

**ESCO version 1.1: state of play**

MSWG 15-02 | MAI 36-02



*29 November 2021*



Contenido

[Background 3](#_Toc86213692)

[State of play 3](#_Toc86213693)

[Use of AI in the development of ESCO v1.1 4](#_Toc86213694)

[Use of AI to process suggestions: analysing feedback 4](#_Toc86213695)

[ESCO versioning 6](#_Toc86213696)

[Data versioning in ESCO 6](#_Toc86213697)

[ESCO minor versus major data versions 7](#_Toc86213698)

[How users can access the different versions of ESCO 8](#_Toc86213699)

[Update on a new structure for transversal skills in ESCO 8](#_Toc86213700)

[Defining a taxonomy of skills for the green transition 11](#_Toc86213701)

[Next steps 13](#_Toc86213702)

[Communication 13](#_Toc86213703)

# Background

For ESCO to remain useful in the labour market and in education and training, it needs to be updated regularly in line with trends of new and changing occupations and skills. Additionally, the quality of ESCO continuously needs to be improved, based on stakeholder needs.

As presented during the 14th meeting of the Member States Working Group on ESCO (MSWG),[[1]](#footnote-2) ESCO v1.1 will bring a substantial number of changes, with the introduction of **515 new concepts** out of which 109 concepts are related to digital technologies.

It will address some of the quality issues flagged during the quality reviews of the ESCO occupations and skills pillars. **35 ESCO v.1.0 concepts have been significantly updated** in terms of their scope and relations between occupations and skills, and the Commission implemented **minor changes to 203 preferred terms of skills and occupations and 331 descriptions of skills and occupations**. **154 ESCO concepts have been fully reformulated** (changing both their label and description). **62 skills will be removed** from the classification as they duplicate other concepts.

It will include a **new model for transversal skills and competences** within the ESCO skills pillar and offer a revised terminology for **transversal skills**. It will support the implementation of the European Skills Agenda by identifying the **skills related to the green transition** and **skills and occupations for researchers**[[2]](#footnote-3).

This note presents the state of play of the release of ESCO version 1.1 and provides detailed information on two important work strands, the introduction of a new terminology for transversal skills and the identification of a taxonomy of skills for the green transition. It provides an overview on how artificial intelligence has been used for the maintenance of ESCO. It also presents the actions to communicate about the launch of ESCO v1.1.

# State of play

The process for updating ESCO started in 2019 with the launch of the community forum for domain experts and continued during 2020. The Commission consulted the ESCO Maintenance Committee on a draft pre-release in January 2021 and the ESCO MSWG during the following months. The consultation on the new content was concluded on 19 March 2021, and translations began in May 2021.



To facilitate the communication between the Commission’s translation services and the Member States, the Commission invited each member of the MSWG to appoint a Single Point of Contact for translations. 21 members (AT, BE, BG, CY, CZ, EI, ES, ET, FI, FR, HR, IS, IT, LT, LV, MT, NO, RO, SI, SK and SV) provided contact information for their SPOCs.

The translations of the new and modified content from English into the remaining ESCO languages were finalised in October 2021. The translation of the new structure and thesaurus of transversal skills will be completed by 30 November 2021. The major upgrade of ESCO to version 1.1 is scheduled for release in December 2021.

# Use of AI in the development of ESCO v1.1

Artificial Intelligence (AI) applications have been developed and employed in various steps of the development of ESCO v1.1. The goal of such applications is to support the Commission in analysing external input, drafting new content, labelling and translating the content.

### Use of AI to process suggestions: analysing feedback

The update of the classification is based on feedback provided by sectoral experts, such as members of the community forum for domain experts, organisations involved in Blueprints for Sectoral Cooperation on Skills, Sectoral Skills Alliances and other European-funded projects, and European sectoral associations. The feedback provided needs significant effort in further processing. AI applications have accelerated and improved the analysis of the suggestions, by helping to remove redundant information and to distinguish the new content from what is already in ESCO.

Machine learning models have been developed to compute semantic similarity between concepts, with the idea that if two terms are likely to have similar meaning or semantic content, this should result in a higher similarity score compared to *another* couple of terms. These models have then been employed to compare concepts suggested by different stakeholders with other new concepts as well as existing ESCO concepts.

As an example, the project [Next Tourism Generation Alliance](https://nexttourismgeneration.eu/) (NTGA), provides research and training modules for green, digital and social skills in the tourism and hospitality sector. About 127 skills and competences included in the NTGA Skills Matrix have been analysed and fed as input for the machine learning models, highlighting similarities between the suggestions and concepts already included in ESCO.

