



EUROPEAN COMMISSION  
DIRECTORATE-GENERAL FOR EMPLOYMENT, SOCIAL AFFAIRS AND INCLUSION

Labour Mobility and International Affairs  
Labour Mobility, Public Employment Services, ELA

Brussels  
EMPL.E.1

## **MEMBER STATES WORKING GROUP ON ESCO**

### **Plenary meeting**

**15 March 2022, 9:30, Brussels**

### **Results of the third phase of the pilot project for linking learning outcomes of qualifications with ESCO skills**

#### **Introduction**

The present document aims at reporting to members of the ESCO MSWG and the EQG AG on the results of the third phase of the pilot project for testing the use of an automated approach for linking learning outcomes of qualifications to ESCO skills.

The document recalls the policy objectives behind the use of an automated approach for linking qualifications to ESCO skills and proposes a way forward for a potential third phase of the project.

Members are invited to provide their input on the content of the note, to comment on the results of the pilot, and to provide their views on the future of the IT tool supporting the automated linking and on its concrete applications.

#### **1. BACKGROUND**

For ESCO to reach its full potential as a bridge between education and training and the labour market, the ESCO skills terminology could be used to systematically identify and analyse which skills are related to a particular qualification.

This can help make qualifications more transparent across Europe. It can also help the translation, comparison and/or review of qualifications. Furthermore, the identification and analysis of the content and profile of a qualification can be used to indicate matches and/or mismatches with skills needs of occupations and sectors. Linking learning outcomes of qualifications to ESCO skills means that employers can more easily grasp the labour market value of a qualification, in particular in a cross border context. Individuals may see their chances on the labour market improved through better matching based on richer qualifications information.

In **June 2021**, the Commission presented to the ESCO Member States Working Group (MSWG) and to the EQF Advisory Group the results of second phase of the pilot project for linking learning outcomes of qualifications to ESCO skills.

The project allowed assessing the required effort for creating links between learning outcomes of qualifications and ESCO skills. To support the linking process, the Commission developed an IT tool (hereinafter referred to as the Linking Service Platform) using artificial intelligence, in particular **Natural Language Processing (NLP)**, and based on requirements expressed by the participating countries. As a result, the exercise confirmed the methodology of an automated approach with an initial degree of human intervention, as suggested by the 2019 ESCO Qualifications Pillar study<sup>1</sup>.

In **October 2021**, the Commission invited the ESCO MSWG and the EQF AG to express their interest in participating in the third phase of the project. The call was also open to other public and private organisations testing the usage of ESCO.

The third phase of the pilot had the following objectives:

1. Gather **verified data** on the links between learning outcomes of qualifications and ESCO skills. This data will be used to further improve the underlying algorithms.
2. Improve the technology supporting the automated linking of learning outcomes of qualifications with ESCO skills, through the implementation of a suggestion engine that consists of **three matching algorithms based on artificial intelligence**.
3. Test the automated linking and the performance of the matching algorithms in **different EU languages**. This comparative data will lead to a choice of the best performing algorithm available at this time, which will then be used in the standalone service.
4. Test how the new **ESCO transversal skills structure** supports the matching with learning outcomes describing transversal skills (e.g. problem solving, critical thinking, life competences).
5. Test how **ESCO skill groups** can be used to introduce a certain level of abstraction for qualifications from higher EQF (European Qualifications Framework) levels.
6. Test the application of **ESCO version 1.1** in the domain of education and training.
7. Make **suggestions on the ESCO skills pillar** that will feed into the ESCO continuous improvement process.
8. Support the publication of information on skills relating to national qualifications in **Europass**.

To this end, participants were invited to test a **minimum of 30 qualifications** covering different EQF levels and including at least some vocational and some academic qualifications, and to **share the linking data** with the Commission for further improving the matching algorithm.

---

<sup>1</sup> <https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8181&furtherPubs=yes>

The study was presented during the joint meeting of the EQF AG and the ESCO MSWG on 6 February 2019.

