Initial and Experimental Ontology Alignment Results in the Circular Economy Domain

Huanyu Li[1,*], Eva Blomqvist[1] and Patrick Lambrix[1,2]

[1]Department of Computer and Information Science, Linköping University, Linköping, Sweden
[2]Swedish e-Science Research Centre, Linköping, Sweden

Abstract
The Circular Economy (CE) domain has a nature of connecting and linking multiple cross-industry domains (e.g., manufacturing and materials) aiming to reduce value loss and avoid waste by building and implementing CE models (i.e., circular value networks) across these domains. In recent years, ontologies have been recognized as a key for representing domain knowledge in CE. Both CE-specific and domain-specific ontologies exist, with more continuously emerging. Matching CE-related ontologies can generate alignments that enhance the interoperability and reusability of such ontologies. In this paper, we present our initial efforts and findings in matching ontologies within the CE domain.

Keywords
Circular Economy, Ontology, Ontology Alignment

1. Introduction

Circular Value Networks (CVN) is a key concept in the Circular Economy (CE) domain, which intends to connect actors involved in a product's entire life cycle to enable and facilitate circular economy strategies. The CE domain has shown significant interest in using ontologies to represent domain knowledge (see overview in [1]). The Circular Economy Ontology Network (CEON)[1] [2] is one such ontology under development in our ongoing *Onto-DESIDE* project.[2] Other CE-specific ontologies exist, as summarized in [1], such as the Circular Exchange Ontology (CEO) [3] and BiOnto [4]. The aims of such existing ontologies differ from CEON. While CEON focuses on modeling CVNs that can be used for different industry use cases, the others were developed with a focus on specific use cases (e.g., materials exchange for CEO, bio-materials and bio-products for BiOnto). Moreover, CEON is intended to represent or reuse cross-industry domain knowledge essential for modeling CVNs. For instance, the 'material' concept in the majority of the domain ontologies in the materials science domain (e.g., MDO [5, 6]) represents domain knowledge relevant to simulations, experiments or processes; while in the context of CE or CVNs, it needs to represent additional knowledge from other domains (e.g., a product's composing materials, recycling techniques for materials). Recognizing the existence of various

The 2nd International Workshop on Knowledge Graphs for Sustainability (KG4S2024) – Colocated with the 21st Extended Semantic Web Conference (ESWC2024), May 27th, 2024, Hersonissos, Greece.
*Corresponding author.
huanyu.li@liu.se (H. Li); eva.blomqvist@liu.se (E. Blomqvist); patrick.lambrix@liu.se (P. Lambrix)
0000-0003-1881-3969 (H. Li); 0000-0003-0036-6662 (E. Blomqvist); 0000-0002-9084-0470 (P. Lambrix)

[1]Circular Economy Ontology Network: http://w3id.org/CEON
[2]Onto-DESIDE: https://ontodeside.eu

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

CE-specific and industry domain-specific ontologies, with more emerging, potential interoperability and reusability challenges appear. Providing alignments among these ontologies offers a solution. Firstly, for CE-specific ontologies focusing on CVNs, understanding their alignments is important. If an industry scenario requires applying multiple CE strategies modeled in different ontologies, alignments can reveal how these strategies connect. Furthermore, as CVNs often connect across industry domains, it is important to understand how CE-specific ontologies align with specific domain ontologies (e.g., materials, products, and manufacturing) so that the relationships between these ontologies can be better understood, and reused if needed. Additionally, aligning CE-specific ontologies with top-level ontologies also allows them to connect with other domains, as top-level ontologies represent universal knowledge.

The work presented in this paper, conducted in the context of the *Onto-DESIDE* project, presents experimental and initial efforts towards aligning ontologies related to the CE domain. The aim is to enhance the interoperability and reusability among CE-related ontologies. To fulfill this goal we expand our previous survey of CE-related ontologies and then establish matching tasks and a working pipeline. Further, we publish these alignments adhering to the FAIR (Findable, Accessible, Interoperable and Reusable) principles [7]. The remainder of the paper is organized as follows. Section 2 introduces related ontologies, existing efforts and experiences in the ontology matching community. Section 3 introduces the method used to produce alignments. In Section 4, we present the initial alignment results.[3] In Section 5, we briefly discuss the current results and outline directions for future work.