**Use of AI for terminological development**

Job vacancies are a key source of information to research emerging occupations and skills, to validate inputs from experts and to ensure wide coverage and labour market representation. For this reason, the Commission chose to use AI to process new content using job vacancies collected via the EURES portal or provided by ESCO implementers. The goal is to detect potential additional NPTs that should be included, as they are part of the terminology used in job vacancies.

This activity consists of comparing Preferred Terms (PTs), Non-Preferred Terms (NPTs) and descriptions of new occupations with job titles and descriptions used in vacancies. The suggestions are then ranked by an algorithm based on two values: the score of semantic similarity obtained via a machine learning model (method explained in the previous sub-section), and the number of times the term was found in the set of vacancies. Results are then analysed by the ESCO Secretariat before their inclusion in the list of proposed alternative labels: by applying this methodology, around 50 additional NPTs have been identified for the English version of ESCO v1.1.

**Use of AI to label new content**

Different AI methods have been tested to build a classifier which allows to distinguish green skills from other skills. This is explained in more detail in a dedicated section later in this document.

**Use of AI to translate Non-Preferred Terms**

The translation of ESCO in 27 languages is crucial for its adoption by Member States and implementers. In view of the release of v1.1, the Commission opted to attempt using AI to support this process, by automating the translation of Non-Preferred Terms (NPTs) of new concepts.

The Commission tested an innovative approach to translate 453 NPTs for occupations and 619 NPTs for skills and knowledge concepts from English into all ESCO languages. To conduct this activity, the following steps were undertaken:

1. Use machine translation tools to generate translations.
2. Conduct major improvements, such as:
	1. Ensure equal gender representation. Automated translation for gendered words usually results in translations limited only to one gender. To comply with ESCO's gender standards, missing gender terms have been added in the vocabulary. ESCO translation memories have been used to ensure consistency of translations between different versions/concepts.
	2. Select terms which should not be translated, such as specific names (e.g., "[tensorflow](https://en.wikipedia.org/wiki/TensorFlow)").
3. Ensure compliance with the ESCO terminological guidelines and conduct minor improvements, such as:
	1. Remove translations of knowledge concepts using verbs in the infinitive-tense.
	2. Apply term-specific rules, such as removing the translation of words which seemed confusing for machine translation tools (e.g., "tanning" can be translated as "converting animal skin into leather" or "expose one's skin to the sun").
	3. Apply language-specific rules (e.g., capitalise first letter of the occupation concepts in German).
	4. Ensure consistency between translated PTs and NPTs for specific cases. For example, some languages translate the term "cloud" (in the context of cloud systems), while other languages adopt the English word. It is important to verify that both PTs and NPTs show the term in the same language, being it translated or in the English form.
4. Use [language models](https://en.wikipedia.org/wiki/Language_model) to test the correctness of translated NPTs. Using machine learning models, this process consists in quantifying the probability of finding a certain sequence of words in one language. This results in a list of NPTs for each language which are flagged as problematic and need further discussion with the Single Points of Contact for translations appointed by Member States.

The steps above imply manual validation to detect issues and evaluate the results of the algorithmic approach.

# ESCO versioning

ESCO will only remain fit-for-purpose and able to deliver a high-quality service to end-users if it is continuously updated. The labour market and the education and training sector are constantly evolving, and these changes need to be reflected in new versions of the ESCO classification. For this purpose, the Commission has set up a process[[3]](#footnote-4) to continuously improve and keep the classification up to date.

### Data versioning in ESCO

Within the working environment of ESCO, versions pertain to resources such as software or data. A **new data version** is created when there is a change in the structure or content of the resource. New versions of an ESCO dataset are created when an existing dataset is reprocessed, corrected or appended with additional data.

The implementation of an efficient versioning system will bring benefits to different actors:

* ESCO implementers should be able to identify the exact dataset in order to support their business operations. This means they need to be able to indicate exactly which version of the ESCO dataset underpins their results. This becomes particularly challenging where the data to be cited are dynamic. For this reason, a versioning indicator and history are important. With these elements in place, implementers could plan development independently, i.e. they can use any ESCO version. In case of a new version, they can decide if and when to implement it.
* An efficient versioning system in ESCO will support EURES countries in updating the mapping tables provided by art. 19 (3) of the EURES Regulation. This will be necessary when one of the two classification systems changes, i.e. the European classification or their national classification (either of occupations or skills).
* Different services of the Commission as well as EU agencies and other international organisations make use of ESCO for a variety of applications. For instance, ESCO was primarily developed to be used by EURES. Cedefop uses ESCO in the Skills-OVATE project[[4]](#footnote-5) for the analysis of online job vacancies. ESCO is also used in Europass and by the Commission Learning and Management System (EU Learn). Ensuring that these services have access to the latest but also to previous versions of ESCO provides flexibility and independence in their operations.

