*The object of this study is the alignment of Indonesian computing graduate learning outcomes (CPL – Capaian Pembelajaran Lulusan) with three heterogeneous competency frameworks: the European skills, competences, qualifications and occupations (ESCO), the occupational information network (O*NET), and the Standar Kompetensi Kerja Nasional Indonesia (SKKNI). The problem addressed is the absence of a unified pipeline capable of simultaneously mapping CPL across these frameworks while accounting for national qualification hierarchies and cross-lingual constraints. An IR-KG (information retrieval–knowledge graph) model is proposed with a seven-stage pipeline using a hybrid scoring function S_final = α·S_sem + + β·S_gr + γ·S_con, integrating TF-IDF (term frequency–inverse document frequency) semantic similarity, ESCO knowledge graph cohesion, and domain constraint scores from ISCED-F 2013 (International Standard Classification of Education: Fields of Education and Training) and APTIKOM 2022 (Asosiasi Pendidikan Tinggi Informatika dan Komputer) classifications. The balanced configuration (α = = β = γ = 0.33) achieves a mean selection objective of 0.537, a 26.1% improvement over the semantic baseline. External consistency validation yields a relaxed consistency rate of 27.1% (8.7× above random baseline), confirming valid alignment signal capture. The CRI-KG (Career Readiness index based on knowledge graph) reveals a gradient R_SKKNI >> R_ONET > R_ESCO, exposing persistent gaps in international framework coverage. The pipeline is applicable for curriculum audit, qualification recognition policy, and national-to-international framework integration where labelled training data are unavailable*

*Keywords: information retrieval, knowledge graph, competency alignment, CPL, SKKNI, O*NET, ESCO, career readiness index*

UDC 004.912:025.4:378.1
DOI: 10.15587/1729-4061.2026.358313

DEVELOPMENT OF A HYBRID INFORMATION RETRIEVAL-KNOWLEDGE GRAPH MODEL FOR CROSS-FRAMEWORK COMPETENCY ALIGNMENT

Halim Maulana
Corresponding Author
Doctoral Candidate, Doctoral Program in Computer Science*
E-mail: halimmaulana@umsu.ac.id
ORCID: https://orcid.org/0000-0003-4361-6377

Poltak Sihombing
Professor, Doctor of Computer Science*
ORCID: https://orcid.org/0000-0001-5348-4537

Amalia
Doctor, Head of Computer Science Study Program*
ORCID: https://orcid.org/0000-0003-0595-8296

Marischa Elveny
Doctor, Lecturer – Master's Program in Data Science and AI*
ORCID: https://orcid.org/0000-0001-5702-8846

*Department of Computer Science
Universitas Sumatera Utara
Dr. T. Mansur str., 9, Kampus Padang Bulan, Medan, Indonesia, 20155

Received 06.02.2026
Received in revised form 25.03.2026
Accepted date 14.04.2026
Published date 30.04.2026

How to Cite: *Maulana, H., Sihombing, P., Amalia, A., Elveny, M. (2026). Development of a hybrid information retrieval-knowledge graph model for cross-framework competency alignment. Eastern-European Journal of Enterprise Technologies, 2 (2 (140)), 32–42. https://doi.org/10.15587/1729-4061.2026.358313*

1. Introduction

The artificial intelligence, automation, and digital platform economies that brought about the structural transformation of labor markets across the globe fundamentally changed what competencies are demanded [1, 2]. The World Economic Forum 2016 report anticipates more than 44% of core workforce skills will change significantly in the next five years while an OECD (Organization for Economic Co-operation and Development) Jobs Study estimates that over a third of all jobs face a high risk of automation [3, 4]. Consequently, the question of skill mismatch has transformed from a quantitative supply-demand issue to a qualitative concern: competency profiles of the graduates must be dynamically matched with the skills required in technology-driven labor markets.

Developed economies have responded by building machine-readable competency taxonomies to facilitate automated, semantically grounded workforce analytics. The European skills, competences, qualifications and occupations (ESCO) taxonomy is a multilingual RDF-SKOS knowledge graph containing information on 13,939 skills and 3,008 occupations [5]. The occupational information network (ONET) serves this need with a series of structured knowledge, skills and abilities (KSA) based profiles on 879 computing-sector occupations [6]. Notwithstanding the complementary nature of their scope, interoperability between ESCO and ONET is limited due to dissimilar granularity, definitional horizons, and relational organization-factors that deny one-to-one alignment and confine cross-framework correspondence to partial semantic equivalence [7, 8].

This challenge is made more complicated in developing economies, where national qualification frameworks have often been developed without institutional, linguistic or ontological alignment with global taxonomies. In the ICT sector, Indonesia's mandatory workforce certification standard SKKNI (Standar Kompetensi Kerja Nasional Indonesia) comprises 1,711 unit-based competency descriptions in Bahasa Indonesia, administered by Badan Nasional Sertifikasi Profesi (BNSP) [9]. Indonesian computing graduate learning outcomes (CPL Capaian Pembelajaran Lulusan) are mandated through the national qualifications framework (KKNI) and implemented via the APTIKOM outcome-based education model as unstructured natural language statements. No existing computational mechanism simultaneously maps CPL against SKKNI and global frameworks within a single pipeline.

This gap has direct consequences. Universities cannot measure how well their curricula prepare graduates for domestic certification or international mobility. Qualification recognition agencies lack empirical evidence for bilateral negotiations. A computational alignment model spanning CPL, SKKNI, ESCO, and O*NET without requiring labelled training data would give curriculum designers, accreditation bodies, and workforce planners a diagnostic tool that does not currently exist.

2. Literature review and problem statement

Paper [7] investigated ESCO-O*NET mapping and found that the list of skills between the two systems only partially aligned. The strength of this study is the identification of a common vocabulary baseline across two major international initiatives. Its major disadvantage is that it does not bridge structural ontological differences, as O*NET operates on macro-level occupational granularity while ESCO is focused at micro-granular skills. The question of how to achieve proper alignment at multiple levels remains unresolved in fully automated alignment, as cross-level granularity mismatches still need to be overcome. This was left unresolved for an objective reason: there was no mechanism to incorporate graph-structural signals together with lexical similarity simultaneously. In paper [8] it was extended by combining large language models (LLMs) with domain knowledge, obtaining better but still limited alignments. The advantage includes the use of domain-specific signals; the disadvantage is that complex cross-ontology relational structures remain uncaptured. The open question is whether structural ontological signals are sufficient to fill the gap of missing relational bridges across frameworks. This remained unresolved because transformer-based approaches were applied without graph-structural augmentation, and no national qualification framework was considered within the alignment scope. In paper [10], documents the relational structure of the ESCO knowledge graph, including its skill hierarchy and relation predicates. The advantage of this work is that it delivers a comprehensive ontological reference for ESCO-based alignment study. The disadvantage is that the relation predicates are designed for human navigation rather than automated semantic processing, resulting in structural difficulties in automating job semantics analysis. The open question is how to leverage ESCO's graph topology computationally for cross-framework alignment. This was left unresolved for an objective reason: the ESCO Handbook was designed as a reference document rather than a computational resource, leaving its graph-structural signals underutilized in existing alignment pipelines.

