A Classification of Data Structures for Process Analysis in Internal Logistics

Maximilian Wuennenberg[1[0000-0002-2036-4624]], Charlotte Haid[1[0000-0001-7047-2343]], and Johannes Fottner[1[0000-0001-6392-0371]]

[1] Technical University of Munich, Boltzmannstrasse 15, 85748 Garching bei Muenchen, Germany

Abstract. Data Science plays a crucial role in driving new approaches to process optimization. With the increasing complexity of internal logistics systems, data-oriented methods have become essential in addressing the challenges that arise. However, standardized process analytics frameworks are lacking due to the heterogeneity of the underlying processes and the resulting data. This article aims to address this complexity by presenting a categorization of internal logistics data, consolidating the current state of the art. The categorization takes into account both real-world and scientifically proposed data architectures, providing a comprehensive overview. It includes a classification of comparative data fields based on their importance, the associated internal logistics processes, and potential usage scenarios. This classification is designed to cater to different use cases, such as diagnostics or prescriptive analytics. By presenting this categorization, the article enables practitioners to effectively leverage generated process data in a more goal-oriented manner. It empowers them to conduct suitable analyses tailored to their specific needs and objectives, based on the provided data architectures. In summary, this article offers valuable insights into internal logistics data categorization, providing a framework for practitioners to make informed decisions and optimize processes using data-driven approaches.

Keywords: Data Analytics, Internal Logistics, Process Analysis.

1 Introduction

1.1 Initial situation and motivation of the topic

Internal logistics systems (ILS) are subject to an increasing level of digitization [1]. Processes are controlled by information technology (IT) systems such as enterprise resource planning (ERP) or warehouse management systems (WMS). The processes, on the other hand, generate data which is subsequently stored and further transferred by these systems [2]. During recent years, a wide range of process types has been adapted to this digital connection, often referred to as Industry 4.0 [3]. This development yields the potential for new types of process analysis, which are the result of recent advances in the field of data science. One example of this analysis is process mining, a method in which data is analyzed from an event-driven perspective, with the goal of obtaining insights and finding process improvement potentials [4]. To that end, the right choice of application is referred to as context-aware process mining [5]. However, in order for analyses to be beneficial from a practitioner's perspective, they rely largely on the data architecture, i.e., in terms of data maturity, but also in terms of the right choice of analyzed data fields and a wise choice of data science applications [6, 7]. The main challenges faced when analyzing process data using digital methods is the heterogeneity of this data, and its high specification with regard to the respective process [8]. Depending on the IT system recording the data, and on the process type, different data structures can be generated in terms of data fields and data types. This issue becomes even more challenging given the inherent complexity and heterogeneity of ILS. Consistent documentation of the ILS process landscape, combined with involved IT systems, the generated data fields, and how they can be brought together with the objective of effective data science applications in mind [9] are all lacking. An interesting scenario from a practitioner's perspective is when a certain dataset is present regardless of any data analysis approaches yet implemented. To that end, the objective is often rather to extract as much value from the given data, instead of acquiring additional data sources.

1.2 Objective

Several recent research contributions have addressed the issue of data science in the field of ILS [2, 10–13]. However, given their focus on certain individual aspects of this ample domain, the applied data structures are highly problem-specific and thus unable to consider all facets of this field. A domain-overarching data architecture is necessary in order to ensure the broader applicability of future research [14]. A link between data creation in different IT systems and data science application with its various goals still needs to be established. To enable this link, all relevant ILS data fields must be classified according to several characteristics, including potential use cases (see Figure 1). Therefore, this article aims to reach the following two objectives:

- Classification of ILS data which considers various IT systems and process types
- Association of relevant data structures with their potential use cases

By fulfilling these objectives, practitioners cannot just decide which data is necessary in order to execute a desired analysis type. They can also deal with the situation that a certain data structure is already given, and that as much value as possible shall be derived from it without additional data gathering effort.

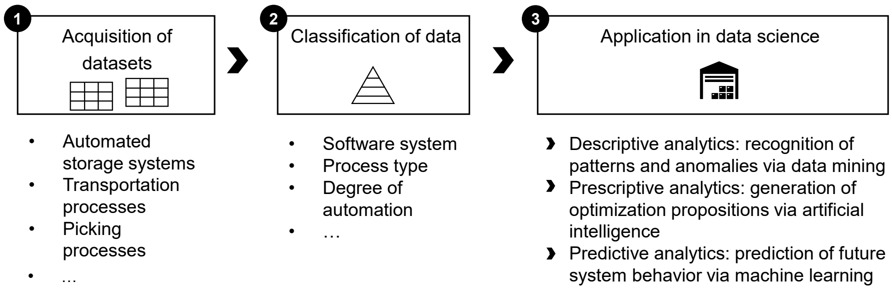

Fig. 1. Procedure for obtaining, classifying, and applying the datasets in this research.

