Results of CMatch in OAEI 2025

Guilherme Sousa[1,*], Rinaldo Lima[2] and Cassia Trojahn[3]

[1]Institut de Recherche en Informatique de Toulouse, France
[2]Universidade Federal Rural de Pernambuco, Recife, Brazil
[3]Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, France

Abstract

This paper presents CMatch (Complex Matcher), an LLM-based ontology matching approach designed for the Complex Track of the Ontology Alignment Evaluation Initiative (OAEI) 2025. CMatch addresses two core challenges in complex ontology matching: the combinatorial explosion of candidate subgraph pairs and the generation of expressive, logic-based correspondences (e.g., n:m correspondences involving class unions or property compositions).

Keywords

Ontology Matching, Complex Alignment, Large Language Models, Ontology Modularization

1. Presentation of the System

CMatch is a system specifically engineered for the challenges of complex ontology matching based on Large Language Models (LLM). To mitigate LLM context limitations and ensure scalability, the architecture employs a search space reduction strategy via the PageRank algorithm [1], which constructs ontology modules around high-centrality entities. These modules are subsequently verbalized and paired using embedding models to isolate semantically relevant ontology fragments. Finally, structured prompts enriched with few-shot task examples guide the LLM to generate precise alignments in the standardized EDOAL format[1].

Performance analysis on the OAEI 2025 Complex track datasets (Conference, GeoLink, Hydrography, Enslaved, and Taxon) reveals that CMatch excels in precision (most notably in GeoLink), but suffers from low recall due to aggressive space reduction. While proving the viability of LLMs for complex alignment without training data, the results highlight the need for enhanced module construction, instance integration, and model fine-tuning to improve overall performance.

2. State, Purpose, General Statement

Traditional alignment techniques, which predominantly rely on lexical metrics or instance co-occurrence, are generally ill-suited for complex ontology matching. Complex alignment tasks (involving $1:n$ or $m:n$ relations) often exhibit semantic heterogeneities that cannot be captured by surface-level syntactic comparison. Furthermore, the efficacy of extensional methods is frequently compromised by the prevalence of unpopulated or sparsely populated ontologies.

Two of the primary challenges in Complex Ontology Matching are the combinatorial explosion of the search space (i.e., the vast number of candidate subgraphs) and the difficulty of synthesizing the correct logical constructors to combine them [2]. For instance, establishing the correspondence $FullName \equiv FirstName \oplus LastName$ requires more than merely locating the constituent entities;

OM 2025: The 20th International Workshop on Ontology Matching collocated with the 24th International Semantic Web Conference (ISWC 2025), November 2nd, 2025, Nara, Japan

*Corresponding author.

guilherme.santos-sousa@irit.fr (G. Sousa); rinaldo.jose@ufrpe.br (R. Lima); cassia.trojahn-dos-santos@univ-grenoble-alpes.fr (C. Trojahn)

0000-0002-2896-2362 (G. Sousa); 0000-0002-1388-4824 (R. Lima); 0000-0003-2840-005X (C. Trojahn)

[1]https://ns.inria.fr/edoal/1.0/

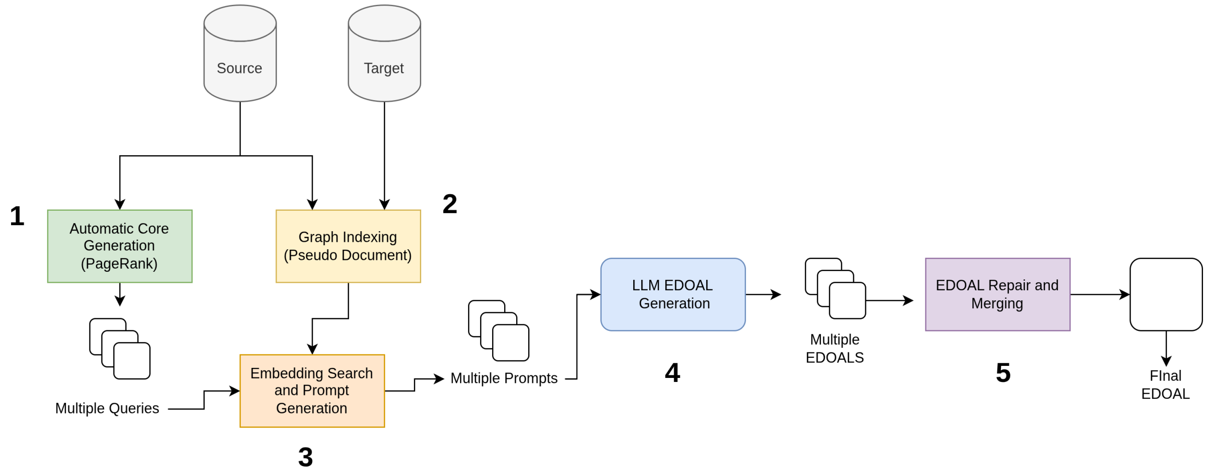

Figure 1: Architecture of CMatch.

it necessitates explicitly defining the aggregation logic (e.g., string concatenation) required to reconstruct the semantics of "FullName".

