.org/ns/prov
- foaf: http://xmlns.com/foaf/0.1/
- org: http://www.w3.org/ns/org#
- rov: http://www.w3.org/ns/regorg#

The exact class and property URIs, cardinality and range constraints of properties will be specified in a later refinement phase of the Qualification Schema (cf. Timetable in section III).

II.   The schema classes
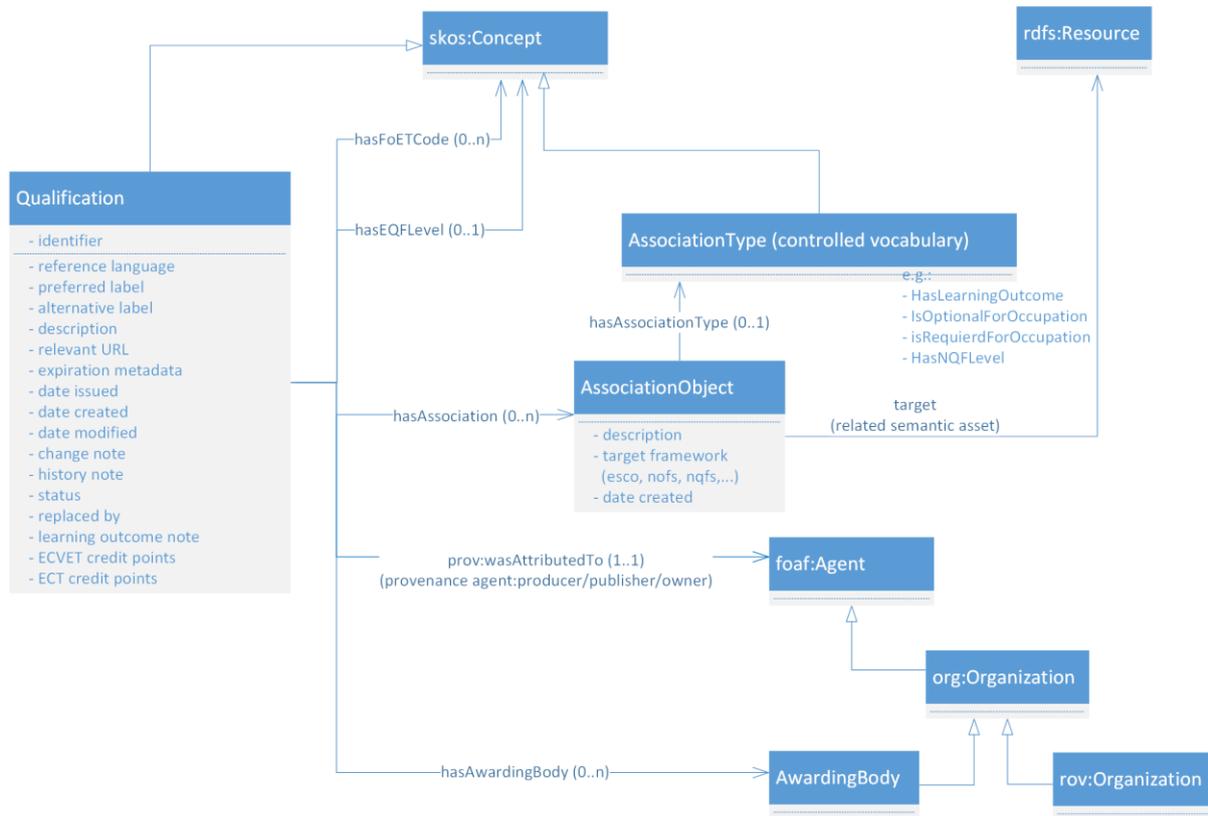
Figure 2- The schema classes

Classes

Class name	Usage note for the schema	Reference
Qualification	A qualification is a formal outcome of an assessment and validation process which is obtained when a competent body determines that an individual has achieved learning outcomes to given standards	http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32008H0506(01)
Agent	An agent (e.g. person, group, organization). Used to define the provenance agent of a qualification. In case the Agent is an organization, the use of the Organization Ontology is recommended.	http://xmlns.com/foaf/spec/#term_Agent, http://www.w3.org/TR/vocab-org/, http://www.w3.org/TR/vocab-regorg/

AwardingBody	A registered institution, organization or company that issues a qualification (certification). The use of the (Registered) Organization Ontology is recommended.	http://xmlns.com/foaf/spec/#term_Agent, http://www.w3.org/TR/vocab-org/, http://www.w3.org/TR/vocab-regorg/
AssociationObject	The generic class of extended directed associative relationships between the subject resource and a semantic asset from another framework).	
AssociationType	The class of concepts representing the type of association between a qualification and a semantic asset.	http://www.w3.org/TR/skos-reference/#concepts
Skos Concept	The generic class of classification codes in the controlled vocabularies used by the Qualification Schema.	http://www.w3.org/TR/skos-reference/#concepts

Table 1

III.   Schema properties per class

Qualification

The class of qualifications.

