A platform architecture for Industry 4.0 driven intelligence amplification in logistics

Jean Paul Sebastian Piest
Department of Industrial Engineering and Business Information Systems
University of Twente
Enschede, the Netherlands
https://orcid.org/0000-0002-0995-6813

***Abstract*— The aim of this doctoral consortium paper is to introduce my doctoral research proposal in the field of enterprise computing. The scientific problem that I address in my research is the limited usage of real-time data, originating from Industry 4.0 (I4.0) technologies (e.g. smart IoT devices and sensors), by Small- and Medium sized Enterprises (SMEs) in the logistics industry. I argue that the development of an industry platform for real-time data streaming and analytics would allow SMEs to benefit from such data and help them streamline their operational processes and overall performance. The main contribution of my research is a reference architecture for such a platform, geared for the needs of SMEs, and incorporating: 1) a logistics canonical data model to collect and harmonize I4.0 data, 2) an automatic schema matcher to map SME data to the logistics canonical data model, 3) autonomous data mining agents, 4) an adoption strategy based on the concept of intelligence amplification and 5) key performance indicators to measure adoption effects on operational and decisional performance.**

Keywords*— *industry 4.0, enterprise architecture, data mining, intelligence amplification, small and medium sized enterprises, logistics

I. INTRODUCTION

The fourth industrial revolution, also referred to as Industry 4.0 (I4.0) or smart industry, drives recent technological developments with respect to automation and data exchange [13]. Cyber Physical Systems (CPS) connect machines, robots, assets, and other objects to the internet and make them situation aware by means of smart sensors and actuators [3]. As these types of systems are equipped with increasingly more sensors and actuators, it is normal to expect that they will generate increasingly larger datasets and real-time data streams [5]. Recent advances in machine learning, deep learning and Artificial Intelligence (AI), together with the possibilities offered by CPSs, provide the technological foundation to design intelligent and autonomous production and transportation systems. This way, I4.0 provides new opportunities for business innovation. However, the adoption of these emerging technologies is a complex process that assumes the existence of specific expertise to manage and process the data, and requires organizations to invest in cloud computing, data analytics tools, and internet of things to enable flexible automation.

The usage of technology and data in transportation and logistics is studied widely in both literature and practice. Transportation companies and logistics service providers have access to real-time data from on-board computers and GPS trackers in trucks and trailers, and from automatic identification systems of vessels and planes. Barcodes, RFID tags and sensors attached to products and their load carriers allow real-time monitoring of the position, movement and condition of cargo. Open data sources can further enrich the available data with real-time traffic information, weather conditions, and network master data. Despite the availability of so much data, a recent survey among 260 organizations in the Dutch logistics industry [11] indicates a limited usage of real-time data and a lack of adoption of I4.0 technologies, in particular by SMEs. The results of the survey show that 23% of the respondents uses real-time data for decision making, 44% uses historical data to make decisions and 19% uses data to explain the past. 67% of the shippers reported in the survey that they are exploring digital platforms, but the majority are not yet looking into developments, such as big data, cloud computing, internet of things and AI.

My doctoral research project is part of the multi-disciplinary research project "Autonomous Logistics Miners for Small and Medium-sized Businesses". On a national level, this research project aims to contribute to the Dutch logistics industry research agenda by improving the usage of data analytics, data mining and the adoption of I4.0 technologies by SMEs. The proposed approach of the research project is to is to increase the competitive power of the Dutch logistics industry by developing intelligent data mining agents, that can autonomously perform the most common data mining functions and require minimal supervision and IT knowledge from the user. More specifically, my aim is to increase the use of real-time data from I4.0 technologies to support decision making in operational logistical business processes in SMEs. Thus, my research helps small- and medium sized businesses that are overwhelmed with data and have limited time, knowledge and resources to analyze it. My focus is on providing insight in the overall performance of supply chains, trends in supply and demand patterns and identifying the critical factors that cause shipment delays and disruptions. Furthermore, I intend to create a symbiotic interaction between humans and intelligent agents, allowing humans to identify conditions of special interest to them and enable the intelligent agents to do the routine work, continuously monitor data streams and raise awareness to the human operator if events occur. This proposal has been inspired by the concept of intelligence amplification (IA).

