Knowledge and Information Systems (2025) 67:4915–4948
https://doi.org/10.1007/s10115-024-02297-y

REVIEW

Process mining on sensor data: a review of related works

Edyta Brzychczy[1] · Milda Aleknonytė-Resch[2] · Dominik Janssen[3] · Agnes Koschmider[3]

Received: 16 May 2024 / Revised: 10 September 2024 / Accepted: 26 November 2024 /
Published online: 3 March 2025

Abstract
Process mining is an efficient technique that combines data analysis and behavioural process aspects to uncover end-to-end processes from data. Recently, the application of process mining on unstructured data has become popular. Particularly, sensor data from IoT-based systems allow process mining to uncover novel insights that can be used to identify bottlenecks in the process and support decision-making. However, the application of process mining requires bridging challenges. First, (raw) sensor data must be abstracted into discrete events to be useful for process mining. Second, meaningful events must be distilled from the abstracted events, fulfilling the purpose of the analysis. In this paper, a comprehensive literature study is conducted to understand the field of process mining for sensor data. The literature search was guided by three research questions: (1) what are common and underrepresented sensor types for process mining, (2) which aspects of process mining are covered on sensor data, and (3) what are the best practices to improve the understanding, design, and evaluation of process mining on sensor data. A total of 36 related papers were identified, which were then used as a foundation to structure the field of process mining on sensor data and provide recommendations and future research directions. The findings serve as a starting point for designing new techniques, enhancing the dissemination of related approaches, and identifying research gaps in process mining on sensor data.

Keywords Process mining · Sensor data · Event logs · Activity discovery · IoT

1 Introduction

Process mining is a widely used technique that employs automatic process analysis methods based on event data. These methods include algorithms for process model discovery, conformance checking between specifications and recorded events, and predictive analytics [36]. Typically, process mining relies on event logs extracted from IT system traces. The

Edyta Brzychczy, Milda Aleknonyte-Resch have contributed equally to this work.

✉ Dominik Janssen
dominik.janssen@uni-bayreuth.de

1 Faculty of Mechanical Engineering and Robotics, AGH University of Krakow, Krakow, Poland

2 Group Process Analytics, Kiel University, Kiel, Germany

3 Group Process Analytics, Bayreuth University, Bayreuth, Germany

advantage of process mining is that it combines data analytics techniques with behavioural process aspects to, for example, identify bottlenecks in the process. Process mining uncovers comprehensive insights regarding the end-to-end processes from data through the exposure and analysis of traces. Thus, process mining enables the discovery of valuable insights that would be challenging or impossible to detect using traditional Business Process Management (BPM) techniques.

Internet of Things (IoT)-based systems, which rely on (distributed) sensors, continuously generate large amounts of unstructured data. Process mining promises benefits from analysing sensor data as it allows for identifying patterns in unstructured data by detecting causal behavioural effects within the data. By applying process mining techniques, novel insights may be gained from sensor data collections related to process prediction or monitoring. In order to efficiently analyse sensor data with process mining methods and uncover hidden insights from data, two main challenges must be addressed:

Data Abstraction: Virtually, all techniques developed in process analytics assume discrete events at a relatively high level of abstraction, close to the business level. In contrast, sensor data are typically unstructured and may even come in continuous time-series data that needs to be transformed into discrete data. Therefore, the gap between “raw” sensor data and discrete events required for process mining has to be overcome.

Data Contextualisation: Sensor data need form and meaning to be understandable. In order to transform sensor data into meaningful information and prepare it for further analysis, it is necessary to process and present the data in a usable format. A systematic understanding is needed of what kinds of sensor data can be tackled by process mining, how to represent sensor data, and which types of sensor data are underrepresented for process mining to date.

To answer both challenges calls for a comprehensive understanding of related sensor types versus existing literature in process mining on sensor data. Some authors already suggested approaches to applying process mining to sensor data. Koschmider et al. [33] Brzychczy and Trzcionkowska [13] and van Eck et al. [23] have designed frameworks which involve variations of the following three main steps:

1. Preprocessing sensor data
2. Discovery/activity recognition
3. Event abstraction

Depending on the type of sensor (mobile or stationary), the approach to these steps may differ. For example, Koschmider et al. [33] focuses on stationary sensors, while van Eck et al. [23] describes mobile sensors, which require more extensive preprocessing due to the continuous nature of the sensor data. Soffer et al. [54] and Bertrand et al. [8] suggest a slightly different approach. They propose a framework based on a hierarchy of events, where a lower-level (less complex) event can be the foundation of one or several higher-level events. Specifically, Bertrand et al. [8] proposes a hierarchy model consisting of IoT (lower-level) events, which are the basis of process and context (higher-level) events. Moreover, some notable instances exist where sensor data are directly used for process mining tasks, namely for conformance checking [56].

Our paper comprehensively analyses previous studies on process mining applied to sensor data. We commence by conducting a systematic literature review that addresses the following research questions:

- **RQ1:** Which sensor types are common, and which are underrepresented for process mining?
- **RQ2:** Which aspects (in terms of use cases) have been covered for process mining on sensor data?
- **RQ3:** What best practices can be derived to improve the understanding, design and evaluation of process mining on sensor data?

Additionally, we elaborated recommendations backed by methodological considerations with the following benefits, which we will summarise in this paper:

Comparison and evaluation of existing literature: We provide the results of a literature review on process mining applied to sensor data. Currently, such an overview is missing. Instead, publications are scattered in various journals and other publication types. Given a use case, finding related work on process mining on that type of sensor data is challenging. Thus, the literature review presented in this paper fills this gap and allows for better comparison and evaluation of existing literature.

Recommendations for process mining on sensor data: These recommendations are suitable as a starting point for designing new techniques and understanding how to describe best-related approaches to increase their dissemination. Additionally, these recommendations aid in comparing and evaluating (new) approaches and publications with existing ones, allowing for an understanding of differences, commonalities, and areas for improvement.

Future research directions: The recommendations rely on a methodological foundation that identifies research gaps in the field to be tackled in the future.

The recommendations are a novel result, and, to the best of our knowledge, this is the only comprehensive overview focusing on process mining for sensor data that has been introduced in the literature before. In order to answer the three research questions, the rest of the paper is structured as follows. Section 2 summarises the terms *process mining*, *event log*, and *sensor data* and discusses the concepts in the light of their combination. This summary aims to partially address **RQ1**. Section 3 describes the structured literature search process and lists the search results in terms of a research classification, which finishes answering **RQ1** and relates to **RQ2**. The recommendations are summarised in Sect. 4, thereby addressing **RQ3**.

2 Concepts Related to Process Mining on Sensor Data

2.1 Process Mining and Event Log

Process mining is a fast-growing discipline related to process analytics and business process improvement. It bridges data and process science [1]. Process mining is a technique that aims to extract knowledge from event logs of various IT systems in order to gain insights into process behaviours and performance. According to van der Aalst, event logs are "*a structured representation of the history of an operational process, which captures the sequence of activities and the context in which these activities were performed*" [1]. We define the key concepts in process mining as follows:

Event: An event represents a single occurrence of an activity execution within a process. It typically includes attributes such as a unique identifier, timestamp, activity name, and additional attributes like the resource involved [1]. Events are the smallest units of analysis in process mining.

Case: A case is an instance of a process. It consists of a sequence of events (a trace) that belong to a single execution of the process. Each case follows a path through the process, depending on various factors such as decisions made and the order of events [2].

Trace: A trace is a sequence of events that occur in a particular case. It reflects the path taken by that case through the process from start to finish [1]. Traces show how individual process executions unfold over time [36].

Event log: An event log is a collection of traces, each representing a different instance (case) of the process. It serves as the primary data input for process mining.

Process: A process refers to a set of coordinated activities aimed at achieving a specific goal. In process mining, processes are analysed by observing patterns in event logs, helping to discover, improve, or verify workflows in real-world scenario [19].

Process Model: A process model is a formal representation of a process [1]. A number of notations exist to model processes, e.g. Petri nets, BPMN, etc.

The main applications of process mining are process discovery, conformance checking, performance analysis, process comparison, and predictive process mining [2].

The event log is the foundation of process mining, as it provides the data necessary to conduct process analysis. In general, event logs describe data on various aspects of process executions, like timestamps, process participants, activities, and life-cycle information of activities [19]. The data in the event log can be analysed using different process mining techniques, such as process discovery, conformance checking, and predictive analytics, to extract valuable insights into the performance of (business) processes [36]. There are three main requirements for the structure of an event log:

Case Identification: Each event in an event log should be linked to a case or process instance. A case identifier is essential for an event log as it enables the grouping of events based on the cases they belong to, making it possible to analyse the behaviour from start to finish [63]. However, in some cases, identifiers may not be available (e.g. industrial data, data streams), and one has to use methods to preprocess data. For instance, knowledge-based labelling [6] or context-based heuristics [13] can be employed to identify cases and link events to these cases.

Event activity correlation: Each event corresponds to an activity executed in the process [63]. More precisely, this means that there is an assumption stating that each event is mapped to a business process activity.

Order of the events: Within an event log, it is possible to establish an order of events within a case. While this can typically be determined by the timestamp of events [63], it is not always the case. Alternatively, the order of the events can be derived from an order of event records in databases.

Many different kinds of recorded events (e.g. sensor data) do not comply with these strict requirements [32]. However, as discussed in Sect. 2.3, preprocessing approaches can transform raw sensor data into a structured event log, fulfilling the above-mentioned requirements.

2.2 Sensor Types

Sensors are devices that detect and measure physical properties, such as temperature, pressure, or motion, and transmit a resulting input, i.e. signal to a control system [42]. There are two main categories of sensors: wearable and stationary. Wearable sensors are small and lightweight and are often used to monitor the health and activity of the wearer. They mainly observe *what* an entity is doing. However, the entity does not necessarily have to be human;

Table 1 Types of sensors

Sensor group	Examples	Description
Pressure & Force	Barometer Load cell Strain gauge	Pressure sensors measure the pressure of gases or liquids. Force sensors, or load cells,measure mechanical forces and provide digital signals that reflect the applied force magnitude
Position & Proximity	Gyroscope Accelerometer Motion sensor	Motion sensors detect the presence of human or objects in a particular area by sensing their infrared signature. Proximity sensors are electronic devices used to detect the presence of nearby objects without any physical contact. They usually emit an electromagnetic field and look for changes in the field
Acoustic & Light	Microphone Photoelectric sensor	Acoustic sensors detect sound waves through its intensity and converting it to electrical signals. Photoelectric sensors detect emitted photons from light sources
Chemical & Radiation	Particle detector Gas detector	Chemical sensors are able to sense chemical reactions, chemical substances or sets of chemicals
Electric & Magnetic	Magnetometer Current sensor	Electrical sensors are electronic devices that sense current, voltage, etc. Magnetic sensor detect magnetic fields in the form of flux, strength and directions
Environmental	Thermometer Humidity Rain sensor	Environmental sensors observe the condition of the surroundings
Flow sensor	Mass flow Water flow	Flow sensors are electronic or electro-mechanical devices used to measures the flow of a fluid such as a gas or liquid
Flaw sensor	Ultrasonic detectors	Flaw sensors detect surface or material inconsistencies

Classification is based on [27, 30, 48, 49]

it can also be livestock or a package in a smart factory. Examples of wearable sensors include heart rate monitors, accelerometers, and electrocardiograms. Wearable sensors often produce continuous recordings such as heart rate and movement acceleration. In process mining, converting this continuous data into discrete events requires extensive preprocessing.