### ESCO minor versus major data versions

Even for minor changes, it is important to keep track of the different dataset versions to make the dataset trustworthy. At the same time, the Commission will take a consistent, informative approach to versioning so that data consumers (incl. EURES countries) are able to understand and work with the changing data, e.g. provision of delta files.

Within the framework of the maintenance of ESCO, the Commission makes a distinction between minor and major versions:

* **Minor releases:** these contain changes that do not affect the concept level, i.e. no concepts are added or removed and the scope of the existing concepts is not changed. Minor releases refer to typos, corrections of translations, changes to the relations between concepts, etc. and do not require any update of mapping tables.
* **Major releases:** these contain changes at the concept level or to the underlying ESCO data model. Changes at the concept level can include added or removed concepts, or a change in the scope of the existing concepts.

The full list of minor and major versions published up to now is present below, and it is published on the ESCO portal[[5]](#footnote-6):

* ESCO v0
* ESCO v0.1
* ESCO v0.2
* ESCO v0.8
* ESCO v.1.0
* ESCO v.1.0.1
* ESCO v.1.0.2
* ESCO v.1.0.3
* ESCO v.1.0.4
* ESCO v.1.0.5
* ESCO v.1.0.6
* ESCO v.1.0.7
* ESCO v.1.0.8
* ESCO v.1.0.9

Version numbers starting with 0 (zero) refer to early ESCO versions that were used for piloting and testing only. The first fully-fledged ESCO version is ESCO v1.0, which was published in July 2017. As of this version, a versioning mechanism keeps track of changes in ESCO throughout its lifecycle. Three-digit versions refers to minor releases while two-digit versions refer to major releases.

### How users can access the different versions of ESCO

The current version of the ESCO portal displays the latest version of ESCO (currently ESCO v1.0.9). Portal users however can access all previous ESCO versions via:

1. the download section of the portal and
2. the API.

The ESCO portal displays by default the latest version of ESCO and will offer the possibility to switch to another version of the classification and display the ancient DS in the home page. Moreover, information about changes in each version will be shown with each concept, and users will be able to access all previous ESCO versions via the downloadable section of the portal and the API.

# Update on a new structure for transversal skills in ESCO

The final report of the expert group on transversal skills and competences was presented during the 14th meeting of the ESCO MSWG.

The report provides the final model of transversal skills, composed by six top-level categories, 24 clusters at the second level and a draft list of skills allocated under each cluster. The new transversal skills model can be displayed as per the image below.



Figure 1 ESCO transversal skills model

The new framework for transversal skills should be systematically integrated in ESCO in order to ensure the overall quality and relevance of the ESCO skills pillar. The transversal skills structure of ESCO version 1.1 in fact will act as a common reference and resource to systematically inform the Commission as well as ESCO implementers.

This will help to:

* Avoid overlaps and duplications.
* Avoid inconsistent naming of terms.
* Identify missing elements in the existing terminology.
* Make visible the role of transversal skills across sectors, occupations and tasks.

The Commission discussed the topic of the integration of the new transversal skills structure into the ESCO skills hierarchy during the 13th meeting of the ESCO MSWG and the 34th meeting of the ESCO Maintenance Committee. A specific question on this topic was included in the [survey on ESCO version 1.](https://ec.europa.eu/eusurvey/runner/MSWGmeeting19November2020)1 addressed to the ESCO MSWG in November 2020. The EQF AG was also consulted on this matter in December 2020.

Based on these discussions and on the contributions submitted in the survey, the option of operating with a separate transversal skills structure was seen as feasible and necessary. To this end, a re-organisation of the current skills hierarchy group is needed. Currently in ESCO v.1.0.9 there are four parallel sub-hierarchies within the overall skills hierarchy:

* Skills
* Knowledge
* Languages
* Attitudes and values

The terms included in the last sub-hierarchy are extensively covered in the new transversal skills hierarchy and therefore the classification of Attitudes and Values will be removed from the ESCO skills hierarchy. In the case of language competences, while languages will also be included in the new structure in the cluster “mastering languages,” a separate entry point in the skills pillar offers an easy access to the language skills for ESCO implementers: the sub-classification of languages will be therefore maintained, alongside the new transversal skills structure.