Characteristics of a qualification:

- The qualification is issued by an awarding body.
- The level of the qualification may be expressed according the European Qualification Framework (EQF Level).
- The subject area of the qualification will be expressed according the UNESCO Institute of Statistics standard for the Fields of Education and Training 2013 (FoET - ISCED, FoET Code)
- Qualification expiration metadata might indicate if the qualification is limited for a period in time and/or if it requires re-certification.
- Learning outcomes of a qualification describe what a learner knows, understands and is able to do on completion of a learning process, defined in terms of knowledge, skills and competence.
- Associations to other semantic assets (e.g. occupations the qualification is required for, the qualification framework level of a country,…)

Property	Usage note	Cardinality
identifier	Unique identifier	1..1
reference language	The language in which information about the qualification is available in the system. This language will be used as default/reference language for the qualification.	1..1
preferred label (title)	The official and/or preferred name of the qualification. This property can be repeated for parallel language versions of the name.	1..n Only one preferred label per language allowed. The preferred label in the reference language must be present.
alternative label	An alternative name of the qualification.	0..n
description	The full description of the qualification, including the learning outcomes of the qualification. This property can be repeated for parallel language versions of the description.	1..n Only one description per language allowed. The description in the reference language must be present.
relevant URL	Website that is relevant for the qualification (e.g. national webpage of the qualification).	0..n
expiration metadata	Records expiration data and recertification needs of a qualification (e.g. period how long a qualification is valid, required recertification,…).	0..1
date issued	Date of formal issuance (e.g., publication) of the qualification.	1..1
date modified	Date on which the qualification was last updated.	1..1
change note	Records information about the qualification version modifications (fine grained changes).	0..n
history note	Records information about significant version changes to the qualification (major lifecycle changes).	0..n
status	The publication status of a qualification (e.g. released, deprecated,…).	
replaced by	In case a qualification is deprecated this property might refer to another qualification that replaces the deprecated one.	
FoET code	Refers to a concept identifying the ISCED 2013/FoET classification code (Field of Education Code). Used to indicate the thematic area of the qualification.	1..1

Property	Usage note	Cardinality
EQF level	Refers to a concept identifying the level (as specified by the European Qualification Framework) applicable to the subject qualification.	1..1
credit points	Credit points as used in the credit system.	0..1
Provenance agent	Refers to the agent (e.g. official authority) responsible for publishing this qualification.	1..1
Awarding body	Refers to an awarding body issuing this qualification.	1..n
(cross framework) association	Refers to an association entity that relates the qualification to another semantic asset (a resource from another framework). See AssociationObject	0..n

Table 2

Agent

An agent (e.g. person, group, software or physical artefact).

The qualification metadata schema uses it to identify provenance agents. Different provenance agents can be identified:

- The provenance agent of a qualification (responsible for publishing the qualification, the owner, e.g. national authority publishing the qualifications)
- Awarding bodies

Agent characteristics used by the qualification metadata schema:

- The (full and formal) name of the agent (required) - see name
- The short name or nickname of the agent (optional) - see nickname
- The phone number(s) of the agent (optional) - see phone
- The home page of the agent (optional) - see homepage
- The e-mail of the agent (optional) - see personal mailbox

If the Agent is a (registered) organization, the use of the (registered) Organization Ontology is recommended. See:

- http://xmlns.com/foaf/spec/#term_Agent
- http://www.w3.org/TR/vocab-org/
- http://www.w3.org/TR/vocab-regorg/

Property	Usage Note	Card.
name	The (full and formal) name of the agent	1..1
nickname	The short (informal) name or nickname of the agent	0..1

Property	Usage Note	Card.
phone number	The phone number(s) of the agent	0..n
home page	The homepage of the agent	0..n
e-mail	The e-mail of the agent	0..n

Table 3

AwardingBody

The class of registered awarding bodies.

