# Leveraging Artificial Intelligence to update the ESCO Occupations Pillar

Report – May 2021

Employment, Social Affairs and Inclusion

# Table of Contents

1. Introduction ..... 2
2. Normalisation of job titles ..... 2
3. Linking new alternative labels and emerging occupations to ESCO ..... 4
4. Labour market job title neighbourhood for a single ESCO occupation ..... 6
5. Evaluating preferred terms for concepts ..... 8

## 1. Introduction

In the previous article[1], we showed how the European Commission uses Artificial Intelligence (AI) to gain additional insights for existing occupations and occupation groups of the ESCO taxonomy. In this article, we will focus on the challenges that the European Commission needs to address when updating the occupations pillar: Should new occupation concepts be added? To which extent are they similar to an existing occupation? To which occupation group do they belong? All these challenges are related to the fundamental problem of mapping textual descriptions of, potentially new, occupations to the existing ESCO occupations hierarchy.

This topic is closely related with the problem that the terminology that is used in various data sources is not always aligned with the terminology that is used in ESCO. For example, the job title '*Senior Android developer*' appears frequently in CVs, user profiles and job vacancies, but there is no exact mapping available to the existing ESCO occupation '*mobile application developer*'. I.e., the job title is not a preferred term or alternative label for an existing ESCO occupation. Therefore, it is fundamental to be able to map new content to ESCO. In the following sections different flavours of this key problem are described.

## 2. Normalisation of job titles

The current English version of the ESCO occupations pillar contains 37,196 preferred terms, alternative labels and descriptions for 2,942 occupations. In general, these labels do not contain elements as '*senior*', '*fixed-term contract*', '*part-time*' or other jargon that is typically found in CVs, user profiles or job vacancies. The process of mapping an occupation term, occupation description or job title to an occupation concept is sometimes described as normalisation or mapping to a canonical form. To address this challenge, the European Commission built a machine learning model to suggest potential matches to existing concepts based on input like job titles and occupation labels. This model was developed based on labour market expert input, Member States' national taxonomies, Qualification Data Register qualifications[2] and job vacancies. The following tables give a number of examples with

1 https://esco.ec.europa.eu/en/about-esco/data-science-and-esco/leveraging-artificial-intelligence-maintain-esco-occupations-pillar

2 https://europa.eu/europass/qdr/

suggested mappings for specific input. The top row of each table is the input for the model and top 5 suggestions with match score are listed underneath:

| **Veterinary receptionist - Brussels** | |
|---|---|
| 98 | Veterinary receptionist |
| 95 | Receptionists (general) |
| 92 | Veterinary nurse |
| 92 | Equine dental technician |
| 92 | Veterinary technician |

| **Experienced tax preparer - work from home** | |
|---|---|
| 97 | Tax advisor |
| 92 | Tax clerk |
| 91 | Tax compliance officer |
| 90 | Tax policy analyst |
| 89 | Tax inspector |

| **First class bricklayer** | |
|---|---|
| 99 | Bricklayer |
| 97 | Bricklayers and related workers |
| 96 | Bricklaying supervisor |
| 93 | Ceiling installer |
| 92 | Plasterer |

| **Daytime Server - Wait staff** | |
|---|---|
| 96 | Waiter/waitress |
| 88 | Head waiter/head waitress |
| 88 | Waiters and bartenders |
| 88 | Waiters |
| 82 | Bartenders |

| **Telesales rep** | |
|---|---|
| 90 | Contact centre salespersons |
| 88 | Call centre agent |
| 85 | Contact centre information clerks |
| 85 | Electricity sales representative |
| 85 | Customer contact centre information clerk |

| **Hairdressing Apprentice** | |
|---|---|
| 96 | Hairdressers |
| 95 | Hairdresser assistant |
| 95 | Hairdresser |
| 95 | Barber |
| 94 | Hairdressers, beauticians and related workers |

| **Caregiver - PRN** | |
|---|---|
| 87 | Home care aide |
| 83 | Companion |
| 82 | Healthcare assistant |
| 80 | Nurse assistant |
| 79 | Companions and valets |

| **Bottling operator - first shift** | |
|---|---|
| 93 | Packing, bottling and labelling machine operators |
| 92 | Wine fermenter |
| 92 | Packaging and filling machine operator |
| 92 | Canning and bottling line operator |
| 90 | Carbonation operator |