Due to their efficiency, interpretability and adaptability to new domains, TF-IDF (term frequency-inverse document frequency) weighted cosine similarity [11] and BM25 (Best Match 25) [12] are commonly used as baselines in skill extraction and matching. In paper [13], effectiveness was demonstrated for ESCO skill extraction in a single-framework setting. The major advantage of these methods is that they are transparent, have low computational cost, and do not require labelled training data. The drawback, however, is their inability to capture lexically oblique semantic equivalences, such as mapping 'pengembangan sistem informasi' (Indonesian) to 'information systems design' (English), in the absence of bridge-text augmentation. The open challenge is how to generalize TF-IDF-based retrieval to cross-framework and cross-lingual settings without sacrificing interpretability. This remained unresolved for a subjective reason: the research community favored neural methods over principled IR-based extensions in multilingual settings. In papers [14, 15], further leveraged transformer architectures and multilingual BERT (bidirectional encoder representations from transformers) for monolingual and cross-lingual skill matching respectively, attaining state-of-the-art results on English benchmarks. The advantage is superior semantic capture across languages; the disadvantage is that both approaches require labelled source-target skill pair datasets, which do not exist for CPL-to-ESCO/O*NET/SKKNI alignment in Indonesia. Additionally, dense embedding approaches output similarity scores in a black-box manner, providing no term-level interpretability that is necessary for curriculum auditors. The challenge of achieving cross-lingual competency alignment in low-resource settings without sacrificing interpretability thus remains open. This has not been resolved for an objective reason: no labelled benchmark dataset exists for Indonesian computing competency alignment, meaning that supervised neural methods cannot be applied in this context.

In paper [16], formal ontology matching methods demonstrated that graph-structural signals are essential for linking concepts across heterogeneous ontological structures. The advantage is a rigorous theoretical foundation for ontology alignment; the disadvantage is that these methods require formally structured ontologies and cannot process unstructured natural language competency statements such as CPL. The open question is how to apply graph-structural alignment signals to competency frameworks that include unstructured text. This remained unresolved for an objective reason: the gap between formal ontology matching and natural language processing had not been bridged in the competency alignment domain. In paper [17], a systematic mapping of computing curricula was performed, providing a reference framework for aligning learning outcomes with international standards. The advantage is a comprehensive and authoritative curriculum taxonomy; the disadvantage is that the mapping is manual, single-framework, and not computationally operationalized, limiting its scalability. The open question is how to achieve scalable automated curriculum-to-framework alignment across multiple heterogeneous frameworks. This remained unresolved for a subjective reason: the educational community has prioritized normative curriculum design over computational alignment tools. Paper [18] employed semantic embedding and distant supervision to improve ontology alignment, showing that keyword matching alone is insufficient for capturing implicit cross-domain conceptual relationships. The advantage is the effective use of complementary signals beyond lexical similarity; the disadvantage is that the approach still relies on partial supervision and does not account for cross-lingual or national standard integration. The common unresolved question across these three studies is how to construct a fully automated, unsupervised, cross-lingual KG-based alignment system that incorporates national competency standards. This remained unresolved for an objective reason: all existing KG-based competency alignment study is confined to two-framework scenarios without cross-lingual dimensions or integration with nationally developed taxonomies in low-resource languages.

Career readiness was examined in papers [19, 20] from a normative theoretical lens, offering conceptual definitions for readiness competencies. The strength is a solid theoretical foundation for conceptualizing graduate employability; the weakness is the total lack of computational operationalization, rendering these frameworks inapplicable in the context of automated curriculum analysis. The open question is how to translate normative career readiness definitions into computable, framework-specific coverage metrics. This remained unresolved for a subjective reason: the career readiness community has emphasized conceptual validation over computational implementation. In [21], it is explicitly called for machine-readable semantic formulations of learning outcomes that can be automatically mapped against global industry standards, yet no computational response to this call exists in the context of developing-country national frameworks. The advantage of the UNESCO report is providing a clear policy direction; the disadvantage is the absence of any implementation pathway for national contexts such as Indonesia. In paper [22], European labor market intelligence tools were developed that operationalize career readiness computationally, but exclusively within European frameworks, making them inapplicable to national standards such as SKKNI. The advantage is demonstrating the feasibility of computational career readiness measurement; the disadvantage is zero applicability outside resource-wealthy European framework settings. The open question is how to construct a career readiness index that is simultaneously grounded in multiple heterogeneous frameworks, calibrated with framework-specific weights, and validated without annotated ground truth. This was left unresolved for an objective reason: combining coverage from both national and international frameworks into a single calibrated readiness metric applicable in low-resource settings has never been successfully performed.

Local problems identified and systematized from these sources point to a common unresolved problem: no computational model exists that is capable of jointly aligning higher education learning outcomes across skill-based (ESCO), occupation-based (O*NET), and national unit-based (SKKNI) frameworks in a cross-lingual, low-resource setting, while incorporating domain boundary knowledge from curriculum standards, providing term-level interpretability, and validating alignment quality without annotated ground truth. all four identified gaps within the context of Indonesian computing education.

Cross-framework competency alignment across heterogeneous taxonomies, languages, and national contexts remains an open problem in workforce intelligence study. The sources reviewed above converge on a single unresolved gap: no computational model exists that jointly aligns higher education learning outcomes across skill-based (ESCO), occupation-based (O*NET), and national unit-based (SKKNI) frameworks in a cross-lingual, low-resource setting while incorporating domain boundary knowledge from curriculum standards, maintaining term-level interpretability, and validating alignment quality without annotated ground truth. The lack of a unified pipeline capable of simultaneously mapping CPL across these frameworks while accounting for national qualification hierarchies and cross-lingual constraints is the central problem this study addresses.

3. The aim and objectives of the study

The aim of this study is to develop a hybrid IR-KG model that maps Indonesian computing graduate learning outcomes (CPL) across ESCO, O*NET, and SKKNI simultaneously combining TF-IDF retrieval, ESCO graph cohesion, and domain constraint scoring without needing labelled training pairs. This will allow universities, accreditation bodies, and qualification agencies to measure how well their CPL covers each framework, pinpoint where the gaps are, and get audit-ready evidence without relying on expert annotation.

To achieve this aim, the following objectives were accomplished:

– to establish and realize a hybrid scoring function purpose $S\_final = \alpha \cdot S\_sem + \beta \cdot S\_gr + \gamma \cdot S\_con$ with three TF-IDF based semantic similarity S_sem, ESCO Knowledge Graph cohesion S_gr, and domain constraint signals (S_con) according ISCED-F 2013 and APTIKOM 2022 curriculum based on the classification;

– to implement an end-to-end seven-stage pipeline performing eight cross-framework and cross-lingual mapping tasks over the CPL-ESCO, CPL-O*NET, CPL-SKKNI, SKKNI-ESCO and SKKNI-O NET framework pairs, and to carry out an ablation study across six alternate parametrized weight configurations measured by a composite selection metric;

– to construct and validate the career readiness index based on knowledge graph (CRI-KG) as a domain-specific coverage metric with calibrated framework weights, and to establish an external consistency validation (ECV) protocol for estimating alignment quality using the officially released ESCO-O*NET crosswalk as a consistency mediator, without relying on annotated ground truth.

4. Materials and methods

4. 1. The object and hypothesis of the study

The object of the study is the alignment of Indonesian computing graduate learning outcomes (CPL) with three competency frameworks ESCO, O*NET, and SKKNI across language and ontological boundaries.

The main hypothesis is that a scoring function combining TF-IDF semantic similarity, ESCO knowledge graph cohesion, and domain constraint scores from ISCED-F 2013 and APTIKOM 2022 will outperform a purely semantic baseline in cross-framework competency alignment especially on tasks involving Indonesian-to-English mapping and multi-framework coverage.

Assumptions made in this study:

– CPL statements are valid proxies for graduate competency profiles of Indonesian computing programs;

– the official ESCO–O*NET crosswalk is a reliable external reference for consistency validation;

– APTIKOM 2022 and ISCED-F 2013 classifications accurately capture the discipline boundaries of Indonesian computing education.