On the other hand, stationary sensors mainly observe *where* entities are moving or collecting information about environmental conditions. Examples of stationary sensors include temperature sensors, pressure sensors, humidity and movement sensors. Stationary sensors may also produce continuous data such as temperature, pressure, and humidity readings. However, stationary sensors (such as movement sensors) often produce discrete data as they are activated only when triggered. While discrete data may present clear events (e.g. presence of movement in a room), for stationary sensors, the challenge in process mining often lies in accurately associating these events with the correct case ID [35].

The sensor type should be carefully selected depending on the analysis's goal. Generally, wearable and stationary sensors can be further classified according to the seven categories depicted in Table 1. The table highlights the vast number of types of sensors. Each sensor type can be used for different use cases such as monitoring, learning about systems [53] or enhancing data (e.g. temperature sensor, camera, heart rate sensor) during process execution in order to improve the process [61]. Furthermore, sensor applications and types can be classified based on the various aspects [27, 53] as shown in Table 2.

Table 2 Overview of classification aspects used to describe sensors

Domain	Smart environments (home, office), health, smart devices, industrial
Application	Process monitoring, testing and qualification, fault prediction, detection, system identification (experimental modelling), product quality assessment, diagnosis, warning generation, surveillance, controlling a system)
Reporting	Active - always gathering data passive - gathering data only when triggered
Detection	Radioactive, biological, electrical, thermal
Conversion	Electro-chemical, electromagnetic, thermo-electric
Output	Analogue, digital
Measure	Biomedical, chemical, electrical/electronic, mechanical, thermo-fluid

Based on [27, 53]

Table 3 An example of industry sensor data

Timestamp	Sensor 1	Sensor 2	Sensor 3	Sensor 4
…	…	…	…	…
12:59:57	237	325	0	122
13:06:21	237	400	1	122
13:06:22	240	450	1	123
…	…	…	…	…

(a) Raw sensor event log

⟶

CaseID	Activity	Timestamp	Duration
…	…	…	…
1	welding	12:59:57	6:25
2	transporting	10:07:22	1:06
…	…	…	…

(b) Aggregated sensor event log

(a) depicts raw data retrieved from sensors, while (b) shows how the data from (a) can be transformed into an event log necessary for process mining

The multitude of different sensor types and the numerous possible fields of application for sensors result in a wide range of use cases for process mining on sensor data. In this paper, we investigate the current extent of process mining's application for sensor data by addressing **RQ1:** Which sensor types are common, and which are underrepresented for process mining? and **RQ2**: Which aspects (in terms of use cases) have been covered for process mining on sensor data?

2.3 Combination of Process Mining on Sensor Data

Raw sensor data typically cannot be used for process mining tasks without appropriate preprocessing. Sensor data are unstructured; this prevents the proper extraction of the necessary information for the desired application of process mining [12]. Most process mining methods require the data (event log) to be of a high abstraction level, i.e. business level. An example of raw sensor data from the industry domain can be seen in Table 3a. These data require preprocessing, including event abstraction and correlation, to derive an event log as seen in Table 3b. Event abstraction aggregates low-level event data into higher-level events that represent the execution of activities. Afterwards, these abstracted events must be correlated to individual cases [21].

The application of process mining on sensor data requires a structured approach, as presented in [34, 66]. The pipeline consists of the following three steps:

1. Preprocessing raw data

2. Abstraction/aggregation of events/activities
3. Process mining on aggregated event data

The first step, data preprocessing, involves data cleaning. Errors, outliers and redundant information are removed, and missing values are handled using various imputation or interpolation techniques. In the second step, data are transformed, including normalisation or discretisation. In the last step, the data are divided into segments using one of three main segmentation approaches: temporal-based, activity-based, or sensor event-based [44]. Each approach identifies smaller blocks of information that can be analysed more efficiently.

In the field of event abstraction for process mining, van Zelst et al. [66] conducted a comprehensive survey and classified related works into four dimensions:

- Supervision strategy (supervised vs. unsupervised)
- Fine-granular event interleaving (e.g. the capability of event abstraction techniques to handle true concurrency at a higher level of granularity)
- Probabilistic nature of outcome (e.g. deterministic vs. probabilistic output)
- Data type (discrete vs. continuous data)

In the papers van Zelst et al. [66] and Koschmider et al. [34], the authors outlined challenges for this stage of the process analytics pipeline. These challenges include a need for more attention given to continuous data sources in existing literature, the tendency for most techniques to produce a discrete output, and a focus on offline rather than online processing. Bertrand et al. [9] also highlight that the described pipeline poses a catch, as the derived high-level event log may lose context information inherent in the original data. De Luzi et al. [18] discuss the interplay between BPM and IoT, outlining key research challenges for realisation. Meanwhile, Koschmider et al. [35] discuss 14 challenges associated with process mining on unstructured data. Among these challenges, they highlight crucial questions, including how to effectively synchronise sensors to ensure accurate data collection, how to contextualise activities correctly to interpret data meaningfully, what constitutes a suitable event to activity abstraction or aggregation to simplify complex data into actionable insights, and how to identify entity centricity to focus on the most relevant data points for process mining. Furthermore, the limitations of XES and OCEL formats have led to a need for research on IoT-enhanced event log models. While IoT-enhanced event log models are outside our work's scope, their main challenges and requirements are identified and presented in [9, 10].

The discovery of process models from raw to event sensor data is in high demand. The BPM-IoT manifesto [28] discusses several challenges concerned with process mining on higher levels of knowledge (i.e. in our context abstracted events):

Integrating IoT into the correctness check of processes: The discovered processes should specifically consider the IoT nature of some components.

Detecting new processes from data: How to consider situational knowledge (specific for IoT data) into process discovery.

Dealing with new situations: How to consider ad-hoc (real-time) decisions into process discovery, involving a continuous change of processes.

Specifying the autonomy of new things: How to integrate the concept of autonomy (i.e. to grant things full autonomy to decide) into process discovery.

We conducted our structured literature review to understand the current state of the art in applying process mining to sensor data. Given the complexities sensor data poses for process mining, it is intriguing to analyse how existing studies have approached these issues. This review not only sheds light on how sensor types are utilised and the use cases explored but also identifies best practices for process mining with sensor data, directly addressing our research questions.

Table 4 Review protocol used to conduct the systematic literature review

			Remaining Papers
Inclusion	Article, conference paper, book chapter		
Search channels	Scopus	*214 papers*	345
	WoS	*131 papers*	
Additional restrictions	Remove duplicates	120 papers	225
	Remove papers published before 2014	11 papers	214
	Remove non-English papers	2 papers	212
Screening for exclusion	Abstract is unrelated	84 papers	128
	Only extended abstract or short papers	9 papers	119
	Conference paper extended in journal	5 papers	114
	Lack of keywords in abstract, title or author keywords	5 papers	109
	Full paper unavailable	6 papers	103
Quality assessment	Paper without sensor data examples	68 papers	35
	Paper without process mining	3 papers	**32**
Backward–forward search	Articles, conference papers, book chapters	*758 papers*	
	Remove irrelevant papers	754 papers	**4**
		Total:	**36**

Numbers in Italic indicate the initial number of papers identified in the described step

3 Systematic Literature Review

This section outlines our methodology for conducting the literature search, including the selection and classification process of sources in our bibliography. This literature review was guided by the search process presented in [65].

3.1 Methodology

Our protocol consists of the following steps: the identification of channels for the literature search, keywords selection, inclusion strategy, refining results with additional restrictions, the definition of exclusion criteria, a screening procedure for exclusion, quality assessment criteria, quality assessment procedure, backward and forward search, data extraction, analysis and reporting of the findings. The search was conducted on the 10$^{\text{th}}$ of November 2023. The review protocol is presented in Table 4.

Channels for literature search: Our primary data source was the Scopus database, the largest database of peer-reviewed literature [67]. To increase the validity of our literature search, we also searched the ISI Web of Science (WoS) database.

Used keywords: We performed the search process incorporating the commonly applied keywords for sensor data and process mining in the search term:

“process mining” AND (“time series” OR “low level” OR “sensor data” OR (“IoT” OR “IIoT”))

Inclusion strategy: In our sample, we included articles, conference papers, and book chapters.

Refining results: Firstly, we removed duplicates from the dataset. We found 120 duplicate items. As a result, we identified 225 unique papers for further analysis. In order to ensure the relevance and contemporary applicability of our findings, we mandated that the selected publications be in English and have a publication date within the last ten years (publication year > 2013). This time frame was chosen to capture recent advancements in the rapidly evolving fields of process mining and sensor data. Consequently, we removed 13 publications older than ten years or two written in Chinese. Eventually, we identified 212 papers for further consideration.

Definition of exclusion criteria: We defined the following exclusion criteria:

- Lack of keywords in abstract
- Keywords appear relevant but are used in a different context or domain than our research focus
- Extended abstracts or short papers
- Conference papers extended in journal (in this case, we favoured journal papers)
- Abstract content is not related to process mining on sensor data
- Full paper being not available

Screening procedure for exclusion: Two researchers independently reviewed all papers in parallel. In cases of doubts, decisions were made after discussing with a third researcher not involved in the initial review process. A total of 109 papers were excluded for the following reasons, listed by frequency:

- Abstract content not relevant to the study (84)
- Extended abstracts or short papers (9)
- Lack of keywords in the abstract, title, or author's keywords or used in other context (5)
- Papers that were not fully available (6)
- Conference papers re-published in a journal (5)

After excluding these papers, 103 publications were identified for further evaluation.

Quality assessment criteria: Our leading quality inclusion criteria for papers were that the paper applied a method or suggested a framework on an exemplary sensor data set in the context of process mining (review papers were discarded). Papers presenting sensor data without explicit connection to process mining are not the subject of our study.

Quality assessment procedure: Similarly to the exclusion procedure, all papers were checked independently by two researchers in parallel. In case of doubts, the decision was made after a discussion with a third researcher not involved in the procedure. We rejected 68 papers without sensor data examples and 3 papers without process mining content.

Backward and forward search: We performed a backward and forward search for papers that passed the initial quality assessment procedure using the *Citationchaser* tool [25]. Consequently, a total of 758 additional papers for which the protocol from the steps "inclusion strategy" to "quality assessment procedure" were repeated. This resulted in 4 additional papers that were included in our literature review.

Data extraction: Data about the topic and form, as well as content, were analysed. Firstly, we analysed the distribution of papers based on the time and place of publication in terms of quantity and quality by extracting the year and type of publication related to our scope. Secondly, we extracted relevant data from each publication for our defined RQs. We considered the aspects shown in Table 5.

Our final data set included 36 papers. The results of our literature analysis are presented in the following subsections.

