Graph Fusion Across Languages using Large Language Models

Kaung Myat Kyaw*, **Khush Agarwal***, **Jonathan Chan**
Innovative Cognitive Computing Research Center (IC2)
School of Information Technology, KMUTT
{kaungmyat.kyaw, khush.agar}@kmutt.ac.th, jonathan@sit.kmutt.ac.th

Abstract

Combining multiple knowledge graphs (KGs) across linguistic boundaries is a persistent challenge due to semantic heterogeneity and the complexity of graph environments. We propose a framework for cross-lingual graph fusion, leveraging the in-context reasoning and multilingual semantic priors of Large Language Models (LLMs). The framework implements structural linearization by mapping triplets directly into natural language sequences (e.g., [head] [relation] [tail]), enabling the LLM to map relations and reconcile entities between an evolving fused graph and a new candidate graph. Evaluated on the DBP15K dataset, this exploratory study demonstrates that LLMs can serve as a universal semantic bridge to resolve cross-lingual discrepancies. Results show the successful sequential agglomeration of multiple heterogeneous graphs, offering a scalable, modular solution for continuous knowledge synthesis in multi-source, multilingual environments. Our implementation and experimental framework are publicly available in our repository: `https://github.com/IC2-Lab-KMUTT/Multilingual-Graph-Fusion`

Keywords: Cross-Lingual Graph Fusion, Large Language Models, Entity Alignment

1. Introduction

The landscape of Artificial Intelligence has been fundamentally reshaped by the emergence of Large Language Models (LLMs) (Grattafiori et al., 2024; Yang et al., 2025; Team et al., 2023). These models, pre-trained on expansive, multilingual corpora, have surpassed their initial role as statistical text predictors to become sophisticated reasoning engines. By internalizing the semantic relationships across hundreds of languages, LLMs have demonstrated an unprecedented capacity for cross-lingual zero-shot transfer (Chirkova and Nikoulina, 2024). In parallel, Knowledge Graphs (KGs) have remained the foundational bedrock of structured, symbolic AI (Hogan et al., 2021). Unlike the probabilistic nature of neural models, KGs provide a deterministic and interpretable representation of real-world facts, serving as the ground truth for domain-specific applications such as biomedical research, legal compliance, and global financial analysis (Ji et al., 2021; Peng et al., 2023).

As the global digital ecosystem becomes increasingly interconnected, a critical challenge has emerged in the management of multilingual knowledge graphs (Perevalov et al., 2024). Knowledge is naturally fragmented across linguistic and cultural boundaries (Huang et al., 2022); for example, the English-centric part of DBpedia (Lehmann et al., 2015) may offer extensive coverage of Western historical events, while its Chinese or Japanese counterparts provide far greater granularity regarding Eastern heritage and regional entities. The task of cross-lingual entity alignment (EA) and graph fusion is therefore essential to bridge this information, facilitating the creation of a truly universal knowledge base. Historically, this task has been dominated by embedding-based strategies (Yang et al., 2019) or complex Graph Neural Networks (GNNs) (Zhang et al., 2023; Tong et al., 2022). However, these traditional methods are often constrained by their reliance on substantial seed data, which comprises manually pre-aligned entity pairs used to learn mapping functions between disparate vector spaces (Huo et al., 2024). Such requirements create a significant barrier to entry, particularly for low-resource languages or emerging domains where seed alignments are non-existent.

While Large Language Models (LLMs) have shown significant promise in graph reasoning (Zhu et al., 2024), their application to N-graph fusion remains underexplored. Unlike binary alignment, sequential N-graph fusion requires maintaining a consistent global state, where a single misalignment can propagate errors across the entire collection. Furthermore, a fundamental architectural tension exists: Knowledge Graphs are large-scale, non-linear structures, whereas LLMs are optimized to process information as linear sequences within finite context windows. This necessitates a framework that can effectively partition graph structures into LLM-readable contexts without losing the global structural integrity required for multi-graph synthesis.