Additional work has been carried out in the second part of 2021 to complete the draft transversal skills structure and in particular:

* Complete the identification of skills for all clusters.
* Compose descriptions and (where necessary) scope notes for skills concepts.
* Allocate alternative labels and hidden terms to skills concepts.

In particular, the formulation of preferred labels has been evaluated against the ESCO terminological guidelines and adjusted to ensure consistency with the terminology used in ESCO. This resulted in the following changes compared to the expert group proposal:

* "Create and edit digital content" was changed into "create digital content"
* "Critically evaluate information and its sources" was changed into "think critically"
* "Act independently" was changed into "work independently"
* "Focus for prolonged periods" was changed into "maintain concentration for long periods"
* "Behave confidently" was changed into "show confidence"
* "Report" was changed into "report facts"
* "Negotiate" was changed into "negotiate compromises"
* "Build and maintain networks" was changed into "build networks"
* " Move and lift objects" was changed into "move objects"
* "React quickly to physical hazards and changes" was changed into "react to physical hazards and changes"
* "Demonstrate awareness of risks to health" was changed into "demonstrate awareness of health risks"
* "Express ideas, experiences and emotions creatively" was changed into "express yourself creatively"

Moreover one additional concept, ''demonstrate curiosity'', was added to the list of terms proposed by the experts in order to better define the scope of the related concept "demonstrate willingness to learn". The final list of transversal skills categories, cluster and concepts is included in Annex 1.

Possible duplicates between the current and the new transversal skills list required a thorough quality check before the dataset is released in ESCO v.1.1. As a result of this analysis:

* 26 existing transversal skills will be part of the new structure.
* 30 existing transversal skills are covered by the new terminology and will be deleted.
* 30 existing transversal skills will be transformed into cross-sectoral skills and linked to occupations.
* 21 existing transversal skills represent the direct implementation of the European Digital Competence Framework (DigComp 2.0) in ESCO.

Moreover, several cross-sectoral, sector- and occupation-specific skills also include an element of transversality and generate unnecessary duplications which hinders the goal of providing a consistent reference terminology on transversal skills. The Commission built on the work carried out during the quality review of the ESCO skills pillar and systematically reviewed the ESCO skills and competences flagged as transversal. 62 concepts were identified as duplicates of the new transversal skills list and therefore deleted.

The integration of the transversal skills thesaurus into the ESCO skills pillar also implies the review of the existing relations between broader and narrower skills. This exercise is needed to ensure that each concept in ESCO is part of a specific sub-section of the ESCO skills pillar: while one concept might be linked to more than one group/category, no concept should be left out of this structure. Therefore, when a transversal skill concept is deleted, all narrower concepts shall be either allocated to a specific group of the skills hierarchy or associated to a new transversal skill. ESCO v1.1 provides a first re-organisation of narrower and contextualised skills within the modified structure of the ESCO skills pillar. A complete review of all broader/narrower relations will be carried out in 2022, following the definition of a comprehensive approach on skills contextualisation.

# Defining a taxonomy of skills for the green transition

Green skills are becoming increasingly important in the debate on the evolution of the labour market. The deployment of green technologies, the transition of our economy towards a resource-efficient society and the ambition of the European Union to become a climate neutral continent all call for investment in competences that will allow to increase the number of professionals who build and master green technologies, develop green products, services and business models, create innovative nature-based solutions and help minimise the environmental footprint of activities[[6]](#footnote-7).

The European Skills Agenda foresees a number of actions related to the skills for the twin transitions, including the development of a taxonomy of skills for the green transition which will allow the statistical monitoring of the greening of our professions. Building on the rich vocabulary of skills and knowledge concepts present in ESCO v.1.0.8 and in the upcoming version 1.1, ESCO will support the identification and classification of green competences.

The definition of green skills builds on the existing literature and on the work done by CEDEFOP and the OECD on this topic. Green skills are defined as "the knowledge, abilities, values and attitudes needed to live in, develop and support a sustainable and resource-efficient society” (Cedefop, 2012). Within the perimeter of this definition, green skills can be further classified in three categories:

* Technical skills: Professional skills related to the mastering of green technologies, depending on the sector/occupation,
* Cross-sectoral skills: professional skills shared across sectors and related to the efficient use of resources or the sustainability dimension of business processes (e.g. innovation skills),
* Transversal skills: skills allowing all citizens to think or act green (e.g. environmental awareness).

To develop the taxonomy of green skills within the ESCO dataset, the Commission opted to combine the outcomes of manual human work and a machine learning driven approach. First, a set of ESCO skills is manually labelled as green (*Human labelling*) and by Artificial Intelligence models (*Machine learning labelling*) in an independent manner. Discrepancies are identified by comparing the two sets, resulting in re-evaluation and applying changes and improvements to the first list.



