A Systematic Literature Review on the Application of Process Mining to Industry 4.0

Katsiaryna Akhramovich[1], Estefanía Serral

Research Centre for Information Systems Engineering (LIRIS), KU Leuven, Belgium

Carlos Cetina

SVIT Research Group, Universidad San Jorge, Spain

[1]Corresponding authors: Katsiaryna Akhramovich
Email addresses: akhr218@gmail.com (Katsiaryna Akhramovich), estefania.serralasensio@kuleuven.be (Estefanía Serral), ccetina@usj.es (Carlos Cetina)

A Systematic Literature Review on the Application of Process Mining to Industry 4.0

Abstract

The transition to Industry 4.0 means a new era in manufacturing with a new level of production automation, human-to-machine cooperation and product customization. It provides many benefits and opportunities to both enterprises and consumers and allows for principally new level of cooperation. At the same time, the complexity of business processes, large volume and the complex structure of data generated and processed by different Industry 4.0 technologies create serious challenges for Business Process Management (BPM). Process mining (PM) can tackle these challenges. PM is a relatively young discipline that is positioned between process-centric and data-centric approaches and focuses on discovering, conformance checking and enhancement of end-to-end business processes. Moreover, new types of PM deal with performance analysis, comparative analysis of several processes, making predictions and triggering improvement actions. This systematic literature review studies the applicability of PM in Industry 4.0 and the benefits that PM can provide to each of the four aspects of Industry 4.0: smart factories, smart products, new business models and new customer services. Approaches of PM proposed in the selected studies are analysed and classified according to two dimensions of the study: PM and Industry 4.0. The research gaps identified while performing the systematic literature review show possible directions for further research in the area.

Keywords: Industry 4.0; Process Mining; Systematic Literature Review.

1 Introduction

After having gone through three industrial revolutions, the world is nowadays entering a new technological era that has been named Industry 4.0, which is a fundamentally new paradigm shift in industrial production [1]. The transition to Industry 4.0 involves an absolutely new level of automation in manufacturing processes [2], a new level of human-to-machine cooperation [3], and a new level of product and service customization [4]. The future vision of manufacturing covers four main aspects: smart factories, smart products, new business models, and new customer services [5, 6].

A prerequisite for the proper functioning of Industry 4.0 technologies is accurate and real-time data sharing, whether during the production cycle or in different product lifecycle stages [2]. Taking into account the huge amount of data that is produced and transported through different channels in Industry 4.0, advanced analytical data techniques have become crucial in uncovering business information that can be used to support efficiency [7]. Machine learning, which is one of the key techniques that are traditionally used for data analytics, gains additional importance in Industry 4.0 [8].

Another crucial point of attention for an organisation that is involved in Industry 4.0 is the focus on business processes [6]. The concept of business processes is central for Industry 4.0, and the management of business processes will have a significant impact on the organizational efficiency. At the same time, Industry 4.0 causes additional complexity and high interconnection of business processes, which is a reason for more strict requirements in Business Process Management (BPM).

Process mining is a relatively young discipline that can be seen as a bridge between data science and BPM [9, 10]. Process mining is closely related to business process intelligence (BPI), and these two terms are often used as synonyms [10, 11]. Throughout this paper, the term “process mining” (PM) is used to address both of these concepts. The goal of PM is to improve operational processes, based on event data and process models [9]. In the context of Industry 4.0, PM becomes a very useful tool that can address both big data analytics and business process complexity challenges and can also contribute to the proper management of business processes, production time improvement, mass customization, and the proper use of technology [6].

In this paper, we discuss and analyze the different approaches of PM used for Industry 4.0. The survey is done in the form of a systematic literature review. The ultimate purpose is to get a general picture of the possibilities of applying either existing or newly discovered approaches of PM to all four aspects of Industry 4.0: smart factories, smart products, new business models, and new customer services. During the process, we have identified the main benefits of the developed approaches for different types of business processes and different aspects of Industry 4.0. We have also identified the main research gaps in terms of aspects of application of PM to Industry 4.0 that are not sufficiently covered by existing studies as well as areas of Industry 4.0 that could potentially benefit from PM but that do not get enough attention by researchers. This study can assist practitioners in identifying the most suitable ways to manage, optimize, and automatize business processes in Industry 4.0, and it can assist researchers in the search for directions of future research in the field.

The remainder of the paper is organized as follows. Section 2 provides definitions and general information on PM, Industry 4.0, and the potential significance of PM for Industry 4.0. Section 3 formulates the purpose of the literature review and summarizes the research questions. Section 4 explains the methodology of the research. Section 5 provides the results of the literature review and identifies the research gaps discovered when performing the review. Section 6 summarizes and discusses the results and conclusions that were made during the review.

2 Background and related work

This section provides background information on the two dimensions of the systematic literature review: PM and Industry 4.0. This section also presents an overview of the existing literature reviews in the field and explains which questions are not answered by the existing studies.

2.1 Process mining

Process mining (PM) is a relatively young discipline that aims at discovering, providing insights, diagnosing problems and supporting the automation of tasks by using combination of data (event logs) and process models. PM establishes a link between the actual processes (and their data) and process models, and it becomes a bridge between process-centric and data-centric approaches [9, 10]. PM overcomes the shortcomings of traditional machine learning methods, which are often limited to analysing separate steps in the business process. PM considers the end-to-end business process and offers the possibility to truly “close” the BPM lifecycle [12, 13].

In PM, the analysis and improvement of processes is based on event data, i.e., collected in event logs. Each event in an event log has a number of attributes. The most commonly used attributes are case (instance), activity (operation, action or task) and timestamp [9].

Processes in PM are represented by process models, which can be viewed as an intersection between the PM and BPM domains [11]. Different languages can be used to present process models. Examples are Petri Nets, transition systems, Business Process Modelling Notation (BPMN), Directly-Follows Graphs (DFG), process trees [9, 10], Event-driven Process Chains (EPCs), Hidden Markov models (HMMs) [9].

Traditionally, the following three types of PM were commonly accepted [10-13]:

- Process discovery: it allows a model to be built based on the examples of behaviour stored in the event log, without using any a priori information.
- Process conformance checking: it checks if the real process recorded in the event log coincides with the model of that process. Process conformance checking allows to determine and explain the deviations in process behaviour and to measure the severance of these deviations.
- Process enhancement: it aims at the extension and improvement of an existing process model based on the event log information.

PM can be applied in offline mode (i.e. use historic data on the processes that have already been terminated), or focus on the online settings (i.e. consider event data in the cases that are still running and thus enable predictive analytics) [10].

In 2022, van der Aalst [9] presents a new framework keeping process discovery and conformance checking, and adding the following PM types:

- Performance analysis is used to identify performance problems, e.g. missed deadlines and recurring quality problems.
- Comparative PM focuses on comparing multiple events logs, e.g. for different locations or time periods and searching for differences and commonalities. The observed differences can be explained with root cause analysis.

- Predictive PMis applied to create a predictive model that can be used in a proactive way to foresee and address compliance and performance problems.
- Action-oriented PM is aimed at triggering improvement actions in case of identified inefficiencies in the process. The two main components of action-oriented PM are: a) constraint monitor, which detects inefficiencies, and b) action engine, which generates action instances when needed [14].

Process discovery, conformance checking, performance analysis and comparative PM are backward-looking techniques, while predictive PM and action-oriented PM are forward-looking [9].

In this paper, we will use the framework proposed in [9] for categorizing the papers as it is the most complete and recent one.

2.2 Industry 4.0

The concept of Industry 4.0 was introduced during the Hannover Fair event in 2011. Since then, it has been the focus of attention of researchers and manufacturers and has been adopted by countries around the world [5, 6]. In different countries, the concept of Industry 4.0 has been given different names: the United States uses the names “Industrial internet” and “Advanced Manufacturing”; the European Commission uses the name “Factories of the Future”; and the United Kingdom uses the name “Future of Manufacturing” [15].

Industry 4.0 provides the methodology for transformation from machine-dominant manufacturing to digital manufacturing [16]. Two key concepts at the center of Industry 4.0 are integration and interoperability [17]. Other features of Industry 4.0 mentioned in the literature include: optimization and customization of production, automation and adaptation, value-added services and businesses, human-machine interaction [17], process integration, real-time information transparency, virtual representation of the real world, and autonomy [18]. Although a generally accepted definition of Industry 4.0 does not yet exist [19], in this paper we understand the concept as defined by Ghobakhloo et al. [20]: “Industry 4.0 may translate into the smartization of all manufacturing and value creation processes, involving the digitalization of plants, suppliers, customers, and distribution channels”.

There are four core aspects of Industry 4.0: smart factories, smart products, new business models, and new customer services [5, 6]. New types of factories under the framework of Industry 4.0, called “smart factories”, enable the automatic exchange of information among manufacturing resources, automatic prediction and maintenance of machines, and control of the production process and factory management [5]. New types of products, called “smart products”, convey guidance to the customers and get the customers’ feedback using sensors, identifiable components, and processors [5]. Smart products are also able to manage their own manufacturing processes [6]. The business perspective of Industry 4.0 includes the communication network of the different actors of the business processes (suppliers, customers etc.), which allows real-time optimization and configuration of the behavior based on the status of the other chains in the network [5, 6]. In addition, the customer perspective of Industry 4.0 embraces new purchasing methods and new levels of product customization, including the possibility to correct orders at any moment in time during the production process [5, 6].

All of the advantages of Industry 4.0 become possible due to the technologies that are used. Among the key technologies of Industry 4.0 are Cyber-Physical Systems (CPS) and the Internet of Things (IoT). Scholars lack mutual understanding of the technologies and paradigms associated with Industry 4.0 [19, 21, 22], but the most often mentioned in literature are: Radio Frequency Identification (RFID), cloud-based manufacturing [17], Enterprise Resource Planning (ERP) systems [17, 23], mobile computing, cloud computing [16], 3D printing, augmented reality, smart predictive decision support tools, wearables [23], visualization technologies, cybersecurity, simulation and modeling, advanced robotics, machine learning, artificial intelligence, and big data analytics [18].

2.3 Related work.

A lot of work has already been done by different authors to systemize research papers concerning different aspects of PM application. Fernandes et al. [24] provide an overview of machine learning and PM techniques for process optimization. Their work systemizes the problems raised by the authors as well as the resources to solve these problems, such as tools, methods, approaches, systems, and frameworks. Garcia et al. [25] make a systematic mapping study on PM, including contribution area (research topic), algorithms, and application domains.

Other authors focus on the practical implementation. Chaydy & Madani [26] provide an overview of application domains of PM and explore the applicability of PM to real-life scenarios in healthcare and finance in more details. Thiede et al. [27] present a systematic literature review of the use of PM in organizations; they make a classification based on the organizational perspective, the service perspective, and the information system perspective. Dakic et al. [28] make a survey on applications of PM in different sectors of the economy, including mapping on PM types, industries, business processes (to which PM is applied), and tools and algorithm that are used. Maita et al. [29] aim at systemizing the interdependencies between PM and data mining and make a systematic mapping between PM and data mining methods and techniques. Other studies focus on specific stages in PM, such as the work by Diba et al. [30], which focuses on the techniques used for data extraction and pre-processing for PM.

There is a range of surveys that provide an overview of approaches for real-time process prediction and monitoring. The work by Marquez-Chamorro et al. [31] presents a systematic review of the approaches for real-time prediction of business processes, making a classification of the methods used for predictive monitoring of business processes. It embraces approaches based on machine learning as well as other approaches such as those based on annotated transition systems and statistical methods. Teinemaa et al. [32] and Di Francescomarino et al. [33] present literature reviews of methods of predictive process monitoring.

In general, the above-mentioned papers do not focus on Industry 4.0. Even though some of them superficially touch upon PM in the domain of Industry 4.0, the surveys do not explain the following: a) which types of processes can be supported in Industry 4.0; b) how the proposed approaches deal with the challenges of data processing associated with Industry 4.0; c) what benefits PM can offer to different aspects of Industry 4.0 and different types of processes.

Different surveys have been presented with the goal of outlining the concept and components of Industry 4.0. Lu [17] conducts a review of the content, scope, and findings of Industry 4.0 and systemizes concepts and perspectives, key technologies, applications, and frameworks. Zhong et al. [7] make a comparison of the concepts of Industry 4.0, such as intelligent manufacturing, the IoT-enabled manufacturing, and cloud manufacturing. That work also reviews the main technologies of Industry 4.0 in terms of their characteristics and application areas.

The above works consider the general scope of Industry 4.0, but they do not analyze applications of PM. Even though the work by Zhong et al. [7] highlights the importance of big data analytics, it does not study which solutions are offered, such as which algorithms, tools, and techniques.

Finally, one more group of surveys investigates the intersection between data analytics and Industry 4.0, and presents overviews of the applicability of data analytics to specific technologies, sources of data, or aspects of Industry 4.0. Ye et al. [34], Mannhardt et al. [35], and Praveen Kumar et al. [36] focus on approaches for processing sensor data. Ye et al. [34] look at the problem through the prism of the concept of situations. Mannhardt et al. [35] analyze approaches that are applied to activity recognition in terms of types of recognition algorithms, process context, and types of sensors. Praveen Kumar et al. [36] present a mapping study of the machine learning algorithms that analyze and monitor the activity of wireless sensor networks. Perera et al. [37] present a literature survey of the approaches to context-aware computing for the IoT. It makes a mapping that is based on the whole life cycle of the context data, including context acquisition, context modeling, context reasoning, and context dissemination. The survey also makes a short analysis of the application of context reasoning techniques, including supervised and unsupervised learning, and it identifies the pros and cons for each technique and the situations where they can be applied. Dalzochio et al. [38] and Çinar et al. [39] study the use of machine learning for predictive maintenance in Industry 4.0. Dalzochio et al. [38] investigate challenges related to machine learning and reasoning for failure prediction. They determine the most popular algorithms and study the contexts in the use of ontology. Çinar et al. [39] determine the algorithms that can be used for sustainable smart manufacturing in terms of their applicability to specific equipment. Ramírez-Gallego et al. [40] consider one of the challenges associated with Industry 4.0, stream mining, but they only focus on data pre-processing. Osman & Ghiran [41] give overview of the Industry 4.0 and highlights the role of PM as a strategic key of the Industry 4.0.

Although the studies on specific facets of data analytics in Industry 4.0 (e.g., processing sensor data, activity recognition, context-data inclusion, and processing event streams) can also be relevant to PM, no previous research article has specifically investigated the application of the BPM aspect of PM.

3 Problem statement

In spite of all potential advantages of the application of PM to business processes in Industry 4.0 and the growing amount of literature in both domains (PM and Industry 4.0), to the best of our knowledge, there is no study yet that synthesizes the different applications of PM to Industry 4.0. In this paper, we

perform a systematic literature review, a well-defined methodology to synthesize the existing work ensuring that its results are less likely to be biased and can help to identify robust trends [42].

