ESCO quality management

Quality dimension

Employment,
Social Affairs
and Inclusion

ESCO (2016) SEC 008 FINAL
Document Date: 01/03/2016
Last update: 14/03/2016

1. Purpose of this document

The purpose of this document is to present to the ESCO Maintenance Committee (MAI):

- The actions taken during development to deliver a consistent and relevant classification
- The vision for the evaluation and management of the quality of ESCO

The ESCO Secretariat (SEC) invites the MAI members to provide their comments on this document.

2. Introduction

ESCO aims to guarantee a high level of quality for its different applications and end-users. For the period prior to publication, the SEC uses quality criteria that are application independent, and those that assure that ESCO will meet the requirements of the EURES business case (see Table 1). Prior to publication, the SEC has processes in place that guarantee that the ESCO classification describes the full scope of occupations (OCC), and that the concepts of the classification are captured in a consistent and comprehensive manner and at a level that is relevant for the labour market. Furthermore, the SEC has organised that the mappings to ISCO-08 are validated. How this is achieved, is described in the section about delivering ESCO v1.

Table 1. Application independent and EURES business case related criteria

Data consistency	The extent to which concepts are described in a consistent manner
Use of language	The extent that the language used in the terms is correct
NOC coverage verified	The extent that all NOC occupations can be mapped on to ESCO concepts
ISCO-08 alignment	The extent that all ESCO occupations correctly map onto a single ISCO-08 unit group

Once the classification is published, quality management for ESCO will be redefined in context of the applications and end-users using it. In the section about post-publication quality management, we describe how this can be achieved for the various types of functionality supported by ESCO. The SEC describes further how this evaluation can be put in context of actual applications, their users, the data sets and the language used.

Table 2. Types of functionality supported by ESCO

1.	Concept extraction
2.	Data entry
3.	Transcoding other classifications
4.	Simple search
5.	Multi-lingual search
6.	Semantic search
7.	Profile comparison
8.	Data collection
9.	Data aggregation

3. Pre-publication quality management for ESCO

The SEC prioritises the EURES business case for the development of ESCO, while at the same time, it manages the general quality criteria for classifications, i.e. on consistency and comprehensiveness.

The EURES business case relies on the functionality domains of data-entry, transcoding and job matching.

This implies that it is key for the classification to have the right level of conceptual coverage, the right level of conceptual granularity, to have the right preferred terms in all the 24 languages, and to have the right knowledge structures. Additionally, to have the appropriate level of terminological richness, and supporting taxonomies is needed to support the usability of the EURES interfaces for this usage scenario.

The SEC achieves this validating the coverage of the classification by comparing it to a number of national classifications, and by having sectoral experts and terminologists to create the terminology sets in the supported languages.

Additionally, the SEC ensures to have with ESCO a full coverage of the ISCO-08 classification. The SEC achieves this by having an ISCO expert involved in the validation of the ISCO-08 mapping, and by evaluating and resolving the gaps with ISCO-08.

The EC is further involved in pilot projects with PESs that have explored the creation of correspondence tables for ESCO, and with organisations that have expressed the wish to develop pilot implementations with ESCO. Based on the discussions within these pilots, the EC is receiving feedback on the fit-for-purpose on some of the quality dimensions of ESCO. To the extent possible, the feedback is taken into account and is used to improve the quality of ESCO prior to publication.

Table 3. Quality management during development of ESCO v1	
Data consistency	Having ESCO guidelines Monitoring concepts using performance indicators Business rule validation Quality assurance by SEC Process validation by the MAI
Use of language	Involving linguists Involving terminologists
Conceptual coverage	Using quality reference classifications Verification with job vacancies Involving sectoral experts ISCO-08 gap analysis National Occupation Classification gap analysis
Conceptual granularity	Using quality reference classifications Verification with job vacancies Involving sectoral experts ISCO-08 gap analysis National Occupation Classification gap analysis
Knowledge structures	Research sector specific sources and classification Verification with job vacancies and curriculum vitae Involving sectoral experts
Supporting taxonomies	Studying other classifications Involving sectoral experts
Right preferred terms	Studying other classifications Involving sectoral experts Involving terminologists Involving linguists
Terminological richness	Involving sectoral experts Involving terminologists Involving linguists

4. ESCO's post-publication quality management

The SEC develops the ESCO quality management framework on the basis of the types of functionality supported by ESCO. For each type of functionality the SEC describes which aspects of the classification (see Table 4) are relevant and to what extent the evaluation of the quality is dependent on the specific application context (See also Table 5). Additionally, the SEC describes how ESCO quality can be evaluated and assured.

