Development of a Semantic Database Model to Facilitate Data Analytics in Battery Cell Manufacturing

Ozan Yesilyurt[1][a], David Brandt[1][b], Julian Joël Grimm[1][c], Kamal Husseini[2][d], Aleksandra Naumann[3,4][e], Julia Meiners[3,4][f] and David Becker-Koch[5][g]

[1]Fraunhofer Institute for Manufacturing Engineering and Automation IPA Nobelstraße 12, 70569 Stuttgart, Germany

[2]Karlsruhe Institute of Technology, Kaiserstraße 12, 76131 Karlsruhe, Germany

[3]Technische Universität Braunschweig, Institute of Machine Tools and Production Technology, Langer Kamp 19b, 38106 Braunschweig, Germany

[4]Technische Universität Braunschweig, Battery LabFactory Braunschweig, Langer Kamp 8, 38106 Braunschweig, Germany

[5]The Centre for Solar Energy and Hydrogen Research Baden-Württemberg, Lise-Meitner-Straße 24, 89081 Ulm, Germany

Keywords: Battery Cell Manufacturing, Database Model, Semantic Model Description.

Abstract: The global demand for batteries is increasing worldwide. To cover this high battery demand, optimizing manufacturing productivity and improving the quality of battery cells are necessary. Digitalization promises to offer great potential to address these challenges. Through data collection along the manufacturing processes, hidden correlations can be identified. However, data is highly diverse in battery cell manufacturing, complicating data analysis. A semantic data storage can increase the understanding of the relationships between the datasets, facilitating the identification of the causes of defects in manufacturing processes. To structure heterogeneous data in a semantically understandable and analyzable form, this paper presents the development of a semantic database model. The realization of this model enables structuring various datasets for simplified access and usage for increasing productivity and battery cell quality in battery cell manufacturing.

1 INTRODUCTION

The global demand for batteries for energy storage is growing due to the continued development of electric vehicles and other mobile devices (Asif & Singh, 2017). The growing number of battery-electric vehicles registered illustrates the increasing global demand for battery cells (Carlier, 2021). To meet this high demand of the battery cell users, digitalization of production offers several opportunities to build a flexible, intelligent, adaptable, and efficient manufacturing system (Zhong, Xu, Klotz, & Newman, 2017). Data is the key element to realizing such manufacturing systems, and its availability has been rapidly growing in the manufacturing industry (Yin & Kaynak, 2015). Smart manufacturing targets transforming data towards manufacturing intelligence to positively affect every manufacturing-related aspect (O'Donovan, Leahy, Bruton, & O'Sullivan, 2015).

[a] https://orcid.org/0000-0003-3002-3230
[b] https://orcid.org/0000-0003-1781-9541
[c] https://orcid.org/0000-0002-0559-3752
[d] https://orcid.org/0000-0002-7110-4697
[e] https://orcid.org/0000-0002-9193-1294
[f] https://orcid.org/0000-0002-6564-6537
[g] https://orcid.org/0000-0001-5921-5768

Yesilyurt, O., Brandt, D., Grimm, J., Husseini, K., Naumann, A., Meiners, J. and Becker-Koch, D.
Development of a Semantic Database Model to Facilitate Data Analytics in Battery Cell Manufacturing.
DOI: 10.5220/0011139500003269
In *Proceedings of the 11th International Conference on Data Science, Technology and Applications (DATA 2022)*, pages 13-20
ISBN: 978-989-758-583-8; ISSN: 2184-285X

The use of data-driven manufacturing technologies is appealing for battery cell manufacturing to improve scrap and product quality. Data in battery cell manufacturing is heterogeneous, resulting from both converging and diverging material flows, continuous and non-continuous processes, and single and batch processes (Turetskyy et al., 2020). On the one hand, it is essential to build data storage in battery cell manufacturing to solve heterogeneous data problems. It is challenging to collect unstructured data from different manufacturing processes and perform informative analyses. On the other hand, it is a major gap in battery cell manufacturing to create and enable context-based data in a semantically understandable way because heterogeneous data consists of different unrelated datasets, which hinders identifying the cause of manufacturing process issues and discovering hidden optimization opportunities in the manufacturing process.