2325-6605/19/$31.00 ©2019 IEEE
DOI 10.1109/EDOCW.2019.00038

Prior IA research has put forward the idea of using human-machine symbiosis to achieve an enhancement of analytic and decision-making capabilities. To this end IA essentially differs from AI, as in IA the argument is that humans and machines, can re-enforce each other's abilities, whereas AI aims at replacing the human with artificial software systems. IA also states that analytic capabilities equals the sum of human capabilities, of machine capabilities, and of capabilities that are the result of collaboration between human and machine. The IA framework proposed in [10] takes as starting point a decision-making process, decomposes its task hierarchy and assigns tasks to the human, machine or human and machine. Based on the task decomposition, intelligent agents can be designed and implemented to machine-tasks assigned to them. Experiments show promising results as far as the decision-making performance of the hybrid solution is concerned, but IA needs further validation, which I am planning to carry out in the logistics industry in the operational context of SMEs. Figure 1 visualizes the high-level concept of the research and I4.0 IA platform architecture.

This first part introduced the context, background, motivation and aim of the research. The remainder of this paper is organized as follows. The second part highlights the main findings from the preliminary literature study. The third part covers the research questions, the research design and methodology. The fourth part gives an overview of the contributions and gives some pointers to related work. The fifth part concludes with the advices I seek at the doctoral consortium.

II. Preliminary Literature Study

Based on a preliminary systematic literature study, several articles were reviewed that address the adoption problem in SMEs and propose some research directions and possible solutions. Next, I discuss the main findings.

A. *Data analytics and -mining in SMEs*

Data analytics and -mining is an established research field that contain numerous methods and techniques. Data mining is aimed at extraction of patterns and relations from large data sets [2]. Common data mining techniques include classification, estimation, segmentation, forecasting, association, and text analysis [2]. Typical use cases of data mining include market basket analysis, customer segmentation, demand forecasting, maintenance prediction and fraud detection. E-commerce and other kinds of digital platforms, together with web mining methods, can be applied for growth prediction models and contribute to the performance of SMEs [21].

The usage of data analytics and data mining in SMEs is discussed by several researchers. The main methods that are applied in SMEs in transportation are "Knowledge Discovery in Data (KDD)", "Sample, Explore, Modify, Model and Assess (SEMMA)" and the "CRoss-Industry Standard Process for Data Mining (CRISP-DM)" [17]. The main limitations of KDD, SEMMA and CRISP-DM for SMEs are the required technical knowledge and the extensiveness of the method. SMEs typically lack the resources and knowledge for data mining and have concerns about data management [17]. [17] classified 10 barriers for SMEs to adopt big data analytics and argues that data mining is critical for SMEs to improve decision making and performance. [9] listed 14, mostly similar, challenges SMEs face concerning big data analytics and presents a maturity model and growth path for SMEs.

The technological requirements and barriers for SMEs are addressed by evaluating open source systems [1,7,16] for SMEs and comparison with regression methods [8]. The knowledge and organizational issues of SMEs are discussed by [15] and a collective approach is proposed for technology watch and competitive intelligence through intermediate R&D centers. This competitive intelligence approach has been implemented by [19] in a web-based decision support system for SMEs based on agent mining. A collective approach is considered relevant for this research project.

