A Novel Ontology-Based Smart Service Architecture for Data-Driven Model Development

Gernot Steindl, Bernhard Heinzl, Wolfgang Kastner

Institute of Computer Engineering
Research Unit Automation Systems
TU Wien, Vienna, Austria
Email: gernot.steindl@tuwien.ac.at

***Abstract*—In the proposed concept, an ontology-based information model is designed to support the data-driven model development process of industrial energy systems. Smart Services for data-driven model development are conceptually deployed in the context of the Reference Architecture Model for Industry 4.0 (RAMI4.0) residing inside the information layer. The ontology-based Smart Service architecture provides interoperability by means of data integration and data exchange between for various applications, part of the functional layer. Due to its modular structure, these Smart Services and their functional enhancements can easily be extended and used by a variety of other applications at different layers of the RAMI4.0.**

Keywords*–*data-driven model development; ontology; Smart Service; industrial energy systems; Industry 4.0; RAMI4.0

I. INTRODUCTION

Currently, a number of research projects and development efforts address the further digitalization in the industrial sector [1], which is often called the fourth industrial revolution or Industry 4.0. Besides digitalization, the main goals of Industry 4.0 are optimization and customization of production as well as enhanced automation and adaption [2]. In this context, Industry 4.0 can also facilitate the sustainable energy transition [3]. Technology of Industry 4.0 is a forerunner to let the plant's energy demand become more controllable and providing means for flexibility, which helps to integrate a higher share of volatile renewable energy sources into the energy system. To utilize this flexibility of industry processes, model-based and predictive control strategies are suitable to handle the conflicts between different objectives and constraints [4]. This allows a factory to react, for example, to a demand response request, while meeting the production constraints. Accurate process models are obligatory for this kind of applications. Developing such process models, however, is the most challenging and time-consuming task in an implementation process [4] [5]. One reason for this is the cumbersome manual analysis of data from heterogeneous sources, which often have insufficient data quality, are only partly available [6] or stored in different formats, encodings and databases. Data-driven model development also relies on implicit domain knowledge [7], which is often scattered and not clearly communicated by the domain experts or hard-coded into various applications. In order to reduce the model development effort and facilitate knowledge reuse, systematic data and information management is a vital part in oder to increase data quality and consistency, as well as make it available over the whole plant life-cycle.

The goal of this work is to develop a framework for systematic support of data-driven model development in the context of industrial energy systems. Comprehensive information management is achieved explicitly by stating domain knowledge and making it available for different services in a formal ontology-based information model. This ontology will act as a semantic abstraction layer, to facilitate integration of different kinds of data.

In Section 2, the data-driven model development process will be discussed. It will be stated how information, which is stored in a formal ontology, could support the individual process steps. A service architecture will be introduced in Section 3, that facilitates the reuse of stored knowledge in the context of Industry 4.0.

II. DATA-DRIVEN MODEL DEVELOPMENT PROCESS

Data-driven models are usually identified during the operational phase of the plant using process data and are developed for a certain purpose or application, like plant control, monitoring, optimization, prediction or performance assessments. The main steps in the data-driven model development process for industrial processes are described in [8].

The first step is to decompose the process and its data based on process knowledge, like location of the components and the plant topology. If this knowledge is available in a machine-readable form or formal ontology, this step can be automated.

Afterwards, pre-processing of the data has to be performed, like removing outliers, normalizing the data, filling or estimating missing data and synchronizing the sampling rates of different process variables. Data analysis methods can be used in combination with domain knowledge from the proposed ontology to analyze and semantically enrich the time series data, e.g. to add reliability metrics to the data points. Also, inconsistencies like falsely labeled data or outliers can be identified automatically, if semantic information of a certain data point is available.

The subsequent step of sample and variable selection is very crucial as different operating conditions can occur in the monitoring data. The samples should be chosen with regard to the scope of application, in order to be able to train an accurate model. Therefore, domain knowledge has to be applied which should also be made explicitly available in the ontology to support this step.

 ISBN: 978-1-61208-690-3

The final model selection and identification can be performed in an iterative way by the control engineer. Design decisions, used training and evaluation data, as well as model performance are also important information, which should be stored in an ontology, to enable comparison and reuse of different component models.

III. Ontology-based Smart Service Architecture

For Industry 4.0, an architectural reference model (RAMI4.0) is available, which combines all elements and IT components in a layered model, taking into account life-cycle aspects [9]. In the context of RAMI4.0, a service-oriented architecture for ”Smart Services” was proposed in [10]. We build on this idea and apply it to support the data-driven model development process by offering services, which can be located on different layers and hierarchies in RAMI4.0, in particular on the Information and Functional Layer.

Figure 1 shows the architecture of the proposed services at different layers of RAMI4.0. The data is available in various databases an formats. An ontology-based information model is used as a foundation for interoperability and data integration. Engineering data as well as operational data are integrated with the help of this ontology. The information model is one of the main components, which consists of smaller, interlinked ontologies. These are specialized for certain services or describe general concepts, like plant topology, plant components, etc. Existing information models, like the P&ID specification of the DEXPI initiative [11] should be reused and adapted. These ontologies manage data access and data exchange between different services. Process data, which typically occurs as time series, will stay in their databases, as these are optimized for this kind of purpose, but are accessible through the ontology to facilitate data integration. Thus, the ontologies can be connected to a ”Linked Factory Data” repository and new information can easily be added into the modular ontology. The available services and their capabilities can also be stated in the ontology itself, to enhance interoperability.