A possible solution to address these challenges is creating a semantic database with a model that facilitates identifying, accessing and processing the appropriate data. Semantically structured data enables data scientists to track and evaluate the manufacturing processes and find optimization potentials in the battery cell manufacturing with various analysis tools.

This paper aims to present the requirements for a specific battery manufacturing scenario and present a possible implementation of such a semantic database model as a solution.

The following chapter describes the data storage transformation. It also presents current data challenges, introduces developed approaches to solve them, and derives the scientific gap. Chapter 3 introduces the specific scenario for which the semantic database model is created. The following two chapters illustrate the model's underlying concept and its realization. The paper concludes with a summary and highlights the future work with the developed semantic database model.

2 RELATED WORK

Data is a key element for smart manufacturing to meet manufacturing needs and inform manufacturing decision-making areas (O'Donovan et al., 2015). Relational databases have been used for decades as regular database systems to store manufacturing data. A relational database is a digital database used for electronic data management in computer systems and is based on a table-based relational database model as proposed by (Codd, 1970). With the introduction of IoT technologies, cloud computing, big data analytics, and AI integrated into manufacturing systems, a high level of multi-source and heterogeneous data is generated (Tao, Qi, Liu, & Kusiak, 2018). Therefore, more and more new database technologies were integrated into the existing data storage architectures, such as NoSQL. With these, the challenges associated with storing the great amount of manufacturing data could be adressed. NoSQL databases became widely used around 2009, which process data faster than relational databases because their data models are built more simply (Leavitt, 2010). They can be categorized into five groups (Column-based, document-based, key-value-based, graph-based, time-series-based) (Cui, Kara, & Chan, 2020; Yen, Zhang, Bastani, & Zhang, 2017). The same big data challenge is also observed in battery cell manufacturing. Various systems (e.g. equipment, controls and simulation models) are involved and generate heterogeneous (e.g. time series, discrete) data in large volumes. Data is distributed in several heterogeneous datasets and databases that need to be linked to enable in-depth analysis for identifying and addressing issues in battery cell manufacturing processes.

Some solution approaches are developed to address this need. A hybrid framework for industrial data storage to utilize zero-defect manufacturing was introduced by (Grevenitis et al., 2019), where unstructured data generated by the IoT devices are processed in a NoSQL database (Apache Cassandra); at the same, time the structured data is stored in a SQL database (MySQL). Then, the filtered data from both databases is converted into knowledge and stored in a triplestore database to be used by experts. Furthermore, (Hildebrand, Tourkogiorgis, Psarommatis, Arena, & Kiritsis, 2019) developed a generic algorithm for the automated conversion of different data types into RDF. With this solution, the researchers seek to enable data mapping without hard coding. This solution is reusable across various data schemas and ontologies, which can be easily modified to fit other data formats. In addition, (Wessel, Turetskyy, Wojahn, Abraham, & Herrmann, 2021) presented and implemented a methodology to develop an ontology-based traceability system in battery cell manufacturing so that the relations between the data sources along the manufacturing chain can be determined. Moreover, (Grimmel, Wessel, Mennenga, & Herrmann, 2022) introduced an ontology-based data processing that enables the creation and distribution of knowledge from decentralized and unstructured data such as warehouse static data, stock exchange data, energy demand data, machine data and the production plan of heterogeneous battery processes in a learning factory. Furthermore, (Malburg, Klein, & Bergmann, 2020) developed 70 semantic web services based on different ontologies for intelligent manufacturing control to enable a near real-time verification for executing cyber-physical workflows. Last, (Kalaycı et al., 2020) created a framework in that manufacturing data from the machines of the Bosch company in Salzgitter for placing electronic components and automated optical inspection are semantically integrated to be used in quality analysis tasks. However, the semantic representation (e.g. relations between process parameters and mapping of various discrete and time-series datasets from equipment, simulation models, and controls) to enable linked and structured datasets, which can be used to analyze optimization potentials in the battery cell manufacturing, has not yet been considered. The next chapter presents the battery cell manufacturing use case scenario adressed in this paper.

3 THE VIPRO PROJECT

The concept of the semantic database model for smart battery cell manufacturing is embedded in the project ViPro – "Virtual Production Systems in Battery Cell Manufacturing for cross-process production control". The project's objective is to develop and validate a concept of cross-process control with a virtual production system. The concepts' envisioned benefits are an increase in battery cell manufacturing productivity and an improvement in the quality of the produced cells through an efficient operation. Realistic and low-risk testing of optimization measures shall be conducted in the virtual space. Once satisfactory results are achieved, these measures can be implemented seamlessly in the battery manufacturing process.

