



Guidelines for manual process discovery for semi-automated processes

Linda Maria Kölbel¹ · Markus Hornsteiner¹ · Stefan Schöning¹

Received: 1 August 2024 / Revised: 7 April 2026 / Accepted: 8 June 2026
© The Author(s) 2026

Abstract

Semi-automated operational environments, where human operators and automated systems coexist, pose significant challenges for process discovery approaches, which typically target either fully manual or fully automated workflows. This paper presents a comprehensive guideline for manual process discovery specifically tailored to semi-automated contexts. Building on established methods (document analysis, direct observation, structured interviews, and collaborative workshops), we adapt and extend each method to capture both human-driven activities and machine-generated events. The resulting artifact comprises four modular guideline packages, offering step-by-step procedures, guiding questions, and optional tooling to uncover control flow, data flow, resource assignments, and system interactions. Use cases in an Industrial Internet of Things (IIoT) setting illustrate how these guidelines can be applied to complex, data-intensive scenarios, though the approach remains domain-agnostic and can be transferred to other semi-automated environments. The development followed the Design Science Research (DSR) cycle, incorporating iterative refinement and validation via expert workshops, surveys, and real-world use cases. Our evaluation confirms that the proposed guidelines enhance the completeness, accuracy, and efficiency of manual process discovery in hybrid human-machine landscapes.

Keywords Business process management · Process discovery · Industrial internet of things · Guideline development · Semi-automated processes

✉ Linda Maria Kölbel
linda.koelbel@ur.de

¹ University of Regensburg Faculty of Informatics and Data Science: Universität, Regensburg Fakultät für Informatik und Data Science, Regensburg, Germany

1 Introduction

Business Process Management (BPM) constitutes a well-established framework for the identification, analysis, optimization, and monitoring of business processes (Dumas et al. 2021). An essential initial phase in any BPM initiative is process discovery, which involves the systematic detection and documentation of processes as they are executed in practice. The acquired process information facilitates the creation of business process models, which serve as the foundation for successful process management (Dumas et al. 2021). Process discovery methodologies are generally categorized into two groups: manual process discovery and process mining. In fully automated environments where the availability of event log data permits the application of process mining, various techniques can be employed to discover process models (Ackermann et al. 2024; Dumas et al. 2021). As shown by dos Santos et al. (2019), numerous process mining techniques exist for specific domains and diverse use cases. However, many real-world processes cannot be identified using process mining alone because event data is often unavailable, unstructured, or incomplete (van der Aalst 2022; Koschmider et al. 2023). Consequently, practitioners report a lack of methodological guidance for manual process discovery in scenarios where automated discovery via process mining is not applicable (Martin et al. 2021). Descriptions of manual discovery methods are often inadequate and lack direct applicability. Specifically, there are no explicit guidelines for conducting document analysis, observations, interviews, or workshops. This shortcoming hinders the successful discovery of process information when process mining techniques cannot be applied. This challenge persists for the discovery of both manual and semi-automated processes.

Numerous real-world workflows, particularly in operational environments, exhibit a semi-automated character; they encompass both machine-driven activities and manual tasks that frequently remain undocumented (Langer and Söffker 2011). We define a semi-automated process as a workflow in which a significant portion of activities is executed automatically by technology (e.g., machines, IIoT devices, or software agents, such as in Robotic Process Automation) yet remains closely integrated with activities performed by human actors (Stoiber and Schöning 2024). These processes are not restricted to manufacturing; they are prevalent in diverse domains such as logistics (Tsai et al. 2025) and healthcare (Wang and Liu 2025). The underlying technological infrastructure is often based on Operational Technology (OT) systems that control machines and physical processes. In contrast to standard IT systems, OT systems are used for machine automation and are designed for long-term operation (Conklin 2016). Although our guidelines focus on the complexities of the Industrial Internet of Things (IIoT) for demonstration and validation, they are designed to be transferable to other semi-automated contexts. Given these conditions, manual process discovery remains indispensable despite the availability of process mining tools because:

- Human activities are often undocumented, not tracked, and hence not apparent from data.
- System logs from legacy OT systems may be incomplete or inaccessible, result-

ing in "black-box" processes.

- Manual methods are required to capture the context, intent, and decision logic underlying human interventions and interactions with automated systems.

Despite the critical importance of process discovery within BPM, practical and detailed guidelines for conducting manual process discovery methods remain rare (Zhang and Zhang 2024). This lack of guidance extends not only to fully manual processes but also to semi-automated contexts. Consequently, practitioners frequently lack concrete instructions on how to execute manual discovery effectively, and specifically, how to capture machine roles, sensor data, data connectivity, and human-machine interactions. These specific details are often missing in manual process discovery methods, particularly when both technical systems and human activities must be discovered and aligned (Rehse et al. 2024). The IIoT serves as a prime example of the relevance of this problem: the presence of smart devices, combined with the often undocumented nature of manual interventions, presents a significant challenge. The IIoT interconnects people and machines to encourage more rapid and integrated industrial processes (Boyes et al. 2018), yet it introduces significant challenges for process discovery: legacy equipment and smart devices coexist (Raptis et al. 2019), manual interventions remain undocumented (Wang et al. 2018), and real-time responsiveness is frequently required (Szelągowski et al. 2022). For instance, an automated production line may generate detailed system logs from robotic stations, whereas human quality inspections are neither digitized nor formally recorded. Such conditions render the IIoT an instructive and demanding context for assessing manual process discovery in complex environments. Therefore, the IIoT serves as the chosen application domain for both the demonstration and evaluation of the proposed guidelines.

To address the lack of well-described and reproducible manual process discovery methods, we present a comprehensive set of guidelines detailing the practical execution of these methods. The primary contribution of this work is the systematic extension and operationalization of established manual process discovery techniques. While specifically tailored for hybrid human-machine environments, they remain applicable to fully manual processes. We detail four complementary guidelines, based on document analysis, observation, interviews, and workshops, that include explicit procedural steps, guiding questions, and optional tools. These are designed to capture control flow, data flow, and system interactions, explicitly incorporating IIoT-specific elements as practical examples of semi-automated processes. The intended outcome of the guidelines is the explicit step-by-step guidance of necessary process information, which serves as the basis to develop a process model in a following step. The formal conversion of this discovered information into graphical or executable process models falls outside the scope of this work.

The development of the guidelines was guided by the Design Science Research (DSR) methodology (Hevner et al. 2004) and involved iterative refinement through multiple evaluation rounds with both academic and industry stakeholders. This multiperspective validation strategy ensures that the guidelines are both practically applicable and methodologically sound.

The structure of this paper is as follows: In Sect. 2, the theoretical background and context of the study are discussed. Section 3 presents the research methodology and the development of the artifact. In Sect. 4, related work is reviewed to situate the contribution within the existing body of knowledge. Section 5 introduces the guidelines themselves, their application, and a combined example of the different methods. In Sect. 6, the guidelines and their effectiveness are evaluated. Finally, in Sect. 7, limitations are summarized and potential directions for future research are outlined.

2 Background

This section provide the theoretical groundwork for the study. It commences with the fundamental principles of BPM and the specific phase of process discovery. The section subsequently introduces the IIoT as an emerging domain characterized by semi-automated processes, emphasizing its relevance to BPM. Finally, it delineates the key process perspectives that form a multidimensional framework for the comprehensive discovery and modeling of complex organizational processes.

2.1 Business process management and process discovery

BPM comprises various methodologies and measures intended to enhance the efficiency and effectiveness of processes (Hansen et al. 2019). It functions as a decision-making instrument for process improvement and supports organizational management (Weske 2012). Specifically, BPM aims to reduce throughput times, augment efficiency, curtail costs, and minimize error rates, which contributes to improved competitiveness (Dumas et al. 2021; Bernardo et al. 2017). While BPM is frequently viewed as a strategy for securing competitive advantage, various definitions exist within the literature (Zur Muehlen et al. 2005). This study adopts the definition by Dumas et al. (2021):

A body of methods, techniques, and tools to identify, discover, analyze, redesign, execute, and monitor business processes to optimize their performance. The emphasis of this study centers on the process discovery phase, explicitly addressing the underlying concepts, methods, and techniques. A comprehensive understanding of the process is a prerequisite to facilitate its further analysis, discussion, and optimization (Jakobs and Spanke 2011). Accordingly, process information is discovered using a variety of methods to capture and document the actual state of situations. Established process discovery methods include document analysis, observation, interviews, and workshops (Dumas et al. 2021; Hansen et al. 2019; Gronau 2017). The gathered data, whether written, visual, or verbal, is then formalized into a process model in a subsequent step (Dumas et al. 2021; Hansen et al. 2019). The objective of process discovery is to aggregate all necessary information for modeling the relevant processes. This encompasses organizational structures, workflows, procedural instructions, assigned responsibilities, used documents, and technical infrastructure (e.g., warehouse layouts, conveyors, production machinery), as well as data and system-related aspects (Dumas et al. 2021; Becker et al. 2012).

To ensure clarity and traceability, modeling is typically executed post-discovery using well-established modeling languages such as Event-driven Process Chains (EPC), Unified Modeling Language (UML), flowcharts, or Business Process Modeling Notation (BPMN) (Dumas et al. 2021). Modeling facilitates process analysis and optimization while serving as a knowledge base for training (Jakobs and Spanke 2011). The outcome of this phase is an "as-is" process model that accurately represents the current status of the process (Dumas et al. 2021). The challenges of process identification and discovery stem from the organization of processes based on the division of labor, the necessity for consistent terminology, and the inherent complexity of the workflows (Hansen et al. 2019). Frequently, multiple iterations are required to fully discover and model a process, as misunderstandings, generalizations, errors, or ambiguities often arise during the data discovery phase (Hansen et al. 2019; Jakobs and Spanke 2011).

2.2 Industrial IoT

The IIoT is defined as a specialized subset of the Internet of Things (IoT) tailored specifically to industrial use cases. While both paradigms consist of networked devices exchanging data, the IIoT is engineered for large-scale, machine-centric environments, such as manufacturing and logistics, where stringent requirements for reliability, security, and real-time performance must be met (Gilchrist 2016; Sisinni et al. 2018). In contrast, the broader IoT landscape encompasses consumer, commercial, and industrial verticals, each characterized by distinct user groups and technical constraints. Consumer-oriented IoT applications, such as smart homes and wearable devices, are primarily designed to enhance user convenience (Gilchrist 2016). By comparison, the IIoT emphasizes the improvement of operational efficiency, automation, and responsiveness through machine-to-machine communication, supporting novel use cases such as predictive maintenance and robotics (Sisinni et al. 2018). IIoT processes are considered a type of semi-automated process.

In the context of BPM, integration with the IIoT is of particular significance. IIoT systems continuously generate large volumes of real-time data, which can be directly incorporated into process models to enable dynamic execution, real-time monitoring, automated decision-making, and data-driven process discovery. This integration transforms BPM methods by enhancing their adaptability and responsiveness to changing environmental and operational conditions. Conversely, IIoT-relevant components within industrial processes must remain manageable and governable through BPM methodologies to ensure reliability, traceability, and compliance. Consequently, there is a growing need for BPM frameworks specifically tailored to the requirements of IIoT environments, capable of discovering and orchestrating complex, data-intensive, and event-driven processes.

2.3 Perspectives

For the complete and accurate discovery and modeling of processes, diverse perspectives must be considered (Jablonski and Götz 2007; Tiftik et al. 2022; van der Aalst 2016). These perspectives provide various aspects that contribute to a comprehensive

understanding of the process. Due to their inherent interdependencies, these perspectives are interconnected and cannot be analyzed in isolation. Consequently, a robust process model should incorporate multiple perspectives. Five relevant perspectives have been identified, each contributing to a holistic view of the process (Jablonski and Götz 2007).

Data perspective includes the input and output objects of a process, such as documents, files, or artifacts, as well as their provisioning time. This perspective illustrates how a task is performed and establishes a link between the process and external data models (Jablonski and Götz 2007; Dumas et al. 2021). Furthermore, it allows for the analysis of data effects on the process (Tiftik et al. 2022). In the context of processes, particularly those within the IIoT, it is crucial to record the connectivity of machines, measurement sensors, and data processing systems, along with their generated data and forwarding paths. Consequently, network traffic data is considered an integral part of the data perspective (Hornsteiner et al. 2024).

Functional perspective defines a process step or subprocess, addressing the question of what needs to be done. It serves as the fundamental unit of process execution and specifies which activity or task is performed, whether by a human agent or an automated system (Jablonski and Bussler 1996; Awadid 2017).

Operational perspective encompasses the tools, programs, applications, or services used within a process step (Schönig et al. 2012; Awadid 2017; Jablonski and Götz 2007). It defines the specific manner in which a process step is executed technologically (Jablonski and Bussler 1996).

Resource perspective facilitates the examination of process execution in relation to the organizational structure. It identifies where activities are performed and who is responsible for their execution (Awadid 2017). This perspective assigns responsibilities for tasks to both human actors and technical systems (Jablonski and Bussler 1996).

Control flow perspective focuses on the logical sequence of process steps, independent of absolute time (Jablonski and Götz 2007; Dumas et al. 2021). It specifies the order in which activities, functional units, and events occur, therefore illustrating the workflow and demonstrating the causal dependencies between all modeling elements (Jablonski and Götz 2007; Awadid 2017). Compliance to control flow specifications is mandatory during process execution. These flows are typically modeled using constructs such as sequences, parallel branching, conditional branching, and merging (Jablonski and Bussler 1996).

3 Research method and artifact development

In this study, the DSR framework by Hevner et al. (2004) is used to develop an artifact that addresses a specific problem within the domain of information systems. The framework involves seven steps, which are outlined below; their specific application in the context of this research is subsequently detailed and depicted graphically in Fig. 1. While the sequence of implementation is not strictly prescribed, this framework ensures a rigorous and traceable research process.

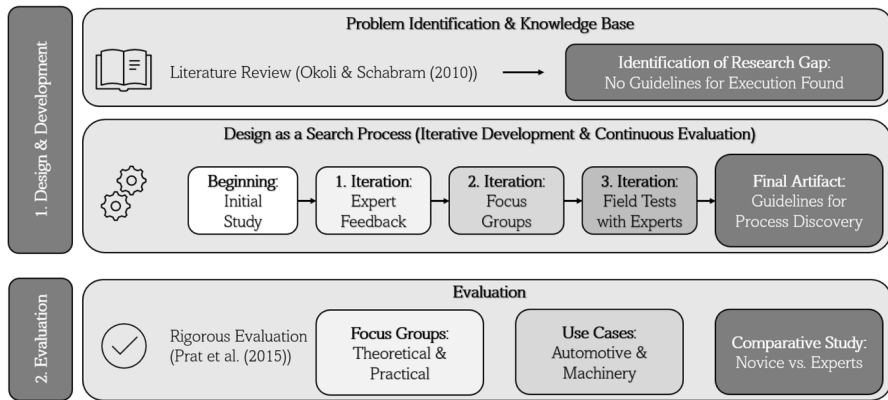


Fig. 1 Structure of the DSR phases conducted in this work

3.1 Design principles

The following paragraphs explain how the seven steps of the DSR are implemented in this study. These steps ensure that the development of our guidelines for process discovery maintains a high degree of scientific rigor while also providing added value to industrial practitioners.

