

Semantic Interoperability for Logistics and Beyond



Marco Franke , Karl Hribernik , and Klaus-Dieter Thoben 

Abstract A Semantic Mediator was conceived in the CRC 637 (The Collaborative Research Centre 637 “Autonomous Cooperating Logistic Processes” focused on adaptive logistic processes including autonomous capabilities for the decentralised coordination of autonomous logistic objects in a heterarchical structure.) to tackle problems of interoperability of heterogeneous information sources in autonomous cooperating logistics processes. Since the conclusion of the CRC 637, the Semantic Mediator has been developed further and successfully transferred to interoperability problems in different domains, including Industry 4.0, the Internet of Things, and Product Lifecycle Management. This paper will introduce the Semantic Mediator and present examples of its successful application.

1 Introduction

Now, more than ever, logistics is a cornerstone of the world’s economy, and its smooth operation is vital to supplying consumers and businesses around the globe. Strain is put on logistics networks, for example by global just-in-time supply chains (Pisch 2020) and the atomization of product-related services by mass customization and mass personalization (Kumar 2007). Logistics processes are challenged by the planning and implementation of the transport and storage of goods, considering time, quality, and costs. Moreover, the growing popularity of online shopping and the desire of customers for transparency and same-day delivery options challenge

M. Franke · K. Hribernik

BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany

e-mail: fma@biba.uni-bremen.de; hri@biba.uni-bremen.de

K.-D. Thoben (✉)

BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany

Faculty of Production Engineering, University of Bremen, Bremen, Germany

e-mail: tho@biba.uni-bremen.de

© The Author(s) 2021

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_6

logistics. For example the number of postal deliveries has increased to over 3.5 billion in 2018—this corresponds to almost 12 million postal deliveries per delivery day. The vast majority, around 84 percent, are parcels (Infografik: Fast 12 Millionen Sendungen pro Zustelltag 2019). In today's world, the pressure of these challenges on logistics networks has become even more obvious. Solutions are sought for a more resilient response to supply chain disruptions caused, for example by COVID-19, the automobile microchip and buildings material shortage, and the Suez Canal blockage.

Concepts of autonomy and self-organization to improve the resilience of logistics networks by enabling decentralized decision-making were investigated in the Collaborative Research Center 637 “Autonomous Cooperating Logistics Processes” at the University of Bremen (CRC 637)¹ and are just as relevant to address today's challenges to logistics. By distributing decision-making over autonomous logistics entities throughout a logistics system, an increase in the robustness, flexibility, adaptability, and reactivity of the overall system was expected (Freitag et al. 2004).

Intelligent logistics objects are a core concept in the realization of autonomous cooperating logistics systems, in which the logistics objects themselves are endowed with the capability to “*process information, to render and to execute decisions on their own*” (Böse and Windt 2007). This means that each physical logistics entity, like a truck, a container, or a package, needs to be equipped with a component capable of processing information, making decisions, and communicating these decisions to other entities in the system. That is, each physical logistics entity requires a digital representation.

These digital representations need access to the right information at the right time to be able to make autonomous decisions towards their logistics goals. They need to access the many different data sources of the logistics IT ecosystem on demand to get that information. Today's logistics IT ecosystem is characterized by countless communication protocols, standards, data formats and terminologies. This so-called heterogeneity is the main challenge of interoperability in logistics and one which needs to be solved for intelligent logistics objects to make autonomous decisions based on accurate and timely information.

A Semantic Mediator was proposed and developed in CRC 637 as an interoperability approach that can contribute to solving this problem. Subproject C2 “Data Integration” successfully demonstrated the viability of the approach for achieving the interoperability of intelligent logistics objects with relevant data sources in the logistics IT ecosystem.

However, the Semantic Mediator's journey did not stop there. In the past decade, similar problems of interoperability have been identified in several different research areas. The applicability of the Semantic Mediator has been successfully investigated in several of these.

¹ The Collaborative Research Centre 637 “Autonomous Cooperating Logistic Processes” focused on adaptive logistic processes including autonomous capabilities for the decentralised coordination of autonomous logistic objects in a heterarchical structure.

For example Industrie 4.0 introduced the concept of Cyber-Physical Systems (CPS) to manufacturing. As autonomous, cooperative elements interacting in a system, CPS share many of the characteristics of intelligent logistics objects. This includes the need to reliably access decision-relevant information from heterogeneous information sources throughout, for example a Smart Manufacturing environment.

Another interoperability problem with similar characteristics can be found in Closed-loop Product Lifecycle Management (Closed-loop PLM). Closed-loop PLM posits closing the loops of all information systems throughout the product lifecycle, bridging the information silos that were previously separated from each other. The aim is to optimize processes by providing access to information from processes previously available.

In this article, the background of the interoperability problem in autonomous cooperating logistics processes is recapitulated. The problem of data heterogeneity identified there is discussed and updated to capture significant developments of the past decade, such as Cyber-Physical Systems, the Internet of Things and Digital Twins. The concept of semantic mediation was introduced as a solution to that problem. An implementation of the concept is presented, the subsequent evolution of which is traced through the different use cases beyond logistics the Semantic Mediator was applied to over the past 10 years. The paper closes with conclusions and an outlook to future work.

2 Background: The Need for Interoperability in Autonomous Cooperating Logistics Systems

For autonomous cooperating logistics processes to be realized, intelligent logistics objects must be able to process the information required to make and execute decisions towards their logistics objectives. This means that the intelligent logistics objects need to be able to access the information relevant to their decisions at any time throughout the logistics processes. Consequently, the objects not only need to be able to communicate with each other, but they need to be integrated into the overall logistics ecosystem in such a way that they can interoperate with any data source required to fulfil their information needs. The following subsections take a closer look at the characteristics of these data sources and what data representation, exchange formats and standards are widespread in logistics that need to be considered to achieve interoperability with those sources. This section attempts to bring the results of investigations into these issues done in CRC 637 up to date by reviewing the major developments in IT of the past decade, such as Cyber-Physical Systems, Digital Twins, and the evolution of the Internet of Things.