The goal of this systematic literature review is to get a general picture of the possibilities of using PM in the context of Industry 4.0. The goal is to capture the general structure of the research area, determine the possible and most common approaches of PM and the most common areas of their application in terms of types of business processes and Industry 4.0 aspects, and determine the benefits that these approaches can provide. In addition, the research gaps must be identified. This means determining the areas of Industry 4.0 that could potentially benefit from applying PM techniques but that are still not the focus of researchers.

The systematic literature review can serve practitioners who consider the application of PM to business processes in Industry 4.0 by providing full information on the proposed approaches, which practitioners can use as a starting point to select an approach that is most relevant and beneficial for their enterprise. The systematic literature review can also serve researchers by providing the general overview of the current state of the art in this area. The identified research gaps can be seen as potential directions for further research in the area.

In alignment with the goal of the survey, the main research question is formulated as follows: How is PM used in the context of Industry 4.0?

This research question is subdivided into the following more specific questions:

Research question 1: What types of processes have been supported by PM in the context of Industry 4.0?

PM can become an extremely useful instrument that can correspond to the complexity of business processes within Industry 4.0 and the importance of their proper management [6]. However, a clear understanding is needed about the range of business processes that can be supported by PM and the ways in which process management can benefit from the application of PM. Answering this question can be useful for practitioners planning to use PM in the organization and can help them to identify potential areas of PM application in the context of the business processes specific for their companies.

Research question 2: What types of PM are used and what are their applications?

Analysis of PM types in the context of the business process types that they support should help for better understanding of the role of PM in the real production environment. Besides, focus on the distinction between backward-looking and forward-looking approaches can help to understand how far the current research is on the way to the realization of the potential of the application of forward-looking PM for the business processes in the context of the Industry 4.0.

Research question 3: Which PM algorithms are used and how can they be classified in the context of PM?

PM algorithms and the types of process models are both inevitable parts of PM and they should be described in detail to ensure the integrity of the proposed approaches and their further application. It is

essential to understand whether existing PM algorithms still remain applicable under Industry 4.0 circumstances or new ones have to be developed.

Research question 4: What aspects of Industry 4.0 do PM approaches address, and what are their benefits?

Potentially PM can provide benefits for each aspect (smart factories, smart products, new business models, and new customer services) of Industry 4.0 [6]. In the context of smart factories, the potential of PM usage is connected with improvement of process efficiency, concepts of interoperability and automation technologies. The concept of smart products can benefit from PM through analyzing production data, as well as data on the usage patterns coming from the customers, that can be used in adapting production processes, reaching new levels of product customization and providing customers with recommendations on the best ways to use the product. Concerning new business models, PM can enable companies to use data coming from other parties in the network for thereal-time adaptation and optimization of business processes. In the sphere of new customer services, PM can help to add value to customers through providing products and related services tailored to customers' needs. In our research, we aim to investigate which aspects of Industry 4.0 have been addressed by PM and what still remains to be done.

Research question 5: Which Industry 4.0 technologies are used in combination with PM, and how they are interconnected?

Enabling technologies is one of the key concepts of Industry 4.0. Technologies make all the advantages of Industry 4.0 possible [17]. Using these technologies requires real-time integration and analysis of massive data. Processing this data and converting it into valuable information for business processes is one of the main challenges of Industry 4.0, which is addressed by PM [6]. However, it is important to understand what is the exact role of PM in the collaboration of different technologies, e.g. which technologies can serve as a source of data for PM and which ones can use process data generated by PM.

Research question 6: What gaps/opportunities still need to be addressed in the context of PM in Industry 4.0?

Identifying gaps in current research is one of the main reasons to perform a systematic literature review [42]. By doing this, it becomes possible to suggest directions for further investigation in the area.

4 Research methodology

The study is undertaken in the form of a systematic literature review. The goal of the review is to: synthesize the literature that has applied PM to Industry 4.0, identify the areas within the field that attract most attention of researchers, areas that lack attention (research gaps) and based on thisdiscuss which questions could become central for future research in the field. We are proceeding from the hypothesis that PM can be widely applied in Industry 4.0. However, we aim to provide a

neutral representation of the domain. We are targeting an exhaustive coverage of literature. However, the question remains to which extent a literature review can be fairly exhaustive. Following Cooper [43], we define “exhaustive” as a comprehensive coverage within the limits that we define. To define the area of PM we use the commonly accepted framework defined in [9]. The definition of the area of Industry 4.0 constitutes more difficulties. Taking into account that for the best of our knowledge there is no commonly accepted definition of the term “Industry 4.0”, we use the approach of including only papers where the authors explicitly define them relevant to Industry 4.0. More details are provided in Subsection 4.1. The methodology of the systematic literature review follows the guidelines formulated by Kitchenham & Charters [42]. In order to answer the formulated research questions, first, the primary studies that are relevant to the research topic must be selected based on a systematic approach. Then, the information required to answer the research questions must be retrieved from the papers and analysed.

4.1 Search process

The search strategy involves performing a manual search in the appropriate databases. To choose the databases where the search must be implemented, the following criteria were taken into account: relevance, volume, popularity, and trustfulness. Based on these criteria, the following databases were chosen: Web of Science, ProQuest, Google Scholar, and Scopus. Scopus embraces data from ACM Library, IEEE Xplore, ScienceDirect, SpringerLink, MDPI, IOS Press, and Plos One. Web of Science, ProQuest, and Scopus were accessed through the LIMO[2] search interface of the KU Leuven. In order to process the search in Google Scholar, Papers[3] software was used.

The search string for retrieving the relevant primary studies must address two main aspects of the review: PM and Industry 4.0. To be in line with this requirement, the search string must consist of two parts that are connected with the Boolean operator AND. For the first part of the search string, we chose the following key words: the term “PM”, the term “business process intelligence”, its abbreviation (BPI). Since we specifically investigate how different types of PM techniques are applied, we use the keywords to address PM types. We combine in our search string both frameworks: framework presented in [10-13] and a new one proposed in van der Aalst [9]. First we include PM types as formulated in [10-13] (“discovery”, “conformance checking”, and “enhancement”) and the term “predictive analytics” to address online PM. Then, in order to address the new classification of van der Aalst [9], we extend the search string with the term “performance analysis”. Other types of PM in the new framework contain “PM” in their names and thus are addresses by the term “processes mining” in the search string. The names of the types of PM and the term “predictive analytics” were used in combination with the term “business process” in order to obtain only the papers that are relevant to BPM in the search results. For the second part of the search string to address Industry 4.0,

[2] LIMO: https://limo.libis.be/primo-explore/search?vid=KULeuven

[3] Papers software: https://www.papersapp.com/

we chose the key word “Industry 4.0”, and its synonyms as specified by Büchi et al. [15]: "Industrial Internet", "Advanced Manufacturing", "Factories of the Future", and "Future of Manufacturing".

As a result, the generated search string was the following:

("process mining" OR "business process intelligence" OR BPI OR ("business process" AND (discovery OR enhancement OR "conformance checking" OR "predictive analytics" OR “performance analysis”)))

AND

("Industry 4.0" OR "Industrial Internet" OR "Advanced Manufacturing" OR "Factories of the Future" OR "Future of Manufacturing")

After performing the search in different databases, the first step is to compare the results from different databases and include each paper only once. The following step is to make decision based on reading the abstracts and applying inclusion and exclusion criteria formulated in 4.1.1. The final decision about inclusion of the paper in the final list is made after reading the whole article. The screening procedure was performed by one author, then checked and verified by the second and the third author. Discrepancies were discussed and the common solution was found.

After the primary studies were selected, they were used as the starting set for backward and forward snowballing, which was performed according to the guidelines provided by Wohlin [44]. First backward snowballing is proceeded, which allows to identify papers that have to be included in the final list based on the reference list of the paper from the start set. Then forward snowballing is implemented, which helps to identify relevant papers choosing from those citing the paper in the start set. Snowballing is proceeded in iterations, meaning that papers identified through backward and forward snowballing are put into a pile to use them as a start set for the next iteration. The loop is ended when no new papers can be identified neither through backward nor through forward snowballing.

4.1.1 Inclusion and exclusion criteria

In order to enable a systematic approach for selecting papers that are relevant to the topic of the review and that can contribute to answering the research questions, the inclusion and exclusion criteria must be applied to the retrieved papers.

The main inclusion criterion was the following: The paper contains an approach for using PM in the context of Industry 4.0. Based on this main criterion, several more specific criteria were formulated:

- The sphere of application of the PM is clearly defined;
- The PM approach is clearly defined;
- The role of the approach in the context of the Industry 4.0 is clearly defined.

The exclusion criteria were formulated as follows:

- The study is not written in English;
- The study was published before 2011 (since the term Industry 4.0 was coined in 2011);

- The study is an older version of a newer one and does not provide relevant details that are not presented in the newer paper;
- The goal of the study is not stated clearly.

We focus our research on the practical aspect of the PM in the context of theIndustry 4.0. This means that we include only papers that provide clear information on which processes are supported by PM and how these processes are relevant tothe Industry 4.0. The paper also should clearly define the way in which PM supports the process. Even if the paper does not specify a PM algorithm, it should provide enough information to be able to classify PM type based on the framework described in Subsection 2.1.

5 Results

This section presents the results of the systematic literature review. The data was extracted from the selected primary studies and systemized in several tables. The research questions were answered based on this data.

5.1 Introduction of the results

The application of the search string described in Subsection 4.1 produces the following search results:

- Scopus – 45 papers;
- Web of science – 14 papers;
- ProQuest – 13 papers;
- Google Scholar – 166 papers.

First, the results from the different databases were compared to ensure that each paper was included in the list only once. This process produced a list of 168 papers. Then, the selection process outlined in Subsection 4.1 was applied to each paper in the list, using the inclusion and exclusion criteria formulated in Subsection 4.1.1. This process produced a list consisting of 15 relevant papers (in Table 1 in Subsection 5.2 rows providing information on these 15 papers are highlighted in grey colour). The selected papers were used as a starting set for forward and backward snowballing, as described in Subsection 4.1. This process helped to identify 32 more papers and to obtain the final list consisting of 47 papers.

The selection procedure is shown in the figure 1.

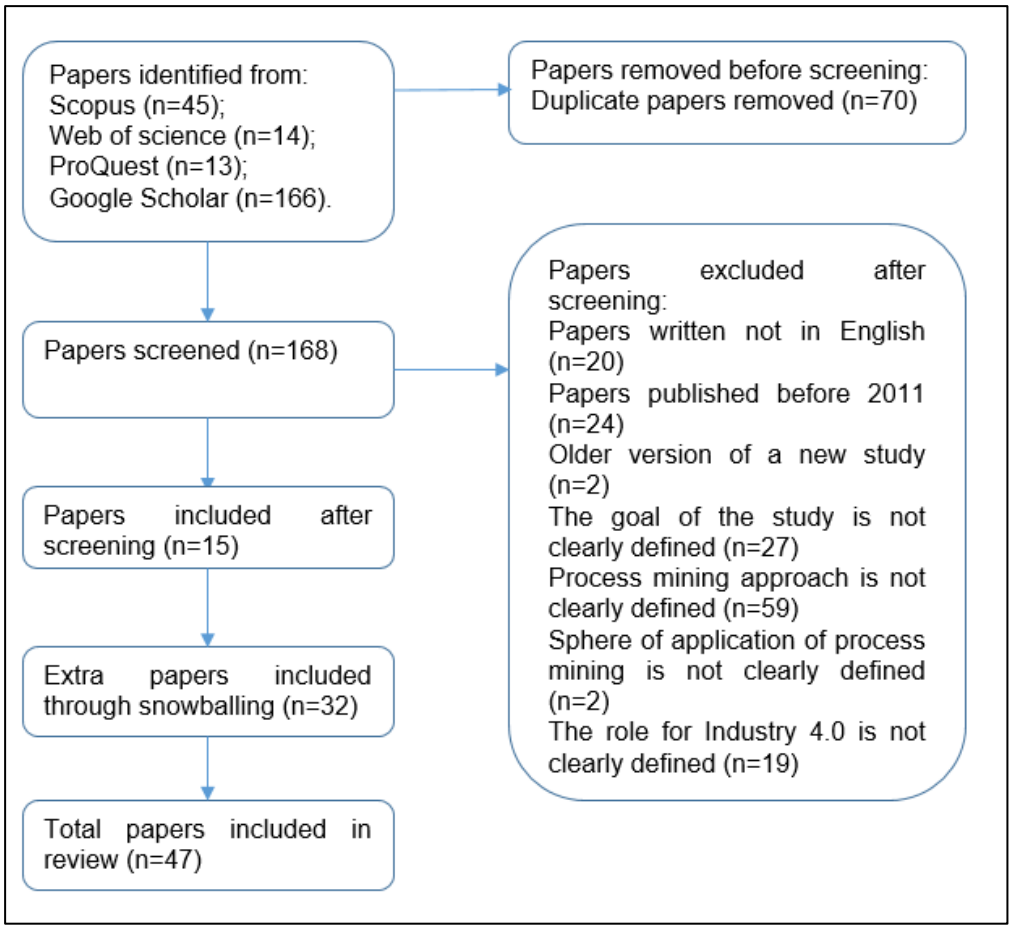


Figure 1. Paper selection procedure

5.2 Synthesis of the results

This subsection summarizes the information extracted from each selected study. The extracted data is grouped into tables. The information that is provided after each table explains the content of the table, summarizes the trends discovered in the data, thereby answering the research questions.

Table 1 gives a full list of the selected papers and their respectiveresearch goal. An ID number is assigned to each paper, which is used in the remainder of this article to identify it.

Figure 2 presents an overview of the application of different types of PM following the framework presented in [9]: process discovery (PD), (online) process conformance checking (PCC(O)), performance analysis (PA), predictive PM (PMM) and action-oriented PM (AOPM) respectively (no paper was found on the Comparative PM type). The figure shows the business processessupported by each type of PM with the number of papers within each PM type(within the circle),and for what the results of PM are applied (outside the circle).

More detailed information on the application of PM to different types of business processes is presented in Tables A-E in the attachment. Each table specifies the type of business process, ID of papers where the support of this process is described, the role of PM and the purpose to apply PM. Based on the data from the tables, Research questions 1 and 2 are answered.

Table 1: List of primary papers(rows that provide information on the 15 papers directly found by the search query are in grey background).