Table 5 Summary of relevant data extracted from each publication in the structured literature review

Category	Options
Domain of Sensor Implementation	Smart environments (home, office), healthcare, smart devices, industrial, other, not specified
Sensor Purpose	Activity recognition, process monitoring, product quality assessment, fault prediction, warning generation/surveillance
Type of Sensor Used	Motion, vital signs, etc.
Type of Sensor Output	Numerical, categorical, mixed
Abstraction of Data	Low-level events, high-level events, activity instance
Processing Methods of Sensor Data	Supervised, unsupervised
Type of Research Artefact	Framework, conceptual
Type of Data Set	Own, publicly available
PM Applicationpresented in the paper	Event log creation, process discovery, conformance checking, enhancement

3.2 Literature Summary

A network of keywords present in the analysed publications can be seen in Fig. 1. The 19 recurring keywords, each appearing in at least two papers, are grouped into three clusters. Process mining, IoT, and sensor data emerge as the most interconnected keywords. Figure 2 shows the number of publications published annually. A growing number of publications can be observed, ranging from one article in 2014 to nine in 2022. It is worth emphasising that most papers throughout the whole time span are conference papers (22 items).

The outcomes of our literature review are summarised in Table 6, detailing the distribution of publications in the field of process mining applied to sensor data. Of these, 17 publications propose frameworks addressing various aspects of process mining on sensor data. The remaining publications focus on specific problem-based or use-case-based artefacts. A significant portion of the literature (14 items) is related to either the industrial domain (including mining [13], manufacturing [41], paper and mill [43], retail [11]), and smart home and office (12 publications) (see Fig. 3 and Table 6 for specific publications).

As can be seen in Table 6 column *Dataset*, most of the publications (23) apply their own datasets, referring mainly to the industrial domain (14), healthcare (4) and other domains such as smart office, smart product (1), and maritime (1). The public sources of the used data are:

Ordonez B [3]: The dataset includes sensor data from 12 sensors and 11 activities related to home residents' behaviour over 21 days.

MIT home dataset [46]: The dataset includes two independent logs from two homes over a period of two weeks. The datasets contain sensor data from two homes, one with 77 and the other with 84 sensors and 16 defined activities.

Kasteren A [59]: The dataset includes sensor data from 14 sensors with ten activities related to the behaviour of home residents over a period of 25 days.

CASAS datasets [17]: The CASAS datasets include sensor data and activities related to the behaviour of home residents for various test settings, e.g. the Milan dataset [15] containing sensor data from 33 sensors and 15 activities related to the behaviour of home residents over a period of 92 days. The other dataset—Kyoto—was used in [29] and Aruba in [38]. https://casas.wsu.edu/datasets/

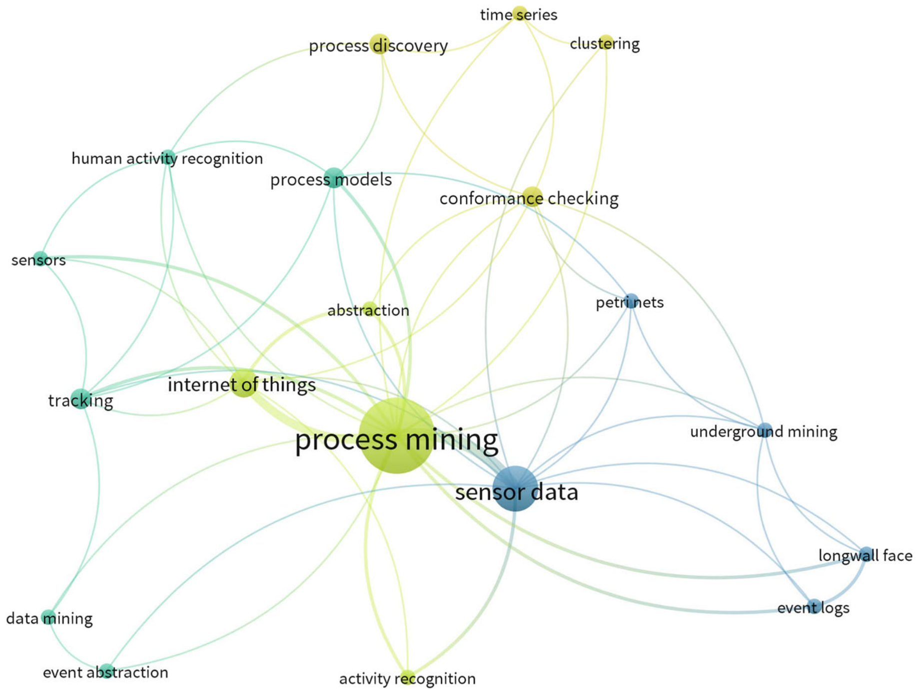
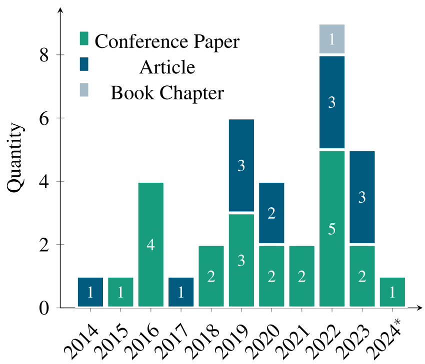


Fig. 1 Network of keywords which occurred at least twice among the 36 papers identified in this review. Visualisation was derived using VOSviewer

Fig. 2 Number of publications over time



MIMIC-III [4, 5]: The dataset includes records referring to demographics, admissions, and clinical observations of almost fifty thousand patients who stayed in critical care units between 2001 to 2012, including data from bedside monitoring (e.g. bed occupancy) and medical tests (e.g. respiratory rate, heart rate, oxygen saturation, blood pressure). https://mimic.mit.edu

Other: e.g. from BP-meets-IoT challenges [24, 37].

Most publicly available datasets are related to the smart home domain, primarily using discrete sensor readings. Thus, datasets containing continuous sensor readings for analysis and experimentation are still rare in the public domain.

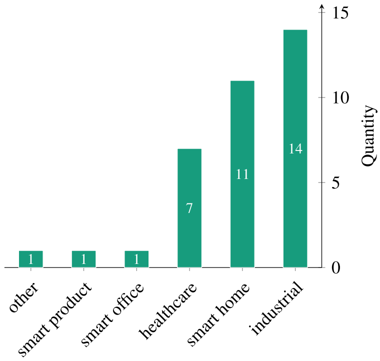


Fig. 3 Number of publications by domain

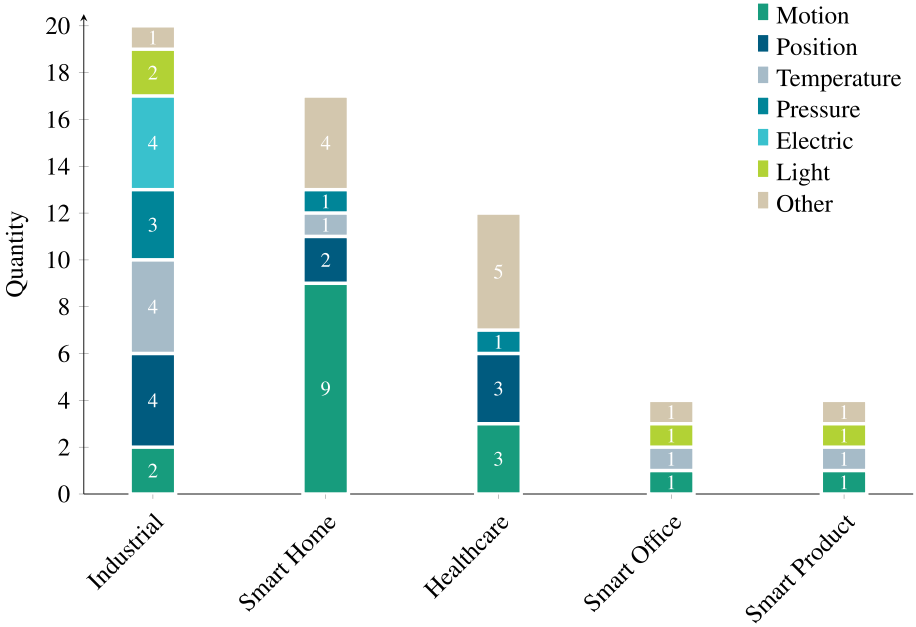


Fig. 4 Number of publications mentioning sensor types used in specific domains

Regarding streaming data, only three of the selected papers consider streaming data. Seiger et al. [50] propose a framework that uses complex event processing (CEP) to handle sensor data streams in real-time, enabling live analysis and correlation of high-volume data without relying on traditional BPM systems. This approach allows for better handling of sensor data variations. It helps generate process event logs from IoT data in real time, addressing the limitations of traditional process mining algorithms in capturing dynamic sensor data. Wolny et al. [64] highlight the challenge that traditional process mining algorithms are often designed for well-structured event logs rather than for the continuous, real-time nature of sensor data streams. The authors propose an approach called Model-driven Runtime State IdEntification (MD-RISE) to address this. This approach derives time-series queries from state machines that monitor and validate system states based on sensor data streams. Mayr et al. [41] claim that their approach can also be applied to streaming data by grouping incoming mini-batches of process-state data with historically derived clusters.

3.3 Main findings on sensors usage in process mining applications

We further identified the types of sensors used in the publications (Fig. 4). The most often used sensors are motion sensors (16 publications), position (9 publications), temperature (7 publications), and pressure (5 publications)—see Table 6, column *Type of Sensor*. Examples of other sensors include occupancy, light, acoustic, water, power, electric, chemical, or vital signs **(RQ1)**. Looking at specific sensor types by domain, we found that motion sensors are primarily used in smart home and smart office applications (10 papers out of 16, see Table 6, column *Type of Sensor* and *Domain*). Position, temperature, and pressure sensors are mostly used in industrial domains but also in healthcare and smart domains. Occupancy sensors are preferred for smart home and office applications. Electric or chemical sensors are used in the industrial domain. Furthermore, we categorised the main objectives for using sensors into three key areas:

Process monitoring. Process monitoring is defined as knowing what the process looks like, and we monitor its behaviour and deviations from it based on sensor data.

Activity recognition. Activity recognition is the process of discovering and modelling unknown activities, aiming to understand their characteristics and behaviour clearly.

Warning generation/surveillance. The primary purpose of warning generation/surveillance is to identify potentially dangerous situations and issue a warning. This type of objective is primarily related to the healthcare domain.

In the collection of relevant papers, we identified 16 cases of process monitoring and activity recognition, along with four cases of warning generation/surveillance. This analysis revealed that specific objectives are closely linked to the domain of application (Fig. 5). For instance, process monitoring is the predominant objective in industrial applications (12 out of 14 cases), whereas activity recognition is most common in smart homes or offices, with 11 out of 13 cases. In the healthcare domain we identified three cases each of activity recognition and warning generation/surveillance. For specific papers, see Table 6, columns *Domain* and *Objective*.

When examining different objectives for using sensors (Table 6, column *Objective*), we saw that the widest variety of sensor types is used for activity recognition (Fig. 6). We noticed that most sensor types have been used for that purpose except for electric and chemical sensors. In contrast, process monitoring predominantly uses position, motion, temperature, electric, and pressure sensors. It is noteworthy that the majority of the papers focus on low-level data abstraction. Specifically, 27 of the publications used data on this level. Additionally, in terms of the type of sensor data output, categorical and numerical data were the most commonly presented types, represented in 18 and 8 papers, respectively. Six publications, primarily related to the industrial domain, presented mixed variables. In [4], numerical variables were used to extend event logs. A detailed description of the variables for each publication can be found in Appendix 1.

