

A Semantic Mediator for Data Integration in Autonomous Logistics Processes

K.A. Hribernik¹, C. Kramer¹, C. Hans¹ and K.-D. Thoben¹

¹ BIBA - Bremer Institut für Produktion und Logistik GmbH
Hochschulring 20
28359 Bremen

Abstract. Autonomous control in logistic systems is characterized by the ability of logistic objects to process information, to render and to execute decisions on their own. This paper investigates whether the concept of the semantic mediator is applicable to the data integration problems arising from an application scenario of autonomous control in the transport logistics sector. Initially, characteristics of autonomous logistics processes are presented, highlighting the need for decentral data storage in such a paradigm. Subsequently, approaches towards data integration are examined. An application scenario exemplifying autonomous control in the field of transport logistics is presented and analysed, on the basis of which a concept, technical architecture and prototypical implementation of a semantic mediator is developed and described. A critical appraisal of the semantic mediator in the context of autonomous logistics processes concludes the paper, along with an outlook towards ongoing and future work.

Keywords: semantic mediator, semantics, ontologies, data integration, autonomous logistics processes

1. Introduction

Autonomous control in logistic systems is characterized by the ability of logistic objects to process information, to render and to execute decisions on their own [1]. The requirements set by complex logistics systems towards the integration of data regarding the individual entities within them prove immensely challenging. In order to implement complex behaviour with regards to autonomous control, dynamism, reactivity and mobility, these entities, including objects such as cargo, transit equipment and transportation systems but also software systems such as disposition, Enterprise Resource Planning (ERP) or Warehouse Management

Systems (WMS) require the development of innovative concepts for the description of and access to data. [2]

2. Related Work

The following sections review related work relevant to data integration in autonomous logistics processes. First of all, an understanding of the term “autonomous logistics processes” is developed and the necessary properties of supporting information systems to support such processes are identified. Subsequently, data integration approaches are reviewed in order to identify one which provides adequate support for the properties identified in the first step.

2.1 Characteristics of Autonomous Logistics Processes

According to [1], “*Autonomous Control in logistic systems is characterized by the ability of logistic objects to process information, to render and to execute decisions on their own.*” They furthermore define logistics objects in this context as “material items (e.g. part, machine and conveyor) or immaterial items (e.g. production order) of a networked logistic system, which have the ability to interact with other logistic objects of the considered system.” In [3], the former are further differentiated as commodities and all types of resources and whilst constraining the immaterial logistics objects to orders. In [4], a catalogue of criteria for autonomous cooperating processes is suggested. Within this catalogue, three criteria explicitly address the “information system” layer, as illustrated in Fig. 1 below. Specifically, these criteria deal with the properties of data storage, data processing and the system’s interaction ability. The first two criteria are directly related to the problem of data integration. The properties of these criteria make apparent that the less central the data storage and processing of an information system is, the higher the level of autonomous control is. The third criterion, interaction ability, relates implicitly to the data integration problem, in that the interaction ability of the information system in question is based upon its ability to access and process data stored according to the initial two criteria.

System layer	Properties				
Information system	Data Storage	Central	Mostly central	Mostly decentral	Decentral
	Data Processing	Central	Mostly central	Mostly decentral	Decentral
	Interaction Ability	None	Data allocation	Communication	Coordination

Increasing level of autonomous control



Fig. 1: Information System Layer Criteria for Autonomous Cooperating Processes [4]

Consequently, an IT system can be said to contribute to the autonomous control of a logistics system if it provides support for information processing and decision-making on the part of logistics objects, both material and immaterial. Furthermore, the information system is required to exhibit the properties data storage, data processing and interaction ability with respect to these types of logistics objects. This paper concentrates on the first of the three properties – data storage – and investigates a semantic mediator can be applied to support a decentral approach to the storage of autonomous logistics data.

2.2 Data Integration Approaches

The diversity of systems involved in complex, autonomous logistics processes defines challenging requirements towards an approach to data integration in this field. The variety of different applications, data sources, exchange formats and transport protocols demand a high level of flexibility and scalability.

