ALIN Results for OAEI 2025

Jomar da Silva[1,*], Kate Revoredo[2], Fernanda Araujo Baião[3] and Cabral Lima[1]

[1]*Graduate Program in Informatics*
Federal University of Rio de Janeiro (UFRJ), Brazil
[2]*Humboldt-Universität zu Berlin, Berlin, Germany*
[3]*Department of Industrial Engineering*
Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Brazil

Abstract

Alin is a system for interactive ontology matching that has been participating in all OAEI editions since 2016. In this new version, we improved the non-interactive version of the tool. We run it like the interactive version but do not use an expert in the interaction; we use ChatGPT in its place.

Keywords

ontology matching, Wordnet, interactive ontology matching, ontology alignment, interactive ontology alignment, lexical analyzer, ChatGPT

1. Presentation of the system

Due to the advances in Information and Communication Technologies (ICT) in general, a large amount of data repositories became available as valuable assets for enabling integrated data exchange platforms across organizations. However, those repositories are highly semantically heterogeneous, which hinders their integration. Ontology Matching has been successfully applied to solve this problem, by discovering mappings between two distinct ontologies which, in turn, conceptually define the data stored in each repository. The Ontology Matching process seeks to discover correspondences (mappings) between entities of different ontologies, and this may be performed manually, semi-automatically or automatically [1]. The interactive approach, which considers the knowledge of domain experts through their participation during the matching process, has stood out among semi-automatic ones [2]. A domain expert is an expensive, scarce, and time-consuming resource; when available, however, this resource has improved the achieved results. Nevertheless, there is still room for improvements [2], as evidenced by the most recent results from the evaluation of interactive tools in the OAEI[1] (Ontology Alignment Evaluation Initiative). ALIN[3][4][5][6] is a system for interactive ontology matching which has been participating in all OAEI editions since 2016.

In an interactive process, besides the F-Measure, which assesses the quality of the generated alignment, the number of interactions with the expert is also important—the fewer the inter-

OM-2025: The 20th International Workshop on Ontology Matching ccollocated with the 24th International Semantic Web Conference (ISWC-2025), November 2nd, 2025, Nara, Japan
*Corresponding author.
jomar.silva@ufrj.br (J. d. Silva); kate.revoredo@hu-berlin.de (K. Revoredo); fbaiao@puc-rio.br (F. A. Baião); cabrallima@ufrj.br (C. Lima)

[1]Available at https://oaei.ontologymatching.org/2025/results/interactive/index.html, last accessed on Nov, 14, 2025.

actions, the better. In the interactive process, a key step is selecting mappings for the expert. An improvement in the alignment process would occur if we could find a way to reduce the number of mappings in this set without lowering the F-Measure, or at least reduce the number of mappings at a greater rate than the decrease in the F-Measure.

ALIN participates not only in the interactive track but also in the anatomy track and the conference track, both non-interactive. Until last year, this participation occurred as follows: before selecting mappings to be presented to the expert, ALIN would select some that went directly into the final alignment. Participation in the non-interactive tracks consisted solely of these automatically selected mappings. This year, we changed this. In addition to the automatically selected mappings, ALIN submitted the mappings chosen to be shown to the expert to ChatGPT. Therefore, ALIN' final alignment is now composed of both the automatic mappings and those approved by ChatGPT.

1.1. State, Purpose and General statement

During its matching process, ALIN handles three sets of mappings: (i) Accepted, which is a set of mappings definitely to be retained in the alignment; (ii) Selected, which is a set of mappings where each is yet to be decided if it will be included in the alignment; and (iii) Suspended, which is a set of mappings that have been previously selected, but (temporarily or permanently) filtered out of the selected mappings.

Given the previous definitions, ALIN procedure follows 5 Steps, described as follows:

1. Select mappings: select the first mappings and automatically accepts some of them. Detailed in the 'Specific techniques used' subsection below;
2. Filter mappings: suspend some selected mappings, using lexical and semantic criteria for that.
3. Ask domain expert: accepts or rejects selected mappings, according to domain expert feedback;
4. Propagate: select new mappings, reject some selected mappings or unsuspend some suspended mappings (depending on newly accepted mappings);
5. Go to step 3 as long as there are undecided selected mappings.

All versions of ALIN (since its first OAEI participation) follow this general procedure.

In this year's version, in item 3, for the non-interactive tracks, the participation of the domain expert was replaced by consulting ChatGPT.

1.2. Specific techniques used

- Step 1. ALIN employs a blocking strategy where it does not consider data and object properties from the ontologies at this step. It selects only concept mappings based on linguistic similarities between previously standardized concept names. ALIN automatically accepts mappings with standardized names that are synonyms, using WordNet and domain-specific ontologies, such as the FMA Ontology in the Anatomy track, to identify these synonyms.