Compared to the tables above, the next examples show results for the same model in case the input is less straightforward to match to ESCO occupations:

| Solution sales specialist - Microsoft Surface | | Operative for groundworks | |
|---|---|---|---|
| 94 | ICT presales engineer | 97 | Civil engineering worker |
| 91 | Information and communications technology sales professionals | 93 | Road maintenance worker |
| 90 | ICT account manager | 93 | Building construction worker |
| 89 | ICT business development manager | 92 | Drainage worker |
| 86 | Sales manager | 92 | Grader operator |

| Tesla Advisor - Sales | | BMW Apprentice Technician | |
|---|---|---|---|
| 85 | Motor vehicles specialised seller | 92 | Vehicle technician |
| 83 | Car leasing agent | 90 | Refurbishing technician |
| 81 | Telecommunications equipment specialised seller | 90 | Automotive brake technician |
| 81 | Motor vehicle aftersales manager | 88 | Motor vehicle mechanics and repairers |
| 81 | Sales engineer | 88 | Restoration technician |

## 3. Linking new alternative labels and emerging occupations to ESCO

In general, the normalisation problem is further complicated by the dynamics of the labour market: new job titles for existing occupations come into fashion and new occupations might emerge. The result is that it is challenging to keep the taxonomy closely aligned. In order to assist in this process, the same model as described above was used to generate mapping suggestions for job titles that are not simply a mix of common labour market jargon with ESCO occupations. The following tables show the top 5 suggestions for some examples where the model had to capture the semantic properties of words:

| Senior Android developer | | Machine learning engineer | |
|---|---|---|---|
| 96 | Mobile application developer | 95 | Data scientist |
| 95 | Industrial mobile devices software developer | 93 | ICT intelligent systems designer |
| 95 | Software architect | 92 | Bioinformatics scientist |
| 94 | Software developers | 91 | Data warehouse designer |
| 94 | Software developer | 89 | Chief data officer |

| Web designer | | Code ninja | |
|---|---|---|---|
| 95 | Web and multimedia developers | 96 | Mobile application developer |
| 95 | User interface designer | 94 | Industrial mobile devices software developer |
| 93 | Web technicians | 93 | Software developers |
| 92 | Webmaster | 93 | Software architect |
| 92 | Web developer | 93 | Software developer |

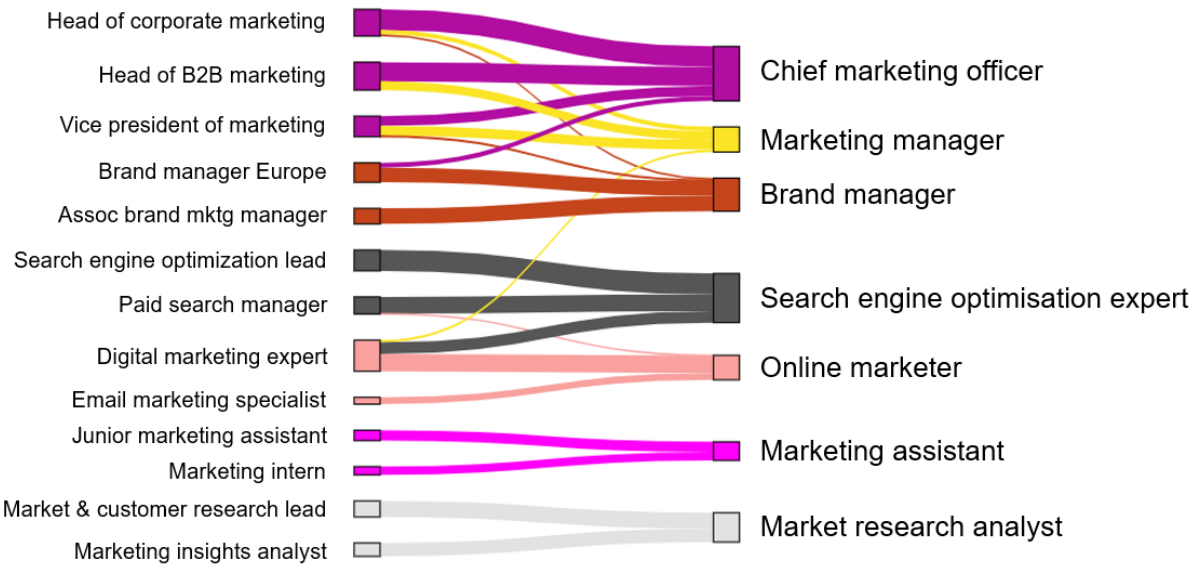
| Beauty consultant | | Medical claims processor | |
|---|---|---|---|
| 96 | Aesthetician | 92 | Insurance claims handler |
| 96 | Pedicurist | 90 | Insurance claims manager |
| 95 | Beauty salon attendant | 90 | Insurance collector |
| 95 | Beauticians and related workers | 90 | Loss adjuster |
| 94 | Tanning consultant | 86 | Insurance clerk |