Overall, **20 Member States and organisations** participated in the second phase of the pilot as full members or observers:

- 7 EU MS: Germany, Ireland, Italy, Poland, Slovenia, Sweden, and The Netherlands.
- 7 Training providers: Riga Stradins University, Hellenic American Union, University of Deusto, Academy of Fine Arts of Vienna, University of Lusofona, Polytechnic of Milan, TU Dresden.
- 4 private companies: Employchain, Hurrah, Nxus, Monster
- 1 third country organisation: the Commonwealth Scientific and Industrial Research Organisation (CSIRO)
- 1 EU agency: the European Training Foundation

Representatives from Czech Republic, SME United and ETUC attended the project meetings as observers.

## **2. PROJECT IMPLEMENTATION**

The third phase of the pilot lasted one year, from January 2022 to February 2023.

A first workshop was organised on 31 January 2022 to present the objectives and main features of the pilot. During the workshop, participating countries described their interest in the project and presented the qualifications that would be linked with ESCO skills.

A technical webinar to present the improved version of the IT tool supporting the automated linking and the technical work behind the training of the algorithms tool took place on 2 June 2022.

The pilot involved **1038 qualifications**, out of which **443 qualifications** established mappings to ESCO skills. The completion rate varies among the participants, with some organisations that could not map all the qualifications within the deadline of end January 2023. In addition, the Commission and the ETF are testing the performance of the AI algorithms to map ESCO skills to **Ukrainian professional qualifications**.

The results of the linking exercise were discussed in a final workshop on **9 February 2023**. The workshop's agenda included a presentation from the Commission of the quantitative results of the pilot and a roundtable discussion on common issues identified by participating countries, on the suitability of the ESCO skills thesaurus for describing learning content, and on the functioning of the IT tool and on the further improvements to be prioritised for future development. To this end, prior to the meeting the Commission circulated a list of questions and points of interest on the perceived quality of the ESCO vocabulary and on the usage of the IT tool.

## **3. ADDITIONAL FEATURES OF THE IT TOOL SUPPORTING AUTOMATED LINKING OF LEARNING OUTCOMES OF QUALIFICATIONS TO ESCO SKILLS**

The revised version of the IT tool used for this third phase included the following additional functionalities:

- The possibility to establish links to ESCO skills and knowledge groups.
- Dashboards displaying the results of the linking activity with the following indicators:
  - number of skills linked to a selected qualification,
  - proportion of exact/narrower/broader/close relations,
  - number of occupations associated to a selected qualification (based on common skills),
  - total number of mapped qualifications/skills,
  - total number and type of matching relation,
  - frequency of skills in the mapping activity.
- An improved mechanism for text splitting.
- Filtering of suggestions according to the skill type (skills or knowledge concepts) and/or key words inserted by users.

A key element of novelty of this third phase of the pilot was the implementation of three different matching algorithms in the backend of the Linking Service Platform.

During the second phase of the pilot, participants worked with a basic search algorithm as available in the ESCO API, and with a first standard NLP algorithm based on word2vec. Thanks to the volume of data collected, it was possible to build a test data set to develop new and improved versions of the algorithm and objectively measure progress.

Two additional models were built based on the Sentence-BERT algorithm and the semantic embedding mapping algorithm. The Sentence-BERT model is a multilingual model that can do semantic comparisons directly in the native language. The semantic embedding mapping model is including a BERT model that only accepts English input text, requiring a translation step.

Finally, a multilingual sentence splitter model was implemented to produce higher quality initial splits. In phase 2 of the pilot the algorithm was splitting only for newlines, resulting in inaccurate results and leading to significant manual corrections. The new sentence splitter that was implemented for phase 3 is a machine learning model that learned to split multilingual qualifications based on a set of already split qualifications.

#### **4. MAIN FINDINGS AND RESULTS FROM THE SECOND PHASE OF THE PROJECT**

The linking exercise was performed on **443 qualifications** out of the 1038 provided, in **12 EU languages**: DE, EL, ES, EN, FR, IT, LV, NL, PL, PT, SE, SI.