The base service for data-driven model development is the ”Smart Data” service, which is able to automatically analyze, pre-process and validate the available data, utilizing information from the ontology. For data-driven model development, a ”Model Identification” service will be established, which uses domain knowledge from the ontology to find related input data to a certain model component and to support the model developer with information of the plant’s operating conditions and perform proper sample selection. Information on the identified model will be stored in the ontology as well, to enable model comparison and reuse. The ”Linked Data Management” service is the central point for entering and maintaining the ontology-based information model itself.

Based on the proposed architecture, additional services and applications on the different layers of RAMI4.0, like ”Optimized Control” or ”Fault Detection”, can be implemented, which are also able to use and contribute to the information stored in the ”Linked Factory Data”.

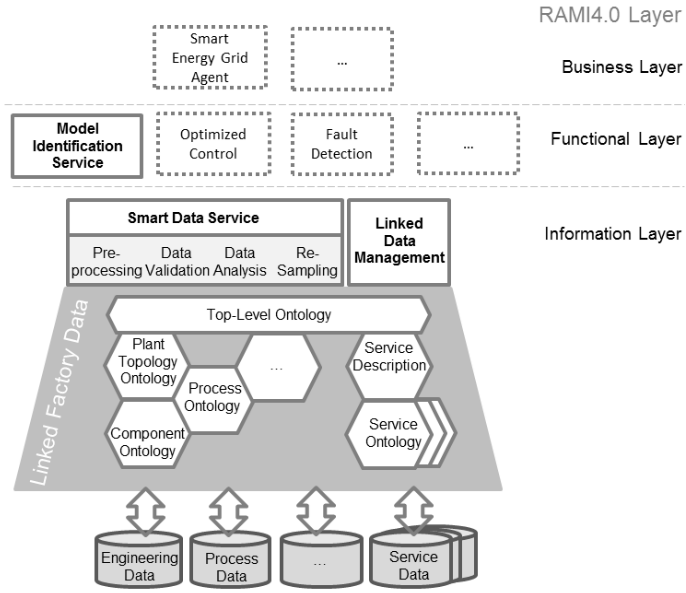

Figure 1. Ontology-based Smart Service Architecture

IV. Conclusion

The goal of this work is to develop a systematic support for a data-driven model development process in the context of industrial energy systems. The proposed ontology-based architecture can help to reduce the manual and error-prone modeling effort and increase the quality, consistency and expressiveness of the available data. The next steps will be to specify the information which has to be included into the ontology and evaluate different methods for ontology-based data access. Afterwards, a real use case of an industrial production plant will be used to implement a ”Smart Data” and ”Model Identification” service as proof-of-concept.

References

[1] Y. Lu, “Industry 4.0: A survey on technologies, applications and open research issues,” Journal of Industrial Information Integration, vol. 6, 2017, pp. 1–10.

[2] V. Roblek, M. Meško, and A. Krapež, “A Complex View of Industry 4.0,” SAGE Open, vol. 6, apr 2016.

[3] UNIDO, “Accelerating clean energy through Industry 4.0: Manufacturing the next Revolution,” United Nations Industrial Development Organization, Vienna, Austria, Tech. Rep., 2017.

[4] M. Killian and M. Kozek, “Ten questions concerning model predictive control for energy efficient buildings,” Building and Environment, vol. 105, 2016, pp. 403–412.

[5] M. Morari and J. H. Lee, “Model predictive control: past, present and future,” Computers & Chemical Engineering, vol. 23, may 1999, pp. 667–682.

[6] Y. Zong, S. You, J. Hu, X. Han, C. Jiang, Y. Zhang, and G. M. Böning, “Challenges of using model predictive control for active demand side management,” 4th International Conference on Microgeneration and Related Technologies, 2015.

[7] G. A. Silver, O. Al-Haj Hassan, and J. A. Miller, “From domain ontologies to modeling ontologies to executable simulation models,” Proceedings - Winter Simulation Conference, 2007, pp. 1108–1117.

[8] Z. Ge, “Review on data-driven modeling and monitoring for plant-wide industrial processes,” Chemometrics and Intelligent Laboratory Systems, vol. 171, 2017, pp. 16–25.

[9] P. Adolphs, H. Bedenbender, D. Dirzus, and E. Martin, “Reference Architecture Model Industrie 4.0 (RAMI4.0),” VDI/VDE-Gesellschaft, Tech. Rep. July, 2015.

[10] G. Koschnick, “Industrie 4.0: Smart services,” German Electrical and Electronic Manufacturers’ Association, Frankfurt am Main, Germany, Tech. Rep. July, 2016.

[11] M. Theißen and M. Wiedau, “DEXPI Data Exchange in Process Industry P&D Specification,” DECHEMA, Tech. Rep., 2016.

 ISBN: 978-1-61208-690-3