2 State of the art

2.1 Internal logistics process landscape

ILS processes consist of several elementary activities: conveying, storing, sorting, and picking. Conveying covers all transportation processes of goods or persons within a locally limited area by technical means. Whenever goods intentionally remain in a certain position, it is considered to be a storage activity [15]. Sorting describes a diverging flow of materials so that transportation units (TU) having certain properties can be separated from others. Apart from diverging material flows, sorting also covers converging ones. All of these processes can be executed with or without the assistance of technical devices that are either human-controlled, mechanized, or self-controlled (automated) [16].

2.2 Process control

In this article, ILS cover all enterprise activities related to conveying, storage, picking, unloading, and loading of goods, including additional processes like packaging. These activities can be executed in a manual, a partly assisted, or an autonomous manner. However, in order to enable the application of data science, each activity must leave a digital footprint. In the case of partly assisted processes, this can be achieved by, e.g., using handheld terminals or barcode scanners that allow workers to confirm executed process steps and thereby generate timestamps, locations, or order information. The generated data is then further transferred and stored, for instance in a WMS or an ERP system.

The various types of IT systems are structured based on the layer-based architecture of the automation pyramid (see Figure 2) [17]. The pyramid contains five layers of abstraction, leading from basic sensor and actuator data up to enterprise-overarching ERP systems.

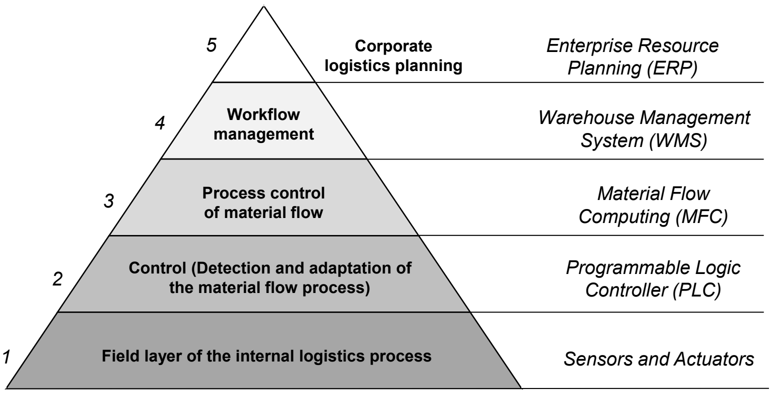

Fig. 2. Automation pyramid, adapted from [17].

The data flow in the automation pyramid is as follows: The highest layer generates a task and sends it to the lower layers, where it becomes more and more detailed. On the lowest layer, the system interacts with the physical environment. After that, the process data is in turn recorded and sent to the higher layers again. The transferred information becomes increasingly abstracted during this procedure (see Figure 3).

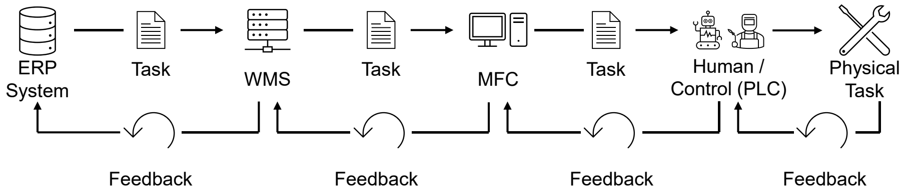

Fig. 3. Cascading data transfer between the layers of the automation pyramid.

The command chain of the automation pyramid is distributed within the following layers: on the lowest layer, the sensors and actuators of the ILS directly interact with their environment [17]. The processing of sensor inputs and generation of actuator outputs is organized on the second layer, mostly by programmable logic controllers (PLC) [17]. The latter fulfill specific material flow operations, which are generated by material flow computers (MFC) on the third layer of the pyramid. MFC decide upon the specific order in which tasks are accomplished and serve as the interface between real-time system components (PLC) and non-real-time system components (higher layers) [17]. Systems on the fourth layer (WMS and MES) are responsible for all operations in a certain subdomain of the overall enterprise process [17]. They thereby offer the possibility for data consolidation, which is necessary for the aggregation of substantial Key Performance Indicators (KPI). Above that, the ERP system is responsible for the long-time planning of enterprise processes and the coordination of all necessary subprocesses [17]. It should be noted that, when a sub-process contains manual activities instead of being fully automated, the human operator takes the role of the two lower layers, i.e., taking commands from a terminal linked to an MFC-equivalent control system and giving confirmations to this terminal.

