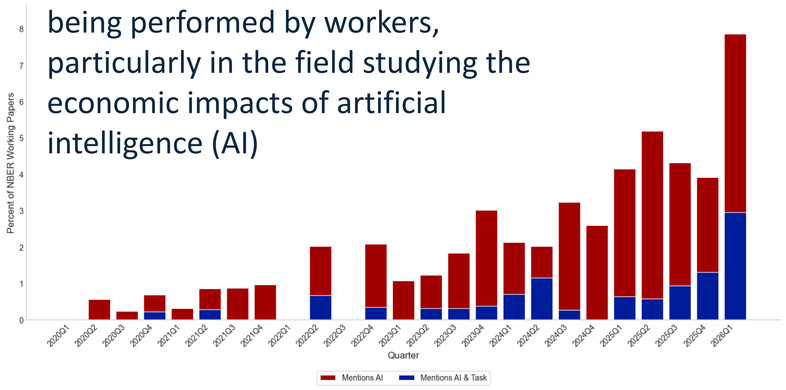
Classifying Job Tasks Using a Sentence-Transformer Model

Nicole Nestoriak & David H. Oh
U.S. Bureau of Labor Statistics

Motivation

A growing interest in the types of tasks being performed by workers, particularly in the field studying the economic impacts of artificial intelligence (AI)



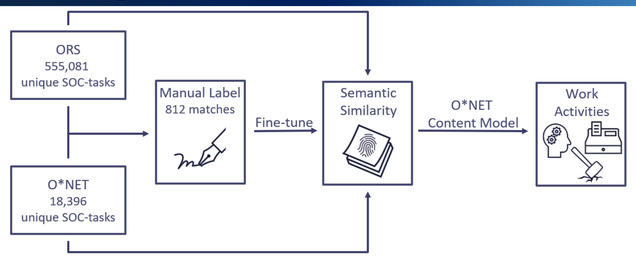
The Occupational Requirements Survey (ORS)

Captures text data describing critical tasks performed by a representative sample of jobs in the U.S. economy

Challenge

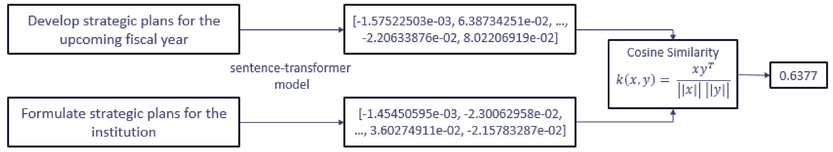
- ORS task data is collected as semi-structured text, not suitable for analysis in its raw form
- No resource to develop a classification system and to classify each task

Method Diagram



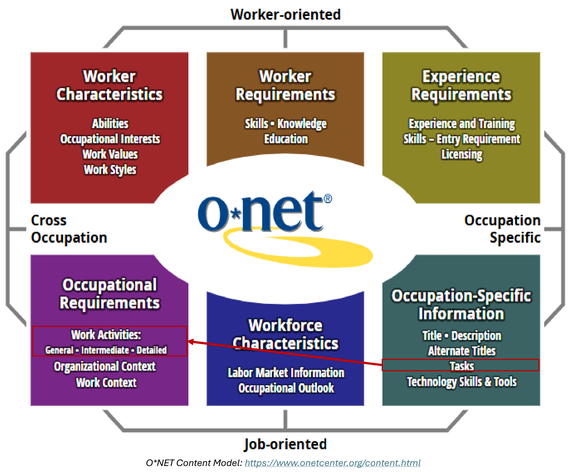
Semantic Similarity

- A fine-tuned sentence-transformer model (*all-MiniLM-L6-v2*) maps ORS and O*NET SOC titles and tasks to a 384-dimensional vector space
- For each vectorized ORS SOC-task, cosine similarity is measured against each vectorized O*NET SOC-task to find the most semantically similar match



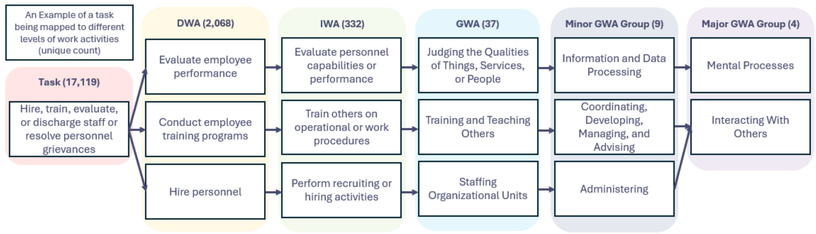
The Occupational Information Network (O*NET)

The Employment and Training Administration (ETA) surveys incumbents and occupational experts/analysts by providing them with a predetermined list of tasks for each occupation



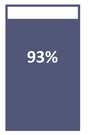
*O*NET Content Model:* https://www.onetcenter.org/content.html

O*NET tasks are mapped to one or more detailed work activity categories, which can be consolidated to a higher aggregate group: intermediate, generalized, etc.

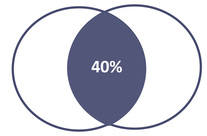


Validation of the Results

Qualitatively,



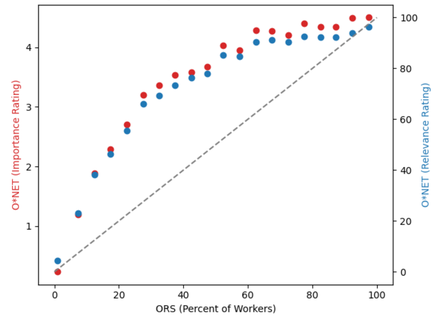
O*NET tasks matched with ORS tasks



Matched tasks have the same SOC

At the task level,

At the occupation level,



Moderate to high correlation when analyzing 17,791 SOC-GWA combinations in both ORS and O*NET

Application of the Results

Benchmark generative AI usage data

- Many studies reporting on the usage of generative AI at the occupation level do not account for the task variation within occupations
- ORS task data can help with that

Explore patterns of tasks associated with telework
(*Work from Home: A Task-based Perspective*, Nestoriak & Oh)

- Telework availability generally aligns with expected task patterns
- There is a meaningful task variation within occupations—especially in interactive tasks—among jobs with and without telework
- Previous telework research results are generalizable beyond specific firms and time periods