Ontology Integration by Semantic Mapping for Solving the Heterogeneity Problem

Moseed Mohammed, Awanis Romli and Rozlina Mohamed

July 16, 2022

Ontology Integration by Semantic Mapping for Solving the Heterogeneity Problem

Moseed Mohammed[1], Awanis Romli[1], Rozlina Mohamed[1]

[1]Faculty of Computing, Universiti Malaysia Pahang, Pahang, Malaysia
qutamee@yahoo.com

Abstract. In recent years, ontology integration has received an increased focus in ontology engineering. Ontology integration is a complex process that has some difficulties such as semantic heterogeneity. The goal of this research is to use semantic mapping to reduce integration complexity and solve semantic heterogeneity. What is ontology engineering? What difficulties haven't been solved until now by ontology integration? What is the effective role of semantic mapping in semantic heterogeneity? This research seeks to address these questions. The expected contribution of this research is to build a comprehensive view of ontology integration and support interoperability. The significance of using semantic mapping to improve interoperability on ontology integration is confirmed by researchers.

Keywords: Ontology engineering, ontology integration, semantic mapping, interoperability

1 Introduction

Ontology is a formal specification of conceptualizations and formal explanation of knowledge [1]. Ontology is created in a branch of artificial intelligence for knowledge-based systems and established to retrieve information problems [2]. Ontology is generally used in several areas such as semantic web [3], engineering systems [4], software engineering [5], healthcare information [6], IoT technology [8], library system [9], knowledge organisation [10], decision-making method [11], and manufacturing systems [12], as ontology decreases the difficulty of information and increases its association [13] as well as eases information sharing. Ontology is used to solve the interoperability problems of multiple domains [14] and create a knowledge-based system [15]. The significance of using semantic mapping to improve interoperability in different areas is confirmed by researchers [16 - 18].
Ontology integration is a procedure to integrate two or more ontologies to build a new integrated ontology [17]. Most present ontology integration methods are restricted for matching between two ontologies [18], and only a few methods manage more than two ontologies simultaneously [19]. There are two basic stages for ontology integration which are the matching stage and merging stage. Ontology integration has been studied over the past two decades, but it remains a stimulating job, where the applica-

tions of ontology integration have been greatly benefited from in the biomedical area [20] and the Internet of Things [21]. This paper is focused on heterogeneity problems in ontology integration. There are two types of heterogeneity in ontology integration, which are schema heterogeneity [22] and semantic heterogeneity [23]; however, the researchers have not focused on semantic heterogeneity [24]. Ontology matching is a real method to address the problem of ontology heterogeneity [25]. Ontology matching is the greatest solution to the heterogeneity problem because it detects matches between semantically related entities in ontologies [20]. Most existential matching solutions depend on schema-level much more than data-level [26]. The goal of this research is to use semantic mapping to reduce integration complexity and solve the heterogeneity. Semantic mapping between concepts is very significant for integration [27], but it is the largest share of unresolved problems and not used much due to their need for a complex process [20]. Syntactic measures are the most similarity used because it is easy for implementation [20]; structural measures are also used while semantic measures are not used much due to their want for difficult operations [28]. This paper is organised as follows. Section 2 defines the study methodology. Section 3 describes the ontology engineering background, explains ontology and the ontology development process. Section 4 presents the concepts used in the integration of ontologies, which are the matching and merging of ontology. Section 5 describes the different existing tools of ontology integration. Section 6 draws the conclusion of this paper.

2 Study Methodology

The guideline that was used to perform the review in this paper was to search for proceedings from conferences and journal papers in Google Scholar, Scopus, and Web of Science. The articles focused on the background of ontology engineering, ontology integration, and semantic mapping. The selected articles were deemed eligible based on their appropriate studies to provide answers to the research questions presented in this research, which are: What is ontology engineering? What difficulties have not been solved until now by ontology integration? What is the effective role of semantic mapping in semantic heterogeneity?