- Step 2. Alin suspends some selected mappings that exhibit low lexical and semantic similarity in their entity names, removing them from the set of selected mappings. We use the Jaccard, Jaro-Winkler, and n-gram lexical metrics to calculate the lexical similarity of the selected mappings. We also used a semantic metric called the Alin metric. These suspended mappings can be further unsuspended later, returning to the set of selected mappings, as proposed in [5]. We employ a threshold for suspension, where we suspend a mapping if all its similarity values are below this threshold. We used a threshold of 0.9 for the Conference track. We adjusted the threshold to 0.96 for the Anatomy track.
- Step 3. At this point, the domain expert interaction begins. Alin sorts the selected mappings in a descending order according to the sum of similarity metric values. The sorted selected mappings are submitted to the domain expert. Alin can present up to three mappings together to the domain expert if a full entity name in a candidate mapping is the same as another entity name in another candidate mapping. In the non-interactive tracks, this year, the expert's participation was replaced by consulting ChatGPT.
- Step 4. Initially, the set of selected mappings contains only concept mappings. At each interaction with the domain expert, if he accepts the mapping, Alin (i) removes from the set of selected mappings all the mappings that compose an instantiation of a mapping anti-pattern [7][8] (we explain mapping anti-patterns below in the 'Mapping anti-patterns' paragraph) with the accepted mappings; (ii) selects data property (as proposed in [6]) and object property mappings related to the accepted concept mappings; (iii) unsuspends all concept mappings whose both entities are subconcepts of the concept of an accepted mapping (as proposed in [5]).
- Step 5. Go to step 3 until there are no selected mappings.

1.2.1. Mapping anti-patterns

An anti-pattern mapping can be a logical inconsistency, a construction constraint on the ontology, or an alignment constraint. An ontology may have construction constraints, such as a concept cannot be equivalent to its superconcept. The alignment between two ontologies can have constraints. For example, an entity of ontology O cannot be equivalent to two entities of the ontology O'. Anti-pattern mapping is a combination of mappings that generates a problematic alignment, i.e., a logical inconsistency or a violated constraint.

1.3. Modifications made in the 2025 version of Alin

Alin participates not only in the interactive track but also in the anatomy track and the conference track, both non-interactive. Until last year, this participation occurred as follows: before selecting mappings to be presented to the expert, Alin would select some that went directly into the final alignment. Participation in the non-interactive tracks consisted solely of these automatically selected mappings. This year, we changed this. In addition to the automatically selected mappings, Alin submitted the mappings chosen to be shown to the expert to ChatGPT. Therefore, Alin's final alignment is now composed of both the automatic mappings and those approved by ChatGPT.

1.4. Link to the system and parameters file

Alin is available [2] as a SEALS package (It can be run with MELT).

2. Results

We will compare Alin's participation this year and last year in the conference track and the anatomy track, both non-interactive. We will not show the results from the interactive track, as the change this year only produced results in the non-interactive tracks.

Table 1
Anatomy Track - OAEI 2024[9]

Matcher	Runtime	Precision	F-measure	Recall
Matcha	42	0.951	0.941	0.931
MDMapper	121	0.926	0.903	0.881
LogMapBio	1346	0.888	0.898	0.908
LogMap	12	0.917	0.881	0.848
ALIN	370	0.984	0.851	0.75
LogMapLite	2	0.962	0.828	0.728
StringEquiv	-	0.997	0.766	0.622
TOMATO	2154	0.955	0.523	0.36

Table 2
Anatomy Track - OAEI 2025[10]

Matcher	Runtime	Precision	F-measure	Recall
Matcha	47	0.951	0.941	0.931
Agent-OM	-	0.959	0.92	0.883
ALIN	1004	0.942	0.912	0.884
LogMapLLM	500	0.964	0.899	0.842
LogMapBio	1750	0.885	0.898	0.911
MDMapper	124	0.899	0.889	0.879
LogMap	8	0.917	0.881	0.848
LogMapKG	9	0.917	0.881	0.848
LogMapLite	2	0.962	0.828	0.728
StringEquiv	-	0.997	0.766	0.622
DRAL-OA	877	0.83	0.828	0.827
LSMatch	16	0.952	0.761	0.634

2.1. Comments on the participation of Alin in OAEI 2025

In Alin's participation in the Anatomy track (Tables 1 and 2), there was a substantial gain in recall and a much less significant loss in precision. As a result, there was a significant gain in

[2] https://osf.io/pu7fv/files/vtczf

Table 3
Conference Track (rar2-M3) - OAEI 2024[11]

Matcher	Precision	F1-measure	Recall
LogMap	0.76	0.64	0.56
MATCHA	0.66	0.64	0.63
MDMapper	0.66	0.59	0.53
ALIN	0.82	0.57	0.44
OntoMatch	0.82	0.56	0.43
edna	0.74	0.56	0.45
LogMapLt	0.68	0.56	0.47
StringEquiv	0.76	0.53	0.41
TOMATO	0.57	0.48	0.42

Table 4
Conference Track (rar2-M3) - OAEI 2025[12]

Matcher	Precision	F1-measure	Recall
ALIN	0.62	0.65	0.68
LogMap	0.76	0.64	0.56
Matcha	0.77	0.63	0.53
Agent-OM	0.64	0.61	0.59
MDMapper	0.69	0.58	0.5
edna	0.74	0.56	0.45
LogMapLt	0.68	0.56	0.47
LSMatch	0.83	0.55	0.41
StringEquiv	0.76	0.53	0.41

Table 5
Anatomy - OAEI 2025[10][13] - Comparison between interactive mode and non-interactive mode

Mode	Error Rate	Precision	F-measure	Recall
non-interactive	-	0.942	0.912	0.884
interactive	0.0	0.986	0.932	0.883
interactive	0.1	0.953	0.909	0.868
interactive	0.2	0.924	0.887	0.852
interactive	0.3	0.895	0.865	0.837

F-measure. Last year, ALIN ranked fifth in F-measure among eight tools, and this year it ranked third among twelve. But this result came at the cost of a much longer runtime.