Simplifications adopted:

– bridge-text augmentation through APTIKOM 2022 and CC2020 vocabulary is sufficient for TF-IDF-based cross-lingual retrieval between Bahasa Indonesia and English;

– the CPL dataset from a single institution is treated as a representative sample for initial pipeline validation.

4. *2*. Research design

The study involves the Indonesian computing graduate learning outcomes (CPL), with 30 CPL items from two computing programs at Universitas Sumatera Utara divided according to the four KKNI competency domains; attitude, knowledge, general skills and specific skills. For this, the study follows a design science research (DSR) methodology [23] that is suitable for research, which main outcome consists of building a computational artefact, in this case, evaluated IR-KG system according to explicit quality criteria. DSR dictates some pseudo-cycles of design, implementation and evaluation which can be seen in the way that an ablation study is designed by varying model parameters to find a performant configuration systematically. The pipeline is built in Python 3.10 (Python Software Foundation, USA) on common laptop hardware (Intel Core i7, Intel Corporation, USA; 16 GB RAM, CPU only) making the methodology scalable without specialized GPU resource.

To disentangle each scoring component's contribution through weight variation, six configurations are shown in Table 2. Configuration v0.9 is the semantic baseline ($\beta = \gamma = 0$), which serves as a reference for hybrid configurations. Configurations v1.0 and v1.4 consider enhancements over the baseline by incrementally introducing *S_gr* and *S_con* weights. Configuration v1.1 eliminates the acceptance gate ($\tau = 0$) to abstract from threshold filtering effects in the comparison. Configuration v1.3 follows a stricter Q3 threshold to evaluate the precision-coverage trade-off at high selectivity. Configuration v1.2 is the Balanced configuration ($\alpha = \beta = \gamma = 0.33$), which tests the hypothesis that equal weighting maximizes domain constraint activation. The selection objective is defined as the harmonic mean of four sub-metrics (3) to prevent any single sub-metric from strongly dictating the evaluation, thereby disincentivizing high precision at the expense of source coverage, or vice versa.

4. 3. Data sources and corpus characteristics

Four main corpora are used in the pipeline:

1) CPL datasets from two computing programs at Indonesian university (each containing 15 CPL items) in Bahasa Indonesia following KKNI/OBE framework at Universitas Sumatera Utara covering four KKNI competency domains; Attitude, Knowledge, General Skills, and Specific Skills for Sistem Informasi and Teknologi Informasi;

2) ESCO v1. 1 (2022: RDF-SKOS vocabulary with English labels of 13,939 entities;

3) ONET v28. 0 (2023): KSA-descriptive 879 computing-sector occupation profiles by major groups of ISCO-08;

4) SKKNI sektor ICT (Indonesia): 1,711 competency units inclusive of unit codes, titles, descriptions and assessment criteria. Interfacing text of ESCO/ONET objectives in the languages of Bahasa Indonesia-English would allow cross-lingual registration, forming keyword-augmented English representations for each CPL item and SKKNI unit from term mappings of APTIKOM study program classifications 2022 BK(Indonesia) database and CC2020 vocabulary. The original text and this one are both concatenated together to generate a bitext that is used as the TF-IDF query vector.

4. 4. Ablation study design

Six configurations were designed to systematically isolate the contribution of each scoring component by varying α, β, γ, and the acceptance threshold strategy τ (Table 1). The evaluation metric *selection_objective* is defined as the harmonic mean of four complementary sub-metrics

$$selection\_objective = \\ = H\begin{pmatrix} accept\_rate, src\_coverage, \\ mean\_S\_final, 1 - forced\_ratio \end{pmatrix}, \quad (1)$$

where *accept_rate* – percentage of source items with at least one accepted non-forced alignment; *src_coverage* – percentage of source items for which the top-1 acceptance mapping is non-forced; *mean_S_final* – average *S_final* of chosen to accepted non-forced alignments; and *forced_ratio* = percentage of source items for which system production was forced to select top-1 regardless threshold (when no candidate surpassed τ). The harmonic mean thus discourages configurations that are merely good on a subset of the sub-clustering metrics, encouraging balanced exploration-exploitation trade-offs. When all K candidates are under τ, *forced_top*1 = True is assigned by the acceptance gate which guarantees that the source will be fully traversed, however note that forced mappings are not used in *mean_S_final* but counted against forced_ratio as the downside being mapping quality violates quality.

Table 1

Ablation study: six weight configurations and design rationale

Config	α	B	γ	τ	Design rationale
v0.9	1.00	0.00	0.00	Median	Pure semantic baseline: *S_sem* only ablates graph and domain components
v1.0	0.60	0.25	0.15	Median	Weak domain constraint: $\gamma = 0.15$ insufficient to overcome *S_sem* dominance
v1.1	0.60	0.25	0.15	0 (none)	Zero threshold: all candidates accepted ablates acceptance gate effect
v1.2 ★	0.34	0.33	0.33	Median	BALANCED: equal weighting maximum domain constraint activation
v1.3	0.60	0.25	0.15	Q3 (0.75)	High precision: strict threshold tests coverage-precision trade-off
v1.4	0.55	0.30	0.15	Median	Graph-dominant: intermediate graph emphasis weak domain constraint

Note: ★ – optimal configuration; Q3 – third quartile of S_final distribution per task.

Table 1 shows six configurations to disentangle the contribution of each scoring component through weights variation. Configuration v0.9 defines the semantic baseline with $\beta = \gamma = 0$, which acts as a reference for hybrid configurations. Configurations v1.0 and v1.4 proposes *S_gr* and *S_con* with uniform weights to evaluate enhancements with respect to the baseline. Configuration v1.1 in which it removes the acceptance gate ($\tau = 0$) to sever the comparison with threshold filtering effects. Configuration v1.3 uses a more stringent Q3 threshold to assess precision-coverage at high selectivity. Configuration v1.2, the balanced configuration ($\alpha = \beta = \gamma = 0.33$) tests the hypothesis that equal weighting leads to maximization of domain constraint activation. The harmonic mean formulation of selection_objective (3) to discourage a single sub-metric from dominating the evaluation, thereby discouraging high precision at the cost of source coverage or vice versa.

Table 3

Domain filter stage 00 output: eligible ESCO skill distribution per study program

Program	Core ($S\_con = 1.0$)	Adjacent ($S\_con = 0.5$)	Total eligible	% of 13,939	Unique URIs	Core/Adj Ratio
SI – information systems	303	190	493	3.54%	493	1.59
TI – information technology	252	189	441	3.16%	441	1.33
CS – computer science	225	209	434	3.11%	434	1.08
SE – software engineering	280	165	445	3.19%	445	1.70
CE – computer engineering	258	185	443	3.18%	443	1.39
DS – data science	222	203	425	3.05%	425	1.09
Total / unique	1,540	1,141	2,681	19.23%	639 unique	1.35 avg

4. 5. Career readiness index formulation

CRI-KG is defined as a weighted sum of framework-specific coverage ratios

$$CRI(c\_i) = w\_E \cdot R\_ESCO(c\_i) + \\ + w\_O \cdot R\_ONET(c\_i) + w\_S \cdot R\_SKKNI(c\_i), \quad (2)$$

where $w\_E = 0.40$, $w\_O = 0.35$, $w\_S = 0.25$.