The average percentage of mapped text (text that was linked with at least one skill or knowledge concept) varies based on the completion rate: participants that completed (or almost completed) the linking exercise showed a percentage of mapped text higher than 90%.

In total **8,499 learning outcomes** were mapped to at least one ESCO concept, generating **29,894 links**. These figures almost triple the result from the second phase. The average of mapped learning outcomes per qualification stands at **11.5**. When looking at the number

of skills linked to a learning outcome entity, the **average for this third phase was of 3.3 ESCO skills per learning outcome entity**: this is coherent with the lessons learned during the second phase and the recommendation of matching between 1 and 3 ESCO skills or knowledge concepts per learning outcome entity.

In terms of mapping relations<sup>2</sup>, data showed a prevalence of the relation type “close match”, suggesting that the selected ESCO concept does not fully cover the scope of a learning outcome. The average percentage of **exact matches was 26%**, however this number varies across projects (ranging from around 43% to 3%). As for the second phase, these statistics should be read taking into account three aspects:

- “close match” is the default relation type in the linking service tool (which implies that a link will be automatically defined as a close match if no action is taken by the reviewer);
- It is not possible to flag a no match but only to leave a learning outcome entity unmapped. Participants highlighted this point as a potential improvement of the tool during the final workshop.
- the definition and interpretation of relation types could differ between participants.

The performance of the mapping algorithm showed promising results. On average, an ESCO concept selected by participants was in **44.65%** of cases **the first concept** suggested by the algorithm (an increase of almost 15 percentage points compared to the second phase), between the **first 3 concepts in the 63.98% of cases, between the top 5 concepts in the 70.35%**. This indicated that lower efforts were required in order to identify relevant skills compared to the second phase.

Based on these analytics and on the comments provided by the participants during the final workshop, the following conclusions can be drawn:

- Linking learning outcomes to skills is done best in the **native language** as this allows the algorithm to learn more about the specifics of a language, thus producing better matches.
- The **structure of learning outcomes** plays a very important role for the correct functioning of the algorithm: on average, 10-15 words are found in a learning outcome, which contributes to returning a high number of concepts, sometimes not (or partially) related to the meaning of the text.
- **Splitting sentences** in a correct way is very important for the functioning of the tool. In this regard, the variety of approaches used in the description of learning outcomes (numbering, punctuation, flat text) represents a challenge.
- It is important to further specify the definitions of the different types of matching relations given the **relevance of exact matches** for training the AI algorithm.