The process of flattening a graph into text which is a necessity for LLM consumption risks the loss of critical "topological neighborhood" context. Without observing the surrounding nodes and edges, an LLM may struggle to disambiguate entities with identical names but different structural roles (Jin et al., 2024; Zhu et al., 2024; Pan et al., 2024). Most

* Equal contribution.

Proceedings of the Knowledge Graphs and Large Language Models Workshop (KG-LLM) @ LREC26, pages 102–109
May 16, 2026. ©ELRA Language Resources Association (ELRA), 2026

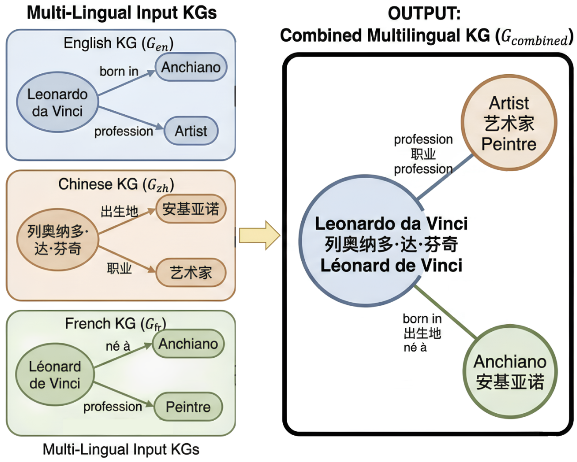

Figure 1: Conceptual framework for LLM-based cross-lingual knowledge graph fusion. Heterogeneous input graphs in English (G_{en}), Chinese (G_{zh}), and French (G_{fr}) are processed through a fusion framework that linearizes structural triplets into natural language and the LLM acts as a semantic bridge to reconcile entities.

current alignment systems (Zhao et al., 2020) are developed as specialized “black boxes” optimized for specific model architectures or fixed language pairs, lacking a modular infrastructure that allows researchers to swap LLM or test various partitioning strategies in a dynamic, plug-and-play fashion.

In this paper, we propose a novel, modular framework for LLM-based N-Graph Knowledge Graph Fusion. As illustrated in Figure 1, we move away from static mappings and position the LLM as a flexible semantic glue that discovers entity and relation alignments through sequential agglomeration and in-context reasoning. An overview of this modular architecture is illustrated in Figure 2. Our work makes several key contributions to the field of knowledge integration.

- We introduce a sequential agglomeration fusion pipeline based on a “rolling” strategy, where N graphs are sequentially integrated into an evolving global graph, $G_{combined}$.
- To address scalability, we implement a context-aware graph partitioner that subdivides large KGs into manageable, readable triplet batches, ensuring the LLM maintains a local structural view of the entities it is aligning.

We provide an evaluation on the multilingual DBP15K (Sun et al., 2017), covering Chinese-English, Japanese-English, and French-English pairs. Our results demonstrate that by optimizing prompt structures and linearization formats, our pipeline achieves superior precision in cross-lingual mapping, proving that LLMs can serve as the primary engine for large-scale, multilingual graph synthesis.

2. Preliminaries & Related Works

A KG is defined as a directed multigraph $G = (E, R, T)$, where E is the set of entities, R is the set of relation types, and $T \subseteq E \times R \times E$ is a set of triplets (h, r, t). In a Multilingual KG (M-KG) setting, we consider N disparate graphs $\{G_1, G_2, \ldots, G_n\}$ where each G_i represents knowledge in a specific language L_i. The objective of Cross-Lingual Entity Alignment (EA) is to find a set of alignment pairs $A_{i,j} = \{(e_i, e_j) \in E_i \times E_j \mid e_i \equiv e_j\}$, where $\equiv$ denotes semantic equivalence of real-world entities across linguistic barriers. Historically, this process relies on seed alignments (or "seeds") which is a small set of manually pre-aligned entity pairs that serve as the ground-truth supervision (Sun et al., 2017; Zhao et al., 2020). These seeds act as anchors to bridge disparate graphs; however, they are often expensive to curate, creating a bottleneck for low-resource languages where such gold standard links are non-existent.