The coverage ratio for framework F is defined as $R\_F(c\_i) = n\_F\ /\ N\_F$, where n_F – the number of accepted non-forced mappings for CPL item c_i, and N_F – the number of possible targets after Stage 00 filtering in framework F. The component weights reflect the semantic richness and granularity of each framework: ESCO receives the highest weight ($w\_E = 0.40$) due to its skill-level granularity; O*NET receives $w\_O = 0.35$ for its occupation-level contextualization; SKKNI receives $w\_S = 0.25$ for its national qualification relevance to domestic employability. Items with $n\_F = 0$ across all frameworks receive $CRI = 0$ and are marked INCOMPLETE. Items with $n\_F = 0$ in at least one framework are marked PARTIAL. Items with $n\_F > 0$ across all three are marked COMPLETE.

4. 6. External consistency validation protocol

ECV determines whether two automatically generated crosswalks of a source item agree by using an authoritative external crosswalk. In chain 2, with sample size $n = 1{,}682$: For a SKKNI unit $u \in$ SKKNI the following holds true: $E(u)$ denotes top-K ESCO skills accepted in T4 (SKKNI → ESCO) and $O(u)$ indicates top-K ONET occupations accepted in T5 (SKKNI → ONET). The corresponding expected ONET set from the ESCO path is shown in (formula), where the crosswalk refers to the official ESCO-ONET occupation-level crosswalk (4,253 exact pairs, European Commission 2022). Consistency metrics are:

– exact consistency rate – the fraction of SKKNI units, which direct O*NET mapping (top-1 of $O(u)$) is found among the ESCO-path predictions: $|\{u: \hat{O}(u) \cap \mathrm{top1}(O(u)) \neq \emptyset\}|\ /\ n$;

– relaxed consistency rate – the fraction of units with at least one matching item, regardless of rank: $|\{u: \hat{O}(u) \cap O(u) \neq \emptyset\}|\ /\ n$;

– top-1 consistency rate – checks whether the system's highest-ranked prediction matches the unit's top-1 direct mapping: $|\{u: \mathrm{top1}(\hat{O}(u)) = \mathrm{top1}(O(u))\}|\ /\ n$;

– top-5 consistency rate – assesses whether any ESCO-path prediction appears among the top-5 direct mappings: $|\{u: \hat{O}(u) \cap \mathrm{top5}(O(u)) \neq \emptyset\}|\ /\ n$.

5. Results of IR-KG pipeline evaluation

5. 1. IR-KG model: structure and components

The IR-KG model is defined as a computational pipeline that aligns source competency descriptions against heterogeneous target frameworks by combining lexical retrieval, ontological graph signals, and curriculum domain constraints into a single scoring function producing ranked, interpretable mappings without requiring annotated training data. The IR-KG model has three parts that work together: a hybrid scoring function, a domain constraint pre-computation stage, and a seven-stage execution pipeline. The scoring function ranks candidate mappings; the pre-computation stage filters the target space before any mapping runs; the pipeline orchestrates everything from candidate generation to final output.

Structure. The scoring function S_final combines three signals semantic similarity S_sem, graph cohesion S_gr, and domain constraint S_con through a weighted linear combination. Stage 00 pre-computes S_con values from ISCED-F 2013 and APTIKOM 2022 classifications once, before any task executes. The seven-stage pipeline (Stages 01–06 + T10) then runs candidate generation, graph enrichment, hybrid scoring, ablation evaluation, ECV, and CRI-KG computation in sequence.

Parameters. Four parameters are configurable: weights α (S_sem), β (S_gr), and γ (S_con) subject to $\alpha + \beta + \gamma = 1$ and all ≥ 0; and acceptance threshold τ, set to median by default. Six weight configurations are compared in the ablation study (Table 2).

Input. Source descriptions from CPL (30 items, Bahasa Indonesia) or SKKNI (1,711 units, Bahasa Indonesia); target corpora from ESCO v1.1 (13,939 skills), O*NET v28.0 (879 occupations), or SKKNI; domain classification metadata from APTIKOM 2022 and ISCED-F 2013.

Output. For each source item: a ranked list of accepted target mappings with S_sem, S_gr, S_con, and S_final scores; a *forced_ratio* flag for audit; aggregate metrics including acceptance rate and source coverage per configuration; CRI-KG score per CPL item; and ECV consistency rates across 1,682 SKKNI units.

At the core of this IR-KG model is a three-component hybrid scoring formula that evaluates each candidate source-target pair (s,t) with a final alignment score:

$$S\_final(s, t) = \alpha \cdot S\_sem(s, t) + \beta \cdot S\_gr(t) + \gamma \cdot S\_con(t, p),$$

$$\text{subject to } \alpha + \beta + \gamma = 1,\ \ \alpha, \beta, \gamma \geq 0, \quad (3)$$

where p identifies the study programme under analysis. The three components capture orthogonal alignment signals:

1. S_sem semantic similarity score: TF-IDF cosine similarity of source bridge-text vector v_s versus target description vector v_t (TF-IDF weights are defined according to the standard formulation $w(\text{term, doc}) = (1 + \log tf(term, doc))$x x $\log(N\ /\ df(term))$, using $sublinear\_tf$ = True and *ngram_*

range = (1, 2) configuration to handle bigram technical phrases ('machine learning', 'cloud computing') Separate task-specific vectorizers (*max_features* = 15,000) were fitted on the full corpus of source and target documents for each task so that the vocabulary across tasks is aligned.

2. *S_gr* graph cohesion score: a structural signal based on the topology of the ESCO knowledge graph. Specifically, for the candidate skill node *t*, $S\_gr(t)$ represents how many of *t*'s ontological neighbors (which are the broader, narrower and related skill nodes within graph radius $r = 2$ from *t*) overlap with currently accepted mapping set A for same task. $Sgr(t) = |N(t, r = 2) \cap A| \ / \ |N(t, r = 2)|$ where $N(t, r)$ is the *r*-hop neighborhood of *t* in the ESCO skill graph encouraging mapping clusters to be more coherent and discouraging them from being isolated nodes. For ESCO-target tasks (T1 → T4), *S_gr* is used ($\beta = 0$ for O*NET, SKKNI); weights α, γ need to be renormalized.

3. *S_con* domain constraint score: a ternary relevance signal {1.0, 0.5, 0.0} reflects the degree of affinity between a target skill t and a study program *p* (in terms of relevant discipline field(*s*) to which each target skill fills criteria). Pre-computed in Stage 00, *S_con* is stored in a lookup table indexed by (program, target_URI). The *S_con* design operationalizes a core theoretical claim: that information regarding domain boundaries explicit in ISCED-F 2013 and APTIKOM 2022 curriculum ontology, but absent from standard IR models forms an independent alignment signal with discriminative properties.

For non-ESCO tasks where $\beta = 0$, the effective formula renormalizes to $S\_final = \alpha\_eff \cdot S\_sem + \gamma\_eff \cdot S\_con$, where $\alpha\_eff = \alpha / (\alpha + \gamma)$ and $\gamma\_eff = \gamma / (\alpha + \gamma)$.

Stage 00. Formal ISCED Whitelist text and TF-IDF domain similarity are used in a two-step process to construct *S_con* lookup table. This precomputation is performed once for each pipeline version and entails 13,939 skill×domain comparisons against six study program domain texts.

Step 1. ISCED Whitelist construction: a set *W* of 463 verified ESCO URIs for skills is obtained by identifying all nodes in ESCO representing skill groups with ISCED-F 2013 classification codes starting with '06' (Information and Communication Technologies) or '0714' (Electronics and Automation) or '0541' (Mathematics), then collecting all skill URIs, which broader URI relation points to a member of this group set. This URI-based extraction also avoids false positives due to substring matching that plagued earlier pipeline iterations (v1.x), which erroneously included skills where UUID hex fragment pseudorandomly began with '06'.

