



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

Towards ESCO version 1.2: key elements and timeline

This document provides a short overview of the methodology used for ESCO v1.2, looking in particular at how data science and artificial intelligence supports the actions undertaken to create new content (occupations, skills and knowledge concepts) in line with labour market trends and improve the quality of existing concepts in the classification.

The continuous improvement of ESCO requires a constant update process where taxonomy enrichment decisions are driven by evidence extracted from data. The ESCO Secretariat has been establishing a process that works towards an approach where:

- 1) Information is extracted from different data sources such as stakeholder input, Member State taxonomies, online job advertisements, workers' profiles and educational data;
- 2) Extracted information is integrated such that evidence from different sources can be compared, allowing to support or complement each other;
- 3) Metadata (e.g. geographical coverage, labour market coverage, sectoral coverage, temporal coverage) about the extracted information is generated to support decision-making;
- 4) Human expertise is pivotal thereby making informed decisions based on the available integrated information and corresponding metadata.

Artificial intelligence techniques, data science and natural language processing represent crucial ingredients to support this process as these techniques allow to extract key information from large amounts of raw information. By applying them to the data sources as described below, it becomes possible to integrate information and perform a meta-analysis. In addition, through a human-in-the-loop approach it is possible to continuously finetune the process over time with the aim to make future updates more efficient while measuring and securing quality.

OVERVIEW OF THE METHODOLOGY FOR v1.2

The visual below summarizes the main steps undertaken for the knowledge management of ESCO v1.2, differentiating between the content update of ESCO (new concepts) and the quality review of the existing concepts:

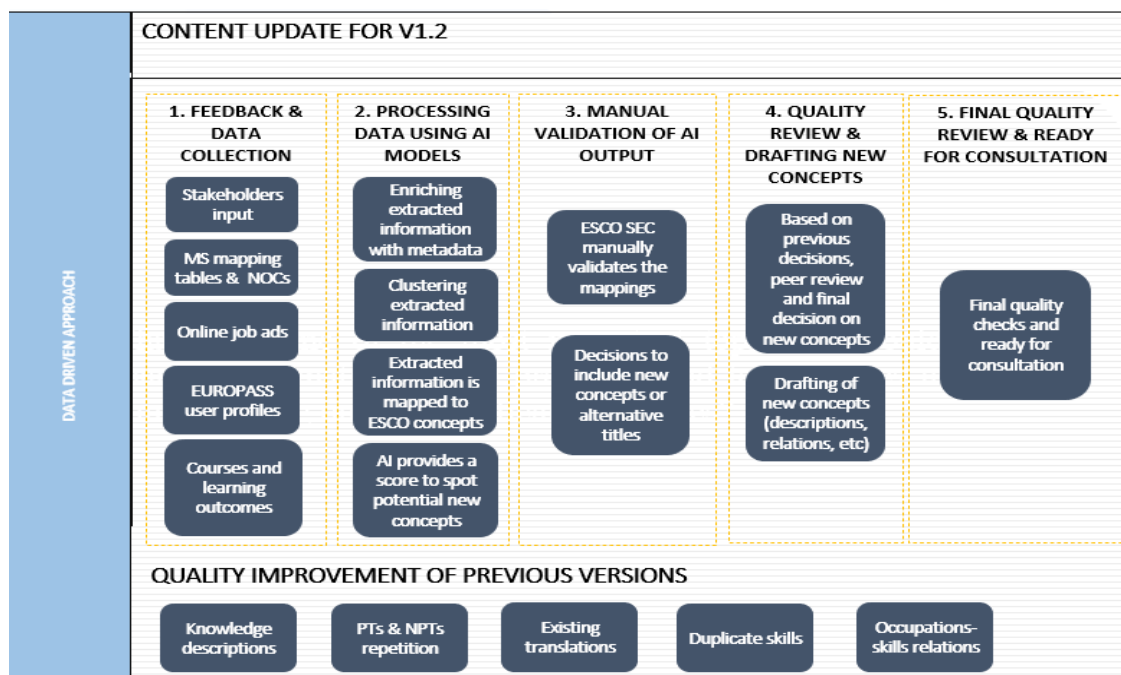


Figure 1. Overview of ESCO v1.2 knowledge management process.

