A Few-Shot Semantic Meta-Learning Framework with CRF for Skill Entity Recognition in Open Innovation Ecosystems

Nurchim[1], Muljono[2], Edi Noersasongko[3], Ahmad Zainul Fanani[4], Deshinta Arrova Dewi[5]
Doctoral Program of Computer Science, Universitas Dian Nuswantoro, Semarang, Indonesia[1, 2]
Faculty of Computer Science, Universitas Duta Bangsa Surakarta, Surakarta, Indonesia[1]
Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang, Indonesia[2, 3, 4]
Faculty of Data Science and Information Technology, INTI International University, Malaysia[5]

***Abstract*—The accelerating pace of digital transformation is reshaping labour-market dynamics, driving the emergence of new competencies, and intensifying the need for scalable skill-intelligence systems within open innovation ecosystems. Yet, research on Indonesian Named Entity Recognition (NER) remains limited for skill-extraction tasks, especially in low-resource contexts where annotated data are scarce and novel skill expressions evolve rapidly. To address this gap, this study contributes to applied Natural Language Processing (NLP) by introducing the Few-Shot Semantic Meta-Learning framework with CRF (FSM-CRF) for Indonesian skill entity recognition, which integrates semantic span representations, episodic meta-learning, and BIO-constrained CRF decoding to enhance prototype stability and entity-boundary precision for complex, multi-token skill expressions. Using the NERSkill.id dataset, the proposed model is evaluated under a 3-way, 10-shot episodic setting and achieves a micro-F1 of 73.84%, outperforming traditional supervised approaches (IndoBERT fine-tuning, BiLSTM-CRF) and existing few-shot baselines. Ablation experiments further demonstrate that semantic span modelling and structured CRF inference play pivotal roles in improving robustness, while meta-learning strengthens adaptability across diverse and evolving skill categories. From an open innovation perspective, this framework offers a data-efficient solution for dynamic competency mapping, reducing dependence on costly annotation pipelines and enabling continuous updates to workforce skill taxonomies. Overall, the findings highlight semantic meta-learning as a promising foundation for next-generation skill-intelligence infrastructures that support AI-enabled innovation management, strategic workforce planning, and evidence-informed policy design.**

Keywords—Few-shot; Named Entity Recognition; skill intelligence; process innovation; open innovation; Natural Language Processing

I. INTRODUCTION

The rapid acceleration of digital transformation has reshaped how organizations, industries, and governments understand, cultivate, and manage human capabilities, as digital technologies reconfigure work processes and employment structures [1], [2], [3]. As emerging technologies continually redefine job roles and competency requirements, skill intelligence systems designed to detect, map, and update workforce skills in real-time have become a critical component of open innovation ecosystems [4], [5]. In this environment, organizations must rely on agile, data-driven mechanisms to monitor labour-market trends, identify emerging competencies, and support continuous learning and upskilling [6], [7]. However, traditional approaches that depend on manual curation of skill taxonomies or large-scale labelled datasets are increasingly unsustainable, especially in dynamic contexts where new skills evolve faster than they can be annotated or embedded into curricula [1], [2]. Gayatri et al. [8] show that this challenge is particularly evident in Indonesian digital labour markets, where projected gaps in advanced digital competences persist through 2025, while Prasetyo [9] illustrates how platform-mediated work and informal-economy dynamics in Indonesia accelerate the diversification of technical and digital skills beyond the reach of conventional competency-based education systems. Together, these findings highlight the urgency of building scalable skill intelligence infrastructures that can operate with minimal manual labelling while remaining sensitive to rapidly evolving, locally embedded skill configurations in digital labour platforms and education ecosystems [1], [7], [8].

Despite growing interest in automated skill extraction, the development of robust NER systems for skill-related Indonesian text remains limited, with existing work on Indonesian NER still concentrating mainly on news or cross-linguistic corpora rather than fine-grained occupational skills [10], [11]. Skill entities in these contexts frequently surface as multi-word expressions such as troubleshooting network, cloud security management, or analysis perform server whose non-compositional and domain-specific semantics are known to challenge sequence labelling models when multiword expressions are not explicitly modelled [12]. These characteristics become even more problematic under low-resource conditions, where only a few annotated examples are available for each skill type and conventional architectures struggle to learn reliable decision boundaries [13], [14]. State-of-the-art NER approaches based on BiLSTM-CRF and Transformer fine-tuning, including IndoBERT-style models, still rely heavily on large, domain-specific annotated corpora and exhibit reduced generalization when confronted with domain-specific phrasing, rich morphological variation, and heterogeneous syntactic structures, as observed in job advertisements, legal/technical documents, and other specialized corpora [15], [16], [17].

Few-shot learning offers a promising alternative for skill entity recognition because it allows models to generalize from only a handful of labelled examples [18], [19], [20]. However, prototype-based few-shot NER methods such as Prototypical Networks and their variants remain highly sensitive to multi-token span representations, which often leads to unstable prototypes and inaccurate entity boundaries in complex skill phrases [19], [21], [22]. Even when enhanced with contrastive or prompt-based learning, many few-shot NER architectures still underutilize structured sequence information, resulting in inconsistent BIO label transitions and degraded precision on span-level prediction [20], [23]. To date, only a limited body of work has jointly integrated semantic span representations, dynamic prototype refinement, and structured decoding layers (e.g., CRF) within a unified framework, and virtually none of these studies has been tailored to Indonesian SkillNER scenarios. At the same time, research on data-efficient NER for skills has only begun to explore its role in labour-market analytics and innovation systems, despite evidence that automated skill extraction can support soft-skill mapping, competency taxonomies, and workforce analytics for open innovation and Industry 4.0 ecosystems [24], [25].

To address these gaps, this study proposes a Few-Shot Semantic Meta-Learning framework with CRF (FSM-CRF) for Indonesian Skill Entity Recognition, building on recent advances in few-shot NER and meta-learning for low-resource settings [18], [19], [26], [27]. The approach integrates three complementary components: 1) semantic span representations derived from IndoBERT to model the rich semantics and label interactions of multi-word skill expressions, inspired by span- and label-aware few-shot architectures [23], [28], [29]; 2) episodic meta-learning to construct stable class prototypes and enhance generalization under limited supervision, in line with recent prototypical and contrastive few-shot NER methods [26], [27], [30]; and 3) BIO-constrained CRF decoding to enforce structural consistency and improve entity boundary detection, following evidence that structured decoding remains beneficial in sequence labelling for NER [31], [32]. Using the NERSkill.id dataset containing annotated Indonesian skill entities across HSkill, SSkill, and Tech categories [33] and extending prior work on Indonesian NER in disaster, financial, and misinformation domains [32], [34], [35]. The framework is evaluated through comprehensive experiments, including baseline comparisons, ablation studies, and statistical significance testing, as recommended in recent few-shot NER evaluations. The results demonstrate that the proposed model consistently outperforms traditional supervised and state-of-the-art few-shot baselines under scarce per-class supervision, achieving notable gains in F1 for complex multi-token skill spans while maintaining BIO-consistent predictions.

The contributions of this study are fourfold. First, it introduces a novel semantic meta-learning architecture specifically designed for low-resource Indonesian SkillNER. Second, it establishes a new benchmark by constructing and analysing the NERSkill.id dataset, which is organised around realistic workforce skill categories. Third, it presents a rigorous empirical evaluation that combines standard performance metrics, ablation studies, and Friedman–Nemenyi statistical testing to clarify the respective roles of the semantic, meta-learning, and structured decoding components. Finally, from an open innovation perspective, this work demonstrates how data-efficient NER models can strengthen adaptive skill intelligence systems, reduce reliance on costly annotation pipelines, and support dynamic workforce innovation strategies. Although the individual components employed in this study semantic span representations, episodic meta-learning, and CRF-based structured decoding have been examined in earlier research, their combination within FSM-CRF goes beyond a straightforward assembly of existing techniques. In this work, Indonesian SkillNER is treated as a span-centric learning problem under few-shot conditions, where the formation of class representations, the identification of entity boundaries, and the enforcement of label consistency are addressed in a unified manner.

Unlike prior span-based or few-shot NER approaches that largely rely on representation similarity alone, or structured decoding methods that are applied after fixed token-level predictions, FSM-CRF brings these processes together during learning. This integration enables more stable handling of multi-token skill expressions while maintaining coherent sequence structure throughout inference. In practical terms, it helps mitigate common failure cases observed in low-resource settings, such as unstable class representations and fragmented entity boundaries. The resulting framework therefore supports more reliable skill extraction in Indonesian texts, not by introducing isolated components, but by reshaping how they interact under limited supervision. Taken together, the core contribution of this work lies in enabling stable, span-level skill extraction under few-shot conditions with structured sequence consistency a capability that prior Indonesian and few-shot SkillNER systems could not achieve reliably while remaining transferable to other low-resource and dynamically evolving skill domains beyond the specific dataset or national context studied here.

