Representing Human Capabilities in Occupational Databases: An Embedding-Based Audit of O*NET, ESCO and Vocational Standards

Ilona Erika Negro
Doctoral Program of Applied Artificial Intelligence
Alma Mater Europaea, Maribor, Slovenia
e-mail: ilona.negro@almamater.si

***Abstract*—Labour-market analyses of automation and artificial intelligence frequently rely on occupational databases such as O*NET and ESCO to characterise the task content of work and model technological substitution or exposure. Yet less attention has been paid to how these data infrastructures represent different dimensions of human expertise. This study uses embedding-based semantic similarity analysis to conduct a representational audit of three widely used occupational taxonomies. Drawing on the Human-Centred Resilience Model (HCRM), the study examines four capability domains: cognitive–technical judgement, embodied–sensorimotor calibration, trust-calibrated interpersonal engagement, and situated adaptive decision-making. Behavioural anchor statements operationalising the four pillars are compared with descriptors from three occupational representation systems, namely O*NET Work Activities, ESCO skills, and national vocational training standards, across three occupations: electrician, dog groomer, and massage therapist. Consistently, O*NET descriptors show the strongest alignment with the capability pillars, followed by vocational standards and ESCO descriptors. A sensitivity analysis confirms that this ordering is robust to the choice of aggregation rule and embedding model. The adaptability pillar shows weaker alignment than cognitive and interpersonal capabilities. Qualitative inspection indicates a more abstract representation of adaptability rather than explicit descriptions of real-time behavioural adjustment. These findings reveal a systematic representational asymmetry in occupational taxonomies: situated adaptive decision-making is consistently encoded in a more compressed and abstract form than cognitive and interpersonal capabilities, suggesting that automation exposure estimates derived from such taxonomies may not fully capture the tacit and context-dependent dimensions of human expertise that are most difficult to replicate computationally.**

Keywords-occupational taxonomies; semantic embeddings; automation research; artificial intelligence; tacit knowledge.

I. Introduction

Artificial intelligence and automation research frequently relies on structured occupational taxonomies as an empirical foundation for estimating labour-market exposure and task automatability. Systems such as O*NET in the United States and ESCO in Europe provide standardised, machine-readable descriptors of occupational skills, abilities, and work contexts.

In previous work, the Human-Centred Resilience Model (HCRM) was developed to conceptualise dimensions of capabilities hypothesised to resist automation, particularly in embodied and relational service roles [1]. The model originally comprised three pillars:

1) Embodied Dexterity and Skilled Movement,
2) Agency and Adaptability, and
3) Emotional Labour.

These dimensions capture forms of expertise associated with tacit knowledge, embodied coordination, and interpersonal interaction that are often difficult to formalise computationally.

When the HCRM was empirically examined in subsequent work [2], the Adaptability pillar produced weaker differentiation than anticipated for occupations expected to rely heavily on situational flexibility and real-time adjustment.

This discrepancy prompted a methodological question. If occupational taxonomies serve as empirical inputs for automation modelling, it becomes important to understand whether they adequately represent the dimensions of work on which such models rely. The present study, therefore, asks: to what extent do occupational representation systems capture tacit, situational, and embodied forms of expertise?

Despite the widespread use of occupational taxonomies in automation modelling, relatively little research has examined how the internal representational structure of these datasets may shape the capability profiles derived from them. The present study addresses this question through a representational audit of three major sources of occupational description:

1) O*NET (United States) [3]
2) ESCO (European Skills, Competences, Qualifications and Occupations) [4]
3) National vocational standards from six European countries, used as comparatively detailed descriptions of occupational practice.

Recent advances in sentence embedding models allow the quantitative evaluation of semantic similarity between conceptual constructs and textual descriptors across large corpora. Using embedding-based semantic similarity analysis, the study evaluates the extent to which these representation systems align with theoretically defined human-centred capability pillars. Rather than redefining the core premises of the HCRM, this research refines its operationalisation for measurement purposes. Specifically, the original adaptability construct is decomposed into analytically distinguishable components that capture different aspects of situated action and decision-making.

The aim of the study is not to critique existing taxonomies normatively, but to quantify how different dimensions of human-centred work are represented within widely used occupational description systems. By examining where representational alignment is strong and where it is weak, the analysis contributes to a more systematic understanding of the data foundations on which empirical automation assessments are built. To the best of the author's knowledge, this study represents one of the first systematic attempts to evaluate occupational taxonomies through embedding-based representational auditing, assessing how conceptual capability constructs align with the descriptor structures of widely used labour-market datasets.

The remainder of this paper is organised as follows. Section II presents the state of the art in automation research and occupational taxonomy use. Section III introduces the theoretical foundations of the HCRM and the empirical anomaly that motivated the present study. Section IV situates the representational audit within the broader context of occupational taxonomy use in automation modelling and describes the three representation systems examined. Section V presents the four-pillar capability framework used as the basis for the analysis. Section VI describes the data sources and corpus preparation procedures. Section VII details the embedding-based methodology. Section VIII reports the results of the representational audit across occupations and capability pillars. Section IX discusses the findings in relation to tacit knowledge, automation modelling, and human-centred work. Section X presents the conclusions and directions for future research.

II. Related Work

Empirical studies estimating automation exposure frequently rely on occupational databases to approximate the task structure of work. For example, influential early estimates of computerisation probabilities trained machine-learning classifiers on O*NET task characteristics and "bottleneck variables" describing perceptual, creative, and social capabilities [5]. Subsequent studies have continued to rely on similar occupational representations when estimating the exposure of occupations to digital technologies or artificial intelligence [6] [7]. More recently, assessments of Large Language Model (LLM) exposure have used O*NET Detailed Work Activities to approximate the task structure of occupations [8]. Beyond these landmark studies, more recent work has examined the labour-market effects of artificial intelligence through task-displacement and task-reinstatement frameworks [9], and through direct assessment of the skills and activities most exposed to AI capabilities [10]. Across this body of work, O*NET remains the dominant empirical input for approximating occupational task structures in the United States, while ESCO fulfils a comparable role in European labour-market research [4].

However, occupational taxonomies are not neutral mirrors of real work; they are representational infrastructures designed to render labour markets legible to statistical and computational analysis by translating complex work practices into standardised descriptors. In order to standardise occupational descriptions across sectors and regions, these systems represent certain aspects of work with high granularity while compressing or simplifying others. Consistent with theories of routine-biased technological change [11] and longstanding observations about the difficulty of formalising tacit knowledge [12], occupational databases may tend to privilege formally articulated, context-stable cognitive activities. By contrast, work that depends on embodied coordination, adaptive judgement under uncertainty, or the relational dynamics of emotional labour may be only partially captured. These structural tendencies have been documented empirically: Handel [13] identified redundancy between importance and level scales, construct overlap, and substantial abstraction in item wording as specific measurement limitations of the O*NET content model — properties that are particularly consequential when fine distinctions between structured cognitive and tacit capabilities are required.

Because automation models frequently rely on these structured representations to approximate occupational capability profiles, any systematic representational imbalance may influence how automation potential is assessed. In particular, forms of expertise that are difficult to codify, such as tacit sensorimotor skill, situated decision-making, and relational coordination, may be underrepresented in the data structures used to model technological substitution. Whether and how these representational properties systematically vary across occupational databases, and what this means for automation modelling, has received limited direct empirical attention. The present study addresses this gap through an embedding-based representational audit of three widely used occupational description systems.

III. Background: The Human-Centred Resilience Model

The theoretical point of departure for the present representational audit is the HCRM, introduced in previous work examining the resilience of embodied service and craft occupations in the context of automation [1] [2]. The HCRM was developed to explain why some occupations, particularly those in the leisure, care, and skilled service economy, appear to retain value despite rapid advances in automation. Examples include wellness practitioners, pet care professionals, and skilled tradespeople, which share characteristics with the occupations analysed in the present study. These roles often combine physical skill, interpersonal interaction, and situational variability, characteristics that are difficult to formalise within conventional task-based automation frameworks.

Conceptually, the model draws on Self-Determination Theory [14] and Bourdieu's concepts of symbolic capital [15] and cultural distinction [16]. From the perspective of Self-Determination Theory, occupations that support autonomy, competence, and relatedness can generate intrinsically motivated forms of work for practitioners while also shaping demand, as clients often value services that provide meaningful interaction, personalised attention, and opportunities for human connection. At the same time, Bourdieu's concepts of symbolic capital and distinction help explain how scarcity, authenticity, and embodied expertise may increase the perceived value of certain forms of human work as automated services become more widespread.

In its original formulation, the HCRM proposed that the resilience of these occupations arises from the combined presence of three capability domains:

1) Embodied Dexterity and Skilled Movement
2) Agency and Adaptability
3) Emotional Labour

The key theoretical insight of the model was that resilience does not emerge from any single attribute in isolation, but rather from the interaction of these dimensions during real-world task execution, particularly in occupations that require close interaction with people, animals, or physical environments.

A. Empirical Anomaly and Measurement Challenge

Initial empirical testing of the HCRM using occupational data revealed a counterintuitive pattern. While the embodied skill and emotional labour dimensions broadly aligned with theoretical expectations, the Agency and Adaptability pillar did not clearly distinguish HCRM occupations from a broader comparison group [2]. At the aggregate level, the comparison group even produced slightly higher overall alignment scores on the adaptability dimension than the occupations expected to rely most strongly on situational flexibility.

This result was surprising because many leisure, care, and personal service roles rely on continuous, real-time judgement exercised in situ, including responding to non-verbal cues, adapting techniques to client preferences or animal behaviour, and operating in variable, often unpredictable environments. Such contexts require ongoing judgement and situational adjustment rather than reliance on predetermined task scripts.

This empirical pattern raised an important methodological question: whether the concept of adaptability was inadequately represented in the occupational taxonomies used for measurement. In particular, descriptors in databases such as O*NET and ESCO appeared to represent adaptability primarily through descriptors of cognitive planning, evaluation, and problem-solving, rather than through descriptions of real-time behavioural adjustment in dynamic environments.

B. Increasing Construct Resolution

To enable a more precise representational analysis of occupational datasets, this study refines the operationalisation of the HCRM. Rather than altering the theoretical foundations of the model, the original Agency and Adaptability pillar is analytically decomposed into two distinct components grounded in the literature on expertise, decision-making, and adaptive performance in complex environments [17] [18] [19].

