Results for BioGITOM in OAEI 2025

Samira Oulefki[1,*], Lamia Berkani[1] and Ladjel Bellatreche[2]

[1]Dep. of Artificial Intelligence and Data Sciences, Faculty of Informatics, USTHB, Bab Ezzouar 16111, Algiers, Algeria
[2]LIAS/ISAE-ENSMA, Poitiers, France

Abstract

BioGITOM is an advanced ontology matching (OM) system tailored for the biomedical domain, aiming to achieve precise ontology alignment and foster semantic interoperability. It integrates Sentence-SapBERT for generating domain-specific semantic embeddings from biomedical concept definitions, capturing fine-grained lexical and contextual nuances. These embeddings are further refined through a hybrid structural encoder that combines Graph Isomorphism Networks and Graph Transformers to jointly model local and global dependencies within ontology graphs. To efficiently handle large-scale biomedical ontologies, BioGITOM employs a FAISS-based Approximate Nearest Neighbor (ANN) retrieval strategy for fast and scalable candidate selection. This synergistic design enables the system to robustly manage the complexity and heterogeneity of biomedical data. In the OAEI 2025 Bio-ML benchmark, BioGITOM ranked first in two tasks (NCIT–DOID and OMIM–ORDO) and second in three tasks (SNOMED–NCIT (Pharm), SNOMED–NCIT (Neoplas), and SNOMED–FMA (Body)), achieving the best overall performance with an average F1-score of 0.804, outperforming ten competing systems.

Keywords

Ontology matching, deep learning, Graph Neural Network, graph transformer, graph isomorphism transformer, FAISS-based Approximate Nearest Neighbor

1. Presentation of the system

The biomedical domain has witnessed an unprecedented expansion of data repositories, each encoding critical knowledge for research, diagnosis, and clinical decision-making. However, these repositories are often modeled through independent ontologies that differ in structure, terminology, and granularity, leading to severe semantic heterogeneity. This heterogeneity hampers data interoperability and limits the potential for cross-repository integration and reasoning. Ontology Matching (OM) addresses this challenge by automatically identifying semantic correspondences between entities across heterogeneous ontologies [1], thereby enabling unified and interpretable biomedical knowledge graphs.

Traditional OM techniques, largely based on handcrafted rules, lexical heuristics, or external dictionaries, have shown limited scalability and adaptability when applied to the complex and large-scale nature of biomedical ontologies. In contrast, recent advances in Deep Learning (DL) and Graph Neural Networks (GNNs) have opened new perspectives for learning expressive concept representations that jointly capture semantics and structure.

BioGITOM (Biomedical Graph Isomorphism Transformer for Ontology Matching) is an advanced OM system that builds upon a series of prior research efforts. It represents the culmination of a progressive line of work initiated with the Multi-Head Attention Graph Isomorphism Network (MHAGINOM) [2] and further refined through the GNN-based hybrid framework [3, 4, 5]. These successive studies laid the foundation for integrating semantic and structural reasoning within OM.

Building on these advances, BioGITOM introduces a hybrid neural architecture that integrates semantic embeddings from Sentence-SapBERT, a sentence-level biomedical encoder fine-tuned on UMLS concept definitions, with a Graph Isomorphism Transformer (GIT) model that unifies the strengths of Graph Isomorphism Networks (GINs) [6] and Graph Transformers (GTs) [7]. This design enables

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.
soulefki@usthb.dz (S. Oulefki); lberkani@usthb.dz (L. Berkani); bellatreche@ensma.fr (L. Bellatreche)

BioGITOM to capture both local structural regularities and long-range semantic dependencies within ontology graphs.

To efficiently scale to millions of biomedical entities, BioGITOM employs a FAISS-based L2 Approximate Nearest Neighbor (ANN) [8] retrieval strategy, enabling rapid and accurate candidate selection during the matching process. This integration of domain-aware semantic encoding, graph-based structural reasoning, and efficient large-scale retrieval allows BioGITOM to achieve state-of-the-art accuracy and robustness.

1.1. State, Purpose, General Statement

BioGITOM is a specialized OM system developed to ensure semantic interoperability across heterogeneous biomedical knowledge sources. Its main goal is to enable consistent integration and alignment of diverse ontological frameworks, supporting large-scale data sharing and reasoning in biomedical research. By leveraging domain-aware semantic encoding and structural learning, BioGITOM provides an effective solution to the challenges posed by the diversity of biomedical terminologies and ontology structures.