Although embedding-based methods represent a significant advancement over lexical metrics, they fail to address the generative aspects of complex alignment. The logical axioms needed to combine subgraphs are extrinsic to the ontology's latent space and cannot be retrieved via simple distance measures. As a result, the matcher must go beyond similarity detection to evaluate candidate subgraphs and construct the appropriate logical rules.

Large Language Models (LLMs) have recently emerged as promising tools for complex ontology matching due to their capacity for flexible prompt-based inference [3]. However, their deployment is challenged by scalability issues (context window limits and high computational resources) and the difficulty of constraining their output to standardized, machine-readable formats like EDOAL employed in the OAEI Complex Matching track.

This work introduces an approach that enhances LLM-based complex ontology matching through a hybrid strategy of search space reduction and structured prompting. It aims to enhance performance for both simple (1:1) and complex (n:m) correspondences without the need for manually annotated training data from the target domain. The architecture and implementation details of the proposed solution are delineated in the next sections.

3. Techniques Used

Building upon [4], the proposed approach implements a two-step workflow consisting of search space reduction and alignment generation (see Figure 1 for an architectural overview).

3.1. Space Reduction

The main hypothesis of this work is that valid correspondences are most probable between entities residing in conceptually aligned modules extracted from the source and target ontologies. The module definition proposed by [5] is adopted, wherein each module is anchored by a central entity and encapsulates its immediate semantic neighborhood. This modular approach effectively prunes the search space by restricting comparisons to intra-module entities; the system eliminates the computational overhead of processing irrelevant entities outside the module boundaries.

As illustrated in Figure 1 (Step 1), the reduction component begins by identifying source-side module seeds. This component relies on the PageRank algorithm to identify the top-k most authoritative entities (seeds) in the source ontology, which function as the 'cores' for module construction. These cores are subsequently expanded via a graph traversal algorithm that incorporates the local neighborhood up to a specific depth. This centrality-based selection is performed asymmetrically (source-side only) to avoid conceptual misalignment, as structural prominence in one ontology does not imply equivalence in another. Additionally, to mitigate the high dimensionality of complex datasets, an upper bound is imposed on the module size.

Following module extraction, a verbalization strategy is employed to transform the graph structure into a textual representation, enabling similarity comparison via embedding models. This process, illustrated as Step 2 in Figure 1, generates a 'virtual document' as proposed by [6]. The document aggregates the core entity with its structural context (including ancestors, descendants, incoming/outgoing properties, and disjoint classes), serialized in the Turtle syntax to preserve structure. While the source ontology relies on PageRank for core selection, the target ontology is processed comprehensively: all target entities are treated as core candidates and subjected to the same verbalization and module construction process to maximize recall.

Following verbalization, Step 3 starts the embedding generation phase for both source and target entities, employing the Qwen/Qwen3-Embedding-8B model[2]. For each source module, its similarity is computed against the target ontology, retrieving all target entities that exceed a predefined similarity threshold. Subsequently, a composite target module is synthesized by aggregating the local neighborhoods (nearby entities) of all retrieved candidates. To conclude this phase, both the source module and the newly constructed target module are serialized in Turtle syntax. These structured representations serve as the context for constructing the final prompts in the subsequent matching step.

3.2. Prompt Generation

To ensure strict adherence to the EDOAL format and enhance task comprehension, two examples are provided within the prompt (few-shot prompting). This strategy constrains the model's output to the required syntax while clarifying the alignment objective. The specific prompt template employed is detailed below.

System: You are a Complex Ontology Matching expert.
User: Based on the examples of the task of complex ontology alignment between the ontologies below, with the results written in EDOAL format (for the sake of space, the examples are committed as the prompt is quite long):
{Example 1}
[...]
{Example N}
[...]
Write a file in EDOAL format containing the complex alignment between the input ontologies <ontology1> and <ontology2>. You don't need to explain yourself. Just give as a response the resulting alignment file without saying anything else. Given the two ontologies below:
<ontology1>
{Ontology 1}
</ontology1>
<ontology2>
{Ontology 2}
</ontology2>

The complete set of few-shot examples employed for the alignment generation is accessible via the associated GitLab repository GitLab[3].