Step 2. Domain text similarity classification: for each of the six computing programs, a domain text *D_p* is created by concatenating the core_keywords a TF-IDF upweighting (repeating them twice) and the adjacent_keywords no repetition and gives documents with 291–340 tokens. Core keywords are drawn from the BK Penciri Pendukung (program-specific Study Materials) of APTIKOM 2022, while adjacent keywords are general literacy keywords sourced from the 21 BK Penciri Bersama (shared Study Materials) that apply across all computing programs. The TF-IDF similarity sim(*t*, *D_p*) is calculated for all the 13,939 ESCO skills and each program domain text. Classification follows the rule

$$S\_con(t,p) = \begin{cases} 1.0 \text{ if } t \in W \wedge sim(t,\ D\_p) \geq 0.050 [Core]; \\ 0.5 \text{ if } t \in W \wedge 0.015 \leq sim(t, D\_p) < 0.050 [Adjacent]; \\ 0.0 \text{ otherwise } [Outside]. \end{cases} \quad (4)$$

For non-whitelist skills ($t \notin W$), the condition $sim(t, D\_p) \geq 0.080$, is required to take any non-zero value, thus placing much higher demands on such a skill in the absence of ISCED certification (as an indicator of related experience). This gate, with two code paths leading to eligibility, generated 2,681 total eligible entries (639 unique skill URIs) across six programs, versus 697 (non-unique and domain-unverified) skills previously from the v1.x percentile-based approach.

The IR-KG pipeline is organized into seven sequential stages. Table 2 summarizes the name, key process, and output of each stage.

Table 2

IR-KG pipeline stages

Stage	Name	Key process	Output
01	Preprocessing	TF-IDF vectorization (8 task mirroring) using concatenated metadata fields as shared bridge texts; training data	8 vectorizer.pkl files
02	Candidate generation	Top-K cosine similarity retrieval over TF-IDF vector space (CPL: K = 20; SKKNI: K = 25), yielding 191,688 candidate pairs under six configurations and eight tasks	191,688 candidate pairs (6 configs × 8 tasks)
03	Graph enrichment	Neighborhood traversal–derived graph score (*S_gr*) from ESCO KG, enabling relational distance signals to be attached to each candidate pair	*S_gr* scores appended to candidate records
04	Hybrid scoring	Median-based acceptance gate (τ = median(*S_final*)) with three-component hybrid scoring ($S\_final = \alpha \cdot S\_sem + \beta \cdot S\_gr + \gamma \cdot S\_con$) per configuration-task combination	Accepted mapping lists per config x task
05	Ablation evaluation	When 6 configs and 8 tasks were systematically ablated across the combinations (48 in total; C[48]) a selection objective matrix was calculated for the data to extract an optimal weight configuration	Selection_objective matrix; optimal config identified
06 ECV	External consistency validation	Chain verification transitive (source → ESCO → O*NET) using official crosswalk; four metrics of consistency rate computed by chain. T10CRI-KG calculated superset from Stage 04 accepted mappings (conf. v1. 2), and aggregated to summary at program level	4 consistency rate metrics per chain
T10 CRI	Career readiness index	CRI-KG computed per CPL item from stage 04 accepted mappings (config v1.2)	Per-item CRI-KG; program-level summary

The pipeline ran across eight cross-framework mapping tasks under six weight configurations, generating 191,688 candidate pairs in total. Before any mapping task executed, Stage 00 pre-computed the S_con lookup table the results of which are presented below.

Stage 00 ran once against the full ESCO skill universe across all six computing programs. Table 3 shows how eligible entries distribute per program across core (*S_con* = 1.0) and adjacent (*S_con* = 0.5) categories.

The SI program produces the biggest eligible pool (493 skills, 3.54%), mirroring the wide interdisciplinary nature of Information Systems. With a score of 1.70, SE has the highest core/adjacent ratio and thus presents the most domain-focused skill profile corroborative with its focus on specialized software development. DS has the smallest eligible pool (425, 3.05%), due to higher terminological specificity for data science keywords (‘dimensionality reduction’, ‘gradient boosting’), which match fewer ESCO skill descriptions. Comparison with the previous v1.x eschewed false positives: domain-agnostic system + global top 5 percentile threshold (697 skills) bottom-up validation of 4 principles (omit v1-worthy skills include supervise aquaculture circulation system, cut wallpaper, and manage silverware collections). v2 correctly reasoned the x UUID hexadecimal prefix collisions with ISCED code ‘06’ are negated 1-aptikom2022.

All six configurations were evaluated against eight cross-framework mapping tasks. As shown in Table 4, the balanced configuration (v1.2, $\alpha = \beta = \gamma = 0.33$) achieves the highest mean selection_objective (0.537) across all tasks and is identified as the optimal configuration.

The mean gain from v0.9 to v1.2 = +26.1% (0.111 absolute) On the task level, gains vary between +3.2% (T1b) and +48.7% (T3b), supporting the hypothesis that the effect of *S_con* is larger on tasks with high mismatching in source language to target domain and corpus heterogeneity.

In Table 5, the incremental contribution of each design element is broken down.

Table 4

Selection objective scores across six configurations and eight tasks

Configuration	T1a	T1b	T2a	T2b	T3a	T3b	T4	T5	Mean
v0.9 (pure *S_sem*)	0.414	0.473	0.440	0.431	0.420	0.411	0.430	0.390	0.426
v1.0 (Weak *S_con*)	0.423	0.462	0.460	0.450	0.465	0.460	0.466	0.416	0.450
v1.1 (max accept)	0.447	0.455	0.491	0.489	0.500	0.510	0.445	0.489	0.478
v1.2 ★ (balanced)	0.447	0.488	0.573	0.555	0.601	0.611	0.493	0.530	0.537
v1.3 (precision τ)	0.428	0.470	0.397	0.397	0.477	0.495	0.455	0.372	0.437
v1.4 (mod. *S_gr*)	0.420	0.459	0.466	0.449	0.499	0.515	0.464	0.424	0.462

As shown in Table 5, the performance improvement of IR-KG is not caused by any single component. The increase from v0.9 to v1.0 (+5.6%) confirms that weak *S_con* activation at $\gamma = 0.15$ is insufficient to overcome *S_sem* dominance. The most impactful gain arises from increasing γ from 0.15 to 0.33 (v1.0 → v1.2, +19.3%), demonstrating that *S_con* requires equal weighting to fully activate domain boundary discrimination. The acceptance gate contributes an additional +12.3% (v1.1 → v1.2) by filtering low-quality candidates below the median threshold. Conversely, tightening the threshold to Q3 (v1.2 → v1.3) incurs a −18.6% penalty, confirming that strict precision filtering sacrifices source coverage. Collectively, these results demonstrate that the three-component structure of *S_final* and the median-based acceptance gate are both necessary and complementary for optimal alignment performance.

The accepted mapping statistics for all eight tasks with the best configuration v1 are shown in Table 6, including overall accepted mappings, forced counts and ratios, as well as mean and maximum per task *S_final* scores.