This decomposition yields a four-pillar capability framework:

1) *Cognitive–Technical Judgement* — The application of structured knowledge, formal rules, and analytical reasoning in diagnosing problems and determining solutions.
2) *Embodied–Sensorimotor Calibration* — Skilled physical performance based on continuous integration of tactile, spatial, and tool-mediated feedback.
3) *Trust-Calibrated Interpersonal Engagement* — The capacity to establish and maintain cooperative, trust-based interactions with clients or stakeholders during service delivery.
4) *Situated Adaptive Decision-Making* — The ability to adjust actions dynamically in uncertain or evolving environments, selecting among multiple possible courses of action without relying on predetermined scripts.

These pillars define the capability constructs used to examine how occupational taxonomies represent different forms of expertise. Fig. 1 illustrates the conceptual structure of the HCRM and the four capability pillars derived from its three foundational domains: embodied dexterity and skilled movement, agency and adaptability, and emotional labour.

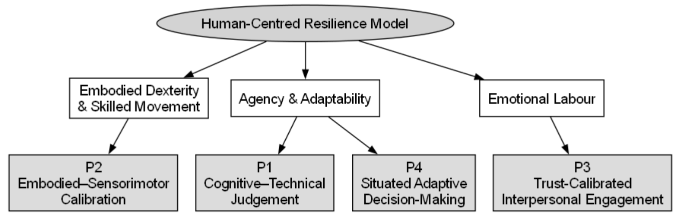


Figure 1. Human-Centred Resilience Model (HCRM) showing the decomposition of the original three-pillar structure into four capability domains.

IV. From HCRM to Representational Audit of Occupational Taxonomies

Automation modelling and algorithmic labour-market analysis rely heavily on occupational description systems that translate work into structured task, skill, and capability descriptors. Such approaches use formalised descriptions of occupational requirements as inputs to empirical automation models to estimate technological substitution or exposure [5] [6] [7] [8] [10]. The validity of these assessments, therefore, does not only depend on model design but also on the representational properties of the occupational data on which they rely.

A. Standard Taxonomies: O*NET and ESCO

The dominant occupational representation systems in this domain are O*NET in the United States and ESCO in the European Union. O*NET provides a large survey-based architecture covering knowledge, skills, abilities, work activities, and work contexts [3]. However, the database also exhibits several measurement challenges, including redundancy between importance and level scales, overlap across constructs, and substantial abstraction in item wording [13]. These properties are particularly relevant in the present context because the analysis focuses on fine distinctions between structured cognitive descriptors and more tacit, situated capabilities.

ESCO, by contrast, provides a multilingual ontology linking occupations to standardised skill and competence descriptors [4]. Its primary function is more to support interoperability among labour-market, education, and policy systems than survey-based measurement. Although its architecture differs from O*NET's, ESCO similarly relies on standardised, context-independent descriptors designed to make occupational characteristics machine-readable and comparable across labour-market systems.

When labour-market and automation studies use such databases as empirical inputs or feature spaces — as seen in influential models of computerisation probability [5] and recent assessments of Large Language Model exposure [8] — they also inherit the representational assumptions embedded in these systems. If an occupational taxonomy primarily represents work through abstract, standardised, or decontextualised descriptors, then downstream models may reflect those abstractions rather than the full situational complexity of occupational practice.

B. Comparative Reference Corpus: National Vocational Standards

To evaluate the representational coverage of standard taxonomies, this study introduces a third source of occupational description: national vocational standards from six countries: Austria, Belgium, France, Germany, Hungary and the Netherlands. Unlike statistical or policy taxonomies, vocational standards are training- and qualification-oriented documents developed to specify the competencies, procedures, safety requirements, and learning outcomes associated with professional practice. In competence-based vocational education systems, training regulations typically define occupational competences and performance requirements in terms of observable learning outcomes linked to workplace activities [20] [21] [22].

Vocational standards should not be treated as direct records of practice. However, because these documents are designed to guide training, assessment, and certification of practitioners, they often include detailed descriptions of procedures, materials, tools, and performance criteria that reflect how work is expected to be carried out in practice.

C. The Representational Audit

The present study, therefore, conducts a representational audit comparing O*NET, ESCO, and national vocational standards across three service and craft occupations.

The audit covers three occupations selected because they rely heavily on embodied skill, interpersonal coordination, and situational variability: electrician, dog groomer, and massage therapist. These occupations are described in detail in Section VI.

V. Pillar Framework for Representational Analysis

To operationalise the capabilities associated with human-centred service and craft work, the HCRM is translated here into four analytically distinct but interrelated capability domains. These pillars capture different forms of expertise through which skilled practitioners engage with technical systems, physical environments, and other people in the course of work. While occupational taxonomies tend to privilege structured cognitive tasks and discrete work activities, research on tacit knowledge, expertise, embodied cognition, and situated action suggests that many forms of skilled performance emerge through the dynamic integration of perception, action, and social interaction in context [12] [18] [19] [23] [24]. This perspective is also reflected in Rasmussen's Skill–Rule–Knowledge framework, which distinguishes between embodied skill-based behaviour, rule-guided action, and knowledge-based reasoning in complex work environments [17]. These traditions collectively emphasise that skilled performance cannot be fully understood as the execution of isolated tasks; it emerges from the coordinated interaction of cognitive judgement, embodied action, and social engagement within dynamic contexts.

The four pillars, therefore, represent complementary dimensions of human capability: cognitive–technical judgement, embodied–sensorimotor calibration, trust-calibrated interpersonal engagement, and situated adaptive decision-making. Table I summarises the four capability pillars used in the representational analysis and the primary forms of expertise associated with each domain.

TABLE I. Capability Pillars Used in the Representational Analysis

Pillar	*Capability Domain*	*Core Capability Mechanisms*
P1	Cognitive–Technical Judgement	Diagnosis of technical problems, interpretation of standards, rule-based and knowledge-based reasoning
P2	Embodied–Sensorimotor Calibration	Skilled bodily movement, tactile feedback, coordination of tools, materials, and physical environments
P3	Trust-Calibrated Interpersonal Engagement	Interpretation of verbal and non-verbal cues, emotional regulation, and maintenance of cooperative interaction
P4	Situated Adaptive Decision-Making	Real-time adjustment to evolving conditions, selection among action paths, integration of situational cues

A. Cognitive–Technical Judgement

Cognitive–Technical Judgement refers to the capacity to understand technical systems, diagnose faults, interpret standards, and apply structured domain knowledge under operational constraints. This pillar captures forms of rule-based and knowledge-based reasoning commonly required in technical occupations, where practitioners must interpret signals, diagnose problems, and apply established procedures or technical standards [17].

In labour-market research, such structured forms of problem-solving align closely with task-based approaches to work, which conceptualise occupations as bundles of tasks that require different combinations of cognitive and manual capabilities [25]. Within this framework, technological change has been shown to substitute for tasks that can be expressed through explicit procedural rules while complementing more complex problem-solving activities [11].

This pillar, therefore, corresponds closely to the kinds of structured cognitive activities typically captured in occupational taxonomies such as the O*NET content model, including analysing information, interpreting technical documentation, and determining solutions to technical problems [3].

B. Embodied–Sensorimotor Calibration

Embodied–Sensorimotor Calibration refers to the capacity to perform skilled physical actions through continuous sensorimotor adjustment, integrating tactile, spatial, and tool-mediated feedback in real time. Skilled manual activity requires the dynamic coordination of perception and movement, allowing practitioners to adapt bodily actions to the constraints imposed by tools, materials, and physical environments. Research in ecological psychology emphasises that such behaviour emerges from direct interaction with environmental affordances rather than from explicit cognitive representations [26].

Similarly, theories of embodied cognition highlight that many forms of practical intelligence arise from the coupling of perception and action within situated bodily experience [27]. In skilled work, this coupling enables practitioners to calibrate force, position, timing, and movement patterns continuously while responding to subtle environmental cues. Classical research in motor control further demonstrates that complex skilled movements are not executed through rigid programmes but are dynamically organised through the coordination of multiple degrees of freedom within the body–environment system [28].

Taken together, these perspectives describe embodied skill as adaptive sensorimotor regulation in which perception, movement, and environmental feedback are tightly integrated during task execution. Such capabilities are central to many craft and service occupations, although they are often described only indirectly in occupational taxonomies, which tend to emphasise tasks and outcomes rather than the underlying sensorimotor processes through which skilled performance is achieved.

C. Trust-Calibrated Interpersonal Engagement

Trust-Calibrated Interpersonal Engagement (P3) refers to the capacity to establish and sustain cooperative, trust-based interactions with clients, colleagues, or other stakeholders in physically co-present environments. This dimension includes the ability to interpret verbal and non-verbal cues, regulate emotional expression, and adjust communicative behaviour in response to the expectations and affective states of others.

Research on emotional labour highlights that many service occupations require workers to actively manage and coordinate emotional displays to produce appropriate interpersonal experiences for clients [29]. In such contexts, effective interaction depends not only on technical competence but also on the ability to sustain socially meaningful encounters through subtle forms of relational coordination and affective regulation.

Sociological studies of face-to-face interaction further emphasise that social encounters rely on continuous mutual monitoring and the maintenance of interactional trust between participants [30]. Workers in client-facing occupations must therefore continuously adjust their communicative behaviour, tone, and attentional focus in order to sustain cooperative engagement.

From a labour-market perspective, these interpersonal capabilities form part of a broader category of social skills that have become increasingly important in modern economies and are associated with occupations involving collaboration, service provision, and client interaction [31]. Such skills rely heavily on situational awareness and tacit interpersonal knowledge, making them difficult to fully capture in standardised occupational descriptors that typically emphasise discrete tasks or communication activities.

D. Situated Adaptive Decision-Making

Situated Adaptive Decision-Making refers to the capacity to operate effectively in complex and uncertain environments without predetermined scripts. It involves selecting among multiple possible courses of action, revising strategies dynamically, and integrating technical knowledge with situational judgement in real time. In many forms of skilled practice, effective performance does not depend solely on the application of explicit rules but also on the ability to interpret evolving environmental cues and adjust behaviour accordingly.

Research in cognitive systems engineering emphasises that practitioners working in complex environments must continuously coordinate multiple interacting constraints while adapting their actions in response to changing conditions [17] [19]. Studies of situated action further demonstrate that real-world activity frequently unfolds through ongoing interaction with the environment rather than through the execution of pre-specified plans [24]. Similarly, research on naturalistic decision making and recognition-primed decision-making shows that experienced practitioners often rely on recognition-based judgement, rapidly identifying patterns in unfolding situations and selecting feasible courses of action without explicitly comparing alternatives [18]. Analyses of expertise further emphasise that such adaptive behaviour relies heavily on tacit, experience-based knowledge that cannot be fully articulated through formal rules or procedures [12] [23].