The methodology for the **content update of ESCO v1.2** is structured as follows:

Step 1: Data collection from a variety of sources such as stakeholders' input (blueprint projects, sectoral associations, social partners, NGOs, private and public stakeholders as well as other Commission services), Member State taxonomies and mapping tables, EURES online job advertisements, Europass user profile data, courses and learning outcomes (e.g. Qualification Dataset Register).

Step 2: Processing of different data sources by enriching extracted information with metadata, clustering of extracted information, mapping content to existing ESCO concepts **through AI models** and identifying potentially new concepts via semantic similarity scores.

Step 3: Manual validation of the processed information thereby deciding whether the input is a potential new concept or related/identical to an existing concept. The validation involves comparing the inputs from the external sources to ESCO concepts to determine whether they are covered or not in ESCO. When an input is relevant (i.e., present in the European labour market) and does not exist in ESCO, the validators suggest adding it either as a new concept or as an alternative label of an existing concept.

Step 4: Quality review following the initial results as obtained from manual validation and **drafting new concepts**, including assigning the allocation in the hierarchy assisted by AI model suggestions. Drafting the new concepts is the outcome of desk research while AI

models assist in linking ESCO skills to new occupation concepts and existing ESCO occupations to new skills and knowledge concepts.

Step 5: Final quality review of drafted concepts and prepare for final publication.

In parallel, the Commission also worked on **improving the quality of the previous version** to tackle inconsistencies and issues that needed to be resolved. In short, the following activities have been carried out to quality assure the existing ESCO data:

- **Quality review of the ESCO knowledge pillar.** Description sentences that start with or include a verb tag were flagged and further analysed through human expert validation, after which appropriate actions were undertaken to correct the identified errors. These include a re-drafting of the concept description or formulating the knowledge as a skill, including changing to the concept type, its preferred and alternative term and allocation to a different subsection of the ESCO skills hierarchy.
- **Removing repetitions between PTs and NPTs.** Alternative labels (NPTs) which had the same title as their Preferred terms (PT) were identified and deleted to avoid duplications and to bring clarity to the concept.
- **Improvements in translations of existing concepts.** Correct and/or improve the translations of certain concepts, identified by stakeholders, to ensure their accuracy.
- **Addressing duplicates skills.** As a result of this process, **159 duplicate skills** are made obsolete in ESCO v1.2 and the corresponding information from the obsolete skill concepts was merged with the retained skill concepts.
- **Improve the skills-occupations relations.** In total, 111 occupations with five or less essential skills and 25 occupations with 70 or more essential and/or optional skills linked to them were identified and corrected.

DATA SOURCES USED FOR V1.2

For ESCO v1.2, the Commission leveraged the above methodology in combination with the following data sources:

Type of data source	Short description/examples
External feedback from sectoral experts	Contributions from external experts for the creation of new occupations/skills in different sectors, indications to improve and update existing concepts in ESCO, proposals to establish new relationships between existing skills-occupations.
National classifications for occupations and skills (NOC)	Input data from national classifications for creating new skills/knowledge or occupation concepts in ESCO. Data is representative of the European labour market.
ESCO-NOC Mapping tables	The mapping tables of the national classifications reflect the work done by the MS to map their own classifications to ESCO. Through this exercise the MS

	identify concepts that have, exact relation, narrower relation, broader relation, or no relation to ESCO. Through this exercise, the MS significantly contribute to the maintenance and update of ESCO by pointing out the concepts that are not covered in the classification.
Europass user profiles	Europass CVs offer valuable data from the labour market supply side. The skills and competences that are reported by users in their profiles provide information about how these are phrased by job seekers and workers.
Scientific research papers	Data on skills from experienced researchers of the labour market and the different economic sectors. The inputs inform the update of ESCO through providing suggestions of concepts potentially not yet covered by the taxonomy.
Online job advertisements (OJA)	Information present in EURES OJAs is representative for the labour market demand side. The data are linked to ESCO to determine whether they are already covered by the taxonomy.
Courses data	Course descriptions (e.g. from QDR) from educational institutions contain valuable information about skills and knowledge that are required in the labour market. Information from those descriptions is compared with existing ESCO concepts to identify new content.