Table 6

Accepted mapping statistics under optimal configuration v1.2. T5 *forced_ratio* = 37.7% is a substantive finding indicating systematic SKKNI-O*NET semantic gap, not model failure

Task	Source items	Accepted total	Forced count	Forced %	Mean *S_final*	Max *S_final*
T1a: CPL-SI → ESCO	15	82	2	2.4%	0.244	0.603
T1b: CPL-TI → ESCO	15	104	1	1.0%	0.265	0.641
T2a: CPL-SI → O*NET	15	61	1	1.6%	0.515	0.782
T2b: CPL-TI → O*NET	15	60	2	2.5%	0.514	0.769
T3a: CPL-SI → SKKNI	15	151	1	0.7%	0.536	0.813
T3b: CPL-TI → SKKNI	15	150	1	0.7%	0.561	0.829
T4: SKKNI → ESCO	1,711	16,516	29	0.2%	0.240	0.712
T5: SKKNI → O*NET	1,711	8,011	646	37.7%	0.515	0.742

The *S_final* mean values (0.240–0.561) are not unrepresentative given that heterogeneous competency text yields expected TF-IDF cosine scores in this range which reflect the underlying terminological distance between frameworks. SKKNI-target tasks (T3a/T3b) obtain the best mean S_final values (0.536/0.561) due to their monolingual design (Indonesian-Indonesian) that avoids cross-lingual noise. The significantly high forced_ratio in T5 (37.7% compared to ≤2.5% for all other CPL-source tasks) indicates a structural semantic gap between SKKNI domain-specific Indonesian vocabulary, and the US-oriented O*NET occupation ontology of an observation with important implications for bilateral qualification recognition policy.

Table 5

Component ablation analysis: incremental contribution of each design element

Ablation comparison	Mean performance gain	Key insight
v0.9 → v1.0: addition of *S_gr* and *S_con* ($\gamma = 0.15$)	+0.024 (+5.6%)	Minimal *S_con* at $\gamma = 0.15$ insufficient; *S_gr* contributes marginally
v0.9 → v1.2: full *S_con* activation ($\gamma = 0.33$)	+0.111 (+26.1%)	Equal γ weight unlocks *S_con* discriminative power fully
v1.0 → v1.2: isolated *S_con* weight increase (γ: 0.15→0.33)	+0.087 (+19.3%)	Isolated effect of *S_con* weight increase without other changes
v1.2 → v1.3: threshold tightening (τ: median→Q3)	−0.100 (−18.6%)	Strict threshold sacrifices coverage; src_coverage drops sharply
v1.1 → v1.2: acceptance gate activation (τ: 0→median)	+0.059 (+12.3%)	Acceptance gate filters low-quality candidates, improves mean_*S_final*

5. 2. External consistency validation results

ECV results for chain 2 (SKKNI-based; $n = 1,682$ units) are presented in Table 7. Thus, all four consistency metrics outperform their random baselines by a factor of 8.7–9.4×, confirming that alignment signal capture is statistically significant above what would be expected due to chance alone.

Table 7

ECV chain 2 consistency metrics for 1,682 SKKNI units with Wilson score 95% CI

ECV metric	IR-KG v1.2	n consistent	Random baseline	Improvement factor	95% CI (approx.)
Exact consistency rate	14.3%	240 / 1,682	~1.6%	8.9×	[12.6%, 16.1%]
Relaxed consistency rate	27.1%	455 / 1,682	~3.1%	8.7×	[24.9%, 29.3%]
Top-1 consistency rate	13.1%	220 / 1,682	~1.4%	9.4×	[11.4%, 14.8%]
Top-5 consistency rate	23.5%	396 / 1,682	~2.7%	8.7×	[21.5%, 25.6%]

As evidenced in Table 7, the IR-KG alignment signal is statistically significant with an 8.7–9.4× gain over random baseline across four consistency metrics. The SKKNI → → ESCO → ONET chain holds semantic continuity strongly above random sampling. Given this mismatch in granularity of the ESCO-ONET crosswalk, a relaxed consistency rate of 27.1% means that over one quarter (26.5%) of SKKNI units are mapped, directly by their paths to ONET paths that have been identified as SKKNI → ONET mappings. The 95% confidence intervals indicate the level of stability across 1,682 units for these estimates.

5. 3. Career readiness index results and curriculum gap analysis

Table 8 presents CRI-KG results for all 15 CPL items of the Sistem Informasi (SI) program under config v1.2 ($w\_E = 0.40$, $w\_O = 0.35$, $w\_S = 0.25$).

Results for all 15 CPL items of the SI program in configuration v1 are shown in Table 8 – CRI-KG scores. 2. This relatively low overall CRI-KG mean score of 0.220 indicates that alignment was, in general, low to moderate with only 4 of the 15 items (26.7%) reaching complete status The prevalence *ig item* mapping is mostly at PARTIAL (60.0%) level, largely due to the expected "sea" of missing ESCO and O*NET mappings, while almost every item still has SKKNI coverage *si* 1 (SI_PLO-3) is classified as INCOMPLETE with CRI-KG = 0.000 SI program learning outcomes that exhibit better alignment with the national framework, as opposed to international standards, suggests that factors such as terminological divergence or a failure in cascade-subject translation may be an obstacle towards adapting curricula from internationally feasible competency frameworks.

A systematic comparison is presented in Table 9 to situate the IR-KG model performance against existing approaches using eight evaluation criteria based on study gaps identified. The four compared approaches are single-framework TF-IDF, ESCO knowledge graph alignment, BERT-based neural methods, and the proposed IR-KG model.

Table 8

CRI-KG scores per CPL item for the SI program under configuration v1.2

CPL ID	Domain	R_ESCO	R_ONET	R_SKKNI	CRI-KG	Level	Flag
SI_PLO-1	Attitude	0.214	0.510	0.527	0.396	Moderate	COMPLETE
SI_PLO-2	Attitude	0.000	0.000	0.528	0.132	Low	PARTIAL
SI_PLO-3	Attitude	0.000	0.000	0.000	0.000	Low	INCOMPLETE
SI_PLO-4	Knowledge	0.290	0.521	0.550	0.436	Moderate	COMPLETE
SI_PLO-5	Knowledge	0.278	0.515	0.538	0.426	Moderate	COMPLETE
SI_PLO-6	Knowledge	0.230	0.000	0.524	0.223	Low	PARTIAL
SI_PLO-7	General skills	0.000	0.000	0.521	0.130	Low	PARTIAL
SI_PLO-8	General skills	0.000	0.000	0.531	0.133	Low	PARTIAL
SI_PLO-9	General skills	0.000	0.000	0.527	0.132	Low	PARTIAL
SI_PLO-10	General skills	0.000	0.000	0.532	0.133	Low	PARTIAL
SI_PLO-11	General skills	0.238	0.508	0.521	0.403	Moderate	COMPLETE
SI_PLO-12	General skills	0.000	0.000	0.527	0.132	Low	PARTIAL
SI_PLO-13	General skills	0.075	0.507	0.563	0.348	Moderate	COMPLETE
SI_PLO-14	General skills	0.000	0.000	0.561	0.140	Low	PARTIAL
SI_PLO-15	General skills	0.000	0.000	0.523	0.131	Low	PARTIAL

Table 9

Systematic comparison of IR-KG against representative baseline approaches

Criterion	TF-IDF single	ESCO KG	BERT-based	IR-KG (proposed)
Framework coverage	Single	Single (ESCO)	Typically single	Triple: ESCO + O*NET + SKKNI
National standard integration	None	None	None	SKKNI ($w\_S = 0.25$, Stage 00)
Domainboundaryoperationalization	None	Manual/expert	Implicit via fine-tuning	Explicit 3-tier S_con (ISCED + APTIKOM)
Cross-lingual alignment	None	None	Via multilingual model	Bridge text + TF-IDF
Training data requirement	None	None	Labeled pairs required	None (unsupervised)
Interpretability	Term-level (high)	Relational (high)	Black-box (low)	Term-level + component weights (high)
Validation without ground truth	None	None	None	ECV transitive crosswalk
Computational requirement	CPU, seconds	CPU, seconds	GPU, hours (fine-tuning)	CPU, 103 s (191,688 pairs)
Reproducibility	Full	Full	Seed-sensitive	Fully deterministic

A systematic comparison is presented in Table 9 to situate the IR-KG model performance against existing approaches using eight evaluation criteria based on study gaps identified. The four compared approaches are single-framework TF-IDF, ESCO knowledge graph alignment, BERT-based neural methods, and the proposed IR-KG model.