3 Ontology Engineering

Ontology is a set of axioms that explains and describes domain entities [26]. Ontology is a 5-tuple $O = (C, P, I, \Lambda, \Gamma)$ [20], where C is a set of classes, P is a set of properties, I is a set of individuals, Λ is a set of axioms, and Γ is a set of annotations. Table 1 describes in detail the components of ontology. Ontology engineering is a branch of knowledge engineering that studies ontology building methods and methodologies [29]. Ontology engineering studies the ontology development process [30], ontology life cycle, ontology construction methods [31], ontology integration [27], and languages that support them. Ontology integration is a significant subject of interest in ontology engineering, as referred to in the next section. Ontology language is a formal language for coding ontology and the user is able to inscribe strong formal representations of domains. There are several languages for ontology, such as Resource Description Framework (RDF) [32], RDF Schema (RDFS) [33], and Ontology Web Language OWL [34].

Table 1. The ontology components

Item	Description
Classes	Set of objects that are grouped according to common features.
Properties	Set of features or characteristics of the object.
Individuals	Set of instances of classes in the real world which are also called terms.
Relations	Set of relationships that provides logical connections between individuals or classes that describe the relation between them.
Axioms	Set of axioms used for checking the consistency of ontology or inferencing new information based on rules in a logical form.
Annotations	Set of annotations that provides metadata for information to be understood.
Function	Set of structures molded by definite relationships that may replace individual terms with extra complex terms.
Restrictions	Set of official declarations that describe what must be true for some declarations to be measured true.
Rules	Set of sentences (if-then statements) which defines inferences that are extracted by confirmation.

Table 1 describes the components of ontology which is a set of objects that has static and dynamic parts. The static part of ontology concerns the structure that is modelled within a particular field such as classes and properties, and the dynamic part revolves around reasoning, inferences, and deriving new facts from already known facts such as axioms and rules.

4 Ontology Integration

Ontology integration is a critical task in ontology engineering. Ontology integration is the procedure to merge two or more ontologies with the goal of building a new integrated ontology [27]. There are many terms regarding ontology integration such as matching, merging, mapping, and relationship that are unclear and at times unused. So, Table 2 provides a description for each term. Ontology integration includes three different cases [27]: (1) Develop a new ontology by reusing ontologies; (2) Create a new unified ontology by integrating different ontologies; and (3) Integrate various ontologies into a single application to describe or apply a knowledge-based system.

Table 2. Ontology integration terms

Terms	Description
Matching	Determining the semantic matches of entities in different ontologies, which is an active way to address the problem of ontological heterogeneity.

Merging	Building complete ontology by integrating knowledge from other ontologies.
Mapping	Mapping an equivalence correspondence which named mapping rules when they are read as ontological declarations or axioms.
Relation	Giving a correspondence for integral relation such as the equivalence, subsumption, and disjointness.

Ontology integration approaches contain two basic stages [11]: First, a matching stage that resolves differences by recognising semantic similarity between the different elements. Second, the merging stage that achieves the outcome of the matching stage by merging or linking matching elements to create a new united vision. Ontology matching approaches are simple matching [35] and complex matching [36]. Ontology merging approaches are simple merge [26], full merge [18], and symmetric merge [37]. Ontology integration has been widely and effectively applied in biomedical [23] and the Internet of Things, while there is a great lack in manufacturing [18].

4.1 Ontology Matching

Ontology matching is the method of identifying the semantic correspondences of entities in different ontologies. Similarity measure is critical for matching ontology methods [24]. There are three categories of similarity measures as shown in Table 2, which are syntactic measure, structure measure, and linguistic measure. These will be presented in detail in the next section.

Table 3 Describes similarity measures categories

Author	Measure 1	Measure 2	Measure 3
[36]	Terminological Mapping	Structural Mapping	Semantic Mapping
[20]	Syntactic measure	Taxonomy measure	Linguistic measure
[38]	Statistics Techniques	Logic Techniques	Linguistics Techniques
[39]	Terminological Techniques	Structural Techniques	Semantic Techniques
[40]	Syntactic Similarity	Structural Similarity	Linguistic measure
[41]	Syntactic techniques	Lexical techniques	Semantic techniques
[42]	Syntactic measure	Structural measure	Linguistic Semantic

4.1.1 Syntactic-based measures:

There are two syntactic measures that are mostly used which are String Metric for Ontology Alignment (SMOA) [43] and Levenshtein [20]. Assumed two strings x1 and x2, the SMOA similarity is defined as follows:

$$SMOA(x1, x2) = comm(x1, x2) - diff(x1, x2) + winklerImpr(x1, x2) \quad (1)$$

where comm(x1, x2) stands for the common length of x1 and x2, while diff(x1, x2) for the different lengths and winklerImpr(x1, x2) is the improved approach proposed in [43].