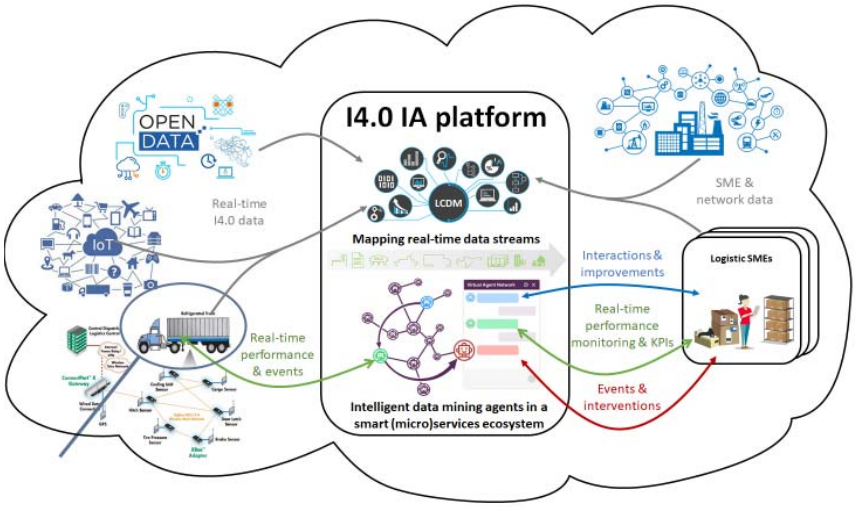

Figure 1: high level concept of the I4.0 IA platform architecture (own drawing)

B. Technology adoption and enterprise architecture in SMEs

Technology adoption and enterprise architecture are widely studied in research and practice. Different methods and techniques exist to model and manage enterprise architecture, e.g. Zachman's framework, TOGAF and ArchiMate. The age and size of the company are important factors related to adoption challenges [12]. A survey among SMEs in Flanders suggests that SMEs do not know nor apply enterprise architecture and argues that SMEs require a different, more simplified, enterprise architecture compared to large organizations [4]. [4] described six characteristics that influence IT adoption by SMEs and created a starting point for enterprise architecture in SMEs based on design science. The four dimensions of the business architecture incorporate a strategic dimension (why), active actor dimension (who), operation dimension (how) and object dimension (what) and are accompanied by supporting criteria (CHOOSE) and software tool support [14]. The approach is also mapped to ArchiMate and is considered relevant for this research project.

III. RESEARCH QUESTIONS, -DESIGN AND -METHODOLOGY

This part provides a synopsis of the research questions, design and methodology of my doctoral research project.

A. Research questions

Based on the context, background and aim, the central research question that I aim to answer is: *How can SMEs in the Dutch logistics industry benefit real-time I4.0 data and utilize autonomous data mining agents in combination with the concept of IA to increase their operational and decisional performance, and overall competitiveness?*

The following research questions are derived from the central research question and aim of the research:

1. Which real-time I4.0 data sources, canonical data models and interoperability standards are available in the logistics industry?
2. Which schema matchers are capable of automatic mapping and harmonization of real-time I4.0 data models with internal data models of SMEs?
3. What are typical usage scenarios in which autonomous data mining agents can support SMEs to increase their operational and decisional performance?
4. What are the requirements for the I4.0 IA platform, the autonomous data mining agents and the level of IT knowledge and resources in SMEs?
5. Which adoption strategies, together with the concept of IA, increase the usage of real-time I4.0 data for data analytics and -mining in SMEs?
6. Which KPIs measure the effects of the I4.0 IA platform on the operational and decisional performance and overall competitiveness of SMEs?

B. Research design and methodology

My doctoral research project consists of an empirical part and an engineering part. Figure 2 visualizes the research design of my doctoral research project and its phases. The empirical part (phase a-c) is aimed at theory development and investigates real-time I4.0 data usage in relation to the performance of SMEs in the logistics industry. The engineering part (phase d) is aimed at the design, development and validation of the I.4.0 IA platform in the operational context of SMEs in the Dutch logistics industry. Both the empirical view and the engineering view are applied to find answers for the research questions.