This task was addressed through knowledge representation learning and translational models such as TransE (Bordes et al., 2013), which assumes that for any valid triplet, the embeddings $\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^d$ satisfy $\mathbf{h} + \mathbf{r} \approx \mathbf{t}$. To bridge linguistic gaps, models like MTransE (Chen et al., 2016) and JAPE (Sun et al., 2017) introduced transition matrices $\mathbf{M}_{ij}$ to map an entity $\mathbf{e}_i$ from G_i into the vector space of G_j, such that $\|\mathbf{M}_{ij}\mathbf{e}_i - \mathbf{e}_j\| \to 0$. However, these distance-based approaches are highly sensitive to the quality and volume of seed alignments, as pre-aligned entity pairs used as supervision are often scarce or non-existent in low-resource or emerging domains. While GNNs such as GCN-Align (Wang et al., 2018) and RDGCN (Wu et al., 2019) utilize neighborhood information to disambiguate entities, they are language-blind without high-quality initial word embeddings. Moreover, GNNs often operate as "black boxes," making it difficult to interpret why an alignment was made or to adapt to new languages without retraining.

With the rise of models like Llama 3 (Grattafiori et al., 2024) and Gemini (Team et al., 2023), the field has shifted from distance-based matching to in-context reasoning. Recent frameworks like ZeroEA (Huo et al., 2024) and Seg-Align (Yang et al., 2024a) demonstrate that LLMs can perform alignment with zero or minimal seed data by treating entities as text. Models like ProLEA (Munne et al., 2025) generate "contextual profiles" for entities, transforming structured triplets into natural language summaries. This "semantic layer" bridges the gap between different linguistic representations.

While binary alignment is well-studied, N-graph

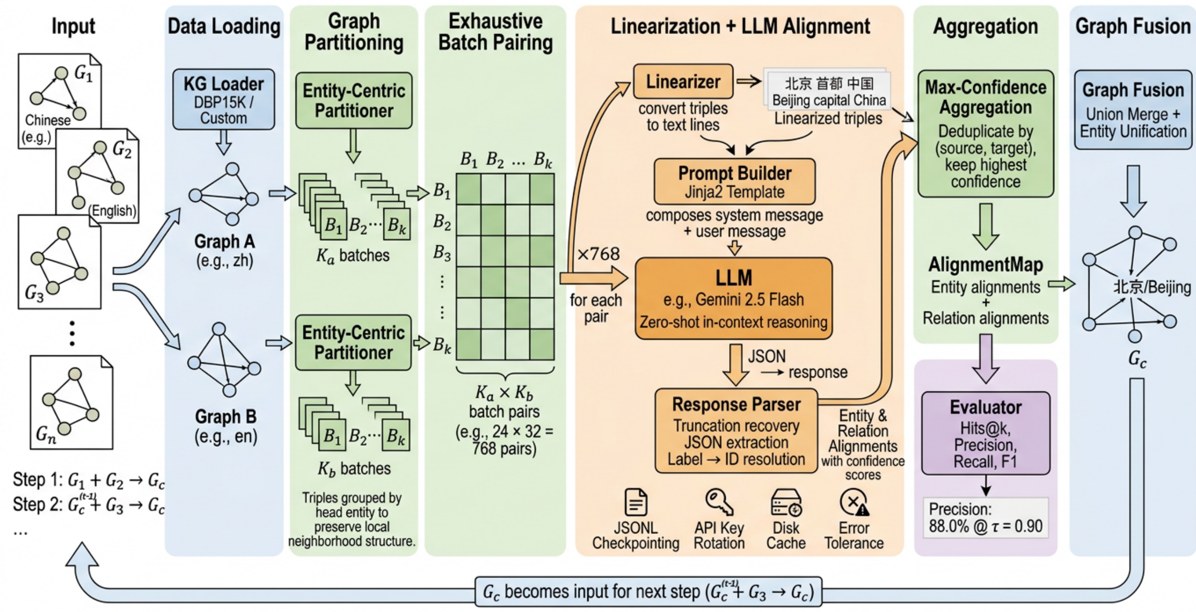

Figure 2: Overview of the proposed modular framework for N-Graph Knowledge Graph Fusion.