ID	Author	Year	Research Goal
[P1]	Halaška&Šperka	2019	Investigate the performance of an automated process discovery technique
[P2]	Knoch et al.	2020	Present an approach to provide the information assistance for the smart assembly station, including the architecture of the station
[P3]	Aehnelt & Bader	2018	Present technologies for recognizing manual activities in the manual assembly process
[P4]	Knoch et al.	2020	Provide a solution to extract information on the process execution
[P5]	Bauters et al.	2018	Introduce a technology to automate the analysis of the process
[P6]	Hutabarat et al.	2016	Combine 3D technologies with simulation to improve decision making
[P7]	Zhuang et al.	2021	Propose an implementation framework for the construction and application of the shop-floor digital twin
[P8]	Zhang et al.	2015	Introduce an architecture of the real-time intelligent navigation of the assembly station to facilitate the real-time process monitoring and control
[P9]	Zhang et al.	2016	Introduce a real-time production performance analysis and exception diagnosis model
[P10]	Fang & Zheng	2020	Introduce a method to optimize the manual assembly planning based on the shop-floor data
[P11]	Roitberg et al.	2015	Present an approach for monitoring and interpreting human activities to improve the efficiency of human-robot interaction
[P12]	Lenz et al.	2011	Present a method for the workflow analysis of an assembly task jointly performed by a human and a robotic system
[P13]	Bleser et al.	2015	Introduce a smart assistance system with emphasis on the algorithms
[P14]	M. Cramer et al.	2018	Introduce an approach to the recognition of human activities in human-robot collaborative processes
[P15]	Roldán et al.	2019	Present a training system for industrial operators in assembly tasks
[P16]	Saraeian & Shirazi	2020	Present an extension of the uncertain BPM system architecture for anomaly detection
[P17]	Myers et al.	2018	Present a novel PM anomaly detection method
[P18]	Setiawan & Yahya	2018	Propose a novel PM technique for human activity detection

ID	Author	Year	Research Goal
[P19]	Tello-Leal et al.	2018	Propose a methodology to use a Long Short-Term Memory Recurrent Neural Network for the prediction of the execution of cases in Industry 4.0
[P20]	Mehdiyev et al.	2017	Operationalize industrial IoT driven predictive and prescriptive analytics
[P21]	Repta et al.	2015	Introduce a semantically enabled system to monitor and analyze flexible processes in manufacturing environments
[P22]	Véjar & Charpentier	2012	Propose a method for the generation of simulation models in production environments
[P23]	Sánchez et al.	2020	Introduce a supervisory system to enable self-coordinated manufacturing processes
[P24]	Brzychczy et al.	2020	Investigate the application of PM to improve operational efficiency in mining industry
[P25]	Ribeiro et al.	2018	Case study to apply PM to the productive process
[P26]	Malburg et al.	2020	Present a factory simulation model for conducting research in BPM and IoT
[P27]	Seiger et al.	2020	Present an approach to derive process events in smart factory context
[P28]	Meroni et al.	2018	Propose an approach to autonomously and continuously monitor multi-party processes using the concept of smart products
[P29]	Ehrendorfer et al.	2019	Use case to analyze the manufacturing data collected on different levels of the automation pyramid
[P30]	Stertz et al.	2020	Provide algorithms to analyze the root cause of a concept drift using sensor event streams
[P31]	Stertz et al.	2020	Propose an approach to quantify the conformance deviation cost in conformance checking
[P32]	Dunkl et al.	2015	Present a method for decision point analysis for exploiting time series data
[P33]	Rodriguez-Fernandez et al.	2021	Present an approach for conformance checking for time-series data
[P34]	Rebmann et al.	2019	Present an approach for capturing manual processes based on the data obtained from IoT
[P35]	Rebmann et al.	2020	Present an approach for the real-time recognition of robot activities
[P36]	Rigger et al.	2020	Propose a method for the continuous process lifecycle management (PLM) improvement
[P37]	Es-Soufi et al.	2017	Present a semi-automated approach for decision making
[P38]	Park et al.	2015	Propose a way to analyze manufacturing processes in make-to-

ID	Author	Year	Research Goal
			order industries
[P39]	Yang et al.	2014	Present a manufacturing data analysis system
[P40]	Lugaresi & Matta	2021	Present a method to automatically discover manufacturing systems and generate digital twins
[P41]	Farooqui et al.	2019	Present an approach to capture data from robots and generate models of behavior
[P42]	Koehler & Jing	2020	Introduce a method to evaluate an activity log in industrial automation domain
[P43]	Lugaresi et al.	2019	Present a method to generate the model of a manufacturing system
[P44]	Viale et al.	2011	Present a methodology to model large-scale manufacturing processes
[P45]	Ruppert & Abonyi	2020	Present a solution for the model-based evaluation of operator performance
[P46]	Dasani et al.	2015	Propose a workflow conformance method to continuously monitor operator actions
[P47]	Hu et al.	2018	Demonstrate how workflow mining can be adopted for alarm management to learn how operators respond to alarms

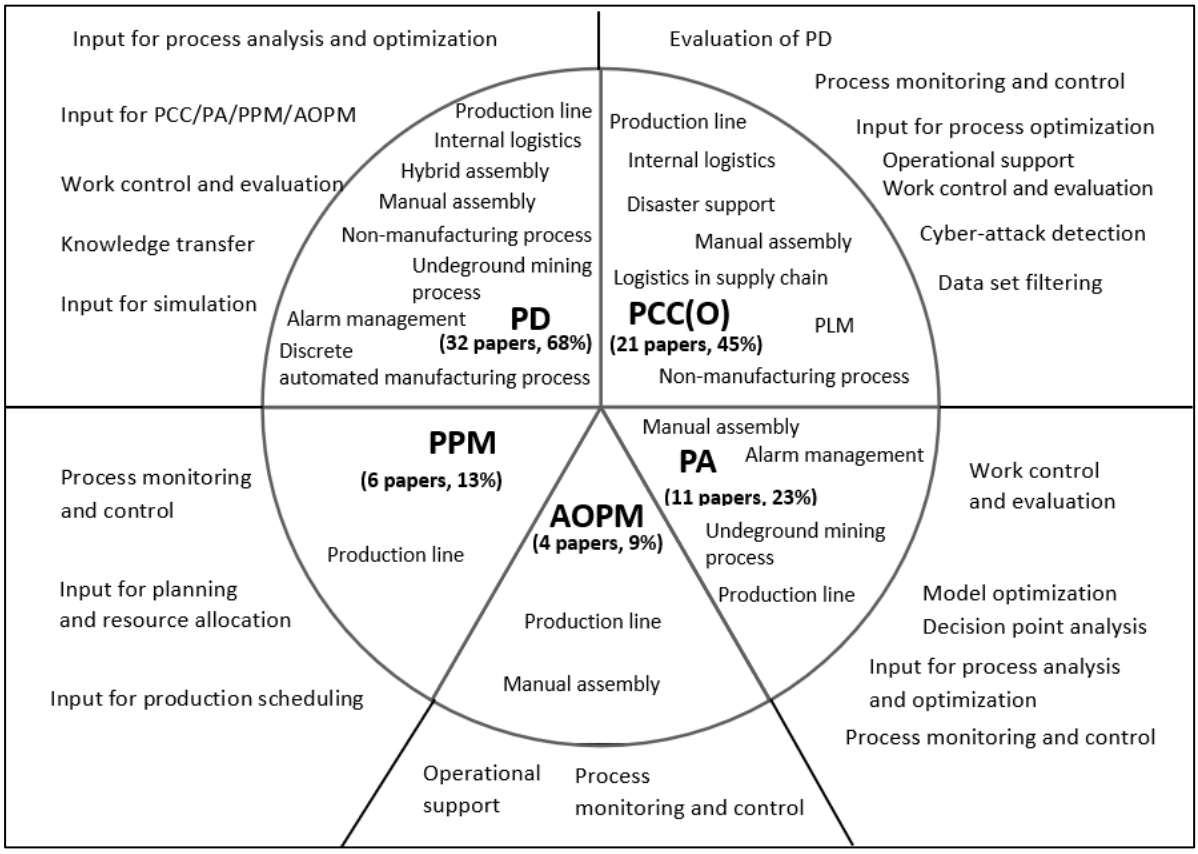


Figure 2. Overview of PM types, purposes of applying each type, and addressed processes

As Figure 2 shows, PM can be applied to different processes and for different purposes (see Tables A-E in the attachment for more details). Table 2 gives a definition of each identified process and shows by how many papers the process is addressed.

Table 2: Summary of types of business processes

Type of process	**Description**	**Number of papers (with %)**
Manual assembly process	Processes which involve manual activity that is performed by operators in order to fulfil a specific part of a production process	12 (26%)
Hybrid assembly process	Processes jointly performed by a human and a robot	3 (6%)
Production line	End-to-end manufacturing processes of a specific product. In some papers, only the production process itself is included; in others, order processing, planning, post-processing, and/or design are also included	23 (49%)
Discrete automated manufacturing process	Processes which involve fulfilling a part of the production process without human intervention	2 (4%)
Underground mining process	Mobile mining asset operations	1 (2%)
Non-manufacturing process	Processes which are not directly linked to the production of a product (e.g., in [P46], it is the start-up of a boiler)	3 (6%)
Product lifecycle management (PLM)	Processes which involve the management of product information, engineering workflows, and related applications	1 (2%)
Internal logistics	The movement of goods within an enterprise	5 (11%)
Logistics in the supply chain	Logistics process which also includes other parties, like suppliers and customers	1 (2%)
Alarm management	These processes start with an alarm notification (e.g., about gas network exploitation) and consist of actions that must be taken to prevent or mitigate the consequences of an accident	2 (4%)
Robot-assisted disaster support	These processes are implemented in case of accident	1 (2%)
Total number of papers		47 (100%)

We can see from Table 2 that the most commonly addressed processes are production line and manual assembly processes. The most commonly addressed types of PM are process discovery and process conformance checking. Process discovery, sometimes in combination with performance analysis is mostly used to: a) derive models for process analysis and optimization, b) as an input for other types of PM or c) for knowledge transfer from more experienced operators to trainees. Challenges for process discovery arise either from the character of the process itself (flexible process, process with multiple instances running in parallel) or from the character of the data (coming from

video, sensors, virtual reality, multiple sources, unstructured data, infinite amount of data, incompleteness of data).

Process conformance checking can be applied in offline or online mode. Offline conformance checking is mostly applied to a) evaluate the quality of the model retrieved by process discovery or b) for a cyber-attack detection. Online conformance checking, often in combination with performance analysis or/and predictive PM is mostly applied to a) control and monitor processes or b) to provide support to operators, performing manual tasks. The purpose of conformance checking is to identify deviations (either in the life process in comparison with the model or in the model in comparison with the real process). Dependent on the goal of the project, deviations identified by conformance checking, serve for one of two purposes. The first purpose is relevant for the projects that propose a) systems of autonomous monitoring of the process or b) systems for the automated support for the process implementation. In this case each identified deviation in real-time serves as a trigger for the action (either in the form of the recommendation for a person implementing the process or in the form of an action performed by the system). The second purpose is relevant for projects which focus on the analysis of processes and their performance. Here not each individual violation is important, but a general cost of violations which serves to evaluate either the process itself or the quality of the model of the process that was discovered in the previous step.

Performance analysis, predictive process monitoring and action-oriented PM are applied much less than process discovery and process conformance checking, and comparative PM is not applied. Performance analysis is mostly seen as an additional step after applying process discovery or process conformance checking to provide additional insights on the process. The most common application of predictive PM is process monitoring and control. For manual assembly processes action-oriented PM triggers instructions or recommendations for the operator. For production lines it triggers automated actions that directly affect the process, e.g. automatic repairing of the process in case of the predicted failure [P23].

In order to answer Research question 3, the algorithms or approaches used for PM are analysed. Figure 3 presents summary of the identified algorithms, classifying them according to the different PM types. Details on these techniques are presented in Tables F-Jin the attachment.

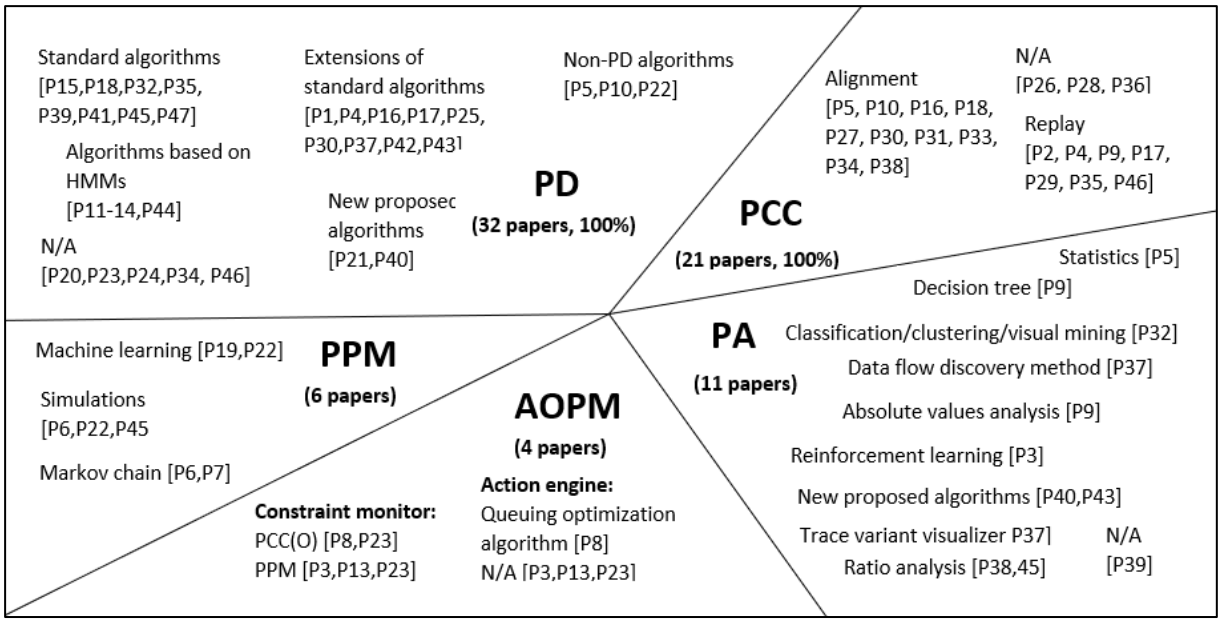


Figure 3. PM algorithms/approaches per PM type

Within process discovery (see Table F in attachment), most of the papers use either one of the standard process discovery algorithms (25%) or an extension of a standard process discovery algorithm (28%). Five algorithms are based on hidden Markov models (HMMs): [P11, P12, P13, P14, P44]. Although HMMs are relevant to PM, HMMs-based methods are considered to have some complications regarding both discovering models and representation of models to the end user [11]. However, extensions of HMMs-based methods have shown their efficiency in discovering flexible processes (e.g., manual assembly processes), since using the distribution of probabilities of the next possible step instead of specifying the order of activities explicitly reflects the flexible structure of the process. Three papers use anon-standard process discovery approach, namely work cycle classification. The work cycle classification represents an approach for determining working cycles based on the trajectories of movement and activities assigned to each working zone.

Process conformance checking approaches (see Table G in attachment) are distributed between replay (33%) and alignment (52%). Performance analysis (see Table H in attachment) is implemented by the use of statistics, calculating ratios or absolute values (cycle time, real-time progress, real-time cost), machine learning methods, decision trees. Several papers also use new proposed algorithms. Within predictive PM (see Table I in attachment) possible approaches are simulations, Markov chain models and machine learning approaches. The approaches to action-oriented PM consist of two parts, as specified in Park & van der Aalst [14]: a constraint monitor and an action engine. In the selected studies (see Table J in attachment), constraint monitors are presented by other PM types: process conformance checking and predictive PM. Techniques for action engines in most cases are not defined. To sum up, while process discovery mostly uses domain-specific process discovery techniques, process enhancement, performance analysis and predictive process monitoring apply techniques from other domains.

