# DogMa results for OAEI 2025

Péter Kardos[1,†], Máté Vass[1,†], Miklós Krész[2,3] and Richárd Farkas[1,*]

[1]*University of Szeged, 6720 Szeged, Dugonics tér 13, Hungary*
[2]*University of Primorska, 6000 Koper, Titov trg 4, Slovenia*
[3]*Innorenew CoE, 6310 Izola, Livade 6a, Slovenia*

**Abstract**
This paper presents the results of the DogMa Matcher in the OAEI 2025 competition. DogMa is a Large Language Model-based system for the Knowledge Graph Entity Matching (KGEM) task in the OAEI 2025 Knowledge Graph Track. The system formulates entity matching as a Retriever–LLM selection problem, where entities are represented as concise, graph-aware textual summaries. Each entity in the source Knowledge Graph is summarized using an instruction-tuned LLM that integrates both textual descriptions and local relational context. Candidate entities are then retrieved from the target graph via dense vector similarity, and a second LLM determines whether any candidate represents the same real-world entity. This combination of graph-informed summarization and selective LLM reasoning yields high precision and recall across heterogeneous Knowledge Graphs, outperforming embedding-based methods in our internal evaluations. DogMa demonstrates that context-engineered LLMs can effectively exploit semi-structured graph information for accurate and interpretable entity alignment.

**Keywords**
Knowledge Graphs, Ontology Alignment, Entity Matching, Large Language Models

## 1. Presentation of the system

### 1.1. State, purpose, general statement

In recent years, large language models (LLMs) such as ChatGPT[1] have gained remarkable popularity across a wide range of natural language processing tasks. However, when applied to complex reasoning problems such as Ontology Alignment, LLMs face several challenges related to limited context window sizes and the degradation of attention mechanisms when processing lengthy or densely structured inputs. As highlighted in recent surveys on context engineering for LLMs [1], the structure and scope of the input context play a crucial role in determining the quality of model outputs. This underscores the need for carefully designed prompts, task decomposition, and structured workflows to ensure reliable and accurate performance.

Several ontology matching systems have begun to incorporate LLMs into their pipelines to address these challenges. For instance, OLaLa [2] and SORBET [3] integrate LLM-based components to enhance candidate generation and evaluation. A typical approach is to adopt a two-stage process: (1) generating candidate correspondence pairs, and (2) leveraging the LLM to assess and select the most plausible alignments. This integration demonstrates the growing potential of LLMs to augment traditional ontology matching techniques, while also revealing open questions regarding scalability, interpretability, and context management.

In this work, we aim to enhance the representation of entites in order to improve the quality of both candidate generation and selection. Our approach leverages LLMs not only in the second stage but also

*OM-2025: The 20th International Workshop on Ontology Matching collocated with the 22nd International Semantic Web Conference ISWC-2025, November 2nd, 2025, Nara, Japan*

* Corresponding author.
† These authors contributed equally.
kardos@inf.u-szeged.hu (P. Kardos); vassmate@inf.u-szeged.hu (M. Vass); miklos.kresz@innorennew.eu (M. Krész); rfarkas@inf.u-szeged.hu (R. Farkas)
0000-0003-4896-6755 (P. Kardos); 0009-0005-7002-7801 (M. Vass); 0000-0002-7547-1128 (M. Krész); 0000-0001-7019-2632 (R. Farkas)