To operationalise this pillar for empirical analysis, adaptability is represented through behavioural anchor statements describing real-time situational adjustment in dynamic environments, derived from the pillar's conceptual definitions. Examples of the behavioural anchors are shown in Table II.

VI. Data Sources

This section describes the three occupational representation systems used in the analysis and the procedures applied to prepare each corpus for embedding-based comparison.

A. O*NET

Occupational descriptors were drawn from the O*NET Database 30.2 [3] Work Activities taxonomy for the selected occupations, with the analysis focusing on the Work Activities component. Each Work Activities descriptor consists of a short activity statement describing a generic work behaviour (e. g., "Monitoring and reviewing information from materials, events, or the environment, to detect or assess problems."). O*NET provides structured occupational descriptions derived from surveys of job incumbents and occupational experts.

In this study, the analysis focused on the 41 descriptors contained in the O*NET Work Activities taxonomy. The taxonomy consists of standardised activity categories applied uniformly across occupations. Because these descriptors are identical across occupations, the textual corpus used in the embedding analysis remains constant for all three cases.

The importance and level scores associated with the Work Activities taxonomy were not incorporated into the similarity analysis, as the objective was to evaluate the semantic representational coverage of capability constructs, not the relative weighting of activities within occupations.

TABLE II. Example Behavioural Anchor Statements Used to Operationalise the HCRM Capability Pillars

Pillar	*Behaviour Anchors*
Cognitive–Technical Judgement (P1)	Diagnoses technical problems using structured reasoning and technical knowledge. ● Interprets technical documentation, diagrams, or operational instructions. ● Applies established procedures and standards to solve technical problems. ● Evaluates system behaviour to determine appropriate corrective actions. ● Analyses technical information to identify faults or system malfunctions. ● Uses domain knowledge to determine appropriate technical solutions. ● Interprets signals, measurements, or system indicators to guide decision-making. ● Applies formal rules or procedures to complete technical tasks.
Embodied–Sensorimotor Calibration (P2)	Adjusts hand movements in response to tactile or spatial feedback during manual tasks. ● Coordinates bodily movement and tool use to perform precise physical actions. ● Continuously adjusts force, position, and timing during manual work. ● Uses tactile and spatial feedback to guide tool manipulation. ● Calibrates bodily actions while interacting with tools or materials. ● Performs skilled manual actions requiring fine motor control. ● Adjusts physical technique based on subtle environmental cues. ● Maintains precise bodily coordination while performing physical tasks.
Trust-Calibrated Interpersonal Engagement (P3)	Establishes trust through clear and respectful communication with clients. ● Interprets verbal and non-verbal cues during interpersonal interaction. ● Adjusts communication style in response to the emotional state of others. ● Maintains cooperative relationships with clients or colleagues. ● Provides reassurance or explanation to reduce uncertainty for clients. ● Listens carefully and responds appropriately to client concerns. ● Maintains calm and professional interaction in emotionally charged situations. ● Coordinates work activities through interpersonal communication.
Situated Adaptive Decision-Making (P4)	Adjusts actions in real time in response to changing environmental conditions. ● Selects among multiple possible courses of action based on emerging cues. ● Revises task strategies when routine procedures prove insufficient. ● Adapts behaviour continuously while interacting with people, animals, or physical environments. ● Responds flexibly to unexpected events during task execution. ● Integrates technical knowledge with situational judgement in real time. ● Modifies task execution as conditions evolve. ● Navigates uncertain environments without relying on predetermined scripts.

B. ESCO

The ESCO skill and competence descriptors used for each occupation were extracted from ESCO v1.2.1 (accessed 8 Feb. 2026) [4]. Individual occupation concept URIs and the full extracted descriptor sets are provided in the supplementary repository: https://github.com/Ilona-woody/Hcrm-representational-audit/tree/main/esco. ESCO provides a multilingual classification of occupations, skills, competences, and qualifications designed to support labour-market interoperability across European policy and employment systems.

For each selected occupation, the associated skill and competence descriptors listed in the ESCO ontology were extracted and treated as individual textual descriptors for embedding comparison.

C. National Vocational Standards

National vocational training and qualification documents for the selected occupations were collected from official national or regional sources in Austria, Belgium, France, Germany, Hungary and the Netherlands. These documents included qualification frameworks, curricula, competency standards, and training specifications describing the competencies, performance criteria, and learning outcomes associated with professional practice. Documents were chosen that provided formal competency specifications for the selected occupations within national qualification frameworks or accredited training programmes. The electrician vocational standards corpus contains substantially more descriptors than the other corpora, reflecting the higher level of procedural specification in electrical installation standards.

Descriptor counts, therefore, differ across representation systems because documentation practices differ, not because the occupations differ in underlying capabilities.

To ensure comparability with O*NET and ESCO, only content-bearing competency and task descriptions were retained. Administrative, legal, and procedural sections that did not describe occupational capability requirements were excluded. For longer documents, only sections containing learning outcomes, competence statements, performance criteria, or task-relevant descriptions were extracted.

Each extracted segment was recorded with provenance metadata, including issuing authority, year, regulatory level, source URL, and language. Each extracted segment was treated as a potential descriptor unit for subsequent semantic analysis. Non-English source documents were translated into English using a consistent machine-translation workflow (DeepL) [32] [33] [34] [35]. The English translations were then used for embedding-based comparison to enable cross-language analysis using an English sentence-transformer model. Full provenance metadata for all vocational standard documents used in the analysis, including issuing authority, year, regulatory level, source URL, and language, is provided in the supplementary repository: https://github.com/Ilona-woody/Hcrm-representational-audit/tree/main/data.

Table III summarises the occupational representation systems examined in the analysis, the type of data each provides, and the textual unit used as the basis for embedding comparison.

TABLE III. Occupational Representation Systems and Units of Analysis

Representation System	*Data Type*	*Textual Unit Used in Analysis*
O*NET Work Activities	structured task descriptors	activity statement
ESCO Skills	skills taxonomy	skill labels
Vocational Standards	vocational training standards	competency statements

D. Corpus Preparation

Each corpus was segmented into descriptor-level textual units to enable embedding-based comparison. Where documents contained longer narrative passages, text was first divided into sentence-level segments using punctuation and line-break boundaries so that similarity scoring could later be performed at the segment level while retaining the descriptor as the reporting unit.

VII. Method

The empirical analysis implements an embedding-based representational audit of occupational databases. Descriptor texts from three representation systems, O*NET Work Activities, ESCO skills, and vocational standards (Regulatory and Professional Documentation Corpus, RPDC), are embedded using a transformer-based sentence embedding model. Behavioural anchor statements operationalising the four human-centred capability pillars are embedded in the same semantic space. Representational alignment is then quantified using cosine similarity between anchors and descriptor segments, with descriptor-level scores defined by the highest observed semantic correspondence. This procedure enables systematic comparison of how strongly different occupational representation systems encode the capability constructs defined by the HCRM. In this sense, the analysis evaluates the extent to which occupational data infrastructures make human-centred capabilities computationally visible.

A. Representational Approach

Embedding-based semantic similarity analysis provides a method for examining how conceptual constructs are represented within textual corpora by mapping linguistic content into high-dimensional vector spaces. In vector-space models of language, words and sentences are represented as numerical embeddings, with their relative positions capturing semantic relationships between expressions [36]. Transformer-based language models, such as BERT, generate contextual embeddings that encode linguistic meaning based on the surrounding context [37]. Sentence-level models, such as Sentence-BERT (SBERT), extend this approach by producing embeddings optimised for semantic similarity comparisons between short texts [38]. In such embedding spaces, semantically related expressions occupy nearby regions. Embedding-based techniques have increasingly been applied in computational social science to analyse conceptual relationships within textual corpora [39].

This study applies embedding-based semantic similarity analysis to evaluate how occupational representation systems encode the four theoretically defined pillars of human-centred capability introduced in Section V. Sentence embeddings were generated using the pretrained Sentence-Transformers model all-MiniLM-L6-v2, a compact transformer architecture designed for efficient semantic similarity and sentence-embedding tasks [40]. The model produces 384-dimensional embeddings optimised for semantic comparison.

Each representation corpus, comprising O*NET Work Activities, ESCO skills, and vocational standards drawn from national training documents (hereafter the Regulatory and Professional Documentation Corpus, RPDC), was converted into descriptor-level textual units. Each pillar was operationalised using a set of behavioural anchor sentences derived from the conceptual definitions of the four capability pillars, as listed in Table II. The complete anchor sets are provided in the supplementary repository: https://github.com/Ilona-woody/Hcrm-representational-audit.

Semantic alignment between pillars and representation systems was estimated using cosine similarity between pillar anchor embeddings and descriptor embeddings. Each pillar is represented by multiple anchor statements. Descriptor-level pillar alignment is operationalised as the maximum cosine similarity between any descriptor segment and any anchor statement within that pillar, capturing the strongest semantic correspondence within the descriptor. Where descriptors contained longer narrative passages, they were segmented into sentence-like units using punctuation and line-break boundaries. Sentence embeddings were computed for these segments, and the maximum similarity value across segments was retained for the original descriptor. This procedure reduces semantic dilution in longer descriptions that contain multiple behavioural components.

Retaining the maximum similarity constitutes a presence-sensitive detection strategy. The objective of the audit is to determine whether construct-relevant capability content appears anywhere within a descriptor, even when embedded within broader narrative descriptions. Maximum similarity, therefore, provides a conservative test of representational coverage: if a representation system fails to produce strong similarity even under this most permissive detection rule, this indicates that the construct is likely compressed, abstracted, or omitted within the representation.

For each pillar $\times$ representation system $\times$ occupation combination, the resulting distribution of descriptor-level similarity scores was summarised using mean, median, 10th percentile (p10), and 90th percentile (p90) statistics. Cosine similarity ranges from -1 to 1. Negative values indicate semantic dissimilarity between descriptor and anchor embeddings and are expected for descriptors unrelated to a given capability domain. Standard deviation and range statistics were also computed as robustness checks and are reported in Table V. Mean similarity is used as the primary summary indicator of overall representational density, while median and percentile statistics are reported to assess distributional robustness and the presence of high-alignment descriptors within each system.