Design as an artifact The research process must produce an artifact, such as a model or a method. The artifact developed in this work comprises a set of process discovery guidelines. These guidelines are designed to structure the execution of established manual discovery methods: document analysis, observation, interviews, and workshops. The objective is to provide a systematic approach to discover process information, which serves as the foundation for modeling a process model in a subsequent or parallel step. While the guidelines are applicable to general process discovery, they are explicitly designed to accommodate semi-automated processes as well. This ensures that the resulting process information captures the details required for advanced modeling applications, such as the utilization of BPMN extensions for the IIoT.

Problem relevance The problem must be relevant to both academia and industry. Given the increasing importance of digitalization and automation in industry, discovering semi-automated processes is becoming increasingly complex. Industrial environments often feature hybrid processes that integrate manual human work with automated machine steps; these often cannot be discovered by process mining techniques due to missing log data or human tasks that generate no logs. The relevance of this problem was empirically confirmed by our *initial study*, a survey of 28 participants from academia and industry. The questionnaire was developed based on the evaluation dimensions proposed by Prat et al. (2015) and designed in accordance with established usability assessment approaches (Laugwitz et al. 2008). The participant group encompassed a wide range of experience levels, with over 70% already experienced in the field of process discovery activities. The results highlighted a critical gap: 85% of participants agreed that guidelines are necessary to ensure completeness and avoid errors, while 96% stated that they would actively use such a structured

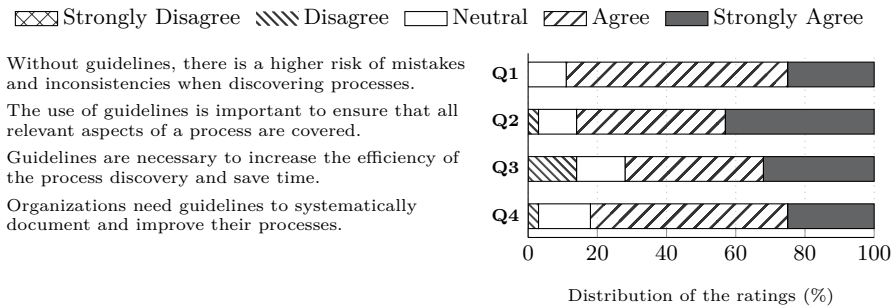


Fig. 2 Questions regarding the necessity of guidelines and distribution of the ratings of the responses

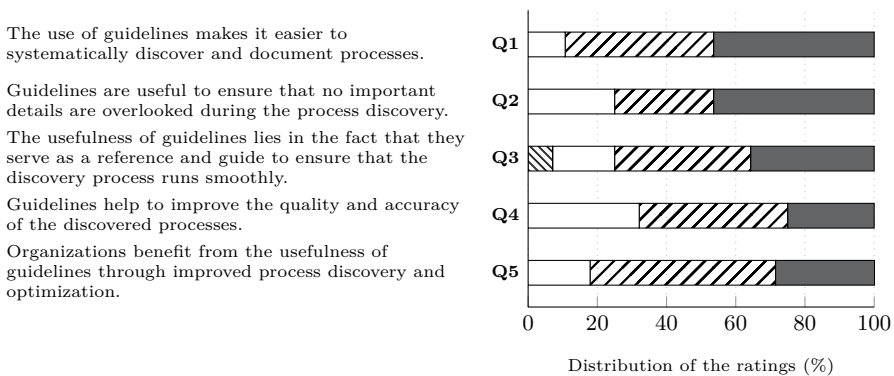


Fig. 3 Questions regarding the usefulness of the guidelines and the distribution of ratings for the answers

guideline. This high level of agreement across all experience levels confirmed that existing methods lack the necessary structure for efficient, error-free documentation in industrial contexts. Consequently, these findings formed the foundation and motivation for our artifact and justified the following rounds of iterative development described in Sect. 3.2. A detailed breakdown of the survey results is presented in Figs. 2, 3, 4 and 5.

Design evaluation The artifact must be rigorously evaluated to demonstrate its utility, quality, and efficacy. While the iterations of the evaluation are detailed in Sect. 3.2, the evaluation itself is described in Sect. 6.

Research contribution The artifact must provide a clear and novel contribution to the state of the art. To determine the existing knowledge base, an SLR was conducted (see Sect. 4). Initially, our objective was to build upon existing guidelines for the execution of manual process discovery. However, neither structured nor unstructured searches yielded publications that described the execution of these methods in detail. While the discovery methods themselves are cited in the literature, no specific guidelines for their application were found. Therefore, the guidelines developed in this paper represent a novel contribution by closing the gap between high-level method

The use of guidelines will lead to a higher quality of the discovered processes.

Guidelines increase the efficiency of process discovery and reduce the time required.

Organizations can expect better transparency and understanding of their processes through the use of guidelines.

The introduction of guidelines for manual process discovery will help to improve the consistency and comparability of the discovered processes.

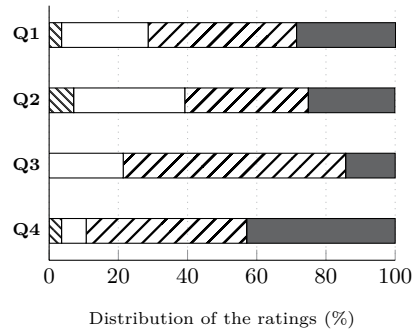


Fig. 4 Questions regarding the expected results when applying the guidelines and the distribution of the ratings of the responses

The provision of guidelines helps employees to learn process discovery more quickly and carry it out more quickly and carry it out more effectively.

The introduction of guidelines can help to ensure a uniform approach to process discovery in the organization.

The leadership team should actively support the use of guidelines for manual process discovery and promote their implementation.

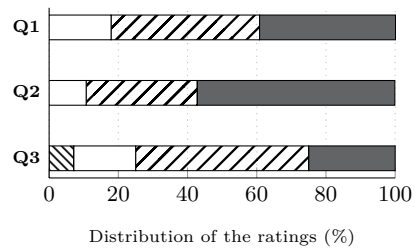


Fig. 5 Questions about the support potential of guidelines and distribution of response ratings

selection and the detailed, practical execution required for modeling both semi-automated industrial processes and manual processes.

Research rigor The artifact must be developed using rigorously defined and reproducible methods. This study started with a literature review conducted according to the guidelines of Okoli and Schabram (2010). Following the identification of the research gap, the artifact was developed through a systematic build-and-evaluate cycle. The robustness of this process is demonstrated by the continuous incorporation of feedback. This responsiveness to draft iterations ensures that the artifact is grounded in both theory and practice.

Design as a search process The development of an artifact is an iterative process characterized by the search for an optimal solution through the generation and testing of design proposals. Because the design and evaluation phases were combined, a detailed description of each step is provided in Sect. 3.2 and summarized in Table 1.

Communication of research The findings of this study are communicated through this publication, which details the methodology, the resulting guidelines, and the evaluation evidence, making the artifact accessible for adoption by process analysts and researchers.

Table 1 Overview of evaluation activities, insights, and improvements aligned with Prat et al. (2015)

Evaluation Phase	Method	Focus Areas	Learnings	Improvements Made
Round 1: Expert Feedback	Expert interviews and follow-up survey	Completeness, clarity, usefulness, necessity, modularity	Generalizability beyond IIoT, need for examples, preference for flexible toolkit structure	Example questions added; structure revised to allow modular use
First Use Case	Application in IIoT project at automotive supplier	Practical applicability of methods; real-world relevance	Observation helpful for control flow; interview better for tools; method mix needed; practical method guidance missing	Added method-specific suggestions; emphasized mixed-method flexibility
Focus Group 1 (Theoretical)	Discussion with researchers with process discovery background	Interpretation, target groups, guideline applicability	Guidelines must work for novices and experts; include support materials (e.g., checklists); discuss integration of automated techniques	Added support tools; opened future extension path (e.g., process mining); clarified audience focus
Focus Group 2 (Practical)	Industry experts with hands-on experience	Practical structure, completeness, real-world usability	Need for compact structure; standard procedure lacking in SMEs; role of psychological factors in interviews/workshops	Added guidance on scope, process context, psychological aspects; promoted top-down approach
Round 2: Expert Field Test	Modeling of IIoT line by experienced analyst	Guideline structure, real-world usability, process logic	process perspective structure not practical	Reorganized structure into intuitive process phases (e.g., Workshop Start, Group Work, etc.)
Second Use Case	Application in new IIoT context	Practical challenges in multi-method discovery, black box effects	Observation timing crucial; live modeling > survey validation; black-box processes challenge completeness	Added recommendations for kickoff; stakeholder involvement, method sequencing and validation techniques
Round 3: Novice vs. Expert	Novices and expert discover same process; validated by company experts	Usability across skill levels, result comparability, output quality	Novices achieved similar structure and completeness; expert added detail; wording differed due to vocabulary	Validated applicability for different user levels; confirmed expert knowledge adds value, not dependency

3.2 Iterative development process

To develop the artifact, we conducted a comprehensive, multi-stage development process, evaluating the results of our artifact in parallel with its development after each step using the evaluation dimensions proposed by Prat et al. (2015). The evaluation process was divided into multiple rounds to assess the artifact from diverse perspectives. The individual steps of the artifact's development and their outputs are shown in Table 1. In the following every step is described in more detail. The final evaluation of the artifact is discussed in Sect. 6.

Iteration 1 – Expert feedback

In the first artifact evaluation round, we conducted expert interviews and a follow-up survey. The focus was on completeness, clarity, necessity, and usability. Experts emphasized the potential generalizability of the approach beyond IIoT, an encouraging signal, though we maintained our focus on this domain for now. The need for example questions and flexible, modular use was raised. Based on this, we integrated illustrative prompts and reframed the guideline as a toolkit rather than a rigid step-by-step manual.

First use case

Subsequently, the guidelines were applied in a real-world IIoT scenario at an automotive supplier (Sect. 6.1). While the guidelines supported the discovery task, uncertainties remained regarding the selection of appropriate discovery methods. This highlighted the need for a more accurate description of method application. Furthermore, previously neglected observation proved valuable, particularly for identifying the control flow. A key finding was that combining observation with interviews increased completeness: observation better captured the control flow, while interviews most effectively clarified tool usage. Consequently, these findings led to the integration of method-specific recommendations into the guideline.

Iteration 2 – Focus group feedback

The adapted guidelines were evaluated in two consecutive focus groups

- **Focus group 1 (Theoretical Perspective):** This group involved researchers and academics with a background in process discovery. Participants emphasized the need for flexible applicability across varying levels of expertise. They also recommended for the inclusion of support tools (e.g., checklists, guiding questions, and literature references). Furthermore, the person-centered discovery approach was subjected to critical analysis, leading to suggestions for integrating automated methods (e.g., process mining) in future iterations.
- **Focus group 2 (Practical Perspective):** This group comprised industry professionals with hands-on experience in process discovery. The experts validated the relevance of structured discovery guidelines, particularly for Small and Medium-sized Enterprises (SMEs). Key feedback included the necessity for a concise and clear structure, the adoption of a top-down discovery approach, the early clarification of scope and objectives, and the consideration of psychological aspects during interviews and observations. These insights prompted several structural

and content-based improvements. **Iteration 3 – expert field test** v An experienced process analyst employed the revised guidelines to document an IIoT production process. This evaluation phase revealed that the initial structure, based on traditional process perspectives such as control flow, data, organization, and resources (Jablonski and Götz 2007), was impractical in real-world settings. Consequently, the structure was improved to bridge prior theoretical insights with the practical progression of a typical process discovery project in industry. Rather than strictly organizing content by process perspectives, the revised version adopts a more intuitive, phase-based structure that reflects the practical workflow of a discovery project. For example, in Sect. 5.4.1, the new sections include *Starting the Workshop*, *Setting up a Process Profile*, and *Working in Groups*. Corresponding changes to align with real-world process discovery procedures were applied consistently across all guideline types. This holistic restructuring enhanced usability and made the guidelines more accessible for practical application across all methods.

Second use case

A second industrial application (see Sect. 6.2) validated and expanded upon previous findings. It became clear that:

- A pre-defined observation schedule significantly improves observation quality.
- Cross-method triangulation is essential for validating the completeness of IIoT models.
- Certain process segments (e.g., those operated by external providers) remain unclear ("black boxes"), requiring early inquiry into participant expertise.
- Live modeling during validation yields richer feedback than post-hoc surveys.
- While combining methods improves quality, it increases effort, highlighting the importance of a thoughtfully defined method sequence.

These insights were integrated into the guidelines, particularly through new sections addressing method selection, sequencing, and stakeholder coordination. The resulting artifact, comprising four distinct guidelines, is presented in the following section. The evaluation of these guidelines is detailed in Sect. 6.

4 Related work

To provide the scientific basis for the design of the artifact presented in this paper, an SLR was conducted following Okoli and Schabram (2010). An SLR ensures that potentially relevant information is reviewed as comprehensively as possible, facilitating the rigorous selection of required facts and details. The SLR presented in this paper follows a step-by-step methodology, encompassing the planning of the search, the identification and selection of relevant literature, and finally, data extraction and analysis.

The *Web of Science* metadatabase and the *ACM Digital Library* were used for the literature search. Within *Web of Science*, the search was restricted to the databases *Springer Nature*, *Elsevier*, *Emerald*, and *IEEE Xplore*. To ensure the most extensive possible coverage of information, the selection included journal articles, conference proceedings, and book chapters. To ensure a transparent and systematic selection process, the following inclusion and exclusion criteria were applied.

Inclusion criteria:

1. **Publication date:** Only literature published from 2007 onward was included. Following the formalization of the de facto process-modeling standard BPMN in 2006, BPM established itself as a holistic discipline with consistent terminology. Similarly, references to the IoT and its industrial counterpart began to become common in the mid-2000s. Consequently, publications predating 2007 were excluded from this study.
 2. **Language:** Only papers written in English were included to ensure consistency and comparability across the reviewed literature.
 3. **Full-text availability:** Only papers where the full text was accessible to the authors were included to enable complete evaluation and data extraction.
- Exclusion criteria:**

1. **Duplicates:** Duplicate publications retrieved from multiple databases were removed to avoid redundancy in the dataset.
2. **Thematic irrelevance:** Papers that did not directly address the research topic or failed to contribute to answering the research questions were excluded.

To find related work, the following research questions were defined: RQ1

Which connections between BPM and the IIoT have already been explored, particularly regarding process discovery?

RQ2 **Are there established approaches for general process discovery that may be transferable to the IIoT context?**

To address RQ1, we searched the selected databases using the iteratively developed search string below. The search was restricted to *Title* and *Abstract* fields to capture the current landscape of the field. This targeted scope ensures that the intersection of BPM and IIoT constitutes the core focus of the retrieved literature, rather than a peripheral mention. In the beginning, six duplicates were eliminated. In a second step, six papers that could not be accessed were excluded. Finally, the content of the identified literature was analyzed regarding its IIoT context and relevance to the present paper. We identified 12 of the 24 retrievable papers as relevant. These selected papers are discussed in more detail below.

("BPM" OR "business process management" OR "process management") AND ("IIoT" OR "Industrial IoT" OR "Industrial Internet of Things" OR "IoT" OR "Internet of Things" OR "Industry 4.0"))

To address RQ2, a broader search was conducted to identify existing process discovery methods. Unlike the restrictive scope of RQ1, this search encompassed *Title*, *Abstract*, and *Keyword* metadata. This expanded approach ensures that established manual discovery frameworks are not omitted due to the high degree of terminological variation common in the process management literature. Initially, one duplicate and two inaccessible papers were excluded. Following an analysis of the remaining 79 publications, 66 papers unrelated to manual process discovery methods were filtered out.