2. Related Work

2.1. Ontologies for the Circular Economy Domain

In our prior work [1], we have conducted a comprehensive survey of ontologies related to the CE domain and multiple cross-industry domains, and 37 ontologies were identified.[4] Aiming to enhance the interoperability and reusability within the CE domain by providing alignments among related ontologies, we expanded our ontology base beyond the ones identified to include newcomers after the previous survey was published, and included the top-level ontology, EMMO (Elementary Multiperspective Material Ontology).[5] Therefore, according to the specified domains [1], we have 6 CE-related ontologies, a number of domain-specific ontologies (5 for sustainability, 13 for materials, 14 for manufacturing, 9 for products, and 8 for logistics), and 1 top-level ontology (EMMO) to conduct ontology matching tasks on in this work.

2.2. Ontology Matching Efforts and Experiences

Since 2004, the Ontology Alignment Evaluation Initiative (OAEI)[6] has organized annual evaluation campaigns for ontology matching technologies.[7] OAEI provides test cases for comparing

[3] The result is published at http://w3id.org/CEON/alignments.
[4] Circular Economy Ontology Catalogue: http://w3id.org/CEON/catalogue
[5] EMMO: https://github.com/emmo-repo/EMMO
[6] OAEI: http://oaei.ontologymatching.org
[7] In this paper, we use "aligning" and "matching" interchangeably, both referring to the process of finding alignments which are sets of "mappings" or "correspondences" among ontologies.

and evaluating ontology matching systems. These test cases include ontologies to be matched and reference alignments, covering a broad range of diverse domains (e.g., biomedical, materials science, nutrition science, and biodiversity use cases). Additionally, OAEI focuses on evaluating how systems handle different matching scenarios, such as T-Box/schema matching, instance matching, multilingual matching, and interaction-based matching.

Most conventional ontology matching systems (although not all), such as AML [8] and LogMap [9], produce alignments based on computing similarity values between entities (e.g., concepts, relationships and instances) in ontologies [10]. A typical ontology matching framework (e.g., as seen in [10, 11]) includes pre-processing, matching based on (combinations of) different strategies including using background knowledge, lexical matching strategies, structure-based strategies and filtering over candidate mappings. Additionally, some systems incorporate reasoning, debugging, and user interaction to detect inconsistencies, remove errors, and potentially add new mappings. Recent years have witnessed the emergence of ontology matching systems based on language models such as AMD [12]. These systems may utilize pre-trained language models or large language models (LLMs), which can better understand and use word semantics for matching tasks [13]. For the former, the systems compute similarity values of entities based on their embeddings, while for the latter, they may verbalize entities into text and incorporate this text into prompts presented to LLMs to generate mappings.

3. Method

3.1. Ontology Matching Tasks and Alignment Producing Pipeline

Based on existing relevant ontologies as presented in Section 2.1 and the *Onto-DESIDE* project's intention of connecting or reusing industry domain knowledge, we establish three ontology matching tasks. They are (a): producing alignments among CE-specific ontologies, (b): producing alignments between CEON and industry domain-specific ontologies, and (c): producing alignments between CEON and top-level ontologies (e.g., EMMO). For each task, we formulate a specific question outlined in Table 1. Upon completing each task, we aim to provide an answer to the corresponding question, advancing our understanding of CE knowledge representation and increasing its interoperability and reusability.

To generate alignments among CE-related ontologies in the context of *Onto-DESIDE*, we set up a pipeline depicted in Figure 1. This pipeline builds upon general ontology matching frameworks (e.g., [11]), and additionally adds a specific step on publishing alignments in a *FAIR*

Task	Research Problem	Aim
a: CE-CE	How are existing CE ontologies aligned to each other?	Enhance interoperability and knowledge exchange among CE-related ontologies.
b: CEON-IndusDom	What are the common concepts between CEON and specific domain ontologies?	Link CEON knowledge to domain specific knowledge.
c: CEON-TopOnto	How is CEON aligned to top-level ontologies?	Link CEON knowledge to universal knowledge in top-level ontologies.

Table 1
Ontology Matching Tasks, Research Problems and Aims.

