HumRRO
HUMAN RESOURCES RESEARCH ORGANIZATION

2025 No. 127

Updates to Related Occupations for the O*NET Program Using the O*NET 30.0 Database

Technical Memorandum

Prepared for:	National Center for O*NET Development 313 Chapanoke Road, Suite 130 Raleigh, NC 27603	**Prepared under:**	Subcontract Number (through RTI International): 1-312-0207142-41224L
Authors:	Jessica L. Harris Jeffrey A. Dahlke	**Date:**	November 17, 2025

Updates to Related Occupations for the O*NET Program Using the O*NET 30.0 Database: Technical Memorandum

Table of Contents

List of Tables

Updates to Related Occupations for the O*NET Program Using the O*NET 30.0 Database: Technical Memorandum

Introduction

The Occupational Information Network (O*NET) is a system developed by the U.S. Department of Labor that provides information on over 900 occupations covering more than 55,000 jobs within the U.S. economy. This information is maintained in a comprehensive database and, at the time we began this research, the most current version was the O*NET 30.0 Database (National Center for O*NET Development, 2025). To keep the database current, the National Center for O*NET Development (hereafter referred to as "the Center") is involved in a continual data collection process aimed at identifying and maintaining current information on the characteristics of workers and occupations. For years, the Human Resources Research Organization (HumRRO) has supported the Center's efforts to maintain the database. This report focuses on work to update the related occupations associated with the 923 data-level occupations included within the O*NET-Standard Occupational Classifications (O*NET-SOC) 2019 taxonomy (Gregory et al., 2019).

The Center has a longstanding history of conducting research to populate related occupations information for the occupations included in the taxonomy (Allen et al., 2012; Drewes et al., 1999). An end user's ability to discover and review related occupations, initially defined within O*NET as occupations with common work attributes and similar worker requirements, is a central component of career exploration, career guidance, and job search. Related occupations information is also often leveraged in a variety of other critical world-of-work-related activities, including human resource functions, workforce development, and basic research.

The Center's current methodology for identifying related occupations was developed and reported by Dahlke et al. (2022). This approach facilitates regular updates to O*NET's related occupations information and operationalizes relatedness using three important contributors to occupational similarity: what people in the occupations *do* (based on task statements and detailed work activities), what they *know* (based on profiles of importance ratings for O*NET's knowledge domains), and what they are *called* (based on lists of job titles). Each occupation is mapped to 10 primary related occupations (i.e., relations that will be displayed in the O*NET database) as well as 10 supplemental related occupations (i.e., the 10 most relevant nonprimary occupations, determined based on structural relations in the O*NET-SOC taxonomy and empirical similarity analyses). Although the approach is primarily based on empirical relatedness information, it also provides an opportunity for the Center to review all lists of empirically related occupations and offer recommendations on which related occupations should be categorized as primary relations.

The last update to the related occupations information occurred in 2023 using data from the O*NET 28.0 Database (Dahlke & Harris, 2024; National Center for O*NET Development, 2023). In the present research, we revisited occupational relationships using data updated by the O*NET Data Collection Program over approximately two years and available in the O*NET 30.0 Database. In the following sections, we describe our technical approach and the high-level outcomes of our updates.

Technical Approach

We followed the same analysis strategy for evaluating relations among O*NET-SOCs as described in Dahlke et al. (2022). Below we offer a brief overview of our technical approach; for more detailed information, please refer to Dahlke et al. (2022).

