



February 17, 2023

Nora Kincaid
BLS Clearance Officer
Division of Management Systems
Bureau of Labor Statistics
2 Massachusetts Avenue NE, Room G225
Washington, DC 20212

**Re: Federal Register / Vol. 87, No. 242 / December 19, 2022 / Notices
Request for Comment re: Occupational Requirements Survey (ORS) Information
Collection – Third Wave**

Dear Ms. Kincaid:

SkillTRAN LLC is a small private company that builds PC and Web-Based solutions to easily navigate data and information about occupations, occupational requirements, and labor market information. We have customers in all 50 states. We serve multiple markets, including public and private vocational rehabilitation organizations, vocational experts, the Veterans Administration, worker compensation programs, long term disability insurance companies, claimant disability representatives and attorneys, and the Social Security Administration. Collectively, tens of thousands of people serving the varied needs of people with disabling conditions continue to rely on the Dictionary of Occupational Titles (DOT) for detailed information about what it takes to perform occupations. We build solutions that integrate all this data to address real world challenges.

SkillTRAN products integrate the rich (though aging) DOT data with its 12,761 unique occupations and 72 discrete occupational characteristics per DOT occupation (a total of 867,748 data elements) with other data resources such as employment numbers (reported at the SOC 2018 classification level in the Bureau of Labor Statistics (BLS) Occupational Employment and Wage Survey (OEWS) program, long term Employment Projections (EP), industry as classified in the North American Industry Classification System (NAICS), County Business Patterns, Current Employment Statistics, and OEWS staffing patterns. It is exactly this kind of data that must be integrated to build solid solutions and opinions with current relevant data. SkillTRAN pioneered a unique method to guide job placement and estimate employment at the DOT level by creating a special cross-reference table from DOT to appropriate NAICS industries.

Since 2019, SkillTRAN is the first company (and only ... to our knowledge) to have integrated the newest ORS data into many of its products for super easy access to this rich new ORS data set. We know these data sets very well. We understand what it takes to both integrate new data with older data sources like the DOT and to maintain it as classification systems change on a regular basis.

Q. Is the proposed “Third Wave” of data collection necessary? ...of practical utility?

A. Given the age of the DOT (last formally updated in 1991, but with 80% of the DOT occupations unreviewed since 1977), there is no question that 30+ and 45+ year old DOT data is long overdue for an update! However, collapsing the 12,761 DOT occupations into a mere 848 civilian SOC 2018 Groups is a major shift in level of precision, particularly when considering the multiple factors by which each group is rated. The reality is that nearly 80% of the SOC groups contain multiple DOT occupations – ranging anywhere from 2 DOTs to one 6-digit SOC group with more than 1500 DOT occupations! You can imagine (and I can show) the diversity of different levels of Strength and Specific Vocational Preparation (SVP) as the number of related DOT occupations increases in a SOC group. This leads to a lot of heterogeneity in many factors reported by SOC Group, particularly SVP, Strength, and many other physical demand factors. To collapse so many occupations into a simpler system (DOT → SOC), it is crucial that valid SOC group data of sufficiently sized sample sizes be collected for ALL six-digit SOC Groups.

When the DOT data was collected by the (now discontinued) Field Analysis Centers, trained BLS Job Analysts (not economists) went on-site to collect data by interview and direct observation. About 75,000 of these on-site job analyses were done, each with full job descriptions and ratings for every one of the 72 different variables. Through a data collection period of about 5 years (late 1972 – 1977), the Fourth edition of the DOT was published in 1977 with data for all 72 variables for each of 12,099 descriptions. This represents data collection of 871,128 data points in 5 years ... by hand ... before computers!

Contrast this with the 4th year of second wave data collection by ORS, which at the end of 2022 reports a total of 60,150 data points for only 426 of the 848 civilian SOC groups. This are 141 data values on average reported (of 343 data points possible to report). Within these 426 reported ORS SOC occupations, sometimes the reported data omits even the very basics essential to SSA (such as SVP and Strength).