ESCO descriptors frequently consist of short skill labels: across the three occupational corpora examined, the median descriptor length was 3–4 words for ESCO, compared with 14 words for O*NET Work Activities and 5–11 words for RPDC vocational standards. Such short phrases contain substantially less contextual information than longer task descriptions and may therefore produce lower semantic similarity scores independently of construct alignment. This compression of contextual information is itself a representational feature of ontology-based taxonomies and may systematically limit the extent to which tacit, situational, or embodied capabilities can be expressed within such systems. The script used to compute descriptor length statistics is provided in the supplementary repository: https://github.com/Ilona-woody/Hcrm-representational-audit/tree/main/scripts.

Because the O*NET descriptor set is identical across occupations, O*NET similarity distributions remain constant across cases; observed cross-occupation variation therefore arises primarily from the RPDC and ESCO corpora.

B. Preprocessing and Unit of Analysis

Occupational representation systems differ substantially in textual granularity. ESCO descriptors typically consist of short skill labels, O*NET Work Activities are structured task descriptions, and vocational standards often contain longer narrative passages describing competencies, training requirements, or performance criteria. To ensure comparability across sources, all corpora were transformed into descriptor-level analysis units prior to embedding.

Text was cleaned through whitespace normalisation and the removal of empty or near-empty strings. O*NET Work Activities and ESCO skills were retained as individual descriptors reflecting their structured format. Vocational standard descriptors were segmented as described in Section VI. Extremely short fragments lacking meaningful lexical content were excluded. For reporting purposes, similarity scores are recorded at the descriptor level; however, for segmented descriptors, the reported value reflects the highest-scoring sentence-level segment contained within that descriptor.

Vocational standards were collected from publicly available national and regional training and qualification documents across six countries and translated into English. Translation was necessary to enable cross-corpus comparison within a single embedding space, though minor linguistic nuances in the original documents may be attenuated in translation. Only sections describing occupational tasks, competencies, or behavioural expectations were retained, while administrative and regulatory sections were excluded.

The resulting descriptor sets served as the unit of analysis for similarity computation. Descriptor counts, therefore, vary across representation systems, reflecting differences in document structure and descriptive granularity.

C. Anchor Construction and Use

Each pillar was represented by an anchor set comprising short behavioural statements designed to capture the conceptual content of the pillar (for example, real-time situated adaptive decision-making in dynamic environments for Pillar 4). Multiple anchor statements were used for each pillar to capture different behavioural expressions of the construct and reduce dependence on any single phrasing. The same anchor sets were applied across occupations and representation systems to ensure comparability.

Anchors were finalised prior to running the cross-occupation comparisons and were not tuned to individual corpora, ensuring that similarity scores reflect representational alignment rather than optimisation artefacts. Table II provides examples of the anchor statements used for each pillar. The complete anchor sets for all four pillars are provided in the supplementary repository: https://github.com/Ilona-woody/Hcrm-representational-audit.

D. Analytical Objective and Interpretation

The objective of the analysis is to assess representational coverage: the extent to which a representation system semantically encodes the capability constructs defined by the four pillars.

Higher similarity scores indicate stronger semantic correspondence between pillar anchors and the descriptors contained in a representation system. Lower similarity scores may indicate abstraction, compression, or omission of construct-relevant content.

Because embedding similarity reflects semantic proximity rather than causal importance, the results are interpreted as indicators of representational emphasis and descriptive granularity. The analysis is descriptive–comparative, not inferential, focusing on systematic differences in representational coverage through occupational databases. The quantitative similarity analysis is complemented by qualitative inspection of high-similarity descriptor matches retrieved from each representation system.

E. Robustness and Sensitivity Analysis

To assess the robustness of the main findings, two sensitivity analyses were conducted. The first varied the aggregation rule: in addition to the maximum cosine similarity used in the primary analysis, mean similarity across all segment-anchor pairs was computed for each descriptor. The second varied the embedding model: a parallel analysis was conducted using the larger all-mpnet-base-v2 model under maximum similarity aggregation, alongside the primary all-MiniLM-L6-v2 model.

The core finding, that O*NET descriptors show the strongest semantic alignment across all pillar–occupation combinations, proved fully stable across all conditions tested: O*NET remained the highest-scoring system in all 36 pillar–occupation–condition combinations examined. The secondary ordering of RPDC above ESCO showed minor variation only in the dog groomer corpus under mean aggregation, consistent with the two exceptions already noted in the main analysis and attributable to the small margin of difference in that corpus. These results indicate that the reported representational ordering is not an artefact of the specific aggregation rule or embedding model selected.

Table IV summarises the results across all three conditions.

VIII. Results

The results are presented in five steps. First, Section VIII.A compares how the three occupational representation systems (RPDC, O*NET, and ESCO) align with the four capability pillars across the selected occupations. Second, Section VIII.B examines pillar-level representation patterns, analysing how each capability dimension is captured by the systems. Third, Section VIII.C focuses on Situated Adaptive Decision-Making (Pillar 4), which exhibits a distinct cross-system representation

TABLE IV. SENSITIVITY ANALYSIS: STABILITY OF REPRESENTATIONAL ORDERING ACROSS AGGREGATION RULES AND EMBEDDING MODELS

Condition	*Model*	*Aggregation*	*O*NET highest (n/12)*	*Full ordering holds (n/12)*
Primary analysis	all-MiniLM-L6-v2	Maximum	12/12	10/12
Sensitivity 1	all-MiniLM-L6-v2	Mean	12/12	9/12
Sensitivity 2	all-mpnet-base-v2	Maximum	12/12	11/12

Note. O*NET highest = O*NET produces the highest mean similarity score across all 12 pillar–occupation combinations. Full ordering holds = O*NET > RPDC > ESCO across all 12 combinations. Deviations occur exclusively in the dog groomer corpus, consistent with exceptions noted in the main analysis. $n = 12$ (4 pillars $\times$ 3 occupations).

profile and therefore requires more detailed examination. Fourth, Section VIII.D presents a qualitative inspection of high-similarity descriptors for Pillar 4, providing interpretive validation of the quantitative patterns. Fifth, Section VIII.E summarises the representational patterns observed across occupations and representation systems. Finally, Section VIII.F reports a robustness and sensitivity analysis examining the stability of the main findings across alternative aggregation rules and embedding models.

O*NET similarity distributions are identical among occupations because the same set of 41 O*NET Work Activity taxonomy descriptors was used for all analyses. Consequently, cross-occupation variation is driven by differences in the occupation-specific corpora (vocational standards and ESCO skills), while the O*NET alignment profile remains constant. The electrician vocational standards corpus contains substantially more descriptors than the other corpora, reflecting the higher level of procedural specification in electrical installation standards.

Fig. 2 shows the mean cosine similarity scores across pillars and representation systems.

Table V summarises the similarity statistics for each pillar–representation system combination for the three occupations. The table reports mean similarity values, the number of descriptors analysed in each corpus, and distributional statistics, including median, standard deviation, and percentile values. Across representation systems, similarity distributions showed moderate dispersion (standard deviation typically between approximately 0.07 and 0.12), indicating that the observed patterns are not driven by a small number of extreme descriptors.

Across occupations and representation systems, several consistent patterns emerge. First, descriptors associated with Trust-Calibrated Interpersonal Engagement (Pillar 3) show strong semantic alignment, particularly within the O*NET activity taxonomy. Second, Cognitive–Technical Judgement (P1) also exhibits relatively high alignment across systems, reflecting the traditional emphasis of occupational classifications on procedural and problem-solving tasks. Third, Embodied–Sensorimotor Calibration (P2) is moderately represented by the systems and is relatively more visible in RPDC, where procedural and bodily techniques are described in operational detail. Finally, Situated Adaptive Decision-Making (P4) shows a more uneven

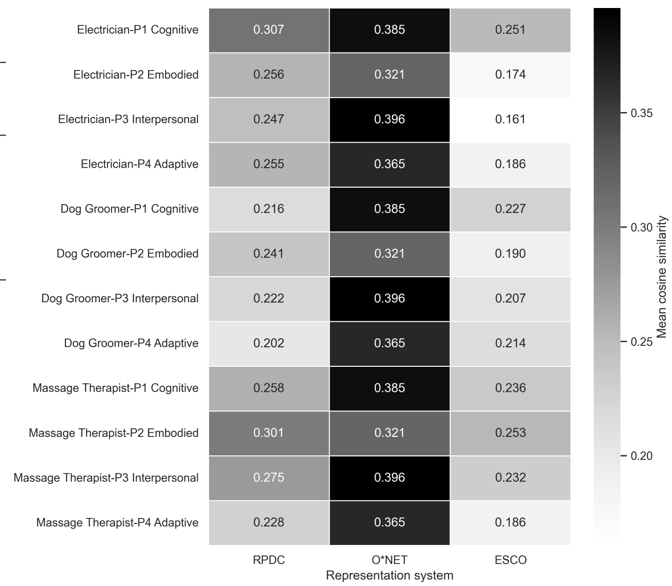


Figure 2. Heatmap of mean cosine similarity scores across pillars and representation systems.

representational profile across systems. While moderately represented in O*NET activities, it exhibits comparatively lower alignment in RPDC and ESCO descriptors. Overall, these patterns suggest that existing occupational taxonomies capture codified tasks, interpersonal interaction, and some forms of procedural embodiment more readily than context-sensitive adaptive behaviour.

The following subsections examine these patterns in greater detail, beginning with cross-system comparisons and then turning to pillar-specific differences.

A. Cross-System Representational Alignment

To evaluate how different occupational representation systems represent the four conceptual pillars, cosine similarity was computed between pillar anchor embeddings and descriptor embeddings extracted from three sources: vocational standards (RPDC), O*NET Work Activities, and ESCO skills. The analysis was conducted for three occupations representing human-centred service and craft work: electrician, dog groomer, and massage therapist.

Across the analysed occupation–pillar combinations, a consistent ordering of representational alignment is observed: O*NET > RPDC > ESCO. In most cases, O*NET descriptors exhibit the highest semantic similarity to the pillar anchors, vocational standards (RPDC) show intermediate alignment, and ESCO descriptors produce lower similarity scores (see Fig. 3a, 3b and 3c). This ordering holds for 10 of the 12 pillar–occupation combinations analysed. Two minor deviations occur in the dog groomer corpus: for Cognitive–Technical Judgement (P1), ESCO (0.227) slightly exceeds RPDC (0.216), and for Situated Adaptive Decision-Making (P4), ESCO (0.214) is marginally higher than RPDC (0.202). These differences are small in magnitude and do not alter the overall cross-system pattern.