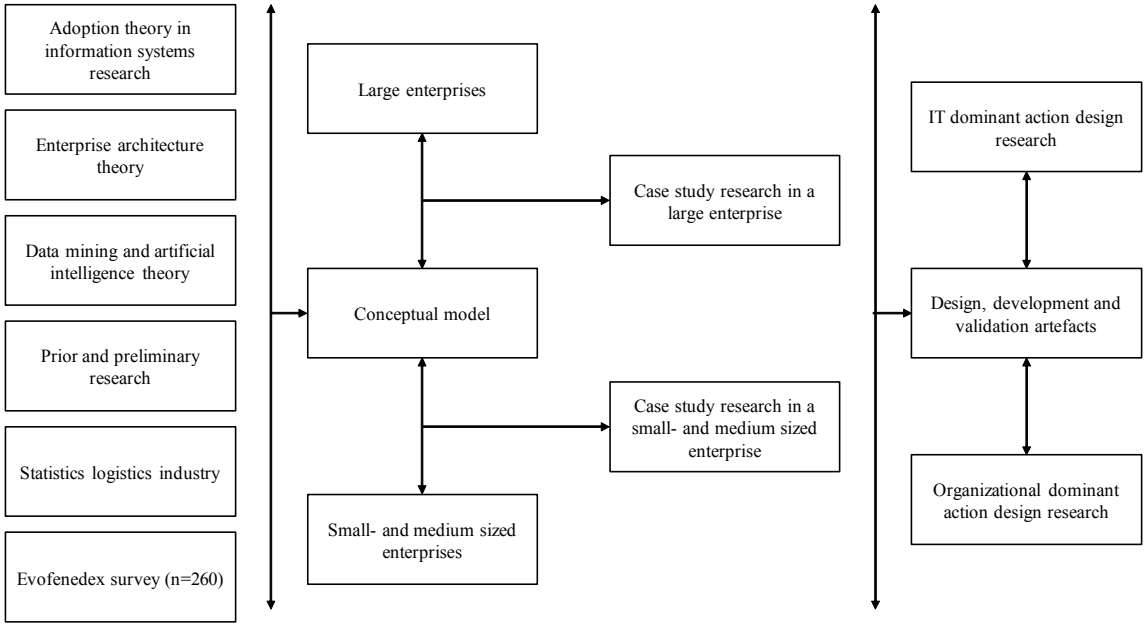

Figure 2: proposed research design and phases, adapted from [22]

The empirical part is based on theories from relevant disciplines, results and findings from prior and preliminary research and available statistics (a). A conceptual model is developed to identify the antecedents of the usage of real-time I4.0 data for data analytics and -mining and the operational and decisional performance (b). The conceptual model will be analyzed from the perspective of SMEs and contrasted with the situation of large enterprises (b). Together with the consortium partners of the research project, which includes small-, medium- and large enterprises from the logistics industry, the conceptual model will be tested and validated using case study research (c).

The engineering part builds upon the results of the empirical part and will result in several artefacts (described in the next section). The research methodology to be followed during this process is the well-established design science methodology for information systems research [18, 23] (d). Action design research [20] is applied as it emphasizes on an iterative process that combines observations with interventions from both an organizational and technological perspective (d). The combination of design science with action design research is expected to improve the outcomes of this research project, because it facilitates an incremental development process. Multiple iterations and interventions will contribute to the improvement of the designed artefacts.

IV. Proposed contributions and related work

This section gives an overview of the proposed contributions, research directions and related work.

A. Proposed contributions

Based on the research design and methodology, the main contribution of my doctoral research project will be the reference architecture that incorporates real-time I4.0 data streaming and analytics in autonomous data mining agents for SMEs in the logistics industry. This research contributes to the development of new knowledge about real-time I4.0 data collection, harmonization and automatic schema matching in the context of SMEs in the logistics industry. The research will build upon IA prior research and contribute to the further development and validation of the IA framework in the context of SMEs in logistics. Figure 3 visualizes the contributions of the research project and artefacts for SMEs.

1) Logistics canonical data model (LCDM)

The starting point of the research project is to collect and harmonize available I4.0 data models in a LCDM for SMEs. My focus is on real-time data from open data sources, IoT devices and logistical assets. My intention is not to develop new interoperability standards, but rather focus on achieving interoperability between existing standards. This first step is linked to research question 1 and will result in a library of I4.0 data models, the LCDM and supporting resources for SMEs.

2) Schema mapper

The second step is to map SME specific data models to the LCDM using a schema mapper. My intention is not to develop a new schema matcher, but rather evaluate existing schema mappers on their capabilities to process heterogeneous data formats and automatically map both structured and unstructured data formats. This second step is related to research question 2 and will result in a comparison of the capabilities and performance of multiple schema matchers.

3) Reference architecture

The third step is to design and develop autonomous data mining agents that can perform common data mining tasks. I aim to design a reference architecture that incorporates the developed autonomous data mining agents and is tailored to the enterprise architecture of SMEs and their data usage requirements. The ArchiMate language and specification will be applied to model the reference architecture of the I4.0 IA platform and its components. This third step is related to research questions 3 and 4 and results in several architecture models, SME viewpoints and prototypical instantiations of the autonomous data mining agents and I4.0 IA platform.