1.2. Specific Techniques Used

BioGITOM leverages an advanced combination of neural and graph-based techniques to achieve high-precision matching by jointly exploiting the semantic and structural dimensions of biomedical concepts. The principal components are summarized below [9]:

1. Preprocessing: This module prepares raw ontology data for downstream processing. It parses OWL (Ontology Web Language) files, constructs RDF (Resource Description Framework) graphs, and extracts labels, definitions, and synonyms of ontology concepts. This stage produces a comprehensive lexical and relational representation of each ontology, forming the basis for subsequent encoding.

2. Semantic Encoder (Sentence-SapBERT): BioGITOM employs Sentence-SapBERT, a sentence-level biomedical encoder fine-tuned on UMLS concept definitions, to generate domain-aware semantic embeddings. These embeddings capture fine-grained lexical and contextual nuances in biomedical terminology, providing robust semantic representations of ontology entities.

3. Graph Isomorphism Transformer (GIT): At the core of BioGITOM lies the GIT model, which integrates GINs and GTs to encode both local and global structural dependencies within ontology graphs:
   - GINs model neighborhood connectivity patterns to preserve discriminative local structures.
   - GTs use multi-head attention to capture long-range relationships among distant nodes.

   This hybrid design produces expressive structural embeddings that represent both hierarchical and cross-graph dependencies among biomedical concepts.

4. Gated Combination Module: This module fuses the semantic embeddings from Sentence-SapBERT with the structural embeddings generated by the GIT model. Through a learnable gating mechanism [10], BioGITOM dynamically adjusts the relative contribution of semantic and structural signals, producing optimized embeddings that balance meaning and topology for each ontology.

5. FAISS-Based L2 Retrieval and Mapping Selection: To efficiently match large-scale biomedical ontologies, BioGITOM integrates a FAISS-based Approximate Nearest Neighbor (ANN) retrieval mechanism using the L2 distance metric. This approach enables fast and scalable similarity search in the joint embedding space, identifying the top-k candidate correspondences for each source

concept. The retrieved pairs are then ranked by similarity scores to produce the final set of mappings with associated confidence values.

1.3. Adaptations made for the evaluation

For this evaluation, BioGITOM was deployed in its standard configuration, without any task-specific adaptations or parameter tuning. This setup highlights the system's intrinsic versatility and robustness, as it achieved consistently high performance across benchmarks without additional customization.

1.4. Link to the system and parameters file

BioGITOM is a fully developed OM system, now publicly available for research and benchmarking purposes. The complete source code, pretrained models, and documentation can be accessed through the official GitHub repository: https://github.com/lias-laboratory/biogitom.

2. Results

The evaluation of BioGITOM on the OAEI 2025 Bio-ML track covers five benchmark datasets. The following subsections summarize its performance and comparative analysis.

2.1. Performance evaluation of BioGITOM using OMIM-ORDO dataset

As reported in Table 1, BioGITOM delivers the best overall results on the OMIM–ORDO dataset. It achieves a precision of 0.845 and a recall of 0.736, leading to an F1-score of 0.787. These figures demonstrate the model's capacity to generate accurate correspondences while maintaining a high coverage of relevant matches.

Although the OMIM–ORDO dataset has limited structural depth (maximum concept depth = 2), BioGITOM effectively leverages rich semantic representations to overcome this constraint and deliver highly reliable alignments.

Table 1
Results of BioGITOM on the OMIM-ORDO dataset.

Tool	P	R	F1
BioGITOM	0.845	0.736	0.787

2.2. Performance evaluation of BioGITOM using DOID-NCIT dataset

On the NCIT–DOID dataset (Table 2), BioGITOM achieves the highest overall performance, ranking first among all participating systems. It attains an F1-score of 0.918, with a precision of 0.924 and a recall of 0.911, reflecting a balanced trade-off between accuracy and completeness. While LogMapLt reaches the highest precision (0.976) and LogMapBio the highest recall (0.959), BioGITOM stands out for its stability across metrics and its ability to align structurally and semantically diverse ontologies with remarkable consistency.

Table 2
Results of BioGITOM on the DOID-NCIT dataset.

Tool	P	R	F1
BioGITOM	0.924	0.911	0.918

2.3. Performance evaluation of BioGITOM using SNOMED-FMA (Body) dataset

Performance on the SNOMED–FMA dataset (Table 3) highlights BioGITOM's strong recall capacity. Although BERTMap and BERTMapLt lead in precision (0.970), BioGITOM achieves a precision of 0.829 combined with the highest recall (0.748) among all systems. This combination produces an F1-score of 0.787, confirming the model's efficiency in capturing a large proportion of valid correspondences across structurally extensive biomedical ontologies.