To address these challenges, the specific requirements complex logistics systems exhibit need to be considered. A number of different, traditional solutions to data integration may be taken into account, foremost amongst these tightly coupled, loosely coupled and object-oriented approaches [5]. Whilst a tightly coupled approach can quickly be dismissed on grounds of its inflexibility, loosely coupled and object-oriented approaches cannot be adopted without critical analysis.

An object oriented approach generally provides good mechanisms for avoiding integration conflicts. However, when considering this approach, one must take into account that a single canonical model is required to describe the entire data model, which clearly restricts its flexibility and scalability. Each time a new stakeholder or data source enters the logistics system, the model needs to be extended. Depending on the dynamics of the logistics system, this may or may not be a disqualifying factor with regards to this approach. As the fluctuation of data sources, stake holders and systems in a complex logistics system with any degree of autonomous control can be assumed to be high, an object-oriented approach to data integration is likely to be unsuitable.

A loosely coupled approach requires detailed knowledge of each of the heterogeneous data sources to be able to be successfully employed. With regards to complex logistics systems, further analysis is required to determine whether this is feasible or not. The possibility of requiring highly flexible, and thus possibly not always pre-determinable, context data, for example from sensor networks, may prove to be an argument against this approach.

Besides the aforementioned traditional approaches to data integration, a number of predominantly semantic approaches remain to be taken into account. Here, the main concepts constituting the architecture of such data integration systems are mediators [6]. In this approach, both syntactic and semantic descriptions of the data to be integrated are applied. The semantic mediator is capable of extracting knowledge regarding the data structures of the underlying data sources and subsequently transforming, decomposing and recomposing data requests according to that knowledge. The mediator relies on semantic descriptions of the data sources. In the case of autonomous logistics processes, this implies a wholly

semantic modeling of the relevant logistics information and data across the distributed, heterogeneous sources, for which a number of approaches, such as ontologies, may be chosen. Here, extensive research is required to determine whether such semantic descriptions of logistics data are feasible and adequate to address the requirements of autonomous logistics processes. However, the application of the semantic mediator to the problem area currently offers the most promising solution candidate — the following paragraphs attempt to sketch a technical concept to the problem on that basis and evaluate it through the appraisal of a prototypical implementation within a defined application scenario.

3. Research Approach

The aim of the research presented in this paper is to develop a concept for the integration of data in autonomous logistics processes, specifically in the field of transport logistics. The semantic mediator was identified as a promising concept based on the analysis of related work outlined above.

In order to achieve a better understanding of the requirements autonomous control in transport logistics processes poses towards semantic mediation, an application scenario was defined and analysed. The scenario selected is based on simulation experiments used to demonstrate the principles of autonomous control using multi-agent systems [7]. It illustrates the application of autonomous control for the transport of fresh fruit products through Europe, taking into account the use of sensors [8] and shelf-life prediction [9] in combination with a multi-agent based system [10] to monitor and control the relevant logistics processes.

Analysis of the scenario focused mainly on identifying the agents' data requirements in the autonomous logistics processes, and by what means that data may be communicated, taking standard data exchange formats and interfaces commonly used in the field of transport logistics into account.

The following section describes the application scenario in detail.

3.1 Application Scenario

This scenario illustrates the application's processes as a story-board exemplifying one possible simulation run. The application scenario's processes are illustrated in a BPMN (Business Process Modeling Notation) diagram in Fig. 2 below.

To begin with, an order is placed with the fruit producer in Poznań, Poland for e.g. a crate of blueberries to be delivered to a restaurant in Paris, France. The order is placed electronically as a business-to-business transaction, from the customer's ERP system to an order management agent, which identifies the goods to be transported and initializes a shipping agent to start the autonomous transport logistics processes.

The shipping agent's initial task is to identify a suitable means of transport for the delivery of the fruit it is responsible for. It does this by making use of the services of a transport brokerage agency. This agency provides an electronic market where freight forwarders can publish details about their available transport

capacities, including their current routes, possible delivery times, cooling capabilities, special equipment such as temperature sensors, and so on. The shipping agent can thus discover which means of transport come into question for the route to Paris, taking into account the requirements resulting from the type of goods to be transported. Fresh fruits require cooling and temperature monitoring and need to be delivered within a defined time frame. The shipping agent begins negotiations with the respective transport agents and chooses and books the most appropriate one. The selection is made the agent on the basis of a number of weighted criteria such as pricing, delivery time guarantees, and quality of service.