Table 1. ESCO v1.2 data sources.

Among those sources, the external feedback is worth detailing as the input varies from detailed comments of occupations and skills to broader labour market analysis.

During the feedback process, these external inputs are analysed and discussed in detail, to select what relevant content could be used for ESCO v1.2. Domain experts are actively involved in this process, including the support to drafting of occupations and skills concepts metadata. This ensures that the implementation of the input received will meet labour market trends.

Furthermore, the process of collaboration and the feedback obtained from external experts is of great importance given the quality of the data that is obtained from these sources. They have a wide expertise in sectors that have become essential for the European labour market. Therefore, the inclusion of their inputs, once the validation process is completed, allows ESCO to fill existing gaps in line with the reality of labour market data.

The following table offers an overview of the 13 experts that shared data and contributed to ESCO v1.2, including their contribution type of and the sectors covered:

Experts who contributed with ESCO v1.2.	Type of Input	Sectors covered
ALBATTIS Blueprint Project	New occupations/skill proposals and quality review of current classification.	Battery
CHAISE Blueprint Project	New occupations/skill proposals and quality review of current classification.	Blockchain
DHBW University	Quality review of current classification.	ICT
ESSA Blueprint Project	New occupations/skill proposals and quality review of current classification.	Software
ETF EU Agency	New occupations/skill proposals and quality review of current classification.	Green skills
EQVEGAN Blueprint Project	New occupations/skill proposals and quality review of current classification.	Food Industry
ENISA EU Agency	New occupations/skill proposals and quality review of current classification.	Cybersecurity
EPBS EU Association	Quality review of current classification.	Biomedical Science
FIELDS Blueprint Project	New occupations/skill proposals and quality review of current classification.	Agriculture & Bioeconomy
METIS Blueprint Project	New occupation-skills relations.	Microelectronics
SAM Blueprint Project	New occupations/skill proposals and quality review of current classification.	Additive Manufacturing
SKILLSEA Blueprint Project	New occupations/skill proposals and quality review of current classification.	Maritime Transport
SPIRE-SAIS Blueprint Project	New occupation-skills relations.	Energy intensive industry

Table 2. Data contribution from external experts and sectors covered.

Moreover, the ESCO Secretariat has been closely collaborating with a total of **19 Blueprint Alliances** from different ecosystems. Apart from those included above, 10 other Blueprint Projects participated in this sectoral feedback process but in a different stage, therefore their input will be taken into account for the upcoming versions. The table below clusters the Blueprints Alliances and their sectors that will be under review for upcoming versions of ESCO:

Blueprint Projects	Sector
ASSETS	Defense
ARISA	AI
B-WISE	Social Economy
BASE	Social Economy
CHARTER	Cultural Heritage
CYANOTYPES	Creative & Cultural Industries
EDDIE	Energy Value Chain
I-RESTART	Agrifood
PANTOUR	Tourism
REWIRE	Cybersecurity

Table 3. Collaboration with Blueprint projects for upcoming ESCO versions.

TIMELINE AND CONSULTATION PROCESS

The figure below shows an high-level overview of the planning for ESCO version 1.2, highlighting the time allocated to receive input from stakeholders, analyse the input and draft the content, conduct the MS consultation, translate the content into the ESCO languages and finally release the new version.