Our analysis revealed that most publications (16) employ supervised methods in their analysis pipeline or a mixture of unsupervised and supervised approaches (11) (see Table 6 column *Processing Methods*). Based on the sensor usage objectives, we assessed these processing approaches (whether supervised, unsupervised or a mixture of both). For activity recognition, both unsupervised and supervised methods are common. Process monitoring and warning generation/surveillance predominantly utilise supervised techniques, particu-

Table 6 Overview of systematic literature review

Citation	Title	Domain	Type of sensor	Processing methods	Objective	Dataset
Al-Ali et al. [3]	A composite machine-learning-based framework for supporting low-level event logs to high-level business process model activities mappings enhanced by flexible BPMN model translation	Smart home	Motion	k-medoids, DT, RF, ANN, MLP, k-NN, SVM	Process monitoring	OrdonezB
Prathama et al. [46]	A multi-case perspective analytical framework for discovering human daily behaviour from sensors using process mining	Smart home	Motion	None	Activity recognition	MIT home
Vitale et al. [62]	A Process Mining-based unsupervised Anomaly Detection technique for the Industrial Internet of Things	Industrial	Vibration, tension, electric	LSTM	Process monitoring	Own
Lofù et al. [40]	A Situation Awareness Computational Intelligent Model for Metabolic Syndrome Management	Healthcare	Weight, chemical, pressure	Supervised	Surveillance / warning generation	Own
Mayr et al. [41]	Abstracting Process Mining Event Logs from Process-State Data to Monitor Control-Flow of Industrial Manufacturing Processes	Industrial	Temperature, gas, density	Cluster Analysis	Process monitoring	Own
Seiger et al. [51]	An Interactive Method for Detection of Process Activity Executions from IoT Data	Industrial	Position, temperature, light, pressure	Supervised	Process monitoring	Own
Bayomie et al. [7]	Analyzing Manufacturing Process By Enabling Process Mining on Sensor Data	Industrial	Temperature, pressure, volume, speed	Supervised	Process monitoring	Own
Szpyrka et al. [56]	Conformance checking of a longwall shearer operation based on low-level events	Industrial	Electric	None	Process monitoring	Own

Table 6 continued

Citation	Title	Domain	Type of sensor	Processing methods	Objective	Dataset
Brzychczy and Trzcionkowska [13]	Creation of an event log from a low-level machinery monitoring system for process mining purposes	Industrial	Electric	Hierarchical clustering	Process monitoring	Own
Nadim et al. [43]	Data-driven dynamic causality analysis of industrial systems using interpretable machine learning and process mining	Industrial	n/a	Inductive rule-based classifier	Surveillance / warning generation	Own
Su et al. [55]	Data-Driven Goal Recognition in Transhumeral Prostheses Using Process Mining Techniques	Healthcare	Motion, vital signs	Aggl. hierarchical clustering, LSTM	Activity recognition	Own
Cameranesi et al. [15]	Discovering process models of activities of daily living from sensors	Smart home	Motion, position, temperature	Supervised	Process monitoring	CASAS
Sztyler et al. [57]	Discovery of personal processes from labeled sensor data - An application of process mining to personalized health care	Healthcare	Motion, position	Supervised	Process monitoring	Own
Elali et al. [24]	Domain Ontology Construction with Activity Logs and Sensors Data - Case Study of Smart Home Activities	Smart home	Occupancy, power, water	Supervised	Activity recognition	BPI 2021 Challenge
van Eck et al. van Eck et al. [23]	Enabling process mining on sensor data from smart products	Smart product	Temperature, motion, light, acoustic	U-Shapelet Clustering	Activity recognition	Own
Rebmann et al. [47]	Enabling the Discovery of Manual Processes Using a Multi-modal Activity Recognition Approach	Industrial	Motion	k-Means	Activity recognition	Own
Tax et al. [60]	Generating time-based label refinements to discover more precise process models	Smart home	Motion, power	Expectation-Maximization	Activity recognition	Kasteren A, MIT home
Dogan et al. [22]	Individual Behavior Modeling with Sensors Using Process Mining	Smart home	Motion, position	Clustering	Activity recognition	Own
Carolis et al. [16]	Learning and recognizing routines and activities in SOFiA	Smart office	Motion, occupancy, light, temperature, pressure	Supervised	Activity recognition	Own

Table 6 continued

Citation	Title	Domain	Type of sensor	Processing methods	Objective	Dataset
Blank et al. [11]	Location-Aware Path Alignment in Process Mining	Industrial	Position	Sweep line algorithm	Process monitoring	Own
Tax et al. [59]	Log-based Evaluation of Label Splits for Process Models	Smart home	Motion	Decision tree learning	Activity recognition	Kasteren A
Wolny et al. [64]	Model-driven Runtime State Identification	Industrial	Position	Supervised	Process monitoring	Own
Ziolkowski et al. [68]	Process Mining for Time Series Data	Other	n/a	Dynamic Time Warping	Process monitoring	Own
Hwang and Jang [26]	Process Mining to Discover Shoppers' Pathways at a Fashion Retail Store Using a WiFi-Base Indoor Positioning System	Industrial	Position	None	Process monitoring	Own
Janssen et al. [29]	Process Model Discovery from Sensor Event Data	Smart home	Motion	Self-Organising Map	Activity recognition	CASAS
Bano et al. [5]	Process-aware digital twin cockpit synthesis from event logs	Healthcare	Occupancy	Supervised	Surveillance / warning generation	MIMIC-III
Brzychczy and Trzcionkowska [14]	Process-oriented approach for analysis of sensor data from longwall monitoring system	Industrial	Electric	hierarchical clustering	Process monitoring	Own
Khowaja et al. [31]	PROMPT: Process Mining and Paravector Tensor-Based Physical Health Monitoring Framework	Healthcare	Motion, position	Quadrature filters	Activity recognition	[58]
Li et al. [39]	Rectify Sensor Data in IoT: A Case Study on Enabling Process Mining for Logistic Process in an Air Cargo Terminal	Industrial	Motion	Supervised	Process monitoring	Own
de Leoni and Pellattiero [37]	The Benefits of Sensor-Measurement Aggregation in Discovering IoT Process Models: A Smart-House Case Study	Smart home	n/a	k-Means, DBScan	Activity recognition	BPI 2020 Challenge

Table 6 continued

Citation	Title	Domain	Type of sensor	Processing methods	Objective	Dataset
Senderovich et al. [52]	The ROAD from sensor data to process instances via interaction mining	Healthcare	Position	Supervised	Activity recognition	Own
Seiger et al. [50]	Towards IoT-driven Process Event Log Generation for Conformance Checking in Smart Factories	Industrial	Temperature, light, pressure	supervised	process monitoring	own
Porouhan[45]	Using process mining for predicting relationships of couples sitting on a sofa	Smart home	Pressure	Supervised	Activity recognition	Own
Di Federico and Burattin [20]	vAMoS: eVent Abstraction via Motifs Search	Smart home	Motion	k-Means	Activity recognition	CASAS
Leotta et al. [38]	Visual process maps: a visualization tool for discovering habits in smart homes	Smart home	Motion	TRACLUS	Activity recognition	CASAS
Banham et al. [4]	xPM: A Framework for Process Mining with Exogenous Data	Healthcare	Vital sign	Supervised	Surveillance / warning generation	MIMIC-III

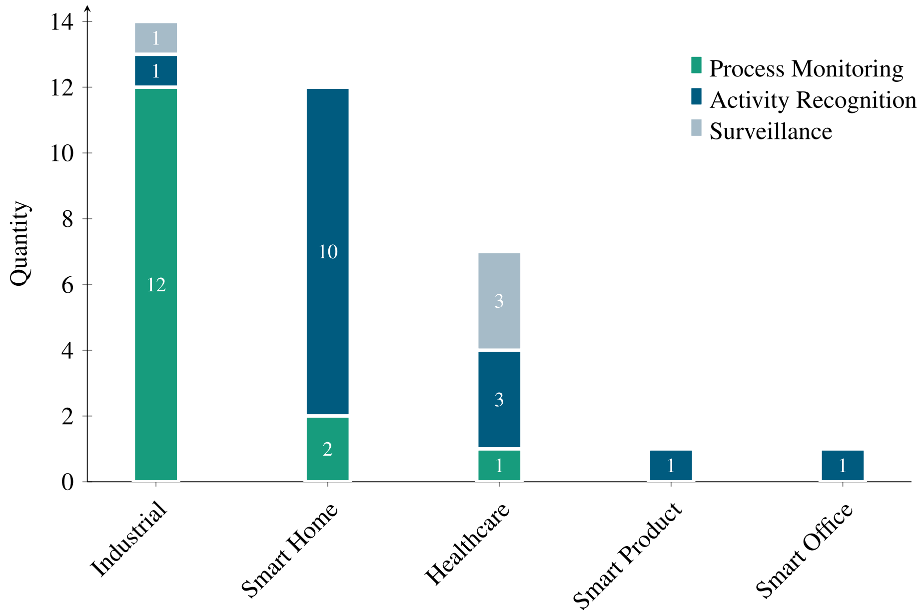


Fig. 5 Number of publications by domain and purpose of analysis

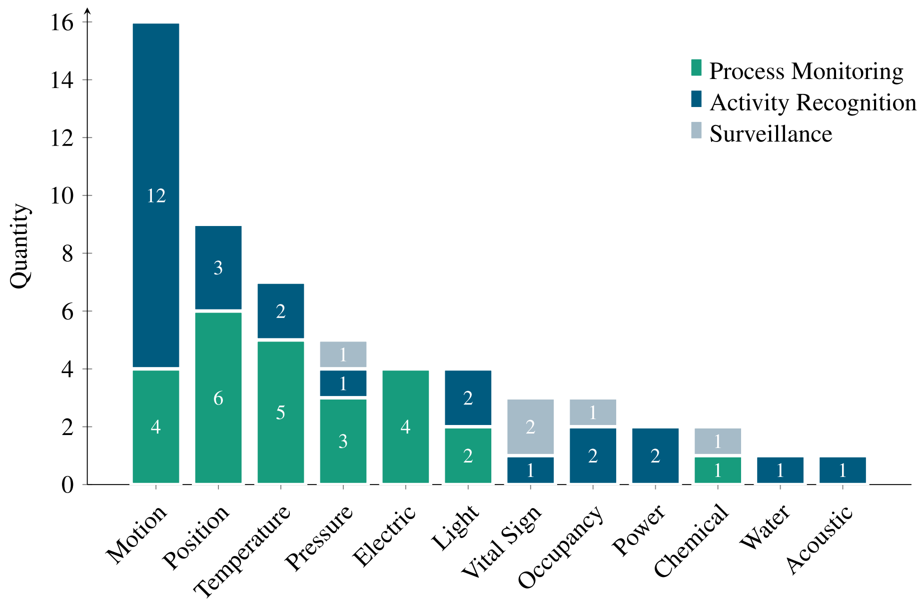


Fig. 6 Number of publications by sensor types and analysis purpose

larly classifiers and expert rules. When examining the application of these methods across different domains, a pattern can be seen. The healthcare domain primarily implements supervised approaches. Conversely, unsupervised and supervised approaches are common in the smart home and industrial domains. Interestingly, in four publications, neither supervised nor unsupervised approaches were implemented, with two cases each in the industrial (specifically retail) and smart home sectors. Object centric process mining does not play a role in any of the examined papers.