The overall system architecture of ViPro consists of different components shown in

Figure 1. The superordinate systems intelligent operation control and cross-process control are connected and perform data processing. The intelligent operation control system includes operator interfaces, where target conditions can be entered and relevant data for decisions is displayed. The cross-process control contains machine learning algorithms evaluating process control values based on intermediate product features of preceding process steps.

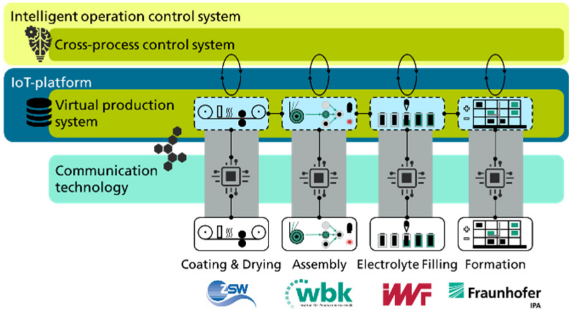

Figure 1: ViPro overall system architecture.

Product features are evaluated within the single process models of the virtual production system based on the respective process input parameters. The virtual production system includes coating, stacking, electrolyte filling, and formation quality prediction models. All of the virtual process representations have a corresponding physical system: the coating machine of the ZSW "research platform for the industrial production of large lithium-ion cells", the stacking machine of the KIT "Battery Technology Center", the electrolyte filling unit of the TU Braunschweig "Battery LabFactory Braunschweig", and the formation equipment of the Fraunhofer IPA "Center for Battery Cell Manufacturing".

To enable communication between the different ViPro components, Virtual Fort Knox (VFK) is implemented as a cloud IoT platform together with the communication middleware Manufacturing Service Bus (MSB) (Schel et al., 2018). The connection from the physical equipment to the VFK-platform is carried out with Station Connector (Defranceski, 2021), which enables a control-independent communication.

To enable and ensure communication between the different services models, the intelligent operation control and cross-process control systems, a representation of the information structure of the underlying complex process behavior and relationships, as well as the heterogeneous data, is needed. This representation of the information structure builds up a network between the different services enabling operable communication and direct data exchange. Furthermore, continuous data exchange from the physical production equipment is needed to evaluate the potential for cross-process control. To address these challenges the semantic database presented in this paper is developed and integrated into the ViPro overall system architecture.

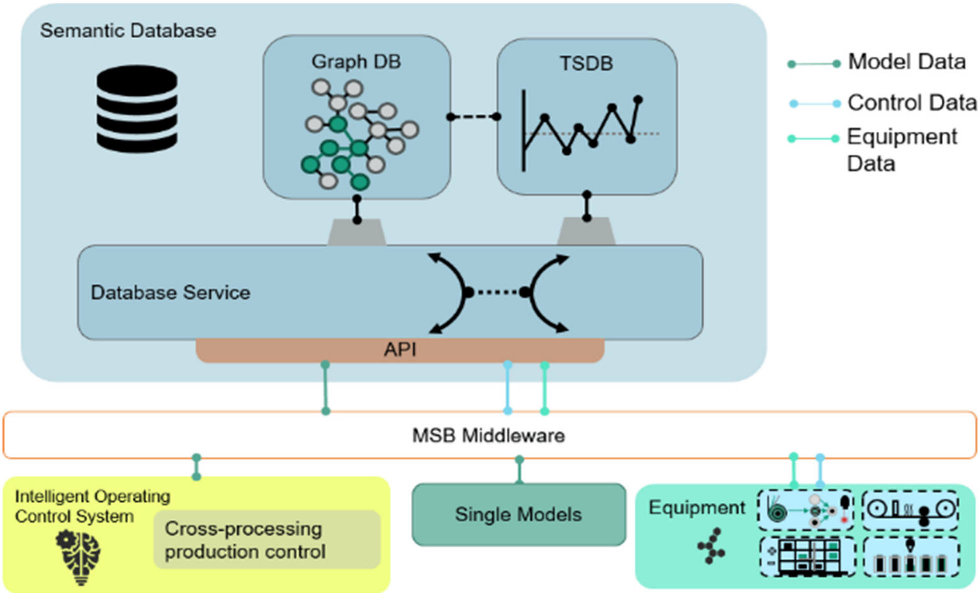

Figure 2: Semantic database concept.