2.3 Data analysis

The data being generated by these systems can contain various types of information, which is usually represented by the choice of appropriate data types. Numeric information on an ordinal scale can be covered as integers, whereas interval scale values need single or double precision float data fields for their representation. Timestamps can be recorded either as integers when a suitable conversion is available, or as datetime data. String data fields are able to cover the broadest range of information, but, since they are unstructured, their subsequent data analysis is most complicated.

Most research publications in the field of ILS processes have considered only a certain subdomain of the entire process and system landscape. For instance, machine learning is used to predict the behavior of inbound unloading processes based on historic data [13]. In a specific application of conveying operations, which are examined on the fifth layer of the automation pyramid (the ERP layer), analyses like Sankey diagrams can be deduced [2]. Another approach combines data from the fifth and fourth (MES) layer in order to set up a simulation model which can be used for process optimization [11]. Furthermore, the consolidation of process data can lead to the application of process mining. However, the underlying data structure is limited to one certain type of subprocess [12]. Also, there is one approach in which data sources from more different layers is considered: in addition to the fourth and fifth, also the second layer (PLC) is taken into account for the prediction of KPI of the ILS [10]. However, this approach only addresses the particularities of process simulation models. Finally, there is the option of deriving process optimization potentials from a combination of business process modeling and simulation models [18]. One main challenge when following this approach is the need for manual parametrization of the models, which could be automated using data science.

2.4 Summary

Condensing the findings from the state of the art leads to the following conclusion: the ILS process landscape is usually vertically structured, following the principle of the automation pyramid. Assuming a sufficient level of automation (and thus digitization) of the subprocesses involved, the generated data can be analyzed starting from a mere description of the database, up to the automated generation of recommendations for process optimization. However, existing research either only deals with isolated considerations of individual subprocesses, or is limited to a specific application scenario for the data, e.g., process simulation parametrization or process mining. As a result, the overarching consolidation of ILS process data aiming for a generalized data analysis application can be identified as a research gap being addressed by this article.

3 Sample datasets from ILS processes

In this article, several ILS were considered in order to obtain a fundamental database able to cover the potentially broad variety of this process landscape. These datasets were extracted both from applications within industrial companies and research projects. They have been pseudonymized and condensed for this research work. All in all, the following processes were covered (for an overview, see Table 1):

1. Automated storage system: high rack with stacker cranes (Elementary ILS process: Storage; Data transfer: between PLC and MFC layer)
2. Automated transportation process: conveying of TU in load units, several units being attributed to single tasks, executed by an automated small-parts warehouse (Elementary ILS process: Storage; Data transfer: between WMS and ERP layer)
3. Data field structure for the information interchange between MFC and WMS (Elementary ILS process: Storage; Data transfer: between MFC and WMS layer)
4. Inbound logistics: forklift trucks for the transportation from goods receipt to the storage (Elementary ILS process: Conveying; Data transfer: between WMS and ERP layer)
5. Manual picking process: picking of various articles in a shelving rack according to pick lists (Elementary ILS process: Picking; Data transfer: between MFC and WMS layer)
6. Multi-process ILS: combination of automated small-parts warehouse and stacker crane-operated high rack storage system for various articles, synchronized with transfer vehicles (Elementary ILS process: Storage, Conveying, Sorting; Data transfer: between MFC and WMS layer)
7. Standardized data transfer protocol for the operation of automated guided vehicles (AGV): communication of AGV with the overarching process control (Elementary ILS process: Conveying; Data transfer: between PLC and MFC layer) [19]
8. String-based data protocol for the information interchange between MFC and PLC (Elementary ILS process: Storage; Data transfer: between PLC and MFC layer)
9. Transportation process executed by tugger trains: data regarding the trips on different transportation routes within a production material replenishment process (Elementary ILS process: Conveying; Data transfer: between WMS and ERP layer)
10. Transportation tasks: data from conveying systems (Elementary ILS process: Conveying; Data transfer: between WMS and ERP layer)

Table 1. Overview of the considered ILS sample datasets.