The variation of tools for applying process mining methods on sensor data used in the analysed publications was limited. ProM was used in most publications (10 publications), followed by Disco (4 publications). A few publications (3 publications) used Pm4Py as their tool of choice. However, many publications (20 publications) did not explicitly mention which tools they used. Table 7 in the appendix shows which publications used which tools.

While addressing **RQ2**, which focuses on the use cases covered by process mining on sensor data within our collection of related publications, we discerned a primary focus on event log creation (29 publications) and process discovery (26 publications)—see Appendix 1. This emphasis underscores a fundamental challenge in process mining applied to sensor data: developing effective methods for generating suitable event logs to facilitate subsequent process mining analyses. Conformance checking and enhancement received attention in six and five papers, respectively, indicating these areas as less explored applications of process mining.

Furthermore, we examined process mining applications by domain (Fig. 7). In the industrial domain, the creation of event logs (12 publications) and process discovery (7 publications) emerged as prevalent tasks, highlighting the complexity of transforming raw data into structured event logs. In smart home/office and healthcare settings, process discovery (17 publications) and event log creation (15 publications) were predominant, reflecting similar challenges across these domains. Conformance checking and enhancement tasks were more frequently associated with the healthcare and industrial domains. Notably, enhancement techniques were mentioned in smart home and healthcare papers, while the industrial domain featured minimal focus on enhancement (1 publication) and conformance checking (2 publications). For specific papers, see Table 6, column *Domain* and Appendix 1. This distribution illustrates the varied application of process mining tasks across domains, with an emphasis on the foundational steps of event log creation and process discovery.

The limitations regarding the application of process mining on sensor data varied greatly in the analysed publications. The following limitations, however, appeared in several publications:

Lack of Domain Knowledge: The approach presented in Ziolkowski et al. [68] depends on manual cluster naming and domain knowledge that may only sometimes be available. Li et al. [39] refer to legacy systems where the limited availability of domain knowledge adds to the complexity of their approach. The interpretation of activities the approach creates, presented in van Eck et al. [23], requires significant domain knowledge. The challenge is finding a connection between measured sensor readings and underlying user behaviour.

Computational Limits: Computational limits are challenging when approaches like the one presented in Su et al. [55] deal with alignment-based conformance checking methods. The authors in Senderovich et al.[52] faced the challenge of exponentially increasing the size of the considered interaction set and the resulting computational limitations.

Data Quality & Quantity: The publication about WiFi-location data Blank et al. [11] states that data quality, especially regarding location accuracy, is a limitation in their current approach. Li et al. [39] saw it as a challenge in their case study to deal with sensor

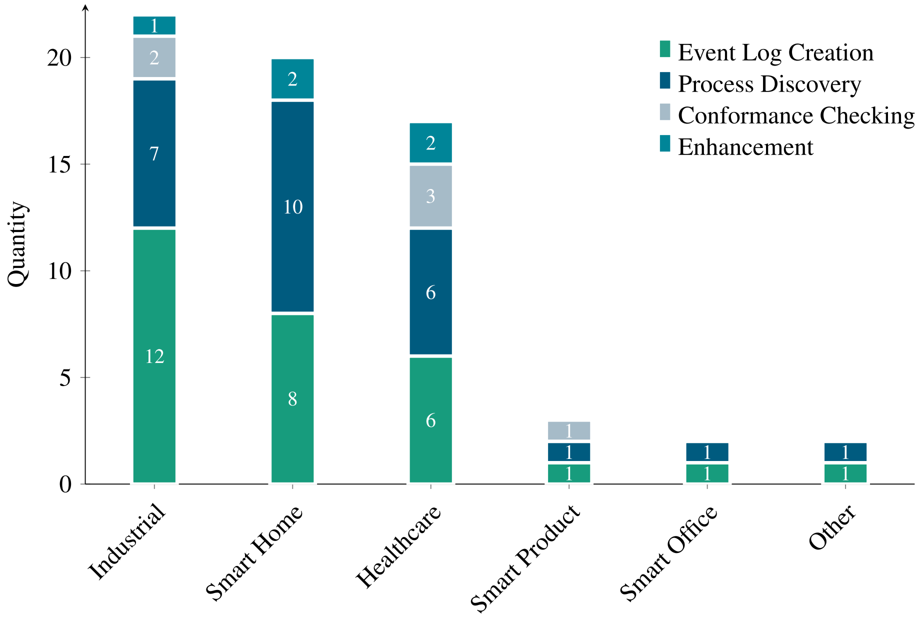


Fig. 7 Number of publications by domain and process mining task implemented

malfunctions, missing readings, and communication overhead. Dogan et al.[22] write that the data quality needs to be assessed further and deep data processing is necessary to remove errors. In Elali et al. [24], the authors state that they would need more data to elaborate further on the connection between activities and sensors. In Rebmann et al. [47], the authors face the challenge that insufficient training data is available for underrepresented classes.

4 Recommendations and future research directions

In order to derive recommendations (addressing **RQ3**) which would improve the understanding, design and evaluation of process mining on sensor data, we analysed the analysis procedures used in the set of related publications. Generally, a process mining pipeline subsumes the following tasks or inputs. We inspected the publications with regard to these tasks:

- **Raw data or event log**—the artefact used as input to analytic workflow, i.e. the initial dataset or (low-level) event log used as input for analysis.
- **Data preprocessing**—activities related to standard data cleaning, e.g. removal of missing data.
- **Event processing**—the processing related to an event level, e.g. trace clustering, sessions of events.
- **Labelling data**—the naming of discovered groups of events (typically a result of unsupervised techniques).
- **Event relabelling**—the renaming of event names in existing event log to reflect new or refined categories, e.g. merging similar activities.

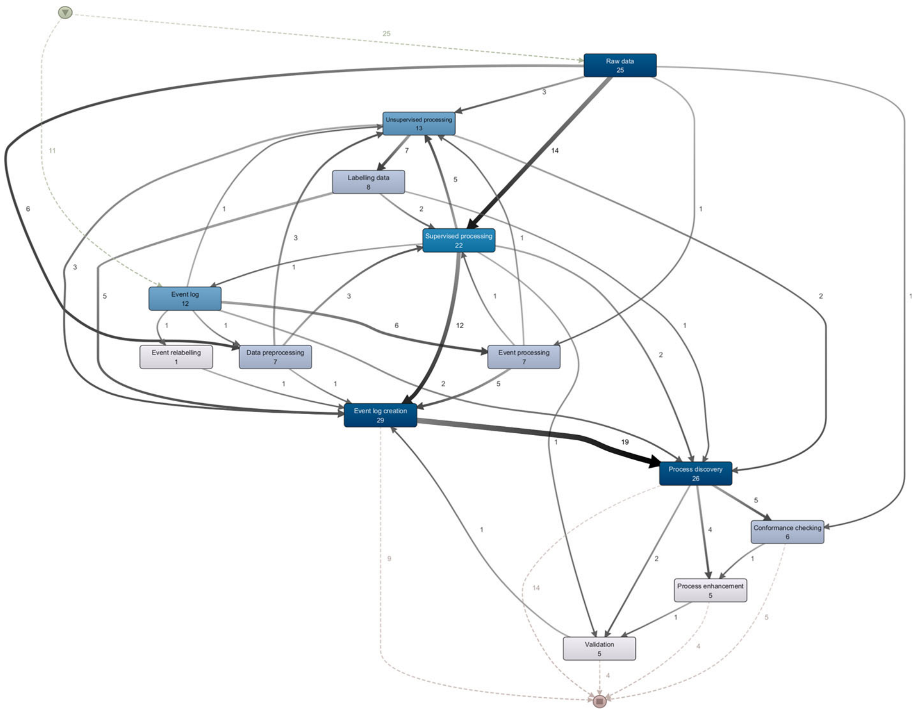


Fig. 8 The process map of analytic workflows presented in papers depicting 90% of paths

- **Supervised processing**—the processing related to raw sensor data with use of supervised approaches, e.g. classification or regression.
- **Unsupervised processing**—the processing related to raw sensor data with use of unsupervised approaches, e.g. clustering or dimensionality reduction.
- **Validation**—the evaluation of proposed method(s) or framework on various datasets.
- **Process mining tasks**—activities such as creating event logs, discovering processes, checking conformance and enhancing processes.

Based on the analysis procedure described in the publications, we designed a process map as shown in Fig. 8. From this process map we can see that the most common stages in the analytics pipelines leading to the process mining tasks included: (gathering) raw data (25 publications), supervised processing (22), in some cases followed by unsupervised processing (13) and labelling data (8 publications), leading to event log creation (requiring additional steps). We observed 28 different paths for the 36 inspected publications.

The paths identified in our analysis, as illustrated in Fig. 9, most often involve the supervised processing of raw data, leading to event log creation and process discovery, as seen in four publications. A second notable path mirrors the first, going directly from raw data to event log creation. Another common approach, noted in two publications, involves event processing that supports the generation of event logs.

We analysed the methods applied across the reviewed literature. We observed both supervised and unsupervised approaches in the initial processing of raw data. Predominantly, supervised processing approaches were used, comprising of a variety of methods including self-annotation, e.g. in [16], leveraging domain knowledge [5], and employing classifica-

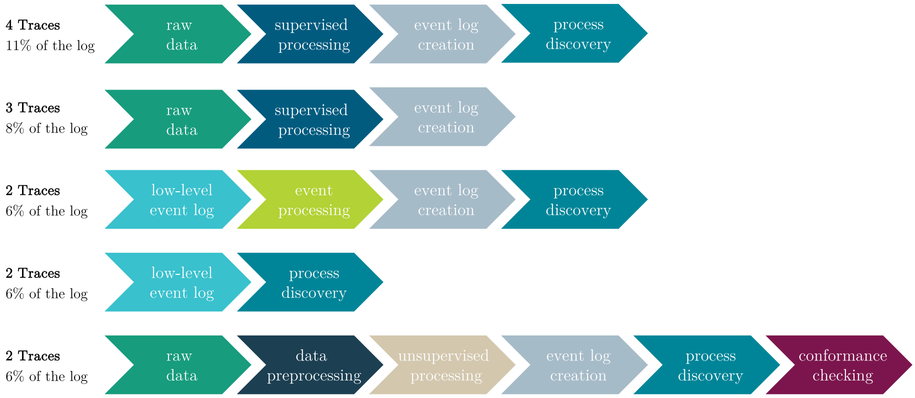


Fig. 9 The most frequently used components to analyse and use sensor data for process mining

tion techniques [43, 47]. Additionally, the use of machine state models [64], attribute-based interactions [52], and complex event processing (CEP) [50] were identified as components of analytical paths. Recent studies have further diversified the approaches by introducing various encoding rules [7, 39].

Beyond these techniques, our review highlighted the implementation of trace annotation [13, 15], event correlation [29], enhanced domain knowledge applications [5, 40], data segmentation [23, 68], and the linking of traces to raw data with event annotation [4] as examples of supervised processing within other analytic procedures. Interestingly, in some cases, supervised processing followed unsupervised processing [3, 13], with classification often being the preferred method. For example, [38] introduced labels for data segments based on predefined rules.