| Scrum master | | Big data developer | |
|---|---|---|---|
| 91 | ICT product manager | 97 | Data warehouse designer |
| 87 | ICT project manager | 96 | Database designer |
| 85 | Software manager | 96 | Database designers and administrators |
| 83 | ICT quality assurance manager | 95 | Database developer |
| 83 | Insurance product manager | 94 | Software architect |

| Auto body technician | | Powertrain engineer | |
|---|---|---|---|
| 96 | Refurbishing technician | 93 | Mechanical engineers |
| 96 | Restoration technician | 91 | Engine designer |
| 94 | Motor vehicle body assembler | 90 | Automotive engineer |
| 93 | Coachbuilder | 90 | Rolling stock engineer |
| 90 | Automotive brake technician | 90 | Electric power generation engineer |

| PPC strategist | | Pharmacy cashier | |
|---|---|---|---|
| 95 | Search engine optimisation expert | 88 | Checkout supervisor |
| 93 | Advertising media planner | 88 | Pharmacy assistant |
| 92 | Advertising manager | 88 | Cashier |
| 91 | Marketeer | 87 | Pharmaceutical technicians and assistants |
| 91 | Advertising and marketing professionals | 87 | Pharmacy technician |

Using this model in combination with data extracted from labour market expert input, national occupation taxonomies, Europass user profiles and job vacancies, provides a powerful approach to assist in enriching the ESCO taxonomy with additional occupations or alternative labels. Frequently observed job titles that are not part of ESCO are good candidates for further investigation in this process. The following figure shows an example for ESCO occupations that are part of the marketing space:




None of the marketing-related job titles on the left are part of the ESCO taxonomy. Each of the ESCO occupations on the right is associated with a certain colour. The thickness of the line that connects it with a job title represents the strength of the association as computed from the model output. The job titles are coloured based on the closest matching ESCO occupation.

## 4. Labour market job title neighbourhood for a single ESCO occupation

Another approach to gain insights in the connection between ESCO occupations and labour market data is to inspect the job titles from Europass user profiles and job vacancies that are closely related to a specific ESCO occupation. The results for ESCO occupation '*mail clerk*' show that there exists a wide range of job titles that potentially refer to this occupation. The example for ESCO occupation '*underground miner*' shows job titles that can be grouped in: underground mining roles, general mining operators, heavy machinery operators, quarry operators and mining technician roles.

| Mail clerk | Underground miner |
|---|---|
| Mail services clerk | Experienced underground coal miner |
| Mail distribution clerk | Experienced underground longwall coal miners |
| Mail room clerk | Experienced underground miner |
| Mail clerk / Mail handler | Experienced underground coal miners |
| Mail clerk I | Mine technician - Underground mining |
| Mail clerk / general clerk I | Underground experienced miner |
| Mailroom clerk | Mine operator |
| Mailroom assistant | Miner operator |
| Mailroom services assistant | Multi skilled mining operator |