Table 3 provides information on the models that are used by the PM techniques.

Table 3: Types of models

Type of model	Paper ID	Comment
Petri Net	[P1, P4, P17, P18, P21, P25, P29, P37, P39, P40, P43]	
Extended stochastic high-level evaluation Petri net (ESHLEP-N)	[P7]	Petri Net variation
Timed Transition Color Petri Net (TTCPN)	[P8]	Petri Net variation
Hierarchical-timed-colored Petri Net (HTCPN)	[P9]	Petri Net variation
Workflow net with time series (TSWF-net)	[P33]	Petri Net variation
Transition system	[P5, P10, P15, P22]	
BPMN	[P26]	
BPMN and CMMN Hybrid Model	[P2]	
Extended-Guard-Stage-Milestone notation (E-GSM)	[P28]	
New proposed model for alarm management	[P47]	
DFG	[P23, P24, P35]	
Other types of process representation:		
HMMs	[P3, P13]	
Hierarchical HMMs	[P11]	
Composite HMMs	[P12]	
Dynamic Bayesian Network	[P14]	
Markov chain	[P6, P7]	
N/A	[P16, P19, P20, P27, P30, P31, P32, P34, P36, P38, P41, P42, P44-46]	

As Table 3 shows, the most popular way of process representation is Petri Nets or Petri Net variations (15 papers out of 47 use models of this type). Some papers use less advanced techniques: DFG (used in 3 papers) or a transition system, which is the simplest way of processing notation, and which represents a process as a sequence of state changes (used in 4 papers). Two papers specify process models of a higher level of abstraction: BPMN [P26], Hybrid model based on BPMN, and Case management model and notation (CMMN) [P2].

Two papers focus on the development of new types of models for process representation. In [P28] the necessity for a new model is explained by the necessity to have a declarative model of the process

(which only requires the identification of conditions that determine when an activity should be executed) to be used in autonomous monitoring of the process. In [P47], a new model notation is presented focused on decision making in a specific area –alarm management. Seven papers present a process using HHMs and their variations, which is not considered to be a standard modeling language. HMMs are mostly used to present manual activities.

In order to answer Research question 4, we identified which Industry 4.0 aspects are addressed by each paper, as specified in Subsection 2.2 (smart factories, smart products, new business models, and new customer services), as well as the contribution of each paper in the context of Industry 4.0.Figure 4 summarizes this information, which is presented in details in Table K in the attachment.

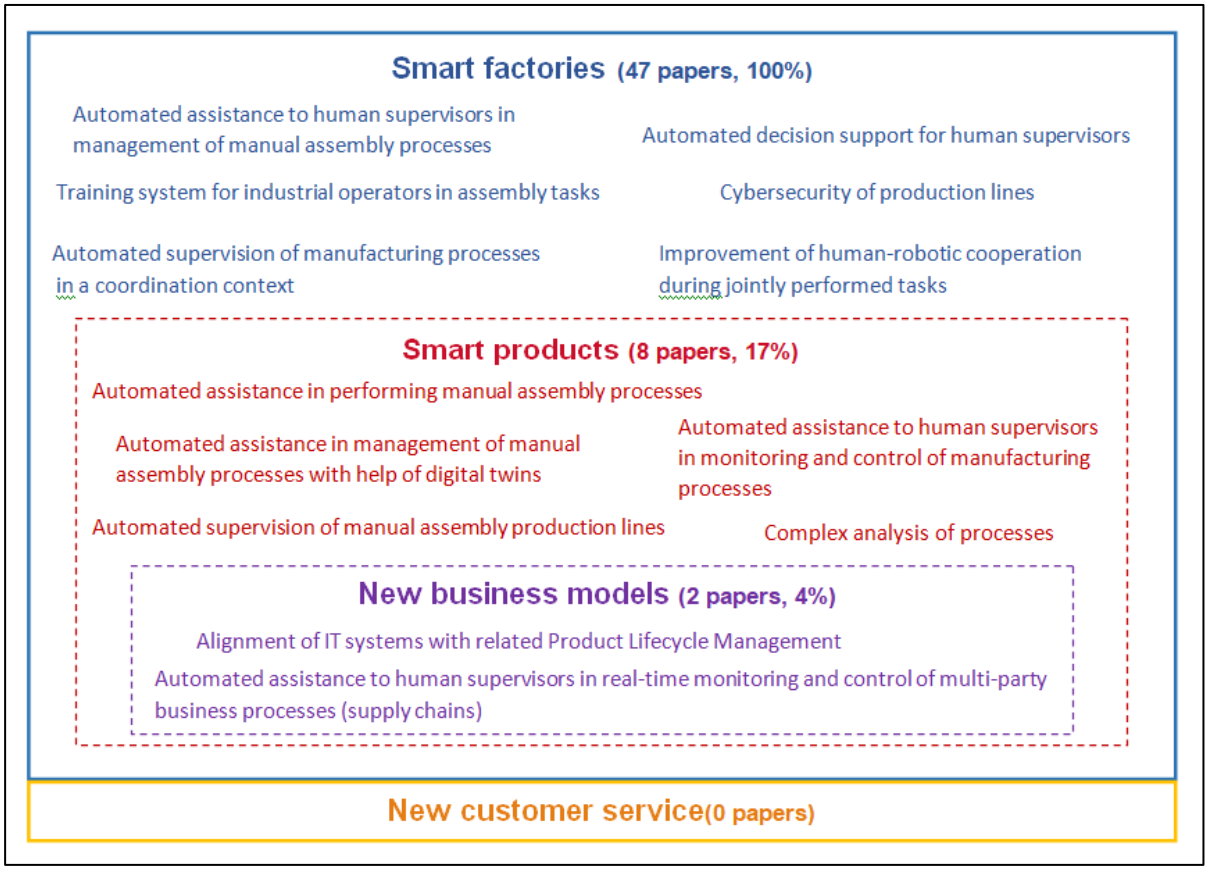


Figure 4. Industry 4.0 specific contribution

The most commonly addressed aspect of Industry 4.0 is smart factories as it is addressed in all papers. Eight papers also address the smart product, and only two papers address new business models. New customer service is not addressed. In the context of the above-mentioned papers, smart product refers to a product that is equipped with chips [P1, P27, P45] or with tags [P7, P8, P9, P22, P28] that allow its location to be tracked during the production process or during the logistics process. In the case of production processes, its location is tracked to identify the level of its completion. New business models are represented by monitoring supply chain logistics [P28] and PLM [P36]. However, the concepts of smart products and new business models in the context of Industry 4.0 have a much broader scope and involve a new level of collaboration with different actors, such as suppliers and

customers, a new level of product customization (including the possibility to change requirements and adapt products in real-time), and the capability of a product to control its own production cycle.

One direction of the contribution of PM in the context of Industry 4.0 is the assistance to humans in either the implementation or supervision of the process: assistance in performing the manual assembly process, assistance in process monitoring and control, assistance in the management of the process. By ensuring a real-time information transparency PM supports decision-making and improves human-machine interaction. PM provides assistance in performing manual assembly processes [P2, P3, P5, P8, P13] by tracking the implementation of the process performed by an operator and giving recommendations on the next step. This can be done in the normal course of the process and in case of detected deviations. The main challenge here is to address the flexible nature of manual processes, as not all activities that deviate from the process model are incorrect. Automated assistance in process monitoring and control [P7, P9, P19, P22, P26, P27, P28, P30, P31, P44, P46], or decision making [P6, P20, P32, P37, P40, P43, P47] is provided by monitoring and analyzing the progress and performance of ongoing or already completed processes. To achieve this, the conformance between an ongoing process and the model of the process is checked in real-time. Deviations are identified in terms of deviated activities or deviated time to perform an activity. Performance analysis can also be performed in real-time. One more way to assist process supervisors is by making predictions on the further flow and performance of the process. Besides, recommendations can be given on the most appropriate decision in the current situation based on the decision point analysis. This way process supervisors are provided with full information on the progress of the process, predictions and with guidelines for decision making. Automated assistance in the management of manual assembly processes [P4, P5, P10, P18, P34, P45] consists of providing the process supervisor with information on the performance of operators, which allows to evaluate the operators' work, and also to find possibilities for process optimization. One more PM contribution in improvement of human-machine interaction is improvement of a human-robotic cooperation. This is assisted by a clearly documented representation and a deep understanding of human behavior [P11, P12, P14] or the behavior of robots [P35] in jointly performed processes. The model of the human behavior is discovered, analysed and downloaded in the robot or, other way around, the behavior of the robot is discovered, analysed and, if necessary, adjusted.

Other direction how PM contributes in Industry 4.0 is ensuring automated supervision of processes. This means that the control of the process and the adjustment of its parameters are implemented by the system itself, without human intervention [P3, P8, P23]. The same as with an automated assistance in process monitoring and control, conformance checking of the ongoing process and predictions are made in real-time. Based on this, decisions are made by the system on the necessary steps. This serves as trigger for actions of the system. Besides, the solution involves the continuous improvement of the decision-making process based on the previous experience. This way, a completely autonomous self-supervised process is ensured.

One more application of PM is complex analysis of processes [P1, P21, P24, P25, P29, P33, P38, P39, P40, P41, P42]. This serves as a tool to understand complex integrated processes, find

inefficiencies, make predictions, and look for possibilities for process improvement. The process analysis in this case is implemented by domain experts. Even if the authors of the papers do not specify this explicitly, it seems possible that the knowledge of the process and the improved process models can be subsequently used in one of the scenarios described in the previous paragraphs.

In order to answer Research question 5, Table 4provides information on Industry 4.0 technologies that are used in combination with PM.

Table 4: Industry 4.0 technologies

Industry 4.0 technology	Link with PM
CPS: [P2-5, P8, P10] – Smart assembly work station [P9, P21, P35, P37] – Automated process monitoring and assistance system/Decision support system [P13] – Smart user-assistance system [P17, P46] - ICS [P20] – Real-time predictive analytics system (iPRODICT) [P3, P8, P23] – Autonomous supervisory system [P26, P27, P43] – Learning factory (manufacturing system model) [P47] – Distributed Control System	PM is implemented by one of the modules of SPC and constitutes an important part of the functionality of the whole system.
IoT: [P19-21, P24, P26, P27, P30-33, P35, P37, P42, P43] [P7-9,P22, P28] – RFID [P2, P3, P10, P12, P13, P18, P34] – Human activity recognition (HAR) system	Serves as a source of data coming directly from the work area.
Vision system: [P4, P5, P6, P11, P14]	Serves as source of data coming directly from the work area.
Visualization technology: [P3, P5, P20, P24, P35, P39, P41]	Provides visual representations of PM results to specialists.
Simulation: [P1] – Agent-based simulation [P6, P43] – Discrete-event simulation [P7, P35, P40, P45] – Digital twin [P10] – Simulation environment with digital human modeling [P26, P27] – Learning factory	Source of data for the analysis of the simulated process by means of PM [P1, P26, P27, P35]; Modelling of a planned process for further conformance checking [P10]; Visualization of processes discovered or updated by means of PM [P6, P7, P40, P43, P45].

Industry 4.0 technology	Link with PM
Information system: [P16] – BPMS [P17, P47] – SCADA [P20, P25, P36, P39] – ERP [P25] – IPSS [P29, P38] – Process aware information system [P39, P40, P44, P45] – MES [P39] – Data analysis system	Serves as a source of data for PM.
Robotics:[P7, P11, P12, P14, P27, P30, P31, P35, P41, P42]	Is involved in automated processes that are supported by PM.
Machine learning/data mining:[P2-8, P11, P13, P14, P20, P23, P24, P29, P34, P35]	An element of event abstraction procedure [P2, P3, P4, P5, P6, P11, P13, P14, P24, P34, P35]; A technique for implementing PA, PPM or AOPM.
Cybersecurity:[P16, P17]	Is supported by PM.
Product equipped with chips:[P1, P27, P45]	Serves as a source of data coming directly from the work area.
Augmented reality: [P13]	Provides a visualization of PM results to provide assistance to users.
Virtual reality: [P6, P15]	Source of data for PM; Visualization of PM results.
Human-robot interaction (HRI) system: [P11, P12]	Source of data for PM; Uses output of PM to improve human-robotic cooperation.
Cloud computing: [P16, P19, P28, P30, P31, P36]	Is used as an infrastructure.

Since PM can collaborate with different types of Industry 4.0 technologies, different types of interconnections between these technologies and PM are possible (see Table 4).

For PCS, e.g. smart assembly station, or autonomous supervisory systems, PM is implemented by one of its modules, which constitutes an important part of the functionality of the whole system. By means of PM the implemented process is compared with the model of the process in real-time. Dependent on the current situation and possible deviations in the process corresponding actions are taken. The actions can be in the form of recommendations to the operator on the next step in the process (e.g., in case of smart assembly stations) or in the form of automated actions taken by the system itself (e.g., in case of autonomous supervisory systems).

IoT and vision technologies serve as a primary source of raw data coming directly from the working area. They can collect data from both automated and manual processes, and in the latter case can be represented by static HAR systems [P2,P3, P4, P5, P11, P14], portable HAR systems [P10, P13,

P34], combined HAR systems [P12], or smart wearables [P18]. Data is also often collected making use of RFID technologies, which allows the location of the object and other data on the product at different moments of time to be identified.

Different types of simulations can provide several functions. It can serve as a source of data for PM. In this case simulations are mostly used for training and research purposes. Also, it can be used for modeling of planned processes and visualization of processes. This type of simulations provide assistance in process monitoring and decision making. Especially, a valuable tool in this sense, is a real-time simulation of an ongoing process [P6, P7, P40, P45].

Process aware information systems serve as a source of data for PM. Information systems, e.g. ERP systems, are considered to be one of the Industry 4.0 technologies, as they contribute in two key features of Industry 4.0: integration and interoperability, by aggregating, systemizing and process-centric analysis of data collected on different levels of complex integrated processes. PM represents in this sense the next step in process analytics, and provides an important extension to information systems functionalities.

Robotics is mentioned in 10 studies as one of the technologies involved in automated processes supported by PM. Moreover, 4 of these papers are devoted to advantages that robotics can bring in case of processes jointly performed with humans [P11, P12, P14, P35]. In this case, PM is used for the improvement of human-robot cooperation.

Machine learning and data mining are components of PM, as was described in subsection 2.1, and in this sense are present in each of the reviewed papers. Table 4 shows machine learning and data mining only if it is either explicitly mentioned as a PM technique, or is used next to PM. Machine learning is used as a technique for performance analysis, predictive PM or action-oriented PM. For action-oriented PM it serves as an action engine [P8]. Machine learning can also be used in data pre-processing for PM as an element of the event abstraction procedure.