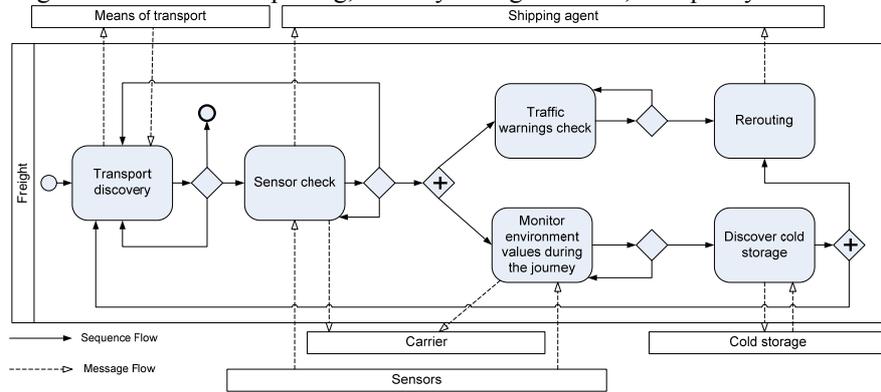


Fig. 2: BPMN-Diagram of the Application Scenario

Once the selected truck has picked the fruits up, it is ready to begin the delivery process. The shipping agent first contacts the agent representing the sensor network in the truck. It ensures the required sensors are present and operating satisfactorily. Having passed the sensor check, the shipping agent arranges for the deployment of a dedicated shelf life monitoring agent, developed and preconfigured by the food producer [11] which automatically connects to the sensor net. The shipping agent then acknowledges that everything is satisfactory and delivery may commence.

The sensor network continuously monitors the shipping conditions and after some time detects an increase in temperature in the front part of the reefer where the crates with berries are located. The monitoring agent evaluates the situation as critical. Its shelf life prediction based on the sensor readings suggest that unacceptable degradation of fruit quality will occur if the current transport plan is not changed. It contacts the shipping agent in order to find a replacement means of transport at a nearby transshipping point.

The shipping agent renegotiates transport according to the process outlined above. A new means of transport is discovered and the handover negotiated. Alternatively, if no suitable transport can be found, the shipping agent may attempt to find a temporary storage facility providing cooling and temperature monitoring facilities. The process for discovering and negotiating warehouse storage space is analogous to that for acquiring a means of transport.

The above autonomous logistics processes iterate until the goods are delivered at the customer's delivery address.

4. Findings

4.1 Scenario Analysis

On the basis of the application scenario, the data integration requirements were identified by first modeling the autonomous logistics processes in BPMN exhibited by the scenario (cf. Fig. 2), and then modeling and analyzing the communication which takes place between the individual actors.

The next step taken was to map the communication between the individual actors to messages and events represented by standard data exchange formats used in transport logistics. The selection of exchange formats was guided by a market study which identified the most prominent exchange formats used in IT systems supporting transport logistics. Consequently, business-to-business communications and agent interactions were mapped to EDIFACT (Electronic Data Interchange For Administration, Commerce and Transport) EANCOM (EAN + Communication) messages whilst information regarding the tracking and tracing of the material flow through the scenario were mapped to EPCIS (Electronic Product Code Information Service) events.