The remainder of this study is organised as follows: Section II reviews related work. Section III describes the proposed method. Section IV outlines the experimental setup and presents and discusses the results. Finally, Section V concludes with theoretical, practical, and policy implications.

II. Related Works

A. Skill Intelligence and Digital Competencies

As digital transformation accelerates, the need for advanced skill-intelligence systems has become increasingly evident. Contemporary studies emphasize that organizations now depend on data-driven mechanisms to identify emerging competencies, support continuous learning, and sustain competitiveness within open innovation ecosystems [1], [4]. Traditional approaches, such as manually curated skill taxonomies or extensively annotated datasets, are no longer adequate in environments where new skills evolve rapidly and unpredictably [4]. In Indonesia, these challenges are even more pronounced due to persistent digital skill gaps and the growing influence of platform-based and informal labour markets [8], [36], [37]. From an open innovation perspective, the ability to extract skill-related information from unstructured text is essential for enabling timely, evidence-based decision-making [25], [38]. Consequently, NLP techniques capable of operating with limited supervision play an increasingly strategic role in supporting dynamic competency mapping across innovation-driven ecosystems [14], [26].

B. *Indonesian Named Entity Recognition*

Research on Indonesian Named Entity Recognition (NER) has expanded across various domains such as news, disaster management, finance, and other publicly accessible texts, employing models like Bi-LSTM–CRF and transformer-based fine-tuning (e.g., IndoBERT) [10], [34], [39]. However, despite this growth, these efforts predominantly target general entity types (persons, organizations, locations), while the domain of skill extraction remains largely unexplored. Skill entities pose additional complexity: unlike conventional NER categories, skill expressions often manifest as multi-word, semantically rich phrases such as troubleshooting network or cloud security management, which require deeper contextual and compositional understanding. Although datasets like NERSkill.id aim to provide foundational resources for Indonesian SkillNER [33], conventional token-level NER methods struggle to represent the full complexity and variability of skill-related language. This gap underscores the need for models specifically optimized for skill extraction, especially in innovation-oriented environments where timely and accurate competency insights are essential.

C. *Few-Shot Named Entity Recognition Approaches*

Few-shot NER methods have gained significant attention for their ability to generalize from minimal annotated data a capability increasingly important in innovation-driven domains where new concepts and competencies continuously emerge [18], [40]. Meta-learning and prototypical-network approaches aim to build class-level representations (prototypes) that support rapid adaptation to new entity categories [18], [27]. Despite their promise, existing few-shot approaches encounter substantial limitations when applied to multi-token or domain-specific entities. Prototype representations often become unstable, resulting in inaccurate span boundaries and misclassification of entity spans, especially when prototypes are too closely distributed, or label dependency is roughly estimated [17], [41]. Although more recent techniques that incorporate prompting or contrastive learning have improved representation quality [19], [42], they frequently overlook structural constraints such as consistent BIO-label transitions or sequence-level dependencies. Notably, the literature lacks a unified approach that combines semantic span modelling, episodic meta-learning, and structured decoding (e.g., CRF), especially tailored for specialized domains such as skill extraction in low-resource languages. This absence represents a significant research gap in developing adaptable, data-efficient competency extraction methods compatible with open innovation and dynamic labour-market ecosystems.

D. *Semantic, Meta-Learning, and Conditional Random Field (CRF)*

Semantic span representations have emerged as a robust alternative to token-level modelling because they better capture the internal structure and meaning of multi-word entities, an advantage especially relevant for complex or domain-specific phrases common in skill-related text [29], [43], [44]. However, their integration within few-shot learning frameworks remains limited even though semantic grounding is essential for forming stable and generalizable prototypes under data scarcity [18], [45]. Meta-learning offers strong potential for rapid adaptation, but without rich semantic span features, prototypes often become inconsistent especially for long or heterogeneous skill expressions which can severely degrade boundary detection accuracy [17], [22].

In parallel, Conditional Random Field (CRF)-based decoding remains highly effective for enforcing BIO-consistent label transitions and ensuring structurally coherent predictions, helping mitigate over-prediction and improving reliability when dealing with multi-token, semantically dense entity spans [46], [47]. Despite the promise shown by each component semantic span modelling, meta-learning, and structured decoding the existing literature typically treats them in isolation. To date, no study (to our knowledge) has proposed a unified few-shot framework that integrates all three components, and none is tailored to the context of Indonesian SkillNER. The present research fills this gap by combining semantic span representations, episodic meta-learning, and BIO-constrained CRF decoding into a cohesive architecture, an integration that we argue enhances prototype stability, improves boundary precision, and strengthens overall robustness, thereby supporting the responsiveness and knowledge adaptiveness central to open-innovation frameworks.

III. PROPOSED METHOD

This section presents a Few-Shot Semantic Meta-Learning Framework with CRF (FSM-CRF) for Indonesian Skill Entity Recognition. The framework (see Fig. 1) integrates BIO-to-span reformulation, semantic prototype modelling, episodic meta-learning, and CRF-based decoding. The methodological design is informed by established practices in deep-learning NER [14], [48], [49], annotation scheme studies [50], and Indonesian transformer-based NER research [51], [52].

A. *Data Preparation and Pre-processing*

1) BIO labelling and reformulation: The Begin–Inside–Outside (BIO) tagging scheme, which is widely used in sequence labelling studies surveyed by Seow et al. [53], is adopted to explicitly represent the boundary structure of skill entities, which in Indonesian texts frequently appear as multi-token expressions rather than isolated terms.

Within this scheme, the B label indicates the beginning of a skill mention, I marks tokens that continue the same skill span, and O is assigned to tokens that do not belong to any skill entity. This explicit distinction allows the model to separate entity boundaries from internal token composition, which is essential for accurately capturing complex skill expressions.

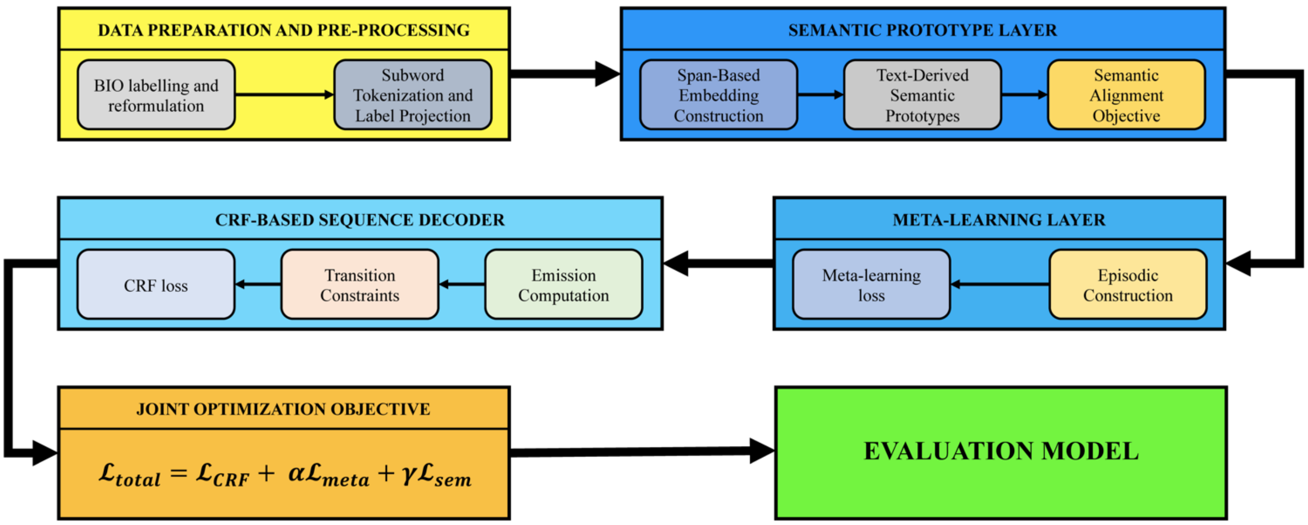

Fig. 1. A few-shot semantic meta-learning framework with CRF for an Indonesian skill entity recognition model.

To illustrate this process, consider the sentence "memahami step troubleshooting network". The token "troubleshooting" is labelled as B-HSkill, followed by "network" as I-HSkill, reflecting that both tokens together form a single Hard Skill entity. In contrast, tokens such as "memahami" and "step" receive the O label. This token-level annotation preserves the internal structure of skill expressions while clearly defining their boundaries.

The BIO scheme is chosen primarily for its robustness and practicality in low-resource annotation settings. Compared to more elaborate schemes such as BILOU, BIO offers a simpler and more consistent annotation process, a trade-off that is commonly recommended in recent NER surveys such as Hu *et al.* [49] when annotation consistency and boundary ambiguity are key concerns. In addition, BIO integrates naturally with structured sequence decoding, allowing label consistency to be enforced during inference without increasing annotation complexity.