With at least \$300 million spent by ORS so far, this is an *average cost of \$704,000 per occupation* and nearly \$5,000 cost per data point. And this data collection is just barely half complete (in terms of reported data) with only 50.2% of SOC occupations reported. I would be delighted if ORS could complete the job by the end of the second wave of data collection (end of FY 2023), but I have no expectation that it will happen. The third wave of data collection is planned now for yet another 5 years (FY 2024-2028 – per section II Current Action, but in Section I – Background, three years of data collection are stated (2023-mid-2026). **Which is it?** As a taxpayer, I have a very difficult time digesting the amount of time and money that this

project is consuming at its current pace and price. I support completion of this essential task because current and complete data is needed for SSA to better adjudicate its 2+ million claims per year, but I expect nothing less than full collection of data for each of the 848 civilian SOC occupations by the end of this “Third Wave”, however long it is planned. Any lesser result would be completely unacceptable and would fail SSA in meeting its obligations.

Q. Evaluate the accuracy of the ... burden, including the validity of the method and assumptions.

Sample Size. The data collection for ORS at present is reported for just over 50% of all the 848 SOC groups. It is difficult to expect that after 8 years of data collection (first + second waves) that magically all these 848 SOC groups will be reported for the final set in Fall, 2023. **The absence of a completed data set is NOT an acceptable outcome of this second wave data set.**

Complex Math. The mathematics involved in implementing this SOC based system of ORS factors with OEWS data is quite complex (see the “Pooled Variance (90%) Math” for just 1 ORS variable here for 1 SOC group: <https://www.bls.gov/ors/factsheet/calculating-occupational-employment-for-job-requirements.htm>) to compute “job numbers”. This result for calculation of this one factor then needs to be repeated for as many additional factors as are involved for each of the hypothetical disabling conditions. A Social Security Administrative Law Judge (ALJ) always asks hypothetical questions that involve many more than a single factor. Because of the mathematical complexity of this calculation, this probability math will have to be automated to be swiftly, properly, and reliably done. Further, each time an additional factor is calculated into the ALJ hypothetical, the standard error (SEM) of the resulting calculation will increase. At some point the SEM will become so wide as to render ALJ decision making about calculated job numbers both confusing and perplexing. This facilitates the ALJ decision making process how?

Inadequate descriptive disclosure by ORS about each reported SOC. Reporting by the ORS does not include sample size by 6-digit SOC group, nor is there any indication of the NAICS that were surveyed, nor the proportion of the survey dedicated to each NAICS. This is basic descriptive statistical reporting about sample size. At a 4-digit level of NAICS coding, there is no chance of disclosing who a surveyed employer might be. **This same depth of reporting is already done by sister agencies within the BLS, including both the OEWS Survey Group and the EP survey group.** Both these groups report data down to 30-50 people nationally by 4-digit NAICS industry coding! There is no reason why ORS cannot similarly disclose both its sample size and the proportion of NAICS industries it has surveyed at the 4-digit level of NAICS coding in its sample collection so that it can be directly compared to the OEWS survey results. Disclosure of numbers at the 4-digit NAICS level will protect employer confidentiality and establish that ORS has indeed sampled appropriately following the same stratification sampling by the excellent OEWS program. **For SSA to use this data confidently, this kind of data must be disclosed by ORS.** Federal constraints on data reporting to protect employer confidentiality apply equally to all BLS programs. ORS has failed to disclose any of this data thus far.

The use of data reported for “rolled up” two-digit SOC codes (e.g. 51-0000 Production Occupations) overestimates employment in basic areas such as Strength (Sedentary and Light) and Specific Vocational Preparation (SVP) because data for the underlying 6-digit SOC codes remains incomplete and because each of the 6-digit SOC codes occurs with different frequency (per OEWS data). Applying rolled up ORS data (which appears to assume equal distribution across all the SOCs reported) is a serious error and could potentially overestimate the number of sedentary unskilled occupations by 50% or possibly more. For example, in the SOC 51-0000 rollup, there are a total of 107 6-digit SOC groups with corresponding OEWS data for 105 of these SOC groups. ORS is reporting data for only 57 of these 6-digit SOC groups, with 50 SOC groups not yet reported by ORS. This is why 2-digit rollups of ORS data cannot be used at this point ... and when they are calculated, they must be reported as proportionately weighted to each of the related OEWS values at the 6-digit SOC level. Unfortunately, BLS economists are making grand and likely inaccurate statements about the proportions of some of these characteristics for “All civilian workers” on its main web page at <https://www.bls.gov/ors> - **Any calculations and percentages must be couched in the context of proportional weighting by the reported frequency in the OEWS survey and for which ORS data is available only. This is not being done and it is leading to exaggerated “facts”.**