4.1.2 Linguistic-based measures

Linguistic similarity between two strings is determined by considering semantic relationships (such as synonyms and hypernym) that typically require the use of thesaurus and dictionaries. WordNet is widely used as an electronic vocabulary database that collects all meanings of different words [24]. For example, two words d1 and d2, Linguistic Similarity (d1, d2) equals:

1 If words d1 and d2 are synonyms in Wordnet.

0.5, if word d1 is the hypernym of word d2 or the opposite is true in Wordnet.

0, otherwise.

4.1.3 Structure-based Measures

Structure-based measures are to make full use of the ontology hierarchy relation to determine the similarity between two entities by considering the similarity of their neighbours (parents, children, and siblings) [44] or have similar instances [42]. For example, if entities e1 in Q1 and e2 in Q2 are properly matched, then the neighbours of e1 are probable match neighbours of e2. When the correspondences linking the neighbours of e1 and e2 have a self-assurance rate, the correspondence (e1 ≡ e2) may be correct. Semantic mapping between concepts is very significant for integration [27]. Syntactic measures are the most similarity used because it is easy for implementation. Structural measures are also used while semantic measures are not used due to their want for complex operations.

4.2 Semantic Mapping

Semantic mapping of a particular correspondence can be a relationship [26], like equivalence relationship (≡), subsumption relationship (⊒ or ⊑), disjointness relationship (⊥), and overlap relationship (≬). Relationships are identified by the next signs: ‘‘=’’ (is equivalent to), ‘‘>’’ (includes or is more general than), ‘‘<’’ (is included by or is more specific than), and ‘‘%’’ (disjointness with).

4.2.1 Equivalence Relationship

The equivalence relationship among two classes C and D indicates that all cases of C are also cases of D, which means that together, the classes have a similar set of entities. The equality relationship that holds between two properties P1 and P2 means that an individual x is linked to an individual or literal data together by P1 and P2. Equivalence relationship between two entities z and w means that entity z is same/equivalent/duplicate to entity w.

4.2.2 Subsumption Relationship

An implicit relationship between classes C and D means that the set of cases of C is a subgroup/super group of the set of cases of D. Subsumption relationship land among two properties P1 and P2 means that if an entity z is linked by P1 to an entity or a data accurate w, then z is linked by P2 to w.

4.2.3 Disjointness Relationship

A disjointness relationship between two classes C and D means that cases of C are absolutely not cases of D. A dissociation relationship between two properties P1 and P2 means that no entity z is linked to a single individual or literal data by P1 and P2.

Equivalence and disjointness are the simplest types of relations, then comes the subsumption relations [45]. Equivalence and subsumption are the simplest relationships, followed by disjointness relationship [46]. Integration approaches must deal with a variety of semantic relationships.

4.3 Ontology Merging

The merging phase is the process of merging the nominated input ontologies into an integrated ontology. The goal of merging is to build a more comprehensive ontology on a topic, and to gather knowledge in a coherent way from other ontologies on the same topic [27]. There are three kinds of ontology merging which are simple merge that is bridge ontology, full merge that is semantically equal, and symmetric merge that is really ontology enhancement. Ontology merging facilitates creating an ontology, support assistance, and growth semantic interoperability. The main violations in ontology merging are [46] incoherence, inconsistency, and redundancy (structural and relational). Ontology incoherence means that there are unsatisfying classes and properties in merging ontology, which reduces its performance and makes it unclear and unusable. An inconsistency in integrated ontology occurs as a result of unintended repercussions of logical inferences that are still hard to discover, understand, clarify, and fix in advance. Structural redundancy or semantic redundancy happens in class hierarchy, where more than one path exists from the root to the leaf. Relational redundancy occurs due to the complete merge of entities or by the adding of equality relationships that connect diverse entities in merging ontology.

5 Ontology Integrating Tools

Several tools have been developed to integrate ontology, particularly for the matching process, such as Graph Theory Model (GTM) [47], Context-Based Measure (CBM) [48], Artificial Neural Networks (ANN) [28], and Protégé [49], as shown in Table 1.