NER data are initially represented as token–label pairs:

$$x = \{w_1, \dots, w_2\}, y = \{t_1, \dots, t_2\}$$

where, each label follows the BIO scheme:

$$t_1 \in \{O, B - E_k, I - E_k\}, \qquad E_k \in \{HSkill, SSkill, Tech\}$$

The BIO scheme has been shown to influence NER performance depending on annotation design [48], [50]. To enable span-level reasoning, sequences of BIO labels are reformulated into entity spans:

$$s = \{i, j, E_k\}, \qquad j = \max\{n : t_n = I - E_k\}$$

yielding, the mapping:

$$f_{BIO \to span}(y) = \{i, j, E_k\} \mid t_1 = B - E_k, t_{i+1:j} = I - E_k$$

This reformulation consolidates multi-token constructs such as $B - HSkill + I - HSkill$ into a single coherent span entity $HSkill$, $B - SSkill + I - SSkill \rightarrow SSkill$, and $B - Tech + I - Tech \rightarrow Tech$, consistent with span-level methodologies in contemporary NER [54], [55].

2) Subword tokenization and label projection: Indonesian text is encoded using a transformer-based subword tokenizer (e.g., IndoBERT), following established practices in Indonesian and low-resource NER [51], [52], [56]. Tokenization yields a mapping:

$$word\_ids[m] = \begin{cases} i, & subword\ m\ originates\ w_1, \\ none, & otherwise \end{cases}$$

BIO labels are projected onto subword units:

$$t_m^{sub} = \begin{cases} t_{word\_ids[m]}, & word\_ids[m] \neq none, \\ O, & otherwise \end{cases}$$

The output of pre-processing is a structured representation:

$$x^{sub}, y^{sub}, span, word\_ids$$

which maintains alignment between spans, subword embeddings, and sequence-level decoding.

B. Semantic Prototype Layer

1) Span-based embedding construction: Each span $s = \{i, j, E_k\}$ is converted into a semantic embedding based on token-level contextual states:

$$v_{span} = \frac{1}{|T_{span}|} \sum_{t \in T_{span}} h_t$$

Span embeddings are widely used in span-based NER due to their representational robustness [49], [55].

2) Text-derived semantic prototypes: Each entity label is verbalized as:

$$d_k = \text{"}Label:\ E_k\text{"}$$

and encoded using the same transformer encoder:

$$p_k^{test} = Encoder(d_k)[CLS]$$

Label-descriptive prototypes leverage semantic priors from pretrained language models, consistent with recent advances in semantic-aware NER [14], [49].

3) Semantic alignment objective: Span embeddings and text-derived prototypes are aligned using a consistency objective. For each span embedding v_i:

$$s_i^{data} = v_i . p^{data}, \qquad s_i^{test} = v_i . p^{test}$$

with probability distributions:

$$q_i = softmax(s_i^{data}), \qquad r_i = softmax(s_i^{test})$$

and their mixture:

$$m_i = \frac{q_i + r_i}{2}$$

The semantic alignment loss is:

$$\mathcal{L}_{sem} = \frac{1}{2}[KL(q_i || m_i)] + [KL(r_i || m_i)]$$

encouraging representational coherence across data-driven and label-informed semantic spaces.

The semantic fusion mechanism is designed to integrate evidence from annotated data with prior semantic information derived from label descriptions, an idea aligned with recent prototype-based few-shot NER studies that explicitly bridge span representations and textual type descriptions [21]. During each training episode, span representations extracted from the support set are first aggregated to form provisional class representations. These representations are then aligned with label-level embeddings obtained from textual descriptions of each skill category, where the use of semantic constraints to stabilize learning under scarce supervision has been shown to be effective in related settings [57].

Through this alignment process, the resulting class representations capture both observed patterns in the data and broader semantic characteristics associated with each skill type. This combined representation provides a more stable basis for classification, particularly when dealing with rare or newly emerging skill expressions that may not appear frequently in the training data; prototype ambiguity under limited samples has been repeatedly highlighted as a key challenge in the literature.

C. Episodic Meta-Learning Layer

FSM-CRF employs N-way K-shot episodic meta-learning, an approach widely used in few-shot NLP tasks and beneficial for low-resource NER [49], [56]. Each training episode follows an N-way K-shot setting that reflects realistic low-resource learning conditions. For instance, in a 3-way 10-shot configuration, three skill categories, HSkill, SSkill, and Tech, are selected, and ten support instances are sampled for each category. A separate set of query sentences is then used to evaluate the model's predictions.

Within each episode, span representations from the support set are used to construct category-level representations, which are subsequently applied to classify spans in the query set. By repeatedly exposing the model to such episodic tasks, the learning process encourages rapid adaptation across different skill categories and supports generalization in scenarios where annotated examples are limited.

Each meta-learning episode samples:

- a class set $C_e = \{E_1, \ldots, E_N\}$
- a support set S_e containing K spans per class,
- a query set Q_e drawn from disjoint samples.

Span embeddings are classified using a softmax classifier:

$$P(y_i = k) = softmax(Wv_i + \mathrm{b})_{k^*}$$

The meta-learning loss is defined as:

$$\mathcal{L}_{meta} = -\frac{1}{M}\sum_{i=1}^{M} \log P(y_i = y_i^{true})$$

Through repeated episodes, the model acquires generalized representations that support effective few-shot generalization across new Indonesian skill-entity contexts [51], [52].

D. CRF-Based Sequence Decoder

Conditional Random Fields (CRF) are integrated to enforce global consistency and BIO structural rules, a proven technique in modern NER architectures [49], [56]. Token-level classification approaches that rely solely on independent predictions, as discussed in a recent comprehensive NER review [53], often struggle to maintain consistency across an entire sequence, particularly when entities span multiple tokens. This limitation becomes evident in skill extraction tasks, where fragmented predictions can lead to incomplete or incorrect entity boundaries. To address this issue, a CRF layer widely adopted in modern NER pipelines [58] is employed as a sequence-level decoder that jointly optimizes label assignments across the sentence.

In this framework, the CRF layer incorporates structural constraints derived from the BIO tagging scheme to guide the decoding process. Certain label transitions are explicitly restricted to prevent invalid sequences. For example, transitions such as O → I-HSkill or B-Tech → I-SSkill are disallowed, while transitions like B-HSkill → I-HSkill or I-Tech → O are permitted. By enforcing these constraints during decoding, the model is encouraged to produce coherent and contiguous skill spans rather than isolated token predictions.

This structured decoding mechanism is particularly beneficial for handling multi-token skill expressions, which are prevalent in the NERSkill.id dataset. By considering both local token information and global sequence structure, the CRF layer helps reduce boundary fragmentation and improves the reliability of span-level predictions, especially for longer and semantically dense skill entities.

1) Emission computation

For each subword embedding h_t:

$$e_{t,k} = \left(W_{tag} h_t + b_{tag}\right)_k$$

2) Transition constraints

BIO constraints are encoded in the transition matrix:

$$A_{i,j} = -10^4, \qquad if\ transition\ i \rightarrow j\ violates\ BIO\ rules$$

Illegal sequences, such as $O \rightarrow I - HSkill$ or $B - Tech \rightarrow I - SSkill$, are strictly discouraged.

3) CRF Loss

The CRF objective for an entire sequence is:

$$\mathcal{L}_{CRF} = -\log P(y \mid e, A)$$

The CRF layer ensures global sequence validity and reduces fragmented entity predictions.

E. Joint Optimization Objective

The overall objective integrates CRF decoding, episodic meta-learning, and semantic alignment:

$$\mathcal{L}_{total} = \mathcal{L}_{CRF} + \alpha\mathcal{L}_{meta} + \gamma\mathcal{L}_{sem}$$

with weighting coefficients:

$$\alpha = 0.8, \qquad \gamma = 0.05$$

Optimization follows *AdamW* with differential learning rates for encoder and classification layers, consistent with best practices in transformer-based NER and low-resource adaptation [49], [56].

F. Evaluation Protocol

Evaluation follows an episodic meta-testing setup, consistent with few-shot NER practices [49], [56]. Each evaluation episode includes:

- a support set for rapid adaptation,
- a query set for prediction,
- CRF decoding to generate sequence-consistent BIO labels.

Predicted subword tags are mapped back to word-level BIO labels via:

$$\hat{t}_i = first\ subworld\ prediction\ mapped\ to\ w_1$$

Performance is measured using precision, recall, and F1-score at the entity level, following widely adopted NER evaluation methodologies [14], [48].

IV. Results

A. Dataset

The NERSkill.id [1] dataset provides a comprehensive benchmark for Few-Shot NER, capturing the diversity of skill expressions needed in fast-evolving digital labour environments, as summarized in Table I. The training set contains 3.440 sentences annotated with 22.798 Hard Skill, 11.852 Soft Skill, and 17.515 Technology entities, offering a rich and varied foundation for models that must generalize from limited examples. Complementing this, the testing set comprises 861 sentences with 5.719 Hard Skill, 2.897 Soft Skill, and 4.199 Technology entities, creating a challenging evaluation setting in which the model must distinguish between abstract soft-skill expressions and more structured technical terms under sparse supervision. This configuration closely reflects real-world innovation dynamics, where emerging competencies appear in fragmented forms yet demand accurate interpretation to support data-driven decision-making and guide strategic responses to technological change.