In ALIN's participation in the Conference track (Tables 3 and 4), there was also a significant gain in recall but a substantial loss in precision as well—though less than the gain in recall—which resulted in an increase in F1-measure. Last year, ALIN ranked fourth in F1-measure among nine tools, and this year it ranked first among the same nine.

When comparing the interactive execution of ALIN with its non-interactive execution (Tables 5 and 6), we see that the interactive execution yields better results than the non-interactive one.

Table 6
Conference - OAEI 2025[12][13] - Comparison between interactive mode and non-interactive mode

Mode	Error Rate	Precision	F-measure	Recall
non-interactive	-	0.62	0.65	0.68
interactive	0.0	0.903	0.79	0.703
interactive	0.1	0.744	0.708	0.676
interactive	0.2	0.631	0.635	0.64
interactive	0.3	0.537	0.574	0.618

This occurs mainly when the expert has a zero error rate. When the error rate reaches 10% in the Anatomy track, the non-interactive execution performs better. In the Conference track, this happens when the error rate reaches 20%.

3. General comments

This new version of Alin uses ChatGPT to simulate the specialist's participation in non-interactive tracks. Its use resulted in a gain in recall, but a loss in precision—smaller than the gain in recall—which led to an overall improvement in F-measure in both the Anatomy track and the Conference track. In both tracks, Alin improved its ranking in the F-measure.

Declaration on Generative AI

During the preparation of this manuscript, the authors used Grok (xAI) and ChatGPT (OpenAI) to assist with text translation, grammar and spelling correction, and improvement of readability. The authors reviewed and edited all outputs and take full responsibility for the content of this publication.

References

[1] J. Euzenat, P. Shvaiko, Ontology Matching - Second Edition, Springer-Verlag, 2013.

[2] H. Li, Z. Dragisic, D. Faria, V. Ivanova, E. Jimenez-Ruiz, P. Lambrix, C. Pesquita, User validation in ontology alignment: functional assessment and impact, The Knowledge Engineering Review (2019). doi:10.1017/S0269888919000080.

[3] J. Silva, K. Revoredo, F. A. Baião, C. Lima, Enhancing ontology matching: Lexically and syntactically standardizing ontologies through customized lexical analyzers, Semantic Web: – Interoperability, Usability, Applicability 16 (2025). URL: https://doi.org/10.1177/22104968251344447. doi:10.1177/22104968251344447.

[4] J. Da Silva, K. Revoredo, F. Baião, J. Euzenat, Alin: improving interactive ontology matching by interactively revising mapping suggestions, The Knowledge Engineering Review 35 (2020). doi:10.1017/S0269888919000249.

[5] J. Silva, F. Baião, K. Revoredo, J. Euzenat, Semantic interactive ontology matching: Synergistic combination of techniques to improve the set of candidate correspondences, in:

OM-2017: Proceedings of the Twelfth International Workshop on Ontology Matching, volume 2032, 2017, pp. 13–24.
[6] J. Silva, K. Revoredo, F. A. Baião, J. Euzenat, Interactive Ontology Matching: Using Expert Feedback to Select Attribute Mappings, in: CEUR Workshop Proceedings, volume 2288, 2018, pp. 25–36.
[7] A. Guedes, F. Baião, R. Shivaprabhu, Revoredo, On the Identification and Representation of Ontology Correspondence Antipatterns, in: Proc. 5th Int. Conf. Ontol. Semant. Web Patterns (WOP'14), CEUR Work. Proc., 2014.
[8] A. Guedes, F. Baião, K. Revoredo, Digging Ontology Correspondence Antipatterns, in: Proceeding WOP'14 Proc. 5th Int. Conf. Ontol. Semant. Web Patterns, volume 1032, 2014, pp. 38––48.
[9] Results for oaei 2024 - anatomy track, 2024. URL: https://oaei.ontologymatching.org/2024/results/anatomy/index.html, accessed: 2025-11-14.
[10] Results for oaei 2025 - anatomy track, 2025. URL: https://oaei.ontologymatching.org/2025/results/anatomy/index.html, accessed: 2025-11-14.
[11] Results for oaei 2024 - conference track, 2024. URL: https://oaei.ontologymatching.org/2024/results/conference/index.html, accessed: 2025-11-14.
[12] Results for oaei 2025 - conference track, 2025. URL: https://oaei.ontologymatching.org/2025/results/conference/index.html, accessed: 2025-11-14.
[13] Results for oaei 2025 - interactive track, 2025. URL: https://oaei.ontologymatching.org/2025/results/interactive/index.html, accessed: 2025-11-14.