("process discover" OR "discovery method") AND ("document analysis" OR "observation" OR "interview" OR "workshop") OR ("business process discovery" OR "process elicitation")*

In Fig. 6, the search process and the steps taken for both research questions are illustrated using a PRISMA diagram (Page et al. 2021). Table 2 presents related works identified for the two research questions described above, along with their respective contents. Our analysis (see Table 2) reveals the absence of specialized manual methods for IIoT process discovery in the existing literature, with current efforts focusing predominantly on process mining approaches. While some studies address IIoT processes, they do not detail the methodologies used for their identification. The key findings from the related literature are discussed below.

The literature review by Giudice (2016) outlines the potential impact of IoT on BPM, both within and beyond the enterprise. Specifically, these impacts pertain to supporting BPM in terms of process optimization and implementation. They analyze the influence of IoT on BPM, rather than the discovery of IoT processes. Furthermore, while the use of IoT implies an industrial context, this is not explicitly specified. Crucially, the paper does not address the phase of process discovery.

Bazan and Estevez (2022) present an approach analogous to that of Giudice (2016). They examine the impact of IoT on BPM, emphasizing that robust organi-

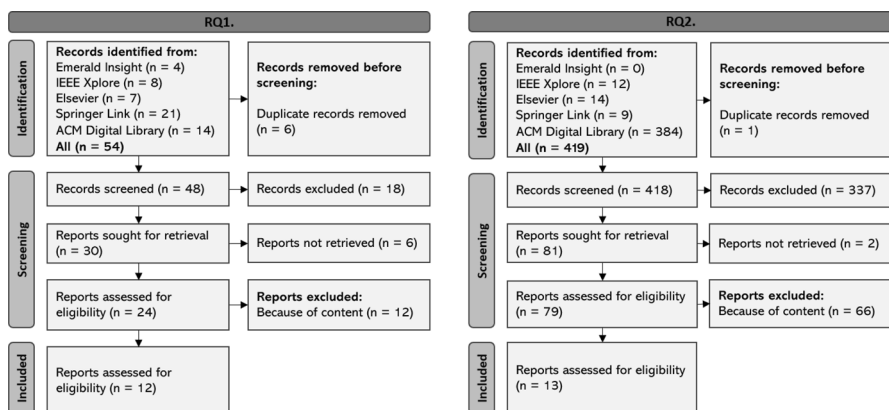


Fig. 6 PRISMA-diagram for RQ1 and RQ2 (Page et al. 2021)

Table 2 Review of thematically relevant literature with regard to the mentioned criteria

Authors	IIoT and BPM included	IIoT process used	Discovery described	Manual method used	Manual method described
Bazan and Estevez (2022)	✓	✓			
Benevento et al. (2022)			✓		✓
Burnay et al. (2020)			✓	✓	✓
Castiglione (2024)	✓	✓	✓		✓
D'hondt et al. (2019)	✓	✓			
Front et al. (2017)			✓	✓	✓
Ghose et al. (2007)					✓
Giudice (2016)	✓				
Grefen et al. (2018)	✓				
Harman et al. (2016)			✓	✓	✓
Han et al. (2020)			✓	✓	✓
Hewelt et al. (2020)			✓	✓	✓
Houy et al. (2010)	✓	✓	✓		
Janiesch et al. (2020)	✓				
Martin et al. (2020)			✓	✓	✓
Mass et al. (2016)	✓	✓			
Schönig et al. (2018)	✓				
Schönig et al. (2020)	✓	✓			
Schönig et al. (2020a)	✓	✓			
Schönig et al. (2020b)	✓	✓			
Seiger et al. (2022)	✓	✓			
Sirgo et al. (2018)			✓		✓

zational BPM facilitates the management of Industry 4.0 challenges. While process discovery is implicitly mentioned in this context, it is not explicitly detailed.

Schönig et al. (2020b) introduce a method for interacting with IIoT data during process execution. This approach leverages portable, real-time user interfaces to notify process participants of current tasks, thereby supporting efficient execution. While they describe an adaptation of process execution for industrial scenarios using IIoT data, they do not explain how the underlying industrial processes were originally discovered.

Seiger et al. (2022) investigate the integration of BPM and IIoT to provide common benefits to both research domains. Their work includes the development of an IIoT system architecture that integrates with BPM systems within smart factories. The primary objective is to enable sensor events to interact smoothly with existing BPM systems. To achieve this, IIoT technologies are integrated with BPM technologies throughout the BPM lifecycle, with a focus on process modeling, automation, and process mining. Although process discovery is not a central topic, the authors mention that process mining techniques can be used to discover processes based on logs and to obtain statistics. However, they do not provide a detailed explanation of the discovery methodology.

Janiesch et al. (2020), similar to Giudice (2016), discuss the relationship between IoT and BPM. Their goal is to highlight the challenges associated with bridging the previously distinct domains of IoT and BPM. Process discovery is also addressed, specifically via the use of log files and process data. However, apart from process mining, no other discovery techniques are described.

Ghose et al. (2007) present a novel method for deriving process models from textual descriptions. This approach is grounded in the assumption that documents serve as a primary source of information for modeling. Although it represents an evolution of document analysis that can speed up process discovery, this approach does not offer a detailed description of a specific manual discovery method. The objective is not to model a process from scratch, but rather to automatically generate an initial model that can subsequently be refined by the modeler.

Similar to Ghose et al. (2007), Han et al. (2020) present an automated business process service aimed at streamlining the process discovery phase of BPM projects. Unlike existing approaches that rely heavily on human knowledge to derive hierarchical structural relationships between activities from textual documentation, this approach leverages a neural network. It is designed to automatically identify latent hierarchical structures within documents. Through process-level language modeling, it accurately identifies process elements and generates BPMN scripts, significantly reducing the time and effort required for process discovery.

Recent literature emphasizes that relying on a single discovery method is often insufficient. In a comparative study of five discovery methods within a safety-critical healthcare setting, Christov et al. (2017) conclude that methodological triangulation is essential to capture the disparity between theoretical "work-as-imagined" and actual "work-as-done". This view is supported by Dolnicar et al. (2024), who criticize the over-reliance on self-reporting (via surveys or interviews) and highlight the superior validity of direct observation for establishing the ground truth of process execution. Despite the rise of automated mining, manual and hybrid methods remain dominant due to data inconsistencies and process complexity.

Martin et al. (2020) demonstrate that qualitative workshops are indispensable for identifying organizational barriers (such as culture and trust) that automated logs inevitably fail to capture. To address the communication gaps existing in standard interviews, Front et al. (2017) propose ISEA, a participative method that utilizes collaborative visual outlining to help stakeholders "co-design" the process. Furthermore, facilitation strategies must be tailored to the audience; Hewelt et al. (2020) found that while "Process-First" strategies suit novices, data-centric methods yield more robust models, nevertheless at the cost of higher cognitive effort.

The review of related work indicates that no manual approaches for process discovery within the IIoT context have yet been described. Janiesch et al. (2020) and Seiger et al. (2022) merely mention the possibility of process discovery based on log files and data generated during execution. Furthermore, we found no general guidelines for process discovery that are independent of the IIoT context. While several studies describe specific components of manual methodology, they are generally limited to the application of isolated techniques in specific contexts. Crucially, the

literature lacks a domain-independent framework for process discovery. This paper addresses the identified research gap by developing structured guidelines for manual process discovery in an industrial context, grounded in the work of Dumas et al. (2021).

Given that the literature review revealed an absence of manual methods for process discovery in the IIoT context, additional sources were consulted. Specifically, we examined Dumas et al. (2021) and Gronau (2017); while these are considered standard literature in BPM and process discovery, they describe only the available methods and offer no step-by-step instructions for their execution.

Dumas et al. (2021) provide insights into *established approaches to general process discovery*. The authors explain the methods of document analysis, observation, interviews, and workshops, offering a description and an example for each. While brief guidelines are provided for conducting interview- and workshop-based discovery, document analysis and observation are covered in less detail. Crucially, the presented methods are primarily contextualized within office environments rather than industrial processes.

Gronau (2017) outlines methods such as interviews, questionnaires, focus groups (conceptualized as workshops), observations, the inventory method (which encompasses document analysis), and automated procedures. The distinctions between these approaches are clarified, though the depth of detail regarding their practical implementation varies. The focus lies primarily on interpersonal components, with no specific reference to industrial contexts.

5 Guidelines for manual process discovery

This section presents an application proposal for the discovery methods *document analysis*, *observation*, *interview*, and *workshop*. Each subsection is structured into two parts: (i) the theoretical background and fundamentals of the process discovery method, and (ii) a description of the specific guidelines for that method.

The descriptions presented here are designed as flexible guidelines rather than rigid procedures. To ensure successful process discovery, the execution sequence may be adapted flexibly. However, upon completion of the discovery phase, it is crucial to ensure that all perspectives outlined in Sect. 2.3 have been captured. Fig. 7 illustrates the requirements that must be met once a process has been identified. Fur-

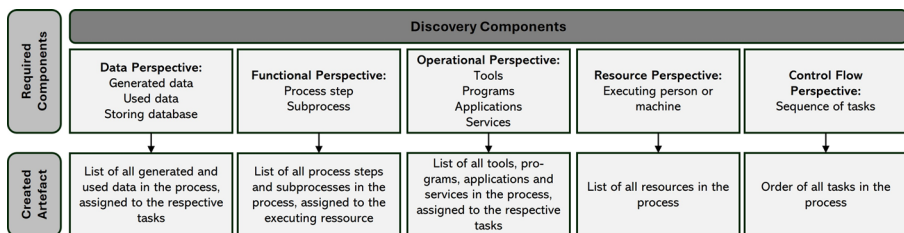


Fig. 7 Required components and artifacts created after successful process discovery

thermore, different discovery methods can be combined, allowing for the selective application of specific components from various methods. An exemplary approach for combining these methods is provided in Sect. 5.5. The individual components of the respective discovery methods are presented here in a suggested sequence, which is particularly beneficial for inexperienced process discoverers (PDs). These guidelines are explicitly designed for practitioners with limited or no prior knowledge of process discovery execution. Individuals with prior expertise may adapt the guidelines or selectively use specific aspects as needed.

We recommend developing a corresponding process map and identifying relevant processes within it before initiating discovery. The process to be discovered should be contextualized in relation to upstream and downstream processes. Furthermore, the objective and purpose of the discovery should be defined in advance to tailor the focus and level of detail accordingly. Depending on the desired granularity, the discovery can be executed using a top-down approach.

5.1 Document analysis

5.1.1 Document analysis method

Existing process descriptions such as internal policies, organizational charts, manuals, models, or instructions can be used for process discovery (Dumas et al. 2021). Similarly, the PD can discover information from business distribution plans, regulations, budget, economic, personnel, and staffing plans, directives, job descriptions, requirement documents, design documents, order statements, or meeting protocols (Federal Ministry of the Interior and Community 2012). The discovered documents, containing information about the current state of the process, are reviewed, and process-relevant information is extracted (Federal Ministry of the Interior and Community 2012). Based on this information, a model of the process can be created. During document analysis, the following information can be extracted for process discovery (Federal Ministry of the Interior and Community 2012): Task structures, areas, and carriers; organizational structures and process workflow; specifications and conditions within the process; decision points; time required; linked and used documents in IT systems.

There is no information on the approach or techniques for extracting information from the documents and converting it into a process model. Documents are usually not process-oriented and therefore not uniformly granular. This leads to different levels of detail in the various documents, making direct derivation of processes from the documents usually not possible (Dumas et al. 2021). Furthermore, the documents used may be outdated or incorrect and do not guarantee the correctness of the information contained therein as they refer to past states (Dumas et al. 2021; Hansen et al. 2019). However, documents are free from personal evaluations and thus provide an objective basis for familiarizing oneself with a process. They serve as a starting point for process discovery (Dumas et al. 2021; Hansen et al. 2019; Federal Ministry of the Interior and Community 2012).

5.1.2 Guideline for document analysis

Preparation

The preparation of the document analysis starts with the collection of all process-relevant documents. These include, for example, business distribution plans, business regulations, budget, economic, personnel, and staffing plans, instructions, job descriptions, requirement and design documents, order declarations, machine manuals and operating instructions, specification documents for involved individuals, site plans of the production facility and other documents related to the process or involved entities. Sort the documents according to their level of detail and check short and structured documents first, followed by more extensive documents. Finally, focus on machine-specific documents due to their specificity. This top-down approach provides a clear overview before delving into details systematically.

Discover resources

In this step, the documents are reviewed from the *resource perspective*. All documents are examined to determine which executing instances, i.e., people and machines (short: resources), are involved in the execution of the process. Each document is worked through individually, and notes are taken as soon as a machine or a role of a person performing a specific activity is mentioned. Person roles are often terms describing activities, i.e. "truck driver". Machines are usually identified by their own names i.e. "welding machine". If explicit names for machines or person roles are not provided in the documents, the PD should assign appropriate designations. It is crucial that these designations are used consistently to avoid confusion, duplicates, or other errors.

Discover tasks

As soon as all resources have been identified and noted, we recommend recording the associated tasks, the *functional perspective*, and assigning them directly to the discovered resources. All documents should be reviewed, and every activity and task (short: task) mentioned should be noted. It should also be noted which resource performs the task that this can be assigned directly during modeling later. A task is one step of work in the process.

It is important to capture all tasks that can be performed by a resource. This applies regardless of whether the tasks appear to be related, occur only in special cases, or have other unusual features. If a task can be performed by different resources, the task is noted for each participating resource, along with an indication which other resources can also perform this task. This will be taken into account later in the modeling by creating an exclusive decision node, which ensures that the task is only executed once. It is also possible to discover the tasks independently of the previous resource acquisition. Nevertheless, it should be ensured that the tasks can be assigned to the correct resource later.

Discover tools

The *operational perspective* is captured by discovering these tools, programs, applications, services, and software components (short: tools) used to perform a task.

For each task, it is determined how it is implemented in practice. All available documents must be reviewed and searched for the mentioned tools and the task where they are used. A direct assignment of the tools to the task is recommended directly when the tool is discovered. For example, if the ERP system is used, this is noted for the corresponding task. It is also possible to detect the tools independently of the previous task detection. Nevertheless, it should be ensured that the tools can be assigned to the correct task later.

Discover data

The data which is produced or used in each task covers the *data perspective*. The documents should be read through chronologically, and any data element mentioned should be noted. For each data element noted, the task in which it is produced and used should be noted. If a data element is produced in a task, details such as the type of sensor used, data capture intervals, data format, and destination are useful to record. If a data element is used in a task, the purpose, format, and source of the data are possibly interesting. In both cases, the storage location of the data, typically a database, is also indicated. To manage the discovered information clearly, it is recommended to use tables that record the aforementioned properties as columns and the data elements as rows.

Discover sequence

The sequence of individual tasks is extracted from the documents to represent the *control flow perspective*. All documents are read through, examining which task the process starts with and which further tasks follow. It is advisable to identify the starting point of the process and then sequentially determine the next step. If there are decision points or intermediate events within the process, these should also be recorded. This may include information such as the duration of certain process steps, the conditions under which certain tasks are executed, and identifying which tasks sensors produce data.