To sum up, the most common ways of cooperation between PM and other technologies are:

- Industry 4.0 technologies serve as a source of data for PM (IoT, vision systems, information systems, products equipped with chips);
- Technologies are used as a part of the PM solution or for the representation of results (machine learning, visualization technologies, simulations);
- Technologies use the output of PM for further actions (CPS, robotics, cybersecurity).

5.2 Research gaps and future research

When the systematic literature review was performed, several research gaps were identified. Research gaps were found in both of the dimensions of the literature review (PM and Industry 4.0) and can serve to identify possible directions for further research in the field.

The first research gap refers to the types of PM applied to support different business processes. We provide the classification of PM types according to the framework presented in [9]. As Tables 1 and 3

show, the majority of papers propose approaches for the application of process discovery or process conformance checking according to both classifications. The application of the new types of PM presented in van der Aalst [9] get much less attention. The new types of PM can have wide application in Industry 4.0. For example, predictive PM can provide a valuable input for real-time monitoring of processes, either by humans or by automated systems. Action-oriented PM can be a part of a CPS and support self-coordinated processes. Future research could focus on further investigating possibilities and advantages of the application of the new types of PM to different types of processes in Industry 4.0.

The second research gap relates to the usage of algorithms to discover flexible processes, such as manual activities of operators. As we can see from Table F in the attachment, many papers use HMMs or work cycle classification methods. These methods, although relevant to PM, are not standard PM approaches. There extensive usage can be an indicator of the lack of standard PM approaches that can be applied to processes where the sequence of steps can vary from one execution to another. Human-machine interaction is one of the key features of Industry 4.0. This interaction can be implemented e.g., in the form of processes jointly performed by humans and robots, by usage of smart assistance systems providing the operator with guidelines about the next step based on the current situation. For these purposes a model is needed that is clear and precise, but also reflects the flexible nature of the process performed by humans [45]. Methods based on HMMs, although successfully implemented in some papers, still have some shortcomings, like the lack of a model representation that is clear and understandable for end users [11]. As such, further work can be done on investigating whether and how standard PM approaches can be used to discover manual processes, or on the inclusion HMM methods in standard PM approaches (e.g., discover new ways to present process models, inclusion of the methods in PM tools, etc.).

The third research gap concerns the aspects of Industry 4.0. In 38 out of 47 papers, the smart factory is the only aspect of Industry 4.0 that is addressed. Only eight papers address smart products [P1, P7, P8, P9, P22, P27, P28, P45], and only two papers address new business models [P28, P36]. New client services are not addressed. Moreover, the dimensions of the concepts of smart products and new business models, which are addressed by the proposed approaches, are limited to a very narrow scope. PM toolkits can facilitate the implementation of these concepts in the broad sense. For example, PM could support the interaction with clients and suppliers by adjusting processes according to changing requirements and feedback from the customers. It can apply complex approaches to process analysis, monitoring, and control taking into account all the business actors involved and all the phases of the product lifecycle. It can enable the product to control its own production cycle by aggregating the information on product requirements and on the steps of the manufacturing process for each specific product. There is enough evidence to indicate that the potential of PM for the concepts of smart products, new business models, and new customer services has not been exploited enough. Future research could be done to discover the ways these concepts can benefit from the application of PM.

Other possible directions for future research include the validation and evaluation of the proposed approaches in more realistic working environments. For example, the approaches proposed in [P1, P26, P27, P35] were validated using data from simulations. Testing these approaches on real-life processes could contribute a great deal in evaluating their efficiency and applicability to Industry 4.0.

6 Discussion and conclusions

In this study, we have investigated the proposed approaches for the application of PM to Industry 4.0. This study is presented in the form of a systematic literature review. The main research question, “How is PM used in the context of Industry 4.0?”, is subdivided into six more specific questions. The systematic literature review has shown that PM has a wide range of applications in the context of Industry 4.0. It can serve to support processes of different types, different levels of automation, and different scope. The most frequently addressed types of processes are production lines and manual assembly processes. PM provides a better understanding of processes, including the complex processes in vertical and horizontal integration. For processes carried out by humans, PM provides partial assistance in the process implementation. For processes supervised by humans, PM provides assistance in process analysis, monitoring, control, and decision making. The use of PM also helps to increase process efficiency and process automation support, including process implementation and supervision. In addition, it facilitates the cooperation of the different actors that are involved in the process implementation, including humans and robots. Most of the proposed approaches address only one aspect of Industry 4.0 - smart factories. Several papers also address smart products and new business models, but with only a narrow understanding of these concepts. PM is applied in combination with other Industry 4.0 technologies. It uses data from such technologies as IoT and vision technologies as input and provides output for the technologies that imply the automated supervision and control of processes or are used by humans as assistance for implementing, analysing and controlling processes (e.g. simulations and smart user-assistance systems).

The identified research gaps can serve as directions for future research. Within the PM dimension, future research can focus on the application of the new types of PM defined in [9]: performance analysis, predictive PM, comparative PM and action-oriented PM. Also, more research can be done to analyse the applicability of process mining algorithms (or to develop new ones) to discover manual processes. Within the Industry 4.0 dimension, future research can focus on exploiting the potential of PM for aspects of Industry 4.0 other than smart factories (i.e., smart products, new business models and new customer services).

The main threats arising from performing a systematic literature review include incompleteness of the search string, selection bias, inaccuracy in data extraction and analysis, and inconsistency in conclusions. To ensure that the search string was as complete as possible, the selection of the key words was based on the study of peer-reviewed literature sources that are relevant to the subject of the review. However, the part of the search string referring to the Industry 4.0 may not contain all possible terms relevant to Industry 4.0. The literature on Industry 4.0 is very broad. At the same time, for the best of our knowledge, there is no yet a common understanding on what exactly the definition

of Industry 4.0 is. Different authors propose different approaches to derive and classify features of Industry 4.0 and its components. For example, Industry 4.0 can be characterized based on the characteristics clusters (context awareness, modularity, heterogeneity, compositionality, interoperability) [46], based on the design features (interoperability, virtualization, decentralization, real-time capability, service orientation, modularity) [46, 47], aspects of Industry 4.0 (smart factory, smart product, new business models and new customer services) [5, 6], intelligence levels (control, integration, intelligence) and automation levels (process, machine, factory) [5], basic components of Industry 4.0 (digitalization and integration of vertical and horizontal value chains, digitalization of product and service offering, digital business models and customer access) [16]. A quite wide-spread approach to characterize Industry 4.0 is based on its enabling technologies [15-17, 18, 46, 48,49]. However, for the best of our knowledge, there is again no agreed list of Industry 4.0 enabling technologies in the literature [19, 21, 22]. Thereby, extending the search string based on any of the above mentioned approach cannot give an exhaustive list of relevant papers. That is why we consider the approach based on including only the term “Industry 4.0” itself and its main known synonyms as specified by Büchi et al. [15].

To ensure that most of the relevant studies were considered while performing the systematic literature review, we implemented the search procedure in well-known data sources, including Web of Science, ProQuest, Google Scholar, and Scopus. Moreover, we also applied backward and forward snowballing. To ensure the accuracy of the selection process, the process of performing the systematic literature review was completely double-checked by the second and third authors of this paper.

We base our classification of the benefits of the PM in the context of Industry 4.0 on the framework that represents four core aspects of Industry 4.0: smart factories, smart products, new business models, and new customer services [5, 6]. However, as we mentioned above, the literature provides other approaches to define the framework of the Industry 4.0. We have selected this one because it is very comprehensive and relevant to the subject of our study. At the same time, as a direction for further research, information of the primary studies can be classified and analysed according to other dimensions of Industry 4.0.

We base our PM classification on the framework presented in [9]. Most papers still use the framework presented in [10-13] for defining PM types. The new types of PM proposed in [9] (performance analysis, predictive PM and action-oriented PM) are applied in several papers, but mostly are not associated with PM types and are seen as a next step after applying PM. Techniques that are used for these new PM types are not specific PM techniques and are taken from other domains, e.g. machine learning or statistics. Several papers use complex multi-step approaches where the output of one step is used as input for another step. For example, the process is checked for compliance with the existing process model and when discrepancies are found, it becomes a trigger for further actions performed by the automated system. In our opinion, it falls under the definition of process-oriented PM according to the framework presented in [9]. However, in older papers authors define the approach as process conformance checking. According to the definition in Park & van der Aalst [14], we define process

conformance checking in such cases as the constraint monitor for the action-oriented PM. Identifying the technique that serves as an action engine in most cases is not possible. Further work can be done on the elaboration on techniques used for the new PM types defined in [9] and examples of their application.

The findings of this systematic literature review can assist practitioners when making decisions on the application of PM to Industry 4.0, and it can assist researchers when identifying directions for further research work in the field. Although this study focuses on the specific scope of PM applications in Industry 4.0, the findings of the study can serve as a starting point for considering PM applications in other domains (e.g., in healthcare or education). Examples of such findings are the following: applications of PM to process analysis and monitoring based on data from technologies such as IoT; applications of PM to capturing and analyzing processes carried out by humans; application of PM to improving human-robot cooperation; and possible interconnections between PM and different technologies (e.g., CPS, IoT, simulations and machine learning).

Acknowledgements

This work was funded by the FWO (FondsWetenschappelijk Onderzoek) research project fundamental research G0B6922N and partially supported by the Ministry of Science and Innovation under the project VARIATIVA (PID2021-128695OB-I00).

References

[1] Lasi, H., Fettke, P., Kemper, H. G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business and Information Systems Engineering*, 6(4), 239–242. https://doi.org/10.1007/s12599-014-0334-4

[2] Mehta, P., Rao, P., Wu, Z. D., Jovanović, V., Wodo, O., & Kuttolamadom, M. (2018). Smart Manufacturing: State-of-the-Art Review in Context of Conventional and Modern Manufacturing Process Modeling, Monitoring and Control. In *Manufacturing Equipment and Systems. ASME 2018 13th International Manufacturing Science and Engineering Conference*: Vol. 3 (pp. 1-21). https://doi.org/10.1115/MSEC2018-6658

[3] Romero, D., Bernus, P., Noran, O., Stahre, J., & Fast-Berglund, Å. (2016). The Operator 4.0 : Human Cyber-Physical Systems Adaptive Automation Towards Human-Automation Symbiosis Work Systems. In: I. Nääs et al. (Eds) *Advances in Production Management Systems. Initiatives for a Sustainable World. APMS 2016. IFIP Advances in Information and Communication Technology*: Vol 488 (pp. 677-686). Cham: Springer. https://doi.org/10.1007/978-3-319-51133-7_80

[4] Fischer, S. (2014). Challenges of the Internet of Services. In W. Wahlster W, HJ. Grallert, S. Wess, H. Friedrich, T. Widenka (Eds), Towards the Internet of Services: The THESEUS Research Program. Cognitive Technologies. (pp. 15-27). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-06755-1_2

[5] Qin, J., Liu, Y., & Grosvenor, R. (2016). A Categorical Framework of Manufacturing for Industry 4.0 and Beyond. In *Procedia CIRP. The Sixth International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV2016):* Vol. 52 (pp. 173–178). https://doi.org/10.1016/j.procir.2016.08.005

[6] Halaška, M., & Šperka, R. (2018). Process Mining - The Enhancement of Elements of Industry 4.0. In 2018 4th International Conference on Computer and Information Sciences (ICCOINS) (pp. 1-6). IEEE. https://doi.org/10.1109/ICCOINS.2018.8510578

[7] 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

[8] Paschek, D., Luminosu, C. T., & Draghici, A. (2017). Automated business process management - in times of digital transformation using machine learning or artificial intelligence. In 8th International Conference on Manufacturing Science and Education – MSE 2017 Trends in New Industrial Revolution: Vol. 121, Article 04007. https://doi.org/10.1051/matecconf/201712104007

[9] van der Aalst, W. M. P. (2022). Process Mining: A 360 Degree Overview. Lecture Notes in Business Information Processing, 448, 3–34. https://doi.org/10.1007/978-3-031-08848-3_1

[10] van der Aalst, W., Zhao, J. L., & Wang, H. J. (2015). Editorial: Business Process intelligence: Connecting data and processes. ACM Transactions on Management Information Systems, 5(4), Article 18e. https://doi.org/10.1145/2685352

[11] van der Aalst, W. (2016). Process mining: Data science in action (2nd ed). Berlin: Springer. https://doi.org/10.1007/978-3-662-49851-4

[12] van der Aalst, W. (2011). Process Mining: Discovery, Conformance and Enhancement of Business Processes. Berlin: Springer. https://doi.org/10.1007/978-3-642-19345-3

[13] van der Aalst, W. (2012). Process mining: Overview and opportunities. ACM transactions on management information systems, 3(2), 1-17. https://doi.org/10.1145/2229156.2229157

[14] Park, G., & van der Aalst, W. M. P. (2022). Action-oriented process mining: bridging the gap between insights and actions. Progress in Artificial Intelligence, 1–22. https://doi.org/10.1007/S13748-022-00281-7

[15] Büchi, G., Cugno, M., & Castagnoli, R. (2020). Smart factories performance and Industry 4.0. Technological Forecasting and Social Change, 150, Article 119790. https://doi.org/10.1016/j.techfore.2019.119790

[16] Oztemel, E., & Gursev, S. (2020). Literature review of Industry 4.0 and related technologies. Journal of Intelligent Manufacturing, 31(1), 127–182. https://doi.org/10.1007/s10845-018-1433-8

[17] Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. Journal of Industrial Information Integration, 6, 1–10. https://doi.org/10.1016/j.jii.2017.04.005

[18] Culot, G., Nassimbeni, G., Orzes, G., &Sartor, M. (2020). Behind the definition of Industry 4.0: Analysis and open questions. International Journal of Production Economics, 226, Article 107617. https://doi.org/10.1016/j.ijpe.2020.107617

[19] Fettermann, D. C., Cavalcante, C. G. S., Almeida, T. D. de, & Tortorella, G. L. (2018). How does Industry 4.0 contribute to operations management? 35(4), 255–268. https://doi.org/10.1080/21681015.2018.1462863

[20] Ghobakhloo, M., Fathi, M., Iranmanesh, M., Maroufkhani, P., & Morales, M. E. (2021). Industry 4.0 ten years on: A bibliometric and systematic review of concepts, sustainability value drivers, and success determinants. *Journal of Cleaner Production*, *302*, 127052. https://doi.org/10.1016/J.JCLEPRO.2021.127052

[21] Riel, A., & Flatscher, M. (2017). A design process approach to strategic production planning for industry 4.0. Communications in Computer and Information Science, 748, 323–333. https://doi.org/10.1007/978-3-319-64218-5_27

[22] Zheng T, Ardolino M, Bacchetti A, Perona M (2020) The applications of Industry 4.0 technologies in manufacturing context: a systematic literature review. https://doi.org/10.1080/00207543.2020.1824085