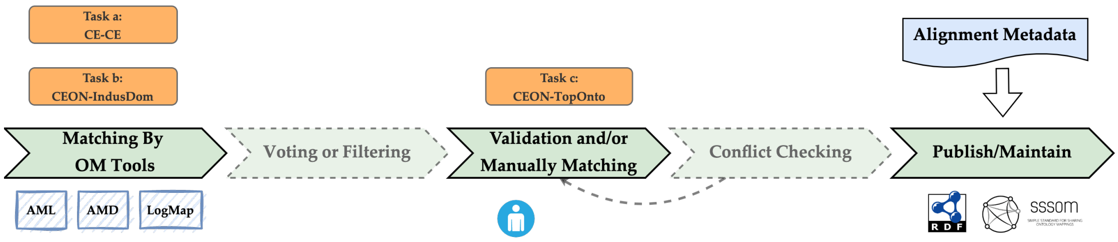

Figure 1: A pipeline of producing alignments based on the general framework outlined in [11].

way. The first step is **matching** ontologies based on three existing matching systems, which are AML [8], LogMap [9], and AMD [12]. AML and LogMap are long-term participants in OAEI, and they show state-of-the-art performance in TBox-matching tracks. AMD is a relatively new system based on pre-trained masked language model. This choice of tools covers state-of-the-art matching strategies. Also, AMD does not require significant computing resources as some other LLM-based tools. Another main step is **validation and/or manually matching** in which users validate candidate mappings or manually create new ones. While **Task a** and **Task b**, start from the first step, we use our prior experience in aligning MDO and EMMO, and start **Task c** from the manual matching step. We note that Figure 1 depicts two optional steps – **voting or filtering** and **conflict checking**, which we do not currently use, but is part of future work. The goal of the voting or filtering step is to refine the initial set of mapping suggestions yielded in the previous step.[8] The conflict checking step aims to detect and address defects (e.g., inconsistency and incoherence) that arise when connecting ontologies through alignments. The final step is publishing and maintenance, which is elaborated in the next section.

3.2. FAIR Ontology Alignments

As discussed in [14] and based on our previous experiences from OAEI, limited attention was paid to generating FAIR ontology alignments. That is, there is a lack of using rich metadata to represent alignments by matching tools. For instance, OAEI participating tools usually represent a mapping as a quadruple in the form of <source_entity, target_entity, confidence, relation_type>. Alignments between two ontologies would be a set of such quadruples. This manner of representing alignments is easy for linking the source and target ontologies with alignments into an ontology network. However, it fails to keep track of information such as matching strategies. Addressing this gap, the recently proposed Simple Standard for Sharing Ontological Mappings (SSSOM) [15, 16] defines a set of metadata to represent alignments, incorporating details such as mapping justifications (lexical matching, manual mapping) and mapping tools (algorithms used and tool versions). It also allows to annotate mappings with provenance information so that candidate mappings and validated mappings can be distinguished. OAEI has started to focus on providing FAIR alignments and adopting the SSSOM schema to an extent.

We see both advantages and disadvantages of these two options mentioned above for representing alignments. Therefore in our work, to adhere to the FAIR principles, the alignments will be represented using both the existing OAEI way and an adapted SSSOM metadata. We leverage

[8] e.g., voting based on the number of tools yielding the same mapping, and filtering based on similarity thresholds.

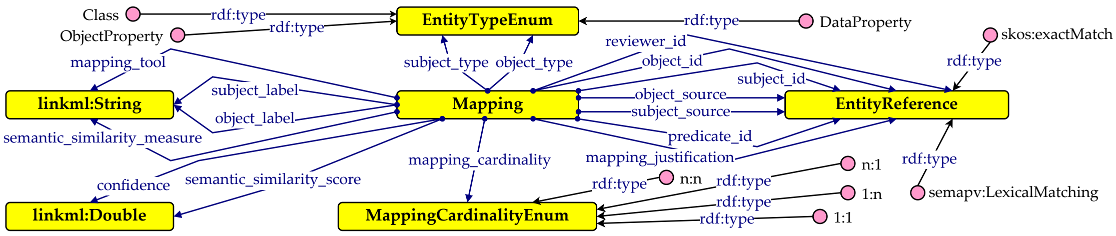

Figure 2: A part of the SSSOM schema.

subject_id	sub_label	predicate_id	object_id	obj_label	justification	confidence	tool	cardinality	reviewer
ceon:Product	Product	exactMatch	dppo:product	product	LexicalMatching	1.0	AML	1:1	HL
ceon:Statement	Statement	relatedMatch	emmo:Information	Information	ManualCuration	–	–	n:1	HL