Table 3 Ontology Integrating Tools

Tools	Description
GTM	Graph Theory Model is a division of separate mathematics which are education graph models and their characteristics. Graphs are mathematical network like models collected of two sets, V (set of apices/nodes) and E (set of edges/arcs).
CBM	Context-Based Measure is to match big rule ontologies, where the measurement of lexical similarity in ontology matching is performed using WordNet.
ANN	Artificial neural networks are computational systems stimulated by the human brain. It has proven its suitability for ontology matching.
Protégé	Protégé is a tool used for matching ontologies to get similar classes, objects, and instances.

6 Conclusion

This paper aims to review ontology integration and some related features that belong to the field of ontology matching. The paper reviewed literature on ideas, methods, several subjects, and future work in the ontology integration field. Most present ontology integration methods are restricted for matching between two ontologies, as only a few methods can manage more than two ontologies simultaneously. The greatest research work in the field of ontology matching remains concentrated on identifying simple equality correspondences among ontological entities which are the easy cases of ontological matching. Limited systems attempt to discover additional difficult correspondences or account for unequal relationships, like subsumption and disjointness. This study is expected to contribute to building a comprehensive view of ontology integration and interoperability support in many areas.

Acknowledgments

The research reported in this study is conducted by the researchers at University Malaysia Pahang (UMP), it is funded by FRGS/1/2018/TK10/UMP/02/3 grant. The researchers would like to thank Ministry of Higher Education and UMP for supporting this research.

References

1. G. Ren, R. Ding, and H. Li, “Building an ontological knowledgebase for bridge maintenance,” Adv. Eng. Softw., vol. 130, pp. 24–40, Apr. 2019.
2. X. Huang, C. Zanni-Merk, and B. Crémilleux, “Enhancing Deep Learning with Semantics: an application to manufacturing time series analysis,” Procedia Comput. Sci., vol. 159, no. 2018, pp. 437–446, 2019.
3. J. Zhang, H. Li, Y. Zhao, and G. Ren, “An ontology-based approach supporting holistic structural design with the consideration of safety, environmental impact and cost,” Adv. Eng. Softw., vol. 115, pp. 26–39, 2018.
4. Z. Shang, M. Wang, and D. Su, “Ontology based social life cycle assessment for product development,” Adv. Mech. Eng., vol. 10, no. 11, pp. 1–17, 2018.
5. M. H. Karray, F. Ameri, M. Hodkiewicz, and T. Louge, “ROMAIN: Towards a BFO compliant reference ontology for industrial maintenance,” Appl. Ontol., vol. 14, no. 2, pp. 155–177, 2019.
6. J. N. Otte et al., “An ontological approach to representing the product life cycle,” Appl. Ontol., pp. 1–19, Apr. 2019.
7. T. Slimani, “A Study on Ontologies and their Classification,” Recent Adv. Electr. Eng. Educ. Technol., pp. 86–92, 2004.
8. M. Mohammed, A. Romli, and R. Mohamed, “Existing Semantic Ontology and its Challenges for Enhancing Interoperability in IoT Environment,” Proc. - 2021 Int. Conf. Softw. Eng. Comput. Syst. 4th Int. Conf. Comput. Sci. Inf. Manag. ICSECS-ICOCSIM 2021, pp. 22–26, 2021.
9. J. Hobbs and T. Fenn, “The Design of Socially Sustainable Ontologies,” Philos. Technol., Jan. 2019.