TABLE V. SIMILARITY STATISTICS FOR PILLAR–SYSTEM COMPARISONS ACROSS OCCUPATIONS

Occupation	*Pillar*	*System*	*n*	*Mean*	*Median*	*Std*	*Min*	*Max*	*p10 / p90*
Dog Groomer	P1	RPDC	161	0.216	0.209	0.106	0.004	0.502	0.072 / 0.372
	P1	O*NET	41	0.385	0.374	0.090	0.219	0.615	0.278 / 0.491
	P1	ESCO	46	0.227	0.218	0.095	0.051	0.457	0.114 / 0.358
	P2	RPDC	161	0.241	0.233	0.097	0.034	0.519	0.118 / 0.372
	P2	O*NET	41	0.321	0.313	0.106	0.153	0.650	0.192 / 0.429
	P2	ESCO	46	0.190	0.184	0.096	0.067	0.518	0.079 / 0.287
	P3	RPDC	161	0.222	0.202	0.104	0.027	0.518	0.106 / 0.370
	P3	O*NET	41	0.396	0.385	0.078	0.247	0.624	0.309 / 0.494
	P3	ESCO	46	0.207	0.173	0.096	0.070	0.506	0.122 / 0.359
	P4	RPDC	161	0.202	0.193	0.090	0.038	0.450	0.089 / 0.341
	P4	O*NET	41	0.365	0.357	0.084	0.235	0.527	0.255 / 0.482
	P4	ESCO	46	0.214	0.202	0.085	0.076	0.459	0.109 / 0.322
Electrician	P1	RPDC	2341	0.307	0.299	0.122	0.001	0.743	0.150 / 0.467
	P1	O*NET	41	0.385	0.374	0.090	0.219	0.615	0.278 / 0.491
	P1	ESCO	41	0.251	0.242	0.106	0.093	0.554	0.125 / 0.390
	P2	RPDC	2341	0.256	0.249	0.098	0.002	0.734	0.135 / 0.380
	P2	O*NET	41	0.321	0.313	0.106	0.153	0.650	0.192 / 0.429
	P2	ESCO	41	0.174	0.166	0.111	0.000	0.540	0.057 / 0.259
	P3	RPDC	2341	0.247	0.245	0.109	−0.003	0.661	0.104 / 0.391
	P3	O*NET	41	0.396	0.385	0.078	0.247	0.624	0.309 / 0.494
	P3	ESCO	41	0.161	0.148	0.082	0.011	0.330	0.063 / 0.264
	P4	RPDC	2341	0.255	0.254	0.100	−0.013	0.634	0.127 / 0.387
	P4	O*NET	41	0.365	0.357	0.084	0.235	0.527	0.255 / 0.482
	P4	ESCO	41	0.186	0.168	0.107	−0.016	0.443	0.076 / 0.329
Massage Therapist	P1	RPDC	180	0.258	0.240	0.120	0.058	0.683	0.123 / 0.412
	P1	O*NET	41	0.385	0.374	0.090	0.219	0.615	0.278 / 0.491
	P1	ESCO	53	0.236	0.245	0.098	0.072	0.480	0.113 / 0.367
	P2	RPDC	180	0.301	0.293	0.090	0.111	0.557	0.192 / 0.425
	P2	O*NET	41	0.321	0.313	0.106	0.153	0.650	0.192 / 0.429
	P2	ESCO	53	0.253	0.229	0.092	0.081	0.493	0.157 / 0.390
	P3	RPDC	180	0.275	0.251	0.123	0.008	0.637	0.135 / 0.448
	P3	O*NET	41	0.396	0.385	0.078	0.247	0.624	0.309 / 0.494
	P3	ESCO	53	0.232	0.201	0.106	0.048	0.506	0.104 / 0.372
	P4	RPDC	180	0.228	0.216	0.097	0.057	0.528	0.110 / 0.370
	P4	O*NET	41	0.365	0.357	0.084	0.235	0.527	0.255 / 0.482
	P4	ESCO	53	0.186	0.187	0.073	0.060	0.359	0.092 / 0.290

For example, in the electrician corpus, mean similarity values for Pillar 1 (Cognitive–Technical Judgement) were 0.384 for O*NET, 0.307 for RPDC, and 0.251 for ESCO. The same ordering is observed for the remaining pillars and across the other two occupations.

At the same time, the similarity distributions vary systematically by pillar. Trust-Calibrated Interpersonal Engagement (P3) generally achieves the highest alignment, especially within O*NET, while Situated Adaptive Decision-Making (P4) tends to be among the weaker represented dimensions in RPDC and ESCO. In O*NET, however, P4 is not the weakest pillar; it scores above Embodied–Sensorimotor Calibration (P2). This indicates that existing occupational taxonomies do not underrepresent all non-cognitive capabilities equally; rather, they vary in how they represent the interpersonal, embodied, and adaptive aspects of work (see Table V). Spanning systems, the pillars therefore indicate distinct representational profiles: cognitive–technical judgement and interpersonal engagement exhibit relatively strong alignment, embodied calibration shows moderate alignment, and situated adaptive decision-making displays the most variable representation. Fig. 3a, 3b and 3c visualise these differences by showing the mean similarity values for each pillar for the three representation systems for the selected occupations.

B. Pillar-Level Representation Patterns

The following subsections examine each capability pillar individually in order to identify systematic differences in how the dimensions are represented in occupational description systems.

1) Cognitive–Technical Judgement (P1): Cognitive–Technical Judgement was relatively well represented across all systems. O*NET produced the highest similarity values for this pillar in all three occupations, with a mean similarity of approximately 0.385 in each case. RPDC corpora also captured this dimension to a substantial degree, particularly for electrician (0.307) and massage therapist (0.258), indicating the presence of analytical and technical reasoning elements in vocational descriptions. ESCO descriptors showed weaker but still measurable alignment, ranging from 0.216 for dog groomer to 0.251 for electrician.

Overall, the results indicate that structured cognitive and technical judgement is comparatively well represented in the language used by existing occupational description systems.

2) Embodied–Sensorimotor Calibration (P2): Embodied–Sensorimotor Calibration showed moderate representation by all sources, with greater variation across occupations than observed for Pillar 1. O*NET produced the highest mean similarity values overall for this pillar (approximately 0.321 across occupations), although within the O*NET profile, this was the lowest-scoring pillar relative to the other three. This indicates that embodied skill is represented within O*NET, but less strongly than cognitive, interpersonal, and adaptability-related descriptors.

Vocational standards came close in occupations involving more explicit bodily technique, particularly for massage therapist, where RPDC achieved 0.301, compared with 0.321 for O*NET. For the electrician and dog groomer, RPDC values were lower (0.256 and 0.241, respectively), but remained above or near ESCO. ESCO showed the weakest representation of embodied skill overall, although the strength of this pillar varied by occupation, from 0.174 for electrician to 0.253 for massage therapist. These results suggest that embodied competence is represented across all three systems, but more strongly where corpora contain explicit references to manual technique, dexterity, bodily coordination, or physical handling.

3) Trust-Calibrated Interpersonal Engagement (P3): Trust-Calibrated Interpersonal Engagement showed the strongest representation in O*NET in all occupations. Mean similarity values were approximately 0.396 for electrician, dog groomer, and massage therapist (reflecting the identical O*NET descriptor set across occupations), making Pillar 3 the highest-scoring O*NET pillar in all three cases. RPDC also captured this dimension, particularly for massage therapist (0.275), but less strongly for electrician (0.247) and dog groomer (0.222). ESCO descriptors showed weaker overall alignment, ranging from 0.162 for electrician to 0.232 for massage therapist. These results indicate that interpersonal coordination, communication, and related social interaction descriptors are represented relatively strongly in O*NET and, to a lesser extent, in vocational standards.

C. Representation of Situated Adaptive Decision-Making (P4)

Because adaptability showed a distinct cross-system pattern of representation, Pillar 4 is analysed separately. Across all three occupations, this pillar produced lower similarity values than the cognitive and interpersonal pillars. Its relation to embodied skill, however, differed across representation systems. In O*NET, Pillar 4 (mean ≈ 0.365 across occupations) scored higher than Pillar 2 (mean ≈ 0.321), whereas in RPDC it was the lowest-scoring pillar for dog groomer and massage therapist and slightly below Pillar 2 in the electrician corpus. This indicates that adaptability is represented differently by the systems: in O*NET, it aligns more closely with evaluative, monitoring, and planning language, whereas in vocational standards, it appears less prominently specified than embodied competence. Table VI reports mean cosine similarity values for Pillar 4 across occupations and representation systems.

TABLE VI. MEAN COSINE SIMILARITY VALUES FOR P4

Occupation	*RPDC*	*O*NET*	*ESCO*
Electrician	0.255	0.365	0.186
Dog groomer	0.202	0.365	0.214
Massage therapist	0.228	0.365	0.186

O*NET still produced the highest values among the three systems, but the representational profile of this pillar remained below that observed for Pillars 1 and 3. Within RPDC, Pillar 4 was the lowest-scoring pillar for both dog groomer and massage therapist, and slightly below embodied capabilities (Pillar 2) for electrician. In ESCO, Pillar 4 was also among the weakest represented dimensions, particularly for electrician and massage therapist.

Fig. 4 presents the representation of Situated Adaptive Decision-Making across occupations and representation systems. It shows mean cosine similarity for Pillar 4 only, highlighting the comparatively weaker and more variable representation of this dimension outside O*NET.

These results indicate that adaptability, when defined as situated, non-scripted, real-time behavioural adjustment under uncertainty, is less directly represented in existing occupational taxonomies than cognitive reasoning or interpersonal coordination, and that its representation depends strongly on the descriptive logic of the system considered.