4.2 System architecture

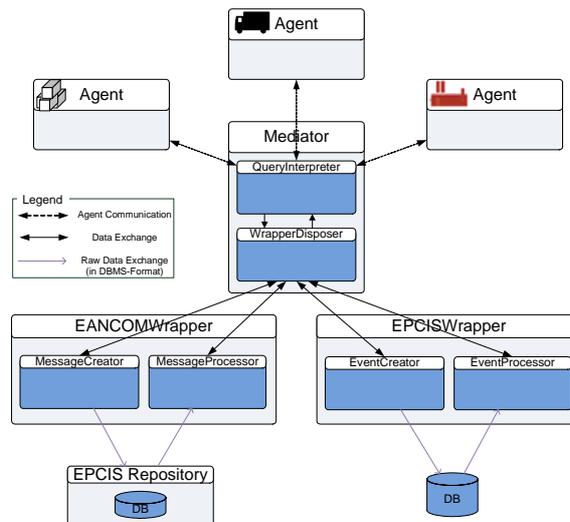


Fig. 3: Mediator Architecture Definition

The proposed system architecture illustrated in Fig. 3 follows the traditional pattern of a semantic mediator - besides the actual mediator component, which possesses a global ontology of the autonomous logistics processes of the application scenario, the wrapper components each contain extension ontologies which fully formalize the data sources they are responsible for as semantic descriptions. Heterogeneity conflicts [21] are solved either by the mediator component itself or by the respective wrapper, depending on the type of conflict. Due to the application

scenario and the selected data exchange formats only very few of the conflicts as described by [21] occur.

4.3 Semantic Data Descriptions – Ontology Specifications

The description of the application domain as well as the objects and data in this domain were implemented as a global ontology. The system, however, is designed to support the use of multiple ontologies in order to leverage modular flexibility and scalability of the mediator approach in further applications.

The Web Ontology Language (OWL-DL) [16] was used for the specification of the ontology which describes the individual data exchange formats. OWL-DL was chosen for three reasons: first of all, it is used in the multi-agent system for the description of the domain of autonomous controlled logistics. Secondly, it was judged to be adequately expressive to cover the semantic description of both the standard exchange formats used in transport logistics and the overarching concepts of autonomous logistics processes. Finally, a number of Java libraries and reasoners are readily available for OWL-DL, which was expected to significantly accelerate the development of a prototypical implementation.

The ontology used is a critical success factor of any semantic mediator. It has to reflect all the characteristics of the application domain and simultaneously has to be as simple and comprehensible as possible. Many existing ontologies may be taken into consideration for the semantic description of the entities in the given transport logistics scenario, such as those used in the fields of product lifecycle and data management as exemplified by [18], [19] or [20]. However, none of these truly reflect the syntax and semantics of autonomous logistics processes whilst encompassing the syntax and semantics of standard logistics data exchange formats. Consequently, as a first step, a new ontology was designed based on both the application scenario and the top-level ontology of the multi-agent system which describes basic concepts of autonomous logistics processes. It can be extended by incorporating additional ontologies into the system. By using the OWL-DL features “owl:equivalentClass” and “owl:equivalentProperty”, the possibility for linking the ontology to other existing ontologies is available.

4.4 Querying the Semantic Mediator

The query language “SPARQL” [12] is used as the query language at the query interface of the system. It was specifically developed for querying ontologies and thus provides an adequate basis for queries to the semantic mediator.

However, SPARQL only offers the possibility to query the system but not write to it. However, SPARQL alone doesn't fulfill all of the requirements, because agents representing autonomous objects also need to be able to create messages and data. To extend the functionality to support bidirectional queries, the “SPARQL Update” [13] language was used to extend the SPARQL query language. This allows for the editing of ontologies with a similar syntax to SPARQL. A combination of both languages was specified and prototypically implemented as the query language in of the semantic mediator.

4.5 Data Transformation in the Wrappers

The wrappers query data from the respective data sources and transform it via an internal format in order to enable the processing of data from heterogeneous data sources and formats. The transformation is carried out within the wrappers and is transparent to the actual mediator component. This allows for a complete abstraction from the data sources. Transformation in the wrappers is rule-based and described and implemented using the business rule management system “Drools” [14] (Drools - Business Logic Integration Platform). The use of Drools offers the possibility to react more quickly and flexibly to modifications to individual data sources. Should a change need to be made, only the rule files need to be updated. Modifications to the source code of the wrappers with subsequent recompilation and deployment can be avoided in this way.

4.6 System Test and Prototypical Implementation

In order to test the semantic mediator with the application scenario described above, all of the mediator components were prototypically implemented. The data sources described in the application scenario were also implemented.