[2]https://huggingface.co/Qwen/Qwen3-Embedding-8B
[3]https://gitlab.irit.fr/melodi/ontology-matching/llm/-/tree/main/prompt_examples?ref_type=heads

3.3. Alignment Generation

The next step of the process is to feed the prompts into Qwen/Qwen3-14B[4] for alignment generation. This model was chosen for its ability to articulate reasoning chains, thereby providing a rationale that supports higher-quality matches. The extraction of these alignments from the reasoning output is illustrated in Step 4 of Figure 1. The pipeline concludes with Step 5, where the data is merged and refined through an error correction strategy to generate a unified alignment file.

4. Results

The source code for the matcher is publicly available via GitLab[5]. In the current OAEI campaign, CMatch was evaluated on five datasets within the Complex Track: Conference, GeoLink, Hydrography, Enslaved, and Taxon, employing two distinct metrics. The Metric 1 presented in [7] and the Metric 2 in [8]. To facilitate a comprehensive overview, the presented results represent the average performance across all datasets. Ontology pairs where matchers failed to execute were treated as zero values for the calculation of the mean. For the sake of brevity, Table 1 lists only those systems that achieved non-zero results.

5. General Comments

5.1. Comments on the Results

The employed search space reduction and modularization strategies effectively mitigated scalability constraints, enabling the LLM to process all target datasets. However, while execution was successful, the system yielded empty alignment sets in certain cases. On the other hand, the structured prompting mechanism successfully enforced strict adherence to the EDOAL format, thereby facilitating the automated evaluation of complex matching tasks.

Results indicate that CMatch favors precision at the expense of recall, with recall values remaining below (< 0.1) for most datasets. In comparative analysis, however, CMatch proves effective in specific domains, yielding the highest precision in Geolink (Metric 1) and achieving higher F-measure scores than participating systems in Conference, Hydrography, and Geolink (Metric 2).

These results highlight CMatch as a promising architecture for instance-free datasets (schema-level matching). However, the consistently low recall suggests that the current search space reduction strategy is overly aggressive, likely discarding valid correspondences that fall outside the boundaries of the selected modules. Despite such limitations, CMatch demonstrates the viability of applying LLMs to complex ontology alignment tasks without the need for domain-specific fine-tuning. The structured prompting strategy proves effective in constraining the LLM to generate syntactically valid, machine-readable alignment formats.

5.2. Future Improvements

Future improvements rely on refining strategies for both space reduction and matching. This includes deploying more sophisticated modularization, partitioning, and growing algorithms, as well as optimizing verbalization and embedding processes. Moreover, a robust mechanism for handling ontology instances must be developed, as the proposed framework currently lacks a dedicated application for such data.

Regarding the matching step, future improvements will also investigate the impact of domain-adapted LLMs via fine-tuning. It is planned to incorporate automated alignment repair mechanisms to filter hallucinations and ensure high-quality, logically consistent alignments.

[4]https://huggingface.co/Qwen/Qwen3-14B
[5]https://gitlab.irit.fr/melodi/ontology-matching/llm

Metric 1				
Dataset	Matcher	Precision	Recall	F-measure
Populated Conference	CANARD (2020)	**0.309**	**0.420**	**0.352**
Conference	AMLC (2020)	0.169	0.031	0.058
	CMatch (2025)	0.231	0.030	0.048
	Matcha (2025)	**0.235**	**0.184**	**0.204**
Hydrography	**CMatch (2025)**	0.157	0.068	0.095
	Matcha (2025)	**0.485**	**0.234**	**0.285**
Geolink	**CMatch (2025)**	**0.644**	0.176	0.276
	Matcha (2025)	0.287	**0.467**	**0.355**
Populated Geolink	AROA (2020)	0.472	**0.492**	**0.482**
	CANARD (2020)	**0.481**	0.453	0.467
	Matcha (2025)	0.245	0.391	0.301
Enslaved	CANARD (2020)	**0.049**	0.188	0.078
	Matcha (2025)	0.048	**0.688**	**0.091**
Bio	Matcha (2025)	**0.516**	**0.516**	**0.516**
Metric 2				
Dataset	Matcher	Precision	Recall	F-measure
Populated Conference	CANARD (2020)	**0.224**	**0.545**	**0.314**
Conference	**CMatch (2025)**	**0.441**	**0.054**	**0.093**
Hydrography	**CMatch (2025)**	**0.350**	**0.069**	**0.103**
	Matcha (2025)	0.181	0.054	0.082
Geolink	AMLC (2020)	0.013	0.004	0.006
	CMatch (2025)	**0.773**	**0.056**	**0.104**
	Matcha (2025)	0.003	0.002	0.002
Populated Geolink	CANARD (2020)	0.448	0.182	0.258
	Matcha (2025)	0.002	0.004	0.003
	AROA (2020)	**0.706**	**0.267**	**0.388**
Enslaved	AMLC (2020)	0.005	0.002	0.003
	CANARD (2020)	**0.218**	**0.166**	**0.189**
	Matcha (2025)	0.001	0.002	0.001
Taxon	CANARD (2020)	**0.367**	**0.287**	**0.308**
	CMatch (2025)	0.135	0.014	0.025

Table 1
Performance of participating systems in the OAEI 2025 Complex Track evaluated using Metric 2. The highest scores in each category are marked in bold, and the proposed CMatch approach is highlighted.

