|  |  |
| --- | --- |
| Logo of the European Commission, 12 yellow stars on a blue background arranged in a circle and framed by two light grey graphic elements representing the Berlaymont building, which is the headquarter of the European Commission. | EUROPEAN COMMISSION  DIRECTORATE-GENERAL FOR EMPLOYMENT, SOCIAL AFFAIRS AND INCLUSION  Labour Mobility and International Affairs  **Labour Mobility, Public Employment Services, ELA** |

Brussels, 28 September 2022

EMPL.E.1

Member States Working Group on ESCO

Focus meeting

Brussels (online), 14 September 2022

Minutes

# Approval of the agenda

The agenda was adopted with no further changes.

# Nature of the meeting

The first focus meeting of the Member States Working Group on ESCO (MSWG) was attended by:

* representatives of 19 Member States’ (MS) authorities on labour market and education and training (AU, BE, BG, DE, DK, EL, ES, FR, HR, HU, IT, LT, NL, PL, PT, RO, SK, SE, SL).
* representatives of one observer country (IS),
* representatives of European social partners (ETUC, SME UNITED),
* Commission and CEDEFOP services.

The following experts were invited by the Commission (COM) to attend the meeting as observers: Cecile Ramombordes (Randstad), David Hunter (Star Class), Gabrielle Fournet (Boost), Marcel Bakker (WCC).

COM chaired and opened the meeting.

The main goal of the meeting was to discuss specific topics related to the maintenance and evolution of ESCO. In particular, the meeting focused on the existing approaches to cluster and classify skills and occupations terms. The Commission invited the input of experts with specific experience on clustering methodologies.

The meeting was hold online and recorded for internal purposes.

# List of points discussed

**3.1 Introduction from the Commission**

COM reiterated the nature of the focus meetings, which seek to discuss concrete topics relevant for the maintenance and evolution of ESCO at technical level. In this regard, participation is reserved to members with specific expertise on the topic on the agenda and willing to actively contribute to the debate. COM will invite external experts to share best practices and provide their expertise based on needs and according to the topic of the meeting.

COM introduced the topic of the first focus meeting: skills and occupations clustering and the use of artificial intelligence to support the transition of ESCO from an ontology into a more connected knowledge graph. The meeting aimed to exchange views on the value of adding additional clusters to the existing ESCO hierarchies, which proved to be effective and therefore will not be abandoned.

**3.2 Introduction to the topic of alternative hierarchies**

COM presented the state of play of skills clustering in ESCO and the different options for grouping skill terms: a top-down approach consisting of assigning concepts to pre-established categories, and a bottom-up approach powered by data science.

COM described the techniques that can be applied following the bottom-up approach:

* co-occurrence driven skill clustering, where terms are connected based on their occurrence in job vacancies and CVs,
* semantic relatedness-driven skill clustering, where terms are grouped based on semantic relatedness and vectorisation,
* and a hybrid methodology combining semantic clustering and co-occurrence of terms in online vacancies, where semantic relatedness is applied to a dataset of co-occurring terms.

Representatives from Pole Emploi and VDAB shared best practices on the topic of skills clustering. In particular, Pole Emploi presented the process followed to cluster competences in the latest version of the ROME classification (ROME 4.0). VDAB presented the approach used to develop a clustering of skills in the Competent classification and how artificial intelligence was used to support human experts in grouping concepts.

The following points were raised during the subsequent discussion:

* Data validation is crucial when working with results provided by artificial intelligence, as validated data serve to improve the performance of the algorithms.
* Human expertise is a fundamental component of any methodology for skills clustering and as such should never be fully replaced by the use of technology.
* It is important to clearly identify the objectives of introducing additional structures in ESCO.
* Online job vacancies might not capture all parts of the labour market. An exclusively data-driven approach might therefore generate gaps in the classification, with the risk of grouping together dissimilar skills.
* Addressing duplicate skills is a crucial challenge for the maintenance of ESCO.
* A hierarchy of skills should help identifying transferable skills, i.e. skills that can be transferred from one activity, occupation, or job, to another.
* A data driven clustering implies higher maintenance costs, as the grouping should be constantly updated based on real data from the labour market.
* Consistency across languages represents a challenge when grouping skills according to the bottom-up approach.
* Mapping of national skills classifications to ESCO should be exploited to inform the maintenance and evolution of both national classifications and ESCO.

COM explained that 120 duplicate skills were made obsolete in ESCO version 1.1.0. The remaining duplicates will be systematically addressed in the next major versions of ESCO.

COM explained that almost 75% of ESCO skills can be matched with text from job vacancies published in EURES, offering a good base to work on a co-occurrence model.

COM confirmed that a hybrid approach combining co-occurrence and semantic similarity will give better results, and that human validation is central to the clustering exercise.

COM proposed to pilot the combined approach with a sample of ESCO skills.

**3.3 Mapping of ESCO occupations to NACE**

COM presented the clustering of ESCO occupations by economic activities and the mapping of ESCO occupations to the revised version of the NACE classification. COM described the methodology to map ESCO to NACE, which includes 5 steps:

1. collect online vacancies published in the EURES portal, which are tagged with ESCO skills and NACE classes
2. run statistical analysis to clean and structure the data
3. build a machine-learning classifier which will predict sectors of economic activities for ESCO occupations
4. select the best performing model
5. validate the results end quality assure the mappings.

An ESCO – NACE mapping will be published on the ESCO portal as a separate dataset, following the request of several ESCO implementers.

The following points were raised during the subsequent discussion:

* grouping occupations by economic activities will lead to a poly-hierarchical model in ESCO. Using NACE for such purpose might be difficult since ESCO occupations are not modelled on industrial activities.
* Mapping to NACE is challenging due to the transversal nature of many ESCO occupations, which do not belong to a single economic sector.
* It could be envisaged to indicate occupations that are more relevant in an industry, without directly classifying such transversal occupations in a specific class.
* While EURES data are already tagged with NACE classes, they might be not representative of the whole labour market.

# Next steps

COM closed the meeting and thanked the participants for their active participation. The next focus meeting is tentatively scheduled for 5 November 2022.