Validation

Whether the process has been accurately discovered cannot be immediately verified. There is the possibility of conducting the entire document analysis again and recreating the model independently of the initial capture. An alternative method for validating the information is to choose another discovery method like observation, interviews, or workshops and execute it. Subsequently, the process information can be compared. If discrepancies are found between them, this indicates that at least one of the discovering was incorrect or incomplete. In such a case, it is necessary to review the documents again to identify and rectify any errors and incompleteness.

To discover the individual perspectives, the documents can be run through in several iterations for one perspective at a time or all perspectives can be run through within one iteration. For clarity, we recommend proceeding in several iterations and identifying the resources first. Then discover the tasks and assign them to the corresponding resources. Discover the data and tools at the same time and assign them directly to the tasks. Finally, the tasks can then be put in the correct order using the documents.

Advantages and disadvantages

The advantages of document analysis include that it does not disrupt ongoing production processes, allowing operations to continue without interruptions. Additionally, it entails no extra personnel resource costs and provides detailed documentation that enables a comprehensive understanding of the processes. These documents remain available at all times, serving as a continuous point of reference.

However, document analysis also presents some disadvantages. It demands a significant time investment from the PD, with analysis time increasing as more documents are included. There is also the risk of misunderstandings arising from complex technical machine descriptions, often requiring the involvement of additional experts for clarification. Furthermore, the accuracy of process models can suffer from outdated or incomplete information. Particularly in IIoT environments, documentation may be lacking or not detailed enough in crucial areas such as data flows and machine communication.

In conclusion, conducting document analysis is sensible when there is a solid base of up-to-date and detailed documentation and when processes cannot be disrupted. It is especially useful in stable, well-documented environments where process understanding is needed without interfering with ongoing operations. However, in highly dynamic or poorly documented IIoT settings, alternative methods or supplementary interviews and observations may be necessary to avoid incomplete or inaccurate results.

5.2 Observation

5.2.1 Observation method

During observation, the PD observes the execution of the process and can directly capture the current process activities (Dumas et al. 2021; Gronau 2017). In business processes, the PD can take an active role (e.g., the perspective of the customer) or a passive role (e.g., the perspective of the employee) to conduct overt or covert observations (Dumas et al. 2021; Gronau 2017). The PD documents their observations and can subsequently translate them into a process model. Similar to document analysis, the literature does not provide specific techniques or instructions for conducting observations. Since observation is passive, no workflows or processes are interrupted. There are no additional costs due to work interruptions or production stops during process discovery (Gronau 2017; Hansen et al. 2019; Dumas et al. 2021). Furthermore, observation enables the PD to gain a better understanding of process boundaries and reduces reliance on information from documents or process participants (Dumas et al. 2021). Observation allows for an objective view of the process from the perspective of the PD and is not exposed to the subjective influences of process participants. However, it should be noted that employees may change their behavior if they know they are being observed and may work in a more structured or faster manner than usual (Dumas et al. 2021; Hansen et al. 2019). Additionally, observation requires that the PD has access to the process, which may require access

to remote facilities or locations (Dumas et al. 2021). It may be temporally or spatially impossible to continuously observe a process. In such cases, the PD may need to observe the process multiple times to ensure that all possible process paths have been observed at least once.

5.2.2 Guideline for observations

It is recommended to model the process during observation. This allows the process to be captured based on the sequence of execution, enabling the components to be recorded directly in the correct order. Thus, the guiding perspective for observation is the *control flow perspective*.

The PD tracks the manufacturing of a particular component from the beginning to the end of the process. For example, if a piece of wire is clamped into a machine at the beginning, the PD follows this piece of wire until the end of the process, where a complete final product is produced. The PD starts at the beginning of the process and records the trigger of the process. Then, they track the first process step. If there are multiple ways the process can start, one possible starting event is chosen, while all other possible starting events are noted. After the observation is completed, the process is observed from another starting event until the first common task is reached after all starting paths. Familiarity with the specific domain where a process takes place can be practical for a PD, as it facilitates the identification, differentiation, and detailed specification of individual tasks. However, such domain-specific expertise is not an essential prerequisite. As demonstrated in the use case presented in Sect. 6.2, a PD can also conduct effective process observation and modeling without prior specialist knowledge.

Discover tasks

For each process step, the PD observes what happens in the process step to cover the *functional perspective*. A process step in the observation is an executed activity and is thus differentiated from the next or previous step. It should be possible to record the observed activities as compactly as possible. The level of detail of a task depends on the previously defined granularity of the process.

Discover resources and tools

At the same time as observing the task, the PD can make a direct note of whom performs this process step (which resource) and whether it is clear which tools are used for execution. This covers the *resource* and *operational perspectives*. The resources and tools can also be discovered independently of each other and of the tasks. However, it is important that they are assigned correctly. We therefore recommend assigning a task directly to the executing resource and noting the tools used. The executing resource can be directly modeled as a lane. If it already exists, the step is skipped. The observed task is modeled as a task and assigned to the corresponding resource. Then, any tools used are linked to the task. Once all components are captured, the PD moves on to the next process step and repeats the process.

Discover sensors for the data perspective

The discovery of data perspective is not possible in the observation. Nevertheless, sensors that are sighted during the process can be included in the model. This allows the model to show where data is generated in the process.

Handling branches

Once the PD recognizes that the process branches, it is advisable to follow one path at a time and not try to capture multiple paths simultaneously, because the size and complexity of the process can lead to confusion. The later model could become faulty or incomplete if the PD forgets parts of the process or neglects a path. To maintain clarity, the following approach is recommended when encountering a process branch:

1. The PD notes in a separate document the task where the process branched and the number of possible additional paths at this point.
2. The PD chooses one of these paths and continues through the process.
3. Once the PD reaches the end of the selected path, they return to the marked point and choose the next path.
4. The process is fully traversed again, and this process is repeated until all noted branches are addressed.

If there are further branches of the process in a path, the task and the number of paths are noted. The processing of the noted branching points is done chronologically according to the notation, with initially noted branches being processed first. This ensures that all branches are addressed, and the process is modeled along the control flow.

Handling overlooking of paths

The overlooking of paths in forward observation occurs when paths are rarely executed. If a path is not activated during observation, it cannot be observed and therefore cannot be captured in the model. However, backward observation does not guarantee the capture of rarely executed paths. To reduce the likelihood of overlooking paths, the following measures can be taken:

- *Multiple iterations:* Observation can be conducted multiple times, both forward and backward. Each additional iteration increases the chance of capturing rarely executed paths.
- *Increased observation duration:* To ensure that even rare paths are captured, the observation duration at each process branch can be extended. This allows for the discovery of unusual or less frequent process flows.

Validation
The accuracy of the observation process discovery can be validated. To do this, the observation is repeated at the end of the process. This ensures that paths or tasks that may have been overlooked in the forward observation are discovered in the backward

observation. The observation usually begins at the finished product. Each previous task is considered and compared with the process model. Special attention is paid to process path mergers. This occurs when a task has multiple possible preceding tasks, resulting in multiple process paths leading to this task. The observed task is noted together with the number of triggering options, as with forward observation. Subsequently, one of these possibilities is selected, and the observation is conducted backward until the beginning of the process is reached. If additional path mergers occur during this backward observation, they are also re-noted. All mergers should be processed chronologically until all have been completed. The goal of backward observation is to uncover paths overlooked during forward observation. For example, if only three possible paths were identified at a process branch, when there are actually four possible paths, the backward observation might uncover four converged paths. Once all these paths have been traversed backward, a path not included in the current model can be identified and added.

Using observation for IIoT processes

Although it is not possible to discover the *data perspective*, observation is suitable for recording IIoT processes. This is because prior observation helps the PD to understand the process and provides initial insights into the process, which the PD can use to develop questions and initial models based on their prior knowledge using other discovery methods, e.g., in an interview. This is shown in Sect. 6.1.

Advantages and disadvantages

The advantages of using observations for process discovery include an objective view of the process and a clear overview of all visible process components. Observation can be carried out within a defined physical space, such as a production facility, without interfering with ongoing operations. It does not require additional resources or costs, and the process flow often becomes quickly clear to the PD. However, this approach also has disadvantages. Certain process operations, especially those occurring inside machines, are not visible. Critical aspects like data usage, network traffic, and machine communication remain hidden. Additionally, physical constraints, such as machine placement, can limit the PD's ability to observe all steps. Rarely executed paths may be missed, leading to incomplete models. Furthermore, measuring waiting times can slow down the process discovery and add to the overall time consumption.

Conducting an observation-based process discovery is advisable when the goal is to quickly gain an initial, unbiased understanding of visible process flows without interrupting the process itself. It is particularly effective when the majority of critical activities are externally visible. However, for complex processes with significant hidden operations or data-driven components, observation should be supplemented by document analysis or expert interviews to achieve a complete and accurate process model.

5.3 Interview

5.3.1 Interview method

In the interview-based method, a subject matter expert (process owner) involved in the process is personally interviewed (Dumas et al. 2021; Gronau 2017). If multiple process owners are involved in the process, multiple different interviews must be conducted.

Dumas et al. (2021) propose two strategies for conducting interviews: (i) backward and (ii) forward process discovery. In (i) the interview starts at the end of the process and works backwards through the process, asking "What happens before?" and continues until the trigger of the process is reached. (ii) operates in reverse, discovering data on the process from the beginning to the end state of the process. Furthermore, Dumas et al. (2021) recommend choosing between a structured or open interview approach. Structured interviews involve predefined questions and hypotheses that need to be confirmed or corrected. This ensures that specific questions of the PD are answered. However, important information may be overlooked in a structured interview as it may not be included in the predefined questions. Additionally, exceptional cases may remain unaddressed. Open interviews allow the process owner to present important aspects of the process from their perspective, potentially revealing information that would have remained hidden in a structured interview. On the other hand, the process owner may skip important information, thus hide it from the PD (Dumas et al. 2021).

To conduct process interviews, the Bundesverwaltungsamt (2013) has created a "*Guideline for the Collection of Business Processes*". To prepare for the interview, available documents may be reviewed to construct a preliminary process plan based on their content. This information requires verification during the interview. It is important to note that this step does not constitute a formal document analysis; rather, it serves as a process familiarization exercise, providing the PD with an initial overview based on available background information. They provide tips for interpersonal aspects that facilitate conducting interviews. No further guidance on structured interview conduct is provided. Gronau (2017) point out general principles for conducting interviews:

- Interview questions must be formulated and asked exactly.
- The PD must be aware of their role as an interviewer. Their interpersonal skills influence the success of the interview and the relationship with their interviewees.
- The information obtained must be carefully documented.
- The interview questions must be tailored to the expertise of the process owner.

The process model is created from the interview notes or parallel during the interview (Dumas et al. 2021). Conducting interviews requires significant time and resource investment. When multiple process owners are involved in a process, inconsistencies may appear that the PD must clarify. This requires verifying the correct process flow, leading to increased time investment (Dumas et al. 2021). Qualified subject matter process owners familiar with the process details are required. During interviews, sub-

ject matter process owners may not be able to perform their regular duties, potentially impacting operations (Krallmann et al. 2002).

5.3.2 Guideline for interviews

One or more interviews can be conducted for the interview-based process discovery. This depends on the complexity and size of the process and the personal preferences of the PD and the process owner. We recommend conducting a separate validation interview after the discovery, which takes place on a separate date, to verify the correctness of the model once it has been fully modeled.

Interviews with multiple process owners

If processes are large and complex (especially IIoT processes) they can be divided into different sections. Each section typically has different process owners. To conduct a process discovery through interviews, the PD should invite all process owners to separate interviews. These interviews can follow or use parts of the guideline described below with each invited participant separately. Once all sections of the process are discovered, the PD must arrange them in the correct sequence. If there are interactions or cross-connections between the interviews or if process owners contradict each other, it is advisable to talk to the process owners involved together to clarify the affected process component.

Preparation through questionnaire

In the beginning, the PD should provide the process owner with a preliminary questionnaire regarding the process. The questionnaire should include questions covering all process perspectives. It is important to use terms in the questionnaire that are familiar to individuals without a background in process modeling. Possible questions for the questionnaire could include:

- What is the main goal or output of the process? What is produced, and what is the outcome of the process?
- Are there specific triggers or conditions starting the process?
- Which individuals, machines, and IT systems are involved in the process? (*Resource perspective*)
- What is the flow of the process? Please describe chronologically what happens in the process. (*Functional and control flow perspective*)
- Are supporting software or other tools used to perform tasks? If yes, please name the tasks where such support is provided and describe the type of software, tools, or other aids involved. (*Operational perspective*)
- Are there specific procedures or measures that are executed in case of an error occurrence?
- What happens to unusable or defective parts identified during the process? Are there mechanisms for monitoring and segregating such parts? Please explain these.
- Which sensors are present within the process? Where in the process are these sensors used (at which tasks)? What data do these sensors each capture and produce?

Please describe this briefly. (*Data perspective*)

- Where are the captured data sent? Where does data processing take place? (*Data perspective*)
- Which machines within the process are interconnected? How do these machines communicate with each other? What information or data do the machines exchange?
- Is there a data network within the process? If yes, please describe and sketch it briefly. (*Data perspective*)

Based on the answers, the PD can create an initial information overview or a process model, considering the various perspectives. The questionnaire allows minimizing the duration of the further interview and gives the PD the opportunity to familiarize themselves with the process before the interview. The PD can already prepare more detailed questions for the first interview, considering the following topics:

- Have any uncertainties or contradictions appeared due to the answers from the first questionnaire?
- Is the process model complete or are there paths that do not have a defined end?
- Does the sequence of tasks in the process make sense?

Starting the interview

We recommend dividing the interview into several parts to allow the PD and the process owner to take breaks. However, it is also possible to conduct only one interview where only parts of the proposed methods are used.

The PD invites the process owner to the interview. Two forms of an interview can be conducted: a closed interview with prepared questions that already explicitly address specific content. This is particularly recommended if a questionnaire has been sent out in advance. During the interview, the PD asks the process owner the prepared questions and notes down the answers. Otherwise, in an open interview, the process owner can be explicitly asked to talk about the process on their own initiative using open questions. This gives the process owner the opportunity to address additional comments and ensures that the PD receives information that may not have been explicitly asked for. Examples of conducting interviews using open and closed interviews can be found in Dumas et al. (2021). Further literature on conducting interviews is provided by Hochstetter-Diez et al. (2025), Burnay et al. (2020), Taherdoost (2022), Turner and Dw (2010), Gubrium et al. (2012) and DiCicco-Bloom and Crabtree (2006).

Check understanding

To check if the PD has understood the answers and process descriptions of the process owner correctly, we recommend carrying out the following step: The PD goes through their notes with the process owner and asks them about each process step, confirming whether the representation and understanding of the process flow are correct to avoid potential errors in modeling. If there are any errors or incompleteness, the process owner corrects them. Alternatively, the PD can integrate the information received during the interview into the existing process model and expand it or, if no model is available yet, model or document it. If a model already exists, the PD

goes through the process model with the process owner and explains the modeled steps. Here again, the process owner should point out any errors or incompleteness and make corrections. It is important to note that the PD should never assume that the process owner is familiar with the modeling language. Therefore, the PD should explain each modeled process step by describing what happens at that point, according to the model. If a correction or addition is necessary, the process owner should communicate this verbally. The PD can either note this information and integrate it into the model later or make the changes directly in the model.

After going through the notes or model once, the PD should ask the process owner if there are any additional information or process components that have not been captured yet (e.g., a process branching that triggers another, previously unrecorded path). If such information is available, it should be integrated into the notes or the model. If desired, a pause can be taken between the interviews at this point.