Table 2
Example SSSOM-based mappings in tabular format (shortcut names are used for some column names).

a subset of the SSSOM schema to annotate the generated alignments. Figure 2 exemplifies a portion of the SSSOM schema[9] used in our alignment publication pipeline. SSSOM schema draws upon vocabularies from other services such as SKOS,[10] LinkML[11] and SEMAPV.[12] Essentially, SSSOM distinguishes mappings over different entity types (e.g., classes, object properties, data properties, named individuals). Each matched entity is represented as an entity reference. Moreover, `EntityReference` can represent source and target ontologies, mapping justifications by specifying their URIs. SSSOM uses LinkML's definition of `String` and `Double` to represent string and double values such as entity labels, mapping tools, confidence scores. Table 2 shows mapping examples based on the SSSOM schema, in tabular format. Moreover, we publish the alignments by employing a permanent URI[3] as an identifier through the w3id service.

4. Initial Alignment Results

Task a: CE-CE. In this task, there are six ontologies that are pairwise matched. Therefore, in total, we have 15 alignments. Equivalence mappings for the *Product* concept appear in all the alignments, which means all the six CE-specific ontologies model *Product*. Equivalence mappings for the *Material* concept appears in 10 alignments. Equivalence mappings for *Manufacturer* or *Manufacturing* are also commonly found (8 alignments). The alignment between CEON and BiOnto has the largest number of mappings, covering more mappings of concepts such as *Reuse Process*, *Production Process*, and *Recycling Process*.

Task b: CEON-IndusDom. The cross-industry domain ontologies are categorized in terms of five domains which are sustainability, materials, manufacturing, products and logistics domains. In general, there are a number of domain ontologies that reuse the PROV-O ontology[13] and/or

[9]SSSOM schema: https://github.com/mapping-commons/sssom/blob/master/project/owl/sssom_schema.owl.ttl
[10]Simple Knowledge Organization System: http://www.w3.org/2004/02/skos/
[11]Linked Data Modeling Language: http://w3id.org/linkml
[12]Semantic Mapping Vocabulary: http://w3id.org/semapv/vocab/semapv.owl
[13]PROV-O ontology: https://www.w3.org/TR/prov-o/

the Basic Formal Ontology (BFO)[14] in terms of *Location*, *Entity*, *Agent*, *Person* and *Activity* concepts. Although we classified domain ontologies into five domains according to their applications, these ontologies have in practice more overlapping conceptualizations based on the alignment results, such as the *Material*, *Product*, *Process*, and *Resource* concepts. Almost all materials-related ontologies intend to model information about the composition of materials in terms of, for instance, composing chemical entities and chemical substances.

Task c: CEON-TopOnto. There are eight mappings created manually between CEON and EMMO. Among these mappings there are three subsumption mappings which are *ceon:Datasheet* ⊑ *emmo:DigitalData*, *ceon:Statement* ⊑ *emmo:Information* and *ceon:Process* ⊑ *emmo:Process*. The remaining ones are equivalence mappings including *ceon:Material* ≡ *emmo:Material*, *ceon:Matter* ≡ *emmo:Matter*, *ceon:ChemicalEntity* ≡ *emmo:ChemicalEntity*, *ceon:ChemicalSubstance* ≡ *emmo:ChemicalSubstance*, and *ceon:MolecularEntity* ≡ *emmo:MolecularEntity*. The relatively large overlap is most likely due to that during the development of CEON, we referred to EMMO's *Matter* branch and followed the same structure.

5. Discussion and Future Work

Within the context of the *Onto-DESIDE* project, our work in this paper contains three alignment tasks with corresponding aims. We explored how existing CE-specific ontologies can be aligned to each other. This helps identify semantic connections within the CE domain. Then we aligned our developed CEON with various industry domain ontologies, and EMMO. This allows CEON to connect with a wider range of domain specific knowledge. Additionally, we made the initial and experimental alignment results in various formats available online, which helps improve the interopeability and reusability of current CE-related ontologies. The initial findings presented in Section 4 contribute to answering the research questions related to the three tasks. In the future, we will further explore and broaden our investigation into ontology alignment within the CE domain. For instance, we will update the ontology base by including new CE-specific and industry domain ontologies as well as other top-level ontologies (e.g., BFO). We are also aware of the issue that aligning ontologies based on different top-level ontologies may bring conflicts since such top-level ontologies may have different ontological commitments. Involving developing teams of top-level ontologies is one way to address the issue as suggested in [17].