No. of dataset	Row No.	Column 1	Column 2										Column n
1	Header	Protocol number	Protocol time	Source position	Current position	Target position	Loading eq. type	Weight					
	1	123456787	07.07.2022 06:59	01-01-017-21-3	RE1-10	02-05-077-02-1	EURO-400	55250					
	2	123456788	07.07.2022 07:01	01-01-024-23-5	RE2-10	02-05-007-18-1	EURO-700	112500					
2	Header	Task number	No. of pos.in task	Material number	Source position	Status	Protocol time [1]	Protocol time [2]	Type / target pos.	Target position			
	1	1234565	2	1120 20050	0000025100	x	10.07.2022	08:42:19	ASS	1000			
	2	1234566	1	2001 41640	5001195520	y	10.07.2022	14:40:46	ASS	1000			
3	Header	Status	User	Protocol time	Current Position	Source position	Target position	Task number	No. of pos. in task	Protocol number	Weight	Loading eq. type	Quant.
	1	0	B	01.07.2022 05:40	1-1-5	0-1-1	5-2-5	120	1	123454	1200	Box	3
	2	0	A	01.07.2022 05:45	1-2-1	0-1-1	5-2-5	122	1	123455	1800	Box	10
4	Header	Task number	Material number	Protocol time [1]	Protocol time [2]	User	Source position	Quantity	Target Position				
	1	123456787	886644220	24.06.2022	19:24:00	A325KP03	102GKF870	12 PCS	082FG003				
	2	123456788	997755331	24.06.2022	19:27:58	A325KP02	161GKF378	200 PCS	061FG374				
5	Header	Task number	Source position	Target position	Material number	Quantity	Protocol time [1]	Protocol time [2]	Status				
	1	123456787	F01-02-05	1OUTBOUND	A5U052125	1	13.07.2022	08:25:29	050				
	2	123456788	F01-02-05	1OUTBOUND	A5U052125	1	13.07.2022	08:26:12	100				
6	Header	Task number [1]	Task number [2]	Protocol number	No. of pos. in task	Weight	Type / target pos.	Current position	Source position	Target position	User	Protocol time	
	1	123456787	987654319	12345678	7654320	5150	2	2521502	0	2530405	WMS-MFC	03.07.2022 09:00	
	2	123456788	987654320	12345678	7654321	5520	1	2020521	0	0	WMS-MFC	03.07.2022 09:00	
7	Header	Protocol number	Protocol time	Loading eq. type	Task number								
	1	12345676	15.04.2017 11:40	Pallet	123455								
	2	12345677	15.04.2017 11:40	Pallet	123456								
8	Header	Source position	Type / target pos.	Loading eq. type	Target position	Current position	Status	Weight					
	1	CAB01	CA1	Pallet	12345	01-IP01	0	1250					
	2	CAB01	CA1	Pallet	12345	Loop01-01	1	51000					
9	Header	Protocol number	Task number	Status	Source position	Type / target pos.	Target Position	Loading eq. type	User	Protocol time			
	1	7986281	761391	0	225	EV-NU	125	VY0105C	SM8-SCAN-04	05.08.2022 14:45			
	2	7986282	761391	10	225	EV-NU	125	VY0105C	SM8-SCAN-04	05.08.2022 14:45			
10	Header	Material number	Quantity	Source position	Target position	Protocol time							
	1	215122	100	Goods receipt	Storage 1-1	24.07.2022 14:01							
	2	215125	150	Goods receipt	Storage 1/2L	24.07.2022 14:15							

The processes considered in this paper cover all of the elementary ILS processes. Figure 4 arranges the available datasets in the automation pyramid. Consolidation of the existing data fields within these datasets thus represents a broad overview of the ILS landscape. From this point, a framework for the classification of ILS data and an assignment to potential use cases can then be developed.

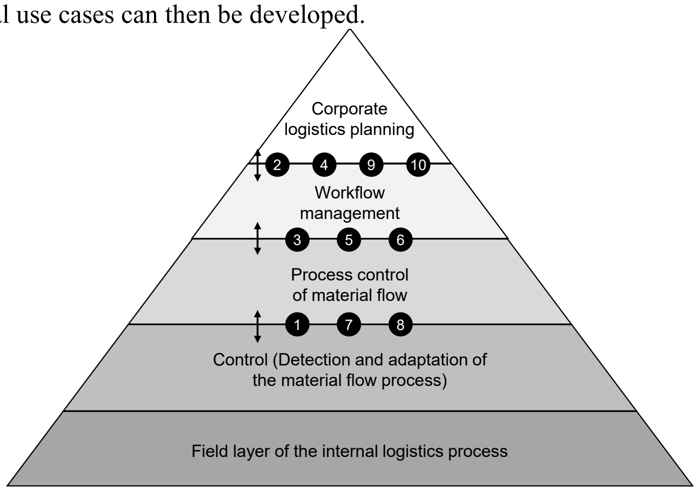

Fig. 4. Sample datasets assigned to the layer transitions of the automation pyramid.

