
Standard Occupational Classification (SOC) 2020 Extended Framework

Methodology for the development of Sub Unit Group (SUG) descriptions within SOC

Background

[SOC 2020 Volume 1](#) SOC 2020 Volume 1 contains the framework structure for 4-digit SOC as well as descriptions for each of the 9 Major Groups, 26 Sub Major Groups, 91 Minor Groups and 412 Unit Groups identified within.

Some existing users of the SOC felt that the current 4-digit structure was not granular enough for their needs. The SOC Extension project, born from this stakeholder demand, has resulted in the addition of a fifth layer to the classification which comprises of 1,463 'Sub Unit Groups' (SUGs).

The development of accurate SUG descriptions is essential in enabling both the automatic and manual matching of data. In turn, this is key to realising the long-term goal of statistical production at the extended level.

Challenge

The Classifications team were keen to avoid duplication and therefore aimed to base the SUG descriptions on existing available information. Numerous potential sources were identified from which descriptions could be compiled. These included international classifications such as [O*net](#) (US) and the European Skills/Competencies, Qualifications and Occupations framework ([ESCO](#)) career advice websites, job vacancy portals and industry specific organisations.

Despite the wealth of data available, there was no identifiable single source to which all SOC 2020 SUGs could be matched. Taking descriptions from multiple sources risked inconsistencies in the formatting and language used. Furthermore, the description of 'job titles' that were found in many of the identified sources rather than 'occupation groups' limited the direct transferral of existing descriptions to the extended SOC framework.

As such, it was determined that whilst information within these existing descriptions could be used to inform the writing of SUG descriptions, these would also require elements of tailoring prior to adoption as SUG descriptors.

Development Process

Step 1: Develop a crosswalk to an existing key source

Following a review of the available sources alongside consultation with existing SOC users, it was concluded that ESCO would provide the most appropriate and complete source for an initial mapping exercise to be conducted.

ESCO contains 2,942 occupation groups making it far more granular than both the current UK version and the experimental extended version of SOC. Linking here enabled SOC to adopt existing descriptions as well as key skills and competencies already available within the European classification.

A manual exercise was undertaken to identify matches between the two classifications. Four fifths of all SUGs (excluding not elsewhere classified (n.e.c.) groups) were able to be mapped to at least one corresponding ESCO group. There were 5 possible ways in which a SOC to ESCO match could be classified.

Mapping type	Description
Unmapped	No appropriate match could be found with ESCO to match the SUG within SOC.
Partial	The SUG was able to be matched to a group in ESCO however the description of the ESCO group only partially covered the SUG.
One to one	A direct match was identified between a SUG and a corresponding occupation group used by ESCO.
One to multiples	Due to the greater level of disaggregation found in ESCO, there were many instances in which a single SUG was found to map to multiple ESCO groups.
n.e.c.	Refers to those groups within SOC that contain occupations not elsewhere classified within the framework.

Step 2: Identify and fill data gaps

Each of the five mapping types required a slightly different approach.

Mapping type	Action required:
Unmapped	Identification of an appropriate alternative source/s from which to adopt a group description. Key skills and competencies were also sought from a range of sources.
Partial	Identification of an appropriate alternative source/s to further inform the development of the group description. In these cases, descriptions are likely to become either bespoke or hybrid in nature. Key skills and competencies were also enhanced from additional sources.
One to one	Where a straight match to ESCO could be identified efforts were made to retain the description to as great a degree as possible. This was measured against also ensuring a consistent use of language across the framework. Key skills and competencies were also harvested from the ESCO match.
One to multiples	A new description was developed which aimed to merge the identified ESCO groups together. In some instances, the description also required the inclusion of information from additional sources. Skills and competencies were merged from the relevant ESCO groups.
n.e.c.	These groups refer to the existing SOC Unit Group descriptions. The often individually distinct set of occupations contained within an n.e.c. group makes description at the extended level problematic. A reverse ESCO to SOC

mapping exercise would also be required to identify the ESCO codes to be assigned here.

Other sources that were used for researching and developing the remaining group descriptions, as well as sense checking those adopted from ESCO included:

- O*net
- Careers advice websites e.g. The National Careers Services (NCS), Prospects, CareerSmart and My World of Work from Skills Development Scotland.
- Industry specific websites – examples included police or army roles and well as occupations within the nuclear sector.
- Job vacancy portals such as indeed.co.uk, totaljobs.com and ziprecruiter.com

Step 3: Edit and review

Following the development of SUG descriptions a further editing exercise was undertaken. The purpose of this was to ensure:

- A commonality of language across the framework. This is particularly necessary to eliminate differences in the language used between O*net and ESCO.
- Application of plain English text throughout.
- Removal of Americanised spelling from the text.
- Standardisation of commonly used terms used throughout.
- Grammar check.