Interoperability in Logistics: An Ontology Alignment Approach

Majid Mohammadi[1], Wout Hofman[2], Yao-Hua Tan[1]

1. TUDelft, Delft, The Netherlands
2. TNO, The Hague, The Netherlands

Corresponding author: m.mohammadi@tudelft.nl

Abstract: *The logistics sector consists of a limited number of large enterprises and many Small and Medium-sized Enterprises (SMEs). These enterprises either have developed proprietary information systems or use Commercial of the Shelve (COTS) systems tailored to their business processes. It is a large number of heterogeneous systems interoperable via a large variety of (subsets of) open -,or proprietary standards. These standards typically reflect the same data sets in distinct ways so that there is a large variation of non-interoperable solutions. As a result, interoperability of information systems of different enterprises takes a lot of development and configuration time leading to high costs, with or without using an intermediate system for data transformations. (Semi-)automatic ontology alignment may solve this issue and support organizations in creating interoperable solutions. This paper presents experiments on applying ontology alignment to logistics.*

Keywords: *ontology alignment, systems integration, logistics*

1. Introduction

The supply and logistics sector consists of millions of large and Small and Medium-sized Enterprises (SMEs). Its size in the EU only is estimated 878 billion Euro (2012)[1] with over 1.2 millions of enterprises (Satta & Parola, 2011). Besides proprietary developed software, these enterprises can choose to use Commercial Off The Shelve (COTS) software from over 200 different suppliers (US)[2]. Each of these systems will have its proprietary database scheme.

The challenge is and has been to integrate the business processes of enterprises and their supporting IT solutions. Development of open standards addresses this challenge. Although these standards were developed, they did not solve the problem. Different implementation guides of (different versions of) open standards have been developed (Hofman, Towards large-scale logistics interoperability based on an analysis of available open standards, 2018)), leading to implementations that are only interoperable with additional efforts and costs. These implementation guides support process interoperability (Wang, Tolk, & Wang, 2009), which implies they will have a particular function for business processes. There are also different open standards providing the same business functionality, e.g., a transport order developed by UN/CEFACT or one of the Uniform Business Language (UBL). To address differences in implementation guides of different open standards with identical or similar functionality, commercial organizations provide transformation services between various data standards.

Applying open standards and commercial transformation services reduces the transformation challenge, but the development and implementation of implementation guides of these open

[1] https://ec.europa.eu/transport/themes/logistics-and-multimodal-transport/logistics_en

[2] https://www.capterra.com/logistics-software/

standards still take too much time for implementing supply chain innovations like agility and resilience (Wieland & Wallenburg), and synchromodal planning (Behdani, Fan, Wiegmans, & Zuidwijk, 2014). To reduce these development and implementation time for interoperability between any two organizations, this paper explores the application of ontology alignment (Euzenat & Shvaiko, 2010). Ontology alignment has resulted in different algorithms, that are tested in a competition, the Ontology Alignment Evaluation Initiative (Shvaiko, Euzenat, Jiménez-Ruis, Cheatham, & Hassanzadeh, 2018). The SANOM algorithm will be applied to ontology alignment for supply and logistics (M. Mohammadi, W. Hofman, and Yao-Hua Tan, 2019). The holy grail of ontology alignment in supply and logistics is to create semi-autonomous alignments between database schemes of different organizations, thus enabling what one could call 'plug and play' (The Digital Transport and Logistics Forum (DTLF), 2017): plug a database scheme into an open data sharing infrastructure and be able to share data with relevant business partners. Plug and play requires an open data sharing infrastructure providing standardized services (Hofman & Dalmolen, 2019).

First of all, this paper introduces ontology alignment, the OAEI, and the SANOM algorithm. Secondly, an experiment for aligning two logistics ontologies is presented. The results of a first experiment will be the basis for upcoming experiments. The result will be discussed in the context of the OAEI competition.

2. Ontology alignment

Ontologies are the proper tool to formalize the objects of a domain along with the relations that these objects have. Ontologies have been used extensively in various domains to model the underlying concepts in a formal manner. In logistics, there are also several efforts to model various aspects of data sharing between business processes of different stakeholders and to take advantages of the benefits of ontologies.