A repository was set up for both EPCIS and EANCOM. The Fosstrak EPCIS repository [15] was used to query and capture the EPCIS events. A minimal EANCOM messaging implementation was developed to relay the respective messages. These repositories formed the basis for the communication. Every message or event created is stored in the corresponding repository to support the validation of the system.

5. Conclusions and Outlook

Validation of the prototypical implementation against the application scenario demonstrated the applicability of the concept of the semantic mediator in the field of autonomous logistics processes. The semantic mediator proved capable of fulfilling the bidirectional data integration requirements set by the application scenario. For this scenario, the semantic mediator supports a decentral approach to the storage of autonomous logistics data, consequently fulfilling the first criterion of the “information system” layer for autonomous cooperating processes [4] and thus contributing to the potential for autonomous control in the transport logistics sector.

From a more practical point of view, one issue is currently in the focus of current work. The extended SPARQL query interface of the semantic mediator is difficult to integrate with existing systems. A loosely-coupled approach is being investigated which aims at providing a generic, service-based interface towards the semantic mediator which can be easily accessed both by enterprise systems and autonomous logistics objects represented, for example, by software agents.

Also, it remains to be shown that the concept for a semantic mediator presented in this paper is adequate for more complex scenarios in autonomous logistics. First of all, the data expressed by the standard data exchange formats EPCIS and EDIFACT EANCOM are, by design, largely disjunct. The former deals mainly with tracking and tracing the logistics material flow whilst the latter is concerned with communicating business-to-business messages. Consequently, the potential

for heterogeneity conflicts in this scenario is quite limited. Future work will analyse whether the presented concept is suitable for more complex scenarios involving more data sources with a higher potential for conflict.

Beyond the current focus on the domain of autonomous control in transport logistics, future work will be directed towards expanding the vocabulary of the mediator into other logistics domains, such as production and reverse logistics. This implies investigating existing ontologies in these areas and how they may be integrated into the current development. Here, one promising approach may be to adopt an ontology from the domain of product lifecycle management in order to link the individual autonomous transport logistics processes into the overall product lifecycle context. A further aspect to be taken into consideration in future work is that of the integration of dynamic data sources. More demanding scenarios of autonomous control in logistics rely heavily on the availability of reliable, real-time data from, for example, sensor networks. Whether and how the presented semantic mediator can be successfully applied to integrating such data will be the focus of future research.

Acknowledgements

This research was supported by the German Research Foundation (DFG) as part of the Collaborative Research Centre 637 “Autonomous Cooperating Logistic Processes”.

References

- [1] Böse, Felix, and Katja Windt. “Catalogue of Criteria for Autonomous Control in Logistics.” In *Understanding Autonomous Cooperation and Control in Logistics - The Impact on Management, Information and Communication and Material Flow*, by Michael Hülsmann and Katja Windt. Berlin: Springer, 2007. 57-72
- [2] Hans, Carl, Karl Hribernik, and Klaus-Dieter Thoben. “An Approach for the Integration of Data within Complex Logistics Systems.” *LDIC2007 Dynamics in Logistics: First International Conference. Proceedings*. Heidelberg: Springer, 2007. 381-389.
- [3] Scholz-Reiter, Bernd, Jan Kolditz, and Thorsten Hildebrandt. “Specifying adaptive business processes within the production logistics domain – a new modelling concept and its challenges.” In *Understanding Autonomous Cooperation & Control in Logistics – The Impact on Management, Information and Communication and Material Flow*, by Martin Hülsmann and Katja Windt, 275-301. Berlin: Springer, 2007.
- [4] Windt, Katja, Felix Böse, and Thorsten Phillipp. "Criteria and Application of Autonomous Cooperating Logistic Processes." *Proceedings of the 3rd International Conference on Manufacturing Research - Advances in Manufacturing Technology and Management*. Cranfield, 2005.
- [5] Wache, H. (2003). *Semantische Mediation für heterogene Informationsquellen*. Akademische V.-G. Aka .