4 SEMANTIC MANUFACTURING DATABASE CONCEPT

The concept described in this paper aims to develop a semantic database that structures heterogeneous data in a semantically understandable and analyzable form, enabling a solid data foundation to analyze and optimize battery cell manufacturing. First, the requirements for developing this database concept are identifieed, and then the designed concept is introduced. After that, the semantic database model is described, consisting of the semantic description of the four simulation models considered in the ViPro project. Then, it is explained why the considered simulation models require this semantic database model and which quality and general parameters of this database model are relevant for the simulation models. Last, the requirements of the intelligent operation control and cross-process control systems are introduced. These requirements define what kind of datasets they need from the semantic database and which features the semantic database should additionally have.

The following general requirements have been identified collaboratively with the project´s process engineers for the semantic database in the ViPro scenario:

- Different heterogeneous data from the models, equipment, and control systems such as experiment, equipment's process data (measurement data), and recipe data should be stored digitally.
- The data should be easily accessible and trackable.
- The relationships between different process parameters of the simulation data should be representable.
- The data from the equipment should be stored in a time frame of one-second intervals.
- The data from the equipment should be linked to the data from the simulation models.
- The database system should be centrally provisioned and enable access for all partners involved (e.g. via virtual private networks) and the possibility to operate across processes.
- The database should provide the interfaces so that required data can be processed via these interfaces by other IT systems in the ViPro project

To cover the requirements, the proposed solution is a combination of several databases for the different types of data and a service that manages the referencing of the data (see Figure 2). Since the relationships between the different types of data and their underlying knowledge need to be stored, a knowledge graph is used to store data. A knowledge graph represents a self-describing knowledge base that stores data and its schema in a graph format and illustrates their relationship (Fensel et al., 2020). A graph database is used in this paper to store the knowledge graph. Due to the requirements of storing and accessing data in real-time, a time-series database (TSDB) is used. The TSDB stores the data with an additional tag such as experiment ID to reference the experiment data, stored in a graph database. The graph database stores this reference key (experiment ID) to describe the simulation data semantically. With this design, the service can store different forms of data and respond to other ViPro IT systems' requests. A middleware enables fast and low-overhead integration of smart objects and IT services (e.g., equipment, simulation models, and database service). Therefore, a middleware that understands different protocols is used to access the data and to connect to the database service.

After introducing the designed concept for the semantic database, the semantic database model is described below. A semantic database model is required to identify the relations between the heterogeneous stored datasets in databases so that they can be easily exported from databases with their linked information to analyzing tools.

Communication between the different services models, the intelligent operation control, and cross-process control systems must be ensured. Therefore, data has to be transferred fast from and to the database.

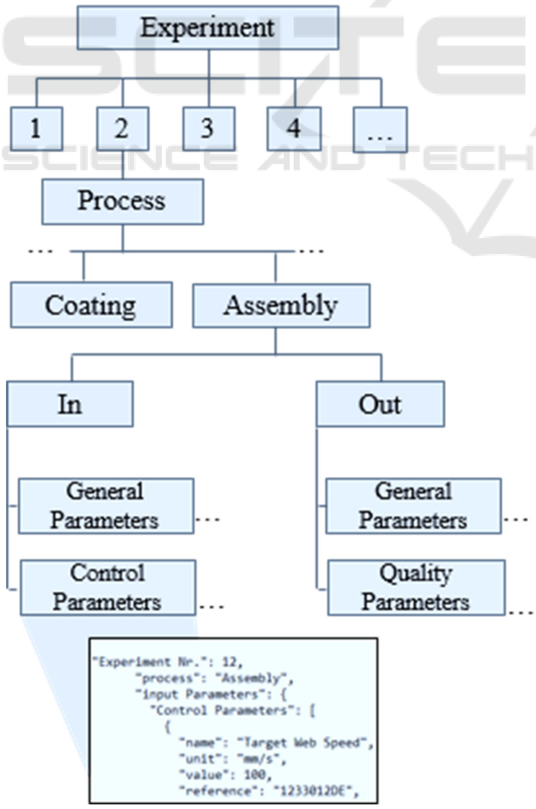

Figure 3: Semantic database model description.