10. M. Mohd, M. Bilo, T. Louge, R. Rai, and M. Hedi, “Computers in Industry Ontology-based approach to extract product ’ s design features from online customers ’ reviews,” Comput. Ind., vol. 116, p. 103175, 2020.
11. H. Cheng, P. Zeng, L. Xue, Z. Shi, P. Wang, and H. Yu, “Manufacturing ontology development based on industry 4.0 demonstration production line,” in Proceedings - 2016 3rd International Conference on Trustworthy Systems and Their Applications, TSA 2016, 2016.
12. Y. He, C. Hao, Y. Wang, Y. Li, Y. Wang, and L. Huang, “An ontology-based method of knowledge modelling for remanufacturing process planning,” J. Clean. Prod., vol. 258, p. 120952, 2020.
13. H. Ostad-Ahmad-Ghorabi, T. Rahmani, and D. Gerhard, “An ontological approach for the integration of life cycle assessment into product data management systems,” CIRP Des. 2012 - Sustain. Prod. Dev., pp. 249–256, 2013.
14. A. N. M Mohammed, A Romli, R Mohamed, “Eco-Ontology for supporting Interoperability in Product Life Cycle within Product Sustainability Eco-Ontology for suapporting Interoperability in Product Life Cycle within Product Sustainability,” IOP Conf. Ser. Mater. Sci. Eng., 2021.
15. M. Mohammed, A. Romli, and R. Mohamed, “Using Ontology to Enhance Decision-Making for Product Sustainability in Smart Manufacturing,” IEEE Xplore ·, pp. 1–4, 2021.
16. F. M. Okikiola, A. M. Ikotun, A. P. Adelokun, and P. E. Ishola, “A Systematic Review of Health Care Ontology,” Asian J. Res. Comput. Sci., vol. 5, no. 1, pp. 15–28, 2020.
17. R. Salman, “Literature Review to Compare Efficiency of Various Machine Learning Algorithms in Predicting Chronic Kidney Disease (CKD),” pp. 1–4.
18. F. Ocker, B. Vogel-Heuser, and C. J. J. Paredis, “A framework for merging ontologies in the context of smart factories,” Comput. Ind., vol. 135, p. 103571, 2022.
19. B. Babalou, S., & König-Ries, “Towards Building Knowledge by Merging Multiple Ontologies with C o M erger : A,” arXiv Prepr. arXiv, 2020.
20. X. Xue, C. Yang, C. Jiang, P. Tsai, G. Mao, and H. Zhu, “Optimizing Ontology Alignment through Linkage Learning on Entity Correspondences,” Complexity, vol. 2021, no. 1, 2021.
21. M. De Roode, A. Fernández-izquierdo, L. Daniele, and M. Poveda-villalón, “SAREF4INMA : a SAREF extension for the Industry and Manufacturing domain,” Semant. Web, vol. 0, no. 0, pp. 1–16, 2020.
22. L. Li, “Technological Forecasting & Social Change China ’ s manufacturing locus in 2025 : With a comparison of “ Made-in-China,” Technol. Forecast. Soc. Chang., vol. 135, no. August 2017, pp. 66–74, 2018.
23. Xue Xingsi, “An Automatic Biomedical Ontology Meta-matching Technique,” J. Netw. Intell., vol. 4, no. 3, pp. 109–113, 2019.
24. H. Zhu, X. Xue, C. Jiang, and H. Ren, “Multiobjective Sensor Ontology Matching Technique with User Preference Metrics,” Wirel. Commun. Mob. Comput., vol. 2021, 2021.
25. X. Xue, H. Wang, J. Zhang, Y. Huang, M. Li, and H. Zhu, “Matching Transportation Ontologies with Word2Vec and Alignment Extraction Algorithm,” J. Adv. Transp., vol. 2021, 2021.
26. I. Osman, S. Ben, and G. Diallo, “Ontology Integration : Approaches and Challenging Issues,” Inf. Fusion, vol. 71, no. December 2020, pp. 38–63, 2021.