Supervised processing, often observed in classification scenarios, employs a range of techniques, including decision trees, random forests, artificial neural networks (including convolutional networks), and support vector machines as some of the most popular methods [3, 31, 43, 47]. In addition to these classification approaches, our analysis identified various clustering techniques for unsupervised sensor data processing. These include hierarchical clustering based on Gower distance [13], k-medoids, k-means [3, 23, 55, 62], and self-organising maps [29]. Furthermore, association rules, particularly the CAPHAR method, have been implemented as part of unsupervised techniques in studies such as [31].

Event log processing also included either supervised or unsupervised approaches. Among supervised approaches we identified "sessions of events" [37], "ontologies" [24] and "rectification of event logs based on domain knowledge" [39]. Event labelling was mainly conducted by domain experts, e.g. [14, 23, 41], based on visualisations [37, 51], or process models, e.g. [3]. Furthermore, we identified various unsupervised techniques in event log processing, mainly clustering, e.g. K-means, DBScan [37], based on quality threshold [22] or fuzzy clustering based on EM algorithm [60]. However, very specific techniques, e.g. motifs identification [20] or trace clustering [11], was also found.

Regarding the process mining tasks, we made the following discovery: the most frequently used algorithms for process discovery were the Inductive Miner (applied in 12 publications, e.g. [26]) and the Fuzzy Miner (applied in 7 publications, e.g. [45]). Other tasks, as mentioned earlier (see Sect. 3.3), were rare, only six papers used conformance checking [23, 40, 55–57, 62] and five presented process enhancement tasks [4, 7, 22, 46, 57]. Other activities in process mining on sensor data included validation [3, 4, 15, 37, 59], data preprocessing, e.g. related to data quality or missing data [13, 15] and event relabelling [59].

The diversity of data types and variations in analytic procedures hampers the formulation of generic recommendations suitable for specific domains. Nonetheless, these analytic procedures follow similar sequential steps, starting with raw data gathering, followed by supervised and unsupervised processing, data labelling, event log creation, and ending with process discovery. In the light of this, we recommend selecting the analytic procedure aligned with the domain's characteristics and the analysis objectives. Table 6 provides a comprehensive overview supporting the identification of techniques tailored to particular data types and domains. For example, an analysis within the industrial domain utilising mixed variables reveals the application of both unsupervised and supervised methods, including scenarios that bypass preprocessing. Our review streamlines the process of identifying, selecting and applying the most relevant techniques for specific analytical contexts.

Taking into consideration the results discussed above, we see the following future research directions:

Creation of more diversified, publicly available datasets The inclusion of different sensor types would increase the amount of data from sensors with continuous output, e.g. temperature, chemical, electric, based on industrial use cases, e.g. manufacturing in Industry 4.0 settings. Extending existing datasets with additional attributes would also be beneficial. This would allow studies on publicly available datasets.

Comparison of machine learning (ML) approaches when using process mining on sensor data. Analysing and comparing various (supervised and unsupervised) ML approaches could derive a default pipeline for efficient sensor data preprocessing and analysis. Furthermore, advanced ML techniques such as deep learning, reinforcement learning and large language models (LLMs) could provide new insights from sensor data through their ability to handle large volumes of complex, unstructured data.

Focus on conformance checking and enhancement with use of sensor data. Conformance checking involves comparing actual sensor-recorded behaviour with a process model to detect deviations. Future research could focus on abstracting low-level sensor data into meaningful events for more accurate conformance checking. Additionally, sensor data could be used to enhance process models, enhancing their precision and adaptability.

Extension of data usage from other sensor types. We identified that primarily environmental (mainly temperature), navigation, and position & proximity sensors (see Table 1) were used. However, we can observe that in industrial processes, process mining could also add benefits for sensor data from chemical and radiation, electric and magnetic, as well as acoustic and light types.

5 Conclusions

This paper presented a systematic literature review on process mining for sensor (event) data. Currently, there is no comprehensive overview of existing literature on the subject. Instead, publications are spread in various journals and other types of publications. The process mining applications and sensor types used in these publications differ. Hence, selecting appropriate techniques for efficient process mining for sensor event data is challenging.

Based on the extensive literature review, we devised a bibliography of existing works on the following criteria: domain of sensor implementation, sensor purpose, type of sensor, type of sensor output, abstraction of data, processing methods of sensor data, type of research artefact, type of dataset, and process mining application addressed in the paper. Our survey aids in understanding the domain better. To the best of our knowledge, our work is the first attempt to condense all related literature on the subject.

Although we made every effort to obtain a comprehensive set of publications for analysis, we are aware that our study has limitations. Firstly, we defined inclusion and exclusion criteria, which can introduce bias. For example, only publications in the English language were included. Excluding publications older than ten years might omit insights on the development and evolution of process mining on sensor data. Furthermore, we acknowledge that the identified publications could be of varying quality. We tried to control this limitation by including papers only from peer-reviewed indexed sources but did not include other quality measures. Finally, the researchers conducting the literature review, being experts in the process mining domain, approached the literature with a specific set of known keywords for abstract screening. This domain-specific perspective could bias the selection process towards specific methodologies or applications, potentially overlooking studies that employ different terminologies.

We are convinced that our paper is, on the one hand, useful for researchers who want to acquire an overview of the current state of research and design techniques to process sensor event data for process mining efficiently. On the other hand, the literature study is helpful for those who want to benchmark their techniques against existing works. Our literature overview enables a quicker finding of appropriate research artefacts.

Generally, the data collected from sensors in IoT environments are heterogeneous in format, structure and semantics, especially when different types of sensors are collecting the raw data. The data are difficult to share and reuse due to the lack of formalised descriptions of them. In future work, we plan to design a framework to capture different types of sensors' different formats and data.

Acknowledgements This project has received funding from the statutory funds of AGH University of Krakow and State of Schleswig-Holstein under the Datencampus project grant no. 220 21 016. We thank Noelle Kist and Emilia Męcfel for their valuable discussions on this work.

Funding Open Access funding enabled and organized by Projekt DEAL.

Appendix

Table 7 Expanded literature overview

Citation	Title	Domain	Event log creation	Process discovery	Conf. checking	Enhancement	Sensor output	Abstraction of data	Tools
Al-Ali et al. [3]	A composite machine-learning-based framework for supporting low-level event logs to high-level business process model activities mappings enhanced by flexible BPMN model translation	Smart home	✓	✓			Categorical	Low-level	
Prathama et al. [46]	A multi-case perspective analytical framework for discovering human daily behavior from sensors using process mining	Smart home	✓	✓		✓	Categorical	Low & high-level	ProM
Vitale et al. [62]	A Process Mining-based unsupervised Anomaly Detection technique for the Industrial Internet of Things	Industrial	✓	✓	✓		Numerical	Low-level	Pm4Py
Lofù et al. [40]	A Situation Awareness Computational Intelligent Model for Metabolic Syndromce Management	Healthcare	✓	✓	✓		Numerical	Low-level	Pm4Py
Mayr et al. [41]	Abstracting Process Mining Event Logs from Process-State Data to Monitor Control-Flow of Industrial Manufacturing Processes	Industrial	✓	✓			Numerical	Low-level	
Seiger et al. [51]	An Interactive Method for Detection of Process Activity Executions from IoT Data	Industrial	✓				Mixed	Low-level	ProM
Bayomie et al. [7]	Analyzing Manufacturing Process By Enabling Process Mining on Sensor Data	Industrial	✓	✓		✓	Mixed	Low-level	Disco
Szpyrka et al. [56]	Conformance checking of a longwall shearer operation based on low-level	Industrial			✓		Mixed	Low-level	

Table 7 continued

Citation	Title	Domain	Event log creation	Process discovery	Conf. checking	Enhancement	Sensor output	Abstraction of data	Tools
Brzychczy and Trzcionkowska (2018) [13]	Creation of an event log from a low-level machinery monitoring system for process mining purposes	Industrial	✓				Mixed	Low-level	
Nadim et al. [43]	Data-driven dynamic causality analysis of industrial systems using interpretable machine Learning and process mining	Industrial	✓	✓			Numerical	Low-level	
Su et al. [55]	Data-Driven Goal Recognition in Transhumeral Prostheses Using Process Mining Techniques	Healthcare	✓	✓	✓		Numerical	Low-level	
Cameranesi et al. [15]	Discovering process models of activities of daily living from sensors	Smart home		✓			Categorical	Low & high-level	ProM
Sztyler et al. [57]	Discovery of personal processes from labeled sensor data - An application of process mining to personalized health care	Healthcare		✓	✓	✓	Categorical	High-level	ProM, Disco
Elali et al. [24]	Domain Ontology Construction with Activity Logs and Sensors Data - Case Study of Smart Home Activities	Smart home	✓	✓			Categorical	Low-level	
van Eck et al. [23]	Enabling process mining on sensor data from smart products	Smart product	✓	✓	✓		Numerical	Low-level	
Rebmann et al. [47]	Enabling the Discovery of Manual Processes Using a Multi-modal Activity Recognition Approach	Industrial	✓	✓			Mixed	Low-level	
Tax et al. [60]	Generating time-based label refinements to discover more precise process models	Smart home	✓	✓			Categorical	High-level	ProM
Dogan et al. [22]	Individual Behavior Modeling with Sensors Using Process Mining	Smart home		✓		✓	Categorical	High-level	ProM, Disco, PALIA Suite

Table 7 continued

Citation	Title	Domain	Event log creation	Process discovery	Conf. checking	Enhancement	Sensor output	Abstraction of data	Tools
Carolis et al. [16]	Learning and recognizing routines and activities in SOFiA	Smart office	✓	✓			Categorical	Low-level	
Blank et al. [11]	Location-Aware Path Alignment in Process Mining	Industrial	✓				Categorical	High-level	
Tax et al. [59]	Log-based Evaluation of Label Splits for Process Models	Smart home	✓				Categorical	High-level	
Wolny et al. [64]	Model-driven Runtime State Identification	Industrial	✓				Numerical	Low-level	
Ziolkowski et al. [68]	Process Mining for Time Series Data	Other	✓	✓			Numerical	Low-level	
Hwang and Jang [26]	Process Mining to Discover Shoppers' Pathways at a Fashion Retail Store Using a WiFi-Base Indoor Positioning System	Industrial		✓			Categorical	High-level	ProM
Janssen et al. [29]	Process Model Discovery from Sensor Event Data	Smart home	✓	✓			Categorical	Low-level	ProM
Bano et al. [5]	Process-aware digital twin cockpit synthesis from event logs	Healthcare	✓	✓			Categorical	High-level	Pm4Py
Brzychczy and Trzcionkowska [14]	Process-oriented approach for analysis of sensor data from longwall monitoring system	Industrial	✓				Mixed	Low-level	
Khowaja et al. [31]	PROMPT: Process Mining and Paravector Tensor-Based Physical Health Monitoring Framework	Healthcare	✓	✓			n/a	Low & High level	ProM
Li et al. [39]	Rectify Sensor Data in IoT: A Case Study on Enabling Process Mining for Logistic Process in an Air Cargo Terminal	Industrial	✓	✓			Categorical	Low-level	