[1] URL: https://chatgpt.com/

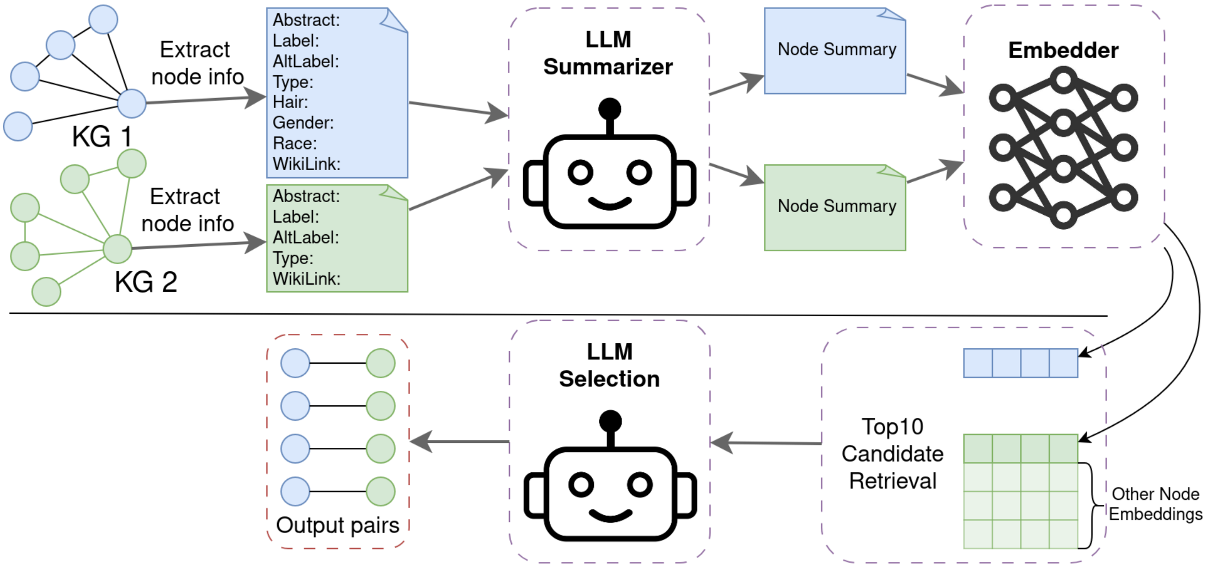


**Figure 1:** Overview of the DogMa workflow.

in the initial candidate selection process, thereby enabling a more robust and comprehensive alignment pipeline.

## 1.2. Specific techniques used

Our proposed system follows a Retriever–Augmented Generation (RAG)-like design with three main stages:

1. **Entity Representation (Dogtag) Generation:** Each node in the input KGs is converted into a compact textual description called a *Dogtag*. These are produced by an LLM-based summarizer that compresses node-level textual content (labels, abstracts, types, infoboxes) and incorporates salient information from local graph neighbors like wiki mention links or other relations.
2. **Candidate Retrieval:** For each entity in the smaller KG (anchor entity), we retrieve the top-$k$ most similar entities from the larger KG using an SBert [4] embedder model (`BAAI/bge-large-en-v1.5`).
3. **LLM Selection:** The retrieved candidates are then passed to an instruction-tuned LLM (`Llama-3.1-70B-Instruct` [5]), which decides whether any candidate refers to the same real-world entity. The LLM can also abstain if no suitable match exists.

All operations are performed in inference mode only, with deterministic decoding (temperature = 0, sampling disabled). The system enforces one-to-one alignments through post-hoc deduplication.

The prompt used for the LLM-based selection step is shown below:

```
TASK: You will be given a description of an anchor entity and a list of candidate entities, all
    formatted with <EXAMPLE> tags. Your task is to:
- Identify the candidate entity that is the same as the anchor entity.
- Return the ID number of the matching candidate entity.
- If none of the candidates match the anchor entity, return -1.
- Answer with the ID (or -1) only, no explanations
###
<ANCHOR>
{anchor}
</ANCHOR>
###
<EXAMPLE>
ID: 1
Summary 1
</EXAMPLE>
<EXAMPLE>
```

```
ID: 2
Summary 2
</EXAMPLE>
...
###
Answer:
```

### 1.3. Adaptations made for the evaluation

For the OAEI 2025 Knowledge Graph Track, the following adaptations were applied:

- **Input preprocessing:** OWL files were parsed into node-level representations using RDFLib. Entity attributes, labels, and neighbor relations were serialized into a textual context for summarization.
- **Summarization model:** Dogtags were generated by prompting `Llama-3.1-70B-Instruct` with a zero-shot summarization template designed to include both textual and relational content.
- **Retriever configuration:** Top-$k$ retrieval was set to $k = 10$ using cosine similarity of `bge-large-en-v1.5` embeddings.
- **Merging alignments:** To improve recall while preserving precision, the final alignment was obtained by merging the high-confidence *Exact match* pairs (entity pairs with corresponding labels) with additional matches from the LLM-based method. A dynamic similarity threshold — set as the median similarity score of the exact matches — filtered secondary pairs, which were only added if they exceeded this threshold and did not violate 1:1 consistency.

### 1.4. Link to the system and parameters file

DogMa can be downloaded from https://github.com/kiscsonti/DogMa-Ontology-Matcher alongside the alignment files.