It should be noted that our current work only utilizes three ontology matching systems with their default settings (e.g., matching strategies). Further exploration is needed and would be worthwhile to fully understand these three ontology matching systems' potential (e.g., different matching strategies) for CE specific ontology matching tasks. Additionally, during the process of providing alignments in different formats (OAEI-based and SSSOM-based), we realize that there still exists a gap to make alignments FAIR. Conventional ontology matching tools provide alignments which are not annotated with rich semantics. Additionally, some information, such as employed matching strategies, is not yielded by tools along with alignments. In future work, we will closely follow the developments in the OAEI and SSSOM communities in terms of FAIR alignments representation, and update our alignment metadata accordingly.

[14]Basic Formal Ontology (BFO): http://basic-formal-ontology.org

References

[1] H. Li, M. Abd Nikooie Pour, Y. Li, M. Lindecrantz, E. Blomqvist, P. Lambrix, A Survey of General Ontologies for the Cross-Industry Domain of Circular Economy, in: Companion Proc. of the ACM Web Conference 2023, 2023.

[2] E. Blomqvist, H. Li, R. Keskisärkkä, M. Lindecrantz, M. Abd Nikooie Pour, Y. Li, P. Lambrix, Cross-domain Modelling–A Network of Core Ontologies for the Circular Economy, in: Proceedings of the 14th Workshop on Ontology Design and Patterns (WOP 2023), 2023.

[3] E. Sauter, R. Lemmens, P. Pauwels, CEO and CAMO ontologies: a circulation medium for materials in the construction industry, in: 6th International Symposium on Life-Cycle Civil Engineering (IALCCE), 2018, pp. 1645–1652.

[4] C. Bicchielli, N. Biancone, F. Ferri, P. Grifoni, BiOnto: An Ontology for Sustainable Bioeconomy and Bioproducts , Sustainability 13 (2021) 4265.

[5] H. Li, R. Armiento, P. Lambrix, An Ontology for the Materials Design Domain, in: The Semantic Web - ISWC 2020 - 19th International Semantic Web Conference, 2020.

[6] P. Lambrix, R. Armiento, H. Li, O. Hartig, M. Abd Nikooie Pour, Y. Li, The materials design ontology, Semantic Web Journal (2023).

[7] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, et al., The FAIR Guiding Principles for scientific data management and stewardship, Scientific data 3 (2016) 160018:1–9.

[8] D. Faria, B. Lima, M. C. Silva, F. Couto, C. Pesquita, AML and AMLC results for OAEI 2021, in: OM 2021: The 16th International Workshop on Ontology Matching, 2021.

[9] E. Jiménez-Ruiz, LogMap Family Participation in the OAEI 2023, in: OM 2023: The 18th International Workshop on Ontology Matching, 2023.

[10] Z. Dragisic, V. Ivanova, H. Li, P. Lambrix, Experiences from the Anatomy track in the Ontology Alignment Evaluation Initiative, Journal of Biomedical Semantics 8 (2017) 56:1–56:28.

[11] P. Lambrix, R. Kaliyaperumal, A session-based ontology alignment approach enabling user involvement, Semantic Web Journal (2017).

[12] W. Zhu, AMD Results for OAEI 2023, in: OM 2023: The 18th International Workshop on Ontology Matching, 2023.

[13] Y. He, J. Chen, D. Antonyrajah, I. Horrocks, BERTMap: A BERT-Based Ontology Alignment System, Proceedings of the AAAI Conference on Artificial Intelligence 36 (2022).

[14] C. Trojahn, FAIR Ontologies, FAIR Ontology Alignments, in: Companion Proceedings of the 23rd International Conference on Knowledge Engineering and Knowledge Management, 2022.

[15] N. Matentzoglu, J. P. Balhoff, S. M. Bello, et al., A Simple Standard for Sharing Ontological Mappings (SSSOM), Database 2022 (2022) baac035.

[16] N. Matentzoglu, I. Braun, A. R. Caron, et al., A Simple Standard for Ontological Mappings 2023: Updates on data model, collaborations and tooling, in: OM 2023: The 18th International Workshop on Ontology Matching, 2023.

[17] S. Borgo, A. Galton, O. Kutz, Foundational ontologies in action, Applied Ontology 17 (2022) 1–16.