6. Discussion of hybrid scoring, validation, and curriculum alignment findings

The model developed produced three results that need unpacking before moving to the broader implications. The balanced configuration ($\alpha = \beta = \gamma = 0.33$) outperforms the purely semantic baseline by 26.1%, with *S_con* alone responsible for +19.3% of that gain (Table 4, 5). The seven-stage pipeline completed 191,688 candidate pair evaluations with *forced_ratio* below 2.5% across all tasks except T5, where it reached 37.7% (Table 6). The domain filter (Stage 00) selected 2,681 eligible entries from 83,634 possible pairs roughly 3.2% (Table 3). What follows explains why these numbers look the way they do, and what they mean for competency alignment research and Indonesian curriculum policy.

The +19.3% gain from increasing γ from 0.15 to 0.33 has a straightforward mechanical explanation. At domain boundaries, *S_sem* loses its discriminating power two candidate skills can differ by just 0.001 in their similarity scores (0.187 vs. 0.188) yet sit on opposite sides of a relevance boundary. When *S_con* shifts from 0.0 to 1.0 for one of them, *S_final* jumps enough to flip the outcome from rejection to acceptance. This is not a side effect of the design it is the point. The three-tier *S_con* signal acts as a precision filter precisely where lexical similarity breaks down, encoding ISCED-F 2013 and APTIKOM 2022 discipline boundaries as computable constraints rather than binary gates. The task-level variation in gains (T3b: +48.7%, T1b: +3.2%) reflects this dynamic directly (Table 4): *S_con* contributes most when the baseline *S_sem* is already decent, as in monolingual Indonesian-to-Indonesian tasks where cross-lingual noise is absent and the model only needs a marginal discriminator to sharpen its selections.

The *forced_ratio* of 37.7% in Task T5 (SKKNI → ONET), versus ≤ 2.5% across all other tasks (Table 6), is the most substantively meaningful empirical finding within this study. The 646 SKKNI units receiving forced top-1 mappings are, ultimately, mostly competency units that delineate domain-specific bureaucratic and regulatory procedures SKKNI codes are usually joined with KBLI sectoral classifications and Indonesian administrative terminology for which there exists no ONET occupation at comparable specificity. This is neither a failure of the IR-KG pipeline, which appropriately reflects that sufficient mapping cannot be derived the forced_ratio signal makes this transparent, nor a lack of coverage depth for this study, but rather visible [[*structural*]] ontological divergence between Indonesia's unit-based SKKNI taxonomy versus the occupation-based ONET paradigm. The finding has direct policy implications, quantifying the volume of existing SKKNI units that currently lack an empirically validated international counterpart in ONET, thus adding empirical weight to Indonesia's continuing national agenda for internationalizing SKKNI and negotiating recognition of bilateral qualifications.

Table 9 presents a systematic comparison with representative baseline approaches. In contrast, single-framework TF-IDF [11] or ESCO Knowledge Graph methods [5] do not provide national standard integration, cross-lingual alignment, nor ground-truth free validation and BERT-based approaches [14, 15], which rely on labeled training pairs produce uninterpretable black-box scores. As shown in Table 9; the IR-KG model addresses all eight of these criteria simultaneously. In the context of low-resource institutions, it provides reproducibility and scalability advantages that are impractical to achieve with transformer-like GPU-dependent methods: the fully deterministic, CPU-only pipeline (103 sec. for 191688 candidate pairs).

This paper has four contributions that advance existing work as follows. As a unique contribution, CPL unifies skill-based (ESCO), occupation-based (O*NET), and national unit-based (SKKNI) on the same scale to contrast tri-laterally in individual scoring as a problem formulation addressing combinatorial complexity not found previously in two-framework studies by [7, 8]. The three-tier *S_con* (based on the continuous examination and relevance process from [3]) operationalizes domain boundary knowledge from ISCED-F 2013 and APTIKOM 2022 into an unbounded recyclability score, providing a support aligned scoring signal between ground truth and potential map Lorenzo between sliding targets replacing the binary preprocessing periodization filters used in previous filtering work. Single-framework or aggregate coverage metrics mask a framework-specific gradient revealed in the CRI-KG (R_SKKNI >> R_ONET > R_ESCO). The ECV method is outperforms random baseline by 8.7–9.4× under the four consistency metric with no annotated ground truth and is demonstrated to be generalizable over any cross-framework alignment system with a valid external crosswalk.

Table 8 shows a consistent gradient R_SKKNI >> >> R_ONET > R_ESCO across both programs, and three structural reasons explain it. R_SKKNI is highest because the mapping is monolingual (Indonesian → Indonesian), SKKNI targets the ICT sector directly, and its unit descriptions share the outcome-based narrative format of CPL all three reduce noise. R_ONET sits in the middle: occupation profiles are broad enough to catch more CPL items than ESCO's skill-level descriptions, but not specific enough to match the depth of SKKNI. R_ESCO is lowest for a straightforward reason 13,939 entities filtered down to ≤ 3.5% eligible targets per program by Stage 00 leaves very little surface for CPL items to land on.

The persistent PARTIAL and INCOMPLETE flags for attitude and general skills items in both programs (Table 8) point to a real curriculum problem. Generic, attitudinal learning outcomes the kind that read well in an accreditation document produce almost no matches in ESCO or O*NET because neither framework has a category for "demonstrates professional ethics" or "communicates effectively." These CPL items need domain-specific, skill-oriented sub-competences to become computationally visible.

These findings are subject to three limitations that limit generalizability. TF-IDF is a bag-of-words model, at odds with the need for contextual neural embeddings to capture even obvious forms of semantic equivalence, and thus while yielding reasonable performance here given low-resource interpretability-constrained setting, future work should consider hybrid retrieval-reranking pipelines that leverage multilingual transformer architectures like XLM-R or mBERT. Second, ECV has an inherent limitation due to the occupation-level granularity of the ESCO-O*NET crosswalk that maps occupations to occupations rather than skills to skills leading to a mismatch in granularity that will automatically suppress consistency scores. Third, the CPL corpus stems from two computing programs across a single institution, making immediate generalizability to other institutions, disciplines or national contexts challenging. The chief drawback of the existing run namely, a missing neural reranking phase can be treated in future work by adding a light cross-encoder to rank candidates according to silver-standard pairs generated from the ECV mechanic itself; an approach that is self-supervised and would not involve human annotation.

7. Conclusions

1. A three-component hybrid scoring function $S\_final = \alpha \cdot S\_sem + \beta \cdot S\_gr + \gamma \cdot S\_con$ was defined and implemented in which TF-IDF semantic similarity, ESCO Knowledge Graph cohesion, and a domain constraint score base on ISCED-F 2013 and APTIKOM 2022 curriculum classifications were integrated together. Evaluation on eight cross-framework mapping tasks finds that the balanced configuration ($\alpha = \beta = \gamma = 0.33$) achieves a mean selection_objective of 0.537 representing a 26.1% improvement over the pure semantic baseline (v0.9). The sole contribution of *S_con* explains +19.3% of this gain (v1.0 → v1.2), confirming that the proposed domain boundary signals operationalized from curriculum ontologies are an independent and discriminative alignment component, thus differentiating the IR-KG approach from previous works employing single-signal TF-IDF and KG-only schemes.