2.1 An Updated Look at Data Sources in Logistics

The major data sources relevant for autonomous cooperating logistics processes were categorized in CRC 637 as (1) Logistics IT Systems, (2) Intelligent Logistics Objects, (3) Digital Counterparts and (4) Sensors and Actuators (Hribernik et al. 2010). Whilst the first and last of the four categories are still appropriate today, the second and third need to be revisited in the light of the development of these topics over the past 10 years. Table 1 shows a restructured list of relevant data sources along with their respective interfaces and standards based on (Hribernik et al. 2011), which has been updated to include Digital Twins, Asset Administration Shell (AAS), CPS, and developments in the field of the Internet of Things. The following subsections discuss the individual categories in more detail.

Table 1 Updated table of major relevant data sources

Data sources	Type(s)	Interface/standard
Logistics IT systems	General	EDIFACT EANCOM
		EANCOM XML
	ebXML	
	SAP compliant	SAP RFC (remote function call)
	Other	REST JSON XML Bespoke proprietary
Digital counterparts	Digital twins Cyber-physical systems Intelligent products Software agents	RAMI 4.0 AAS, others RAMI 4.0 AAS Dialog, PMI, EPCIS, OSGi ACL (agent communication language) Agent proxies Dialog
	Other	Bespoke proprietary
Internet of things	Internet of things	OPC-UA
		OPC DA
		OPC XML DA
		MQTT
		PMI
	Java-based	OSGi
	OGC compliant	SensorML
	General	GDI ORIN API
Other sensors	Proprietary formats	
Other actuators	Proprietary formats	

2.1.1 Logistics IT Systems

With regards to the first category, the IT systems used in the logistics sector today remain as varied, complex, and heterogeneous as investigated in CRC 637. Many different systems from different vendors are in use, many with their own proprietary data representations and interface schemas. To name but a few examples, they include Enterprise Resource Planning (ERP), Warehouse Management Systems (WMS), Transport Management Systems (TMS), Supply Chain Management (SCM), and Disposition Systems (DSS). In addition to solutions developed by solution providers, proprietary in-house solutions are widespread.

Exchange formats are the primary factors driving standardization in the field of logistics and the main means of interoperability between the relevant IT systems. Specific EDIFACT subsets like FORTRAS or the interface languages of popular ERP (Enterprise Resource Management) systems are predominant examples (Ballnus 2000). In addition, REST services and micro-service architectures have established XML and JSON as de facto data formats for sharing data. Moreover, document-based databases have supported these data formats to reach the persistence layer in legacy systems. Thus, application scenarios related to Industry 4.0 and IoT need to support these data formats.

2.1.2 Digital Counterparts

The second category originally encompassed Intelligent Logistics Objects, which constituted uniquely and automatically identifiable physical logistics objects. Category three included the decision-making components of intelligent logistics objects, such as software agents or holons. Over the past decade, these concepts have been consolidated together with similar concepts like Smart and Intelligent Products, Closed-loop PLM and others into notions such as CPS (Cyber-Physical Systems) and Digital Twins. For that reason, both have been consolidated here into Digital Counterparts.

CPS build upon concepts such as autonomous control, holonic manufacturing systems, intelligent manufacturing, IoT and embedded systems and are thus closely related to intelligent logistics entities (Lee 2006; Monostori 2014). CPS are “*systems of systems of autonomous and cooperative elements connecting with each other in situation dependent ways, on and across all levels of production, from processes to machines up to production and logistics networks, enhancing decision-making processes in real-time, response to unforeseen conditions and evolution along time*”. (Cardin 2019). Like autonomous control in logistics, CPS systems promise to improve the adaptability, scalability, resiliency, safety, and security of industrial systems (Muller 2017), by contributing to making decentralized decisions (Lu 2017).

Digital Twins can be understood as comprehensive digital representations of physical assets, comprising their design and configuration, state, and behaviour. The term Digital Twin is still, however, used inconsistently throughout literature,

and comprehensive reference models are lacking (Kritzinger et al. 2018; Lu et al. 2020). The RAMI 4.0 reference architecture now promotes the AAS interface as a single point of entry to the digital representations of physical assets and has been put forward as the link to Digital Twins. While more work is required before this is widely accepted, AAS needs to be taken into consideration when looking at interoperability with digital representations of physical assets like logistics entities going forward (Tantik and Anderl 2017; Anderl et al. 2018). Even though Digital Twins show clear parallels to concepts such as software agents or holons, the concept of autonomy has not yet been investigated adequately in conjunction with Digital Twins.

2.1.3 The Internet of Things

The fourth category includes sensors and actuators not covered by the previous three categories. In the past decade, the success of the Internet of Things has led to a mushrooming of IoT standards, with many hundreds competing for primacy in the marketplace. Even though the most recent years have shown a trend towards consolidation of standards in the industry around OPC-UA and MQTT, the multitude of implemented proprietary standards have only made the problem of interoperability more difficult. These changes have been reflected in the list of relevant data sources shown in Table 1.