3.1 Classification of data structures

As explained in the previous section, the various elementary processes of ILS are often embedded in different IT landscapes. The generated data is consolidated on several layers of the automation pyramid. In addition, the ILS processes usually differ by various characteristics. From the perspective of a practitioner, a classification of data according to a maturity model must be executed individually for each subprocess. If the data is inconsistent with its description (i.e., existing metadata), its meaningfulness is diminished [20]. The analysis of data can be grouped into three levels (see Figure 4).

Descriptive analytics cover the consolidation and preprocessing (e.g., outlier elimination) of data fields with the objective of identifying and visualizing patterns and anomalies in the data [21]. Obtaining and consolidating the data is covered by the extract-transform-load (ETL) process [2]. More sophisticated yet, predictive analytics enable the diagnosis and monitoring of the system, thereby explaining certain phenomena in the ILS behavior [21].

As Figure 5 shows, such objectives can still be achieved by applying data mining scenarios. Further on, a prediction model can be set up that uses algorithms to predict future behavior and the development of relevant KPI based on historic values. The key to this approach is a model using optimization metrics such that the algorithm can iteratively improve the accuracy of the prediction. [22]. Therefore, machine learning approaches are necessary [23]. The mightiest and most complex level of data science, prescriptive analytics often incorporates simulation to deduce suggestions for system optimization [21]. If the model is able to identify and apply optimizations on its own, it can execute tasks similar to those of a human operations manager. The behavior of this highest level of data science can thus be classified as deep learning [22].

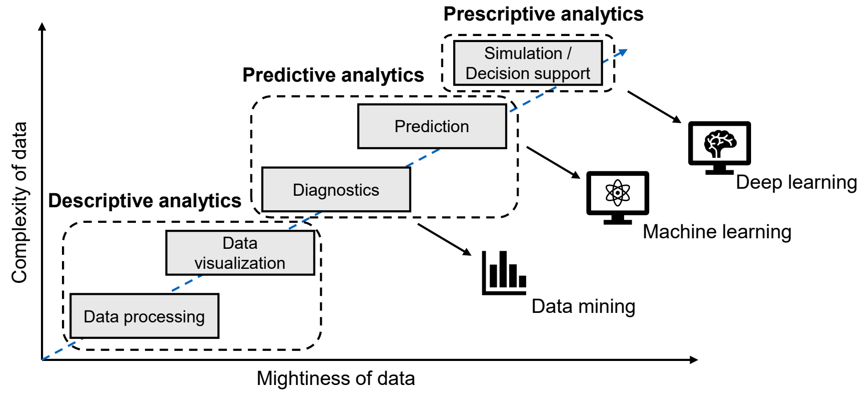

Fig. 5. Evolutionary steps of data science, adapted from [21].

3.2 Application on the sample data

The consolidation of all data fields from the ten ILS process datasets introduced in the previous section leads to the classification shown in Table 2, Table 3, and Table 4.

Table 2. Data classification (Part I: core data).

No.	Name of information	Exemplary data type	Process types
1	Task number	String	Storage, Picking, Conveying, Sorting
2	Protocol time	Datetime	Storage, Picking, Conveying, Sorting
3	Source position	String	Storage, Picking, Conveying, Sorting
4	Target position	String	Storage, Picking, Conveying, Sorting

The tables are sorted (column: No.) by decreasing frequency of the respective data field, No. 1 appearing in 9/10 and 26-33 in only 1. The name of information is a generalization of the (mostly different) names the respective type of process information is given in the datasets. The exemplary data type column shows how the information is represented. If different types were used among the datasets, the most general one was considered in the table (integer – float – datetime – string, with ascending generalization). The process types are created as a reference between the elementary processes covered by a certain dataset, and the datasets that cover a certain data field. Again, if several process types apply, then all are mentioned in the table.

Table 3. Data classification (Part II: extended selection No. 1).