[23] Sanchez, M., Exposito, E., & Aguilar, J. (2020). Industry 4.0: survey from a system integration perspective. International Journal of Computer Integrated Manufacturing, 33(10-11), 1017–1041.

https://doi.org/10.1080/0951192X.2020.1775295

[24] Fernandes, E. C., Fitzgerald, B., Brown, L., & Borsato, M. (2019). Machine learning and process mining applied to process optimization: Bibliometric and systemic analysis. Procedia Manufacturing. 29th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM 2019): Vol. 38, (pp. 84–91). https://doi.org/10.1016/j.promfg.2020.01.012

[25] Garcia, C.D.S., Meincheim, A., Faria Junior, E.R., Dallagassa, M.R., Sato, D.M.V., Carvalho, D.R., Santos, E.A.P., & Scalabrin, E.E. (2019). Process mining techniques and applications – A systematic mapping study. Expert Systems With Applications, 133, 260–295. https://doi.org/10.1016/j.eswa.2019.05.003

[26] Chaydy, N., & Madani, A. (2019) An overview of Process Mining and its applicability to complex, real-life scenarios. In 2019 International Conference on Systems of Collaboration, Big Data, Internet of Things and Security (SysCoBIoTS) (pp. 1-9). IEEE. https://doi.org/10.1109/SysCoBIoTS48768.2019.9028024

[27] Thiede, M., Fuerstenau, D., & Barquet, A.P.B. (2017). How is process mining technology used by organizations? A systematic literature review of empirical studies. Business Process Management Journal, 24(4), 900-922. https://doi.org/10.1108/BPMJ-06-2017-0148

[28] Dakic, D., Stefanovic, D., Cosic, I., Lolic, T., & Medojevic, M. (2018). Business process mining application: A literature review. In B. Katalinic (Ed.), Proceedings of the 29th International DAAAM Symposium.: Vol. 1. (pp. 0866-0875). Vienna, Austria: DAAAM International. https://doi.org/10.2507/29th.daaam.proceedings.125

[29] Maita, A.R.C., Martins, L.C., López Paz, C.R., Rafferty, L., Hung, P.C.K., Peres, S.M., & Fantinato, M. (2018). A systematic mapping study of process mining. Enterprise Information Systems, 12(5), 505–549. https://doi.org/10.1080/17517575.2017.1402371

[30] Diba, K., Batoulis, K., Weidlich, M., & Weske, M. (2020). Extraction, correlation, and abstraction of event data for process mining. WIREs Data Mining and Knowledge Discovery, 10(3), Article e1346. https://doi.org/10.1002/widm.1346

[31] Marquez-Chamorro, A. E., Resinas, M., & Ruiz-Cortes, A. (2018). Predictive monitoring of business processes: A survey. IEEE Transactions on Services Computing, 11(6), 962–977. https://doi.org/10.1109/TSC.2017.2772256

[32] Teinemaa, I., Dumas, M., La Rosa, M., & Maggi, F. M. (2019). Outcome-Oriented Predictive Process Monitoring: Review and Benchmark. ACM Transactions on Knowledge Discovery from Data, 13(2), 1–57. https://doi.org/10.1145/3301300

[33] Di Francescomarino, C., Ghidini, C., Maggi, F. M., & Milani, F. (2018). Predictive Process Monitoring Methods: Which One Suits Me Best? In M. Weske, M. Montali, I. Weber, J. vomBrocke (Eds) Business Process Management. BPM 2018. Lecture Notes in Computer Science: Vol. 11080 (pp. 462-479). Cham: Springer. https://doi.org/10.1007/978-3-319-98648-7_27

[34] Ye, J., Dobson, S., & McKeever, S. (2012). Situation identification techniques in pervasive computing: A review. Pervasive and Mobile Computing, 8(1), 36–66. https://doi.org/10.1016/j.pmcj.2011.01.004

[35] Mannhardt, F., Bovo, R., Oliveira, M. F., & Julier, S. (2018). A taxonomy for combining activity recognition and process discovery in industrial environments. In Intelligent Data Engineering and Automated Learning – IDEAL 2018 19th International Conference: Vol. 11315, (pp. 84–93). https://doi.org/10.1007/978-3-030-03496-2_10

[36] Praveen Kumar, D., Amgoth, T., & Annavarapu, C. S. R. (2019). Machine learning algorithms for wireless sensor networks: A survey. Information Fusion, 49, 1–25. https://doi.org/10.1016/j.inffus.2018.09.013

[37] Perera, C., Zaslavsky, A., Christen, P., & Georgakopoulos, D. (2014). Context aware computing for the internet of things: A survey. IEEE Communications Surveys and Tutorials, 16(1), 414–454. https://doi.org/10.1109/SURV.2013.042313.00197

[38] Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., & Barbosa, J. (2020). Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. Computers in Industry, 123, Article 103298. https://doi.org/10.1016/j.compind.2020.103298

[39] Çinar, Z. M., Nuhu, A. A., Zeeshan, Q., Korhan, O., Asmael, M., & Safaei, B. (2020). Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. Sustainability, 12, Article 8211. https://doi.org/10.3390/su12198211

[40] Ramírez-Gallego, S., Krawczyk, B., García, S., Woźniak, M., & Herrera, F. (2017). A survey on data preprocessing for data stream mining: Current status and future directions. Neurocomputing, 239, 39–57. https://doi.org/10.1016/j.neucom.2017.01.078

[41] Osman, CC., Ghiran, AM. (2019) When Industry 4.0 meets Process Mining. Procedia Computer Science, 159, 2130–2136. https://doi.org/10.1016/J.PROCS.2019.09.386

[42] Kitchenham. B.A., & Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in software engineering. EBSE Technical Report EBSE-2007-01. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.471

[43] Cooper HM (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. Knowledge in Society, 1, 104–126. https://doi.org/10.1007/BF03177550

[44] Wohlin, C. (2014). Guidelines for Snowballing in Systematic Literature Studies and a Replication

in Software Engineering. In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE) (pp. 1-10). http://dx.doi.org/10.1145/2601248.2601268

[45] Bertrand, Y., Van den Abbeele, B., Veneruso, S., Leotta, F., Mecella, M., Serral, E. (2023). A Survey on the Application of Process Mining to Smart Spaces Data. In M. Montali, A. Senderovich, M.Weidlich, (eds) *ICPM 2022:Process Mining Workshops. Lecture Notes in Business Information Processing*: Vol 468 (pp. 57-70). Cham: Springer. https://doi.org/10.1007/978-3-031-27815-0_5

[46] Mittal S, Khan MA, Romero D, Wuest T (2017) Smart manufacturing: Characteristics, technologies and enabling factors. Smart manufacturing and digital factory, 233(5), 1342-1361. https://doi.org/10.1177/0954405417736547

[47] Butt J (2020) A conceptual framework to support digital transformation in manufacturing using an integrated business process management approach. Designs, 4(3), 1–39. https://doi.org/10.3390/designs4030017

[48] Fatorachian H, Kazemi H (2018) A critical investigation of Industry 4.0 in manufacturing: theoretical operationalisation framework. Production planning and control. 29(8), 633–644. https://doi.org/10.1080/09537287.2018.1424960

[49] Butt J (2020) A strategic roadmap for the manufacturing industry to implement industry 4.0. Designs, 4(2), 1–31. https://doi.org/10.3390/DESIGNS4020011

[50] Guo, Q., Wen, L., Wang, J., Yan, Z., & Yu, P. S. (2015). Mining invisible tasks in non-free-choice constructs. In H. Motahari-Nezhad, J. Recker, M. Weidlich (Eds) BPM 2016: Business Process Management. LectureNotes in Computer Science: Vol 9253 (pp. 109-125). Cham: Springer. https://doi.org/10.1007/978-3-319-23063-4_7

[51] Leemans, S. J. J., Fahland, D., & van der Aalst, W. (2014). Discovering block-structured process models from incomplete event logs. In G. Ciardo, E. Kindler (Eds) PETRI NETS 2014: Application and Theory of Petri Nets and Concurrency. LectureNotes in Computer Science: Vol. 8489 (pp. 91–110). Cham: Springer.https://doi.org/10.1007/978-3-319-07734-5_6

[52] Leemans, S. J. J., Fahland, D., & van der Aalst, W. (2016). Using life cycle information in process discovery. In M. Reichert M., H. Reijers H. (Eds) BPM 2016: Business Process Management Workshops. Lecture Notes in Business Information Processing: Vol. 256 (pp. 204-217). Charm: Springer. https://doi.org/10.1007/978-3-319-42887-1_17

[53] Stertz, F., & Rinderle-Ma, S. (2018). Process Histories - Detecting and Representing Concept Drifts Based on Event Streams. In: H. Panetto, C. Debruyne, H. Proper, C. Ardagna, D. Roman, R. Meersman (Eds) On the Move to Meaningful Internet Systems. OTM 2018 Conferences. OTM 2018. Lecture Notes in Computer Science: Vol 11229 (pp. 318-335). Cham: Springer. https://doi.org/10.1007/978-3-030-02610-3_18

[54] Leemans, S. J. J., Fahland, D., & van der Aalst, W. (2015). Exploring processes and deviations.

In F. Fournier, J. Mendling (Eds) BPM 2014: Business Process Management Workshops. Lecture Notes in Business Information Processing: Vol. 202 (pp. 304–316). https://doi.org/10.1007/978-3-319-15895-2_26

[55] Augusto, A., Conforti, R., Dumas, M., La Rosa, M., & Bruno, G. (2018). Automated discovery of structured process models from event logs: The discover-and-structure approach. Data and Knowledge Engineering, 117, 373–392. https://doi.org/10.1016/j.datak.2018.04.007

[56] Augusto, A., Conforti, R., Dumas, M., & La Rosa, M. (2017). Split miner: Discovering accurate and simple business process models from event logs. In 2017 IEEE International Conference on Data Mining (ICDM) (p.1–10). IEEE. https://doi.org/10.1109/ICDM.2017.9

[57] vanden Broucke, S. K. L. M., & de Weerdt, J. (2017). Fodina: A robust and flexible heuristic process discovery technique. Decision Support Systems, 100, 109–118. https://doi.org/10.1016/j.dss.2017.04.005

[58] Golchi, M. M., Saraeian, S., & Heydari, M. (2019). A hybrid of firefly and improved particle swarm optimization algorithms for load balancing in cloud environments: Performance evaluation. Computer Networks, 162, Article 106860. https://doi.org/10.1016/j.comnet.2019.106860

[59] van der Aalst, W., Rubin, V., Verbeek, H. M. W., Van Dongen, B. F., Kindler, E., &Günther, C. W. (2010). Process mining: A two-step approach to balance between underfitting and overfitting. Software and Systems Modeling, 9(1), 87–111. https://doi.org/10.1007/s10270-008-0106-z

[60] Viale, P., Benayadi, N., Le Goc, M., & Pinaton, J. (2010). Modeling large scale manufacturing process from timed data: Using the TOM4L approach and sequence alignment information for modeling STMicroelectronics' production processes. In ICEIS 2010 - Proceedings of the 12th International Conference on Enterprise Information Systems, 2 AIDSS (pp. 129–138). https://doi.org/10.5220/0002971801290138

[61] de Leoni, M., & van der Aalst, W. (2013). Data-Aware Process Mining: Discovering Decisions in Processes Using Alignments. In SAC '13: Proceedings of the 28th Annual ACM Symposium on Applied Computing (pp. 1454-1461). https://doi.org/10.1145/2480362.2480633

[62] JagadeeshChandraBose, R. P., & van der Aalst, W.M.P. (2009). Abstractions in process mining: A taxonomy of patterns. In U. Dayal, J. Eder, J. Koehler, H.A. Reijers (Eds) BPM 2009: Business Process Management. Lecture Notes in Computer Science: Vol 5701 (pp. 159-175). Berlin: Springer. https://doi.org/10.1007/978-3-642-03848-8_12

Primary studies

[P1] Halaška, M., & Šperka, R. (2019). Performance of an automated process model discovery - The logistics process of a manufacturing company. *Engineering Management in Production and Services*, 11(2), 106–118. https://doi.org/10.2478/emj-2019-0014

[P2] Knoch, S., Herbig, N., Ponpathirkoottam, S., Kosmalla, F., Staudt, P., Porta, D., Fettke, P., & Loos, P. (2020a). Sensor-based Human–Process Interaction in Discrete Manufacturing. *Journal on Data Semantics*, 9(1), 21–37. https://doi.org/10.1007/s13740-019-00109-z

[P3] Aehnelt, M., & Bader, S. (2018). Providing and adapting information assistance for smart assembly stations. In Y. Bi, S. Kapoor, R. Bhatia (Eds) *Proceedings of SAI Intelligent Systems Conference (IntelliSyst) Lecture Notes in Networks and Systems*: Vol. 15 (pp. 540–562). Cham: Springer. https://doi.org/10.1007/978-3-319-56994-9_37

[P4] Knoch, S., Ponpathirkoottam, S., & Schwartz, T. (2020b). Video-to-model: Unsupervised trace extraction from videos for process discovery and conformance checking in manual assembly. In: D. Fahland, C. Ghidini, J. Becker, M. Dumas (Eds) *BPM 2020: Business Process Management. Lecture Notes in Computer Science*: Vol 12168 (pp. 291–308). Cham: Springer. https://doi.org/10.1007/978-3-030-58666-9_17

[P5] Bauters, K., Cottyn, J., Claeys, D., Slembrouck, M., Veelaert, P., & van Landeghem, H. (2018). Automated work cycle classification and performance measurement for manual work stations. *Robotics and Computer-Integrated Manufacturing*, 51, 139–157. https://doi.org/10.1016/j.rcim.2017.12.001

[P6] Hutabarat, W., Oyekan, J., Turner, C., Tiwari, A., Prajapat, N., Gan, X. P., & Waller, A. (2016). Combining virtual reality enabled simulation with 3D scanning technologies towards smart manufacturing. In *2016 Winter Simulation Conference (WSC)*, (pp. 2774-2785). IEEE. https://doi.org/10.1109/WSC.2016.7822314

[P7] Zhuang, C., Miao, T., Liu, J., & Xiong, H. (2021). The connotation of digital twin, and the construction and application method of shop-floor digital twin. *Robotics and Computer-Integrated Manufacturing*, 68, Article 102075. https://doi.org/10.1016/j.rcim.2020.102075

[P8] Zhang, Y., Xu, J., Sun, S., & Yang, T. (2015). Real-time information driven intelligent navigation method of assembly station in unpaced lines. *Computers and Industrial Engineering*, 84, 91–100. https://doi.org/10.1016/j.cie.2014.09.007

[P9] Zhang, Y., Wang, W., Wu, N., & Qian, C. (2016). IoT-Enabled Real-Time Production Performance Analysis and Exception Diagnosis Model. *IEEE Transactions on Automation Science and Engineering*, 13(3), 1318–1332. https://doi.org/10.1109/TASE.2015.2497800

[P10] Fang, W., & Zheng, L. (2020). Shop floor data-driven spatial–temporal verification for manual assembly planning. *Journal of Intelligent Manufacturing*, 31(4), 1003–1018. https://doi.org/10.1007/s10845-019-01491-y