Since ontologies are created subjectively by different experts of a domain, the discrepancy among different ontologies of a domain is inevitable. Ontology alignment is the effort to reconcile the heterogeneity between different ontologies which state similar concepts of a real-world domain. In this regard, one needs to consider the sources of heterogeneity in different ontologies and use appropriate strategies to find similar concepts in two ontologies. Basically, there are two strategies for finding similar concepts coming from two types of heterogeneity. The first strategy is to consider the name of concepts and try to align the concepts with similar names. It is due to the fact that domain's experts might use different terminology for the same concepts in that domain. To gauge the similarity among concepts, one technique is to use string similarity metrics to consider the sole. Yet this technique cannot detect identical concepts with, for instance, synonymous names. For this type of heterogeneity, the use of an external resource is important. To date, WordNet is the most comprehensive resource that can boost significantly the result of ontology alignment. Another similarity is computed based on the position of the concept in the ontologies. For instance, if two concepts have many superclasses in common, then chances are that those classes are identical even though they are not named similarly.

Simulated annealing-based ontology matching (M. Mohammadi, W. Hofman, and Yao-Hua Tan, 2019), or SANOM, is an ontology alignment system which takes advantage of the well-known simulated annealing to find similar concepts in two given ontologies. Figure 1 displays the architecture of SANOM for aligning two ontologies. According to this figure, it is first required to parse the ontologies and store them into an object called *Lexicon*. Then, the similarity computations module calculates the similarity of each pair of concepts from two ontologies and store them in a hashtable called *matrix*. Prior to executing the simulated annealing, the *warm initialization* is conducted which produces a good initial guess. Then, both the initial guess and the hashtable is given to the simulated annealing and it will finally reach to a proper alignment by considering both the name similarity in *matrix*, and the structural similarity in the fitness evaluation of states in the simulated annealing.

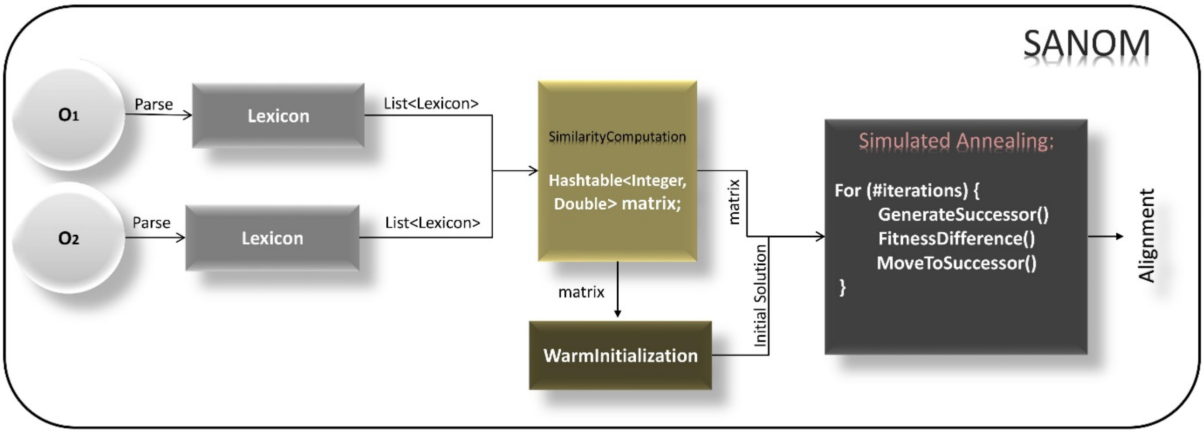

Figure 1The architecture of SANOM

3. The experiment

This section introduces the experiment where the SANOM algorithm is applied for aligning two ontologies. There are still choices to be made with respect to the experiments that will be discussed first.

3.1. Choices for the experiment

The OAEI considers a common ontology that is used to generate other ontologies. Ontology alignment algorithms are applied to align the latter ontologies to the one that is used for generating them. This context differs for supply and logistics. First of all, ontologies are not common in supply and logistics. Open standards, their implementation guides, and database schemes have to be transformed into ontologies to enable alignment. Secondly, the following alignment choices need to be considered:

1. Database scheme alignment – one could consider the alignment between database schemes of different organizations. This option is not considered feasible since databases provide more functionality than interoperability between two organizations; they support an organization in its business.