fusion remains a frontier. Recent joint approaches, such as MultiEA (Yang et al., 2024b), attempt to align multiple KGs in a single pass by clustering entities in a shared feature space to avoid the transitive error propagation inherent in sequential chains. However, such monolithic strategies are often difficult to scale to truly heterogeneous multilingual scenarios, as they require all N graphs to be present simultaneously during the optimization process. This lacks the modularity needed for incremental graph integration, where new languages or data sources may be introduced sequentially.

By combining the structured reliability of KGs with the flexible, multi-step reasoning abilities of large language models, this work outlines a practical path toward a scalable, multilingual knowledge base. Rather than treating knowledge as fixed binary mappings, our approach supports an evolving synthesis of information that can grow across languages and domains. In this way, the study aims to help move the field toward more adaptive and globally integrated knowledge systems.

3. Methodology

3.1. Problem Formulation

We define the Multilingual N-Graph Fusion task as the sequential agglomeration of a global, unified knowledge graph $G_c^{(N)}$ from a set of N heterogeneous graphs $\{G_1, G_2, \ldots, G_n\}$ expressed in different languages. At each iteration $t \in [2, N]$, the framework seeks to align a new candidate graph G_t with the previously synthesized global state $G_c^{(t-1)}$. For each type $S \in \{E, R\}$, we identify the optimal alignment set A_t^S by maximizing the sum of correspondence probabilities across all candidate pairs $(u_i, u_j) \in S_c^{(t-1)} \times S_t$:

$$A_t^S = \arg\max_A \sum_{(u_i,u_j)\in A} P(u_i \equiv u_j \mid \text{context}_t)$$

subject to an one-to-one mapping constraint. Here, context_t denotes the linearized states of the global and candidate graphs. We define the confidence score $\sigma(u_i, u_j)$ as the computed probability $P(u_i \equiv u_j \mid \text{context}_t)$ for each candidate pair. An alignment is accepted into the unified graph if its confidence score satisfies:

$$\sigma(u_i, u_j) \geq \tau,$$

where τ is a predefined threshold (e.g., $\tau = 0.90$ for high-precision synthesis).

3.2. Entity-Centric Graph Partitioning

To manage the high dimensionality of G and the finite context window W of the LLM, we implement an entity-centric partitioning strategy. Given a set of triplets $T = \{(h, r, t)\}$, we partition T into k batches $\{B_1, B_2, \ldots, B_k\}$ such that each batch B_m is formed by grouping triples by their head entity h:

$$B_m = \{(h, r, t) \in T \mid h \in E_{sub}\},$$

where E_{sub} is a subset of entities assigned to the m-th partition. This ensures that the local topological

```
You are an expert at knowledge graph entity
alignment. Your task is to identify corresponding
entities and relations between two knowledge
graphs that may be in different languages. You
analyze triplets from both graphs and discover
which entities refer to the same real-world
concept and which relations express the same
meaning.

You must respond with valid JSON only, no
additional text. Use this exact format:
{
    "entity_alignments": [
        {"entity_a": "label from Graph A",
        "entity_b": "label from Graph B",
        "confidence": 0.9,
        "reason": "brief explanation"}
    ],
    "relation_alignments": [
        {"relation_a": "label from Graph A",
        "relation_b": "label from Graph B",
        "confidence": 0.8,
        "reason": "brief explanation"}
     ]
}
```

(a) System Prompt

```
Here are triplets from Knowledge Graph A (zh_en):
Graph: zh_en
- 北京 首都 中国
- 中国 官方语言 汉语
...

Here are triplets from Knowledge Graph B (zh_en):
Graph: zh_en
- Beijing capital China
- China official_language Chinese_language
...

Identify which entities from Graph A correspond to
entities in Graph B.
Also identify matching relations between the two
graphs.
Return your answer as JSON with "entity_alignments"
and "relation_alignments" arrays.
```

(b) User Prompt

Figure 3: Details of the LLM prompting interface. The system prompt (left) defines the operational constraints and confidence thresholds, while the user prompt (right) serves as the data payload, containing the linearized triples from the source and target knowledge graphs.

neighborhood of an entity is preserved within a single processing unit, which is critical for structural disambiguation.