Calculation of Relatedness Scores

For each pair of data-level O*NET-SOCs, we computed an empirical relatedness score as a function of three similarity analyses:

- **Work-Based Occupational Similarity Metric (WB-OSM):** A similarity index that represents the average of (a) the mean cosine similarity among the O*NET-SOCs' task statements and (b) the mean cosine similarity among the O*NET-SOCs' Detailed Work Activities (DWAs). We based these cosine analyses on quantitative embeddings derived from tasks and DWAs using the "nli-distilroberta-base-v2" Sentence-BERT (SBERT) model (Reimers & Gurevych, 2019). We limited our analyses to tasks that were classified as "core" when possible. When an O*NET-SOC had no core tasks (30 of the 923 O*NET SOCs), we analyzed all tasks associated with that O*NET-SOC.
- **Job Titles (JT) Cosine:** The estimated cosine similarity between the O*NET-SOCs' lists of job titles. We based these cosines on word-level GloVe ("global vectors for word representation") embeddings (Pennington et al., 2014) that we averaged into O*NET-SOC-level embeddings using Term Frequency-Inverse Document Frequency (TF-IDF) weights. The IDF components of the weights were computed within each O*NET-SOC's respective job family.
- **Knowledge Importance Profile (KIP) Cosine:** The estimated cosine similarity between the O*NET-SOCs' profiles of knowledge importance ratings. There are 29 O*NET-SOCs missing knowledge importance ratings. This metric was not estimable when one or both of the O*NET-SOCs in a pair were missing these ratings.

For all three metrics, we standardized the similarity values according to the norms displayed in Table 1. After combining these three standardized variables into an equally weighted composite, we standardized the composite using the norms from Table 1 to give it a more interpretable scaling. When expressed as a single formula, we computed relatedness scores as follows[1]:

$$Relatedness = \frac{\left(\frac{\frac{WBOSM-0.302}{0.063}+\frac{Cosine_{JT}-0.182}{0.173}+\frac{Cosine_{KIP}-0.924}{0.035}}{3^*}\right)+0.002}{0.842}$$

*If knowledge importance ratings were unavailable for at least one O*NET-SOC involved in a given pair, this 3 was replaced with a 2 to ensure the available values were averaged correctly.

[1] The formula for calculating relatedness scores uses exact unrounded values for the three similarity indices and associated norms. The values presented in the formula and Table 1 are rounded to three decimal places for presentation purposes. Use of rounded values versus unrounded values will result in slightly different relatedness scores.

Table 1. Norm Distributions Used to Scale Related Scores

Score Type	Mean	SD
Work-Based Occupational Similarity Metric (WB-OSM)	0.302	0.063
Job Titles (JT) Cosine	0.182	0.173
Knowledge Importance Profile (KIP) Cosine	0.924	0.035
Average of WB-OSM, AT Cosine, and KIP Cosine *Z* Scores	-0.002	0.842

Note. These norms were based on the O*NET 28.0 Database, which was the first database version used to update the related occupations matrix following the initial implementation of Dahlke et al.'s (2022) methods.

Identification of O*NET-SOCs Requiring Review

An O*NET-SOC's related occupations information required review if there were changes to the empirically predicted primary relations for the O*NET 30.0 Database compared to the predicted primary relations from the previous time the occupation was reviewed. If the list of the 10 most similar related O*NET-SOCs was unchanged since a target O*NET-SOC's related occupations were reviewed, that target O*NET-SOC's published list of primary occupations was carried forward. However, even if a target O*NET-SOC was not reviewed, the primary related O*NET-SOCs could move freely between the "short" and "long" primary related occupations lists as a function of their empirical similarity levels (the "short" list represents the top five most empirically similar primary related O*NET-SOCs, and the "long" list represents the remaining five primary related O*NET-SOCs). After comparing results from analyses performed on the 30.0 Database to prior analyses, we identified 635 O*NET-SOCs that required review by staff at the Center.

Review Process

For each target O*NET-SOC requiring review, HumRRO constructed a review worksheet listing the following information:

- The target O*NET-SOC code.
- The target O*NET-SOC's description.
- The 10 current primary related O*NET-SOCs for the target O*NET-SOC, obtained from the target O*NET-SOC's most recent publication matrix.
- The 15 most relevant related O*NET-SOCs for the target O*NET-SOC that were not part of the primary related O*NET-SOC set.
  - These were O*NET-SOCs that were structurally related to the target O*NET-SOC in the current O*NET taxonomy or had the next-highest empirical similarity after the primary related O*NET-SOCs. We labeled these as "alternate" related O*NET-SOCs for the purposes of the review.