Furthermore, the simulation models' heterogeneous input and output data need to be considered. That is time-series data from the process execution and control data from the cross-process control and the intelligent operation control systems. To do so, a representation of the information structure of the simulation models concerning behavior and relationship has to be implemented in the semantic database. This representation of the simulation models' information structure is the semantic database model. It is developed based on the description of the four simulation models considered in the ViPro architecture and described in the following paragraphs.

Figure 3 shows the structure of the semantic database model description. Various experiments are carried out in which data of input and output parameters are stored in the graph database. Each experiment contains the considered battery cell manufacturing processes. The input data are divided into general input parameters and control parameters. The general input parameters include, for example, material parameters and specifications regarding the cell format, whereas the control parameters include the setting parameters on the respective production machine. The output parameters are further subdivided into general output parameters and quality parameters. The general output parameters include values that each process step contains. These are, for example, statements about the energy requirements of the process and the throughput. The quality parameters provide information about the respective intermediate product properties or process quality. The structure described can be transferred into a data exchange format such as JSON or XML.

Four processes of battery cell manufacturing are considered for the semantic description of the models, which need to be integrated into the overall system of cross-process control: Electrode coating, assembly, electrolyte filling, and formation. The models are based on different approaches, and each of them focuses on specific key quality parameters.

The model of the electrode coating process is based on historical data. A Kernel density estimation is used to determine the relationships between input and output parameters and tested via cross validation and optimized with a grid search (Hasilová & Vališ, 2018). The key parameter of this model is the coating weight per unit area, which has a major impact on final cell performance. The next model considers the process of cell assembly. For this purpose, a simulation is implemented using Simcenter Amesim, which depicts the separation of electrodes and the stacking process. In particular, the target web tension

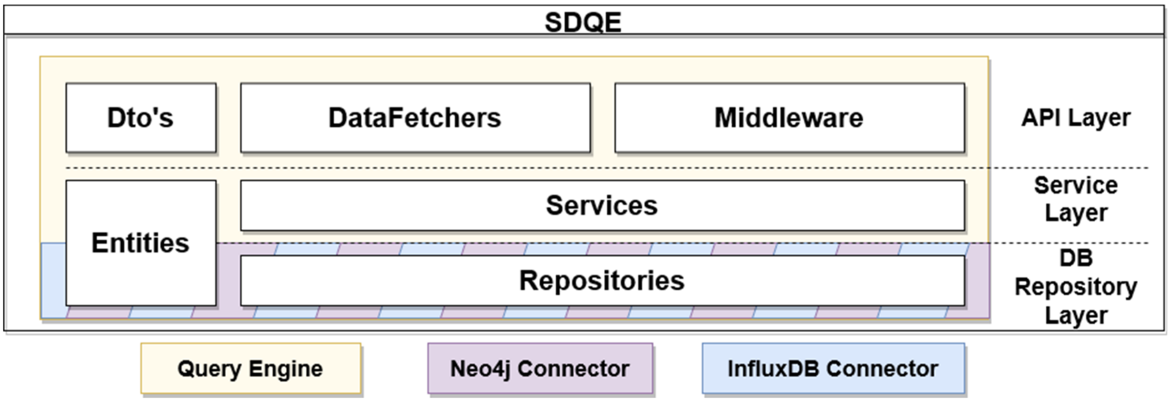

Figure 4: Semantic database components.

and the web speed are the parameters that determine the final dimensional accuracy of sheets. Furthermore, the electrolyte filling of battery cells is considered. Due to the lack of quality data on the filling process, known historical relationships of various publications are used to map the process. The amount of electrolyte filled in is a decisive factor in determining the quality parameter wetting degree. The last model represents the process of formation. It is implemented as a grey-box model in MATLAB/Simulink and consists of a discrete-event-simulation and a database model part. The key parameters of the formation process equal the parameters for final cell performance, which are cell capacity and efficiency.