No.	Name of information	Exemplary data type	Process types
5	Protocol number	String	Storage, Conveying, Sorting
6	Number of position in task	String	Storage, Conveying, Sorting
7	Current position	Integer	Storage, Conveying, Sorting
8	Type of target position	String	Storage, Conveying, Sorting
9	Loading equipment type	String	Storage, Conveying
10	Material number	String	Storage, Conveying, Sorting
11	Quantity	Integer	Storage, Conveying, Sorting
12	Weight	Integer	Storage, Conveying, Sorting
13	Status	Float	Storage, Conveying, Sorting
14	User	String	Storage, Conveying, Sorting
15	Next position	String	Storage, Conveying, Sorting

As the separation of data fields into three different tables (the latter ones being called extended selection) indicates, not all types of process information share the same relevance for applications of data science. The information represented in Table 2 should always be present when examining the respective subprocess on any particular layer of the automation pyramid, even if the methods applied only cover elementary data mining scenarios (i.e., descriptive analytics or diagnostics). The data covered by this table describes atomic material movements within ILS processes. Albeit being rudimentary, insights can be gathered for example by creating a histogram that depicts the time spent by TU between certain positions over a list of tasks. Data from Table 3 becomes more relevant when addressing more sophisticated process analytics approaches, i.e., it is necessary to apply approaches like machine learning models in predictive analytics scenarios to turn data into insights. Therefore, data fields in this table cover information that can be used to interpret the overall process rather than individual movements. For the creation of predictive analytics models, there is a need for data fields which describe more than just the output behavior to be predicted (e.g., the time spent within the system). Information such as material number or weight allow the prediction model to deduce an output parameter behavior that depends on those process information inputs. The information covered in this table is also typically collected by a conventional, analog value stream analysis [24]. Finally, Table 4 contains supplementary information. It is not compulsory for initial data exploration or description, and various analysis types can be used without having access to it. However, with the most complex and mighty data science applications (such as artificial intelligence as an enabler for prescriptive analytics), additional insights can be generated by considering this data. A prescriptive analytics model will usually not rely on all those fields to deduce an optimization potential, but it cannot be determined in advance which information is actually required. Just like a human process optimization expert, improvements can often only be reached when all information is made available.

Table 4. Data classification (Part III: extended selection No. 2).

No.	Name of information	Exemplary data type	Process types
16	Time of task creation	Datetime	Storage, Picking, Conveying, Sorting
17	Priority	Integer	Storage, Conveying, Sorting
18	Type of source position	String	Storage
19	Quantity before task fulfillment	Integer	Storage
20	Type of task	String	Storage, Conveying
21	Previous position	String	Storage, Conveying, Sorting
22	Type of current position	String	Storage, Conveying, Sorting
23	Type of next position	String	Storage, Conveying, Sorting
24	Loading equipment ID	String	Storage, Conveying, Sorting
25	Quantity after task fulfillment	Integer	Storage
26	Length	Integer	Storage, Conveying, Sorting
27	Width	Integer	Storage, Conveying, Sorting
28	Height	Integer	Storage, Conveying, Sorting
29	Volume	String	Conveying
30	Target system of data transfer	String	Conveying
31	User person	Integer	Conveying
32	Assignment of user and person	Datetime	Conveying
33	Route	Datetime	Conveying
34	Tour	String	Conveying
35	Duration of tour	String	Conveying
36	Due date	String	Conveying
37	Starting time	String	Conveying

3.3 Results of the application

A summarizing association of these data structures with appropriate application scenarios can be derived on the basis of the consolidation and classification of various data fields. Therefore, the user must follow several steps when implementing data science-based process analyses building on an existing database: First of all – for all existing ILS subprocesses – the existing data transfer interfaces need to be merged according to the respective layer of the automation pyramid. The data maturity must then be assessed [7]. One can then determine which subprocesses should be considered for data-based process analytics. After that, the user must check which data fields are available on the different layer transitions for each subprocess being analyzed. The potential scope of data science applications can be determined based on the importance of the data fields indicated by Table 2, Table 3, and Table 4. Finally, a suitable set of process analytics algorithms can be selected. In this context, core data must be available for the most basic descriptive and diagnostic data mining scenarios. For more sophisticated predictive analytics use cases, the more important fields from the extended data selection must be available so that machine learning algorithms can be trained and thus generate meaningful insights. If all of the extended data fields are available, artificial intelligence is then able to determine prescriptive analytics in the form of improvement suggestions.