- [6] Ullman, J.D.: Information integration using logical views. In F.N. Afrati and P. Kolaitis, editors, Proceedings of the 6th International Conference on Database Theory (ICDT'97), 1997.
- [7] Jedermann, R., Gehrke, J. D., Becker, M., Behrens, C., Morales Kluge, E., Herzog, O., et al. (2007). Transport Scenario for the Intelligent Container. In M. Hülsmann, & K. Windt, Understanding Autonomous Cooperation & Control in Logistics. The Impact on Management, Information and Communication and Material Flow (pp. 393-404). Berlin: Springer.
- [8] Jedermann, R.; Behrens, C.; Laur, R.; Lang, W.: Intelligent containers and sensor networks, Approaches to apply autonomous cooperation on systems with limited resources. In: Hülsmann, M.; Windt, K. (eds.): Understanding Autonomous Cooperation & Control in Logistics – The Impact on Management, Information and Communication and Material Flow. Springer, Berlin, 2007, pp. 365-392
- [9] Jedermann, R.; Edmond, J.P.; Lang, W.: Shelf life prediction by intelligent RFID. In: Haasis, H.D.; Kreowski, H.J.; Scholz-Reiter, B. (eds.): Dynamics in Logistics. First International Conference, LDIC 2007, Springer, Berlin Heidelberg, 2008, pp. 231-240
- [10] Gehrke, J.D.; Ober-Blöbaum, C.: Multiagent-based Logistics Simulation with PlaSMA. In: Koschke, R.; Herzog, O.; Rödiger, K.-H.; Ronthaler, M. (eds.): Informatik 2007 - Informatik trifft Logistik, Band 1. Beiträge der 37. Jahrestagung der Gesellschaft für Informatik e.V. (GI), GI, Bonn, 2007, pp. 416-419
- [11] R. Jedermann, C. Behrens, D. Westphal, W. Lang, Applying autonomous sensor systems in logistics; Combining Sensor Networks, RFIDs and Software Agents, Sensors and Actuators A (Physical) 132 (1), 2006, pp. 370-375.
- [12] Prud'hommeaux, E., & Seaborne, A. (2008, Januar 15). SPARQL Query Language for RDF. Retrieved 07 15, 2009, from <http://www.w3.org/TR/rdf-sparql-query/>
- [13] Seaborne, A., Manjunath, G., Bizer, C., Breslin, J., Das, S., Harris, S., et al. (2008, July 15). SPARQL Update. A language for updating RDF graphs. Retrieved 07 28, 2009, from <http://www.w3.org/Submission/SPARQL-Update/>
- [14] Drools Community Documentation. Drools Introduction and General User Guide 5.0.1 Final. JBoss Enterprise, 2009. Retrieved 07.12.09 from http://downloads.jboss.com/drools/docs/5.0.1.26597.FINAL/drools-introduction/html_single/index.html
- [15] Roduner, C., Steybe, M.: Fosstrak EPCIS. ETH Zurich & University St. Gallen. Retrieved 07.12.09 from <http://www.fosstrak.org/epcis/index.html>
- [16] Smith, M. K., Welty, C., & McGuinness, D. L. (2004, Februar 10). OWL Web Ontology Language Guide. Retrieved 07 15, 2009, from <http://www.w3.org/TR/owl-guide/>
- [17] Wiederhold, G. (1992). Mediators in the Architecture of Future Information Systems. Computer , pp. 38-49.
- [18] Terzi, S. (2005). Elements of Product Lifecycle Management: Definitions, Open Issues and Reference Models. PhD – University Henri Poincaré Nancy 1 and Politecnico di Milano, May
- [19] Tursi, A. (2009). Ontology-approach for product-driven interoperability of enterprise production systems. PhD - University Henri Poincaré Nancy 1 and Politecnico die Bari, November
- [20] Jeongsoo Lee, Heekwon Chae, Cheol-Han Kim, Kwangsoo Kim. (2009). Design of product ontology architecture for collaborative enterprises. In Expert Systems with Applications, Volume 36, Issue 2, Part1, pp. 2300-2309
- [21] Wache, H., Vögele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., Hübner, S. (2001). Ontology-Based Integration of Information – A Survey of Existing Approaches