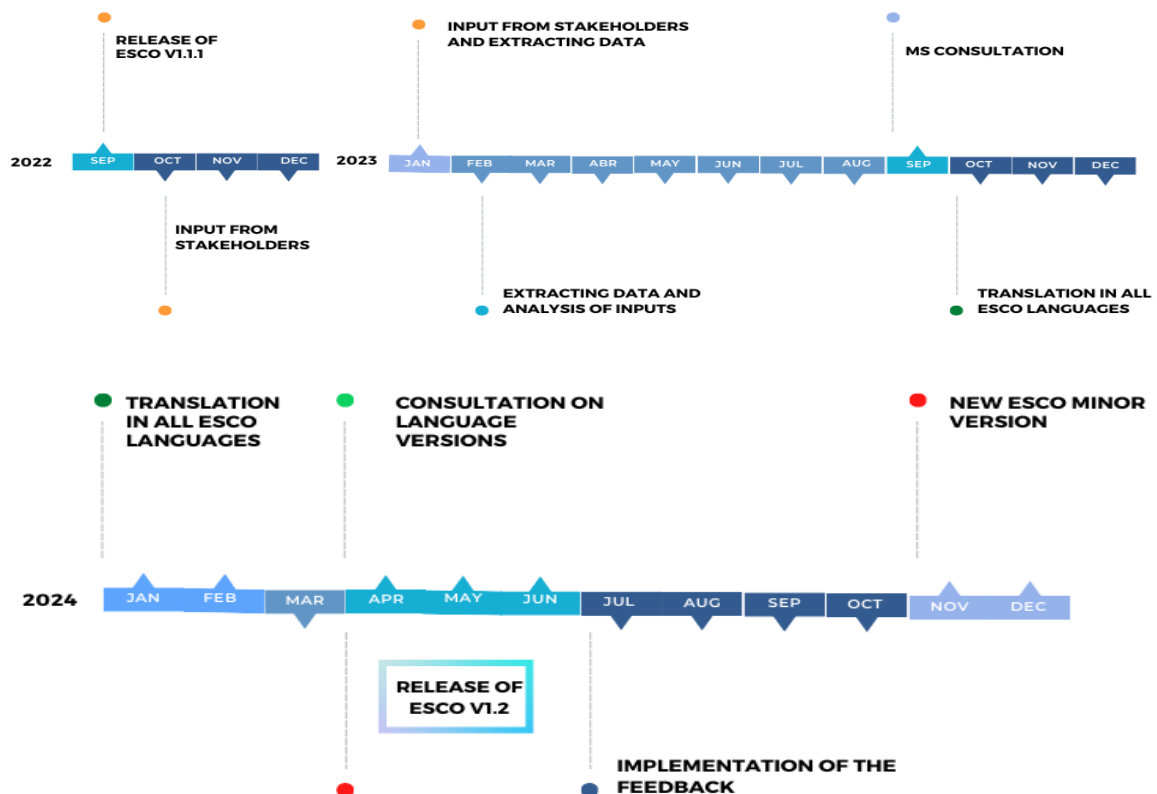


Figure 2. Summarized overview of planning for ESCO version 1.2

Overall, the analysis of input, creation and validation of new content will be completed by late **July 2023**.

Member States will then be consulted on the improvements to existing ESCO concepts and on the new content in the **first half of September 2023**. The consultation will last 1 month after which the input received will be analysed and processed by the ESCO Secretariat.

Since the last MSWG meeting in December 2022, the Commission decided to change the approach for consulting the group. For this version and ideally for the upcoming versions to follow, the Member States Working Group will be consulted using a consultation module of an existing taxonomy management tool used internally by the ESCO Secretariat to manage the classification. This facilitates the storage and processing of feedback received by MSWG in a single platform, ensuring the consultation is done a user-friendly way.

A mock-up of the consultation platform will be presented in the MSWG meeting to provide already a look and feel of the platform and an explanation on how Member States should provide their feedback. Detailed instructions on how to provide feedback will also be provided in writing at the start of the consultation. Users of the platform will be able to visualise the whole classification (including older versions) but they will only be able to share comments and feedback on the new content or the improved data that falls under the scope of this new version.

Translations will then start by **beginning of the last quarter of 2023** so that the new version could be published in **March 2024**.

The consultation of the ESCO Member States Working Group on the different language versions will run in parallel to the publication process. Member States will have three months to provide feedback on translations of new and modified concepts, which will then be implemented in a subsequent minor version in the **end of 2024**.