Table 4. Description of ESCO's classification aspects	
Conceptual coverage	The extent to which ESCO's concepts cover the scope of a domain
Conceptual granularity	The level specificity on which ESCO describes the concepts
Knowledge structures	The relations between the ESCO concepts
Supporting taxonomies	Additional structures to support specific applications or use-cases
Right preferred terms	The term defined to represent the concept (language specific)
Terminological richness	All terms defined for the concepts (language specific)

Table 5. Relevancy of ESCO's classification aspects for each type of functionality

	Conceptual coverage	Conceptual granularity	Knowledge structures	Supporting Taxonomies	Right preferred terms	Terminological richness
Concept extraction	**Key**	**Key**	**Supporting**	**Supporting**		**Key**
Data entry	**Key**	**Key**	**Supporting**	**Supporting**	**Key**	**Supporting**
Transcoding	**Key**	**Key**			**Supporting**	**Supporting**
Simple search	**Key**	**Key**				**Key**
Multi-lingual search	**Key**	**Key**				**Key**
Semantic search*	**Key**	**Key**	**Key**	**Supporting**		**Key**
Profile comparison	**Key**	**Key**	**Supporting**			
Data collection	**Key**	**Key**				**Key**
Data aggregation	**Key**					

* Job-matching is an application of semantic search

4.1. ESCO to support concept extraction

Description

The EC publishes with ESCO concept definitions for occupations (OCC), knowledge, skills and competences (KSC). Each of these concepts will be published with sets of terms in 24 languages[1]. Applications can use these sets of terms to extract the relevant concepts from e.g. plain-text curriculum vitae, job vacancies or other relevant documents. Concept extraction applications depend on having the ESCO terms and concepts aligned with the terms and concepts used in the corpus of documents that the application has to parse. For ESCO this means that quality is defined here as having all concepts needed (concept coverage), with all the appropriate terms used (terminological richness), for all relevant documents and in each of the languages.

Relevant for e.g. Curriculum vitae parsers, search engines, etc.

N.B. ESCO's quality is to be evaluated in context of language and a relevant dataset.

[1] Bulgarian, Spanish, Czech, Danish, German, Estonian, Greek, English, French, Irish, Croatian, Italian, Latvian, Lithuanian, Hungarian, Maltese, Dutch, Polish, Portuguese, Romanian, Slovak, Slovenian, Finnish, Swedish.

Example

Companies that are involved in the automatic interpretation of curriculum vitae (curriculum vitae parsing) will use different parse techniques to extract the relevant data-elements (e.g. name, address, birthdate, previous employers, qualifications, skills, previous occupations etc.) from the plain-text CV. For recognising the ESCO concepts in the documents it can use the terms defined in ESCO. The quality of ESCO in context of supporting this process depends on (1) ESCO having all the relevant terms used in the CV, and (2) ESCO's concepts aligning with the needs of the application.

Types of functionalities involved

- Conceptual coverage (key)
- Conceptual granularity (key)
- Terminological richness (key)
- Knowledge structures (supporting)
- Supporting taxonomies (supporting)

Quality management approach

The SEC organises that ESCO will be published with sets of terms for each concept in all supported languages. To ensure quality for the terms, the SEC involves linguistic and labour market experts in the term definition process. To support further alignment of the conceptual coverage, conceptual granularity and the terminological richness of ESCO for CV parsing, the SEC will cooperate with the actual users of the classification on the basis of concrete applications. This way, the SEC will learn where ESCO can be better aligned with a specific application. On the basis of this assessment, the SEC can define further improvement actions.

4.2. ESCO to support data entry applications

Some organisations that will use ESCO have applications in which job seekers, employees, employers or other interested parties maintain information on e.g. themselves, or job vacancies. The usability aspects of these applications are dependent on the alignment of ESCO's vocabulary with the vocabularies used by the end-user. In terms of ESCO this translates into:

1. Having the right preferred terms in all the languages defined for the concepts, i.e. terms that make sense for the end-users. (The preferred terms are used in user interfaces to represent the concept.)
2. Having the terminological richness in all the languages to cover all relevant terms used by the end-users (to support type-ahead functionality in case an end-user specifies a search-term).
3. Having the concept coverage required by the data entry application.
4. Having supporting hierarchical structures that are aligned with how the end-user thinks about concepts in the user interface (to support navigation functionality).
5. Having meaningful associative relationships in the classification, e.g. between occupations and skills (to support suggestion functionality).

Relevant for e.g. EURES, EUROPASS, employment services, job boards, etc.

N.B. ESCO's quality has to be evaluated in light of the language and vocabulary used by the end-user.