For each related O*NET-SOC, we indicated whether the empirical similarity analyses suggested a potential justification for demoting primary related occupations to supplemental occupations or elevating alternate related occupations to primary related occupations. The reviewer's task was to evaluate the lists of primary and alternate related occupations and, if necessary, recommend which alternate related occupations should replace the primary related occupations they deemed unsuitable. If a reviewer decided that a primary related occupation was not suitable as

either a primary or supplemental related occupation, they had the option to exclude the related occupation from use by marking it as “do not display.”

We organized the review sheets for the target O*NET-SOCs requiring review into a collection of Excel workbooks, each of which included up to 15 target O*NET-SOCs. For the present update, we constructed 43 workbooks for the Center to review.

Summary of Updates

The Center made replacements to the primary related occupations lists for 149 of the target O*NET-SOCs (23.46% of the 635 target O*NET-SOCs included in the review and 16.14% of the 923 data-level O*NET-SOCs in the O*NET 30.0 Database). Across these 149 target O*NET-SOCs, the review resulted in a total of 190 replacements to the lists of primary related occupations (2.99% of the 6,350 primary related O*NET-SOCs included in the review). The Center did not identify any primary related O*NET-SOCs to add to the cumulative set of “do not display” related occupations that are ineligible to appear in the O*NET operational related occupations matrix. See Table 2 for a summary of the number of replacements made per target O*NET-SOC.

Table 2. Cumulative Frequencies of Empirical Top-10 Related Occupations Replaced by Reviewers

Frequency of Replacements per Target O*NET-SOC	Non-Cumulative Summary of Target O*NET-SOCs		Cumulative Summary of Target O*NET-SOCs with a Number of Replacements Less Than or Equal to Row Frequency		Cumulative Summary of Target O*NET-SOCs with a Number of Replacements Greater Than or Equal to Row Frequency	
	k	%	*k*	%	*k*	%
0	486	76.54	486	76.54	635	100.00
1	118	18.58	604	95.12	149	23.46
2	24	3.78	628	98.90	31	4.88
3	4	0.63	632	99.53	7	1.10
4	3	0.47	635	100.00	3	0.47

Conclusion

We updated the related occupations information for all 923 data-level O*NET-SOCs based on data from the O*NET 30.0 Database. This update included changes to the lists of primary related occupations for 149 of the 635 O*NET-SOCs reviewed by the Center. We also refreshed the lists of supplemental related occupations for all data-level O*NET-SOCs and reallocated primary related O*NET-SOCs between the “short” and “long” lists of related O*NET-SOCs based on updated empirical relatedness scores.

After the Center reviewed the related occupations lists and we accounted for their recommended edits, we constructed the following final work products for inclusion in the O*NET 30.1 Database (see the Appendix for more information about these files’ contents):

- An “Operational Related Occupations Matrix” that can be used to identify which related O*NET-SOCs to display on O*NET’s suite of websites, which includes 10 primary related occupations and 10 supplemental primary related occupations for each O*NET-SOC.

- A “Related Occupations Research Dataset” that contains quantitative similarity information about how each of the 923 data-level O*NET-SOCs relates to each of the other 922 data-level O*NET-SOCs.

References

Allen, M.T., Waugh, G.W., Shaw, M., Tsacoumis, S., et al. (2012). *The development and evaluation of a new O*NET Related Occupations Matrix (ROM)* (FR-12-29). Human Resources Research Organization. https://www.onetcenter.org/reports/Related.html

Dahlke, J. A., & Harris, J. L. (2024). *Updates to related occupations for the O*NET program: Technical memorandum* (2023 No. 157). Human Resources Research Organization. https://www.onetcenter.org/reports/Related_2024.html

Dahlke, J. A., Putka, D. J., Shewach, O. R., & Lewis, P. (2022). *Developing related occupations for the O*NET program* (2022 No. 036). Human Resources Research Organization. https://www.onetcenter.org/reports/Related_2022.html

Drewes, D.W., Tarantino, J.W., Atkins, S.G., & Paige, B.A. (1999). *Development of the O*NET Related Occupations Matrix (ROM)*. National Center for O*NET Development.