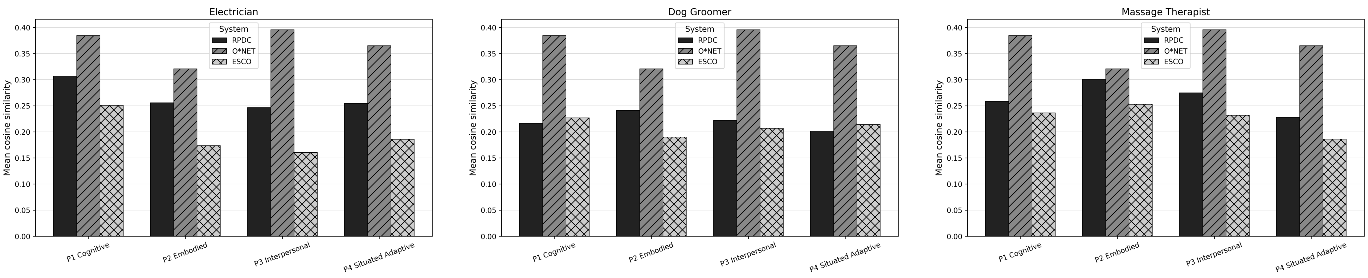
D. Qualitative Inspection of High-Similarity Descriptors for Pillar 4

To complement the quantitative similarity analysis shown in Fig. 2–4, the top 15 descriptors per occupation and representation system with the highest cosine similarity to Pillar 4 were examined qualitatively. This step provides interpretive validation by allowing the semantic structure captured by the embeddings to be examined directly in the descriptor language. The descriptors reported here correspond to the highest-ranked matches produced by the similarity computation. Qualitative inspection of nearest neighbours is a common interpretive step in embedding-based analyses, where retrieved matches are examined to interpret the semantic structure captured by vector similarity measures [40] [41]. Similar approaches have also been used in computational analyses of occupational task structures and automation exposure [42].

For all three occupations, O*NET retrieved highly similar descriptors centred on analytical evaluation, monitoring, and planning, including examples such as: “Using relevant information and individual judgment to determine whether events or processes comply with laws, regulations, or standards”, “Monitoring and reviewing information from materials, events, or the environment, to detect or assess problems”, “Analysing information and evaluating results to choose the best solution and solve problems”, “Identifying information by categorising, estimating, recognising differences or similarities, and detecting changes in circumstances or events”, “Developing specific goals and plans to prioritise, organise, and accomplish your work”.

These examples suggest that O*NET primarily encodes adaptability through descriptors of evaluation, information processing, judgment, and planning.

Vocational standards showed a different pattern. In the electrician corpus, high-similarity RPDC matches included more explicit references to situational adjustment and troubleshooting, such as “Takes environmental conditions into account (temperature, humidity, etc.)”, “Easily adapting behaviour or working methods to changing situations and different people”, “Technical understanding (e. g., analysing and troubleshooting faults in electrical and electronic systems)”. In the dog groomer



(a) Electrician (b) Dog Groomer (c) Massage Therapist

Figure 3. Mean cosine similarity scores for the four capability pillars across occupational representation systems for (a) electrician, (b) dog groomer, and (c) massage therapist. Bars compare RPDC, O*NET, and ESCO descriptors.

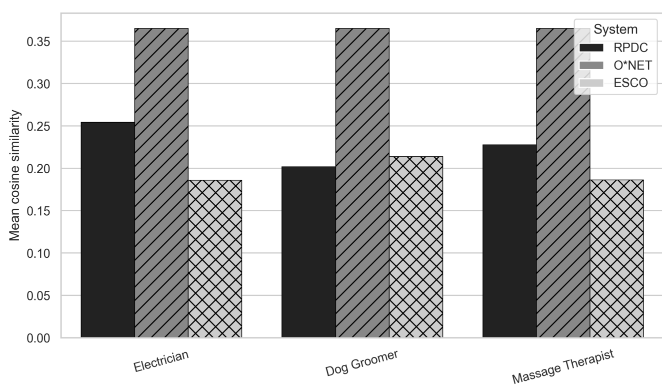


Figure 4. Representation of Situated Adaptive Decision-Making (Pillar 4) across occupations and representation systems. Bars represent mean cosine similarity between pillar anchor sets and occupational descriptors extracted from RPDC, O*NET Work Activities, and ESCO skills.

and massage therapist corpora, high-similarity RPDC matches included descriptors such as “Reaction skills”, “Dealing with change”, “Ability to work independently”, “Goal-oriented working”.

Other RPDC matches combined broader professional characteristics with technical or organisational language, for example, references to analytical skills, individual treatment plans, appointment coordination, and flexibility. All in all, these descriptors acknowledge adaptability, but often do so through a combination of broader professional attributes, technical planning, and occupation-specific competence formulations rather than by consistently explicit descriptions of real-time behavioural adjustment.

ESCO descriptors were more compressed and typically appear as short skill labels rather than process-oriented descriptions. High-similarity examples included “handle customer complaints”, “control animal movement”, “develop personalised massage plan” and “manage a small-to-medium business”. In some cases, ESCO also retrieved descriptors that were relevant but highly condensed, such as “cope with challenging circumstances in the veterinary sector” or “make decisions regarding the animal’s welfare.” Largely, however, the ESCO matches function primarily as categorical labels rather than explicit descriptions of dynamic adaptive behaviour.

Overall, the qualitative inspection shows that all three representation systems capture aspects of adaptability but represent it differently: O*NET through evaluative and planning language, vocational standards through a mixture of occupation-specific planning, adjustment, and broader professional formulations, and ESCO through short categorical skill labels.

E. Summary of Representational Patterns

Across the three occupations, three consistent patterns emerge. First, O*NET provides the strongest overall semantic alignment with the four capability pillars, followed by RPDC and then ESCO, a pattern stable across occupations and distributional statistics. Second, Cognitive–Technical Judgement and Trust-Calibrated Interpersonal Engagement show comparatively strong alignment, while Embodied–Sensorimotor Calibration shows moderate alignment with relatively stronger representation in RPDC for occupations involving explicit bodily technique. Third, Situated Adaptive Decision-Making exhibits the most uneven representational profile, scoring below cognitive and interpersonal pillars quantitatively and differing in semantic form across systems, as confirmed by qualitative inspection.

F. Robustness and Sensitivity Analysis

To assess whether the observed representational ordering is robust to methodological design choices, two sensitivity analyses were conducted as described in Section VII.E. Table IV summarises the results across all three conditions. The finding that O*NET descriptors show the strongest semantic alignment proved fully stable: O*NET remained the highest-scoring system across all 36 pillar–occupation combinations examined, regardless of aggregation rule or embedding model. The secondary ordering of RPDC above ESCO held in the large majority of combinations and showed minor variation only in the dog groomer corpus, consistent with the two exceptions already documented in Section VIII.A and attributable to the small margin of difference in that corpus.

These results indicate that the reported representational ordering is not an artefact of the specific aggregation rule or embedding model selected, and that the core findings are robust to the methodological choices examined.

IX. Discussion

The following subsections interpret the empirical findings in light of the theoretical literature on tacit knowledge, automation modelling, and human-centred work, and consider implications for occupational taxonomy design and AI labour-market research.

A. Adaptability, Tacit Knowledge, and the Limits of Occupational Representations

This analysis was motivated by an unexpected finding in the HCRM. Based on the theoretical and empirical literature, occupations characterised by intensive interpersonal interaction, embodied skill, and situational variability were expected to exhibit high scores on adaptability. Professions such as electricians, dog groomers, and massage therapists routinely operate in environments that require real-time adjustment to changing physical, social, and contextual conditions. Practitioners must simultaneously interpret signals from clients, animals, or physical environments while executing skilled manual tasks, often in private, uncontrolled settings. Such contexts require continuous decision-making without predetermined scripts and involve the coordination of cognitive, embodied, and interpersonal processes.

From the perspective of automation and expertise research, such open-ended adaptive behaviour has long been recognised as difficult to formalise computationally. Early discussions of this challenge appear in the literature on artificial intelligence and skill acquisition, particularly in the work of Dreyfus and Dreyfus [23], who argued that expert human performance often depends on tacit, context-sensitive judgment rather than explicit rule-following. Similarly, Polanyi's well-known observation that "we know more than we can tell" highlights the central role of tacit knowledge in skilled activity [12]. Within labour economics, this insight has been reflected in discussions of Polanyi's paradox, which emphasises the difficulty of automating tasks that rely on implicit knowledge and situational judgement [43].

Similar observations appear in cognitive systems research on human expertise. In the Skill–Rule–Knowledge framework, Rasmussen [17] describes skilled performance as operating through multiple levels of control, ranging from automatic sensorimotor skill execution to rule-guided procedures and analytical problem-solving. Importantly, experts flexibly shift between these modes depending on the demands of the situation. Such dynamic coordination between embodied action, procedural knowledge, and analytical judgement is characteristic of many craft and service occupations, yet it is difficult to capture within static occupational taxonomies that describe work primarily through discrete tasks or procedural categories, as illustrated by the present analysis.

Against this background, the HCRM framework led to the expectation that occupations characterised by direct human interaction and embodied service provision would display particularly high scores on the adaptability dimension when measured using O*NET adaptability-related indicators. Contrary to this expectation, the empirical analysis revealed slightly lower adaptability scores for the selected occupations. This discrepancy prompted a closer examination of the representational systems used to describe work, specifically the occupational taxonomies and databases that frequently underpin empirical automation analyses.

The results of the present study suggest that the unexpected HCRM finding may reflect a limitation of occupational representation systems, rather than an absence of adaptive capacity in the occupations themselves. Across the three occupations and representation systems examined (O*NET, ESCO, and vocational standards), the pillar representing situated adaptive decision-making generally showed lower semantic alignment than the cognitive and interpersonal pillars and, in several cases, was among the weakest represented dimensions.

Qualitative inspection of high-similarity descriptors further illustrates how adaptability is represented in these systems. The descriptors most closely aligned with the adaptability anchors often emphasised evaluating information, monitoring conditions, planning, or broad professional flexibility, rather than explicit descriptions of real-time behavioural adjustment within dynamic environments. While such descriptors capture aspects of decision-making, they predominantly frame adaptability as deliberative evaluation and planning rather than as real-time behavioural adjustment within dynamic environments.

This distinction is theoretically significant. As Suchman [24] demonstrated in her critique of classical AI planning models, human expertise rarely relies on the execution of formal, prespecified cognitive scripts. Instead, action emerges through ongoing interaction with the immediate conditions of the environment, a phenomenon she describes as situated action. Similarly, the ecological psychology tradition emphasises that skilled behaviour emerges from the coupling between agents and environmental affordances, enabling fluid adjustment to changing constraints [26]. From this perspective, expertise often manifests not as explicit problem solving but as situated responsiveness embedded within continuous perception–action loops.

The present findings, therefore, indicate that such forms of situational adaptability may be compressed or indirectly represented within occupational taxonomies. Because taxonomies must translate complex activities into discrete, standardised categories, they tend to privilege formally articulated procedures and explicit decision points. In the terminology of resilience engineering, this reflects the well-known distinction between work-as-imagined (formal procedures and plans) and work-as-done (the situated, adaptive practices through which work is actually carried out) [44].