Example

EURES supports people active on the labour market to create and maintain a curriculum vitae online. For this purpose a data-entry application is used. ESCO can support this kind of data entry application:

- To increase the value of the data captured by providing a rich controlled set of well-understood knowledge, skill and competence concepts,
- To increase the usability by providing knowledge structures that support the provision of user-relevant choices for selection (e.g. suggesting related skills, etc.),
- To increase the usability by having the capability to recognise the use of a concept in case synonyms are provided.

Types of functionalities involved

- Conceptual coverage (key)
- Conceptual granularity (key)
- Right preferred terms (key)
- Knowledge structures (supporting)
- Supporting taxonomies (supporting)
- Terminological richness (supporting)

Quality management approach

The SEC envisions working with organisations deploying ESCO based data-entry applications to learn where ESCO can be improved in terms of its different aspects. The organisations can learn about the usability aspects of their user interface by studying how it is used and the quality of the data captured. The study will reveal if there are requirements for improvement of the classification, or if other supporting data structures could be used to improve the user's experience. These findings can be used to plan improvement actions for the classification.

4.3. ESCO to support transcoding other classifications

In many situations information on OCC and KSC are captured using other classifications than ESCO. If the information needs to be passed on to EURES, the information needs to be transcoded into ESCO. To support these cases, the provider of the information needs to create correspondence tables that describe how the information is represented in ESCO.

Transcoding data always comes with a certain level of information loss. Since ESCO needs to support many classifications, transcoding quality for ESCO means that it needs to minimise the comprehensive information loss for all relevant classifications.

Relevant for e.g. EURES, employment services to exchange data, statistical bureaus, and researchers

N.B. Quality has to be evaluated in context alignment of the meaning and granularity of the concepts used in the classifications.

Example

The public employment services (PES) of Europe capture information on occupations using their national occupational classification. To support EURES they are asked to create correspondence tables to ESCO so that they can transcode and deliver the information to EURES using the ESCO concepts.

Types of functionality involved

- Conceptual coverage (key)
- Conceptual granularity (key)
- Preferred term (supporting, i.e. only used at the phase of defining the correspondence table)
- Terminological richness (supporting, i.e. only at the phase of defining the correspondence table)

Quality management approach

The structure of the correspondence tables defined for ESCO captures information about the quality of the alignment of the concepts in the two classifications. Because of this feature of the correspondence tables, the SEC can evaluate the alignment of the various classifications used with ESCO. The SEC will use this information to investigate where and how ESCO coverage and granularity can be revisited to reduce information loss when transcoding classifications.

4.4. ESCO to support simple search

The ESCO term sets can be used in search engine configurations to support two key search-engine features:

1. Type-ahead functionality during query specification
2. Synonym recognition

These functionalities will improve the usability aspects for the end-user if:

- the vocabulary of the end-users is covered by ESCO's terminological richness
- the granularity and coverage of the concepts in ESCO match what is needed by the end-user

Relevant for: search engines and job boards providing searches for any of the concepts provided by ESCO

N.B. Quality has to be evaluated in light of the language and vocabularies used by the end-users, required level of specificity of the concepts needed, and the language and vocabulary used in the indexed documents.

Example

Generic search-engines make use of synonym dictionaries to bridge the vocabulary used in the expression of a query and the actual words used in documents. Additionally search-engines frequently come with support for type-ahead functionality to support the user to formulate relevant queries. ESCO's terms sets can be converted into search engine dictionaries that support these functionalities.

Types of functionality involved

- Conceptual coverage (key)
- Conceptual granularity (key)
- Terminological richness (key)

Quality management approach

If optimisation of ESCO's support for simple search is required, the SEC can support comparison of actual search-engine results with expected results for relevant queries. This type of investigations will reveal what type of actions can be taken to further optimise the ESCO classification.

4.5. ESCO to support multi-lingual search

The terms of the ESCO classification can be used to support multi-lingual search for the domains covered by ESCO. The fact that for each concept relevant terms are captured in 24 languages empowers the user to use these as a multi-lingual synonym dictionary for a search engine.

Relevant for e.g. Web search engines, multi-lingual systems

N.B. Quality has to be evaluated in light of the languages and vocabularies used by the end-users, required level of specificity of the concepts needed, and the language and vocabulary used in the indexed sources.

Example

Multi-lingual search engines use multi-lingual synonym databases in a similar fashion as generic search-engines do for dictionaries for a single language. Key difference is their ability to execute cross-language searches.