Gregory, C., Lewis, P., Frugoli, P., & Nallin, A. (2019). *Updating the O*NET®-SOC taxonomy: Incorporating the 2018 SOC structure*. National Center for O*NET Development. https://www.onetcenter.org/reports/Taxonomy2019.html

National Center for O*NET Development. (2023). O*NET 28.0 database. https://www.onetcenter.org/dictionary/28.0/excel/

National Center for O*NET Development. (2025). O*NET 30.0 database. https://www.onetcenter.org/dictionary/30.0/excel/

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. https://doi.org/10.3115/v1/D14-1162

Reimers, N., & Gurevych, I. (2019). *Sentence-Bert: Sentence embeddings using Siamese BERT-networks. ArXiv:1908.10084 [Cs]*. http://arxiv.org/abs/1908.10084

Appendix A: Data Dictionaries/Codebooks for the Operational Related Occupations Matrix and Related Occupations Research Dataset

Table A-1. Data Dictionary/Codebook for the Operational Related Occupations Matrix

Variable Name	Variable Type	Variable Description
O*NET-SOC Code	Nominal String	O*NET-SOC codes for target occupations.
Title	Nominal String	O*NET titles for target occupations.
Related O*NET-SOC Code	Nominal String	O*NET-SOC codes for related occupations.
Related Title	Nominal String	O*NET titles for related occupations.
Relatedness Tier	Ordinal String	Ordinal categories indicating level of relatedness after expert review. Determined separately within each target O*NET-SOC. Used to determine which related occupations to display on web resources. - Primary-Short = Five most strongly related occupations after expert review. - Primary-Long = 6th to 10th most strongly related occupations after expert review. - Supplemental = 11th to 20th most strongly related occupations after expert review.
Index	Integer	Ordering of related occupations after expert review. Determined separately within each target O*NET-SOC.

Table A-2. Data Dictionary/Codebook for the Related Occupations Research Dataset

Variable Name	Variable Type	Variable Description
O*NET-SOC Code	Nominal String	O*NET-SOC codes for target occupations.
Title	Nominal String	O*NET titles for target occupations.
Related O*NET-SOC Code	Nominal String	O*NET-SOC codes for related occupations.
Related Title	Nominal String	O*NET titles for related occupations.
Relatedness Tier	Ordinal String	Ordinal categories indicating level of relatedness after expert review. Determined separately within each target O*NET-SOC. Used to determine which related occupations to display on web resources. - Primary-Short = Five most strongly related occupations after expert review. - Primary-Long = 6th to 10th most strongly related occupations after expert review. - Supplemental = 11th to 20th most strongly related occupations after expert review. - N/A = Not categorized.
Index	Integer	Ordering of related occupations after expert review. Determined separately within each target O*NET-SOC. Values are only displayed for related occupations that were included in the Operational Related Occupations Matrix.
Work-Based Occupational Similarity	Numeric	Simple average of SBERT-based WB-OSMs computed using (a) task statements and (b) detailed work activities (DWAs).
Knowledge Cosine	Numeric	Cosine similarity between occupation's profiles of importance ratings for O*NET's 33 knowledge domains.
Job Titles Cosine	Numeric	Cosine similarity between the TF-IDF-weighted GloVe embeddings for occupations' job titles.
Relatedness Score	Numeric	Final quantitative metric for determining occupational similarity/relatedness.
Empirical Relatedness Rank	Integer	Rank ordering of related occupations based on descending values of Relatedness Scores within each target O*NET-SOC.