2.2 *Summarizing the Interoperability Problem in Complex Logistics Systems*

The distributed, heterogeneous IT ecosystems outlined above, which prevail in complex logistics systems, demand a unique approach to interoperability. The heterogeneity and distribution of the data sources as well as their syntactic and semantic representations, pose a significant challenge. In addition to their geographical distribution, the data sources are distributed across multiple stakeholders in the logistics value network. Nevertheless, autonomous control cannot be realized without providing access to the required information in a predictable and reliable manner, regardless of the formal, structural, and physical properties of the underlying individual systems, standards, and formats (Hans et al. 2008). Some constraints identified in the CRC 637 no longer pose as much of a challenge as they did a decade ago. Specifically, concerns about the limited availability of intelligent logistics objects are almost negligible in today's world of ubiquitous wireless internet connectivity. The imminent widespread deployment of 5G networks will make this issue even less problematic, with the possible exception of logistics objects moving through remote areas.

3 The Problem of Data Heterogeneity

This section takes a closer look at the problem of data heterogeneity in distributed IT ecosystems such as complex logistics networks. It will furthermore introduce semantic mediation as the solution approach proposed in CRC 637 to achieve interoperability of intelligent logistics objects under these conditions.

3.1 Heterogeneity Classification

Data is represented using data formats that define the data’s syntax. The syntax defines how the information has to be represented. The syntactical elements define which semantics of an item of information can be represented. For example the goal is to store a family tree in a data format. A representation in the CSV format could store a person, children, and parents as columns. In this case, additional knowledge is required to interpret the columns as parents and children. In contrast, XML or JSON can represent the family tree structure natively. The example demonstrates that different data formats have different syntax and capabilities. The interoperable representation of an item of information or the transformation between the data formats needs to consider the different capabilities of these data formats. Data integration conflicts may arise due to heterogeneity. A data integration conflict occurs if an item of information is extracted from one data source and inserted into another data source (Goh 1996; Wache 2003). For example the transformation of the family tree from JSON into CSV results in schematic conflicts because the taxonomy cannot be represented natively. An overview of possible data integration conflicts is shown in Table 2.

3.2 Interpretation Levels

Semantic mediation focuses on the interoperable representation of heterogenous data, which stores tuples containing data and meta-information. Using tuples prevents a false interpretation of data by different stakeholders. Information can be defined as “...data that has been given meaning by way of relational

Table 2 Types of possible data integration conflicts (Goh 1996)

Schematic Conflicts	Semantic Conflicts	Intensional Conflicts
Data type	Naming	Domain
Labelling	Scaling	Integrity constraint
Aggregation	Confounding	
Generalization		

Fig. 1 JSON example

```
{
  "parcel" :{
    "to":{
      "name" : "Franke" ,
      "street" : "Muster str."
    }
  }
}
```

connection . . .” (Bellinger et al. 2011). An unambiguous interpretation of the information inside each data source is possible on different levels: lexical, syntactical, morphological, semantic, and pragmatic understanding (Ören et al. 2007). Interpretation on lexical, syntactic, and morphological levels of understanding allows the structure of the information to be recognized and relevant entities to be grouped together. However, the meaning remains unclear. Common data formats for data exchange in Industrie 4.0 and IoT include CSV, JSON and XML. These data formats do not include enough meta-information to enable the false-free interpretation for the semantic understanding. An example of a parcel’s address in JSON, including the data and its meta-information, is shown in Fig. 1.

The parcel’s address defines on the morphological understanding that an addressee is defined by two attributes. The meaning of these two attributes is unclear without any additional meta-information. The reader cannot know that the first attribute describes the surname of a person and that the string “str.” inside the value of the second attribute street is an acronym for the street. This meta-information is not contained in the data. While a human reader immediately correctly interprets the JSON attributes name and street, this semantic understanding needs to be added programmatically as part of a data integration solution.

4 Solution Approach: A Semantic Mediator for Complex Logistics Systems

The goal is to represent the information along with all required meta-information to facilitate interpretation on the level of semantic understanding, rendering the information independent of its original data format and neutralizing data format restrictions with respect to syntax and semantics. To do so, each information needs to be represented in the interoperability model supporting the appropriate amount of meta-information.

The proposed approach uses a network of ontologies as an interoperability model. Ontologies are formal, partial specifications of an agreement over the description of a domain (Guarino 1998). Specific domains are modelled since broader ontologies are generally not convincing or feasible (Masolo et al. 2002).

Ontologies are modelled for different levels with specific purposes. Ontologies that model basic concepts and properties are called upper ontologies and are

important for integrating heterogeneous knowledge from different sources (Mascardi et al. 2007). Examples of upper ontologies include Basic Formal Ontology (BFO), Business Objects Reference Ontology (BORO), Conceptual Reference Model (CIDOC), and Descriptive Ontology for Linguistic or Cognitive Engineering (DOLCE). Domain ontologies make use of upper ontologies for basic concepts and properties but describe all specifics with their own concepts. Domain ontologies in logistics include GenCLOn for urban freight transport (Anand et al. 2012) or ontologies for logistics services (Scheuermann and Hoxha 2012; Deng et al. 2019).

Whilst the interoperability approach proposed here uses upper and domain ontologies, it does not use a fixed ontology network or monolithic ontology since the data sources it needs to support are varied and may change over time. It proposes a loose coupling of existing domain ontologies to create a network of ontologies based on data sources such as intelligent logistics objects, IoT, CPS, and legacy systems. The adaptive ontology network proposed is appropriate for the movement of objects throughout a complex network and across stakeholder and process boundaries. Moreover, the approach needs to support a bidirectional flow of information between systems while allowing each system to access the information in its own data format. To achieve this, the role of the ontology network is as an intermediate representation, which is shown in Fig. 2. Each system transforms its data into the interoperability model. Based on the interoperability model, the transformation is possible with one further transformation step into each target data format. This approach guarantees scalability with respect to the number of involved systems.