Depending on the duration of the first interview and the extent of the model changes, it may be useful to end the interview at this point and conduct a second part in a separate session. This allows the PD to integrate new insights from their notes into the model without causing unnecessary waiting time for the process owner. Updating the model can take several hours depending on the extent of the notes and the complexity of the process. Additionally, the concentration of the participants decreases after a certain duration.

Detail contents

In a second interview or as a second step, the process can be detailed and process exceptions identified. It can be assumed that the first model does not have a high level of detail yet and that none of the rarely executed paths have been captured. These details should be worked out in the second interview. This includes identifying subtasks, defining decision rules at relevant branches, integrating sensors at suitable locations in the model for data discovery, documenting the data forwarding and usage at the respective goals, and verifying that all possible process paths are mapped in the model, regardless of their rarity of execution. To discover this information in a structured manner, the PD can question the process owner based on the revised model from the first interview. It is recommended to go through the following questions for each task to capture all necessary information:

- Does the task divide into further subtasks? If yes, into which ones?
- Are data discovered at this point in the process? If yes, which ones, why are they discovered, where are they sent (to which other process task), and where are they stored (at which storage location)?
- Are there additional paths that can follow this task that have not been mapped so far?

The PD may further divide tasks and create multiple tasks or model subprocesses if necessary. Data is attached or referenced as data objects to the respective tasks, and connections are documented. Once all tasks have been explored, all branches are examined chronologically, and for each path, it is determined what conditions must be met for the path to be executed, the interview ends. The PD can add the informa-

tion into the model now. To pursue the top-down approach, the interview can also be conducted several times to discover the process more and more granularly until the desired level of detail of the process is achieved.

Area expert

If the process owner (contrary to the assumption) does not have enough detailed knowledge of the process to answer the questions and describe the process at a certain point, an area expert must be involved. This person is also familiar with the process but has specific expertise in a particular area. The area expert is called in for a discussion and supports the process owner where they lack detailed knowledge of the process. The area expert is explicitly involved only in this specific section of the process.

Validation

We recommend using a validation interview where the PD and the process owner again go through the extended process model to check for correctness and completeness. The process should proceed chronologically, and at each task it is checked whether the task is correct, the associated data is correct, the executing resource is correctly assigned, and the task order is correct. It should be verified whether the tasks before and after are actually executed at this point in the process flow. When the process reaches a branching point, its paths and the conditions required for them are checked. Subsequently, the cooperation between the individual machines and roles is examined, with each connection being verified to ensure it actually exists and connects the correct resources. If the process owner has no corrections and confirms that the model is complete and correct, the process discovery is completed. If errors are found by the process owner, they should be corrected after the interview. The interview can be repeated until the process owner confirms the model as correct and complete. However, the model can also be finished after the corrections of the validation interview or terminated after a certain number of validation steps. The PD decides when the model and the process discovery are complete.

Advantages and disadvantages

The advantages of conducting interviews for process discovery include gathering up-to-date and detailed information directly from the process owner. Interviews allow for immediate clarification of content-related questions and misunderstandings, making certain that even hidden aspects, which are not visible through observation or documentation, can be addressed. Additionally, interviews make it possible to inquire about all process paths, data flows, and network interactions, offering extensive opportunities also to gather IIoT-related insights. This method significantly increases the likelihood of developing a complete and accurate process model.

However, interviews also have notable disadvantages. They can be both time- and cost-intensive, disrupting normal process flows as the process owner is occupied and unable to perform their usual tasks. The accuracy of the discovered information heavily depends on the assumption that the process owner is thoroughly familiar with the process or can consult area experts. There is also the risk of receiving incomplete or incorrect information, especially if the process owner unintentionally presents an

idealized version of the process rather than its current state, resulting in inaccuracies in the model.

Conducting interviews is particularly sensible when in-depth knowledge, hidden process details, and IIoT-specific information (e.g., sensor data, data acquisition/storage, and decisions based on IIoT data or machine communication) are required, especially in complex systems where neither observation nor documentation alone can capture the full process. Interviews are most effective when combined with other methods and when process owners are well-informed and able to provide an honest and realistic depiction of the current process.

5.4 Workshop

5.4.1 Workshop method

The workshop-based discovery method is similar to the interview method but involves several process experts at the same time (Dumas et al. 2021; Jadhav 2011). When selecting participants, it should be considered that small, hierarchically homogeneous groups work more efficiently than large groups or groups with hierarchical heterogeneity. It is not recommended to include executives and their employees in the same workshop (Richerzhagen 2015; Dumas et al. 2021). Technical staff involved in the process indirectly, for example, through the management of supporting systems (e.g., ERP systems) should also be involved. However, the maximum number of twelve experts in a workshop should not be exceeded (Dumas et al. 2021). Jadhav (2011) recommends conducting review workshops, where subject matter experts must approve the modeled process afterwards to ensure that the process is consistent and correct. Dumas et al. (2021) suggest that a moderator moderate the workshop and should coordinate the contributions of the participants and equally involve the process experts in the workshop. Involving a moderator allows the PD to design a model during the workshop. A process analyst can be included as a third person to note relevant statements made by the participants, which may need to be followed up on later that no interruptions of the discussion for questions are needed. This approach is particularly suitable for complex processes involving many experts (Dumas et al. 2021). If there are different opinions within the group regarding the content or presentation of the process, it is the moderator's task to guide the group discussion that ultimately the correct process can be identified (Bundesverwaltungsamt 2013).

Since a detailed model cannot be created in one session due to complexity, multiple sessions are required (Dumas et al. 2021). The first session is used to establish the context, communicate the workshop's goals, and identify the participants' expectations (Jadhav 2011; Dumas et al. 2021). In addition, the scope of the organization to be considered must be defined, whereby the use of process maps, if available, is appropriate (Richerzhagen 2015).

5.4.2 Guideline for workshops

A process discovery through workshops should involve the preparation and at least one or more workshops. The workshop-based discovery can be carried out in one or in several workshops. The PD can incorporate the information into the process model between the workshops. Otherwise, the model can be created in parallel to the workshop or afterwards. For the workshop, we also recommend analyzing the individual perspectives and pursuing a top-down approach. For guidance on moderation and interaction within a workshop, see the literature by Lauttamäki (2014); Linds and Gee (2023); Pavelin et al. (2014); Jolles (2017) and Storvang et al. (2018).

Preparation

To conduct a process discovery through a workshop, the PD has the opportunity to decide whether to moderate the workshop and gather information simultaneously or delegate one of these tasks to another person. Additionally, as suggested by Dumas et al. (2021), an analyst may be involved. The planning of the workshop begins with the selection of relevant participants. The PD needs to determine which participants have a detailed understanding of the process. It is important to ensure that at least one participant from each process area attends the workshop to ensure that the entire process is adequately represented by different participants. Participants can include employees who perform direct tasks in the process, as well as those who are indirectly involved in the process, i.e., machine operators or developers of the programs executed in the machines. The PD schedules the dates for the workshop and invites the selected participants, taking care to find dates when all invited participants are available. For the execution of the workshops, it is advisable to follow the suggested approach of Camunda BPM by Richerzhagen (2015).

Starting the workshop

In the beginning, the PD should introduce the process being discovered and briefly explain the planned approach for the discovery. Then, all participants should introduce themselves and explain their involvement in the process by describing their roles and providing rough descriptions of their tasks.

Set up a process profile

We recommend starting the workshop by creating a process profile to get an overview of the process. To do this, the participants are asked to list the roles involved in the process (*resource perspective*) and jointly describe a rough process flow (*control flow perspective*). When describing the rough process flow, the participant responsible for each area should explain what happens in that section of the process. The moderator should ensure that all participants contribute equally to the discussion and ask targeted questions. The PD records the information gathered and may directly capture it as a model, while the analyst takes note of additional information that may require further investigation through follow-up questions.

Once all necessary information has been gathered, the analyst can raise any follow-up questions. These questions should be answered by the participants, allowing the PD to directly capture the information. Then, the PD presents their findings and shares their understanding of the process flow, the roles involved, and process outcomes with the participants. Participants are encouraged to make any necessary corrections and provide additional comments. The moderator also guides this discussion. This part of the workshop is the ideal place to take a break that the PD then can document the results in the form of a process model.

Working in groups

Depending on the group size, the process discovery can be done collectively with the moderator in the entire group or, for a large number of participants, in separate groups working independently. Working in small groups is advisable if the moderator considers the overall group to be too large. When working in small groups, the process should be divided into several sections, with each small group assigned a specific section of the process. Each small group should include participants who hold roles in that section of the process. If a role appears in multiple sections of the process, the participant is assigned to the section they consider to be the most complex or the group work is not carried out at the same time to allow the participant to be present in both groups. Once the work in the small groups is completed, they present their results. This is done in the order of the process flow. After all small groups have presented their results, further additions and corrections are discussed in the entire group. Participants, who were required in several stages of the process, were asked to contribute to this discussion. The PD notes the comments for each section. If the workshop is conducted with all participants at the same time, the work is similar to that in small groups.

Discovering tasks, tools and data

We recommend focusing on the *functional*, *operational*, and *data perspectives* after discovering the process profile. The PD should present the current process model if this is already existing. Subsequently, the refinement of the process and the identification of operational support are carried out.

The whole group or each small group discusses, based on the current process model or information, the refinements needed in each task and whether all tasks for each role in that section are mentioned (*functional perspective*). The additions and refinements should be documented in writing.

For each task, the software and other support tools used should be noted chronologically (*operational perspective*).

A list can be created that lists all data generated or used during the process to focus the *data perspective*. It is then recorded for each data record which task generated it, what type of information it contains, and where it is used in the process. If the data originates from another process or is sent to another process, this must also be documented. Together with the participants, the PD then assigns the respective data to the producing and consuming tasks, thus establishing the connections between the machines. Finally, the storage location for all data objects (usually the respective database) is determined and noted. The process steps could be detailed and refined,

and the support tools used are added. Subsequently, the PD should integrate the new information into the process model. The PD can directly represent this in the modeling tool and make the process model visible to all participants, for example, by using a canvas projection. The order and level of detail in which the various perspectives are recorded can be customised.

The moderator has to ensure that each discussion is continued until the participants agree on the correctness of the discussion component. It is important to ensure that participants are not persuaded to make a statement or agree due to hierarchy, as this could compromise the correctness of the process. Higher-ranking employees should not have decisive authority.

Complete workflow

To complete the discovery, all decision nodes in the model should be reviewed. This can be done during the creation of the process profile, separately at the end of the workshop or in a separate workshop. When this discovery should take place can be made dependent on the size and complexity of the process and the number of decision nodes. At each decision node, discussions are held regarding the conditions triggering the paths, and the triggering data values are noted. Once all decision nodes have been examined, the analyst asks questions about the recorded information. Once these questions are answered, the workshop is concluded, and the PD can finalize the process model.

Validation

In the final workshop, the PD presents the completed information or model and explains the entire process flow to the participants. If all participants agree with the presented information or process model, the workshop, and thus the process discovery, is considered complete. If there are any improvements, corrections, or additions, these are noted and added. The final workshop can be repeated until all participants agree with the presented information or process model or, as with the interview, can be terminated after a certain number of runs or other termination criteria.

Advantages and disadvantages

Workshops offer several advantages for process discovery. They enable the involvement of multiple individuals at once, allowing for the immediate resolution of disagreements and examination of the process from various expert perspectives. This collaborative environment ensures a more complete understanding of semi-automated related information, data flows, and machine connectivity. As a result, workshops significantly increase the likelihood of developing an accurate and comprehensive process model compared to relying on the input of a single process owner.

However, workshops also come with some disadvantages. They are resource-intensive, requiring participants to step away from their regular duties, leading to higher personnel costs. Scheduling becomes more complex and time-consuming as the number of participants grows, and process disruptions may occur during the workshop. Additionally, there's always the risk that certain technical aspects remain unclear if no participant has complete knowledge of those areas.

Conducting a workshop is particularly sensible for complex processes that span multiple areas and require input from various specialists. They are most effective when comprehensive knowledge from different domains is essential, and when the effort and cost of organizing the workshop are justified by the need for accuracy and completeness in modeling the process.

5.5 Combined guideline method application

As previously established, the components of the guidelines are designed for modular application. This flexibility allows researchers to mitigate the limitations of individual methods and adjust their influence. This section introduces a proposed hybrid configuration, detailing the underlying rationale and objectives. By integrating these techniques, we aim to generate a comprehensive, accurate, and verifiable opportunity that captures all relevant IIoT data. Crucially, this approach seeks to minimize resource consumption (time and cost) without compromising model quality. Consequently, as part of our DSR approach, we developed an exemplary framework that sequences these methods strategically. Developed in collaboration with domain experts, this configuration serves as an illustrative instantiation. As the subsequent use cases demonstrate, alternative combinations may be employed depending on resource availability and information constraints. This proposed sequence begins with process observation, followed by targeted document analysis and an initial interview. Finally, a follow-up interview and a collaborative workshop are conducted for validation, as illustrated in Fig. 8.

The observation phase is purposed to provide the PD with initial insights into the process, facilitating the design of a preliminary model that includes the involved resources and their visibly executed activities. Through observation, aspects of the *resource*, *control flow*, and *functional perspectives* are partially covered. To conduct the observation, the PD follows the instructions outlined in Sect. 5.2.2 and goes through the process, recording the sequence of performed tasks and the corresponding executing resources. Process branches are handled as previously explained. The result of the observation is an initial process model where resources are represented as lanes and their executed tasks are modeled as activities in the correct sequence.

Subsequently, the PD proceeds to document analysis, reviewing existing documents to extract the following information in the designated order:

1. Additional tasks and paths that are not captured during observation. (Completion of the *control flow* and *functional perspective*.)
2. Software and supporting tools that are used for each task. (Discovering the *operational perspective*.)

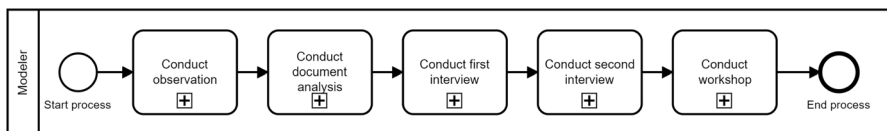


Fig. 8 Combination of the individual methods into one process discovery

3. Data produced and consumed in relation to each task. (Discovering the *data perspective*.)
4. Interconnections between the involved machines. (Discovering the *connectivity*.)

To gather this information, the PD iterates through the documents multiple times. In the first iteration, the focus is on the *control flow* and *functional perspective*. In the second iteration, the *operational* and *data perspectives* are discovered. In the third iteration, information about *machine connectivity* is gathered. Subsequently, the PD also notes any uncertainties and questions arising from the existing information or model. For the document analysis, the PD should refer to Sect. 5.1.2. If the PD finds inconsistent information in the documents compared to what was discovered during the observation, this information from the documents should be ignored as its accuracy is not guaranteed, unlike the observation. Upon completion of the document analysis, the PD already possesses a comprehensive and detailed process model. The correctness of the model has been partially ensured through observation. So far, only costs and time commitments for the PD have been incurred.