Our analysis of the SOC 51-0000 Production occupations is interesting in that this single 2-digit SOC grouping cross-references to 60.1% of all occupations in the DOT (n=7,663/12,761). Yet SOC 51-0000 covers only about 5% of the entire civilian labor force. This confirms that the DOT went “overboard” in data collection in the Manufacturing Sector. This supports exactly why ORS must report the 4-digit NAICS for each SOC to appropriately proportional weight by OEWS.

Q. Enhance the quality, utility, and clarity of the information collected.

A. Not collected in this ORS survey is any information about the WORK Fields or MPSMS codes to provide appropriate methods for transferability of skills searches, a federal Code of Federal Regulations (CFR) requirement [(http://www.ssa.gov/OP_Home/cfr20/404/404-1568.htm - Section (d) (4)] for certain kinds of claims decision making by SSA. There is no equivalent concept in the SOC 2018 system, in this ORS Survey, nor in O*NET itself. **This is a glaring oversight to not find some way to collect and confirm this kind of data in the data collection.**

ORS data collection forms include areas for job description and lists of tasks performed. Yet nowhere in released ORS data is this information being publicly shared. Task data should be reported, but so far has not been disclosed in either the final First Wave or Second Wave data to date.

ORS data is being collected from employers and human resources people. This is very different from direct job observation, for which economists are not truly trained to collect nor to recognize the difference between what is said vs. how jobs actually do get done, which trained job analysts detect and objectively report.

Business owners and HR respondents may not be truly knowledgeable about all the areas for which ORS data is collected. An employer's judgment/rating may be clouded by their experience rate of worker compensation claims incidents filed, which could artificially alter the true requirements of a job.

One of the key differences between the DOT data collection and the ORS data collection is the nature and methods of data collection. The DOT was collected by BLS trained job analysts, collecting about 75,000 different job analyses, which included both interview with the employer PLUS direct observation (and rating) of the occupation being performed, which often revealed subtle differences NOT KNOWN by the human resources (HR) respondent. With every job analysis conducted by SkillTRAN customers in their unique circumstances (particularly private rehabilitation practitioners, worker compensation and long-term disability), **on-site job analysis with direct observation and measurement of worker requirements is the standard practice.** The on-site visit almost always reveals many more details about the occupation that are NOT known by HR personnel.

In ORS, economists are the data gatherers, who for the most part, do not typically observe the occupation being performed. Economists focus on high level aspects of various conditions. There is a **SUBSTANTIAL QUALITATIVE DIFFERENCE** in the methodology of each type of occupational observation. Will economists replace vocational experts in subsequent SSA vocational expert consultations for ALJ claimant hearings?

Q. Minimize respondent burden, including electronic submission of responses.

A. ORS estimates that it takes approximately 66 minutes to complete a single survey. On average, it is my understanding (no formal published data) that the average number of SOC occupations surveyed per employer is about 5. In 66 minutes (average) there are 3960 seconds. Each survey (example Form PPD-4PF) has 70 variables for data collection. With 5 occupations on average per employer survey, that is 350 data points to be collected in those 3960 seconds. The average amount of time for the employer to hear the question, ponder, respond, and record the response is about 11 seconds per question. It does not seem reasonable that high quality data can be gathered this quickly for each of the 70 questions per occupation in just 66 minutes. Ideally, electronic reporting, capture and submission of this information would reduce the burden on ORS staff, but any electronic version used by employers must be very clear, simple, and articulate, with excellent tool tip (hover over help) and help file explanation of sometimes obscure questions so that the respondent is able to discern a proper response. This will take a lot of careful user testing to assure ease of completion.

Please accept my responses here as evidence of my continuing very strong interest in the best possible collection of ORS data for the proper use of the fully collected data set to support the accurate determination of disability claim outcomes for persons with disabilities.

Jeffrey A. Truthan – MS – Rehabilitation Counseling, Certified Vocational Evaluator
President – SkillTRAN LLC