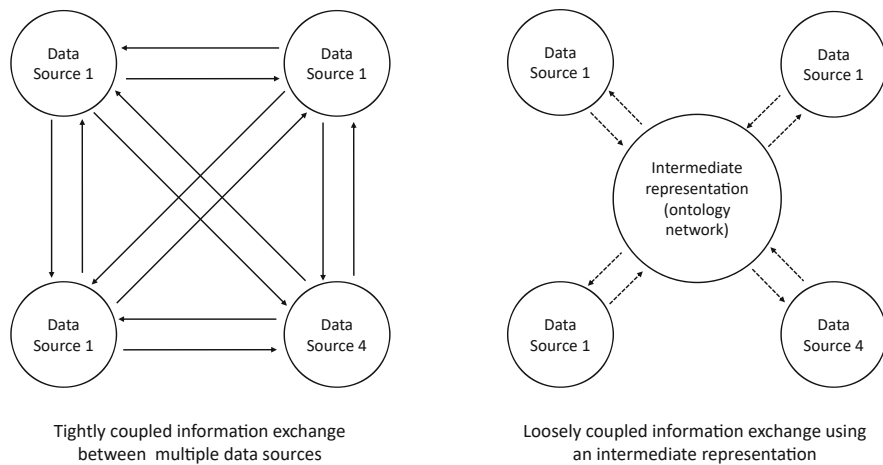


Fig. 2 Tight coupling vs. integration approach with an intermediate representation

4.1 Data Integration Approach

The information gathered from a data source can be mapped onto the interoperability model using Global as View (GAV), Local as View (LAV) or hybrid (GLAV) approaches (Lenzerini 2002). The approach proposed here uses GAV. A data source schema is mapped onto a global view, which is the interoperability model as a network of ontologies. Each information query in GAV needs to be formulated based on the global view based on ontological concepts and properties. This allows a query to be designed independently of specific data sources. The proposed approach uses the query language SPARQL to support this.

A virtual data integration approach, which aggregates the data on the fly and does not store the result, is followed. It is executed each time a SPARQL query is issued. The SPARQL query defines which amount of information should be requested from all connected heterogeneous data sources. The query-based data integration uses a mediator-based approach that collects and aggregates data from heterogeneous data sources. The approach, shown in Fig. 3, uses the component Semantic Mediator to aggregate partial results from the data sources. Each data source is connected to the Semantic Mediator via a wrapper, which translates data from the local view into the global view. In the following, both components are presented in detail.

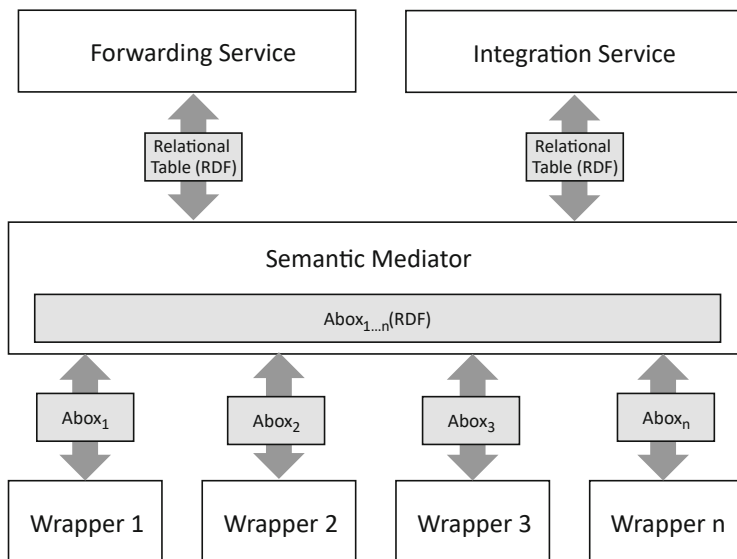


Fig. 3 The semantic mediator approach

4.1.1 Semantic Mediator Core Component

The core component implements Semantic Mediator functionality. Its input is a SPARQL query, and its output is an intermediate representation as ontology. The data integration starts with the identification of relevant data sources. The identification process maps the query's ontological concepts and properties on the local ontologies provided by the data sources. The Semantic Mediator can thus derive data source specific queries and issue them to the IT systems in question. These queries only include the ontological concepts that can be addressed by the data source. Each data source responds to the respective query by providing the Semantic Mediator with a set of individuals for the requested concepts. The semantic mediator collects all individuals and joins them. The resulting single ontology includes both the individuals and the definition of the concepts. This ontology can be delivered as a result or used as an intermediate representation for post-processing.

4.1.2 Wrapper

Wrappers implement the links between data sources and the Semantic Mediator. These links are based on configurations that map the local views of the data sources onto the global view of the Semantic Mediator. There is no need to implement interfaces to bind a data source to a wrapper or implement the data acquisition and the data transformation for each linked data source. Different kinds of wrappers are required to bind specific types of data sources. Wrappers have been prepared to connect to different relevant data sources listed in Table 1, such as EDIFACT EAONCOM, SQL databases, CSV files, XML files, REST services or streaming platforms such as Kafka. Each wrapper configuration contains an ontology defining the content of the data source and a wrapper type-specific mapping file defining the mapping between the source schema and the ontology schema.

5 Data Transformation Examples

The following example shows how a JSON Wrapper transforms JSON data into an RDF/XML ontology. Figure 4 shows the company name represented in different data formats. The company's transformation of all shown data source snippets would need the application of the EDIFACT, ANSI ASC X12, XML, and JSON wrappers. The outcome of these wrappers would be four individuals of the concept Company. The Semantic Mediator receives the four individuals and performs a join. In this example, all four data snippets contain a similar company name. Thus, the data integration would result in only one individual inside the ontology (see Fig. 6). The following shows the transformation of a JSON String into an ontological individual.