A registered awarding body is an institution, organization or company that is recognized (official or otherwise) for issuing formal certificates of skills and competences.

An awarding body is a subclass of Organization. (See Agent)

AssociationObject

The generic class of extended directed associative relationships between the subject resource (e.g. a qualification) and a semantic asset from another framework and/or classification system (e.g. between a qualification and an occupation).

The extension allows to specify metadata on the association.

The association:

- is typed by the tagging property 'has association type'
- starts from the subject resource (e.g. qualification) identified by 'is association for'.
- is detailed using a descriptive text
- ends in the resource identified by 'target resource'
- specifies the framework (semantic asset) to which the resource being described is associated

Property	Usage Note	Card.
has association type	Refers to a tagging concept identifying the type of association (the category of alignment). See Association Types	0...1
description	Descriptive text that details the association. This property can be repeated for parallel language versions of the description.	0..n
date created	Date of creation of the association.	
is association for	The subject resource. The resource being described (e.g. a qualification).	1..1
target framework	The framework (semantic asset) to which the resource being described is associated (e.g. the ESCO classification system, a national qualification framework system). See Alignment Frameworks	0..1

Property	Usage Note	Card.
target resource	Refers to a node (rdfs:resource e.g. a skos:concept) in an established framework (e.g a skill or occupation from the national classification system, the qualification level according the national qualification framework,…)	0..1
target description	The description of a node in an established framework. This property can be repeated for parallel language versions of the target description. (This property can be used if the node in the target framework is not published as rdfs:resource)	0..1
target name	The name of a node in an established framework. This property can be repeated for parallel language versions of the target name. (This property can be used if the node in the target framework is not published as a rdfs:resource)	0..1
target url	The URL of a node in an established framework. (This property can be used if the node in the target framework is not published as rdfs:resources)	0..1

Table 4

AssociationType
The class of association types used in the Qualification Schema.

The qualification relations to other semantic assets (resources from other semantic frameworks and or classification systems) are "typed relations"[15]. The different association types in the qualification metadata schema are represented and identified by concepts. The tagging concepts are part of a taxonomy managed by the qualification metadata schema. See Association Types

Property	Usage Note
preferred label	Preferred or formal name of the association type. This property can be repeated for parallel language versions of preferred label.
description	Description of the association type. This property can be repeated for parallel language versions of the description.
notation	String of characters used to uniquely identify the association type. A notation is different from a lexical label in that a notation is not normally recognizable as a word or sequence of words in any natural language.

Table 5

---

15 A relation between the content item represented by this element's owning content element, and another content item

IV. Controlled vocabularies

To improve semantic interoperability and support multilingual purposes, the qualification metadata schema links to semantic assets from other established frameworks and classification systems. Besides that the schema also defines its own code lists (small supporting taxonomies) to be used as a value list for some properties in the schema.

Controlled vocabularies to be used
The table below is an indication of the properties and controlled vocabularies that will be used. The exact properties and controlled vocabularies that must be used will be defined at a later refinement phase of the Qualification Schema (cf. *Table 9: Actions plan* in section 3).

Property	Used for Class	Vocabulary name/description	Usage note
FoET code	Qualification	ISCED 2013/FoET (International Standard Classification of Education: Fields of Education and Training 2013).	Used to identify the FoET code of a qualification.
EQF level	Qualification	The European Qualification Framework (EQF) levels published as recommendation by the EU Publication office.	Used to identify the EQFLevel of a qualification.
has association type	AssociationObject	Qualification association types. Supporting taxonomy defined by the Qualification Metadata Schema.	Used to identify the association type to another semantic asset.
target framework	AssociationObject	List of established semantic frameworks to which qualifications might be associated with.	Used to identify the framework and/or classification system the semantic asset belong too when associating a qualification to another semantic asset.

Table 6

Supporting code and value lists
To improve semantic interoperability and support multilingual purposes the qualification metadata schema will also define its own supporting code and value lists that must be used for certain properties. Both of them will be discussed during workshops with Stakeholders (EQF AG, Member States and the Commission) that will take place in Q4 of 2015 (cf. *Table 9: Actions plan* in section 3).