To access data fast, the described structure of the database needs to be built up logically and according to the intelligent operation control and cross-process control systems that work together closely. They both need similar individually compiled datasets for different applications. One important requirement is that certain data types can be extracted from other experiments. For the human-machine interface, the control and machine learning modules, default datasets of process parameters, and historical and real-time data are needed from the database and the cross-process control system, respectively. The datasets are used for monitoring and decision support. Since the intelligent operation control system should be able to access and present all types of data related to the whole manufacturing process, including different analyses, the interfaces must be defined and designed accordingly.

5 IMPLEMENTATION

This chapter introduces the selected database software systems first. Then, the realization of the database systems for ViPro is described. At last, it is examined whether this realization fulfills the predefined requirements and can be implemented in all battery cell manufacturing processes.

Neo4j is used as a graph database for storing the model data and its references to the control data and equipment data in a synchronized manner. The determining factor was Neo4j's numerous advantages in points of performance, flexibility, and interoperability.

The process data of four different ViPro equipment should be stored in one-second intervals in a database. According to the executed benchmark and the test results of (Hao et al., 2021), InfluxDB has the best compression performance, the highest performance at writing data at high concurrency and handles queries faster compared to the other three TSDBs.

The realization of the database systems for ViPro is illustrated in Figure 4. A database service with InfluxDB and Neo4j connectors which can process data according to the relevant database, was developed and implemented. It is called Semantic Database Query Engine (SDQE). Using the Neo4j connector, the SDQE can store the simulated model data in a knowledge graph, as shown in Figure 2. The InfluxDB connector was developed in SDQE for storing the control and equipment data. The SDQE controls the underlying data with the service layer and decides which data structure is stored in which database. An application programming interface

(API) layer was defined that can process the requests from the MSB middleware and the databases.

According to the defined data structure in Figure 3, a schema was created in Neo4j. The structure in Figure 5 shows the different information and their relationships. It shows that the "Experiment" node is connected to the "Process" node, which in turn may have the "InputParameter" and "OutputParameter" nodes. The parameter nodes then store the value and its reference to the TSDB. The "Unit" node is not stored directly in the parameter to make the nodes as atomic as possible for better scalability.

As described above, the implementation solution consists of a Neo4j graph database and Influx TSDB. Neo4j graph database contains the data of simulation models, while InfluxDB allows storing the data from the equipment and control systems. To query these databases, two database interfaces with InfluxQL and GraphQL were developed in the SDQE. Furthermore, every stored data is identified with a reference key in both databases so that the data from the databases can be accessed and used more efficiently. The data stored in the Neo4j graph database is built after the semantic database model description. This design aims to visualize the relationships of the simulation data with each other and to create a fundamental database for further analysis.

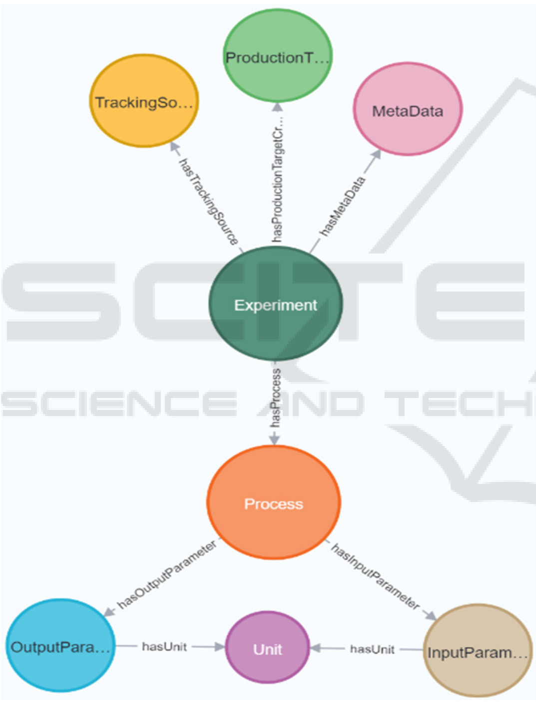

Figure 5: Data schema in Neo4j.

Moreover, both databases are deployed in a cloud platform so that only certain services and users have access to these systems via VPN. In addition, the first tests are performed with the equipment. The equipment's process data (measurement data) can be read out in a one-second time cycle and stored with a tag (experiment ID, which is created by the Neo4j graph database) in the InfluxDB.