Table 3
Results of BioGITOM on the SNOMED-FMA (Body) dataset.

Tool	P	R	F1
BioGITOM	0.829	0.748	0.787

2.4. Performance evaluation of BioGITOM using SNOMED-NCIT (Pharm) dataset

For the SNOMED–NCIT (Pharm) task (Table 4), BioGITOM displays a consistent balance between precision and recall. While LogMapLt reaches the highest precision (0.994), BioGITOM secures the best recall (0.779) and an F1-score of 0.786. These outcomes emphasize the model's adaptability in identifying a broad spectrum of correct mappings and its reliability in managing heterogeneous biomedical ontologies.

Table 4
Results of BioGITOM on the SNOMED-NCIT (Pharm) dataset.

Tool	P	R	F1
BioGITOM	0.793	0.779	0.786

2.5. Performance evaluation of BioGITOM using SNOMED-NCIT (Neoplas) dataset

As displayed in Table 5, BioGITOM achieves competitive results on the SNOMED–NCIT (Neoplas) dataset. With a recall of 0.719, it ranks among the strongest systems in terms of coverage, despite LogMapLt attaining the highest precision (0.931). The resulting F1-score of 0.745 illustrates BioGITOM's resilience in dealing with semantically intricate and hierarchically layered ontologies, reaffirming its reliability across different biomedical alignment scenarios.

Table 5
Results of BioGITOM on the SNOMED-NCIT (Neoplas) dataset.

Tool	P	R	F1
BioGITOM	0.793	0.774	0.745

3. General Comments

3.1. Comments on the Results (Strengths and Weaknesses)

The experimental results clearly highlight the competitive advantage and robustness of BioGITOM compared to other top-performing systems in the OAEI 2025 Bio-ML track. One of the principal strengths of the proposed approach lies in the Graph Isomorphism Transformer (GIT) architecture, which effectively integrates local neighborhood modeling through Graph Isomorphism Networks (GINs) with global dependency capture via Graph Transformers (GTs). This hybrid design enables the model to generate contextually enriched and structurally coherent representations, allowing BioGITOM to handle the inherent complexity, heterogeneity, and hierarchical diversity of biomedical ontologies.

Another major strength is the Sentence-SapBERT encoder, which leverages domain-specific language modeling on biomedical concept definitions to capture subtle lexical and semantic nuances. The synergy between semantic encoding and graph-based structural reasoning proved crucial for achieving high recall across multiple datasets. Furthermore, the incorporation of a FAISS-based L2 retrieval mechanism ensures scalability and efficiency, allowing BioGITOM to manage millions of biomedical entities without compromising accuracy.

However, a limitation of the current version lies in its focus on equivalence mappings. While this design ensures strong precision and balanced recall in equivalence-based tasks, it does not yet extend to other semantic relationships such as subsumption or part-of relations, which are critical in hierarchically rich biomedical ontologies. Addressing these relationships remains an open challenge for future work, as also discussed in our recent publication [9].

3.2. Discussion on Improvements for the Proposed System

Building upon these results, several enhancements are being explored to further increase the versatility and accuracy of BioGITOM. First, we aim to extend the system's matching scope beyond equivalence to encompass hierarchical (subsumption) and associative relationships, enabling richer and more semantically expressive alignments. This expansion will make BioGITOM suitable for broader ontology integration tasks, such as knowledge graph fusion and reasoning-based biomedical data linking.

Second, we are investigating the transfer of learned concept representations into a hyperbolic embedding space, motivated by the observation that Euclidean spaces inadequately preserve hierarchical geometries inherent to biomedical ontologies. Hyperbolic spaces, in contrast, provide a more natural geometric foundation for modeling tree-like and taxonomic structures, thereby reducing distortion and improving the representation of complex ontological hierarchies. As reported in our recent publication [9], such geometric transformations are expected to significantly enhance BioGITOM's capacity to capture nuanced structural dependencies and improve the accuracy of hierarchical ontology alignment.

4. Conclusion

BioGITOM represents a novel and advanced approach to biomedical ontology matching, built upon a hybrid Graph Isomorphism Transformer (GIT) architecture that integrates the strengths of Graph Isomorphism Networks (GINs) and Graph Transformers (GTs). Supported by Sentence-SapBERT for domain-specific semantic encoding and a FAISS-based retrieval mechanism, BioGITOM effectively combines semantic and structural information to produce precise and scalable ontology alignments.