TABLE I. DISTRIBUTION OF SENTENCES AND ENTITIES IN THE NERSKILL.ID FEW-SHOT NER BENCHMARK

Data	Sentences	Entity		
		HSkill	SSkill	Tech
Train	3440	22798	11852	17515
Test	861	5719	2897	4199

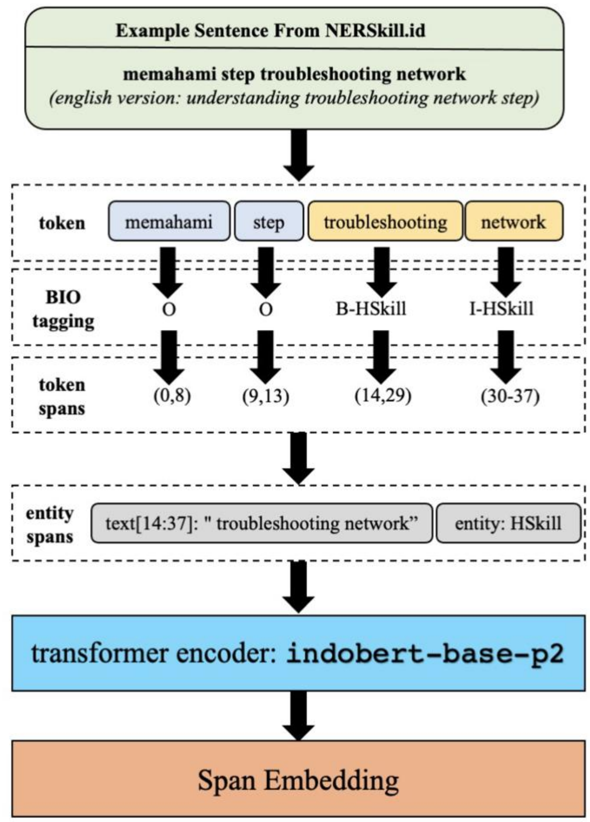

Fig. 2. Illustration of tokenization, BIO tagging, and span construction from the NERSkill.id dataset prior to transformer-based encoding.

The example in Fig. 2 illustrates how entities from the NERSkill.id dataset is processed within the Few-Shot NER pipeline, showing the step-by-step transformation from raw text into span-level embeddings. In the sentence "memahami step troubleshooting network", the tokens "troubleshooting" and "network" are correctly recognized as a continuous Hard Skill through the BIO tagging scheme, producing token spans that remain faithful to their semantic boundaries. These spans are then merged into a coherent entity representation "troubleshooting network" before being encoded by an Indonesian transformer backbone (indobert-base-p2). This encoding captures both contextual nuance and the internal structure of the skill expression, enabling the model to generalize effectively even when annotated examples are limited. The transition from token-level labels to enriched span embeddings demonstrates the dataset's ability to support robust

[1] https://data.mendeley.com/datasets/5s8r9ndfvc/2

Few-Shot NER modelling, particularly in identifying complex technical skills embedded in practical, real-world language.

TABLE II. HYPERPARAMETERS USED IN THE FEW-SHOT NER EXPERIMENT FOR SKILL EXTRACTION

Component	Hyperparameter	Value
Pretrained encoder	Model name	indobert-base-p2
Input	Max sequence length	96
Few-shot episode (train)	N-way	3
	K-shot (support per class)	15
	Q (query sentences)	30
Few-shot episode (test)	N-way	3
	K-shot	10
	Q	20
Training schedule	Training episodes	3000
	Evaluation episodes	50
Optimization	Optimizer	AdamW
	Encoder learning rate	1×10^{-5}
	Head learning rate	3×10^{-4}
	Weight decay	0.01
	Warmup ratio	0.1
Architecture and regularization	Dropout	0.1
	Frozen encoder layers	6
Prototypical learning	Prototype dimension	Encoder hidden size
	Temperature τ	0.1
	Prototype loss weight α	0.8
Semantic alignment	JS alignment weight γ	0.05
	Alignment temperature τ_sem	0.07
Fusion mechanism	Semantic fusion	Enabled
Decoding	CRF layer	Enabled
Class balancing	Class weights	$\sqrt{\max\_count/count}$
Rare entity boost	Rare entity boost	All 1.0

B. Experiment Setup

The experimental setup (as shown in Table II) was designed to examine how a semantic meta-learning framework with CRF can extract skill information from limited examples within a fast-changing and innovation-driven labor ecosystem. We implemented an episodic Few-Shot NER architecture based on a semantic meta-learning variant of the Episodic Span Metric NER model, using `indobert-base-p2` as the underlying encoder. Each episode incorporated all three skill categories in the NERSkill.id dataset (HSkill, SSkill, Tech) in a 3-way configuration, with up to 15 support sentences and 30 query sentences, and a maximum sequence length of 96 subword tokens. BIO-labelled tokens were transformed into span-level representations to construct data-driven class prototypes, which were then adaptively combined with text-based label prototypes through a semantic fusion mechanism. A JS-divergence alignment term was added to strengthen semantic coherence between the two sources of prototypes. To ensure training stability in a low-resource setting typical of emerging skills and novel competency expressions the lower six transformer layers were frozen, while task-specific layers were optimized with differential learning rates using AdamW. A CRF decoder with strict BIO transition constraints was used to produce structured predictions. During evaluation, the same episodic protocol was applied to the test set using 3-way episodes, 10-shot support, and 20 query sentences across 50 episodes, enabling a realistic assessment of how effectively the semantic meta-learning model generalizes across diverse skill types encountered in complex digital labour markets.

C. Main Results

The evaluation outcomes, summarized in Table III, indicate that the proposed approach demonstrates strong generalization under the 3-way, 10-shot, 20-query episodic configuration evaluated across 50 episodes. Overall, the model achieves a micro-F1 of 73.84% and a macro-F1 of 74.33%**,** reflecting balanced performance across entity types. Among the three categories, the model performs best on Tech entities (F1 = 84.10%**,** Precision = 84.23%, Recall = 83.96%), suggesting that technology-related terms are typically more concrete and consistently benefit from the span-based prototypical representation and the constrained CRF decoding mechanism. Soft Skill (SSkill) extraction also yields strong results (F1 = 76.79%**),** which underscores the effectiveness of the semantic fusion component in capturing more abstract and context-dependent skill expressions. In contrast, Hard Skill (HSkill) shows lower performance (F1 = 62.12%**),** with recall slightly lagging behind precision, indicating that the model is more conservative and tends to miss instances of hard skills that often exhibit greater lexical variability. The close alignment between micro, macro, and weighted averages (around 74% F1**)** suggests that performance remains relatively stable across classes without overfitting to any particular entity type. From an open innovation perspective, these findings highlight the potential of semantic meta-learning as a scalable approach for mapping emerging competencies in digital labour ecosystems, while also revealing the need for further refinement, particularly in handling more heterogeneous hard-skill expressions to improve the robustness of AI-driven skill intelligence systems.

TABLE III. MODEL PERFORMANCE ACROSS ENTITIES UNDER THE SEMANTIC META-LEARNING FRAMEWORK WITH CRF

Entity	Precision	Recall	F1
HSkill	63.16 %	61.12 %	62.12%
SSkil	78.22 %	75.41 %	76.79 %
Tech	84.23 %	83.96 %	84.10 %
micro avg	74.64 %	73.06 %	73.84 %
macro avg	75.20 %	73.49 %	74.33 %
weighted avg	74.53 %	73.06 %	73.78 %

The comparative results presented in Table IV, highlight the contrasting learning behaviours of traditional supervised architectures and episodic meta-learning approaches. The BiLSTM-CRF and IndoBERT fine-tuning baselines, each trained for 5 supervised epochs, deliver strong F1 scores of 70.55% and 72.80%**,** respectively. Their performance reflects the advantages of dense, token-level supervision, which enables stable optimization and effective pattern learning when sufficient labelled data is available. However, these models inherently depend on extensive annotation and remain less flexible when encountering emerging or low-frequency skill expression situations that commonly arise in dynamic, innovation-driven labour markets.

TABLE IV. PERFORMANCE COMPARISON BETWEEN BASELINE AND THE PROPOSED MODEL

Model	Precision	Recall	F1
BiLSTM-CRF	74.55 %	66.96 %	70.55 %
Transformer Fine-Tuning(IndoBERT)	70.36 %	75.41 %	72.80 %
Prototypical Networks	1.97 %	3.48 %	2.51 %
Frozen ProtoNet	3.40 %	5.02 %	4.05 %
FSM-CRF (Proposed model)	**74.64 %**	**73.06 %**	**73.84 %**

In comparison, few-shot baselines, such as Prototypical Networks (F1 = 2.51%) and Frozen ProtoNet (F1 = 4.05%), both trained for 3000 episodic iterations, perform poorly despite their substantially longer training schedule. This gap demonstrates that simply increasing episodic cycles does not guarantee effective learning in domains characterized by heterogeneous multi-token skill spans and nuanced semantic structures. Without mechanisms to incorporate span-level context or semantic grounding, these models struggle to form reliable prototypes, resulting in extremely low predictive accuracy.