EDIFACT	ANSI ASC X12
NAD+BT+MyCompany::91	PER*CR*MyCompany*TE
O-MI	Custom
<pre> <omi:msg xmlns="odf.xsd" xsi:schemaLocation="odf.xsd"> <Objects> <Object> <id>My- Company</id> </Object> </Objects> </omi:msg> </pre>	<pre> { "company":{ "name": "MyCompany" } } </pre>

Fig. 4 Name of a company represented in different data formats

```

<DatatypePropertyMapping propertyName="Name"
Unit="String">
  <SubClass>Company</SubClass>
  <Tag>Root/company/name</Tag>
</DatatypePropertyMapping>

```

Fig. 5 Transformation rule for the property name

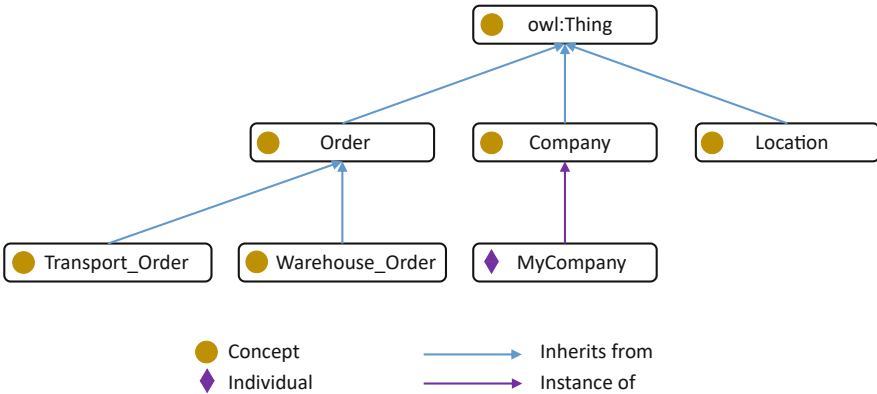


Fig. 6 The result of the transformation into an ontology

The JSON Wrapper follows the configured transformation rules and applies them to the data. Figure 5 shows a transformation rule that transforms the JSON String of the example in Fig. 4 into the ontology of Fig. 6.

For that purpose, the transformation rule defines that the rule is applicable for the property Name of the concept Company. Subsequently, the rule defines that the value for property Name is located in the path Root/company/name in the JSON file. Thus, the wrapper applies all rules sequentially, and the application of each rule extends the ontology with additional statements.

The presented example for the transformation of data into ontologies works similar for all wrappers. The only exception is that the structure of the transformation rules is dependant on the data source.

6 Evolution of Semantic Mediator

The semantic mediation approach for interoperability in complex logistics systems was proposed in 2009. Since then, the approach has been successfully applied to different interoperability problems in various industrial sectors. While the core design of the approach stood the test of time, and the overall system has evolved. The following describes a series of select application scenarios along with their impact on the mediator as a historical outline. Subsequently, the evolution steps are presented as a timeline.

6.1 Application Scenarios as a Historical Outline

In the subproject C2 of the Collaborative Research Centre SFB 637, research was carried out between 2008 and 2012 on developing an interoperability approach for the integration of logistics data to support self-controlling logistics processes. The main result of the work was a prototypical mediator component called “Semantic Mediator“. It focused on a virtual data integration approach whereby the logistics data sources were the primary data sources. This prototype was applied as preliminary work for later research projects.

6.2 Interoperability for Cyber-Physical Systems

The goal of the research project CyProS was the development of a representative spectrum of CPS modules for production and logistics systems. For that purpose, a framework for designing CPS-based solutions in production environments was in the focus.

The framework emphasized the need for interoperable information exchange for the seamless collaboration between CPS and legacy systems. The Semantic Mediator made it possible to transform data from the CPS and the legacy systems into one ontology to create a common global view (as motivated in Fig. 7) for

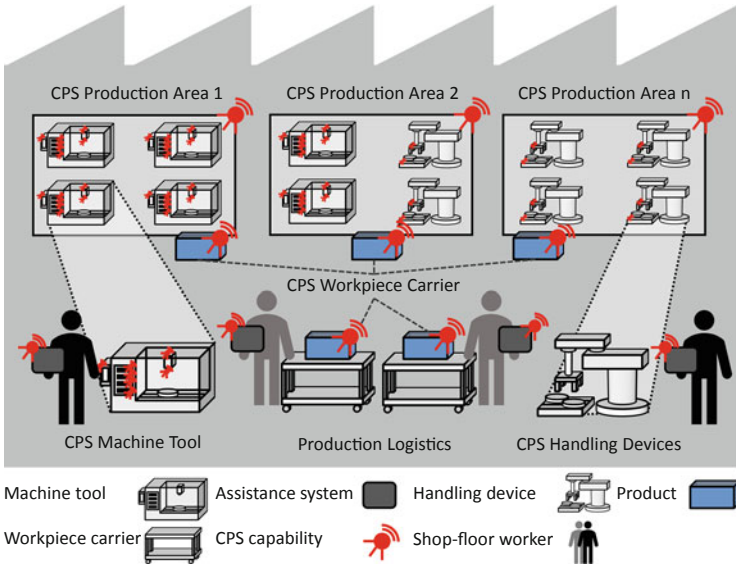


Fig. 7 Enabling interoperability between CPS and legacy systems (adapted from Reinhart et al. 2013)

autonomous and decentral decision-making processes. The dynamic registration and deregistration of data sources for an open-world application scenario was the focus. Here, the adaptation of the domain ontology in the role of a shared global view was developed depending on the connected systems. In addition, wrappers for CPS-specific interfaces were developed.