Table 7 continued

Citation	Title	Domain	Event log creation	Process discovery	Conf. checking	Enhancement	Sensor output	Abstraction of data	Tools
de Leoni and Pellattiero [37]	The Benefits of Sensor-Measurement Aggregation in Discovering IoT Process Models: A Smart-House Case Study	Smart home		✓			n/a	High-level	
Senderovich et al. [52]	The ROAD from sensor data to process instances via interaction mining	healthcare	✓				Categorical	Low-level	ProM
Seiger et al. [50]	Towards IoT-driven Process Event Log Generation for Conformance Checking in Smart Factories	industrial	✓				n/a	Low-level	
Porouhan [45]	Using process mining for predicting relationships of couples sitting on a sofa	Smart home		✓			Categorical	High-level	Disco
Di Federico and Burattin [20]	vAMoS: eVent Abstraction via Motifs Search	Smart home	✓				Categorical	Low-level	
Leotta et al. [38]	Visual process maps: a visualization tool for discovering habits in smart homes	Smart home	✓	✓			Categorical	Low-level	
Banham et al. [4]	xPM: A Framework for Process Mining with Exogenous Data	Healthcare	✓	✓		✓	n/a	Low & high level	ProM

References

1. van der Aalst WMP (2016) process mining—data science in action, Second Edition. Springer. https://doi.org/10.1007/978-3-662-49851-4,
2. van der Aalst WMP (2022) Process mining: a 360 degree overview. Springer International Publishing, Cham, pp 3–34. https://doi.org/10.1007/978-3-031-08848-3_1,
3. Al-Ali H, Cuzzocrea A, Damiani E, Mizouni R, Tello G (2020) A composite machine-learning-based framework for supporting low-level event logs to high-level business process model activities mappings enhanced by flexible BPMN model translation. Soft Comput 24(10):7557–7578. https://doi.org/10.1007/s00500-019-04385-6, http://link.springer.com/10.1007/s00500-019-04385-6
4. Banham A, Leemans SJJ, Wynn MT, Andrews R (2021) xpm: A framework for process mining with exogenous data. In: Munoz-Gama J, Lu X (eds) Process mining workshops—ICPM 2021 international workshops, eindhoven, the netherlands, october 31 - november 4, 2021, revised selected papers, Springer, Lecture Notes in Business Information Processing, vol 433, pp 85–97. https://doi.org/10.1007/978-3-030-98581-3_7
5. Bano D, Michael J, Rumpe B, Varga S, Weske M (2022) Process-aware digital twin cockpit synthesis from event logs. J Comput Lang 70(101):121. https://doi.org/10.1016/j.cola.2022.101121
6. Bayomie D, Helal IMA, Awad A, Ezat E, ElBastawissi A (2016) Deducing case ids for unlabeled event logs. In: Reichert M, Reijers HA (eds) Business process management workshops. Springer International Publishing, Cham, pp 242–254
7. Bayomie D, Revoredo K, Bachhofner S, Kurniawan K, Kiesling E, Mendling J (2022) Analyzing manufacturing process by enabling process mining on sensor data. In: Bork D, Barat S, Asprion PM, Marcelletti A, Morichetta A, Schneider B, Kulkarni V, Breu R, Zech P (eds) Proceedings of the poem 2022 workshops and models at work co-located with practice of enterprise modelling 2022,London, United Kingdom, 23–25 Nov 2022, CEUR-WS.org, CEUR Workshop Proceedings, vol 3298. https://ceur-ws.org/Vol-3298/paper_DTE_625.pdf
8. Bertrand Y, De Weerdt J, Serral E (2022) A bridging model for process mining and iot. In: Munoz-Gama J, Lu X (eds) Process mining workshops. Springer International Publishing, Cham, pp 98–110
9. Bertrand Y, De Weerdt J, Serral E (2023) Assessing the suitability of traditional event log standards for iot-enhanced event logs. In: Cabanillas C, Garmann-Johnsen NF, Koschmider A (eds) Business process management workshops. Springer International Publishing, Cham, pp 63–75
10. Bertrand Y, Veneruso S, Leotta F, Mecella M, Serral E (2024) Nice: the native iot-centric event log model for process mining. In: De Smedt J, Soffer P (eds) Process mining workshops. Springer Nature Switzerland, Cham, pp 32–44
11. Blank P, Maurer M, Siebenhofer M, Rogge-Solti A, Schönig S (2016) Location-aware path alignment in process mining. In: Dijkman RM, Pires LF, Rinderle-Ma S (eds) 20th IEEE international enterprise distributed object computing workshop, EDOC workshops 2016, vienna, austria, 5–9 Sept 2016, IEEE Computer Society, pp 1–8, https://doi.org/10.1109/EDOCW.2016.7584367,
12. Boulton D, Hammersley M (2006) Analysis of unstructured data. Data Collection Anal 2:243–259
13. Brzychczy E, Trzcionkowska A (2018) Creation of an event log from a low-level machinery monitoring system for process mining purposes. In: Yin H, Camacho D, Novais P, Tallón-Ballesteros AJ (eds) Intelligent data engineering and automated learning - IDEAL 2018 - 19th international conference, madrid, spain, november 21-23, 2018, proceedings, part II, Springer, Lecture Notes in Computer Science, vol 11315, pp 54–63, https://doi.org/10.1007/978-3-030-03496-2_7,
14. Brzychczy E, Trzcionkowska A (2019) Process-oriented approach for analysis of sensor data from longwall monitoring system. In: Burduk A, Chlebus E, Nowakowski T, Tubis A (eds) Intelligent systems in production engineering and maintenance. Springer International Publishing, Cham, pp 611–621
15. Cameranesi M, Diamantini C, Potena D (2017) Discovering process models of activities of daily living from sensors. In: Teniente E, Weidlich M (eds) Business process management workshops—BPM 2017 international workshops, Barcelona, Spain, 10–11 Sept 2017, revised papers, Springer, Lecture Notes in Business Information Processing, vol 308, pp 285–297, https://doi.org/10.1007/978-3-319-74030-0_21,
16. Carolis BD, Ferilli S, Mallardi G (2014) Learning and recognizing routines and activities in sofia. In: European conference on ambient intelligence, Springer, pp 191–204
17. Cook DJ, Schmitter-Edgecombe M (2009) Assessing the quality of activities in a smart environment. Methods Inf Med 48(05):480–485
18. De Luzi F, Leotta F, Marrella A, Mecella M (2024) On the interplay between business process management and internet-of-things: a systematic literature review. Bus Inf Syst Eng, pp 1–24
19. De Weerdt J, Wynn MT (2022) Foundations of process event data. In: Process mining handbook, Springer International Publishing Cham, pp 193–211

20. Di Federico G, Burattin A (2023) vamos: event abstraction via motifs search. In: Cabanillas C, Garmann-Johnsen NF, Koschmider A (eds) Business process management workshops. Springer International Publishing, Cham, pp 101–112
21. Diba K, Batoulis K, Weidlich M, Weske M (2020) Extraction, correlation, and abstraction of event data for process mining. WIREs Data Min Knowl Dis 10(3). https://doi.org/10.1002/widm.1346
22. Dogan O, Martinez-Millana A, Rojas E, Sepúlveda M, Munoz-Gama J, Traver V, Fernandez-Llatas C (2019) Individual behavior modeling with sensors using process mining. Electronics 8(7):766
23. van Eck ML, Sidorova N, van der Aalst WMP (2016) Enabling process mining on sensor data from smart products. In: Tenth IEEE international conference on research challenges in information science, RCIS 2016, grenoble, france, june 1-3, 2016, IEEE, pp 1–12, https://doi.org/10.1109/RCIS.2016.7549355,
24. Elali R, Kornyshova E, Deneckère R, Salinesi C (2022) Domain ontology construction with activity logs and sensors data–case study of smart home activities. In: International conference on advanced information systems engineering, Springer, pp 49–60
25. Haddaway NR, Grainger MJ, Gray CT (2022) Citationchaser: a tool for transparent and efficient forward and backward citation chasing in systematic searching. Res Synthesis Methods 13(4):533–545
26. Hwang I, Jang YJ (2017) Process mining to discover shoppers' pathways at a fashion retail store using a wifi-base indoor positioning system. IEEE Trans Autom Sci Eng 14(4):1786–1792. https://doi.org/10.1109/TASE.2017.2692961
27. Jamshed MA, Ali K, Abbasi QH, Imran MA, Ur-Rehman M (2022) Challenges, applications, and future of wireless sensors in internet of things: a review. IEEE Sens J 22(6):5482–5494. https://doi.org/10.1109/JSEN.2022.3148128
28. Janiesch C, Koschmider A, Mecella M, Weber B, Burattin A, Di Ciccio C, Fortino G, Gal A, Kannengiesser U, Leotta F, Mannhardt F, Marrella A, Mendling J, Oberweis A, Reichert M, Rinderle-Ma S, Serral E, Song W, Su J, Torres V, Weidlich M, Weske M, Zhang L (2020) The internet of things meets business process management: a manifesto. IEEE Syst Man Cybernet Mag 6(4):34–44. https://doi.org/10.1109/MSMC.2020.3003135
29. Janssen D, Mannhardt F, Koschmider A, van Zelst SJ (2020) Process model discovery from sensor event data. In: Leemans SJJ, Leopold H (eds) Process mining workshops—ICPM 2020 international workshops, padua, italy, october 5-8, 2020, revised selected papers, Springer, Lecture Notes in Business Information Processing, vol 406, pp 69–81, https://doi.org/10.1007/978-3-030-72693-5_6,
30. Javaid M, Haleem A, Singh RP, Rab S, Suman R (2021) Significance of sensors for industry 4.0: Roles, capabilities, and applications. Sens Int 2:100–110. https://doi.org/10.1016/j.sintl.2021.100110
31. Khowaja SA, Khuwaja P, Dev K, Jarwar MA (2023) Prompt: process mining and paravector tensor-based physical health monitoring framework. IEEE Sens J 23(2):989–996. https://doi.org/10.1109/JSEN.2022.3195613
32. Koschmider A, Mannhardt F, Heuser T (2019) On the contextualization of event-activity mappings. In: Daniel F, Sheng QZ, Motahari H (eds) Business process management workshops. Springer International Publishing, Cham, pp 445–457
33. Koschmider A, Janssen D, Mannhardt F (2020) Framework for process discovery from sensor data. EMISA Workshop. https://ceur-ws.org/Vol-2628/paper5.pdf
34. Koschmider A, Oppelt N, Hundsdörfer M (2022) Confidence-driven communication of process mining on time series. Inform Spektrum 45(4):223–228. https://doi.org/10.1007/s00287-022-01470-3
35. Koschmider A, Aleknonyte-Resch M, Fonger F, Imenkamp C, Lepsien A, Apaydin K, Harms M, Janssen D, Langhammer D, Ziolkowski T, et al. (2023) Process mining for unstructured data: challenges and research directions. arXiv preprint arXiv:2401.13677
36. Laue R, Koschmider A, Fahland D (2020) Prozessmanagement und Process-Mining - Grundlagen. De Gruyter Studium, De Gruyter,. https://doi.org/10.1515/9783110500165
37. de Leoni M, Pellattiero L (2021) The benefits of sensor-measurement aggregation in discovering iot process models: A smart-house case study. In: Marrella A, Weber B (eds) Business process management workshops - BPM 2021 international workshops, Rome, Italy, 6–10 Sept 2021. revised selected papers, Springer, Lecture Notes in Business Information Processing, vol 436, pp 403–415, https://doi.org/10.1007/978-3-030-94343-1_31,
38. Leotta F, Mecella M, Sora D (2020) Visual process maps: a visualization tool for discovering habits in smart homes. J Ambient Intell Human Comput 11(5):1997–2025. https://doi.org/10.1007/s12652-019-01211-7
39. Li CY, Joshi A, Tam NTL, Lau SSF, Huang J, Shinde T, van der Aalst WMP (2024) Rectify sensor data in iot: A case study on enabling process mining for logistic process in an air cargo terminal. In: Sellami M, Vidal ME, van Dongen B, Gaaloul W, Panetto H (eds) Cooperative information systems. Springer Nature Switzerland, Cham, pp 293–310