The proposed semantic database model is a generic model for all battery cell manufacturing processes. Only the input and output parameters should be adjusted for the new manufacturing processes. Thanks to the Neo4j IT architecture, new processes can be easily added to the database with new input and output parameters. Additionally, the equipment data of the new processes can be stored in the InfluxDB as new measurements. With a new reference key, which SDQE creates, the data from Neo4j can be linked to the equipment data from the InfluxDB.

6 CONCLUSIONS

A semantic database with a model is developed considering the ViPro use case for smart battery cell manufacturing. First, requirements that enable managing heterogeneous data in a semantically comprehensible and analyzable format are identified. Then, the designed concept is introduced. Last, the realization of two different database technologies and a connecting middleware was presented.

The developed concept shows that the heterogeneous data from the simulation models, equipment, and controls can be stored in a semantically understandable way. Various datasets are linked and structured for all manufacturing processes, which can be used later to analzye and exploit battery cell manufacturing optimization potentials. In future work, the communication solution is built for the ViPro use cases so that the ViPro IT-components exchange data with each other and retain the data as in the created concept. Future developments can include the inclusion of further processes and models or a transfer to different use cases and industries.

ACKNOWLEDGEMENTS

This research was conducted within the scope of the project “Virtuelle Produktionssysteme in der Batteriezellfertigung zur prozessübergreifenden Produktionssteuerung” (ViPro). The authors gratefully acknowledge the financial support of the German Federal Ministry of Education and Research (BMBF). D.B.-K. would like to thank Steffen Stökler-Thurn for the introduction to the subject and the support.

REFERENCES

Asif, A., & Singh, R. (2017). Further Cost Reduction of Battery Manufacturing. *Batteries*, *3*(4), 17. https://doi.org/10.3390/batteries3020017

Carlier, M. (2021). Worldwide number of battery electric vehicles in use from 2016 to 2020: (in millions). Retrieved from https://www.statista.com/statistics/270603/worldwide-number-of-hybrid-and-electric-vehicles-since-2009/

Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of the ACM*, *13*(6), 377–387. https://doi.org/10.1145/362384.362685

Cui, Y., Kara, S., & Chan, K. C. (2020). Manufacturing big data ecosystem: A systematic literature review. *Robotics and Computer-Integrated Manufacturing*, *62*, 101861. https://doi.org/10.1016/j.rcim.2019.101861

Defranceski, M. (2021). Auslesen und Nutzen von Maschinendaten einfach gemacht. *Wt Werkstattstechnik Online*, *111*(03), 159–160. https://doi.org/10.37544/1436–4980–2021–03–67

Fensel, D., Şimşek, U., Angele, K., Huaman, E., Kärle, E., Panasiuk, O., Toma, I., Umbrich, J., Wahler, A. (2020). Why We Need Knowledge Graphs: Applications. In D. Fensel, U. Şimşek, K. Angele, E. Huaman, E. Kärle, O. Panasiuk, Toma, I., Umbrich, J., A. Wahler (Eds.), *Springer eBook Collection. Knowledge Graphs: Methodology, Tools and Selected Use Cases* (1st ed., pp. 95–112). Cham: Springer International Publishing; Imprint Springer. https://doi.org/10.1007/978-3-030-37439-6_4

Grevenitis, K., Psarommatis, F. [F.], Reina, A., Xu, W., Tourkogiorgis, I. [I.], Milenkovic, J., Cassina, J., Kiritsis, D. [D.] (2019). A hybrid framework for industrial data storage and exploitation. *Procedia CIRP*, *81*, 892–897. https://doi.org/10.1016/j.procir.2019.03.221

Grimmel, P., Wessel, J., Mennenga, M., & Herrmann, C. (2022). Potentials of ontology-based knowledge discovery in data bases for Learning Factories. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.4073026

Hao, Y., Qin, X., Chen, Y., Li, Y., Sun, X., Tao, Y., Zhang, X., Du, X. (2021). TS-Benchmark: A Benchmark for Time Series Databases. In *2021 IEEE 37th International Conference on Data Engineering (ICDE)* (pp. 588–599). IEEE. https://doi.org/10.1109/ICDE51399.2021.00057