6.3 Interoperability of Product Usage Information for Product-Service-System Improvement and Design

A further interoperability problem was found in the field of Product-Service Systems (PSS) (Hribernik et al. 2018). These combinations of services and products are reliant on product usage information (PUI)—information about how a product is used—to help companies provide services offers for their products throughout their lifecycles (see Fig. 8). Previously, PUI was collected in processes such as customer relations management and MRO (maintenance, repair and overhaul) via repair logs, call centres or helpdesks. The digitalization of the product lifecycle, the Internet of Things, CPS, and social media have introduced new sources of valuable PUI which can be used in product and service development systems, e.g. Product Data Management (PDM), Product Lifecycle Management (PLM), Computer-Aided Technologies (CAx) and simulation tools, and the related processes and

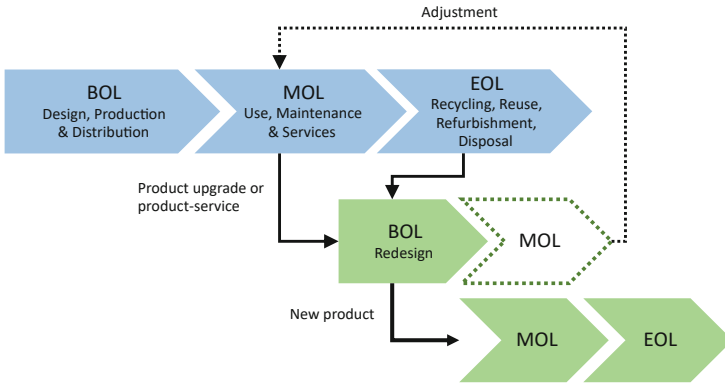


Fig. 8 PUI feedback loops in PSS Improvement and Design (adapted from Hribernik et al. 2016)

methodologies to improve current PSS and design new ones. The Semantic Mediator was consequently extended to efficiently and seamlessly integrate and apply PUI to PSS (re-)design processes. A major extension in this context was the development of a wrapper that uses text mining algorithms and NLP (natural language processing) to extract information from social media and integrate it with other PUI for use in PSS development and design.

6.4 Sustainable Manufacturing: Extending the Useful Life of Major Capital Investments and Large Industrial Equipment

A further use case for the Semantic Mediator was found in the field of sustainability in manufacturing, more specifically in extending the useful life of major capital investments and large industrial equipment. Here, a platform is being developed which offers services, ranging from the Digital Twin’s setup, modernization actions to diagnose and predict the operation of physical assets to the refurbishment and remanufacturing activities towards End-of-Life (see Fig. 9). These services depend heavily on information gathered from machines and legacy systems. The interoperability challenge here is to achieve a scalable and extendable service platform. For that purpose, the Semantic Mediator enables the translation of the streaming data as well as legacy data via specific ontologies into the input data formats of the service layer. To support this feature, there was a paradigm shift in the data integration approach. Static data sources were until then the primary goal for the data integration, which stored the data permanently and can be requested on the user demand. This assumption changed to support also dynamic data sources whose data are volatile and must be transformed immediately independently from the current user demand. Examples of these new types of data sources are comments



Fig. 9 A platform for extending the useful life of major capital investments and large industrial equipment (The LEVEL-UP Project 2019)

in social media distributed by websites or sensor values distributed by distributed event streaming platforms such as Apache Kafka.

The results were new wrapper types, and the support of interoperable information flows based on streaming data. Moreover, the virtual data integration approach shifted to hybrid approach that combines the virtual and the physical data integration approaches. Current research projects suggest another paradigm shift. The focus of the ontology shifts from the target result to an intermediate representation. The uses cases need to consume the integrated information not as ontology but rather requested the translation into the target data formats of the use cases. For example sensor values should no longer be provided as an ontology but should be uploaded directly to an influx database for AI methods.

7 Outlook and Conclusion

The original aim of the Semantic Mediator was to solve the interoperability problem of intelligent logistics objects in autonomous cooperating logistics processes. The Semantic Mediator was successfully shown to be able to contribute to solving the problem with its virtual data integration approach, which focuses on providing exactly the right amount of information required on demand and defined by each individual query. The Semantic Mediator was shown to be able to provide intelligent logistics objects with the information required for taking autonomous logistics

decisions. That information is stored in heterogeneous, highly distributed systems such as logistics enterprise systems, Internet of Things data sources, sensors, and embedded systems. It was shown that the chosen interoperability approach was capable of solving the heterogeneity conflicts that may arise in the heterogeneous and distributed IT ecosystem found in complex logistics systems. The resulting software solution, the Semantic Mediator, comprised the core mediator component along with a number of wrappers focused on the logistics application domain. These included EDIFACT EANCOM, CSV, and EPCIS wrappers, which were themselves configurations of generic wrappers developed to handle semi-structured texts, database management systems, and enterprise systems. In addition, a gateway module was developed to provide the Semantic Mediator with capabilities to access the Internet of Things and other embedded devices.

The subsequent research in complementary problem areas and industrial sectors led to an extension of the capabilities of the Semantic Mediator. These included wrappers for the connection of SQL and XML data sources which were added up to 2015. The next evolutionary step was taken in 2017 when data integration was expanded to support also physical data integration. For this purpose, the possibility was created to forward data to a data sink. A triple store was selected as a data sink for this purpose. Furthermore, a new generation of wrappers has been introduced that significantly improved the capabilities of the mediator to handle dynamic data sources, such as streaming sensors or social media data. The current expansion stage supports Kafka as a dynamic data source and provides the feature that transformed data can also be published on streaming platforms such as Kafka. This enables the mediator to homogenize and stream data from heterogeneous data sources using a network of ontologies within the framework of a continuous flow of information. Possible consumers of these streams are services such as predictive maintenance or anomaly detection, which rely on data from various sources.