3.3. Exhaustive Batch Pairing

To ensure that every potential entity correspondence is evaluated, we implement an Exhaustive Batch Pairing strategy. Given the partitioned batch sets $\{B_1, \ldots, B_k\}$ from $G_c^{(t-1)}$ and $\{B'_1, \ldots, B'_{k'}\}$ from G_t, we construct a Cartesian product of all possible batch pairs:

$$\mathcal{P} = \{(B_i, B'_j) \mid 1 \le i \le k, 1 \le j \le k'\}$$

This results in $k \times k'$ discrete reasoning tasks. While computationally intensive, this exhaustive approach guarantees that every entity neighborhood in the source graph is compared against every neighborhood in the target graph, maximizing the discovery of cross-lingual semantic anchors without relying on pre-computed candidate selection or heuristic blocking.

3.4. Linearization and LLM Reasoning

The core alignment engine operates via a Linearization Function $\mathcal{L}(B)$, which transforms structural triplets into a plain-text sequence $\mathcal{S}$. For each triplet $(h, r, t) \in B$, the transformation is defined as:

$\mathcal{L}(h, r, t) \rightarrow$ "[head_label] [relation_label] [tail_label]"

The resulting strings from a specific batch pair (B_i, B'_j) are concatenated into a prompt P consisting of a system persona π and the linearized data. The prompts used for the LLM can be seen in Figure 3. The LLM performs zero-shot reasoning to discover a local mapping set $\hat{A}_{i,j}$ by maximizing the posterior probability of the alignment sequence:

$$\hat{A}_{i,j} = \arg\max_{\hat{A}} P(\hat{A} \mid \mathcal{L}(B_i), \mathcal{L}(B'_j), \pi)$$

where $B_i \subseteq G_c^{(t-1)}$ and $B'_j \subseteq G_t$. This generative approach allows the model to leverage its internal cross-lingual semantic priors to bridge the gap between disparate surface forms without requiring explicit seed translations.

3.5. Robust Response Parsing and Resolution

Due to the probabilistic nature of Large Language Model (LLM) generation, we implement a multi-stage Response Parser to handle failure modes such as output truncation or syntax errors. The parsing function f_{parse} recovers partial JSON objects by identifying the last complete element in the sequence and closing unclosed braces to salvage alignment items that would otherwise be lost to token limits. The resolved textual labels are subsequently mapped back to internal dataset identifiers via a case-insensitive lookup function f_{res} : Label $\rightarrow$ ID. To ensure global consistency, we apply a Max-Confidence Aggregation rule to deduplicate predictions appearing across multiple batch pairs. For any source entity or relation u_i, the final alignment u_j^* is determined by:

$$u_j^* = \text{arg max}_{u_j} \{\sigma(u_i, u_j) \mid (u_i, u_j, \sigma) \in \hat{A}_{total}\}$$

This aggregation effectively filters low-confidence noise and ensures that only the most probable semantic correspondences which of those that appear with the highest model certainty across various structural contexts are integrated into the unified global graph.

3.6. Agglomeration Fusion and Schema Unification

The framework performs a dual-track unification of the entity and relation sets. For every identified pair in A_t, the framework collapses the language-specific nodes into a single unified node that inherits a multilingual label set and all incident edges. To maintain schema consistency, we perform Relation Folding, where aligned predicates are merged into a consistent cross-lingual representation. The updated graph $G_c^{(t)}$ then serves as the base for the next iteration, maintaining a linear computational complexity $O(N)$ relative to the number of graphs.