2. Functional view alignment – this option considers creating a functional view of for instance a transport order on two database schemes that will be aligned. If ontology alignment would provide optimal results, this would be an ideal situation since it does not require any formulation of open standards. It is however also complex, while it requires to align many structures all using potentially different terminology.

3. Open standard alignment – alignment of two open standards. This could be a first start which does not require any involvement of organizations (yet). Open standards are publicly available. However, the development of an ontology from an open standard

might be complex, depending on the supported functionality. An open standard for a transport order may for instance cover all transport modalities and all types of cargo.

4. Implementation guide alignment – alignment of implementation guides of an open standard. For this purpose, organizations will have to provide their implementation guides.

5. Alignment with a Canonical Information Model –integration of IT applications of one organization can be via a so-called Canonical Information Model (CIM, (Hohpe & Woolf, 2004)). This approach can also be applied for external integration, i.e. between IT applications of different organizations. It requires time for constructing a CIM, but in case the CIM can be used for automatic alignment between functional views of database schemes, it will support the aforementioned ‘plug and play’. There are different options using a CIM, like:

    a. Alignment of a functional view with the CIM;

    b. Alignment of an open standard with the CIM;

    c. Alignment of an implementation guide of an open standard with the CIM.

The proposal is to conduct the first experiment by aligning implementation guides of open standards with a CIM that has been developed in EU funded projects. This experiment can be completely controlled. The CIM is already represented as ontology, called LogiCo[3] and an implementation guide of an existing open standard will be produced that is expected to contain concepts represented by the CIM.

3.2. Setting of the experiment

Like stated, the experiment is conducted by the alignment of ontologies derived from implementation guides of open standards, with and without using LogiCo as a CIM. These choices will be further elaborated.

Using LogiCo will have some risks with respect to the experiment; it might not support the functionality of an implementation guide. To reduce this risk, an implementation guide of an open standard needs to be aligned as much as possible with LogiCo. Therefore, it is worthwhile to list the foundational concepts of LogiCo:

- **Activity** denotes some action and is relevant for the purpose of logistics, such as, for example, the activities of transport, storage, transshipment, loading, and unloading. Some activities are atomic and can be used to compose more complex activities.

- **Event** represents an occurrence of interest for the execution of a certain activity. In contrast to an activity, which denotes an action that is continuous in time, an event denotes an occurrence at a specific moment in time. For example, the departure of transport means from a location of origin and its arrival to the destination can be regarded, respectively, as starting and ending events for the transport activity.

[3] Ontology.tno.nl

- **Actor** represents organizations, authorities or individuals that offer or require activities and operate on resources related to these activities. An actor can have a **Role**, for example, customer and service provider, or shipper, consignee, forwarder, and carrier.

- **Entity** represents something that is used or exchanged during an activity. We specialize an Entity in a **Spatial Entity**, which represents tangible objects, such as an equipment or a person, and an **Intangible Entity**, which represents intangible objects, such as a modality, a characteristics or a dimension. We also define a **Temporal Entity**, which represents the start time, end time or time interval associated to activities and events. To this regard, since time is a basic (foundational) concept relevant for logistics, but common to other domains, we re-use the time ontology proposed by W3C (http://www.w3.org/TR/owl-time), instead of specifying our time ontology from scratch.

- **Location** represents the geographical area or geographical point used to define the place of origin and destination for entities and activities. Location can have different levels of granularity. Location can be coarse-grained for scheduling, since in long term planning it is sufficient to specify approximately the place of origin and destination, such as, for example, the Netherlands or the port of Rotterdam.

- **Moveable Resources** are characterized by the capability of moving on their own or being contained in another entity for the purpose of movement, and **Static Resources** are used to host and/or handle moveable resources.

An implementation guide has been constructed for an open standard representing document data for road transport. The open standard has been developed by UN/CEFACT for electronic waybills, with a specialization to the eCMR for road transport. The eCMR is an XML Schema Definition (XSD) that imports a number of other XSDs, where these XSDs import others in their turn. The eCMR XSD assigns one specific document type, the CMR, to a generic representation of data that can be stored by all types of transport documents. Thus, the core structure should as well be applicable for documents shared in other modalities. To conduct the experiments, part of the eCMR XSD is represented by an ontology. Representing the complete eCMR XSD by an ontology would be too time-consuming, we have not found open source tools to generate an ontology from an XSD. Only those parts of the eCMR XSD were chosen that could be aligned to LogiCo.