The human labelling process consists of labelling ESCO skills as green based on the above definition. A draft list of green concepts in ESCO includes 421 skills (325 skills already included in ESCO v1.0.8 and 95 skills that will be added in ESCO v1.1.) and 185 knowledge concepts (out of which 140 are already included in the ESCO dataset).

The machine labelling process consists of training machine learning models to distinguish green skills from non-green skills. For this purpose, an independent training set has been manually developed, including two main types of information: skills that have a positive impact on the environment (such as " Design water conservation systems") and skills that have a negative impact on the environment (such as "Coal Mining"). The training set is developed in such a way that it reflects as close as possible the ESCO terminology for skills and knowledge concepts. Sources employed for building the green and non-green sets are summarised in the table below.

|  |  |
| --- | --- |
| **Source category** | **Examples of sources** |
| Taxonomies from Member States and other national taxonomies | * Onemev (France)
* Empleo Verde (Spain)
* Environmental Goods and Services Sector (United Kingdom)
* Low Carbon and Renewable Energy Economy Survey (United Kingdom)
* O\*NET (United States)
* United States Environmental Protection Agency (United States)
* Australian Skills Classification (Australia)
 |
| Publications by the European Commission | * Commission Delegated Regulation C(2021) 2800 final
* Directive 2003/87/EC of the European Parliament and of the Council
* PROF/TRAC Blueprint Project
* JRC
 |
| Publications by International Organisations | * International Labour Organisation
* United Nations Industrial Development Organisation
* OECD
 |
| Job boards | * GoodWork.ca
* Environmental Careers Organisation
* Indeed
* LinkedIn
 |
| Other | * Wikipedia
 |

# Next steps

The Commission will consult the ESCO MSWG on the translation of the new content and of the revised occupations, skills and knowledge starting from 30 November 2021. The consultation will last until 31 January 2022.

To this end, and similarly to the consultation process on the translation of concept descriptions, the Commission will circulate three separate excel files for occupations, skills and knowledge including both the new and the modified content. The files will include the following information: concept URI, preferred label, alternative labels, translated descriptions.

Member States should provide the suggested translation next to the original version. Feedback should be returned to the ESCO Secretariat: EMPL-ESCO-SECRETARIAT@ec.europa.eu

A separate consultation on the translation of the transversal skills structure will be organised in December. Feedback from the Member States will be implemented in a minor version which will be released in April 2022.

# Communication

To promote the launch of ESCO v.1.1, the Commission foresees the following communication activities:

* **News item**: a dedicated news article will be published in the ESCO portal
* **ESCO newsletter and partnering newsletters**: the news of the launch of v1.1 will take central stage in the ESCO newsletter; also, the news will be distributed to other partnering newsletters (EURES, CEDEFOP, ETF, etc.) for further dissemination to their networks.
* **Social media:** distribution on Twitter and LinkedIn
* **Online launch event:** an online webinar where stakeholders can learn about what is new in v1.1, ask questions and get inspired by concrete showcases of ESCO usage. The event is foreseen in the first quarter of 2022 and it will require pre-registration. The pre-registration link will be available at least 3 weeks before the event and it will be distributed through the above-mentioned channels. The webinar is open to all ESCO stakeholders interested in getting news about the classification and a practical understanding of how ESCO is used in HR or career management digital services, different EU or national projects or labour market research.

Member States are invited to share the news in their networks and circulate widely the invitation to participate in the online launch event.

1. see the note MSWG 14-05 Report on the ESCO continuous update cycle -ESCO v1.1 and v1.2. The final list of new occupations, skills and knowledge concepts is available in Annex 2 to the MSWG 14-05 document. [↑](#footnote-ref-2)
2. see Action 5 of the European Skills Agenda and Action 8 of the new ERA communication.
 [↑](#footnote-ref-3)
3. https://ec.europa.eu/esco/portal/document/en/f834e202-0ebf-461a-9249-a00e91d86e94 [↑](#footnote-ref-4)
4. [Skills-OVATE: Skills Online Vacancy Analysis Tool for Europe | CEDEFOP (europa.eu)](https://www.cedefop.europa.eu/de/tools/skills-online-vacancies) [↑](#footnote-ref-5)
5. [<https://ec.europa.eu/esco/portal/version>](https://ec.europa.eu/esco/portal/version) [↑](#footnote-ref-6)
6. Communication on a European Skills Agenda for sustainable competitiveness, social fairness and resilience, pag. 12. [↑](#footnote-ref-7)