Experimental evaluations on the OAEI 2025 Bio-ML benchmark demonstrate that BioGITOM consistently outperforms competing systems across multiple biomedical datasets, achieving superior overall performance and confirming its robustness and generalization capability. Nevertheless, the current version focuses primarily on generating equivalence mappings, which limits its ability to capture other semantic relations such as subsumption or part-of.

Future work, will focus on extending BioGITOM to support a broader spectrum of semantic relationships and exploring hyperbolic representation spaces to better model hierarchical ontological structures. These enhancements aim to further improve the system's versatility, interpretability, and accuracy, reinforcing its role as a comprehensive solution for large-scale biomedical ontology alignment.

Declaration on Generative AI

During the preparation of this work, the authors used Grammarly in order to grammar and spell check, and improve the text readability. After using the tool, the authors reviewed and edited the content as needed to take full responsibility for the publication's content.

References

[1] J. Euzenat, P. Shvaiko, Ontology Matching, 2nd ed., Springer-Verlag, 2013. doi:10.1007/978-3-642-38721-0.

[2] S. Oulefki, L. Berkani, N. Boudjenah, I. E. Kenai, A. Mokhtari, Ontology Matching Using Multi-head Attention Graph Isomorphism Network, in: M. Mosbah, T. Kechadi, L. Bellatreche, F. Gargouri (Eds.), Model and Data Engineering, Springer Nature Switzerland, Cham, 2024, pp. 200–213. doi:10.1007/978-3-031-49333-1_15.

[3] S. Oulefki, L. Berkani, A GNN-Based Framework for Ontology Matching, in: Model and Data Engineering: 13th International Conference, MEDI 2024, Naples, Italy, November 18–20, 2024, Proceedings, Springer-Verlag, Berlin, Heidelberg, 2025, p. 183–197. doi:10.1007/978-3-031-87719-3_14.

[4] S. Oulefki, L. Berkani, N. Boudjenah, L. Bellatreche, A. Mokhtari, Concepts and relations features are all you need for embedding-based ontology matching, in: Web Information Systems Engineering – WISE 2024: 25th International Conference, Doha, Qatar, December 2–5, 2024, Proceedings, Part I, Springer-Verlag, Berlin, Heidelberg, 2024, p. 416–430. doi:10.1007/978-981-96-0579-8_29.

[5] S. Oulefki, L. Berkani, L. Bellatreche, N. Boudjenah, A. Mokhtari, Results for BioGITOM in OAEI 2024, in: Proceedings of the 19th International Workshop on Ontology Matching co-located with the 23rd International Semantic Web Conference (ISWC 2024), Baltimore, USA, volume 3897, 2024, pp. 104–109. URL: https://ceur-ws.org/Vol-3897/oaei2024_paper2.pdf.

[6] K. Xu, W. Hu, J. Leskovec, S. Jegelka, How Powerful are Graph Neural Networks?, 2019. URL: https://arxiv.org/abs/1810.00826. arXiv:1810.00826.

[7] E. Min, R. Chen, Y. Bian, T. Xu, K. Zhao, W. Huang, P. Zhao, J. Huang, S. Ananiadou, Y. Rong, Transformer for Graphs: An Overview from Architecture Perspective, 2022. URL: https://arxiv.org/abs/2202.08455. arXiv:2202.08455.

[8] J. Johnson, M. Douze, H. Jégou, Billion-Scale Similarity Search with GPUs, IEEE Transactions on Big Data 7 (2021) 535–547. doi:10.1109/TBDATA.2019.2921572.

[9] S. Oulefki, L. Berkani, N. Boudjenah, L. Bellatreche, A. Mokhtari, BioGITOM: Matching Biomedical

Ontologies with Graph Isomorphism Transformer, The VLDB Journal 34 (2025). doi:10.1007/s00778-025-00943-7.

[10] Y. Gu, X. Qu, Z. Wang, Y. Zheng, B. Huai, N. J. Yuan, Delving Deep into Regularity: A Simple but Effective Method for Chinese Named Entity Recognition, in: M. Carpuat, M.-C. de Marneffe, I. V. Meza Ruiz (Eds.), Findings of the Association for Computational Linguistics: NAACL 2022, Association for Computational Linguistics, Seattle, United States, 2022, pp. 1863–1873. doi:10.18653/v1/2022.findings-naacl.143.