A second experiment was conducted by creating an ontology for shipping instruction, which is a transport order for sea transport. The functionality should thus be similar to that of an eCMR, with the exception that the transport mode differs. The data structure of an international booking site for sea transport provides the basis for this shipping instruction.

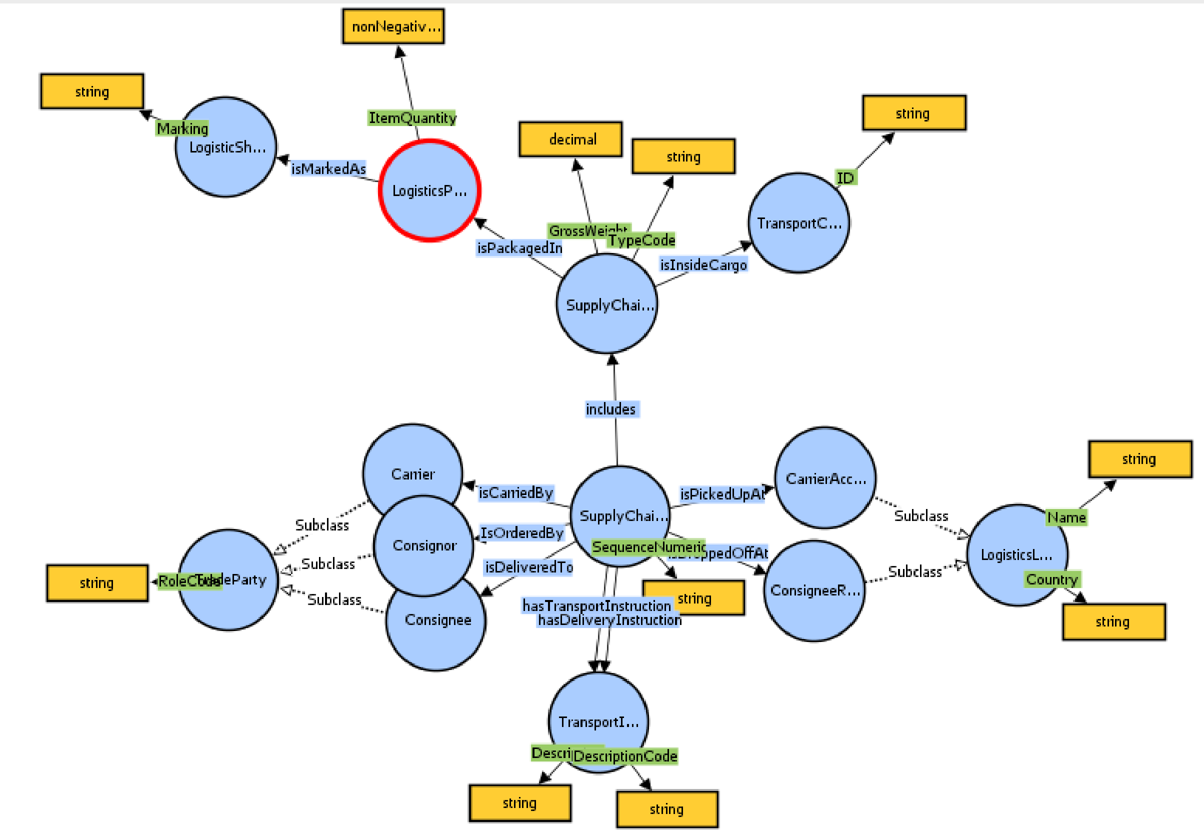

Figure 2 eCMR ontology

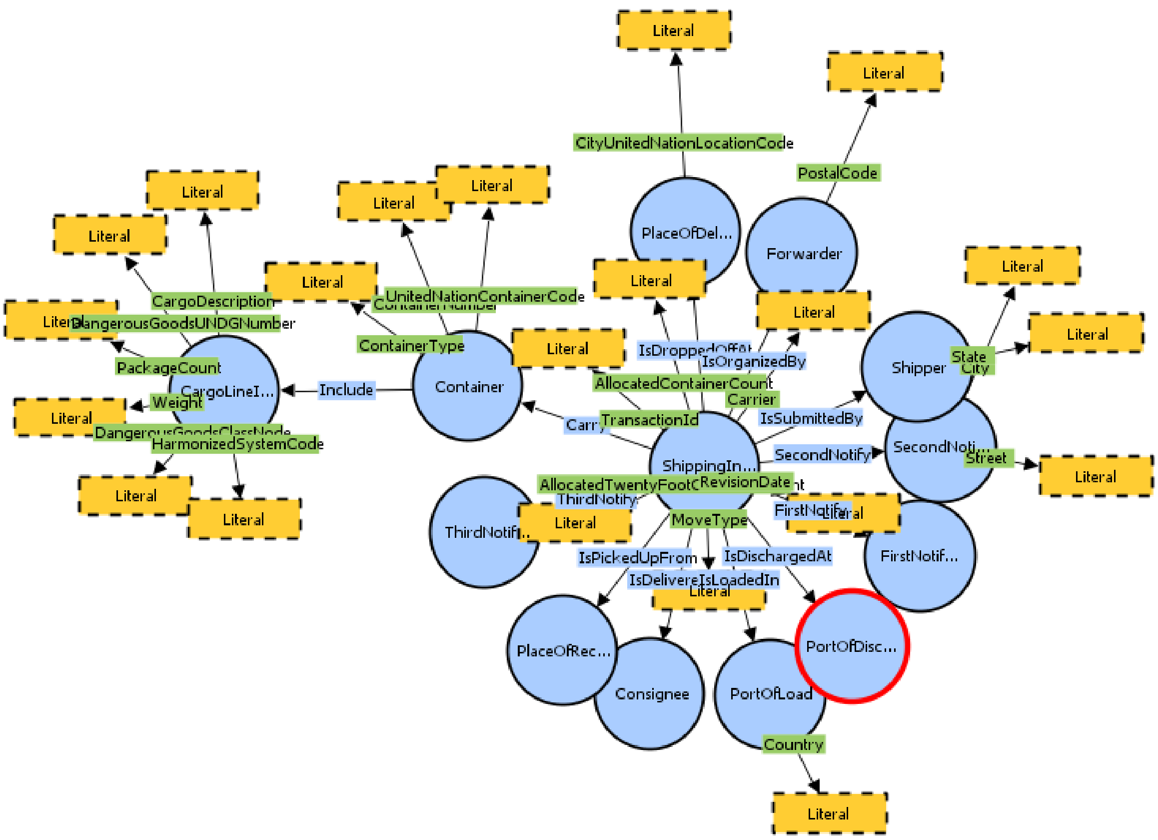

Figure 3 Shipping Instrcution Ontology

3.3. Results of the experiment

The first experiment was the alignment of the two ontologies representing implementation guides of open standards. The second experiment was to align these ontologies with LogiCo. The alignment is not satisfactory. Only concepts representing common roles of organizations in the three ontologies can be aligned with each other, but not other concepts. This is the same for alignment between the two implementation guides of open standards and these implementation guides with LogiCo.

This mismatch of alignments is due to naming conventions that differ between the three ontologies used. For instance, the eCMR has concepts like ‘SupplyChainConsignment’ and ‘LogisticsPackage’ where these concepts are not present in the other ontologies. The concept ‘SupplyChainConsignment’ is also not expected to be part of a shipping instruction ontology, since the latter represents a consignment. In general, it is not common in supply and logistics standards to use a type of prefix ‘SupplyChain-‘ for naming concepts, which makes alignment only possible to those open standards that use the same prefix for naming. The same applies to ‘LogisticsPackage’. In LogiCo, this concept is ‘Moveable Resource’ whereas, in a shipping instruction for containers, it should equal ‘container’.

Furthermore, shipping instruction has additional roles, due to delivery conditions. Besides a consignor (which is equal to a shipper) and consignee, also notifies will be mentioned. These notify have to be informed when containers arrive at a port of discharge.