B. Occupational Taxonomies and Automation Modelling

These representational dynamics are particularly relevant in the context of labour-market automation research. Empirical research on automation has increasingly emphasised that the economic effects of new technologies depend not only on their technical feasibility but also on how work is organised into tasks within occupations [9]. This perspective builds on the task-based framework in labour economics, which conceptualises occupations as bundles of heterogeneous tasks rather than fixed job categories [25]. Within this framework, occupational databases such as O*NET are frequently used to approximate the task structure of work, providing the empirical inputs for models estimating technological substitution, task exposure, or augmentation potential. Many influential studies estimating the susceptibility of occupations to automation rely on occupational databases such as O*NET to characterise task content and skill requirements. For example, Frey and Osborne [5] used descriptors derived from the O*NET database to train machine-learning classifiers estimating the probability of computerisation across occupations. Related approaches have been adopted in subsequent studies analysing the effects of artificial intelligence, robotics, and digital technologies on employment structures [6] [7] [10]. More recently, analyses of generative AI exposure have relied on O*NET work activities to approximate the task structure of occupations [8].

In such frameworks, occupational characteristics derived from taxonomies function as empirical inputs to models estimating automation risk. If certain dimensions of work are under-represented or abstracted in these representations, automation estimates may partly reflect the representational structure of the taxonomy used to describe work, rather than the full complexity of real-world work activities.

An important contrast emerged between embodied skill and situated adaptability. As shown in the qualitative inspection (Section VIII.D), adaptability in O*NET is primarily represented through abstract descriptors related to evaluation, monitoring, and planning.

In the original HCRM analysis, the embodied dexterity pillar produced the strongest differentiation for the selected occupations. This suggests that although O*NET represents manual activity through general descriptors, such as manual dexterity and finger dexterity, these indicators still capture core aspects of skilled manual work and therefore contribute to higher scores for occupations involving hands-on practice. By contrast, adaptability in O*NET is largely represented by cognitive descriptors related to problem-solving, planning, and decision-making. This formulation may favour professional occupations and may partly explain why the expected adaptability advantage did not appear for the selected service and craft occupations.

In vocational standards, by contrast, embodied skill was often more strongly represented than adaptability, while situated real-time adjustment remained weakly specified. This divergence indicates that different representation systems privilege different aspects of work: survey-based taxonomies more readily represent abstract decision processes, whereas training-oriented documents more readily encode physical technique than unscripted behavioural adaptation.

Taken together, these findings suggest that situated adaptability is one of the capability domains most vulnerable to under-specification in occupational taxonomies. Many service and craft occupations rely on adjustment to evolving biological, interpersonal, and physical conditions that cannot be fully specified in advance. Such adaptive processes have historically posed major challenges for automation. Robotics research has long recognised the difficulty of replicating human adaptability in unstructured environments, an asymmetry famously captured by Moravec's paradox, which observes that high-level reasoning is computationally easier than sensorimotor interaction with complex physical environments [45] [46]. Despite major advances in machine learning and robotics, the real-time coordination of perception, action, and social interaction in dynamic environments remains a central technical challenge.

From this perspective, the representational structure of occupational taxonomies becomes particularly important. If adaptive work practices are primarily represented through abstract cognitive descriptors, occupations that rely heavily on embodied or situational adaptation may appear more amenable to automation than they are under real deployment conditions. This does not invalidate taxonomy-based automation research; rather, it suggests that such models should be interpreted with awareness of the representational assumptions embedded in their input data. The structure of occupational representation systems may influence how automation risk is estimated, highlighting the importance of examining how human capabilities are encoded in the datasets used to model technological change.

C. *Implications for Human-Centred Work and AI Integration*

More broadly, these findings contribute to ongoing debates about the resilience of human-centred service occupations in the context of artificial intelligence and technological change. Occupations involving close physical proximity, interpersonal trust, and real-time interaction with unpredictable environments occupy a distinctive position within the contemporary labour market. Such roles often require workers to coordinate technical skill, social communication, and embodied responsiveness simultaneously.

In many service contexts, practitioners must continuously interpret subtle signals, such as a client's discomfort, an animal's behavioural cues, or constraints imposed by the physical environment, while adapting their actions accordingly. These adaptive processes are frequently acquired through experience rather than formal instruction and may therefore remain only partially codified in occupational standards, consistent with longstanding insights on the role of tacit knowledge in professional practice [47] [48]. The present analysis suggests that such forms of situated adaptability may therefore be only indirectly represented within existing occupational taxonomies.

These observations do not imply that such occupations are immune to technological change. Rather, they suggest that the interaction between artificial intelligence and human-centred work may be better understood through augmentation frameworks than simple substitution models. Recent labour-economic analyses increasingly emphasise that digital technologies may restructure tasks within occupations rather than eliminate occupations entirely [49]. In this perspective, AI systems may automate certain cognitive components of work while increasing the relative importance of human capabilities that remain difficult to formalise computationally.

Accordingly, in many HCRM-type occupations, AI technologies are likely to function primarily as complementary tools supporting planning, documentation, scheduling, or information retrieval, leaving situational judgement and embodied interaction relatively more dependent on human practitioners. Understanding how occupational representation systems capture, or fail to capture, these capabilities is therefore important for designing AI systems that complement rather than attempt to replace human expertise. In this sense, the HCRM provides a framework for identifying capability domains that may remain under-specified in occupational datasets. The present analysis constitutes an initial step toward a broader empirical programme examining how human-centred capabilities are encoded and measured within the data infrastructures used to model technological change in labour markets.

D. Limitations

This study provides an exploratory analysis of how occupational taxonomies represent different dimensions of human work and encode occupational capabilities in datasets commonly used to model automation risks. The empirical analysis was limited to three occupations and three representation systems. Expanding the range of occupations would help determine whether the observed representational patterns generalise across broader segments of the labour market.

A further limitation concerns the institutional and geographical scope of the occupational representation systems examined. The analysis combines O*NET, which reflects the structure of the U.S. labour market, with ESCO and national vocational standards drawn primarily from European training frameworks. While these datasets represent some of the most widely used occupational taxonomies in labour-market research, they may not fully capture how skills and competencies are defined in other institutional or cultural contexts. In particular, vocational training systems outside Europe and North America may structure occupational knowledge differently.

The embedding-based methodology employed in this study estimates semantic similarity between textual descriptors and conceptual anchor sets, but does not analyse the full structural organisation of occupational taxonomies. A sensitivity analysis reported in Section VIII.F confirms that the core findings are robust to the choice of aggregation rule and embedding model, though the scope of this robustness check remains limited to the conditions examined.

X. Conclusion and Future Work

This study examined how occupational representation systems encode different dimensions of human work, with particular attention to situational adaptability. The analysis was motivated by an unexpected finding in the HCRM: occupations characterised by highly variable interpersonal and physical environments did not exhibit the anticipated high scores on the adaptability pillar. Because occupations such as electricians, dog groomers, and massage therapists require continuous real-time adjustment to changing circumstances, this discrepancy prompted a closer examination of the occupational representational systems used to describe work activities.

Using embedding-based semantic similarity analysis, the study compared three widely used occupational representation systems: O*NET Work Activities, ESCO skills, and national vocational training standards. Across all three occupations, a consistent pattern emerged. O*NET descriptors exhibited the strongest semantic alignment with the conceptual capability pillars, vocational standards showed intermediate alignment, and ESCO descriptors generally produced lower similarity values, with minor exceptions in the dog groomer corpus. At the same time, the pillar representing non-scripted adaptive decision-making showed a weaker alignment than the cognitive and interpersonal pillars and was among the weaker represented dimensions across the examined representation systems.

Qualitative inspection of high-similarity descriptors further clarified this pattern. Adaptability in occupational taxonomies is typically represented through abstract cognitive processes, such as analysing information, evaluating outcomes, or planning actions. While these descriptors capture important aspects of decision-making, they tend to frame adaptability primarily as deliberative reasoning rather than as situated behavioural adjustment within dynamic environments. In many service and craft occupations, however, expertise manifests through continuous perception–action coordination in response to evolving physical, biological, and interpersonal conditions.

Taken together, the findings point to a representational asymmetry in occupational description systems. Dimensions of work that can be explicitly articulated, such as analytical reasoning, planning, and communication, are comparatively well represented, whereas tacit and situational forms of adaptive expertise appear more weakly specified. Because many labour-market automation studies rely on occupational taxonomies to operationalise task content, these representational asymmetries may influence automation exposure estimates for any occupation containing tacit, situational, or context-dependent components of expertise, not only those characterised by highly interpersonal or embodied work, though the present analysis focuses on three illustrative cases from this domain.

Importantly, the results should not be interpreted as evidence that existing taxonomies are inadequate. Systems such as O*NET provide rich and systematically structured descriptions of work activities and remain indispensable resources for labour-market research and policy analysis. Rather, the findings highlight the inherent difficulty of representing certain forms of human expertise, particularly those grounded in tacit knowledge, embodied interaction, and context-sensitive judgement, within standardised classification frameworks.

More broadly, the study contributes to emerging discussions on the data infrastructures underlying automation research. A sensitivity analysis confirmed that the core representational ordering is robust to the choice of aggregation rule and embedding model, supporting the reliability of the findings within the scope of the present study. As artificial intelligence systems perform an expanding range of cognitive tasks, understanding how occupational capabilities are represented in the datasets used to model technological change becomes increasingly important. Embedding-based representational audits, such as the one presented here, offer a systematic method for examining how different dimensions of human work are encoded in occupational databases and for identifying capability domains that may remain underspecified in existing representations of work.

Future research should extend this representational audit to a broader range of occupations and occupational databases, including taxonomies from labour-market regimes outside Europe and North America. Complementary methodological approaches, such as network analysis, ontology comparison, and multi-modal observational datasets including task recordings and sensor-based studies of manual work, could provide additional insight into how embodied and situated capabilities are formally represented and organised within occupational systems. Extending embedding-based audits across multiple labour-market datasets may contribute to a more systematic understanding of how occupational knowledge is encoded in the data infrastructures that increasingly inform economic forecasting and AI capability assessment.

Acknowledgement

The author thanks Prof. Matjaž Gams for his insightful guidance and valuable feedback on the original HCRM. The author is also grateful to Professors Dr Matej Mertik and Dr Maciej Wielgosz for their continued intellectual support throughout the development of this research. Their stimulating discussions, thoughtful questions, and encouragement helped refine many of the ideas presented in this work.