The Semantic Mediator is well prepared for future digitization challenges beyond Industry 4.0, Autonomous Logistics, Digital Factories and Closed-loop PLM. The challenging requirements of interoperability for autonomous cooperating logistics processes on the one hand, and on the other the concept of intelligent logistics objects which foreshadowed later developments such as CPS, Digital Twins, Smart and Digital Factories influenced the design of the Semantic Mediator in such a way that it proved applicable to many novel and challenging interoperability problems, a selection of which are outlined above. It is a testimony to the research done in CRC 637 and more broadly in LogDynamics that even after 10 years, the principles underlying the Semantic Mediator can still be considered novel and relevant to today's interoperability problems.

Even though the Semantic Mediator has evolved significantly over the past decade, there remain research challenges that have not yet been solved. Apart from tapping additional application scenarios, the capability for the mediator to (semi-)automatically integrates new and unknown data sources has not yet been realized. This adaptive, (semi-)automatic approach to semantic data integration could significantly improve the flexibility of autonomous processes in logistics, manufacturing, the product lifecycle, and other domains. Starting points for research

towards adaptive semantic data integration have already been identified and include applying principles of ontology learning, algorithmic ontology mapping, and methods of artificial intelligence and software engineering for automatic configuration and deployment. That means the next 10 years of research into interoperability in complex industrial systems will remain as exciting and productive as the last decade.

Acknowledgements This research was supported by the German Research Foundation (DFG) as part of the Collaborative Research Centre 637 “Autonomous Cooperating Logistic Processes”, the German Federal Ministry of Education and Research (BMBF) as a part of the project CYPROS (Grant Agreement No. 02PJ2477), and the European Commission through the projects FITMAN (Grant Agreement No. 604674), FALCON (Grant Agreement No. 636868) and LEVEL-UP (Grant Agreement No. 869991). The contents of this paper reflect only the authors’ view, and the respective funding organizations are not responsible for any use that may be made of the information it contains. The authors wish to acknowledge the funding organizations and all project partners for the fruitful collaboration.

References

- Anand, N., et al.: GenCLON: an ontology for city logistics. *Expert Syst. Appl.* **39**(15), 11944–11960 (2012). <https://doi.org/10.1016/j.eswa.2012.03.068>
- Anderl, R., et al.: Digital twin technology – an approach for Industrie 4.0 vertical and horizontal lifecycle integration. *Inf. Technol.* **60**(3), 125–132 (2018). <https://doi.org/10.1515/itit-2017-0038>. De Gruyter Oldenbourg
- Ballnus, R.: Erfolg mit EDI und E-Commerce: Handlungsempfehlungen für die Abstimmung und Organisation interorganisationaler Netzwerke.. Available at: https://books.google.de/books?hl=en&lr=&id=7tQGL6hi19IC&oi=fnd&pg=PA11&dq=Erfolg+mit+EDI+und+E-commerce+-+Handlungsempfehlungen+f%C3%BCr+die+Abstimmung+und+Organisation+interorganisationaler+Netzwerke&ots=ecGT8N0-jo&sig=JNdAjWpAUQvkjxMVEuoTI_hPtCM (2000). Accessed 5 July 2021
- Bellingier, G., Castro, D., Mills, A.: Data, Information, Knowledge, and Wisdom (2011)
- Böse, F., Windt, K.: Catalogue of criteria for autonomous control in logistics. In: *Understanding Autonomous Cooperation and Control in Logistics: the Impact of Autonomy on Management, Information, Communication and Material Flow*, pp. 57–72. Springer, Berlin (2007). https://doi.org/10.1007/978-3-540-47450-0_5
- Cardin, O.: Classification of cyber-physical production systems applications: proposition of an analysis framework. *Comput. Ind.* **104**, 11–21 (2019). <https://doi.org/10.1016/j.compind.2018.10.002>
- Deng, Q., et al.: An ontology framework for multisided platform interoperability. In: *Proceedings of the I-ESA Conferences*. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-13693-2_36
- Freitag, M., Herzog, O., Scholz-Reiter, B.: Selbststeuerung logistischer Prozesse—ein Paradigmenwechsel und seine Grenzen. *Industrie Management*, 20:23–27. Available at: <http://www.sfb637.uni-bremen.de/pubdb/repository/SFB637-A5-04-002-II.pdf> (2004). Accessed 5 July 2021
- Goh, C. H.: Representing and reasoning about semantic conflicts in heterogeneous information systems. Massachusetts Institute of Technology. Available at: <https://elibrary.ru/item.asp?id=5551531> (1996). Accessed 6 July 2021
- Guarino, N.: Formal ontology and information systems. In: *Formal Ontology in Information Systems*, pp. 3–15. IOS Press, IOS Press, (1998). Available at: <https://books.google.de/>