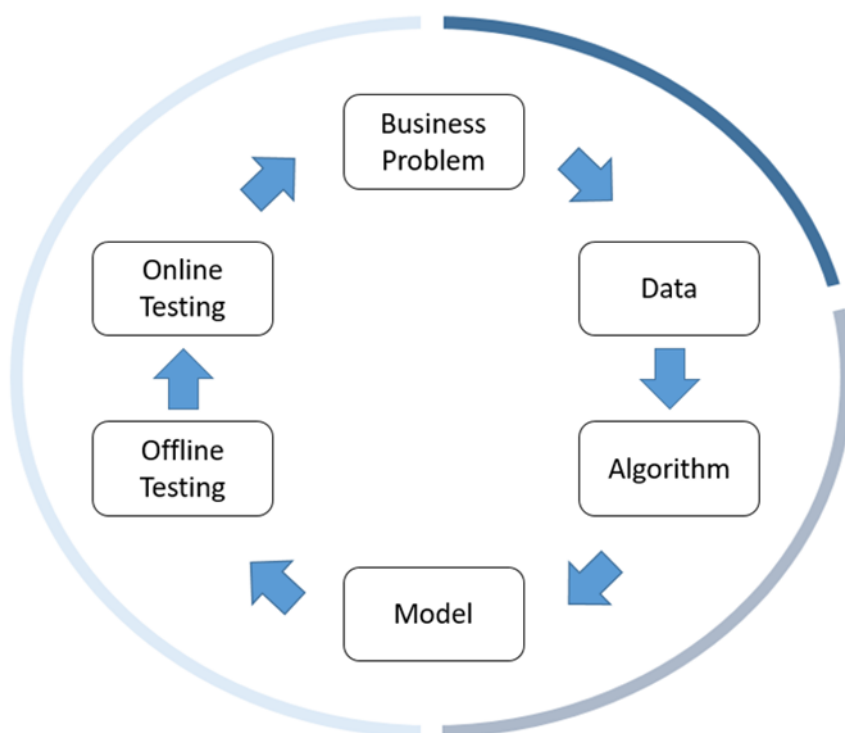
---

<sup>2</sup> The tool allowed four types of mapping relations: exact match, narrower than, broader than, close match.

- Linking is **time-consuming** due to the need of validating the results. Scaling up this exercise would require dedicated resources with **professional knowledge of the interested qualifications**.
- Structuring the ML suggestions based on **professional fields** would improve the selection of relevant skills.
- In most cases **several ESCO concepts** can be matched to one learning outcome entity, suggesting that learning outcomes could be decomposed to avoid merging several competences into one.
- Introducing **proficiency levels** for language skills is recommended, given that the learning outcomes related to language competences refer to a specific proficiency.

## 5. PERSPECTIVES OF AI USAGE AND FURTHER IMPROVEMENTS OF THE MATCHING ALGORITHM

To continuously improve natural language processing algorithms, large amounts of data are needed. Especially **verified or validated data**, in this case the links between learning outcomes and skills, is crucial. This validated data allows the Commission to define a test data set to compare different algorithms and quantify their relative performance in an objective way.



As mentioned above, this third phase of the pilot aimed to compare the performance of three different algorithms and to test a mechanism built on a multilingual model and departing from the translation of the original text into English.

While challenges related to text fragmentation, comprehension of contextual information, misinterpretation of key words and performance across languages persist, the results of this third phase can be seen as encouraging.

**More data** are needed to improve the accuracy of the model and reduce manual browsing of the ESCO skills thesaurus, which was still very much used during this last phase. However, the volume of labelled data needed to finalise the development of the algorithm cannot be met only by organising a fourth phase of the pilot, given the lack of resources available at national level.

Making this technology publicly available would allow the collection of large volumes of labelled data while minimising the resources needed on both the Commission and the organisations interested in mapping learning content with skills terms.

## 6. NEXT STEPS

Based on the results described above, the Commission believes that the ESCO skills taxonomy is a valuable tool to identify which skills are linked to a particular qualification and support the overall goal of promoting transparency of information on qualifications.

At the same time, the Commission believes that further investing in using AI to link learning outcomes of qualifications to ESCO skills is a precondition for developing an AI infrastructure that can promote skills-based job matching, reduce skills mismatches, support re- and up-skilling based on the individual's skillset and bridge the terminological distance between the world of education and training and the labour market.

The Commission is therefore working to develop an open version of the linking application available to all ESCO stakeholders, without requiring user authentication. The current linking tool would remain available to institutional stakeholders that want to link high volumes of qualifications.

The public version of the linking tool will allow users to copy the text of the qualifications to be mapped and received suggestions of ESCO skills in return. The system will be based on a multilingual model to offset the need to translate the content from the original language into English.

At the same time, the Commission is exploring synergies between the Linking Service Tool and the infrastructure supporting the publication of qualifications and learning opportunities in Europass (QDR). Linking learning outcomes with ESCO skills in fact would further improve the current system recommending courses to Europass users, and support career guidance and upskilling –reskilling. On the other end, it requires the availability of datasets of qualifications and learning opportunities mapped to ESCO skills: while the European Learning Model provides for this possibility, the effort required by the publishing countries is still to be determined.

The Commission invites therefore the feedback of the ESCO MSWG and the EQF AG on these two topics.