TABLE V. ABLATION RESULTS OF THE FSM-CRF MODEL

Model Few Shot	k-shot = 10		
	Precision	Recall	F1
FSM-CRF (proposed method)	**74.64 %**	**73.06 %**	**73.84 %**
Meta-Learning + CRF	40.41 %	37.34 %	38.81 %
Semantic + CRF	71.51 %	68.58 %	70.02 %
Semantic + Meta-Learning	43.48 %	57.54 %	49.53 %

The proposed Few-Shot Semantic Meta-Learning model with CRF, also trained for 3000 episodic iterations, achieves a significantly higher F1 score of 73.84%, approaching the performance of fully supervised fine-tuning despite operating under few-shot constraints. This improvement stems from its ability to integrate span-based representations with semantic fusion and BIO-constrained decoding, enabling the model to generalize effectively from limited examples. From an open innovation perspective, these findings suggest that semantic meta-learning offers a more adaptive and scalable solution for identifying emerging competencies in complex digital labour ecosystems. While traditional models excel when labelled data is abundant, the proposed approach demonstrates superior resilience and adaptability qualities increasingly essential for monitoring skill evolution in technology-mediated, rapidly transforming markets.

The ablation results presented in Table V clarify the individual and combined contributions of the semantic module, meta-learning strategy, and CRF-based structured decoding within the proposed framework. The FSM-CRF full model (Semantic + Meta-Learning + CRF) delivers the highest performance 74.64% precision, 73.06% recall, and 73.84% F1 demonstrating that these three components operate synergistically rather than serving as direct substitutes. When the semantic module is removed (Meta-Learning + CRF), performance falls sharply to an F1 of 38.81%, indicating that meta-learning alone is insufficient for constructing stable span prototypes without explicit semantic grounding. In contrast, retaining the semantic module, but disabling meta-learning (Semantic + CRF) yields a considerably stronger F1 of 70.02%, showing that semantic fusion together with BIO-constrained CRF decoding provides a solid foundation for low-resource skill extraction. The Semantic + Meta-Learning configuration (without CRF) produces an F1 of 49.53%, with notably higher recall (57.54%) than precision (43.48%), suggesting a tendency toward over-prediction when sequence-level constraints are absent. These results collectively highlight that semantic representation and structured decoding are central to robust Few-Shot NER on complex skill data, while meta-learning contributes critical adaptability when supported by strong semantic and structural priors.

TABLE VI. NEMENYI TEST *P*-VALUES FOR FSM-CRF ABLATION MODELS

	FSM-CRF*	Meta-Learning + CRF	Semantic + CRF	Semantic + Meta-Learning
FSM-CRF*	1	0.354	0.947	0.692
Meta-Learning + CRF	0.354	1	0.692	0.947
Semantic + CRF	0.947	0.692	1	0.947
Semantic + Meta-Learning	0.692	0.947	0.947	1

The statistical analysis applied to the four few-shot configurations FSM-CRF full model (Semantic + Meta-Learning + CRF), Meta-Learning + CRF, Semantic + CRF, and Semantic + Meta-Learning was conducted using a non-parametric Friedman test followed by a Nemenyi post-hoc procedure. The Friedman test produced a chi-square value of $\chi^2(3) = 3.00$ with $\rho = 0.392$, indicating no statistically detectable difference among the models under the current number of evaluation runs. The post-hoc Nemenyi comparison, summarised in Table VI, reports ρ-values ranging from 0.35 to 0.95, confirming the absence of significant pairwise differences. Although the descriptive performance metrics clearly position the full Semantic + Meta-Learning + CRF model as the strongest configuration, the statistical results suggest that the observed differences represent consistent performance trends rather than formally significant improvements. Such outcomes are common in few-shot learning experiments, where limited observations and episodic variability often reduce the statistical power of non-parametric tests. From an open innovation perspective, this interplay is essential: only when all three components are integrated does the model achieve the resilience and flexibility required to identify emerging skills and competencies in rapidly evolving, technology-intensive labour markets.

D. Discussion

The experimental results highlight that combining semantic span representations, meta-learning, and CRF-based structured decoding yields an effective framework for Few-Shot Named Entity Recognition in skill-focused Indonesian text. The full model integrating semantic fusion and episodic meta-learning with BIO-constrained CRF achieves the highest overall performance (F1=73.84%), demonstrating stronger generalization than all ablated variants despite learning from limited examples per entity type. This behaviour is consistent with recent evidence that hybrid meta-learning architectures and semantic-aware prototype representations can substantially improve performance in low-resource NER, especially when dealing with domain-specific entities that exhibit multi-word compositions and rich semantic dependencies [18], [21], [22], [29]. More broadly, the observed gains also align with survey findings that emphasize the advantage of combining deep contextual encoders with structured decoding and metric- or prototype-based components for complex NER scenarios [14], [53].

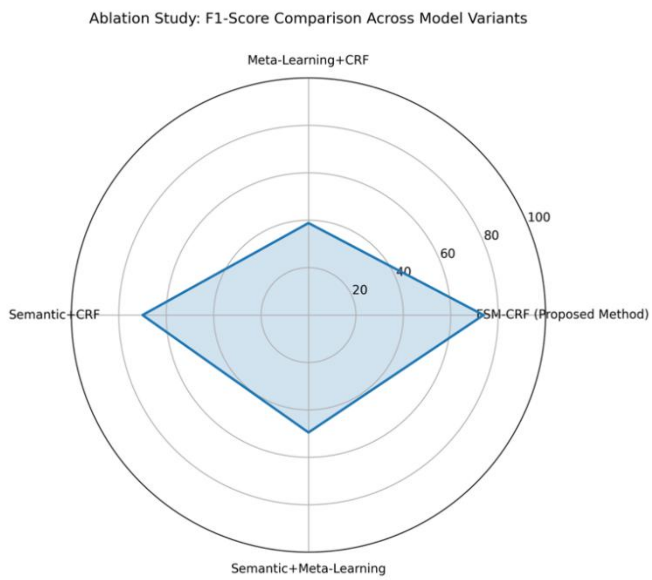

Fig. 3. F1-score comparison of ablated model configurations for FSM-CRF (k-shot = 10).

Ablation analysis further clarifies the role of each component in the proposed framework (refer to Fig. 3) and aligns with recent findings in learning-based information extraction, particularly within Indonesian legal NER contexts, where structured prediction and contextualized representations are shown to substantially enhance extraction accuracy [59]. Retaining semantic span representations together with CRF decoding ("Semantic + CRF") yields strong performance (F1 = 70.02%), indicating that contextualized semantic embeddings are crucial for modelling skill-related entities, which typically encode functional, behavioural, and domain-specific meaning consistent with recent work showing that span-level and semantically enriched representations substantially improve robustness over purely token-level models [60], [61], [62], [63]. In contrast, removing semantic grounding while preserving meta-learning ("Meta-Learning + CRF") produces a marked drop in performance (F1 = 38.81%), suggesting that prototype formation becomes unstable without semantic anchors, in line with critiques that standard prototypical networks struggle with multi-token entities and dispersed prototype distributions in few-shot NER [17], [64]. The "Semantic + Meta-Learning" configuration, which omits CRF, exhibits inflated recall but reduced precision, highlighting a tendency to over-predict entities when sequence-level constraints are absent and reinforcing evidence that CRF-based or constrained decoding remains essential for controlling label noise and enforcing valid BIO transitions in practical NER systems [65], [66], [67], [68].

The statistical evaluation complements these descriptive findings. Although the full model consistently outperforms the other configurations, the Friedman test yields no statistically significant differences (p = 0.392), and none of the Nemenyi post-hoc comparisons indicate pairwise significance. This pattern is consistent with reports that the Friedman test has limited statistical power when only a small number of datasets or experimental conditions are available, which makes it difficult to detect differences even when effect sizes are non-trivial [69]. Similar challenges are highlighted in recent surveys on learning with limited labelled data and few-shot learning, which show that episodic evaluation and small numbers of runs often lead to high variance and unstable significance outcomes, especially in meta-learning and low-resource settings [70], [71]. Even so, the consistent directional gaps between semantic-enabled and semantic-disabled variants in our experiments suggest meaningful functional differences that may become statistically robust when evaluated with more episodes, additional datasets, or cross-domain repetitions, in line with recommendations to increase repetitions and task diversity when benchmarking few-shot and meta-learning algorithms [72], [73].