Functional aspects involved

- Conceptual coverage (key)
- Conceptual granularity (key)
- Terminological richness (key)

Quality management approach

The typical quality management approach for the scenario of multi-lingual search is very similar to the one of simple search, however in this case the set of terms to be investigated has to cover all languages covered by the repository and the queries.

4.6. ESCO to support semantic search

The SEC delivers ESCO with a rich set of terms for each concept, and with relations that put the concept in context of the other concepts. These features make that ESCO can be used for semantic search. Semantic search engines can use these features to come to sophisticated evaluations of the relevance for the results. The quality of ESCO to support semantic search is therefore not only dependent on the terminological coverage and the conceptual coverage of the classification, but also on the richness and appropriateness of the relations between the concepts of the ESCO classifications.

Relevant for e.g. Web semantic search engines, job matching algorithms

Example

A job-matching engine typically uses the complete profile of a job seeker to search for the best matching profiles in a collection of vacancies (or vice versa). It uses all the concepts captured in the profile to identify potential matches, and rank the results according to their relevancy parameters captured with the profiles. In this process the job-matching engine uses not only concepts expressed in the profile, but can use as well the related concepts, like the related concepts provided by ESCO.

Functional aspects involved

- Conceptual coverage (key)
- Conceptual granularity (key)
- Knowledge structures (key)
- Terminological richness (key)
- Supporting taxonomies (supporting)

Quality management approach

The SEC can assess the quality of ESCO for semantic search by working with the organisations that use ESCO in this context to establish a feedback mechanism that provides detailed information about the performance of the classification in their environment. This information will clarify which parts of ESCO perform well in for semantic search, and which areas need attention.

4.7. ESCO to support profile comparison

The EC delivers the ESCO classification as a means to support comparing different sets of labour market data. Using this feature one can e.g. compare the profile of a job seeker with the profile of an ideal candidate for a job. Such a comparison will reveal how well on which capabilities and the short comings of the job seeker for the job.

Relevant for e.g. career guidance providers

Example

Career guidance applications compare the set of KSC, and Q required for a desired or vacant position with the concepts captured for a student, employee or job seeker. Based on the KSC and Q that the student, employee or job seeker does not have, career guidance can be formulated.

Key functional aspects involved

- Conceptual coverage
- Conceptual granularity
- Knowledge structures (supporting)

Quality management approach

The SEC can assess the quality of ESCO for career guidance by working with the career guidance providers to establish a feedback mechanism that provides information about KSC usage, and where KSC are missing. This information will clarify which parts of ESCO are used for career guidance, and where possible gaps occur.

4.8. ESCO to support data collection

One of the usage scenarios for ESCO is to use it to support data collection. Researchers can use the concept extraction functionality to identify occurrences of the concept and count these occurrences in a data collection (e.g. online newspaper, EURES log file, twitter stream). The (change of) frequency of occurrence of an OCC, KSC or Q in a repository or stream provides additional information on the dynamics within education and the labour market.

Relevant for e.g. statisticians and researchers

Example

Applications like Google Trends can measure the frequency of use for certain terms or concepts in documents on the web for a given period. Using these type of tools, statisticians can use ESCO concepts to learn about changes in the frequency of mentions of the individual KSC. Researchers can use this mechanism to e.g. monitor which programming languages are trending and which ones are abandoned.

Key functional aspects involved

- Conceptual coverage (key)
- Conceptual granularity (key)
- Terminological richness (key)

Quality management approach

The SEC can assess the quality of ESCO for data collection by establishing working relationships with the statistical bureaus and domain experts to ensure that the classification covers the concepts used by statisticians and have the information on new emerging concepts.

4.9. ESCO to support data aggregation

One of the business cases for ESCO is defined around statistical reporting. In this case ESCO is used to group collected data points and aggregate the numbers on a level relevant for the report. In this context ESCO must be comprehensive for the domain of the report (see also concept extraction). Additionally it will need to have correspondence tables to the classification used for the data analysis (see also ESCO to support transcoding). The required correspondence tables are not part of the ESCO deliverables.

Relevant for e.g. statisticians, researchers

N.B. Quality has to be evaluated in light of the types of statistics being produced, and the appropriateness of the concepts defined in ESCO to support the analysis.

Example

Many reports on the labour market are provided using ISCO-08. ESCO data can be aggregated using ISCO-08 since ESCO is delivered natively with an ISCO-08 correspondence table.

Key functional aspects involved

- Conceptual coverage (key)

Quality management approach

The quality of ESCO for data aggregation can be managed by assuring that the classification covers comprehensively the domain of data. The SEC can evaluate the comprehensiveness of the dataset with domain experts and with other classifications to ensure that this criterion is met.