[P11] Roitberg, A., Somani, N., Perzylo, A., Rickert, M., & Knoll, A. (2015). Multimodal human activity recognition for industrial manufacturing processes in robotic workcells. In *ICMI 2015 - Proceedings of the 2015 ACM International Conference on Multimodal Interaction* (pp. 259–266). https://doi.org/10.1145/2818346.2820738

[P12] Lenz, C., Sotzek, A., Röder, T., Radrich, H., Knoll, A., Huber, M., & Glasauer, S. (2011). Human workflow analysis using 3D occupancy grid hand tracking in a human-robot collaboration scenario. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3375-3380). IEEE. https://doi.org/ 10.1109/IROS.2011.6094570

[P13] Bleser, G., Damen, D., Behera, A., Hendeby, G., Mura, K., Miezal, M., Gee, A., Petersen, N., Maçães, G., Domingues, H., Gorecky, D., Almeida, L., Mayol-Cuevas, W., Calway, A., Cohn, A. G., Hogg, D. C., & Stricker, D. (2015). Cognitive Learning, Monitoring and Assistance of Industrial Workflows Using Egocentric Sensor Networks. *PLOS ONE*, 10(6), Article e0127769. https://doi.org/10.1371/journal.pone.0127769

[P14] Cramer, M., Cramer, J., Kellens, K., & Demeester, E. (2018). Towards robust intention estimation based on object affordance enabling natural human-robot collaboration in assembly tasks. In *Procedia CIRP. 6th CIRP Global Web Conference – Envisaging the future manufacturing, design, technologies and systems in innovation era (CIRPe 2018)*: Vol. 78, (pp. 255–260). https://doi.org/10.1016/j.procir.2018.09.069

[P15] Roldán, J. J., Crespo, E., Martín-Barrio, A., Peña-Tapia, E., & Barrientos, A. (2019). A training system for Industry 4.0 operators in complex assemblies based on virtual reality and process mining. *Robotics and Computer-Integrated Manufacturing*, 59, 305–316. https://doi.org/10.1016/j.rcim.2019.05.004

[P16] Saraeian, S., & Shirazi, B. (2020). Process mining-based anomaly detection of additive manufacturing process activities using a game theory modeling approach. *Computers and Industrial Engineering*, 146, Article 106584. https://doi.org/10.1016/j.cie.2020.106584

[P17] Myers, D., Suriadi, S., Radke, K., & Foo, E. (2018). Anomaly detection for industrial control systems using process mining. *Computers and Security*, 78, 103–125. https://doi.org/10.1016/j.cose.2018.06.002

[P18] Setiawan, F., & Yahya, B. N. (2018). Improved behavior model based on sequential rule mining. *Applied Soft Computing Journal*, 68, 944–960. https://doi.org/10.1016/j.asoc.2018.01.035

[P19] Tello-Leal, E., Roa, J., Rubiolo, M., & Ramirez-Alcocer, U. M. (2018). Predicting activities in business processes with LSTM recurrent neural networks. In *2018 ITU Kaleidoscope: Machine Learning for a 5G Future (ITU K)* (pp. 1-7). https://doi.org/10.23919/ITU-WT.2018.8598069

[P20] Mehdiyev, N., Emrich, A., Stahmer, B., Fettke, P., & Loos, P. (2017). iPRODICT-Intelligent Process Prediction based on Big Data Analytics. In T. Hildebrandt, M. Brambilla (Eds) *CEUR Workshop Proceedings*: Vol. 1985 (pp. 13-24). CEUR-WS.
Retrieved from: http://ceur-ws.org/Vol-1985/BPM17industry02.pdf

[P21] Repta, D., Moisescu, M. A., Sacala, I. S., Dumitrache, I., & Stanescu, A. M. (2015). Towards the development of semantically enabled flexible process monitoring systems. *International Journal of Computer Integrated Manufacturing*, 30(1), 1–13. https://doi.org/10.1080/0951192X.2015.1107914

[P22] Véjar, A., & Charpentier, P. (2012). Generation of an adaptive simulation driven by product trajectories. *Journal of Intelligent Manufacturing*, 23(6), 2667–2679. https://doi.org/10.1007/s10845-011-0504-x

[P23] Sanchez, M., Exposito, E., & Aguilar, J. (2020b). Implementing self-* autonomic properties in self-coordinated manufacturing processes for the Industry 4.0 context. *Computers in Industry*, 121, Article 103247. https://doi.org/10.1016/j.compind.2020.103247

[P24] Brzychczy, E., Gackowiec, P., & Liebetrau, M. (2020). Data Analytic Approaches for Mining Process Improvement—Machinery Utilization Use Case. *Resources*, 9(2), 17. https://doi.org/10.3390/resources9020017

[P25] Ribeiro, R., Analide, C., & Belo, O. (2018). Improving productive processes using a process mining approach. *Advances in Intelligent Systems and Computing*, 746, 736–745. https://doi.org/10.1007/978-3-319-77712-2_69

[P26] Malburg, L., Seiger, R., Bergmann, R., & Weber, B. (2020). Using Physical Factory Simulation Models for Business Process Management Research. In A. Del Río Ortega, H. Leopold, F.M. Santoro (Eds) *BPM 2020: Business Process Management Workshops. Lecture Notes in Business Information Processing*: Vol 397 (pp. 95-107). Cham: Springer. https://doi.org/10.1007/978-3-030-66498-5_8

[P27] Seiger, R. ,Zerbato, F., Burattin, A., Garcia-Banuelos, L., & Weber, B. (2020). Towards IoT-driven Process Event Log Generation for Conformance Checking in Smart Factories. In *2020 IEEE 24th International Enterprise Distributed Object Computing Workshop (EDOCW)* (pp. 20-26). IEEE. https://doi.org/10.1109/EDOCW49879.2020.00016

[P28] Meroni, G., Baresi, L., Montali, M., & Plebani, P. (2018). Multi-party business process compliance monitoring through IoT-enabled artifacts. *Information Systems*, 73, 61–78. https://doi.org/10.1016/j.is.2017.12.009

[P29] Ehrendorfer, M., Fassmann, J. A., Mangler, J., & Rinderle-Ma, S. (2019). Conformance checking and classification of manufacturing log data. In *21st IEEE Conference on Business Informatics (CBI)* (pp. 569–577). IEEE. https://doi.org/10.1109/CBI.2019.00072

[P30] Stertz, F., Rinderle-Ma, S., & Mangler, J. (2020a). Analyzing process concept drifts based on sensor event streams during runtime. In D. Fahland D., C. Ghidini, J. Becker, M. Dumas (Eds) *BPM 2020: Business Process Management. Lecture Notes in Computer Science*: Vol. 12168 (pp. 202-219). Cham: Springer. https://doi.org/ 10.1007/978-3-030-58666-9_12

[P31] Stertz, F., Mangler, J., & Rinderle-Ma, S. (2020b). Data-driven Improvement of Online Conformance Checking. In *2020 IEEE 24th International Enterprise Distributed Object*

Computing Conference (EDOC) (pp. 187-196). IEEE. https://doi.org/10.1109/EDOC49727.2020.00031

[P32] Dunkl, R., Rinderle-Ma, S., Grossmann, W., & Fröschl, K. A. (2015). A method for analyzing time series data in process mining: Application and extension of decision point analysis. In S.Nurcan, E.Pimenidis (Eds) *Information Systems Engineering in Complex Environments. CAiSE 2014. Lecture Notes in Business Information Processing*: Vol. 204 (pp. 68-84). Cham: Springer. https://doi.org/10.1007/978-3-319-19270-3_5

[P33] Rodriguez-Fernandez, V., Trzcionkowska, A., Gonzalez-Pardo, A., Brzychczy, E., Nalepa, G. J., & Camacho, D. (2021). Conformance Checking for Time-Series-Aware Processes. *IEEE Transactions on Industrial Informatics*, 17(2), 871–881. https://doi.org/10.1109/TII.2020.2977126

[P34] Rebmann, A., Emrich, A., & Fettke, P. (2019). Enabling the Discovery of Manual Processes Using a Multi-modal Activity Recognition Approach. In C. Di Francescomarino, R. Dijkman, U. Zdun (Eds) *BPM 2019: Business Process Management Workshops. Lecture Notes in Business Information Processing*: Vol 362 (pp. 130–141). Cham: Springer.

https://doi.org/10.1007/978-3-030-37453-2_12

[P35] Rebmann, A., Rehse, J. R., Pinter, M., Schnaubelt, M., Daun, K., & Fettke, P. (2020). Iot-based activity recognition for process assistance in human-robot disaster response. In D. Fahland, C. Ghidini, J. Becker, M. Dumas (Eds) *BPM 2020: Business Process Management Forum. Lecture Notes in Business Information Processing*: Vol 392 (pp. 71-87). Cham: Springer. https://doi.org/10.1007/978-3-030-58638-6_5

[P36] Rigger, E., Vosgien, T., Bitrus, S., Szabo, P., & Eynard, B. (2020). Enterprise Architecture Method for Continuous Improvement of PLM Based on Process Mining. In F. Nyffenegger, J. Ríos, L. Rivest, A. Bouras (Eds) *PLM 2020: Product Lifecycle Management Enabling Smart X. IFIP Advances in Information and Communication Technology*: Vol 594 (pp. 563-575). Cham: Springer. https://doi.org/10.1007/978-3-030-62807-9_45

[P37] Es-Soufi, W., Yahia, E., & Roucoules, L. (2017). A process mining based approach to support decision making. In J. Ríos, A. Bernard, A. Bouras, S. Foufou (Eds) *PLM 2017: Product Lifecycle Management and the Industry of the Future. IFIP Advances in Information and Communication Technology*: Vol 517 (pp. 264-274). Cham: Springer.
https://doi.org/10.1007/978-3-319-72905-3_24

[P38] Park, M., Song, M., Baek, T. H., Son, S. Y., Ha, S. J., & Cho, S. W. (2015). Workload and delay analysis in manufacturing process using process mining. In J. Bae J, S. Suriadi, L. Wen (Eds) *AP-BPM 2015: Asia Pacific Business Process Management. Lecture Notes in Business Information Processing*: Vol 219 (pp. 138–151). Cham: Springer.
https://doi.org/10.1007/978-3-319-19509-4_11

[P39] Yang, H., Park, M., Cho, M., Song, M., & Kim, S. (2014). A system architecture for

manufacturing process analysis based on big data and process mining techniques. In *2014 IEEE International Conference on Big Data (Big Data)* (pp. 1024-1029). IEEE. https://doi.org/10.1109/BigData.2014.7004336

[P40] Lugaresi, G., & Matta, A. (2021). Automated manufacturing system discovery and digital twin generation. *Journal of Manufacturing Systems*, 59, 51–66. https://doi.org/10.1016/j.jmsy.2021.01.005

[P41] Farooqui, A., Bengtsson, K., Falkman, P., & Fabian, M. (2019). From factory floor to process models: A data gathering approach to generate, transform, and visualize manufacturing processes. *CIRP Journal of Manufacturing Science and Technology*, 24, 6–16. https://doi.org/10.1016/j.cirpj.2018.12.002

[P42] Koehler, W., & Jing, Y. (2020). Trace induction for complete manufacturing process model discovery. International *Journal of Advanced Manufacturing Technology*, 110, 29–43. https://doi.org/10.1007/s00170-020-05747-3

[P43] Lugaresi, G., Zanotti, M., Tarasconi, D., & Matta, A. (2019). Manufacturing systems mining: Generation of real-time discrete event simulation models. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)* (pp. 415-420). IEEE. https://doi.org/10.1109/SMC.2019.8914025

[P44] Viale, P., Frydman, C., & Pinaton, J. (2011). New methodology for modeling large scale manufacturing process: Using process mining methods and experts' knowledge. In *9th IEEE/ACS International Conference on Computer Systems and Applications (AICCSA)* (pp. 84–89). IEEE. https://doi.org/10.1109/AICCSA.2011.6126595

[P45] Ruppert, T., & Abonyi, J. (2020). Integration of real-time locating systems into digital twins. *Journal of Industrial Information Integration*, 20, Article 100174. https://doi.org/10.1016/j.jii.2020.100174

[P46] Dasani, S., Shah, S. L., Chen, T., Funnell, J., & Pollard, R. W. (2015). Monitoring safety of process operations using industrial workflows. In B. Huang, R. Findeisen, M. Guay, B. Gopaluni (Eds) *9th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM 2015. IFAC-PapersOnLine*: Vol. 28(8) (pp. 451–456). https://doi.org/10.1016/j.ifacol.2015.09.009

[P47] Hu, W., Al-Dabbagh, A. W., Li, D., & Chen, T. (2018). Extraction and Graphical Representation of Operator Responses to Multivariate Alarms in Industrial Facilities. In *10th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM 2018. IFAC-PapersOnLine*: Vol. 51(18) (pp. 25–30). https://doi.org/10.1016/j.ifacol.2018.09.240

Attachment

Table A: Application of process discovery

Type of process	Paper ID	Role of PM	Application purpose
Manual assembly process	[P4]	Discovers the manual process from videos	Input for process analysis and optimization
	[P5]	Discovers the manual process from videos	Input for PCC(O)/PA
	[P10]	Discovers flexible (manual) process from sensors in wearable tracking system	Input for process analysis and optimization, input for PCC
	[P13]	Discovers flexible (manual) process from sensor data	Knowledge transfer, input for PCC(O)/AOPM
	[P15]	Discovers flexible (manual) process from data in virtual reality	Knowledge transfer
	[P18]	Discovers flexible (manual) process from sensor data	Work control and evaluation
	[P23]	Discovers process from sensors to see the production's flow, bottlenecks, average processing time	Input for AOPM
	[P34]	Discovers process of manual activity from sensor data, images and feedback from the worker	Input for process analysis and optimization
	[P45]	Discovers process and updates the model based on information from manufacturing execution system and real-time locating system	Input/configuration for simulation
Hybrid assembly process	[P11, 12]	Discovers flexible (manual) process from sensor data	Input for process analysis and optimization
	[P14]	Discovers flexible (manual) process from sensor data	Transfer model of human behavior into robot to improve human-robot cooperation
Discrete automated manufacturing process	[P41]	Discovers and visualizes process of the behavior of a robot from data from Programmable Logic Controllers (PLCs)	Input for process analysis and optimization,
	[P42]	Discovers process from PLCs, identifies bottlenecks, predict time for an activity to be executed	Input for process analysis and optimization