From an open innovation perspective, these findings carry important implications for how skills are monitored and governed in digitally mediated labour markets. Recent studies show that organizations increasingly rely on rich skill taxonomies, digital platforms, and AI-based analytics to track emerging competencies and support innovation strategies, rather than focusing solely on formal qualifications [74], [75], [76]. In this context, the proposed semantic meta-learning framework with CRF offers a data-efficient mechanism for automatically identifying new or evolving skills from unstructured text, complementing ongoing efforts to map digital competencies and open innovation capabilities in organizations [77], [78]. The ability to extract fine-grained, semantics-aware skill entities directly supports emerging work on formalized skill taxonomies and soft-skill labelling, such as ESCO-oriented approaches and big-data-based workforce analytics and reduces dependence on costly manual annotation pipelines [75], [79]. Moreover, Few-Shot NER for skill intelligence aligns with AI-enabled competency mapping and assessment tools that aim to provide continuous, adaptive views of workforce capabilities for education providers, firms, and policy actors [3], [80]. Collectively, these connections position the proposed model as a technical enabler for open, innovation-driven skill ecosystems in which capabilities can be updated dynamically in response to technological change.

V. Conclusion

This study demonstrates that combining semantic span representations, meta-learning, and CRF-based structured decoding offers an effective framework for Few-Shot Named Entity Recognition in skill-extraction tasks. The proposed semantic meta-learning model consistently outperforms all ablated variants and achieves performance comparable to fully supervised baselines, even with limited annotated examples. These results underscore the importance of semantic grounding and structured inference in stabilizing prototype formation and improving entity boundary detection in low-resource conditions. They also show that meta-learning delivers its strongest benefits when supported by robust semantic and structural priors.

From an open innovation perspective, the model's ability to generalize from minimal supervision provides meaningful value for adaptive skill-intelligence systems. As industries evolve and new competencies emerge, Few-Shot NER enables organizations to update skill taxonomies efficiently without extensive annotation pipelines. This aligns with the principles of open innovation, where rapid knowledge flows, flexible learning mechanisms, and resource-efficient processes are essential for maintaining competitiveness. The findings therefore support emerging applications in workforce analytics, domain-adaptive NER, and AI-driven competency mapping across digital innovation ecosystems.

While limitations include the modest number of episodic evaluations and the single-domain focus, addressing these areas in future work will enhance the robustness and broader impact of Few-Shot NER within innovation-driven ecosystems. Future research should broaden evaluation across diverse domains, increase episodic trial counts to strengthen statistical validity, and explore multilingual or cross-domain transfer settings where skill definitions may vary. Integrating large language models or hybrid symbolic–neural approaches could further enhance adaptability and interpretability, especially in environments where competency requirements change rapidly. Such developments would extend the model's relevance for educational systems, employment services, and digital governance.

Beyond its empirical contributions, this study offers important implications for theory, practice, and policy. Theoretically, it reinforces the value of hybrid semantic–meta-learning architectures for modelling complex, domain-specific entities. Practically, the model provides an efficient solution for refining skill taxonomies and supporting adaptive learning and talent management. For policymakers, the approach enables evidence-informed monitoring of emerging competencies, improving alignment between innovation strategies and labor-market needs.

ACKNOWLEDGMENT

The authors would like to thank the help provided by Universitas Duta Bangsa Surakarta Indonesia, Universitas Dian Nuswantoro Semarang Indonesia, and INTI International University Malaysia.

REFERENCES

[1] M. Bouwmans, X. Lub, M. Orlowski, and T.-V. V. Nguyen, "Developing the digital transformation skills framework: A systematic literature review approach," PLoS One, vol. 19, no. 7, p. e0304127, Jul. 2024, doi: 10.1371/journal.pone.0304127.

[2] E. van Laar, A. J. A. M. van Deursen, J. A. G. M. van Dijk, and J. de Haan, "Determinants of 21st-Century Skills and 21st-Century Digital Skills for Workers: A Systematic Literature Review," SAGE Open, vol. 10, no. 1, 2020, doi: 10.1177/2158244019900176.

[3] J. Hochstetter-Diez, M. Negrier-Seguel, M. Diéguez-Rebolledo, E. Candia-Garrido, and E. Vidal, "From Mapping to Action: SmartRubrics, an AI Tool for Competency-Based Assessment in Engineering Education," Sustain., vol. 17, no. 13, pp. 1–27, 2025, doi: 10.3390/su17136098.

[4] B. Audrin, C. Audrin, and X. Salamin, "Digital skills at work – Conceptual development and empirical validation of a measurement scale," Technol. Forecast. Soc. Change, vol. 202, no. February, p. 123279, 2024, doi: https://doi.org/10.1016/j.techfore.2024.123279.

[5] K. Maldonado-Mariscal, M. Cuypers, A. Götting, and M. Kohlgrüber, "Skills Intelligence in the Steel Sector," Machines, vol. 11, no. 3, p. 335, 2023, doi: 10.3390/machines11030335.

[6] M. Bone, E. González Ehlinger, and F. Stephany, "Skills or degree? The rise of skill-based hiring for AI and green jobs," Technol. Forecast. Soc. Change, vol. 214, no. February, p. 124042, 2025, doi: https://doi.org/10.1016/j.techfore.2025.124042.

[7] P. Mhaske, B. Bhattacharjee, N. Haldar, P. Upadhyay, and A. Mandal, "Bridging digital skill gaps in the global workforce: A synthesis and conceptual framework building," Res. Glob., vol. 11, no. June, p. 100311, 2025, doi: 10.1016/j.resglo.2025.100311.

[8] G. Gayatri, I. G. N. M. Jaya, and V. M. Rumata, "The Indonesian Digital Workforce Gaps in 2021–2025," Sustain., vol. 15, no. 1, 2023, doi: 10.3390/su15010754.

[9] E. H. Prasetyo, "Digital platforms' strategies in Indonesia: Navigating between technology and informal economy," Technol. Soc., vol. 76, no. November 2023, p. 102414, 2024, doi: 10.1016/j.techsoc.2023.102414.

[10] J. Santoso, E. I. Setiawan, C. N. Purwanto, E. M. Yuniarno, M. Hariadi, and M. H. Purnomo, "Named entity recognition for extracting concept in ontology building on Indonesian language using end-to-end bidirectional long short term memory," Expert Syst. Appl., vol. 176, no. August 2020, p. 114856, 2021, doi: 10.1016/j.eswa.2021.114856.

[11] D. A. Sulistyo, A. P. Wibawa, D. D. Prasetya, and F. A. iini Ahda, "Indonesian cross-linguistic named entity recognition," Res. Methods Appl. Linguist., vol. 4, no. 3, p. 100236, 2025, doi: 10.1016/j.rmal.2025.100236.

[12] A. M. Avram, V. B. Mititelu, V. Păiș, D. C. Cercel, and Ștefan Trăușan-Matu, "Multilingual Multiword Expression Identification Using Lateral Inhibition and Domain Adaptation," Mathematics, vol. 11, no. 11, pp. 1–18, 2023, doi: 10.3390/math11112548.

[13] S. Chen, Y. Pei, Z. Ke, and W. Silamu, "Low-resource named entity recognition via the pre-training model," Symmetry (Basel)., vol. 13, no. 5, pp. 1–15, 2021, doi: 10.3390/sym13050786.

[14] J. Li, A. Sun, J. Han, and C. Li, "A Survey on Deep Learning for Named Entity Recognition," IEEE Trans. Knowl. Data Eng., vol. 34, no. 1, pp. 50–70, 2022, doi: 10.1109/TKDE.2020.2981314.

[15] O. Ozcelik and C. Toraman, "Named entity recognition in Turkish: A comparative study with detailed error analysis," Inf. Process. Manag., vol. 59, no. 6, p. 103065, 2022, doi: 10.1016/j.ipm.2022.103065.

[16] M. Gavrilescu, F. Leon, and A. A. Minea, "Techniques for Transversal Skill Classification and Relevant Keyword Extraction from Job Advertisements," Inf., vol. 16, no. 3, pp. 1–24, 2025, doi: 10.3390/info16030167.

[17] B. Ji et al., "Few-shot Named Entity Recognition with Entity-level Prototypical Network Enhanced by Dispersedly Distributed Prototypes," in Proceedings of the 29th International Conference on Computational Linguistics, Oct. 2022, vol. 29, no. 1, pp. 1842–1854, [Online]. Available: https://aclanthology.org/2022.coling-1.159/.

[18] V. Moscato, M. Postiglione, and G. Sperlí, "Few-shot Named Entity Recognition : Definition , Taxonomy and Research Directions," ACM Trans. Intell. Syst. Technol., vol. 14, no. 5, 2023, doi: 10.1145/3609483.

[19] Z. Yang, Y. Liu, C. Ouyang, S. Zhao, and C. Zhu, "Improving Few-Shot Named Entity Recognition with Causal Interventions," Big Data Min. Anal., vol. 7, no. 4, pp. 1375–1395, 2024, doi: 10.26599/BDMA.2024.9020052.