40. Lofù D, Pazienza A, Ardito C, Di Noia T, Di Sciascio E, Vitulano F (2022) A situation awareness computational intelligent model for metabolic syndrome management. In: 2022 IEEE conference on cognitive and computational aspects of situation management (cogsima), pp 118–124, https://doi.org/10.1109/CogSIMA54611.2022.9830673
41. Mayr M, Luftensteiner S, Chasparis GC (2021) Abstracting process mining event logs from process-state data to monitor control-flow of industrial manufacturing processes. In: Longo F, Affenzeller M, Padovano A (eds) Proceedings of the 3rd international conference on industry 4.0 and smart manufacturing (ISM 2022), virtual event / upper austria university of applied sciences - hagenberg campus - linz, austria, 17-19 november 2021, Elsevier, Procedia Computer Science, vol 200, pp 1442–1450, https://doi.org/10.1016/j.procs.2022.01.345,
42. Mukhopadhyay SC, Tyagi SKS, Suryadevara NK, Piuri V, Scotti F, Zeadally S (2021) Artificial intelligence-based sensors for next generation iot applications: a review. IEEE Sens J 21(22):24920–24932. https://doi.org/10.1109/JSEN.2021.3055618
43. Nadim K, Ragab A, Ouali M (2023) Data-driven dynamic causality analysis of industrial systems using interpretable machine learning and process mining. J Intell Manuf 34(1):57–83. https://doi.org/10.1007/s10845-021-01903-y
44. Ni Q, García Hernando AB, De la Cruz IP (2015) The elderlyâ€™s independent living in smart homes: A characterization of activities and sensing infrastructure survey to facilitate services development. Sensors 15(5):11312–11362. https://doi.org/10.3390/s150511312, https://www.mdpi.com/1424-8220/15/5/11312
45. Porouhan P (2019) Using process mining for predicting relationships of couples sitting on a sofa. In: 2019 17th international conference on ict and knowledge engineering (ictke), pp 1–9, https://doi.org/10.1109/ICTKE47035.2019.8966893
46. Prathama F, Nugroho Yahya B, Lee SL (2021) A Multi-case Perspective Analytical Framework for Discovering Human Daily Behavior from Sensors using Process Mining. In: 2021 IEEE 45th annual computers, software, and applications conference (COMPSAC), IEEE, Madrid, Spain, pp 638–644, https://doi.org/10.1109/COMPSAC51774.2021.00093, https://ieeexplore.ieee.org/document/9529405/
47. Rebmann A, Emrich A, Fettke P (2019) Enabling the discovery of manual processes using a multi-modal activity recognition approach. In: Francescomarino CD, Dijkman RM, Zdun U (eds) Business process management workshops—BPM 2019 international workshops, Vienna, Austria, 1–6 Sept 2019, revised selected papers, Springer, Lecture Notes in Business Information Processing, vol 362, pp 130–141, https://doi.org/10.1007/978-3-030-37453-2_12,
48. Ripka P, Arafat M (2019) Magnetic sensors: Principles and applications. In: Reference module in materials science and materials engineering, Elsevier. https://doi.org/10.1016/B978-0-12-803581-8.11680-7. https://www.sciencedirect.com/science/article/pii/B9780128035818116807
49. Sehrawat D, Gill NS (2019) Smart sensors: Analysis of different types of iot sensors. In: 2019 3rd international conference on trends in electronics and informatics (icoei), pp 523–528, https://doi.org/10.1109/ICOEI.2019.8862778
50. Seiger R, Zerbato F, Burattin A, García-Bañuelos 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. https://doi.org/10.1109/EDOCW49879.2020.00016
51. Seiger R, Franceschetti M, Weber B (2023) An interactive method for detection of process activity executions from iot data. Future Internet 15(2). https://doi.org/10.3390/fi15020077, https://www.mdpi.com/1999-5903/15/2/77
52. Senderovich A, Rogge-Solti A, Gal A, Mendling J, Mandelbaum A (2016) The road from sensor data to process instances via interaction mining. In: International conference on advanced information systems engineering, Springer, pp 257–273
53. de Silva C (2017) Sensor systems: fundamentals and applications. CRC Press. https://doi.org/10.1201/9781315371160
54. Soffer P, Hinze A, Koschmider A, Ziekow H, Di Ciccio C, Koldehofe B, Kopp O, Jacobsen A, Sürmeli J, Song W (2019) From event streams to process models and back: Challenges and opportunities. Inf Syst 81:181–200. https://doi.org/10.1016/j.is.2017.11.002. https://www.sciencedirect.com/science/article/pii/S0306437917300145
55. Su Z, Yu T, Lipovetzky N, Mohammadi A, Oetomo D, Polyvyanyy A, Sardiña S, Tan Y, van Beest N (2023) Data-driven goal recognition in transhumeral prostheses using process mining techniques. In: 5th International conference on process mining, ICPM 2023, Rome, Italy, 23–27 Oct 2023, IEEE, pp 25–32. https://doi.org/10.1109/ICPM60904.2023.10271945,

56. Szpyrka M, Brzychczy E, Napieraj A, Korski J, Nalepa G (2020) Conformance checking of a longwall shearer operation based on low-level events. Energies 13(24). https://doi.org/10.3390/en13246630, https://www.mdpi.com/1996-1073/13/24/6630
57. Sztyler T, Völker J, Carmona J, Meier O, Stuckenschmidt H (2015) Discovery of personal processes from labeled sensor data - an application of process mining to personalized health care. In: van der Aalst WMP, Bergenthum R, Carmona J (eds) Proceedings of the international workshop on algorithms and theories for the analysis of event data, ATAED 2015, satellite event of the conferences: 36th international conference on application and theory of petri nets and concurrency petri nets 2015 and 15th international conference on application of concurrency to system design ACSD 2015, Brussels, Belgium, 22–23 June 2015, CEUR-WS.org, CEUR Workshop Proceedings, vol 1371, pp 31–46. http://ceur-ws.org/Vol-1371/paper03.pdf
58. Sztyler T, Carmona J, Völker J, Stuckenschmidt H (2016) Self-tracking reloaded: applying process mining to personalized health care from labeled sensor data. Transactions on Petri nets and other models of concurrency XI, pp 160–180
59. Tax N, Sidorova N, Haakma R, van der Aalst WMP (2016) Log-based evaluation of label splits for process models. In: Howlett RJ, Jain LC, Gabrys B, Toro C, Lim CP (eds) Knowledge-based and intelligent information & engineering systems: Proceedings of the 20th international conference kes-2016, York, UK, 5–7 Sept 2016, Elsevier, Procedia Computer Science, vol 96, pp 63–72, https://doi.org/10.1016/j.procs.2016.08.096,
60. Tax N, Alasgarov E, Sidorova N, Haakma R, van der Aalst WM (2019) Generating time-based label refinements to discover more precise process models. J Ambient Intell Smart Environ 11:165–182. https://doi.org/10.3233/AIS-190519
61. Valderas P, Torres V, Serral E (2022) Towards an interdisciplinary development of iot-enhanced business processes. Bus Inf Syst Eng. https://doi.org/10.1007/s12599-022-00770-y
62. Vitale F, De Vita F, Mazzocca N, Bruneo D (2023) A process mining-based unsupervised anomaly detection technique for the industrial internet of things. Internet Things 24(100):993. https://doi.org/10.1016/j.iot.2023.100993. https://www.sciencedirect.com/science/article/pii/S2542660523003165
63. Weerdt JD, Wynn MT (2022) Foundations of process event data. In: van der Aalst WMP, Carmona J (eds) Process mining handbook, Lecture Notes in Business Information Processing, vol 448, Springer, pp 193–211. https://doi.org/10.1007/978-3-031-08848-3_6,
64. Wolny S, Mazak A, Wimmer M, Huemer C (2019) Model-driven runtime state identification. In: Emisa forum: vol. 39, no. 1, De Gruyter
65. Xiao Y, Watson M (2019) Guidance on conducting a systematic literature review. J Plann Educ Res 39(1):93–112. https://doi.org/10.1177/0739456X17723971
66. van Zelst SJ, Mannhardt F, de Leoni M, Koschmider A (2021) Event abstraction in process mining: literature review and taxonomy. Granular Comput 6(3):719–736. https://doi.org/10.1007/s41066-020-00226-2
67. Zerbino P, Stefanini A, Aloini D (2021) Process science in action: A literature review on process mining in business management. Technological Forecasting and Social Change 172(121):021. https://doi.org/10.1016/j.techfore.2021.121021. https://www.sciencedirect.com/science/article/pii/S0040162521004534
68. Ziolkowski T, Koschmider A, Schubert R, Renz M (2022) Process mining for time series data. BPMDS 2022/EMMSAD 2022 450:347–350

Edyta Brzychczy graduated PhD and DSc with distinction in Management and Marketing in Industry at the Mining Faculty of AGH University of Science and Technology, Krakow, Poland. Currently, she is a university professor at the Faculty of Mechanical Engineering and Robotics at her alma mater and Head of the Data & Process Mining Group. Her main research interests are related to the modelling and optimisation of industrial processes. She authored and co-authored over 100 publications, mainly in the field of modelling and optimisation of mining production. She also coordinated projects related to data mining and analytics in the mining domain. Currently, she is interested in process re-engineering based on sensor data and process mining in heavy industries.

Milda Aleknonytė-Resch is a postdoctoral researcher in business informatics at Kiel University. She completed her PhD in natural sciences in 2021. Milda's research focuses on developing data-driven methods with a strong focus on increasing their explainability. She emphasizes interdisciplinary research, with interests spanning process mining for unstructured data, particularly in the medical and healthcare domains, as well as time-series analysis, interaction effects, and simulation studies.

Dominik Janssen is pursuing a PhD in business informatics. His research focuses on developing and implementing a framework that combines various machine learning techniques to discover process models from low-level IoT sensor data. In previous projects, he worked on transforming sensor data from smart homes and wearables into process models representing daily activities. These insights are now being further developed in collaboration with industry partners to optimize production processes and implement predictive maintenance solutions.

Agnes Koschmider is a professor of business informatics at the University of Bayreuth. She completed her PhD in 2007 and her habilitation in applied informatics in 2015 at KIT. Agnes researches methods for data-driven analysis and explanation of processes (process mining), based on artificial intelligence. At the centre of her research is process analytics: developing a pipeline to efficiently process the complete chain from raw data to process discovery.