In the third step, the process owner is invited to an initial interview session. As mentioned in Sect. 5.3.2, it may be necessary to interview multiple process owners. The following formulations refer to a single process owner but can also be optionally conducted with multiple process owners for different process areas. A detailed questionnaire is not sent in advance to avoid further increasing the time commitments of the process owner. Instead, the process owner is simply informed about the process that will be discovered and provided with the questions that the PD has noted for preparation. In the interview, the PD proceeds according to the instructions in Sect. 5.3.2. The PD asks the previously noted questions and records the answers from the process owner. The process owner then has the opportunity to contribute additional relevant aspects. Afterwards, the PD presents the current process information or model. The process is then worked through chronologically, and the following questions are clarified for each task:

- Does the task break down into further sub-tasks? If yes, in which ones?
- Are data discovered at this point in the process? If yes, what data is discovered, where is it sent (to which other process task), and where is it stored (at which storage location)?
- Are data used at this point in the process? If yes, what are these data used for, where were they previously discovered (at which process task), and where are they stored (at which storage location)?
- Are the tools used correctly assigned? Are tools missing from the task?
- Are there any additional paths that can occur after this task that have not been depicted so far?

If questions cannot be answered or only partially answered, area experts can also be consulted here. Once the process has been walked through, the first interview ends. The PD processes all the information obtained during the interview and incorporates it into the model.

In the fourth process step, the process is validated through another interview with the process owner. In the second interview, the entire process is presented to the process owner again. If the process owner has no comments, the interview and the first validation unit are completed. However, if the process owner still has suggestions for improvements and corrections, the PD incorporates these into the model during the meeting. This avoids the need for further interview sessions with the process owner. The interview ends once the process owner perceives the model as complete and correct. Completeness and correctness also refer to the capturing of data and machine connectivity.

The second validation unit comprises a workshop with the relevant process participants. As outlined in Sect. 5.4.2, several relevant process participants are identified and invited to the workshop. In this workshop, the process model previously created with the process owner's assistance is reviewed. The workshop follows the structure of the last workshop presented in Sect. 5.4.2. This is to ensure that all parties involved agree on the completeness and correctness of the model, including data collection and connectivity. Upon completion of this workshop, there should be a process model that is highly likely to be correct and complete. Figure 9 shows the combined process discovery guideline in BPMN.

6 Evaluation

This section evaluates the applicability of the proposed guidelines by demonstrating their implementation in two real-world industrial use cases. Both use cases are detailed below and constitute a key component of the evaluation. Furthermore, it presents a comparative study that assesses the efficacy of the artifact by contrasting the process discovery results obtained from novice and expert users.

6.1 Use case to discover a semi-automated process in automotive industry

In the first use case, we analyzed a manufacturing IIoT process within the automotive industry by independently applying a combination of our guideline components. This application served as part of the evaluation, potentially initiating a further phase in the DSR cycle.

The identity of the company remains confidential for competitive reasons; consequently, the resulting process model is published in a redacted format. The production line manufactures an automotive component, involving the collaboration of multiple machines and human actors. The final product consists of 27 individual parts, which are assembled through six high-level, sequential process steps: bending, welding, assembly, molding, testing, and packaging.

The process discovery began with a document analysis. This was followed by a series of interviews and later observations, combined with a workshop dedicated to validation.

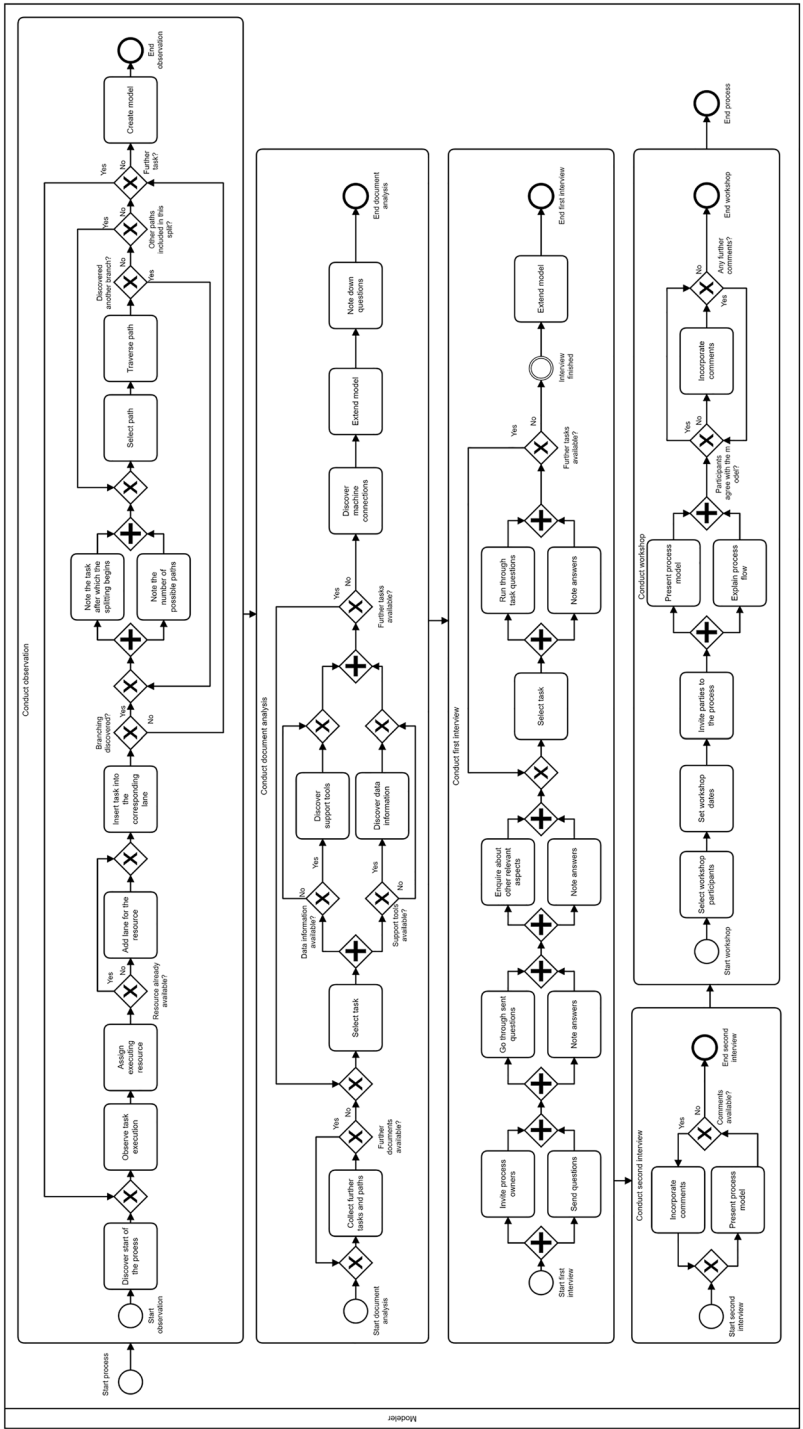


Fig. 9 Combined process discovery guideline, presented in BPMN

Document analysis - preparation

The process discovery started with a document analysis, as documentation was available only to a limited extent. The PD used the document analysis guideline to gain initial insights into the process and formulated interview questions based on the reviewed documents. These questions followed a top-down approach and were initially kept generic, avoiding in-depth process details. Consequently, the questions were derived from the suggestions provided for the pre-interview questionnaire.

Interview - discover resources, tasks and work flow

In the initially closed interview, the PD asked about the resources involved, their tasks, and the overall process flow. The PD then explained their understanding of the process to the process owner and recorded any necessary corrections. Following the interview, the PD constructed an initial process model based on the information provided, noting any open questions or ambiguities that emerged during the modeling phase.

Interview - discover data and tools

In a second interview, the PD discussed the outstanding questions with the process owner, using the preliminary model as a reference. Once these queries were clarified, the investigation focused on identifying the tools and data employed within the process. However, it emerged that neither the process owner nor the participants consulted as area experts could explain the connectivity or data transmission mechanisms. Subsequently, the process owner was asked in an open interview if any other process components remained overlooked. The PD then integrated the information gathered during this second session into the process model.

Interview - validation

The third and final interview served as the primary mechanism for process validation. The PD systematically reviewed the model with the process owner on a task-by-task basis, verifying the accuracy and completeness of the sequence, tasks, tools, data, and resources. The process owner's feedback and corrections were incorporated directly into the model during the session.

Observation and workshop - validation

For validation, the PD conducted an on-site observation of the process. This was combined with a workshop in which employees explained exactly what occurred during the individual process steps observed, allowing the PD to compare these details with the model in parallel. The observation demonstrated that the results of the interviews already corresponded almost entirely to the actual process flow. The chronological order was completely consistent, and all process steps described in the interviews were visible during observation. There were no substantive errors (e.g., incorrectly described tasks). The process steps not recorded during the interview phase amounted to only five tasks in relation to the overall size of the model. Of these, only three were considered relevant to the process.

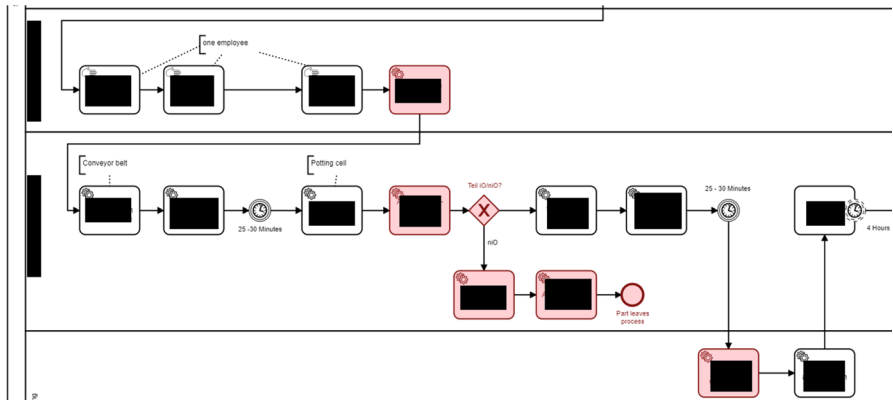


Fig. 10 Red tasks were detected during validation

Table 3 Table to handle process observation information

Task	Number of triggers	Associated resource	Associated tools	Previous task	Subsequent task	Modelling status
01 - Start Commissioning	1	R04	Order list		02	x
02 - Accept order	1	R04	Label	01	03	x
03 - Robot starts	1	R01, R02		02	04a/04b	x

The resulting process model can be viewed in a redacted format in the Fig. 11. The tasks inserted following validation are highlighted in Fig. 10. The constructed model contains no gaps or incomplete paths. Based on the validation results, it can be assumed that the model is both correct and complete. While the current level of detail is enough, the granularity of the process could be further refined.

6.2 Use case to discover a semi-automated warehouse management process

In a second use case, the *Autostore warehouse management* process was discovered at SOMIC, a company specializing in the manufacturing of packaging machines. The process entails warehouse management using mobile robots that retrieve required components from storage bins based on production needs and transport them to the output portal. The robots traverse an aluminum grid construction, which serves as a rail system, to extract individual bins from their stacks as required.

The guidelines were provided to three novices with no prior knowledge of process discovery. Working collaboratively, they discovered the process using the following procedure:

Preparation

The preparation phase began with contacting the company to schedule appointments for the observation and interview. Additionally, requests were made for pro-

cess documentation to facilitate document analysis and for employee participation in the workshop. The interview questionnaire was transmitted in advance. However, due to the unavailability of process documents and time constraints, the document analysis and workshop were not executed.

Observation

The observation was conducted independently of other discovery methods. To achieve this, each process step was observed individually during a forward observation phase and recorded in a tabular format, as shown in Table 3. Consequently, all process perspectives were captured during the observation of each task. Modeling was performed at the same time with the observation. Subsequently, a backward observation was undertaken to validate the observed tasks against the generated model. Fig. 12 in the Appendix depicts the process model of the process following the observation phase.

Interview

To facilitate the interview, a questionnaire was distributed to the process owner in advance; however, it could not be completed due to time constraints. Consequently, the questionnaire served as the basis for discussion during the first interview, where inquiries focused on resources, tasks, control flow, data, and tools. Concurrently, an initial process model was constructed. A second interview was conducted to provide further details and clarify outstanding queries. The resulting model, discovered independently of the prior observation, is presented in Fig. 13 of the Appendix. Finally, the model was transmitted to the process owner via email for validation.

Combination

Finally, the findings from the observation and the interview were synthesized. This integration allowed for the addition of new tasks and a resource to the observation-based model. Overall, this resulted in a more detailed process model, which is depicted in Fig. 14 in the Appendix.

The constructed model is free of gaps or incomplete paths. Due to the further merging of findings, it can be assumed that the model is correct and substantially complete. While the current granularity is adequate, the process could be deepened further. The IIoT perspective was addressed through the integration of a database, which is linked to the respective tasks. The objective is to organize a workshop to potentially finalize the model.

6.3 Comparative evaluation

In the final phase of development, we engaged groups of novice and expert users to discover the same industrial process using the updated guidelines, facilitating a comparative analysis of the resulting information. Due to non-disclosure agreements, we can report only limited details regarding the specific industrial process and the

identities of the subject matter experts. The primary objective of the study design was to ensure strict experimental equivalence between the two groups. To minimize confusing variables, both groups operated under identical experimental conditions, specifically concerning the provided materials, the interview partner, and the workshop participants. The procedure followed the structure outlined in Sect. 5.5. Novices and experts performed these steps independently, with no interaction permitted between groups during the experiment. The procedure was executed as follows:

Observation

The observation phase was iterative and time-constrained. The process simulation was repeated multiple times, allowing both groups to observe the workflow until they self-determined that their discovery was complete.

Document analysis

Both groups received identical reference materials and conducted the analysis independently to prevent cross-group influence.

Interview

Interviews were conducted separately for each group by the same process-responsible employee. To ensure comparability, this employee was briefed on a standardized response protocol, ensuring that experts and novices received consistent information for similar inquiries. This measure aimed to prevent the employee from providing varying levels of detail based on the group's expertise.

Workshop

The validation workshops were held on separate dates. Crucially, the composition of the participating employee group remained constant; the same individuals attended both sessions to ensure consistency in group dynamics and contributions. Like the interviewee, workshop participants were briefed beforehand to mitigate bias.

The results indicate that novices were able to discover information comparable to that of experts in terms of completeness and structure. Observed differences were primarily lexical and granular, reflecting variations in domain vocabulary and experience. Notably, process owners validated all discovered information, confirming its accuracy and utility. These findings substantiate the guidelines' usability across varying levels of expertise while highlighting that domain knowledge serves to enhance the granularity of the results.

In conclusion, the evaluation process was characterized by an iterative and rigorous design that synthesized both theoretical and practical feedback. Each phase facilitated the progressive refinement of the artifact, as detailed in Sect. 3.2 and summarized in Table 1. Ultimately, this process yielded flexible, effective guidelines applicable to real-world IIoT environments, laying a foundation for future generalization beyond this specific domain.

7 Conclusion

The primary objective of this study was to determine how organizations can systematically and reproducibly capture process information when relying on established manual discovery methods, such as document analysis, observation, interviews, or workshops. Although these methods are ubiquitous in both research and practice, the academic literature has lacked operational guidelines specifying their execution.

Empirical evidence from a survey of 28 academic and industrial participants indicates that over 80% of respondents perceive clear guidelines as a driver for higher-quality process models; notably, there was unanimous consensus that such guidelines improve the overall outcomes of manual process discovery. Controlled experiments with both novice and expert modelers corroborate this perception: when applying our guideline, both groups yielded highly comparable results, with divergences limited to nomenclature and minor structural details. Furthermore, while informal, in-house guidelines exist within select companies, a comparable artifact has been conspicuously absent from the scholarly domain.