[20] K. He, R. Mao, Y. Huang, T. Gong, C. Li, and E. Cambria, “Template-Free Prompting for Few-Shot Named Entity Recognition via Semantic-Enhanced Contrastive Learning,” IEEE Trans. Neural Networks Learn. Syst., vol. 35, no. 12, pp. 18357–18369, Dec. 2024, doi: 10.1109/TNNLS.2023.3314807.

[21] J. Feng, G. Xu, Q. Wang, Y. Yang, and L. Huang, “Note the hierarchy: Taxonomy-guided prototype for few-shot named entity recognition,” Inf. Process. Manag., vol. 61, no. 1, p. 103557, 2024, doi: 10.1016/j.ipm.2023.103557.

[22] P. Chen, J. Wang, H. Lin, D. Zhao, and Z. Yang, “Few-shot biomedical named entity recognition via knowledge-guided instance generation and prompt contrastive learning,” Bioinformatics, vol. 39, no. 8, p. btad496, Aug. 2023, doi: 10.1093/bioinformatics/btad496.

[23] W. Fang, Y. Liu, C. Ouyang, L. Ren, J. Li, and Y. Wan, “Joint span and token framework for few-shot named entity recognition,” AI Open, vol. 4, no. July, pp. 111–119, 2023, doi: 10.1016/j.aiopen.2023.08.009.

[24] F. Chiarello, G. Fantoni, T. Hogarth, V. Giordano, L. Baltina, and I. Spada, “Towards ESCO 4.0 – Is the European classification of skills in line with Industry 4.0? A text mining approach,” Technol. Forecast. Soc. Change, vol. 173, 2021, doi: 10.1016/j.techfore.2021.121177.

[25] S. Fareri, N. Melluso, F. Chiarello, and G. Fantoni, “SkillNER: Mining and mapping soft skills from any text,” Expert Syst. Appl., vol. 184, no. June, p. 115544, 2021, doi: https://doi.org/10.1016/j.eswa.2021.115544.

[26] M. Košprdić, N. Prodanović, A. Ljajić, B. Bašaragin, and N. Milošević, “From zero to hero: Harnessing transformers for biomedical named entity recognition in zero- and few-shot contexts,” Artif. Intell. Med., vol. 156, no. August, p. 102970, 2024, doi: 10.1016/j.artmed.2024.102970.

[27] J. Li, B. Chiu, S. Feng, and H. Wang, “Few-Shot Named Entity Recognition via Meta-Learning,” IEEE Trans. Knowl. DATA Eng., pp. 1–12, 2022.

[28] Z. Ren, X. Qin, and W. Ran, “SLNER: Chinese Few-Shot Named Entity Recognition with Enhanced Span and Label Semantics,” Appl. Sci., vol. 13, no. 15, pp. 1–19, 2023, doi: 10.3390/app13158609.

[29] D. Zhou, S. Li, Q. Chen, and H. Yao, “Improving few-shot named entity recognition via Semantics induced Optimal Transport,” Neurocomputing, vol. 597, no. May, p. 127938, 2024, doi: 10.1016/j.neucom.2024.127938.

[30] E. Zha, D. Zeng, M. Lin, and Y. Shen, “CEPTNER : Contrastive learning Enhanced Prototypical network for Two-stage few-shot Named Entity Recognition,” Knowledge-Based Syst., vol. 295, no. April, p. 111730, 2024, doi: 10.1016/j.knosys.2024.111730.

[31] T. Zhang, S. Zhang, X. Liu, B. Cao, and J. Fan, “Multimodal Data Fusion for Few-shot Named Entity Recognition Method,” Int. J. Softw. Informatics, vol. 14, no. 1, pp. 73–96, 2024, doi: 10.21655/ijsi.1673-7288.00318.

[32] E. Dave and A. Chowanda, “IPerFEX-2023: Indonesian personal financial entity extraction using indoBERT-BiGRU-CRF model,” J. Big Data, vol. 11, no. 1, 2024, doi: 10.1186/s40537-024-00987-6.

[33] M. N. Tentua, Suprapto, and Afiahayati, “NERSkill.Id: Annotated dataset of Indonesian’s skill entity recognition,” Data Br., vol. 53, p. 110192, 2024, doi: 10.1016/j.dib.2024.110192.

[34] G. F. Shidik et al., “Indonesian disaster named entity recognition from multi source information using bidirectional LSTM (BiLSTM),” J. Open Innov. Technol. Mark. Complex., vol. 10, no. 3, p. 100358, 2024, doi: 10.1016/j.joitmc.2024.100358.

[35] P. W. Cahyo, U. S. Aesyi, W. A. Setianto, and T. Sulaiman, “A Novel Named Entity Recognition approach of Indonesian fake news using part of speech and BERT model on presidential election,” Int. J. Inf. Manag. Data Insights, vol. 5, no. 2, p. 100354, 2025, doi: 10.1016/j.jjimei.2025.100354.

[36] Syukur, A. I. Faried, and L. N. Nasution, “Labor Gap Analysis in Supporting Digital Economic Growth in Indonesia,” in International Conference Epicentrum of Economic Global Framework, 2025, pp. 91–106.

[37] M. S. Silitonga, “The Public Sector’s Digital Skills Gap in Indonesia: The Challenges and Opportunities,” J. Good Gov., pp. 70–79, 2023, doi: 10.32834/gg.v19i1.585.

[38] A. Singh, A. Kanaujia, and V. K. Singh, “Data to Decisions: A Computational Framework to Identify Skill Requirements from Advertorial Data,” in Advanced Network Technologies and Intelligent Computing, 2025, pp. 435–458.

[39] Z. Zainuddin, Mudassir, and Z. Tahir, “Entity Extraction in Indonesian Online News Using Named Entity Recognition (NER) with Hybrid Method Transformer, Word2Vec, Attention and Bi-LSTM,” Int. J. Informatics Vis., vol. 9, no. 3, pp. 964–973, 2025, doi: 10.62527/joiv.9.3.2902.

[40] S. K. Alqaaidi, E. Bozorgi, A. Shams, and K. J. Kochut, “A Few-Shot Learning-Focused Survey on Recent Named Entity Recognition and Relation Classification Models,” in Proceedings ofthe 13th International Conference on Data Science, Technology and Applications (DATA 2024), 2024, pp. 450–459, doi: 10.5220/0012791600003756.

[41] J. Wang et al., “SpanProto: A Two-stage Span-based Prototypical Network for Few-shot Named Entity Recognition,” Proc. 2022 Conf. Empir. Methods Nat. Lang. Process. EMNLP 2022, pp. 3466–3476, 2022, doi: 10.18653/v1/2022.emnlp-main.227.

[42] Y. Xiao, Q. Yang, J. Zou, and S. Zhou, “LACNER: Enhancing Few-Shot Named Entity Recognition with Label Words and Contrastive Learning,” in 2024 International Joint Conference on Neural Networks (IJCNN), 2024, pp. 1–8, doi: 10.1109/IJCNN60899.2024.10650197.

[43] Y. Xu et al., “An Effective Span-based Multimodal Named Entity Recognition with Consistent Cross-Modal Alignment,” 2024 Jt. Int. Conf. Comput. Linguist. Lang. Resour. Eval. Lr. 2024 - Main Conf. Proc., pp. 1063–1072, 2024.

[44] H. Mao et al., “Span-based unified named entity recognition framework via contrastive learning,” 2024, doi: 10.24963/ijcai.2024/708.

[45] G. Dong et al., “A Multi-Task Semantic Decomposition Framework with Task-specific Pre-training for Few-Shot NER,” in Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, 2023, pp. 430–440, doi: 10.1145/3583780.3614766.

[46] J. Yu, B. Ji, S. Li, J. Ma, H. Liu, and H. Xu, “S-NER: A Concise and Efficient Span-Based Model for Named Entity Recognition,” Sensors, vol. 22, no. 8, 2022, doi: 10.3390/s22082852.

[47] W. J. Xu and J. Q. OuYang, “A Streamlined Span-based Factorization Method for Few Shot Named Entity Recognition,” 2024 Jt. Int. Conf. Comput. Linguist. Lang. Resour. Eval. Lr. 2024 - Main Conf. Proc., pp. 1673–1683, 2024.

[48] B. Jehangir, S. Radhakrishnan, and R. Agarwal, “A survey on Named Entity Recognition — datasets, tools, and methodologies,” Nat. Lang. Process. J., vol. 3, no. 10, p. 100017, 2023, doi: 10.1016/j.nlp.2023.100017.

[49] Z. Hu, W. Hou, and X. Liu, “Deep learning for named entity recognition: a survey,” Neural Comput. Appl., vol. 36, no. 16, pp. 8995–9022, 2024, doi: 10.1007/s00521-024-09646-6.

[50] N. Alshammari and S. Alanazi, “The impact of using different annotation schemes on named entity recognition,” Egypt. Informatics J., vol. 22, no. 3, pp. 295–302, 2021, doi: 10.1016/j.eij.2020.10.004.