6. Conclusion

This paper presented the performance of CMatch, a novel LLM-based matching system, within the context of OAEI 2025. To address the intricacies of complex alignment, the proposed approach combines ontology modularization for efficient search space reduction with precise prompt engineering. This architecture allows CMatch to harness the power of LLMs while overcoming inherent limitations regarding large-scale processing and output structure enforcement.

Future work will explore supervised fine-tuning on alignment data, optimize module extraction algorithms, and integrate A-Box, instance-level information. These enhancements aim to significantly improve recall while preserving the high precision baseline of the current architecture.

In summary, CMatch constitutes a promising step in LLM-driven ontology alignment. It offers a generalized and flexible framework capable of handling diverse heterogeneity scenarios, effectively operating across a wide spectrum of ontologies.

Declaration on Generative AI

During the preparation of this work, the authors used Grammarly to grammar and spell check, and improve text readability. The authors reviewed and edited the content as needed to assume full responsibility for the published content.

References

[1] L. Page, S. Brin, R. Motwani, T. Winograd, The PageRank citation ranking: Bringing order to the web., Technical Report, Stanford infolab, 1999.

[2] É. Thiéblin, O. Haemmerlé, N. Hernandez, C. Trojahn, Survey on complex ontology matching, Semantic Web 11 (2020) 689–727. URL: https://doi.org/10.3233/SW-190366. doi:10.3233/SW-190366.

[3] R. Amini, S. S. Norouzi, P. Hitzler, R. Amini, Towards complex ontology alignment using large language models, in: S. Tiwari, B. Villazón-Terrazas, F. Ortiz-Rodríguez, S. Sahri (Eds.), Knowledge Graphs and Semantic Web - 6th International Conference, KGSWC 2024, Paris, France, December 11-13, 2024, Proceedings, volume 15459 of *Lecture Notes in Computer Science*, Springer, 2024, pp. 17–31. URL: https://doi.org/10.1007/978-3-031-81221-7_2. doi:10.1007/978-3-031-81221-7_2.

[4] G. H. S. Sousa, R. Lima, C. Trojahn, Towards generating complex alignments with large language models via prompt engineering, in: E. Jiménez-Ruiz, O. Hassanzadeh, C. Trojahn, S. Hertling, H. Li, P. Shvaiko, J. Euzenat (Eds.), Proceedings of the 19th International Workshop on Ontology Matching co-located with the 23rd International Semantic Web Conference (ISWC 2024), Baltimore, USA, November 11, 2024, volume 3897 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024, pp. 43–56. URL: https://ceur-ws.org/Vol-3897/om2024_LTpaper4.pdf.

[5] C. Shimizu, K. Hammar, P. Hitzler, Modular ontology modeling, Semantic Web 14 (2023) 459–489. URL: https://doi.org/10.3233/SW-222886. doi:10.3233/SW-222886.

[6] F. Schneider, S. Dash, S. Bagchi, N. Mihindukulasooriya, A. M. Gliozzo, NLFOA: natural language focused ontology alignment, in: K. B. Venable, D. Garijo, B. Jalaian (Eds.), Proceedings of the 12th Knowledge Capture Conference 2023, K-CAP 2023, Pensacola, FL, USA, December 5-7, 2023, ACM, 2023, pp. 114–121. URL: https://doi.org/10.1145/3587259.3627560. doi:10.1145/3587259.3627560.

[7] M. C. Silva, D. Faria, C. Pesquita, Cmomgen: Complex multi-ontology alignment via pattern-guided in-context learning, arXiv preprint arXiv:2510.21656 (2025).

[8] G. H. S. Sousa, R. Lima, C. T. dos Santos, On evaluation metrics for complex matching based on reference alignments, in: E. Curry, M. Acosta, M. Poveda-Villalón, M. van Erp, A. K. Ojo, K. Hose, C. Shimizu, P. Lisena (Eds.), The Semantic Web - 22nd European Semantic Web Conference, ESWC 2025, Portoroz, Slovenia, June 1-5, 2025, Proceedings, Part I, volume 15718 of *Lecture Notes in Computer Science*, Springer, 2025, pp. 77–93. URL: https:doi.org/10.1007/978-3-031-94575-5_5. doi:10.1007/978-3-031-94575-5_5.