Hasilová, K., & Vališ, D. (2018). Non-parametric estimates of the first hitting time of Li-ion battery. *Measurement*, *113*, 82–91. https://doi.org/10.1016/j.measurement.2017.08.030

Hildebrand, M., Tourkogiorgis, I. [Ioannis], Psarommatis, F. [Foivos], Arena, D., & Kiritsis, D. [Dimitris] (2019). A Method for Converting Current Data to RDF in the Era of Industry 4.0. In F. Ameri, K. E. Stecke, G. von Cieminski, & D. Kiritsis (Eds.), *IFIP Advances in Information and Communication Technology. Advances in Production Management Systems. Production Management for the Factory of the Future* (Vol. 566, pp. 307–314). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-30000-5_39

Kalaycı, E. G., Grangel González, I., Lösch, F., Xiao, G., ul-Mehdi, A., Kharlamov, E., & Calvanese, D. (2020). Semantic Integration of Bosch Manufacturing Data Using Virtual Knowledge Graphs. In J. Z. Pan, V. Tamma, C. d'Amato, K. Janowicz, B. Fu, A. Polleres, O. Seneviratne, L. Kagal (Eds.), *Springer eBook Collection: Vol. 12507. The Semantic Web – ISWC 2020: 19th International Semantic Web Conference, Athens, Greece, November 2–6, 2020, Proceedings, Part II* (1st ed., Vol. 12507, pp. 464–481). Cham: Springer International Publishing; Imprint Springer. https://doi.org/10.1007/978-3-030-62466-8_29

Leavitt, N. (2010). Will NoSQL Databases Live Up to Their Promise? *Computer*, *43*(2), 12–14. https://doi.org/10.1109/MC.2010.58

Malburg, L., Klein, P., & Bergmann, R. (2020, November 2 - 2020, November 4). Semantic Web Services for AI-Research with Physical Factory Simulation Models in Industry 4.0. In *Proceedings of the International Conference on Innovative Intelligent Industrial Production and Logistics* (pp. 32–43). SCITEPRESS - Science and Technology Publications. https://doi.org/10.5220/0010135900320043

O'Donovan, P., Leahy, K., Bruton, K., & O'Sullivan, D. T. J. (2015). An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities. *Journal of Big Data*, *2*(1). https://doi.org/10.1186/s40537-015-0034-z

Schel, D., Henkel, C., Stock, D., Meyer, O., Rauhöft, G., Einberger, P., Stöhr, M., Daxer, M.A., Seidelmann, J. (2018). Manufacturing Service Bus: An Implementation. *Procedia CIRP*, *67*, 179–184. https://doi.org/10.1016/j.procir.2017.12.196

Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems*, *48*, 157–169. https://doi.org/10.1016/j.jmsy.2018.01.006

Turetskyy, A., Thiede, S., Thomitzek, M., Drachenfels, N. von, Pape, T., & Herrmann, C. (2020). Toward Data-Driven Applications in Lithium-Ion Battery Cell Manufacturing. *Energy Technology*, *8*(2), 1900136. https://doi.org/10.1002/ente.201900136

Wessel, J., Turetskyy, A., Wojahn, O., Abraham, T., & Herrmann, C. (2021). Ontology-based Traceability System for Interoperable Data Acquisition in Battery Cell Manufacturing. *Procedia CIRP*, *104*, 1215–1220. https://doi.org/10.1016/j.procir.2021.11.204

Yen, I.-L., Zhang, S., Bastani, F., & Zhang, Y. (2017). A Framework for IoT-Based Monitoring and Diagnosis of Manufacturing Systems. In *Sose 2017: 11th IEEE International Symposium on Service-Oriented System Engineering: Proceedings: 6-9 April 2017, San Francisco, California* (pp. 1–8). Piscataway, NJ: IEEE. https://doi.org/10.1109/SOSE.2017.26

Yin, S., & Kaynak, O. (2015). Big Data for Modern Industry: Challenges and Trends [Point of View]. *Proceedings of the IEEE*, *103*(2), 143–146. https://doi.org/10.1109/JPROC.2015.2388958

Zhong, R. Y., Xu, X., Klotz, E., & Newman, S. T. (2017). Intelligent Manufacturing in the Context of Industry 4.0: A Review. *Engineering*, *3*(5), 616–630. https://doi.org/10.1016/J.ENG.2017.05.015