[51] I. Budi and R. R. Suryono, “Application of named entity recognition method for Indonesian datasets: a review,” Bull. Electr. Eng. Informatics, vol. 12, no. 2, pp. 969–978, 2023, doi: 10.11591/eei.v12i2.4529.

[52] E. Yulianti, N. Bhary, J. Abdurrohman, F. W. Dwitilas, E. Q. Nuranti, and H. S. Husin, “Named entity recognition on Indonesian legal documents: a dataset and study using transformer-based models,” Int. J. Electr. Comput. Eng., vol. 14, no. 5, pp. 5489–5501, 2024, doi: 10.11591/ijece.v14i5.pp5489-5501.

[53] W. L. Seow, I. Chaturvedi, A. Hogarth, R. Mao, and E. Cambria, “A review of named entity recognition : from learning methods to modelling paradigms and tasks,” Artif. Intell. Rev., vol. 58:315, 2025, doi: https://doi.org/10.1007/s10462-025-11321-8 A.

[54] Y. Wang, H. Tong, Z. Zhu, and Y. Li, “Nested Named Entity Recognition: A Survey,” ACM Trans. Knowl. Discov. Data, vol. 16, no. 6, 2022, doi: 10.1145/3522593.

[55] Y. Park, G. Son, and M. Rho, “Biomedical Flat and Nested Named Entity Recognition: Methods, Challenges, and Advances,” Appl. Sci., vol. 14, no. 20, 2024, doi: 10.3390/app14209302.

[56] F. Ullah, A. Gelbukh, M. T. Zamir, E. M. F. Riverón, and G. Sidorov, "Enhancement of Named Entity Recognition in Low-Resource Languages with Data Augmentation and BERT Models: A Case Study on Urdu," Computers, vol. 13, no. 10, 2024, doi: 10.3390/computers13100258.

[57] D. Zhou, S. Li, Q. Chen, and H. Yao, "Improving few-shot named entity recognition via Semantics induced Optimal Transport," Neurocomputing, vol. 597, no. May, p. 127938, 2024, doi: https://doi.org/10.1016/j.neucom.2024.127938.

[58] B. Duc Tho, M. Nguyen, D. Tien Le, L. Ying, S. Inoue, and T.-T. Nguyen, "Improving biomedical Named Entity Recognition with additional external contexts," J. Biomed. Inform., vol. 156, no. May, p. 104674, 2024, doi: 10.1016/j.jbi.2024.104674.

[59] F. Solihin, I. Budi, E. M. S. Rochman, F. A. Mufarroha, A. A. Ramdlany, and D. A. Dewi, "Learning-Based Information Extraction to Obtain Prominent Named Entities in Indonesian Court Decision Documents," Math. Model. Eng. Probl., vol. 12, no. 5, pp. 1627–1633, 2025, doi: 10.18280/mmep.120517.

[60] Z. Li, S. Cao, M. Zhai, N. Ding, Z. Zhang, and B. Hu, "Multi-level semantic enhancement based on self-distillation BERT for Chinese named entity recognition," Neurocomputing, vol. 586, no. March, p. 127637, 2024, doi: 10.1016/j.neucom.2024.127637.

[61] Y. Sun, X. Wang, H. Wu, and M. Hu, "Global Span Semantic Dependency Awareness and Filtering Network for nested named entity recognition," Neurocomputing, vol. 617, no. May 2024, 2025, doi: 10.1016/j.neucom.2024.129035.

[62] E. Zhu, Y. Liu, and J. Li, "Deep Span Representations for Named Entity Recognition," Proc. Annu. Meet. Assoc. Comput. Linguist., no. 2, pp. 10565–10582, 2023, doi: 10.18653/v1/2023.findings-acl.672.

[63] Y. Xu, C. Mao, Z. Wang, G. Jin, liangji Zhong, and T. Qian, "Semantic-enhanced graph neural network for named entity recognition in ancient Chinese books," Sci. Rep., vol. 14, no. 1, pp. 1–12, 2024, doi: 10.1038/s41598-024-68561-x.

[64] Y. Xiao, J. Zou, and Q. Yang, "Advancing Few-Shot Named Entity Recognition with Large Language Model," Appl. Sci., vol. 15, no. 7, 2025, doi: 10.3390/app15073838.

[65] B. Lester et al., "Constrained Decoding for Computationally Efficient Named Entity Recognition Taggers," in Findings of the Association for Computational Linguistics: EMNLP 2020, Nov. 2020, pp. 1841–1848, doi: 10.18653/v1/2020.findings-emnlp.166.

[66] T. Wang, Y. Liu, C. Liang, B. Wang, and H. Liu, "XLNet-CRF: Efficient Named Entity Recognition for Cyber Threat Intelligence with Permutation Language Modeling," Electronics, vol. 14, no. 15, p. 3034, 2025, doi: 10.3390/electronics14153034.

[67] X. Xu, Z. Li, H. Zhang, and K. Ma, "Named entity recognition for Chinese electronic medical records by integrating knowledge graph and ClinicalBERT," Front. Artif. Intell., vol. 8, no. September, 2025, doi: 10.3389/frai.2025.1634774.

[68] R. A. A. Jonker, T. Almeida, R. Antunes, J. R. Almeida, and S. Matos, "Multi-head CRF classifier for biomedical multi-class named entity recognition on Spanish clinical notes," Database, vol. 2024, p. baae068, Feb. 2024, doi: 10.1093/database/baae068.

[69] L. F. Fernández-Salvador, B. Vilallonga Tejela, A. Almodóvar, J. Parras, and S. Zazo, "Attentive Neural Processes for Few-Shot Learning Anomaly-Based Vessel Localization Using Magnetic Sensor Data," J. Mar. Sci. Eng., vol. 13, no. 9, pp. 1–24, 2025, doi: 10.3390/jmse13091627.

[70] W. Zeng and Z.-Y. Y. Xiao, "Few-shot learning based on deep learning: A survey," Math. Biosci. Eng., vol. 21, no. 1, pp. 679–711, Jan. 2024, doi: 10.3934/mbe.2024029.

[71] B. Pecher, I. Srba, and M. Bielikova, "A Survey on Stability of Learning with Limited Labelled Data and its Sensitivity to the Effects of Randomness," ACM Comput. Surv., vol. 57, no. 1, Oct. 2024, doi: 10.1145/3691339.

[72] H. Gharoun, F. Momenifar, F. Chen, and A. H. Gandomi, "Meta-learning Approaches for Few-Shot Learning: A Survey of Recent Advances," ACM Comput. Surv., vol. 56, no. 12, Jul. 2024, doi: 10.1145/3659943.

[73] J. Y. Lim, K. M. Lim, C. P. Lee, and Y. X. Tan, "A review of few-shot image classification: Approaches, datasets and research trends," Neurocomputing, vol. 649, no. May, 2025, doi: 10.1016/j.neucom.2025.130774.

[74] Z. Musinszki, E. Horváthné Csolák, and K. Lipták, "Analysis of Labour Market Expectations in the Digital World Based on Job Advertisements," Adm. Sci., vol. 15, no. 7, 2025, doi: 10.3390/admsci15070282.

[75] F. Gurcan, A. Soylu, and A. Q. Khan, "Towards a Sustainable Workforce in Big Data Analytics: Skill Requirements Analysis from Online Job Postings Using Neural Topic Modeling," Sustain., vol. 17, no. 20, pp. 1–24, 2025, doi: 10.3390/su17209293.

[76] A. Dewalska-opitek, J. S.- Łucka, A. Szejniuk, and D. Kalita, "Competencies of the Future : How Artificial Intelligence Has Been Shaping Skills and Labour Market Transformation," vol. XXVIII, no. 4, pp. 151–170, 2025.

[77] M. L. Prasetyo, R. A. Peranginangin, N. Martinovic, M. Ichsan, and H. Wicaksono, "Artificial intelligence in open innovation project management: A systematic literature review on technologies, applications, and integration requirements," J. Open Innov. Technol. Mark. Complex., vol. 11, no. 1, p. 100445, 2025, doi: 10.1016/j.joitmc.2024.100445.

[78] L. Espina-Romero, D. Ríos Parra, J. G. Noroño-Sánchez, G. Rojas-Cangahuala, L. E. Cervera Cajo, and P. A. Velásquez-Tapullima, "Navigating Digital Transformation: Current Trends in Digital Competencies for Open Innovation in Organizations," Sustain. , vol. 16, no. 5, 2024, doi: 10.3390/su16052119.

[79] C. Panzaru and A. Grama, "Towards explicit soft skills labelling in ESCO through semantic NLP analysis," J. Labour Mark. Res., vol. 59, no. 1, 2025, doi: 10.1186/s12651-025-00409-x.

[80] S. Talodhikar and S. Farooqui, "Utilizing Artificial Intelligence for Competency Mapping and Personalised Skill Development in IT Organizations," J. Inf. Syst. Eng. Manag., vol. 10, no. 15s, pp. 61–76, 2025, doi: 10.52783/jisem.v10i15s.2430.