2. An end-to-end pipeline over seven stages was established and executed for eight cross-framework and cross-lingual mapping tasks CPL-ESCO, CPL-ONET, CPL-SKKNI, SKKNI-ESCO, and SKKNI-ONETs resulting in 191,688 candidate pairs across six weight configurations. The ablation study reveals that the median-based acceptance gate adds +12.3% of independent gain, and that v1.2 → v1.3) incurs a −18.6% penalty, suggesting that balanced acceptance is jointly optimal with component weights of $n1 = n2$. *Forced_ratio* for Task T5 (SKKNI → ONET) is 37.7% vs. ≤ 2.5% for all other tasks and represents a significant finding: it quantifies the structural ontological divergence between Indonesia's unit-based SKKNI taxonomy, according to units of competency and the occupation based ONET paradigm, that potentially informs bilateral qualification recognition policy negotiations as concrete empirical evidence.

3. The career readiness index (CRI-KG) was calculated from 30 CPL items across two computing programs, showing a systematic gradient R_SKKNI >> R_ONET > R_ESCO with COMPLETE rates of 26.7% (SI) and 33.3% (TI), indicating that Indonesian computing curricula were meaningfully more aligned with the national framework than other international taxonomies. External consistency validation applied to $n = 1{,}682$ SKKNI units yielded a relaxed consistency rate of 27.1% (95% CI: 24.9–29.3%), an 8.7× improvement over the random baseline. The SKKNI alignment is monolingual advantage which explains these results and the ESCO domain filtering has granularity constraints.

Financing

The study was performed without financial support.

Data availability

Data will be made available on reasonable request. CPL datasets are available from the corresponding author; ESCO and O*NET data are publicly available from their respective institutional repositories.

Use of artificial intelligence

The authors declare the use of AI for grammar and typographical error checking under full human supervision. The authors bear full responsibility for all manuscript content.

Authors' contributions

Halim Maulana: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Writing – Original Draft, Writing – Review & Editing, Visualization; **Poltak Sihombing**: Conceptualization, Supervision, Writing – Review & Editing; **Amalia**: Supervision, Writing – Review & Editing, Project Administration; **Marischa Elveny**: Validation, Writing – Review & Editing, Formal Analysis.

References

1. Acemoglu, D., Restrepo, P. (2019). Automation and New Tasks: How Technology Displaces and Reinstates Labor. Journal of Economic Perspectives, 33 (2), 3–30. https://doi.org/10.1257/jep.33.2.3
2. Autor, D. H. (2019). Work of the Past, Work of the Future. AEA Papers and Proceedings, 109, 1–32. https://doi.org/10.1257/pandp.20191110
3. The Future of Jobs Report 2023 (2023). Available at: https://www.weforum.org/reports/the-future-of-jobs-report-2023
4. OECD Employment Outlook 2021 (2021). In OECD Employment Outlook. OECD Publishing. https://doi.org/10.1787/5a700c4b-en
5. European Skills, Competences, Qualifications and Occupations (ESCO). ESCO v1.1. Available at: https://esco.ec.europa.eu/en/about-esco/escopedia/escopedia/esco-v11
6. Occupational Listings. O*NET 28.0 Database – occupations updated with job incumbent or occupational expert (OE) data. Available at: https://www.onetcenter.org/listings/28.0/updated.html
7. Neutel, S., de Boer, M. H. T. (2021). Towards Automatic Ontology Alignment using BERT. Proceedings of the AAAI 2021 Spring Symposium on Combining Machine Learning and Knowledge Engineering (AAAI-MAKE 2021). Available at: https://ceur-ws.org/Vol-2846/paper28.pdf
8. Snijder, L. L., Smit, Q. T. S., De Boer, M. H. T. (2024). Advancing Ontology Alignment in the Labor Market: Combining Large Language Models with Domain Knowledge. Proceedings of the AAAI Symposium Series, 3 (1), 253–262. https://doi.org/10.1609/aaaiss.v3i1.31208

9. Panduan Kurikulum Berbasis OBE/KKNI/SKKNI APTIKOM: Program Studi Sarjana Sistem Informasi, ver. 1.1 (2022). Jakarta: APTIKOM. Available at: http://elibrary.almaata.ac.id/3475/1/19%20Jan%202023%20V1.1%20-%2013%20jan%202023%20-%20FORDI%20SI%20-%20BUKU%20KURIKULUM%20BERBASIS%20OBE.pdf
10. European Commission: Directorate-General for Employment, Social Affairs and Inclusion (2017). ESCO handbook: European skills, competences, qualifications and occupations. Publications Office. https://doi.org/10.2767/934956
11. Manning, C. D., Raghavan, P., Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press. https://doi.org/10.1017/cbo9780511809071
12. Robertson, S., Zaragoza, H. (2009). The Probabilistic Relevance Framework: BM25 and Beyond. Foundations and Trends® in Information Retrieval, 4 (1-2), 1–174. https://doi.org/10.1561/1500000019
13. Fareri, S., Melluso, N., Chiarello, F., Fantoni, G. (2021). SkillNER: Mining and mapping soft skills from any text. Expert Systems with Applications, 184, 115544. https://doi.org/10.1016/j.eswa.2021.115544
14. Decorte, J.-J., Van Hautte, J., Demeester, T., Develder, C. (2021). JobBERT: Understanding Job Titles Through Skills. arXiv. https://doi.org/10.48550/arXiv.2109.09605
15. Chiarello, F., Fantoni, G., Hogarth, T., Giordano, V., Baltina, L., Spada, I. (2021). Towards ESCO 4.0 – Is the European classification of skills in line with Industry 4.0? A text mining approach. Technological Forecasting and Social Change, 173, 121177. https://doi.org/10.1016/j.techfore.2021.121177
16. Euzenat, J., Shvaiko, P. (2013). Ontology Matching. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-38721-0
17. Computing Curricula 2020: Paradigms for Global Computing Education (2021). Association for Computing Machinery. https://doi.org/10.1145/3467967
18. Chen, J., Jiménez-Ruiz, E., Horrocks, I., Antonyrajah, D., Hadian, A., Lee, J. (2021). Augmenting Ontology Alignment by Semantic Embedding and Distant Supervision. The Semantic Web, 392–408. https://doi.org/10.1007/978-3-030-77385-4_23
19. Savickas, M. L. (2005). The Theory and Practice of Career Construction. Career Development and Counseling: Putting Theory and Research to Work. Wiley, 42–70.
20. Development and Validation of the NACE Career Readiness Competencies (2022). NACE. Available at: https://www.naceweb.org/uploadedFiles/files/2022/resources/2022-nace-career-readiness-development-and-validation.pdf
21. Reimagining our futures together: a new social contract for education (2021). UNESCO. https://doi.org/10.54675/asrb4722
22. European skills and jobs survey (ESJS). CEDEFOP. Available at: https://www.cedefop.europa.eu/en/projects/european-skills-and-jobs-survey-esjs#:~:text=The%201st%20ESJS%20collected%20information,facilitated%20numerous%20research%20studies%2C%20including:
23. Hevner, A. R., March, S. T., Park, J., Ram, S. (2004). Design Science in Information Systems Research1. MIS Quarterly, 28 (1), 75–106. https://doi.org/10.2307/25148625