4. Experiments and Results

This section provides a detailed quantitative and qualitative analysis of our LLM-based N-Graph fusion framework. We evaluate the system's ability to discover cross-lingual correspondences in a zero-shot environment, focusing on the trade-offs between precision, recall, and model confidence.

4.1. Experimental Setup

To assess the framework's efficacy, we utilize the DBP15K (Chinese-English) (Sun et al., 2017) dataset, which serves as the gold standard for cross-lingual entity alignment.

- Dataset Composition: The zh_en pair consists of two knowledge graphs with 15,000 ground-truth alignment pairs.
- Zero-Shot Evaluation: Unlike traditional supervised methods, our approach requires no training data. We evaluate against the full 15,000 pairs (combining both training and test sets) to measure the system's total discovery capability.
- Model Configuration: The pipeline employs Gemini 2.5 Flash with the temperature set to 0.0 to ensure reasoning consistency and minimize hallucinatory variations.
- Runtime and Infrastructure: The pipeline completes the exhaustive batch processing in approximately 5–6 hours.

4.2. Quantitative Performance Metrics

The system demonstrated a robust ability to identify semantic equivalencies without any prior exposure to the dataset. Table 1 outlines the primary results for the zh_en language pair.

Table 1: Primary Alignment Metrics (DBP15K zh_en)

Metric	Value
Total Batch Pairs	768
Unique Alignment Predictions	5,416
True Positives (TP)	3,543
False Positives (FP)	1,873
Precision	**65.4%**
Recall	**23.6%**
F1 Score	**34.7%**

The model's recall is currently bounded by its total prediction volume; the system generated 5,416 unique alignment correspondences out of the 15,000 possible ground-truth pairs. Despite this coverage bottleneck, the precision of the generated predictions remains robust at 65.4%, demonstrating strong initial performance for a purely zero-shot configuration.

4.3. Hits@k Analysis and Source Accuracy

The Hits@k metrics provide insight into the model's ranking capabilities, with Hits@1 measuring whether the correct target entity is identified as the top prediction.

Table 2: Hits@k Evaluation

k	Hits / 15,000	Rate	Acc. Predicted
1	3,516	23.4%	**88.3%**
5	3,543	23.6%	89.0%
10	3,543	23.6%	89.0%

As detailed in Table 2, when the model generates a prediction for a source entity, it identifies the correct target with an accuracy of **88.3%** at Hits@1. The minimal performance gap between Hits@1 and Hits@5 (a 0.7 percentage point difference) further underscores the model's high precision in its top-ranked predictions.

4.4. Impact of Confidence Filtering

The model's internal confidence score (σ) serves as a strong quality discriminator. True Positives exhibited a mean confidence of **0.980**, whereas False Positives averaged **0.738**.

As illustrated in Table 3, by applying a threshold of $\tau = 0.90$, the framework achieves a precision of

Table 3: Confidence Threshold (τ) Sensitivity Sweep

τ	Pred.	TP	Prec.	Rec.	F1
0.00	5,416	3,543	65.4%	23.6%	34.7%
0.80	4,309	3,540	82.2%	23.6%	36.6%
0.90	**4,002**	**3,520**	**88.0%**	**23.5%**	**37.0%**
0.95	3,483	3,219	92.4%	21.4%	34.8%

88.0% with negligible recall loss compared to the baseline.

5. Conclusion

This preliminary study introduced a modular, zero-shot framework for multilingual N-graph fusion leveraging the in-context reasoning capabilities of Large Language Models. By implementing an entity-centric partitioning strategy and a max-confidence aggregation rule, we successfully mitigated the challenges of structural heterogeneity and transitive error propagation that typically plague N-graph scenarios. Our results on the DBP15K dataset demonstrate that while achieving high recall remains a challenge, likely due to the inherent trade-offs between precision and recall in zero-shot alignment, the precision of LLM-based alignments remains robust. This is particularly evident when filtered by a 0.90 confidence threshold, suggesting that LLMs can indeed serve as a universal semantic bridge for synthesizing global knowledge bases without the need for expensive, manually curated seed alignments.