Type of process	Paper ID	Role of PM	Application purpose
Production line	[P22]	Discovers process based on product location data stream	Input for PPM
	[P23]	Discovers process from sensors to see the production's flow, bottlenecks, average processing time	Input for AOPM
	[P39]	Discovers process from combination of structured and unstructured data	Input for process analysis and optimization
	[P40]	Discovers manufacturing system structure and characteristics from data generated by Manufacturing Execution system	Input/configuration for simulation, input for process analysis and optimization
	[P41]	Discovers and visualizes process of the behavior of a robot from data from PLCs	Input for process analysis and optimization,
	[P43]	Discovers model with possibility to fix the maximum number of activities allowed (decide on the level of details)	Input/configuration for simulation, input for process analysis and optimization
	[P44]	Discovers generic model from activity and from expert knowledge	Input for process optimization
	[P45]	Discovers process and updates the model based on information from manufacturing execution system and real-time locating system	Input/configuration for simulation
Production line +planning+post-processing	[P16]	Discovers process from the Business Process Management System (distributive event log)	Input for PCC
	[P25]	Discovers process with lots of parameters and high cost of failure from data coming from internal process support system and process management reports	Input for process analysis and optimization
Production line+order processing	[P20]	Discovers process based on data from different sources: sensor data, mobility data, video data, ERP systems etc.	Input for PCC, Input for PPM
	[P30]	Discovers process from process history with use of sliding window approach to deal with infinite amount of data	Input for PCC(O)

Type of process	Paper ID	Role of PM	Application purpose
Underground mining process	[P24]	Discovers process from big volume of data coming from sensors and PLC-readers with a short time interval with lots of outliers	Input for process analysis and optimization
Non-manufacturing process	[P17]	Discovers process from industrial control systems device log	Input for PCC
	[P46]	Discovers process from the data in the information system	Input for PCC, knowledge transfer
Internal logistics	[P1]	Discovers the process addressing scalability issues, incompleteness and lack of information in datasets	Input for process analysis and optimization
	[P21]	Discovers flexible process from sensor data taking into account context data	Input for process analysis and optimization
	[P32]	Discovers process from data from simulation environment	Input for PA
Alarm management	[P37]	Discovers process taking into account context data	Input for process analysis, input for PA
	[P47]	Discovers process with focus on dependencies between alarm and operator's response, model represented with the new graphical visualization framework	Input for process analysis and optimization, knowledge transfer
Robot-assisted disaster support	[P35]	Discovers process of robot's activity from sensor data	Input for PCC, documentation, knowledge transfer

Table B: Application of process conformance checking

Type of process	Paper ID	Role of PM	Application Purpose
Offline			
Manual assembly process	[P4]	Measures the quality of the model generated by PD	Evaluation of PD, evaluation of procedure to generate event log
	[P10]	Identifies deviations in terms of trajectories of movement and time differences	Input for optimization of the planned process
	[P18]	Measures the quality of the model generated by PD (alignment cost)	Evaluation of PD, evaluation of procedure to generate event log
	[P34]	Measures the quality of the model generated by PD	Evaluation of PD
Production line	[P29]	Identify special instances in the dataset (accidentally included, have been aborted, or have many ad-hoc changes or changes to the thread of control)	Filtering manufacturing datasets for further analysis
Production line +planning+post-processing	[P16]	Detects anomalous events	Cyber-attack detection
Production line+order processing	[P20, P38]	Detects violation is process execution	Input for process optimization
Non-manufacturing process	[P17]	Detects anomalous events	Cyber-attack detection
	[P33]	Identifies discrepancies in terms of absent events and time mismatches for time series data	Input for process analysis and optimization
PLM	[P36]	Quantifies the cost of deviations between a model and a real process	Monitoring alignment of IT systems with related PLM processes
Online			
Manual assembly process	[P2]	Detects deviations in ongoing process	Operational support, in-time material supply, input for process optimization
	[P5]	Detects deviations in ongoing process	Operational support, process monitoring and control,

Type of process	Paper ID	Role of PM	Application Purpose
			work control and evaluation, knowledge transfer, input for process optimization
Production line	[P9]	Detects deviations in ongoing process	Process monitoring and control
	[P26]	Detects deviations in ongoing process base on data from sensors with multiple instances running in parallel	Process monitoring and control
Production line+order processing	[P27]	Detects deviations in ongoing process based on data from sensors	Process monitoring and control
	[P30]	Detects concept drift	Process monitoring and control, input for process optimization
	[P31]	Detects deviations in ongoing process, quantifies conformance deviation cost	Process monitoring and control
Non-manufacturing process	[P46]	Detects deviations in ongoing process	Work control and evaluation
Internal logistics	[P9]	Detects deviations in ongoing process	Process monitoring and control
	[P27]	Detects deviations in ongoing process based on data from sensors	Process monitoring and control
Logistics in supply chain	[P28]	Detects compliance violations, even if they do not explicitly violate the control flow, based on data from IoT and declarative process model	Process monitoring and control
Robot-assisted disaster support	[P35]	Detects deviations in ongoing process	Process monitoring and control, input for process optimization

Table C: Application of performance analysis

Type of process	Paper ID	Role of PM	Application purpose
Manual assembly process	[P3]	Controls and change error handling behavior with machine learning	Input for process optimization
	[P5]	Identifies bottlenecks in terms of cycle time above average	Operational support, process monitoring and control, work control and evaluation, knowledge transfer, input for process optimization
	[P45]	Analyses human resource effectiveness as a function of working time and idle time	Work control and evaluation
Production line	[P3]	Controls and change error handling behavior with machine learning	Model optimization
	[P9]	Measure time for jobs to wait in the queue, queue length, production cycle time, finds causes of exception events	Process monitoring and control
	[P39]	Analyses manufacturing time, waiting time, lead time	Input for process analysis and optimization
	[P40]	Analyses needed level of details of the model	Model optimization
	[P43]	Obtains formally correct model of the system	Model optimization
	[P45]	Analyses human resource effectiveness as a function of working time and idle time	Work control and evaluation
Production line+order processing	[P38]	Analyses workload analysis, delay analysis	Input for process analysis and optimization
Underground mining process	[P24]	Analyses performance parameters, e.g. minimum, maximum, mean value of duration between activities	Input for process analysis and optimization
Internal logistics	[P9]	Measure time for jobs to wait in the queue, queue length, production cycle time,	Process monitoring and control
	[P32]	Discovers decision rules	Decision point analysis

Alarm management	[P37]	Discovers decision rules	Decision point analysis

Table D: Application of predictive PM

Type of process	Paper ID	Role of PM	Application purpose
Production line	[P6, P7]	Predicts possible performance shortfalls and bottlenecks in real-time	Process monitoring and control
	[P19]	Predicts execution of cases	Input for planning and resource allocation
	[P20]	Predicts failure, predicts post-processing step	Process monitoring and control
	[P22]	Predicts next step in the process	Process monitoring and control
	[P45]	Generates plant simulation model	Input for production scheduling
Production line+order processing	[P20]	Predicts failure, predicts post-processing step	Process monitoring and control

Table E: Application of action-oriented PM

Type of process	Paper ID	Role of PM	Application purpose
Manual assembly process	[P3]	Makes decisions on the recommended next work step based on the current situation, process model and contextual background knowledge	Operational support, in-time material supply, documentation, work control and evaluation, knowledge transfer, input for task scheduling
	[P8]	Identifies current step in the process and provides operator with instruction on the next step	Operational support
	[P13]	Identifies current step, identifies deviations and provides operator with instruction on the next step	Operational support
	[P23]	Predicts future failures based on predictive model of a process and provides operator with instruction on the next step	Operational support
Production line	[P3]	Makes decisions on the recommended next work step based on the current situation, process model and contextual background knowledge	Operational support
	[P8]	Captures exceptions, e.g. changed delivery time, new tasks with high priority, and makes decisions on queueing tasks in real-time	Process monitoring and control, task scheduling
	[P23]	Predicts future failures based on predictive model of a process and model of a normal process and evokes automatic cycle of self-healing	Process monitoring and control

Table F: Process discovery algorithms

Type of algorithm. N. of papers (%)	**Algorithm**	**Paper ID**	**Comments**	**Identified Capabilities**
Standard process discovery algorithms. 8 (25%)	α-algorithm	[P32]		
	Inductive miner	[P15, P41]		Best algorithm in the context of robot mission, remarkable fitness with good balance between generalization and precision [P15]
	Heuristic miner	[P18, P45, P47]		Suitable to discover human behavior because it takes frequencies, sequences, casual dependencies of events into account [P18] Robust to noises, more appropriate for the case with multivariate alarms: [P47]
	One of the standard PM algorithms	[P35, P39]	without specifying which one in particular	
Extensions of standard algorithms. 9 (28%)	α$ [50]	[P1]	extension of α-algorithm	Good performance when discovering simple process with both big number and small number of cases
	α_{LC} algorithm	[P42]	extension of α-algorithm for industrial automation process (definition of event sequences based on timestamps)	Less complicated than other PM algorithms Can derive a model based on minimum number case traces.
	Extensions of inductive miner	[P1, P4]		Good performance when discovering simple and more complex process with both big number and small number of cases [P1].

Type of algorithm. N. of papers (%)	Algorithm	Paper ID	Comments	Identified Capabilities
	[51]			
	Inductive miner with Perfect fitness	[P17]	extension of inductive miner	The most suitable algorithm to discover model of an industrial control systems
	Extension of inductive miner [52]	[P25]		Ability to distinguish concurrency and interleaving, necessity for high precision of the model
	Process history mining based on inductive miner [53]	[P30]	extension of inductive miner	
	Inductive visual miner [54]	[P37]	extension of inductive miner with a decision mining algorithm	Allows inclusive gateways to be identified
	Structure Heuristic Miner [55]	[P1]	extension of heuristic miner	Good performance when discovering simple processes with both big and small number of cases, and when discovering complex processes with big number of cases
	Split Miner [56]	[P1]	extension of heuristic miner	Good performance when discovering simple processes with both big and small number of cases
	Fodina [57]	[P1]	extension of heuristic miner	Good performance when discovering simple and complex processes with big number of cases
	Hybrid algorithm based on	[P16]	extension of heuristic miner	High sensitivity to initial conditions of IPSO algorithm, high convergence of the Firefly algorithm, high accuracy of

Type of algorithm. N. of papers (%)	Algorithm	Paper ID	Comments	Identified Capabilities
	combination of Improved Particle Swarm Optimization (IPSO) and Firefly algorithms [58]			combined algorithm
	Extension of heuristic miner, fuzzy miner, and the genetic approach.	[P43]	combines frequentist approach (heuristic miner) with solving optimization problem (genetic approach) and classification at input phase (fuzzy miner)	Discovers process with multiple paths with possibility to fix the number of events (level of details)
New proposed algorithms. 2 (6%)	Enhancement of two-step PM algorithm [59]	[P21]	generates a transition system as an intermediate representation and converting transition system into a Petri Net using regions as the final representation	Improve performance for the case of large number parallel activities
	Combination of a new algorithm for graph model generation with a standard PM for	[P40]	PM	Reproduces characteristics and estimate the performance of the system

Type of algorithm. N. of papers (%)	Algorithm	Paper ID	Comments	Identified Capabilities
	conversion into a Petri Net. PM			
Algorithms based on hidden Markov models (HMMs). 5 (16%)	HMMs with new proposed algorithm based on discriminative and generative classification methods	[P11]		Discovers flexible (manual) processes by estimating the most likely sequence of events. Models not only temporal aspects between the activities, but also hierarchical relationships between different levels (large-scale activities and simpler sub-activities)
	HMMs with Baum-Welch algorithm	[P12]		Discovers flexible (manual) processes by estimating the most likely sequence of events.
	HMMs with Probabilistic multi-class Support Vector Machine	[P13]		Discovers flexible (manual) processes by estimating the most likely sequence of events.
	HMMs with supervised learning	[P14]	using supervised learning allows to acquire and apply the model for robot assistance from the very beginning and then improve and attune to individual operator over time	Discovers flexible (manual) processes by estimating the most likely sequence of events.
	TOM4L [60]	[P44]		

Type of algorithm. N. of papers (%)	Algorithm	Paper ID	Comments	Identified Capabilities
Non-process discovery algorithms. 3 (9%)	Work cycle classification	[P5, P10, P22]		Discovers flexible (manual) processes based on the trajectories of movement of operator [P5, P10]. Discovers process based only on the location of the product in the production line (product ID, location, time) [P22]
N/A. 5 (16%)		[P20, P23, P24, P34, P46]		

Table G: Process conformance checking approaches

Approach	Number of papers (with %)	Paper ID
Replay	7 (33%)	[P2, P4, P9, P17, P29, P35, P46]
Alignment	11 (52%)	[P5, P10, P16, P18, P27, P30, P31, P33, P34, P38]
N/A	3 (14%)	[P26, P28, P36]
Total	21 (100%)	

Table H: Performance analysis approaches

Approach	Paper ID
Reinforcement learning	[P3]
Statistics	[P5]
Absolute values analysis	[P9]
Decision tree	[P9]
Part of process discovery	[P24]
Classification/clustering/visual mining	[P32]
Data flow discovery method [61]	[P37]
Trace variant visualizer [62]	[P37]
Ratio analysis	[P38, 45]
New proposed depth-first local search algorithm	[P40]
New proposed method based on searching procedure	[P43]
N/A	[P39]

Table I: Predictive PM approaches

Approach	Paper ID
Discrete event simulation	[P6]
Generic or template-based simulation	[P22]
Monte Carlo simulation	[P45]
Markov chain model	[P6, P7]
Machine learning	[P19, P20]

Table J: Action-oriented PM approaches

Constraint monitor		Paper ID	Action engine	Paper ID
PCC(O)	Replay	[P8, P23]	Queuing optimization algorithm	[P8]
PPM	HMMs	[P3, P13]	N/A	[P3, P13, P23]
	Machine learning	[P23]		

Table K: Industry 4.0 specific contribution (none for New customer service)

Industry 4.0 aspect	**Smart factories**	**Smart product**	**New business models**
Automated assistance in performing manual assembly processes	[P2, P3, P5, P8, P13]	[P8]	
Automated supervision of manual assembly production lines	[P3, P8]	[P8]	
Automated assistance to human supervisors in management of manual assembly processes	[P4, P5, P10, P18, P34]		
Automated assistance in management of manual assembly processes with help of digital twins	[P45]	[P45]	
Automated decision support for human supervisors	[P6, P20, P32, P37, P40, P43, P47]		
Automated assistance to human supervisors in monitoring and control of manufacturing processes	[P7-9, P19, P22, P26, P27, P30, P31, P44, P46]	[P7, P9, P22, P27]	
Automated assistance to human supervisors in real-time monitoring and control of multi-party business processes (supply chains)	[P28]	[P28]	[P28]
Automated supervision of manufacturing processes in a coordination context	[P23]		
Improvement of human-robotic cooperation during jointly performed tasks	[P11, P12, P14, P35]		
Training system for industrial operators in assembly tasks	[P15]		
Cybersecurity of production lines	[P16, P17]		
Complex analysis of processes: In logistics In manufacturing	 [P1, P21] [P24, P25, P29, P33, P38-42]	[P1]	
Alignment of IT systems with related Product Lifecycle Management for complex and dynamic IT landscape	[P36]	[P36]	[P36]
Number of papers (with %)	47 (100%)	8 (17%)	2 (4%)