Besides these differences in the naming of concepts, which will require a common data dictionary like the United Nations Trade Data Elements Directory (UNTDED)[4]. This naming difference might be solved by annotating the CIM with terms used by other ontologies. A next experiment is required with an annotated CIM.

Another difference is that these open standards represent the transport services of enterprises like carriers. For instance, a shipping line is able to transport a container between the hinterland and a port and position a container for stuffing at the location of a shipper or only transport a container between two ports. The difference is known as carrier - and merchant haulage respectively and is, in fact, a combined service. This combined service is however not represented by an eCMR, nor by LogiCo. However, the eCMR contains another modeling issue, namely that of modeling a shipment that can consist of more than one consignment. The shipment is used for data sharing with a carrier; the consignment for data sharing between a shipper and a forwarder.

These differences cannot be solved by ontology alignment. It would require agreement on modeling. Two modeling principles could be to specify transport services that cannot be decomposed in other services and to model data sharing between two stakeholders only (Hofman & Dalmolen, 2019). Associations between a shipment and its consignments are thus only internal by for instance a forwarder combining consignments into one shipment (i.e., a Less than Container Load or LCL container). An authority like customs that wants to access data of consignments in containers, the so-called data pipeline (Heshket, 2010) can do so by monitoring data sharing between shipper and forwarder and forwarder and carrier (Hofman,

[4] https://www.unece.org/tradewelcome/un-centre-for-trade-facilitation-and-e-business-uncefact/outputs/standards/untded-iso7372/introducing-untded-iso7372.html

Dalmolen, & Spek, 2019). Notice also that a carrier is never allowed to know the exact content of a container[5].

A third difference is the representation of cargo. There are three different concepts used by these three ontologies, namely LogisticsPackage, Moveable Resource, and container. One could argue that a container represents the actual cargo, but it also has packages stuffed inside. The container can be a specialization of a Moveable Resource or LogisticsPackage', although a Moveable Resource might not be equal to a LogisticsPackage. What is required besides agreement on the naming of concepts, is the associations between those concepts. The following figure shows an example of proposed concepts and their associations, where each association can carry properties. It is based on a concept of Digital Twin from simulation (Boschert & Rosen, 2016), where a digital twin can be specialized and will be in a place at a time. The same place can contain more than one digital twin, for instance containers stacked in a yard.

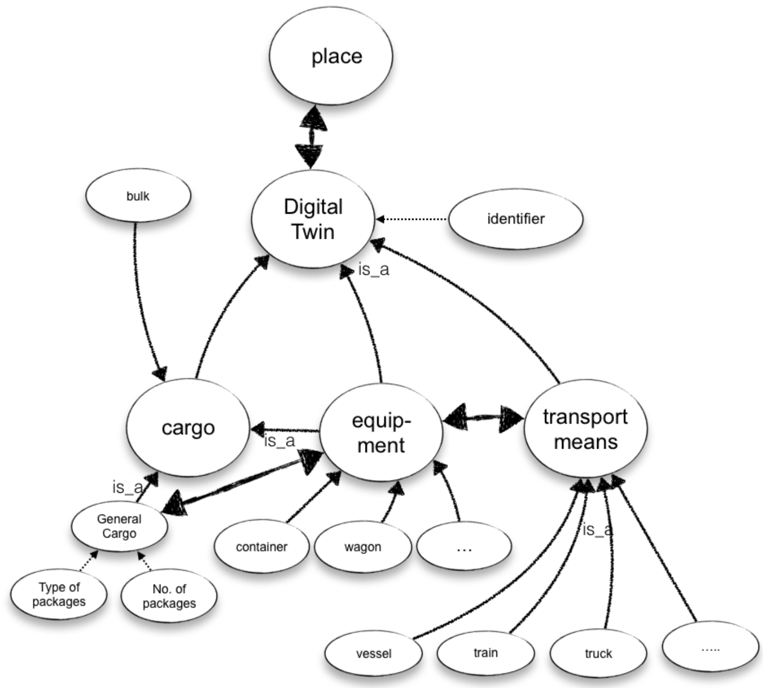

Figure 2: Modelling cargo and transport means

A question that arises is: why do these differences not arise in the context of the OAEI? Although we did not study this in detail, two potential answers could be given. First of all, the aligned ontology all stem from one ontology. Thus, the ontologies that have to be aligned with the upper one used for generating them, will all contain similar functionality. A second difference could be that the ontologies represent a domain that uses common agreed dictionaries, potential implicit ones formulated by scientific papers and books. Everyone modeling the domain will thus develop similar concepts. This could be for further study.