4 Conclusion

4.1 Interpretation

The preceding section of this article developed a classification of process data and the association of suitable data science applications for process analytics. Intended for practitioners in the domain of process analysis and optimization, this article contributes to a systematic assessment of the existing database and supports in selecting suitable algorithms able to fulfill their purpose. In an early phase of the process analysis, it can already be determined which subprocesses can be analyzed in a reasonable way, and what degree of data analysis is appropriate. Reciprocally, it can be deduced which additional data would be necessary so that a desired stage of data analysis could be executed. This means it is not necessary for practitioners to integrate specific data sources to obtain a highly specific set of analyses, but instead, the optimum of potential insights from a given data base can be drawn.

The consolidation of various sets generated by several industrial enterprises and research projects ensures the necessary generalizability of the findings. The gap in existing data science publications within the ILS domain was able to be addressed.

However, in order for data science projects to be successful, the need for process experts with a reasonable amount of implicit process knowledge still exists. That is, without this expert knowledge, an appropriate assessment of the database is not possible, and the conclusions drawn by the application of the presented findings could be misleading. This means that this research cannot fully substitute humans in ILS process analysis and optimization, but rather play an important role in supporting humans and reducing work effort, while at the same time ensuring a timely estimation of target attainment.

4.2 Limitations

The findings presented herein allow for a classification of process data enabling the reasonable selection of applications scenarios. However, target-oriented data science frameworks for all of these application scenarios are not fully present. In other words, further data science methods should be developed in addition to the research works discussed in the state of the art. Furthermore, the classification presented in this article must be tested – using various datasets and data science applications – in order to determine whether it is fully applicable in every possible scenario. Specifically, the transition between the three data tables presented is not sharp and clear. Depending on the individual circumstances, data science models to be applied might differ from the classification provided in this article. This situation leads to research tasks that need to be covered by future works.

Furthermore, the availability of data fields alone is necessary, but it is not sufficient for the application of certain data science methods. As indicated in data maturity models, the columns must be consistently filled with entries, and these entries must be meaningful with respect to the definition of the data architecture. Especially when applying methods with optimization metrics, e.g., machine learning or artificial intelligence, the generated results can only be as good as the input data.

4.3 Conclusion and outlook

Given the data classification presented herein for optimization scenarios in the domain of internal logistics, a reasonable tool for practitioners in the operations management can be provided. Deduced from a broad set of data sources, the findings can help to successfully implement data science projects given that the knowledge of process experts is present as well. Thereby, the first research objective has been addressed. The classification allows a differentiation between descriptive analytics, predictive analytics, and prescriptive analytics approaches. Sensible analysis methods can thus be deduced following this classification. At the same time, the different types of process in the ILS domain are considered. With that in mind, also the second research objective can be considered as achieved. Industrial application projects can be set up with no need for an objective-specific integration of additional data sources in advance.

In the future, a further detailing of the entire framework by developing target-oriented process analytics applications could help to enhance the generalizability and applicability of the findings presented. Furthermore, by applying the framework to different industrial use cases, the validity of the approach could be examined with the perspective of discovering potential issues for improvement.

References

1. Vogel-Heuser, B., Konersmann, M., Aicher, T., Fischer, J., Ocker, F., Goedicke, M.: Supporting evolution of automated Material Flow Systems as part of CPPS by using coupled meta models. In: IEEE Industrial Cyber-Physical Systems (ed.) Proceedings 2018 IEEE Industrial Cyber-Physical Systems (ICPS): ITMO University, Saint Petersburg, Saint Petersburg, Russia, 15-18 May 2018, pp. 316–323. IEEE, Piscataway, NJ (2018).
2. Knoll, D., Prüglmeier, M., Reinhart, G.: Materialflussanalyse mit ERP-Transportaufträgen: Automatisierte Ableitung und Visualisierung von Materialflüssen in der Produktionslogistik. Werkstattstechnik online 107(3), pp. 129–133 (2017).
3. Gehlhoff, F., Fay, A.: On agent-based decentralized and integrated scheduling for small-scale manufacturing. at - Automatisierungstechnik 68(1) 15–31 (2020).
4. van der Aalst, W.:: Process Mining: Data Science in Action. Springer-Verlag, Berlin, Heidelberg (2016).
5. Becker, T., Intoyoad, W.: Context Aware Process Mining in Logistics. Procedia CIRP 63 557–562 (2017).
6. Wuennenberg, M., Muehlbauer, K., Fottner, J., Meissner, S.: Towards Predictive Analytics in Internal Logistics – an approach for the Data-driven determination