## 2. Results

Our system was evaluated on the Knowledge Graph Track KG pairs. We present our results in Table 1.

| | mcu-marvel | | | memoryalpha-memorybeta | | | stexpanded-memoryalpha | | | swg-starwars | | | swtor-starwars | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Method** | **P** | **R** | **F1** | **P** | **R** | **F1** | **P** | **R** | **F1** | **P** | **R** | **F1** | **P** | **R** | **F1** |
| ExactMatch | **88.39** | 66.22 | 75.72 | 94.49 | 88.23 | 91.25 | **96.59** | 90.89 | 93.66 | **93.67** | 62.00 | 74.61 | **94.20** | 88.66 | 91.35 |
| DogMa | 52.89 | 45.05 | 48.66 | **95.46** | 86.90 | 90.98 | 95.22 | 91.74 | 93.44 | 93.44 | 80.11 | 86.26 | 90.76 | 86.56 | 88.61 |
| DogMa Merged | 82.61 | **78.70** | **80.61** | 92.93 | **93.64** | **93.28** | 94.83 | **94.94** | **94.89** | 92.06 | **85.82** | **88.83** | 93.67 | **93.21** | **93.44** |

**Table 1**
Results comparison across the Knowledge Graph track graph pairs and methods

## 3. General Comments

### 3.1. Comments on the Results

**Strengths:** DogMa demonstrates that high-quality textual summarization of graph nodes enables robust retrieval and matching. The LLM Selector effectively reasons over candidate sets and leverages implicit world knowledge to resolve difficult equivalences.

**Weaknesses:** Our approach is computationally expensive: each node in the smaller KG triggers one LLM call. This limits scalability for very large graphs.

### 3.2. Ways to Improve the System

Future improvements include:

- Reducing cost via lightweight selectors or small LLMs distilled from the current system.
- Incorporating global structural information (e.g., communities, influence paths) beyond local neighbors.
- Exploring prompt-order-invariant selection mechanisms to mitigate positional bias.

### 3.3. Comments on the OAEI Procedure

The OAEI framework provides a solid and reproducible evaluation setup. However, due to partial gold standards, true precision and recall are underestimated. We recommend augmenting gold alignments to include additional verified correspondences.

## 4. Consclusions

We introduced **DogMa**, an LLM-based system for Knowledge Graph Entity Matching in the OAEI 2025 track. DogMa formulates alignment as a Retriever–LLM selection task, where entities are represented through graph-aware textual summaries and matched via dense retrieval followed by LLM-based verification.

Results indicate that integrating local relational context into entity representations substantially improves matching quality, yielding high precision and recall across heterogeneous graphs. Despite its computational overhead, DogMa shows that context-engineered LLMs can effectively reason over semi-structured data, bridging symbolic and neural alignment methods.

## Acknowledgments

This research has been supported by the European Union project RRF-2.3.1-21-2022-00004 within the framework of the Artificial Intelligence National Laboratory. Miklós Krész was supported by the Slovenian Research and Innovation Agency (ARIS) through research program P1-0404 and by the research program Cognicom (0013103) at the University of Primorska. He is also grateful for the support of ARIS grant J7-60122.

## Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT for rephrasing, grammar and spell checking. After using these tools, the generated text was reviewed and edited and the authors takes full responsibility for the publication's content.

## References

[1] L. Mei, J. Yao, Y. Ge, Y. Wang, B. Bi, Y. Cai, J. Liu, M. Li, Z.-Z. Li, D. Zhang, C. Zhou, J. Mao, T. Xia, J. Guo, S. Liu, A survey of context engineering for large language models, 2025. URL: https://arxiv.org/abs/2507.13334. arXiv:2507.13334.

[2] S. Hertling, H. Paulheim, Olala: Ontology matching with large language models, in: Proceedings of the 12th Knowledge Capture Conference 2023, K-CAP '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 131–139. doi:10.1145/3587259.3627571.

[3] F. Gosselin, A. Zouaq, Sorbet: A siamese network for ontology embeddings using a distance-based regression loss and bert, in: The Semantic Web – ISWC 2023: 22nd International Semantic Web Conference, Athens, Greece, November 6–10, 2023, Proceedings, Part I, Springer-Verlag, Berlin,

Heidelberg, 2023, p. 561–578. URL: https://doi.org/10.1007/978-3-031-47240-4_30. doi:10.1007/978-3-031-47240-4_30.

[4] N. Reimers, I. Gurevych, Sentence-bert: Sentence embeddings using siamese bert-networks, 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), Association for Computational Linguistics, Hong Kong, China, 2019, pp. 3982–3992. doi:10.18653/v1/D19-1410.

[5] A. Grattafiori, A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Vaughan, et al., The llama 3 herd of models, arXiv preprint arXiv:2407.21783 (2024). doi:10.48550/arXiv.2407.21783. arXiv:2407.21783.