This research was conducted as part of the Applied Artificial Intelligence doctoral programme at Alma Mater Europaea, and the supportive academic environment is gratefully acknowledged. Any opinions, findings, or conclusions expressed in this paper are those of the author and do not necessarily reflect the views of the affiliated institutions.

References

[1] I. E. Negro, 'The resilience of the leisure and care economy: Human-centred niches in an AI-driven labour market', in *AISyS 2025, The Second International Conference on AI-based Systems and Services*, Lisbon, Portugal: IARIA, 2025, pp. 1–3, ISBN: 978-1-68558-303-3.

[2] I. E. Negro, 'The slow food of labour: A task-structural analysis of Human-centred resilience in AI-driven labour markets', in *Proceedings of the Future Technologies Conference (FTC) 2026*, in press, Berlin, Germany, 2026.

[3] N. G. Peterson et al., 'Understanding work using the occupational information network (O*NET): Implications for practice and research', *Personnel Psychology*, vol. 54, no. 2, pp. 451–492, Jun. 2001. DOI: 10.1111/j.1744-6570.2001.tb00100.x

[4] European Commission, 'European skills, competences, qualifications and occupations (ESCO)', 2025, Accessed: 2026-02-08. [Online]. Available: https://esco.ec.europa.eu/en

[5] C. B. Frey and M. A. Osborne, 'The future of employment: How susceptible are jobs to computerisation?', *Technological Forecasting and Social Change*, vol. 114, pp. 254–280, Jan. 2017. DOI: 10.1016/j.techfore.2016.08.019

[6] M. Arntz, U. Zierahn and T. Gregory, 'The risk of automation for jobs in OECD countries: A comparative analysis', OECD, Social, Employment and Migration Working Papers 189, May 2016. DOI: 10.1787/5jlz9h56dvq7-en

[7] E. Brynjolfsson, T. Mitchell and D. Rock, 'What can machines learn, and what does it mean for occupations and the economy?', *AEA Papers and Proceedings*, vol. 108, pp. 43–47, May 2018. DOI: 10.1257/pandp.20181019

[8] T. Eloundou, S. Manning, P. Mishkin and D. Rock, 'GPTs are GPTs: An early look at the labor market impact potential of large language models', 2023, Accessed: 2026-03-07. [Online]. Available: https://arxiv.org/abs/2303.10130

[9] D. Acemoglu and P. Restrepo, 'Artificial intelligence, automation and work', National Bureau of Economic Research, Cambridge, MA, Tech. Rep. w24196, Jan. 2018. DOI: 10.3386/w24196

[10] M. Webb, 'The impact of artificial intelligence on the labor market', Nov. 2019, Accessed: 2026-03-07. [Online]. Available: https://ssrn.com/abstract=3482150

[11] D. H. Autor, F. Levy and R. J. Murnane, 'The skill content of recent technological change: An empirical exploration', *Quarterly Journal of Economics*, vol. 118, no. 4, pp. 1279–1333, Nov. 2003. DOI: 10.1162/003355303322552801

[12] M. Polanyi, *The Tacit Dimension*. Chicago: University of Chicago Press, 1966.

[13] M. J. Handel, 'The O*NET content model: Strengths and limitations', *Journal for Labour Market Research*, vol. 49, no. 2, pp. 157–176, Oct. 2016. DOI: 10.1007/s12651-016-0199-8

[14] R. M. Ryan and E. L. Deci, 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being', *American Psychologist*, vol. 55, no. 1, pp. 68–78, 2000. DOI: 10.1037/0003-066X.55.1.68

[15] P. Bourdieu, *Distinction: A Social Critique of the Judgement of Taste*. Cambridge, MA: Harvard University Press, 1984.

[16] P. Bourdieu, *The Logic of Practice*. Stanford, CA: Stanford University Press, 1990.

[17] J. Rasmussen, 'Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models', *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-13, no. 3, pp. 257–266, 1983. DOI: 10.1109/TSMC.1983.6313160

[18] G. Klein, *Sources of Power: How People Make Decisions*. Cambridge, MA: MIT Press, 1998.

[19] D. D. Woods, 'The theory of graceful extensibility: Basic rules that govern adaptive systems', *Environment Systems and Decisions*, vol. 38, no. 4, pp. 433–457, Dec. 2018. DOI: 10.1007/s10669-018-9708-3

[20] M. Mulder, T. Weigel and K. Collins, 'The concept of competence in the development of vocational education and training in selected EU member states: A critical analysis', *Journal of Vocational Education & Training*, vol. 59, no. 1, pp. 67–88, 2007. DOI: 10.1080/13636820601145630

[21] C. Bohne, F. Eicker and G. Haseloff, 'Competence-based vocational education and training (VET): An approach of shaping and networking', *European Journal of Training and Development*, vol. 41, no. 1, pp. 28–38, 2017. DOI: 10.1108/EJTD-07-2016-0052

[22] International Labour Organization, 'Competency-based training (CBT): An introductory manual for practitioners', Oct. 2020, Accessed: 2026-03-11. [Online]. Available: https://www.ilo.org/publications/competency-based-training-cbt-introductory-manual-practitioners

[23] H. L. Dreyfus and S. E. Dreyfus, *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. New York: Free Press, 1986.

[24] L. A. Suchman, *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge, UK: Cambridge University Press, 1987.

[25] D. Autor, 'The "Task approach" to labor markets: An overview', National Bureau of Economic Research, Cambridge, MA, Tech. Rep. w18711, Jan. 2013. DOI: 10.3386/w18711

[26] J. J. Gibson, *The Ecological Approach to Visual Perception: Classic Edition*, 1st ed. New York: Psychology Press, 1979.

[27] F. J. Varela, E. Rosch and E. Thompson, *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press, 1991. DOI: 10.7551/mitpress/6730.001.0001

[28] N. A. Bernstein, *The Co-ordination and Regulation of Movements*. Oxford: Pergamon Press, 1967.

[29] A. R. Hochschild, *The Managed Heart: Commercialization of Human Feeling*. Berkeley: University of California Press, 1983.

[30] E. Goffman, *Interaction Ritual: Essays in Face to Face Behavior*. New York: Doubleday, 1967.

[31] D. J. Deming, 'The growing importance of social skills in the labor market', *Quarterly Journal of Economics*, vol. 132, no. 4, pp. 1593–1640, Nov. 2017. DOI: 10.1093/qje/qjx022

[32] M. Deng and X. Fan, 'Post-editing efficiency and quality assessment: A comparative analysis of Google translate and DeepL', *IOSR Journal of Humanities and Social Science*, vol. 30, no. 1, pp. 10–24, Jan. 2025. DOI: 10.9790/0837-3001021024

[33] M. I. Kamaluddin, M. W. K. Rasyid, F. H. Abqoriyyah and A. Saehu, 'Accuracy analysis of DeepL: Breakthroughs in machine translation technology', *Journal of Education, English, and Foreign Language*, vol. 4, no. 2, pp. 122–126, Jun. 2024. DOI: 10.29303/jeef.v4i2.681

[34] L. Li, 'Artificial intelligence translator DeepL translation quality control', *Procedia Computer Science*, vol. 247, pp. 710–717, 2024. DOI: 10.1016/j.procs.2024.10.086

[35] S. Sabrina, M. Fadhillah, F. O. Chawdri, C. Safara and L. Juniarti, 'Contrastive analysis of DeepL translation vs. Google translate's performance in rendering academic texts', *Jurnal Sinestesia*, vol. 26, no. 1, pp. 73–82, Apr. 2025. DOI: 10.32672/jsi.v26i1.2502

[36] T. Mikolov, K. Chen, G. Corrado and J. Dean, 'Efficient estimation of word representations in vector space', 2013, Accessed: 2026-03-07. [Online]. Available: https://arxiv.org/abs/1301.3781

[37] J. Devlin, M.-W. Chang, K. Lee and K. Toutanova, 'BERT: Pre-training of deep bidirectional transformers for language understanding', 2018, Accessed: 2026-03-07. [Online]. Available: https://arxiv.org/abs/1810.04805

[38] N. Reimers and I. Gurevych, 'Sentence-BERT: Sentence embeddings using siamese BERT-networks', in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, China: Association for Computational Linguistics, 2019, pp. 3980–3990. DOI: 10.18653/v1/D19-1410

[39] A. C. Kozlowski, M. Taddy and J. A. Evans, 'The geometry of culture: Analyzing meaning through word embeddings', 2018, Accessed: 2026-03-07. [Online]. Available: https://arxiv.org/abs/1803.09288

[40] W. Wang et al., 'MiniLM: Deep self-attention distillation for task-agnostic compression of pre-trained transformers', 2020, Accessed: 2026-03-07. [Online]. Available: https://arxiv.org/abs/2002.10957

[41] W. L. Hamilton, J. Leskovec and D. Jurafsky, 'Diachronic word embeddings reveal statistical laws of semantic change', in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, Berlin, Germany: Association for Computational Linguistics, 2016, pp. 1489–1501. DOI: 10.18653/v1/P16-1141

[42] P. Brandes and R. Wattenhofer, 'Opening the Frey/Osborne black box: Which tasks of a job are susceptible to computerization?', 2016, Accessed: 2026-03-07. [Online]. Available: https://arxiv.org/abs/1604.08823

[43] D. Autor, 'Polanyi's paradox and the shape of employment growth', National Bureau of Economic Research, Cambridge, MA, Tech. Rep. w20485, Sep. 2014. DOI: 10.3386/w20485

[44] E. Hollnagel, 'Resilience engineering: A new understanding of safety', *Journal of the Ergonomics Society of Korea*, vol. 35, no. 3, pp. 185–191, Jun. 2016. DOI: 10.5143/JESK.2016.35.3.185

[45] H. Moravec, *Mind Children: The Future of Robot and Human Intelligence*. Cambridge, MA: Harvard University Press, 1988.

[46] R. A. Brooks, 'Intelligence without representation', *Artificial Intelligence*, vol. 47, no. 1–3, pp. 139–159, Jan. 1991.

[47] I. Nonaka and H. Takeuchi, *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. New York: Oxford University Press, 1995.

[48] M. Eraut, 'Non-formal learning and tacit knowledge in professional work', *British Journal of Educational Psychology*, vol. 70, no. 1, pp. 113–136, Mar. 2000. DOI: 10.1348/000709900158001

[49] D. Autor, 'Applying AI to rebuild middle class jobs', National Bureau of Economic Research, Cambridge, MA, Tech. Rep. w32140, Feb. 2024. DOI: 10.3386/w32140