of Key Performance Indicators. CIRP Journal of Manufacturing Science and Technology 44, pp. 116–125 (2023).
7. Muehlbauer, K., Wuennenberg, M., Meissner, S., Fottner, J.: Data driven logistics-oriented value stream mapping 4.0: A guideline for practitioners. 18th IFAC Workshop on Control Applications of Optimization (2022).
8. Burow, K., Franke, M., Deng, Q., Hribernik, K., Thoben, K.-D.: Sustainable Data Management for Manufacturing. In: IEEE Conference on Engineering, Technology and Innovation (ed.) 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC). IEEE (2019).
9. Wünnenberg, M., Hujo, D., Schypula, R., Fottner, J., Goedicke, M., Vogel-Heuser, B.: Modellkonsistenz in der Entwicklung von Materialflusssystemen: Eine Studie über Entwicklungswerkzeuge und Einflüsse auf den Produktentstehungsprozess. ZWF 116(11) 820–825 (2021).
10. Vernickel, K., Brunner, L., Hoellthaler, G., Sansivieri, G., Härdtlein, C., Trauner, L., Bank, L., Fischer, J., Berg, J.: Machine-Learning-Based Approach for Parameterizing Material Flow Simulation Models. Procedia CIRP 93 407–412 (2020).
11. Milde, M., Reinhart, G.: Automated Model Development and Parametrization of Material Flow Simulations. in *2019 Winter Simulation Conference (WSC)*In: N. Mustafee (ed.), [Piscataway, NJ]. IEEE (2019, pp. 2166–2177.
12. Knoll, D., Reinhart, G., Prüglmeier, M.: Enabling value stream mapping for internal logistics using multidimensional process mining. Expert Systems with Applications 124 130–142 (2019).
13. Knoll, D., Prüglmeier, M., Reinhart, G.: Predicting Future Inbound Logistics Processes Using Machine Learning. Procedia CIRP 52 145–150 (2016).
14. Wuennenberg, M., Vollmuth, P., Xu, J., Fottner, J., Vogel-Heuser, B.: Transformability in Material Flow Systems: Towards an Improved Product Development Process. in *Lecture Notes in Networks and Systems, Managing and Implementing the Digital Transformation: Proceedings of the 1st International Symposium on Industrial Engineering and Automation ISIEA 2022*In: D. T. Matt, R. Vidoni, E. Rauch, and P. Dallasega (eds.), Cham. Springer International Publishing (2022, pp. 3–14.
15. Hompel, M. ten, Schmidt, T., Dregger, J.:: Materialflusssysteme. Springer Berlin Heidelberg, Berlin, Heidelberg (2018).
16. Wiendahl, H.-P., Reichardt, J., Nyhuis, P.:: Handbuch Fabrikplanung.: Konzept, Gestaltung und Umsetzung wandlungsfähiger Produktionsstätten. 2nd edn. Hanser, München (2014).
17. Verband Deutscher Maschinen- und Anlagenbau: Datenschnittstellen in Materialflußsteuerungen., VDMA 15276 (1994).
18. Wuennenberg, M., Wegerich, B., Fottner, J.: Optimization of Internal Logistics using a combined BPMN and Simulation Approach. In: I. A. Hameed, A. Hasan, Alaliyat, and Saleh Abdel-Afou (eds.) Proceedings of the 36 th ECMS International Conference on Modelling and Simulation ECMS 2022, pp. 13–19. Pirrot, Saarbrücken (2022).
19. Verband der Automobilindustrie: Schnittstelle zur Kommunikation zwischen Fahrerlosen Transportfahrzeugen (FTF) und einer Leitsteuerung., VDMA 5050 (2022).

20. Klare, H., Kramer, M. E., Langhammer, M., Werle, D., Burger, E., Reussner, R.: Enabling consistency in view-based system development — The Vitruvius approach. Journal of Systems and Software 171(110815) 1–35 (2021).
21. Kargul, A.: Entwicklung eines Baumaschinenmanagements zur integrativen und adaptiven Steuerung des Maschinenbestandes über den Lebenszyklus. . Dissertation, Lehrstuhl für Fördertechnik Materialfluss LogistikTechnische Universität München, Garching b. München (2020).
22. Schuh, G., Reinhart, G., Prote, J.-P., Sauermann, F., Horsthofer, J., Oppolzer, F., Knoll, D.: Data Mining Definitions and Applications for the Management of Production Complexity. Procedia CIRP 81 874–879 (2019).
23. Rebala, G., Ravi, A., Churiwala, S.:: An Introduction to Machine Learning. Springer International Publishing, Cham (2019).
24. Rother, M., Shook, J.:: Learning to see.: Value-stream mapping to create value and eliminate muda. 1st edn. Lean Enterprise Inst, Boston (2018).