6. Limitations and Future Work

This study presents a preliminary evaluation of the proposed N-graph fusion framework, and several constraints must be acknowledged. First, due to computational resource constraints and the current availability of standardized multi-graph benchmarks, our experiments were conducted on binary language pairs ($N = 2$) from the DBP15K dataset. While this validates the core entity-centric partitioning and confidence-based alignment, it does not fully capture the cumulative complexity of higher-order sequential fusion. Consequently, the empirical extent to which our "rolling" strategy mitigates transitive error propagation in deep fusion chains ($N > 2$) remains a subject for future investigation.

Second, our current problem formulation focuses on strict semantic equivalence. In "true" multilingual scenarios, cross-lingual discrepancies often involve nuanced subsumption hierarchies and differences in semantic granularity (e.g., a concept in one language may broader or narrower than its counterpart in another). Because our zero-shot prompting does not yet explicitly instruct the LLM to detect these partial equivalences, the framework may experience information loss when synthesizing ontologies with divergent expressive granularities. Furthermore, the inherent Anglo-centric bias of many multilingual LLMs may impair alignment precision for low-resource language pairs where implicit world knowledge is sparse.

To address these limitations, several avenues for optimization remain. A primary direction involves transitioning from exhaustive batch pairing to a multi-stage heuristic blocking and semantic indexing strategy. By utilizing lightweight multilingual embedding models to pre-index entity neighborhoods, the framework could perform targeted candidate selection, directing LLM reasoning only toward high-probability pairs. This would reduce the computational overhead of the search space while improving recall by focusing resources on semantically ambiguous clusters.

Furthermore, future work will move beyond localized triplet-level views by incorporating global structural context through hierarchical prompting. Providing the LLM with a schema-level overview or a summary of high-degree hub nodes could significantly mitigate disambiguation errors for entities with identical labels but distinct ontological roles. Beyond architectural shifts, exploring hybrid architectures offers a path toward self-improving systems, where LLM-discovered correspondences serve as "silver seeds" to bootstrap and fine-tune traditional Graph Neural Network (GNN) aligners. Finally, evaluating the framework across a broader spectrum of domain-specific datasets (e.g., medical or legal ontologies) will be essential to assess how variations in specialized pre-training influence the precision of large-scale, cross-lingual knowledge synthesis.

7. Bibliographical References

Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. *Advances in neural information processing systems*, 26.

Muhao Chen, Yingtao Tian, Mohan Yang, and Carlo Zaniolo. 2016. Multilingual knowledge graph embeddings for cross-lingual knowledge alignment. *arXiv preprint arXiv:1611.03954*.

Nadezhda Chirkova and Vassilina Nikoulina. 2024. Zero-shot cross-lingual transfer in instruction tuning of large language models. In *Proceedings of the 17th International Natural Language Generation Conference*, pages 695–708.

Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Alex Vaughan, et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard De Melo, Claudio Gutierrez, Sabrina Kirrane, José Emilio Labra Gayo, Roberto Navigli, Sebastian Neumaier, et al. 2021. Knowledge graphs. *ACM Computing Surveys (Csur)*, 54(4):1–37.

Zijie Huang, Zheng Li, Haoming Jiang, Tianyu Cao, Hanqing Lu, Bing Yin, Karthik Subbian, Yizhou Sun, and Wei Wang. 2022. Multilingual knowledge graph completion with self-supervised adaptive graph alignment. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 474–485.

Nan Huo, Reynold Cheng, Ben Kao, Wentao Ning, Nur Al Hasan Haldar, Xiaodong Li, Jinyang Li, Mohammad Matin Najafi, Tian Li, and Ge Qu. 2024. Zeroea: A zero-training entity alignment framework via pre-trained language model. *Proceedings of the VLDB Endowment*, 17(7):1765–1774.

Shaoxiong Ji, Shirui Pan, Erik Cambria, Pekka Marttinen, and Philip S Yu. 2021. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE transactions on neural networks and learning systems*, 33(2):494–514.