Alignment Frameworks

Qualifications can be related to semantic assets published by other frameworks and classification systems (e.g. national occupation classification system, national qualification system level, …)

The exact list of recommended and established frameworks a qualification might be associated with will be defined in the later refinement phase of the schema. The table below is just an indication.

Framework description	Usage note
National Occupation Classification System	Used to identify the national occupation classification system when associating a qualification to an occupation from the national classification that's being used, if relevant.
National Qualification Framework	Used to identify the national qualification framework when associating a qualification to its national qualification framework level if relevant

Table 8

3. FINALISING THE SCHEMA

I.   Open issues and future actions

At the moment, the schema as presented in section 2 is not complete. The topics that are still open and that will be elaborated in more detail include the following:

- Final metadata URIs, their range and cardinality
- Definition of provenance criteria e.g. licensing and quality assurance
- Versioning metadata
- Location metadata e.g. jurisdiction area
- Identifying metadata
- Expiration and recertification metadata
- Accreditation metadata
- Agent metadata (e.g. provenance agents, accreditation agents (AA) and awarding bodies)
- Association metadata
- Dependencies between qualifications
- Quality assurance and accreditation process analysis based on the referencing reports
- Comparative overview of the previous XML (XSD) schema with the metadata schema described in this document

The Commission is going to organize workshops in Q4 2015 between the developer of the metadata schema presented in this document and technical experts from the interested Member States in order to benefit from the inputs of the latter to finalize the technical specifications of the current schema. During Q1 and Q2 of 2016, visits from experts assigned from the Commission will take place to all interested countries. During these visits a small team of experts will assess the technical maturity level of all countries and the status of the process of publishing qualifications. This team will also provide technical consultancy and training to appointed staff.

In addition to the above an upcoming analysis of the quality assurance and accreditation processes based on the referencing reports will also take place during Q1 and Q2 of 2016. The final version of the metadata schema is foreseen for the end of the 2nd quarter 2016.

II.   The description of learning outcomes – a forthcoming European format for presenting learning outcomes

A transparent description of the learning outcomes of a qualification is essential for being able to exchange information on qualifications. While the learning outcomes approach is increasingly used to define and describe qualifications across Europe, no common solution on how to **present and share** these learning outcomes descriptions has been developed and agreed. The resulting diversity makes it difficult to compare qualifications, for example in qualifications databases. While meta-data schemes and linked open data approaches can be very helpful, the original input of information regarding learning outcomes need to be agreed on.

To improve the comparability of qualifications the EQF Advisory Group has developed a common format for the presentation of learning outcomes in the national and European databases and portals[16]. This common format, **to be used on a voluntary basis, will add to but not replace existing**

---

[16] "Work programme to promote the implementation of learning outcomes approaches across Europe – further proceeding" (Note EQF AG 29-4).

learning outcomes formats currently used at sectoral, regional, national or international level. The format will make it possible to strengthen the independent ability of the EQF and NQFs to provide comparable information on the content and profile of qualifications, directly benefitting individual citizens and employers. Given that consultation on this format is still ongoing, detailed information on this will be distributed late autumn 2015.

III.   Summary of actions

The following table summarizes the action plan defined in the previous section.

Action ID	Description	When
1	Series of workshops with MS to refine the data model	Q4 2015
2.	Refinement of the metadata schema	Q4 2015
3.	Visits to the member states - assessment of maturity level	Q1 and Q2 2016
4.	Provision of technical training and support on transform and upload of XML files	Q1 and Q2 2016
5.	Analysis of the quality assurance and accreditation processes based on the referencing reports	Q1 and Q2 2016
6.	Finalized metadata schema	Q2 2016

Table 9: Actions plan

ANNEX I: LIST OF ABBREVIATIONS

AA	Accreditation Agent
AB	Awarding Body
DC	Dublin Core, a small set of vocabulary terms to describe web and physical resources
ECTS	European Credit Transfer and Accumulation System
EQF AG	European Qualification Framework (Advisory Group)
EURES	European Employment Services
FoET	Fields of Education and Training
GLDWG	Government Linked Data Working Group
ISA	Interoperability Solutions for European Public Administrations
ISCED	International Standard Classification of Education
LOD	Linked Open Data
NQF	National Qualification Framework
RDF	Resource Description Framework, a W3C standard and general method for information modeling
SKOS	Simple Knowledge Organization System
W3C	World Wide Web Consortium: develops protocols and guidelines for the web
XML	eXtensible Markup Language