By presenting a domain-independent, step-by-step guideline for manual and semi-automated process discovery, this work bridges the gap between fragmented practitioner know-how and the lack of a consolidated methodological reference in academic research. The guideline synthesizes theory and practice by *(i)* codifying tacit practitioner knowledge, *(ii)* aligning it with established modeling principles, and *(iii)* validating it through two real-world use cases in the IIoT domain. Although the IIoT provided a rigorous context for validation, the applicability of the guideline extends beyond this specific domain.

For organizations, the guideline provides an actionable framework suitable for integration into quality management or digital transformation initiatives. Standardizing manual discovery methods enables teams to minimize resource expenditure, achieve comprehensive process visibility, and establish shared vocabulary across departments, prerequisites for successful automation, compliance auditing, and continuous improvement.

Despite these contributions, exhaustive coverage cannot be guaranteed, particularly in automated or poorly documented environments. Consequently, future research should focus on *(i)* combining this guideline with process-mining techniques to integrate machine-generated event traces, and *(ii)* employing NLP and computer-vision tools to accelerate the analysis of documents and observations (Hornsteiner et al. 2024; Fichtner et al. 2020; Ackermann et al. 2021; Neuberger et al. 2023). Evaluating the guideline in additional sectors (e.g., healthcare or logistics) and across varying levels of process maturity also represents a promising avenue for future inquiry.

In conclusion, by formalizing and empirically validating a practical guideline for manual process discovery, this study establishes a vital methodological link between academic research and industrial practice. This contribution empowers both communities to capture complex processes with greater speed, consistency, and quality.

Appendix A extended data

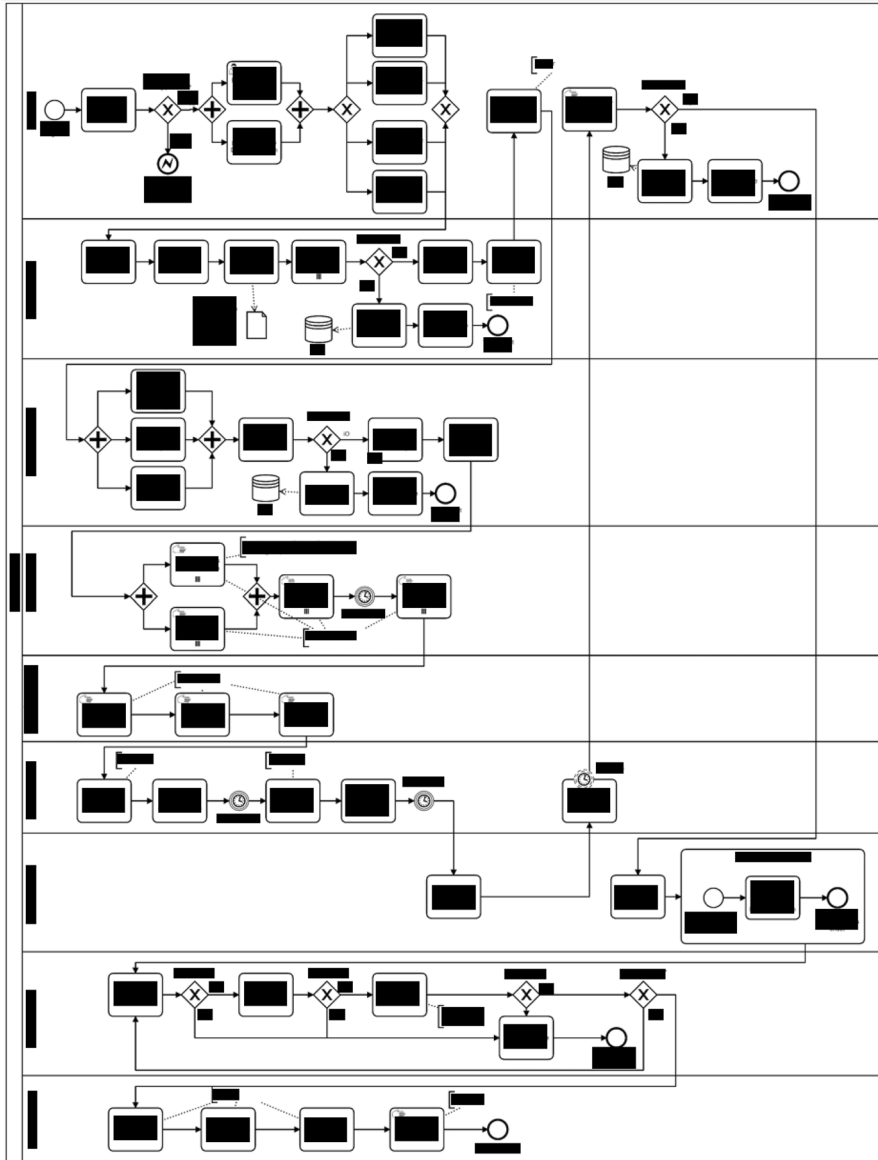


Fig. 11 Blackened industrial process for the manufacture of engine components

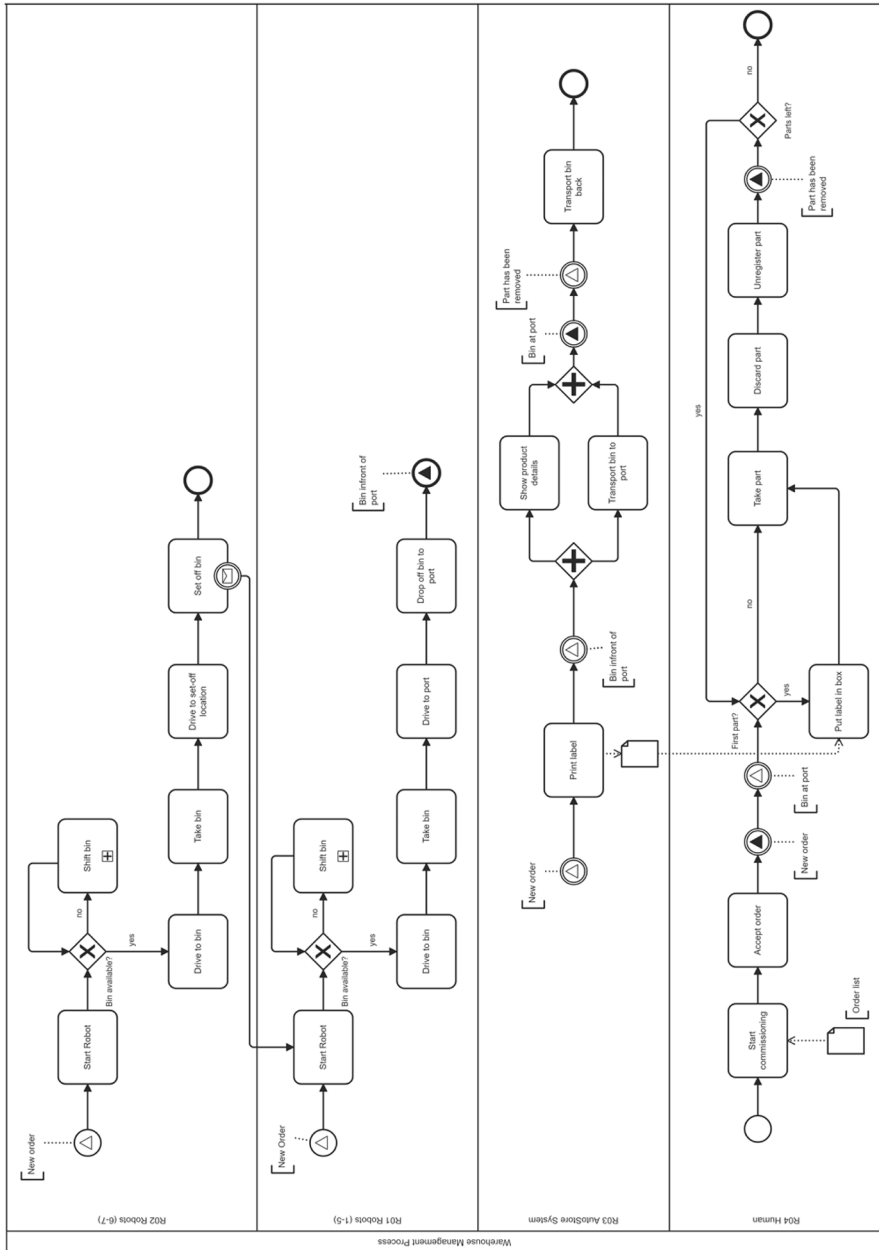
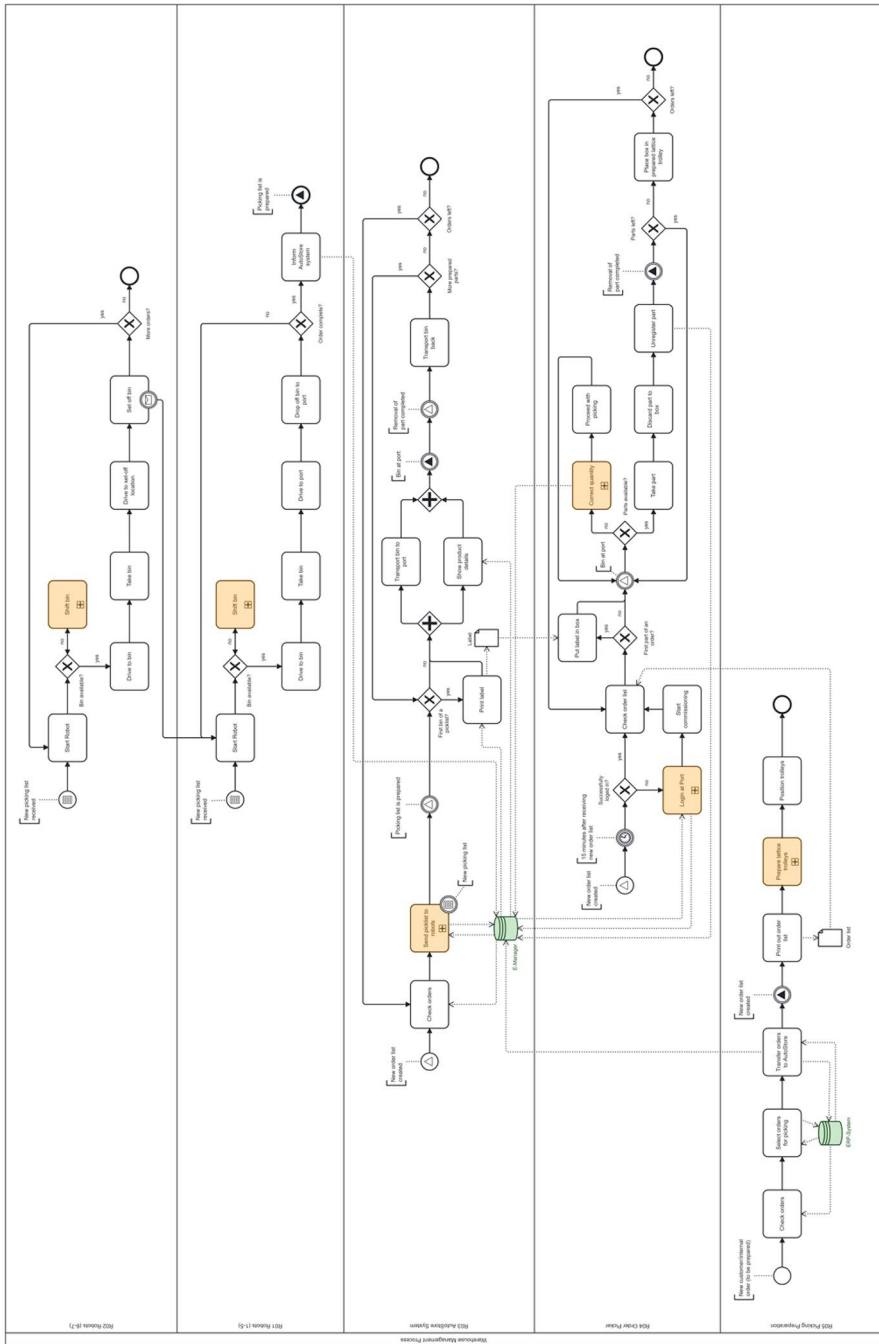


Fig. 12 Model of the autostore process at Somic after observation

Fig. 13 Model of the autostore process at Somic after interview



 Springer

Funding Open Access funding enabled and organized by Projekt DEAL. This work is funded by the "Bavarian Ministry of Economic Affairs, Regional Development and Energy" within the project Security Iot pRocEss Notation (SIREN) (Grant No. DIK-2308-0012//DIK0511).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- van der Aalst WMP (2016) Process Mining - Data Science in Action, Second Edition. Springer. <https://doi.org/10.1007/978-3-662-49851-4>
- van der Aalst WMP (2022) Foundations of Process Discovery. Springer International Publishing, Cham, pp 37–75
- Ackermann L, Neuberger J, Jablonski S (2021) Data-driven annotation of textual process descriptions based on formal meaning representations. In: Advanced Information Systems Engineering - 33rd International Conference, CAiSE 2021, Melbourne, VIC, Australia, June 28 - July 2, 2021, Proceedings, Lecture Notes in Computer Science, vol 12751. Springer, pp 75–90, https://doi.org/10.1007/978-3-030-79382-1_5
- Ackermann L, Käppel M, Marcus L, et al (2024) Recent advances in data-driven business process management. [arxiv:2406.01786](https://arxiv.org/abs/2406.01786), 2406.01786
- Awadid A (2017) Supporting the consistency in multi-perspective business process modeling: A mapping approach. In: 11th International Conference on Research Challenges in Information Science, RCIS 2017, Brighton, United Kingdom, May 10-12, 2017. IEEE, pp 414–419, <https://doi.org/10.1109/RCIS.2017.7956568>
- Bazan P, Estevez E (2022) Industry 4.0 and business process management: state of the art and new challenges. *Bus Process Manag J* 28(1):62–80. <https://doi.org/10.1108/BPMJ-04-2020-0163>
- Becker J, Kugeler M, Rosemann M (2012) Prozessmanagement: Ein Leitfaden zur Prozessorientierten Organisationsgestaltung, 7th edn. Springer Verlag, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-642-33844-1>
- Benevento E, Aloini D, van der Aalst WM (2022) How can interactive process discovery address data quality issues in real business settings? Evidence from a case study in healthcare. *J Biomed Inform* 130:104083
- Bernardo R, Galina SVR, de Pádua SID (2017) The BPM lifecycle: How to incorporate a view external to the organization through dynamic capability. *Bus Process Manag J* 23(1):155–175. <https://doi.org/10.1108/BPMJ-12-2015-0175>
- Boyes H, Hallaq B, Cunningham J et al (2018) The industrial internet of things (iiot): An analysis framework. *Comput Ind* 101:1–12. <https://doi.org/10.1016/J.COMPIND.2018.04.015>
- Bundesverwaltungsamt (2013) Leitfaden für die erhebung von geschäftsprozessen im bundesministerium des innern und seinen nachgeordneten behörden. https://www.bva.bund.de/SharedDocs/Downloads/DE/Behoerden/Beratung/Prozessmanagement/Leitfaeden/Erhebung_Prozesse.pdf?__blob=publicationFile&v=2, Accessed 16 Feb 2024
- Burnay C, Bouraga S, Gillain J et al (2020) What lies behind requirements? a quality assessment of statement grounds in requirements elicitation. *Software Qual J* 28(4):1615–1643
- Castiglione C (2024) Automated generation of digital models for manufacturing systems: The event-centric process mining approach. *Computers & Industrial Engineering* 197:110596