27. J. S. Salamon, C. C. Reginato, and M. P. Barcellos, “Ontology Integration Approaches : A Systematic Mapping,” ONTOBRAS, no. October, pp. 161–172, 2018.
28. M. A. Khoudja, “Ontology Matching using Neural Networks : Survey and Analysis,” 2018 Int. Conf. Appl. Smart Syst. (pp. 1-6). IEEE., no. November, pp. 24–25, 2018.
29. M. Mohammed, A. Romli, and R. Mohamed, “Eco-design based on ontology : Historical evolution and research trends Eco-Design Based on Ontology : Historical Evolution and Research Trends,” AIP Conf. Proc. (Vol. 2339, No. 1, p. 020166). AIP Publ. LLC., vol. 020166, no. May, 2021.
30. A. Fernández-Izquierdo and R. García-Castro, “Ontology verification testing using lexico-syntactic patterns,” Inf. Sci. (Ny)., vol. 582, pp. 89–113, 2022.
31. S. Tartir, I. B. Arpinar, and A. P. Sheth, “Ontological Evaluation and Validation,” pp. 115–130, 2010.
32. T. Berners-lee et al., “Tabulator: Exploring and Analyzing linked data on the Semantic Web,” Swui, vol. 2006, no. i, p. 16, 2006.
33. F. T. Fonseca, M. J. Egenhofer, C. A. Davis, and K. A. V. Borges, “Ontologies and knowledge sharing in urban GIS,” Comput. Environ. Urban Syst., vol. 24, no. 3, pp. 251–272, 2000.
34. S. Lemaignan, A. Siadat, J. Y. Dantan, and A. Semenenko, “MASON: A proposal for an ontology of manufacturing domain,” Proc. - DIS 2006 IEEE Work. Distrib. Intell. Syst. - Collect. Intell. Its Appl., vol. 2006, pp. 195–200, 2006.
35. J. M. J. F. Aldana-montes, “Evaluation of two heuristic approaches to solve the ontology meta-matching problem,” Knowl. Inf. Syst. 26(2)., pp. 225–247, 2011.
36. Y. K. Hooi, M. F. Hassan, and A. M. Shariff, “A Survey on Ontology Mapping Techniques,” Adv. Comput. Sci. its Appl., pp. 829–836, 2014.
37. S. Raunich and E. Rahm, “Towards a Benchmark for Ontology Merging Towards a Benchmark for Ontology Merging,” OTM Confed. Int. Conf. Move to Meaningful Internet Syst. (pp. 124-133). Springer, Berlin, Heidelberg., no. September 2012, pp. 124–133, 2016.
38. A. Konys, “Knowledge Systematization for ontology learning methods,” in Procedia Computer Science, 2018, vol. 126, pp. 2194–2207.
39. J. Gracia, I. Kernerman, and J. Bosque-gil, “Toward Linked Data-native Dictionaries,” Electron. Lexicogr. 21st Century Lexicogr. from Scratch. Proc. eLex 2017 Conf. (pp. 19-21)., pp. 550–559, 2017.
40. Y. Lv, “2010 Third International Conference on Intelligent Networks and Intelligent Systems A Framework for Ontology Integration and Evaluation,” 2010.
41. S. Chaabane, W. Jaziri, and F. Gargouri, “A proposal for a Geographic Ontology Merging Methodology,” 2009.
42. S. F. Pileggi, H. Crain, and S. Ben Yahia, “An Ontological Approach to Knowledge Building by Data Integration,” Int. Conf. Comput. Sci. (pp.479-493). Springer, Cham, Switz., pp. 479–493, 2020.
43. G. Stoilos, G. Stamou, and S. Kollias, “A String Metric for Ontology Alignment,” pp. 624–637, 2005.
44. S. P. Juárez, H. E. Esquivel, M. C. Suárez-figueroa, C. De Montegancedo, and B. Monte, “CreaDO – A Methodology to Create Domain Ontologies using Parameter-based Ontology Merging Techniques,” 2011 10th Mex. Int. Conf. Artif. Intell. IEEE., pp. 23–28, 2011.
45. M. Cheatham and C. Pesquita, “Semantic Data Integration.”

46. A. Solimando, G. Guerrini, and E. Jiménez-ruiz, "Minimizing conservativity violations in ontology alignments : algorithms and evaluation," Knowl. Inf. Syst., vol. 51, no. 3, pp. 775–819, 2017.
47. P. Petrov, M. Krachunov, E. Todorovska, and D. Vassilev, "An Intelligent System Approach for Integrating Anatomical Ontologies AN INTELLIGENT SYSTEM APPROACH FOR INTEGRATING ANATOMICAL," Biotechnol. Biotechnol. Equipment, 26(4), vol. 2818, pp. 3173–3181, 2012.
48. E. Ndip-agbor, J. Cao, and K. Ehmann, "Towards smart manufacturing process selection in Cyber-Physical Systems," Manuf. Lett., vol. 17, pp. 1–5, 2018.
49. J. Kumar and S. Reddy, "Implementation Of Ontology Matching Using Protégé," Int. J. Comput. Appl. Techn*ol. Res.*, vol. 2, no. 6, pp. 723–725, 2013.