| | |
|---|---|
| Mailroom processing clerk | Miner |
| Mailer | Mining operators |
| Mailroom helper | Operator, mine services |
| Mail services associate | Underground construction miner |
| Mailroom service clerk | Haul truck operator - Mining |
| Clerk, mail room | Dozer operator - Mining |
| Mail associate | Loader operator - Mining |
| Mail processor | Quarry operator |
| Clerk, mailroom | Production miner |
| Intercept mailroom clerk | Mine support |
| Mail sorter | Quarrying specialist (12g) |
| Part time mailroom processing clerk | Quarry labourer (excavator licence) |
| Mail specialist I | Coal technicians |
| Postal clerk | Maintenance / operator (aggregate, mining) |
| Mail room associate | 12g quarrying specialist |
| Mail specialist | Entry-level underground miner |
| Mail services specialist I | Construction miner |
| Mailroom associate | Shaft construction miner |
| Mailroom operator | Mining field technician |
| Outgoing mailroom technician | Quarry operator / maintenance |
| Mailroom coordinator I | Blast hole driller - Mining |
| Mail processing specialist | Coal technician |
| Postal clerk (retail/operations) | Quarry technician |
| Mail handler 3rd shift | Dozer operators |
| Part time mailroom processing assistant | Underground operator |
| Mail sorter operator | Ancillary operator (dozer, loader) |
| Mailroom coordinator | Quarry maintenance |
| Warehouse mail sorter | Mining supervisor |
| Postal technician I | Mine/ore control open pit geologist |
| Mail room attendant | Plant operator/groundsperson-Quarry (day shift) |
| Mailroom coordinator II | Quarry loader - Quarry |
| Mail coordinator | Equipment operator - Quarry (day shift) |
| Part-time office services representative (mailroom) | Scraper operators |
| Mail handler 1st shift | Quarry worker |
| Mail services coordinator | Entry level driller offsiders / labourer (mining) |
| Mailing operator | Plant operator - Quarry (night shift) |
| Back up mail carrier | Mining site supervisor |
| Customer service associate - Mailroom | Scraper operator |
| Warehouse mail sorter - 2nd shift | Mainline excavator operator |
| Warehouse mail sorter - 3rd shift | Dozer / loader operator |

## 5. Evaluating preferred terms for concepts

This last section focuses on investigating whether an ESCO concept has a representative preferred term by leveraging the occupation mapping model. For example, the ESCO concept with preferred term '*ICT intelligent systems designer*' has the following description: '*ICT intelligent systems designers apply methods of artificial intelligence in engineering, robotics and computer science to design programs which simulate intelligence including thinking models, cognitive and knowledge-based systems, problem solving, and decision making. They also integrate structured knowledge into computer systems (ontologies, knowledge bases) in order to solve complex problems normally requiring a high level of human expertise or artificial intelligence methods*'. This concept has alternative labels: '*AI system designer*', '*linked data engineer, intelligent systems developer*', '*semantic technologies engineer*', '*designer of intelligent systems'*, '*machine learning systems designer*', '*intelligent systems engineer*', '*intelligent system designer*', '*intelligent systems designers*', '*artificial intelligence designer*'. Based on frequency data from CVs and vacancies, it can be derived that the preferred term '*ICT intelligent systems designer*' is not a common label to describe this occupation. Feeding the machine learning model with the alternative labels as input, results in the following job vacancy/CV job title suggestions:

| **ICT intelligent systems designer** | | **Medical administrative assistant** | |
|---|---|---|---|
| 95 | Machine learning engineer | 95 | Medical secretary |
| 95 | Data modeller | 92 | Medical administrative assistant |
| 95 | Data architect | 90 | Practice assistant |
| 95 | Data modeler | 90 | Medical office assistant |
| 94 | BI solution designer | 90 | Medical support assistant |
| 94 | Data scientist - Machine learning | 90 | Medical reception technician |
| 94 | Data engineer | 89 | Practice office associate |
| 94 | Data scientist | 89 | Advanced medical support assistant |
| 94 | Data visualisation specialist | 89 | Medical office specialist |
| 94 | Staff data scientist | 89 | Clinical support assistant |

On the other hand, repeating the same exercise for the concept with preferred term '*Medical administrative assistant*' results in suggestions that do include and/or are very well aligned with the expected preferred term.

This article explained how ESCO is leveraging artificial intelligence through occupation mapping to enrich the occupations pillar. Occupation normalisation and mapping new job titles for existing/new occupations is pivotal in this process. In the next article we will focus on the importance of going one step further than only considering job titles.

If you are an ESCO implementer and want to share your feedback, please get in touch via email at EMPL-ESCO-SECRETARIAT@ec.europa.eu, use our hashtag #ESCO_EU and join our online forum.