4) Adoption strategy

The fourth step is to develop an adoption strategy, that could be entirely or partly supported by the I4.0 IA platform, for SMEs in the logistics industry. The ArchiMate language and specification will be applied together with the business process modeling notation. This fourth step is related to research question 5 and will extend the architecture models with strategy and implementation viewpoints for SMEs and supporting process descriptions.

5) KPIs

The fifth and last step is to evaluate the adoption effects on the operational and decisional performance of SMEs by

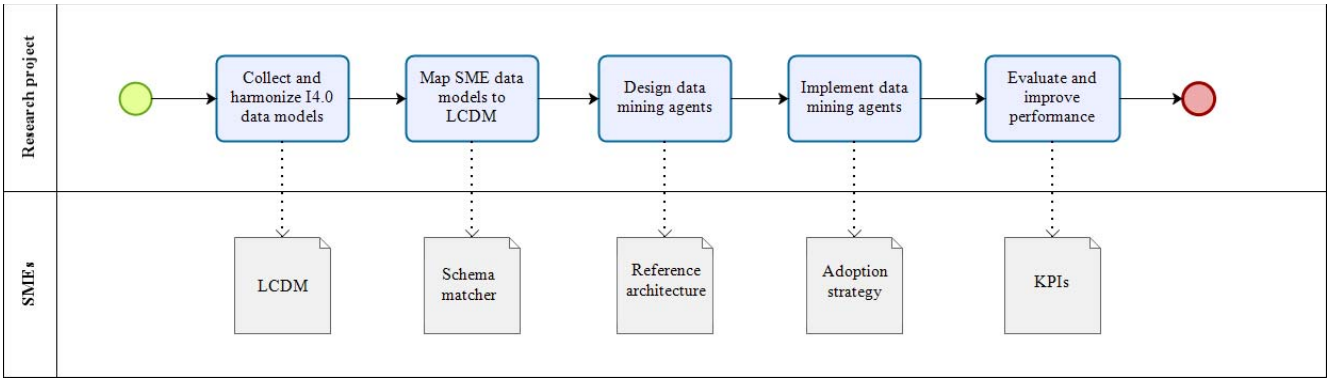

Figure 3: overview of contributions and artefacts

means of several KPIs, such as efficiency, costs, service levels, customer satisfaction, CO_2 emissions, etc. This fifth step is related to research question 6 and will result in a set of KPIs that together can measure, evaluate and improve the effects of the I4.0 IA platform and adoption strategy in relation to the operational and decisional performance of SMEs.

B. Related earlier work

My research continues the work done in the research project Synchromodal IT. This previous work illustrates a typical usage scenario in the logistics industry, in which automated retrieval and processing of multiple streaming open data sources improves the operational performance. The developed enterprise architecture models, prototype and common data model described in [6] serve as a starting point for the development of the LCDM, schema matcher and reference architecture.

V. ADVICE SOUGHT

Participating in this doctoral consortium offers me a unique opportunity to get introduced to the scientific community active in the field of enterprise computing. Gaining expert advice in this early phase of my doctoral research enables me to better position my work and helps me reach a more precise definition of my research questions. I am seeking advice concerning the development of hypotheses and research directions to develop an integrated approach for the Dutch logistics industry.

ACKNOWLEDGMENT

My doctoral research project is financially supported by the Dutch Ministry of Economic Affairs and co-financed via the Top consortia Knowledge and Innovation of the Dutch Institute for National Advanced Logistics. I would like to thank my promotors Prof. Dr. M.E. Iacob and Prof. Dr. J. van Hillegersberg and the involved consortium partners for their support and contributions to the research project.

REFERENCES

[1] Almeida, P., & Bernardino, J. (2016). A survey on open source Data Mining Tools for SMEs. In New Advances in Information Systems and Technologies (pp. 253-262). Springer, Cham.

[2] Arnth-Jensen, N. (2006). Applied Data Mining for Business Intelligence (Master's thesis, Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby, Denmark).