[5] www.rotterdamrules.com

4. Conclusions

This paper presents an experiment of ontology alignment for supply and logistics. The experiment considers the alignment of ontologies representing implementation guides of two open standards and the alignment of each of these ontologies with a Canonical Information Model represented as ontology (LogiCo). The experiments did not give satisfactory results due to differences in naming convention and systems modeled by the ontologies.

A second experiment will be performed by annotating the CIM with the terminology used in one of the implementation guides, namely the one of representing the shipping instruction. We assume that this will improve alignment results.

In view of the challenges encountered for alignment of implementation guides of open standards, it is safe to assume that alignment of (functional views of) database schemes represented as ontologies will even be more difficult. We cannot expect that database schemes use the same naming conventions and they will also have different structures.

Furthermore, it is our expectation that ontology alignment will only improve if there is a common understanding of what needs to be represented by an ontology or the CIM used for alignment has to be extended with knowledge of business service composition to address all possible standards. In the latter case, the alignment algorithm may also have to be extended.

What is not yet addressed in alignments, is the possibility that two or more concepts are merged into one or one concept is split into two or more concepts. This type of merging (and splitting) can be on a concatenation of strings, but could also consider calculation functions (e.g. length x width x height = volume). Data transformations between for instance units or code values are not part of ontology alignment and are thus not considered.

References

Behdani, B., Fan, Y., Wiegmans, B., & Zuidwijk, R. (2014). Multimodal Schedule Design for Synchromodal Freight Transport Systems. *Information Systems Logistics And Supply Chain.* Breda: Dinalog.

Boschert, S., & Rosen, R. (2016). Digital Twin - The Simulation Aspect. In P. Hehenberger, & D. Bradley, *Mechatronics Future* (pp. 55-74). Cham: Springer.

Euzenat, J., & Shvaiko, P. (2010). *Ontology Matching.* Springer-Verlag.

Heshket, D. (2010). Weakness in the supply chain: who packed the box? *World Customs Journal, 4*(2).

Hofman, W. (2018). Towards large-scale logistics interoperability based on an analysis of available open standards. *Interoperability for Enterprise Systems and Applications.* Berlin: Springer.

Hofman, W., & Dalmolen, S. (2019). Data sharing in supply and logistics networks – development and implementation of extendable, standardized platform services for the Physical Internet in an open dynamic ecosystem of organizations. *International Physical Internet Conference.* London.

Hofman, W., Dalmolen, S., & Spek, J. (2019). Supply Chain Visibility Ledger. *6th International Physical Internet Conference.* Londen.

Hohpe, G., & Woolf, B. (2004). *Enterprise Integration Patterns - designing, building, and deploying messaging solutions.* Addison-Wesley.

M. Mohammadi, W. Hofman, and Yao-Hua Tan. (2019). Simulated Annealing-based Ontology Matching. *ACM Transactions on Management Information Systems*.

Satta, G., & Parola, F. L.-w. (2011). The EU-27 Logistics Industry: structure and trends of major subsectors and national markets. *KMI International Journal of Maritime Affairs and Fisheries, volume 3, Issue 1*, 001-032.

Shvaiko, P., Euzenat, J., Jiménez-Ruis, E., Cheatham, M., & Hassanzadeh, O. (. (2018). Ontology Matching OM-2018. *Proceedings of the ISWC Workshop.*

The Digital Transport and Logistics Forum (DTLF). (2017). *An outline for a generic concept for an innovative approach to interoperability in supply and logistics chains.* Discussion Paper, EC DG Move, Brussels.

Wang, W., Tolk, A., & Wang, W. (2009). The levels of conceptual interoperability model: applying systems engineering principles to M&S. *Spring Simulation Multiconference.* Society for Computer Simulation International.

Wieland, A., & Wallenburg, C. M. (n.d.). The influence of relational competencies on supply chain resilience: A relational view. *International Journal of Physical Distribution & Logistics Management*, 300-320.