Bowen Jin, Gang Liu, Chi Han, Meng Jiang, Heng Ji, and Jiawei Han. 2024. Large language models on graphs: A comprehensive survey. *IEEE Transactions on Knowledge and Data Engineering*, 36(12):8622–8642.

Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick Van Kleef, Sören Auer, et al. 2015. Dbpedia–a large-scale, multilingual knowledge base extracted from wikipedia. *Semantic web*, 6(2):167–195.

Rumana Ferdous Munne, Md Mostafizur Rahman, and Yuji Matsumoto. 2025. Entity profile generation and reasoning with llms for entity alignment. *Findings of the Association for Computational Linguistics: EMNLP*, pages 20073–20086.

Shirui Pan, Linhao Luo, Yufei Wang, Chen Chen, Jiapu Wang, and Xindong Wu. 2024. Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering*, 36(7):3580–3599.

Ciyuan Peng, Feng Xia, Mehdi Naseriparsa, and Francesco Osborne. 2023. Knowledge graphs: Opportunities and challenges. *Artificial intelligence review*, 56(11):13071–13102.

Aleksandr Perevalov, Andreas Both, and Axel-Cyrille Ngonga Ngomo. 2024. Multilingual question answering systems for knowledge graphs–a survey. *Semantic Web*, 15(5):2089–2124.

Zequn Sun, Wei Hu, and Chengkai Li. 2017. Cross-lingual entity alignment via joint attribute-preserving embedding. In *International semantic web conference*, pages 628–644. Springer.

Gemini Team, Rohan Anil, Sebastian Borgeaud, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M Dai, Anja Hauth, Katie Millican, et al. 2023. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*.

Vinh Tong, Dat Quoc Nguyen, Trung Thanh Huynh, Tam Thanh Nguyen, Quoc Viet Hung Nguyen, and Mathias Niepert. 2022. Joint multilingual knowledge graph completion and alignment. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 4646–4658.

Zhichun Wang, Qingsong Lv, Xiaohan Lan, and Yu Zhang. 2018. Cross-lingual knowledge graph alignment via graph convolutional networks. In *Proceedings of the 2018 conference on empirical methods in natural language processing*, pages 349–357.

Yuting Wu, Xiao Liu, Yansong Feng, Zheng Wang, Rui Yan, and Dongyan Zhao. 2019. Relation-aware entity alignment for heterogeneous knowledge graphs. *arXiv preprint arXiv:1908.08210*.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, et al. 2025. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*.

Hsiu-Wei Yang, Yanyan Zou, Peng Shi, Wei Lu, Jimmy Lin, and Xu Sun. 2019. Aligning cross-lingual entities with multi-aspect information. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4431–4441.

Linyan Yang, Jingwei Cheng, and Fu Zhang. 2024a. Advancing cross-lingual entity alignment with large language models: Tailored sample segmentation and zero-shot prompts. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 8122–8138.

Yaming Yang, Zhe Wang, Ziyu Guan, Wei Zhao, Weigang Lu, Xinyan Huang, Jiangtao Cui, and Xiaofei He. 2024b. Aligning multiple knowledge graphs in a single pass. *arXiv preprint arXiv:2408.00662*.

Xiaoming Zhang, Wencheng Zhang, and Huiyong Wang. 2023. Cross-language entity alignment based on dual-relation graph and neighbor entity screening. *Electronics*, 12(5):1211.

Xiang Zhao, Weixin Zeng, Jiuyang Tang, Wei Wang, and Fabian M Suchanek. 2020. An experimental study of state-of-the-art entity alignment approaches. *IEEE Transactions on Knowledge and Data Engineering*, 34(6):2610–2625.

Yuqi Zhu, Xiaohan Wang, Jing Chen, Shuofei Qiao, Yixin Ou, Yunzhi Yao, Shumin Deng, Huajun Chen, and Ningyu Zhang. 2024. Llms for knowledge graph construction and reasoning: Recent capabilities and future opportunities. *World Wide Web*, 27(5):58.