[3] Baum, G., Borcherding, H., Broy, M., Eigner, M., Huber, A., Kohler, H., Stümpfle, M. (2013). Industrie 4.0 - Beherrschung der industriellen Komplexität mit SysLM. Verlag Berlin Heidelberg: Springer.

[4] Bernaert, M., Poels, G., Snoeck, M., & Backer, M. De. (2000). Enterprise Architecture for Small and Medium- Sized Enterprises : A Starting Point for Bringing EA to SMEs , Based on Adoption Models, 1–33.

[5] Brühl, V. (2015). Wirtschaft des 21. Jahrhunderts - Herausforderungen in der Hightech-Ökonomie. Wiesbaden: Springer Fachmedien.

[6] Bol Raap W., Iacob ME., van Sinderen M., Piest S. (2016) An Architecture and Common Data Model for Open Data-Based Cargo-Tracking in Synchromodal Logistics. In: Debruyne C. et al. (eds) On the Move to Meaningful Internet Systems: OTM 2016 Conferences. OTM 2016. Lecture Notes in Computer Science, vol 10033. Springer, Cham

[7] Chen, X., Ye, Y., Williams, G., & Xu, X. (2007, May). A survey of open source data mining systems. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 3-14). Springer, Berlin, Heidelberg.

[8] Chertchom, P. (2018, May). A comparison study between data mining tools over regression methods: Recommendation for SMEs. In 2018 5th International Conference on Business and Industrial Research (ICBIR) (pp. 46-50). IEEE.

[9] Coleman, S., Göb, R., Manco, G., Pievatolo, A., Tort Martorell, X., & Reis, M. S. (2016). How can SMEs benefit from big data? Challenges and a path forward. Quality and Reliability Engineering International, 32(6), 2151-2164.

[10] Dobrkovic, A., Liu, L., Iacob, M. E., & van Hillegersberg, J. (2016, November). Intelligence amplification framework for enhancing scheduling processes. In Ibero-American Conference on Artificial Intelligence (pp. 89-100). Springer, Cham.

[11] Evofenedex (2018). Onderzoeksrapport data en digitalisering. Source: https://www.ttm.nl/wp-content/uploads/2018/11/Onderzoeksrapport-Data-en-digitalisering-in-de-logistiek-2018.pdf

[12] Grube Hansen, D., Malik, A. A., & Bilberg, A. (2017). Generic Challenges and Automation Solutions in Manufacturing SMEs, 1161–1169.

[13] Halang, W. A., & Unger, H. (2014). Industrie 4.0 und Echtzeit. Echtzeit 2014.

[14] Ingelbeen, D., & Bernaert, M. (2013). Enterprise architecture software tool support for small and medium-sized enterprises: EASE.

[15] Izquierdo, J., & Larreina, S. (2005). Collective SME approach to technology watch and competitive intelligence: the role of intermediate centers. In Knowledge Mining (pp. 181-189). Springer, Berlin, Heidelberg.

[16] Kalan, R. S., & Ünalir, M. O. (2016, October). Leveraging big data technology for small and medium-sized enterprises (SMEs). In 2016 6th International Conference on Computer and Knowledge Engineering (ICCKE) (pp. 1-6). IEEE.

[17] Mohd Selamat, S. A., Prakoonwit, S., Sahandi, R., Khan, W., & Ramachandran, M. (2018). Big data analytics—A review of data mining models for small and medium enterprises in the transportation sector. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(3), e1238.

[18] Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of management information systems, 24(3), 45-77.

[19] Ponis, S. T., & Christou, I. T. (2013). Competitive intelligence for SMEs: a web-based decision support system. International Journal of Business Information Systems, 12(3), 243-258.

[20] Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. Management Information Systems Quarterly, 35(1), 37-56.

[21] Te, Y. F., & Cvijikj, I. P. (2017, June). Design of a small and medium enterprise growth prediction model based on web mining. In International Conference on Web Engineering (pp. 600-607). Springer, Cham.

[22] Verschuren, P. J. M., & Doorewaard, H. (2007). Het ontwerpen van een onderzoek. Lemma.

[23] Wieringa, R. J. (2014). Design science methodology for information systems and software engineering. Springer.