- books?hl=en&lr=&id=Wf5p3_fUxacC&oi=fnd&pg=PR5&dq=Guarino,+N.:+Formal+Ontology+in+Information+Systems&ots=noSG1StCNL&sig=OiecmvXspIaw_6l42X2OaW7_kr4. Accessed 5 July 2021
- Hans, C., Hribernik, K.A., Thoben, K.-D.: An approach for the integration of data within complex logistics systems. In: *Dynamics in Logistics*, pp. 381–390. Springer, Berlin, Heidelberg (2008). https://doi.org/10.1007/978-3-540-76862-3_38
- Hribernik, K.A., et al.: A semantic mediator for data integration in autonomous logistics processes. In: *Enterprise Interoperability IV*, pp. 157–167. Springer, London (2010). https://doi.org/10.1007/978-1-84996-257-5_15
- Hribernik, K.A., et al.: Impacts of data integration approaches on the limitations of autonomous cooperating logistics processes. In: *Autonomous Cooperation and Control in Logistics*, pp. 247–267. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-19469-6_18
- Hribernik, K.A., et al.: Feedback mechanisms across the lifecycle for customer-driven optimisation of innovative product-service design. In: Zelm, M., Doumeingts, G., Mendonca, J.P. (eds.) *Enterprise Interoperability in the Digitized and Networked Factory of the Future (I-ESA 16 Workshop Proceedings)*. ISTE Ltd, Washington, DC (2016)
- Hribernik, K. et al.: Towards a platform for integrating product usage information into innovative product-service design. In: *2017 International Conference on Engineering, Technology and Innovation: Engineering, Technology and Innovation Management Beyond 2020: New Challenges, New Approaches, ICE/ITMC 2017 – Proceedings*, pp. 1407–1413 (2018). <https://doi.org/10.1109/ICE.2017.8280047>
- Infografik: Fast 12 Millionen Sendungen pro Zustelltag.: Statista.. Available at: <https://de.statista.com/infografik/9992/in-deutschland-von-den-paket-und-kurierdiensten-befoerderten-sendungen/> (2019). Accessed 6 July 2021
- Kritzinger, W., et al.: Digital twin in manufacturing: a categorical literature review and classification. *IFAC-PapersOnLine*. **51**(11), 1016–1022 (2018). <https://doi.org/10.1016/j.ifacol.2018.08.474>
- Kumar, A.: From mass customization to mass personalization: a strategic transformation. *Int. J. Flex. Manuf. Syst.* **19**(4), 533–547 (2007). <https://doi.org/10.1007/s10696-008-9048-6>
- Lee, E. A.: Cyber-physical systems – are computing foundations adequate? In: *NSF Workshop on Cyber-Physical Systems: Research Motivation, Techniques and Roadmap*. Austin, Texas, pp. 1–9. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.84.8011&rep=rep1&type=pdf> (2006). Accessed 18 July 2019
- Lenzerini, M.: Data integration. In *Association for Computing Machinery (ACM)*, p. 233 (2002). <https://doi.org/10.1145/543613.543644>
- Lu, Y.: Industry 4.0: a survey on technologies, applications and open research issues. *J. Ind. Inf. Integr.* **6**, 1–10 (2017). <https://doi.org/10.1016/J.JII.2017.04.005>. Elsevier
- Lu, Y., et al.: Digital twin-driven smart manufacturing: connotation, reference model, applications and research issues. *Robotics and Computer-Integrated Manufacturing*. **61**, 101837 (2020). <https://doi.org/10.1016/J.RCIM.2019.101837>. Pergamon
- Mascardi, V., Cordi, V., Rosso, P.: A comparison of upper ontologies. In: *WOA 2007 – 8th AI*IA/TABOO Joint Workshop ‘From Objects to Agents’: Agents and Industry: Technological Applications of Software Agents*, pp. 55–64. Available at: https://www.researchgate.net/profile/Viviana-Mascardi/publication/228674060_A_Comparison_of_Upper_Ontologies_Technical_Report_DISI-TR-06-21/links/0fcfd512d2286ae6de000000/A-Comparison-of-Upper-Ontologies-Technical-Report-DISI-TR-06-21.pdf (2007). Accessed 5 July 2021
- Masolo, C. et al.: The {W}onder{W}eb Library of Foundational Ontologies.. Available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.4243> (2002). Accessed 5 July 2021
- Monostori, L.: Cyber-physical production systems: roots, expectations and R&D challenges. In: *Procedia CIRP*, vol. 17, pp. 9–13. Elsevier, Amsterdam (2014). <https://doi.org/10.1016/J.PROCIR.2014.03.115>
- Muller, H.A.: The rise of intelligent cyber-physical systems. *Computer*. **50**(12), 7–9 (2017). <https://doi.org/10.1109/MC.2017.4451221>

- Ören, T. I., Ghassem-Aghaee, N., Yilmaz, L.: An ontology-based dictionary of understanding as a basis for software agents with understanding abilities. In Agent Directed Simulation Symposium, ADS 2007 – Proceedings of the 2007 Spring Simulation Multiconference, SpringSim 2007, pp. 19–27. Available at: https://www.researchgate.net/profile/Tuncer-Oeren/publication/322399131_An_Ontology-Based_Dictionary_of_Understanding_as_a_Basis_for_Software_Agents_with_Understanding_Abilities/links/5a577c10aca2726376b65703/An-Ontology-Based-Dictionary-of-Understanding- (2007). Accessed 7 July 2021
- Pisch, F. Managing global production: theory and evidence from just-in-time supply chains Department of Economics, CEP Discussion Paper. Available at: <https://www.alexandria.unisg.ch/260088/> (2020). Accessed 7 July 2021
- Reinhart, G. et al.: Cyber-Physische Produktionssysteme – Produktivitäts- und Flexibilitätsteigerung durch die Vernetzung intelligenter Systeme in der Fabrik. *wt Werkstattstechnik online*, Jahrgang 1(H. 2):84–89 (2013)
- Scheuermann, A., Hoxha, J.: Ontologies for intelligent provision of logistics services. ICIW 2012 : The Seventh International Conference on Internet and Web Applications and Services, 2(c), 106–111. Available at: http://72.52.166.99/articles/iciw_2012_4_20_20101.pdf (2012). Accessed 5 July 2021
- Tantik, E., Anderl, R.: Potentials of the asset administration Shell of Industrie 4.0 for service-oriented business models. In: *Procedia CIRP*, pp. 363–368. Elsevier B.V., Amsterdam (2017). <https://doi.org/10.1016/j.procir.2017.03.009>
- The LEVEL-UP Project.: Available at: <http://www.levelup-project.eu/project> (2019). (Accessed 28 July 2021)
- Wache, H.: *Semantische Mediation für heterogene Informationsquellen*, KI. Akademische Verlagsgesellschaft Aka GmbH (DISKI Dissertationen zur Künstlichen Intelligenz), Berlin (2003)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