- Christov SC, Marquard JL, Avrunin GS et al (2017) Assessing the effectiveness of five process elicitation methods: A case study of chemotherapy treatment plan review. *Appl Ergon* 59:364–376
- Conklin W (2016) It vs. ot security: A time to consider a change in cia to include resilienc. pp 2642–2647, <https://doi.org/10.1109/HICSS.2016.331>
- D'Hondt T, Wilbik A, Grefen P, et al (2019) Using BPM technology to deploy and manage distributed analytics in collaborative iot-driven business scenarios. In: Proceedings of the 9th International Conference on the Internet of Things, IoT 2019, Bilbao, Spain, October 22–25, 2019. ACM, pp 19:1–19:8, <https://doi.org/10.1145/3365871.3365890>
- DiCicco-Bloom B, Crabtree BF (2006) The qualitative research interview. *Med Educ* 40(4):314–321. <https://doi.org/10.1111/j.1365-2929.2006.02418.x>
- Dolnicar S, Zinn AK, Demeter C (2024) A typology of quantitative approaches to discovery. *Ann Tour Res* 104:103704
- Dumas M, La Rosa M, Mendling J, et al (2021) Grundlagen des Geschäftsprozessmanagements, 1st edn. Springer Verlag, Berlin, Heidelberg, <https://doi.org/10.1007/978-3-662-58736-2>
- Federal Ministry of the Interior and Community (2012) 6.1.1 dokumentenanalyse. https://www.orghandbuch.de/OHB/DE/Organisationshandbuch/6_MethodenTechniken/61_Erhebungstechniken/611_Dokumentenanalyse/dokumentenanalyse_inhalt, Accessed 16 Feb 2024
- Fichtner M, Schöning S, Jablonski S (2020) Using image mining techniques from a business process perspective. In: Enterprise Information Systems - 22nd International Conference, ICEIS 2020, Virtual Event, May 5–7, 2020, Revised Selected Papers, Lecture Notes in Business Information Processing, vol 417. Springer, pp 62–83, https://doi.org/10.1007/978-3-030-75418-1_4
- Front A, Rieu D, Santorum M et al (2017) A participative end-user method for multi-perspective business process elicitation and improvement. *Software & Systems Modeling* 16(3):691–714
- Ghose A, Koliadis G, Chueng A (2007) Rapid business process discovery (R- BPD). In: Conceptual Modeling - ER 2007, 26th International Conference on Conceptual Modeling, Auckland, New Zealand, November 5–9, 2007, Proceedings, Lecture Notes in Computer Science, vol 4801. Springer, pp 391–406, https://doi.org/10.1007/978-3-540-75563-0_27
- Gilchrist A (2016). Springer, (Industry 4.0.)
- Giudice MD (2016) Discovering the internet of things (iot) within the business process management: A literature review on technological revitalization. *Bus Process Manag J* 22(2):263–270. <https://doi.org/10.1108/BPMJ-12-2015-0173>
- Grefen P, Ludwig H, Tata S, et al (2018) Complex collaborative physical process management: A position on the trinity of bpm, iot and DA. In: Collaborative Networks of Cognitive Systems - 19th IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2018, Cardiff, UK, September 17–19, 2018, Proceedings, IFIP Advances in Information and Communication Technology, vol 534. Springer, pp 244–253, https://doi.org/10.1007/978-3-319-99127-6_21
- Gronau N (2017) Geschäftsprozessmanagement in Wirtschaft und Verwaltung: Analyse, Modellierung und Konzeption, 2nd edn. GITO mbH Verlag, Berlin
- Gubrium JF, Holstein JA, Marvasti AB (2012) The SAGE Handbook of Interview Research: The Complexity of the Craft, 2nd edn. Sage Publications, Thousand Oaks, CA. <https://doi.org/10.4135/9781452218403>
- Han X, Hu L, Mei L, et al (2020) A-BPS: automatic business process discovery service using ordered neurons LSTM. In: 2020 IEEE International Conference on Web Services, ICWS 2020, Beijing, China, October 19–23, 2020. IEEE, pp 428–432, <https://doi.org/10.1109/ICWS49710.2020.00063>
- Hansen H, Mendling J, Neumann G (2019) Wirtschaftsinformatik, 12th edn. De Gruyter Oldenbourg, Berlin. <https://doi.org/10.1515/9783110608731>
- Harman J, Brown R, Johnson D et al (2016) Augmenting process elicitation with visual priming: An empirical exploration of user behaviour and modelling outcomes. *Inf Syst* 62:242–255
- Hevner AR, March ST, Park J et al (2004) Design science in information systems research. *MIS Q* 28(1):75–105. <https://doi.org/10.2307/25148625>
- Hewelt M, Pufahl L, Mandal S et al (2020) Toward a methodology for case modeling. *Softw Syst Model* 19(6):1367–1393
- Hochstetter-Diez J, Negrier-Seguel M, Navarro-Cruces C (2025) Challenges and opportunities in business process elicitation. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2025.3590295>
- Hornsteiner M, Empl P, Bunghardt T et al (2024) Reading between the lines: process mining on OPC UA network data. *Sensors (Basel)*. <https://doi.org/10.3390/s24144497>

- Houy C, Fettke P, Loos P, et al (2010) Bpm-in-the-large - towards a higher level of abstraction in business process management. In: E-Government, E-Services and Global Processes - Joint IFIP TC 8 and TC 6 International Conferences, EGES 2010 and GISP 2010, Held as Part of WCC 2010, Brisbane, Australia, September 20-23, 2010. Proceedings, IFIP Advances in Information and Communication Technology, vol 334. Springer, pp 233–244, https://doi.org/10.1007/978-3-642-15346-4_19
- Jablonski S, Bussler C (1996) Workflow Management: Modeling Concepts, Architecture and Implementation, 1st edn. Cengage Learning Verlag, Holborn
- Jablonski S, Götz M (2007) Perspective oriented business process visualization. In: Business Process Management Workshops, BPM 2007 International Workshops, BPI, BPD, CBP, ProHealth, RefMod, semantics4ws, Brisbane, Australia, September 24, 2007, Revised Selected Papers, Lecture Notes in Computer Science, vol 4928. Springer, pp 144–155, https://doi.org/10.1007/978-3-540-78238-4_16
- Jadhav S (2011) Business process discovery. Tech. rep, BP Trends (February) Citeseer
- Jakobs EM, Spanke J (2011) Sprache als erfolgsmotor industrieller prozessmodellierung. Evolution der Informationsgesellschaft: Markenkommunikation Im Spannungsfeld der Neuen Medien. VS Verlag für Sozialwissenschaften, Wiesbaden, pp 181–195. https://doi.org/10.1007/978-3-531-92860-9_12
- Janiesch C, Koschmider A, Mecella M et al (2020) The internet of things meets business process management: A manifesto. IEEE Systems Man and Cybernetics Magazine 6(4):34–44. <https://doi.org/10.1109/MSMC.2020.3003135>
- Jolles RL (2017) How to Run Seminars and Workshops: Presentation Skills for Consultants, Trainers, Teachers, and Salespeople, 4th edn. John Wiley & Sons, Hoboken, NJ
- Koschmider A, Aleknytytė-Resch M, Fonger F (2023) Process mining for unstructured data: challenges and research directions. arXiv:2401.13677 arXiv preprint
- Krallmann H, Frank H, Gronau N (2002) Systemanalyse Im Unternehmen: Vorgehensmodelle, Modellierungsverfahren und Gestaltungsoptionen, 4th edn. Oldenbourg Verlag, München
- Langer M, Söffker D (2011) Human guidance and supervision of a manufacturing system for semi-automated production. In: 2011 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), pp 1–6, <https://doi.org/10.1109/AEECT.2011.6132501>
- Laugwitz B, Held T, Schrepp M (2008) Construction and evaluation of a user experience questionnaire. In: HCI and Usability for Education and Work, 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings, Lecture Notes in Computer Science, vol 5298. Springer, pp 63–76, https://doi.org/10.1007/978-3-540-89350-9_6
- Lauttamäki V (2014) Practical guide for facilitating a futures workshop. Finland futures research centre, pp 2–11
- Linds W, Gee T (2023) What Is Workshop? Springer Nature Singapore, Singapore. https://doi.org/10.1007/978-981-99-2291-8_3
- Martin N, De Weerd J, Fernández-Llatas C et al (2020) Recommendations for enhancing the usability and understandability of process mining in healthcare. Artif Intell Med 109:101962
- Martin N, Fischer DA, Kerpedzhiev GD et al (2021) Opportunities and challenges for process mining in organizations: results of a delphi study. Business & Information Systems Engineering 63(5):511–527
- Mass J, Chang C, Srirama SN (2016) Wiseware: A device-to-device-based business process management system for industrial internet of things. In: 2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Chengdu, China, December 15-18, 2016. IEEE, pp 269–275, <https://doi.org/10.1109/ITHINGS-GREENCOM-CPSCOM-SMARTDATA.2016.69>
- zur Muehlen M, Ho DT (2005) Risk management in the BPM lifecycle. In: Business Process Management Workshops, BPM 2005 International Workshops, BPI, BPD, ENEI, BPRM, WSCOBPM, BPS, Nancy, France, September 5, 2005, Revised Selected Papers, pp 454–466
- Neuberger J, Ackermann L, Jablonski S (2023) Beyond rule-based named entity recognition and relation extraction for process model generation from natural language text. In: Cooperative Information Systems - 29th International Conference, CoopIS 2023, Groningen, The Netherlands, October 30 - November 3, 2023, Proceedings, Lecture Notes in Computer Science, vol 14353. Springer, pp 179–197, https://doi.org/10.1007/978-3-031-46846-9_10
- Okoli C, Schabram K (2010) A guide to conducting a systematic literature review of information systems research. SSRN Electron J 10. <https://doi.org/10.2139/ssrn.1954824>

- Page MJ, McKenzie JE, Bossuyt PM et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. <https://doi.org/10.1136/bmj.n71>
- Pavelin K, Pundir S, Cham JA (2014) Ten simple rules for running interactive workshops. *PLoS Comput Biol*. <https://doi.org/10.1371/JOURNAL.PCBI.1003485>
- Prat N, Comyn-Wattiau I, Akoka J (2015) A taxonomy of evaluation methods for information systems artifacts. *J Manag Inf Syst* 32(3):229–267. <https://doi.org/10.1080/07421222.2015.1099390>
- Raptis TP, Passarella A, Conti M (2019) Data management in industry 4.0: State of the art and open challenges. *IEEE Access* 7:97052–97093
- Rehse JR, Leemans SJ, Fettke P et al (2024) On process discovery experimentation: addressing the need for research methodology in process discovery. *ACM Transactions on Software Engineering and Methodology* 34(1):1–29
- Richerzhagen B (2015) Königsdisziplin Prozess-Workshops. Camunda Inc,
- dos Santos GC, Meincheim A, Faria Junior ER et al (2019) Process mining techniques and applications – a systematic mapping study. *Expert Syst Appl* 133:260–295
- Schönig S, Ackermann L, Jablonski S (2018) Internet of things meets BPM: A conceptual integration framework. In: *Proceedings of 8th International Conference on Simulation and Modeling Methodologies, Technologies and Applications, SIMULTECH 2018*, Porto, Portugal, July 29–31, 2018. SciTePress, pp 307–314, <https://doi.org/10.5220/0006824803070314>
- Schönig S, Ackermann L, Jablonski S et al (2020) Iot meets BPM: a bidirectional communication architecture for iot-aware process execution. *Softw Syst Model* 19(6):1443–1459. <https://doi.org/10.1007/S10270-020-00785-7>
- Schönig S, Jasinski R, Ermer A (2020b) Data interaction for iot-aware wearable process management. In: *Service-Oriented Computing - ICSOC 2020 Workshops - AIOps, CFTIC, STRAPS, AI-PA, AI-IOTS, and Satellite Events*, Dubai, United Arab Emirates, December 14–17, 2020, *Proceedings, Lecture Notes in Computer Science*, vol 12632. Springer, pp 67–71, https://doi.org/10.1007/978-3-030-76352-7_10
- Schönig S, Günther C, Jablonski S (2012) Process discovery and guidance applications of manually generated logs. In: *Proceedings of the 7th International Conference on Internet Monitoring and Protection (ICIMP 2012)*. IARIA, Stuttgart, Germany, pp 61–67
- Schönig S, Ackermann L, Jablonski S et al (2020) Iot meets bpm: a bidirectional communication architecture for iot-aware process execution. *Softw Syst Model* 19(6):1443–1459
- Seiger R, Malburg L, Weber B et al (2022) Integrating process management and event processing in smart factories: A systems architecture and use cases. *J Manuf Syst* 63:575–592. <https://doi.org/10.1016/j.jmsy.2022.05.012>
- Sirgo G, Esteban F, Gómez J et al (2018) Validation of the icu-dama tool for automatically extracting variables for minimum dataset and quality indicators: The importance of data quality assessment. *Int J Med Informatics* 112:166–172
- Sisinni E, Saifullah A, Han S et al (2018) Industrial internet of things: Challenges, opportunities, and directions. *IEEE Trans Ind Informatics* 14(11):4724–4734. <https://doi.org/10.1109/TII.2018.2852491>
- Stoiber C, Schönig S (2024) Leveraging the industrial internet of things for business process improvement: a metamodel and patterns. *IseB* 22(2):285–313
- Storvang P, Mortensen B, Clarke AH (2018) Using workshops in business research: A framework to diagnose, plan, facilitate and analyze workshops. *Collaborative Research Design* pp 155–174. https://doi.org/10.1007/978-981-10-5008-4_7
- Szelagowski M, Lupeikiene A, Berniak-Woźny J (2022) Drivers and evolution paths of bpmns: State-of-the-art and future research directions. *Informatica* 33(2):399–420
- Taherdoost H (2022) How to conduct an effective interview; a guide to interview design in research study. *International Journal of Academic Research in Management* 11(1):39–51
- Tiftik MN, Erdogan TG, Tarhan AK (2022) A framework for multi-perspective process mining into a bpmn process model. *Math Biosci Eng* 19(11):11800–11820. <https://doi.org/10.3934/mbe.2022550>
- Tsai WH, Lee KH, Yang CF (2025) Design and implementation of a semi-automated warehouse system for smart logistics in urban infrastructure. *Sens Mater*. <https://doi.org/10.18494/sam5896>
- Turner DW III (2010) Qualitative interview design: a practical guide for novice investigators. *Qual Rep* 15(3):754. <https://doi.org/10.46743/2160-3715/2010.1178>

- Wang J, Zhang W, Shi Y, et al (2018) Industrial big data analytics: Challenges, methodologies, and applications. CoRR <https://doi.org/10.48550/arXiv.1807.01016>
- Wang W, Liu L (2025) Advances in the application of human-machine collaboration in healthcare: insights from china. *Front Public Health* 13:1507142
- Weske M (2012) *Business Process Management - Concepts, Languages, Architectures*, 2nd Edition. Springer. <https://doi.org/10.1007/978-3-642-28616-2>
- Zhang Y, Zhang Y (2024) A review of process discovery methods and conformance checking methods. *EAI*

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.