Pradhan *et al. Process Science* (2025) 2:25
https://doi.org/10.1007/s44311-025-00032-6

Process Science

RESEARCH Open Access

Process mining starts here: on expertise, exchange, and the gaps between

Shameer K. Pradhan[1*], Mieke Jans[1,2] and Niels Martin[1]

*Correspondence:
Shameer K. Pradhan
shameer.pradhan@uhasselt.be
[1]Digital Future Lab, Hasselt University, Wetenschapspark 2, Diepenbeek 3590, Belgium
[2]School of Business and Economics, Maastricht University, Minderbroedersberg 4-6, Maastricht
6211 LK, The Netherlands

Abstract
In the pre-analysis stage of process mining projects, experts from diverse backgrounds—such as process owners, information system specialists, and process analysts—contribute knowledge shaped by their educational and professional experience. This stage is inherently collaborative, relying not only on individual expertise but also on effective knowledge exchange among stakeholders. Yet, despite its critical role, this exchange has received limited scholarly attention. Drawing on a qualitative, interview-based study, we investigate the key knowledge units required during pre-analysis. We also examine the mechanisms through which knowledge is exchanged, as well as the challenges that impede this exchange. Our findings identify seven categories of knowledge units, including business knowledge, process understanding, and impact awareness. Additionally, we uncover four types of knowledge exchange mechanisms and twelve challenges grouped into four overarching categories. These insights offer practical guidance for improving collaboration, training, and knowledge management in process mining initiatives, with a particular emphasis on the pre-analysis stage.

Keywords Data preparation, Expert knowledge, Knowledge exchange, Process mining

Introduction

Effective *knowledge management* (KM) is essential in complex, multidisciplinary initiatives (Randeree 2006; Edvardsson and Durst 2013). KM refers to the processes and mechanisms through which experts organize, share, and apply knowledge to achieve common goals in complex settings (Galagan 1997; Davenport et al. 1998; Rowley 1999). It helps align objectives, define requirements, and achieve desired outcomes in a project. However, differences in terminologies and disciplinary perspectives regarding key concepts, goals, and methodologies frequently create communication barriers, leading to misunderstandings, errors, and inefficiencies (Brock et al. 2024). Such challenges are particularly pronounced in domains requiring specialized expertise and collaboration, such as process mining (Brock et al. 2024).

Process mining is a powerful suite of techniques designed to analyze and improve organizational processes by systematically using data stored in various information systems (IS), such as Enterprise Resource Planning (ERP) systems (Aalst 2022). However, process mining algorithms cannot directly use raw data from these systems. Instead, the data must first undergo a series of preparatory tasks, collectively known as the *pre-analysis stage* (Pradhan et al. 2025). This stage involves data identification, extraction, and event log-building that transforms raw data into structured event logs suitable for process mining. These preparatory steps are not only highly labor-intensive, consuming up to 80% of the total time and effort in process mining initiatives (DeWeerdt and Wynn 2022), but also demand significant manual intervention (Stein Dani et al. 2022).

The complexity of the pre-analysis stage necessitates collaboration among experts from diverse knowledge domains. The multidisciplinary nature of process mining (Eck et al. 2015; Al-Dowail and Al-Hashedi 2021) introduces challenges for KM, particularly in facilitating effective knowledge exchange. Experts often struggle with formulating appropriate project questions, understanding source systems and data structures, transforming data into event logs, and navigating specific process domains (Zimmermann et al. 2023). These challenges point to a broader gap in the literature: although the importance of expert collaboration is well acknowledged, systematic insights into the specific knowledge units (i.e., types of expert knowledge) required during the pre-analysis stage of process mining remain scarce. Against this backdrop, the goal of this research is to systematically identify the expert knowledge units required in the pre-analysis stage—particularly those related to process context, contextual understanding, and stakeholder interaction—as well as associated aspects such as knowledge exchange mechanisms and challenges.

This paper presents the findings of a qualitative interview-based study conducted at five companies to explore the expert knowledge dynamics within the pre-analysis stage of process mining. Through 15 in-depth interviews, we identified seven broad expert knowledge units. We also identified four mechanisms of knowledge exchange in this stage. Additionally, our analysis uncovered twelve challenges in knowledge exchange, which we categorized into four overarching themes. The insights gained through this study provide a nuanced understanding of the interplay between expert knowledge, management practices, and collaborative processes during the pre-analysis stage of process mining. Building on this foundation, the study aims to systematically analyze expert knowledge and related aspects such as knowledge exchange mechanisms and challenges, which could support more effective decisions related to data selection, event log-building, and project strategy formulation in this critical early phase.

The paper is structured as follows. Section "Background and Research Questions" provides background by introducing the experts involved in process mining initiatives, outlining their roles and the knowledge-related concepts that underpin their work. It also defines the research objectives. Section "Research Methodology" details the methodological approach used to investigate expert knowledge in process mining pre-analysis. Section "Results" presents our findings, categorizing the units of knowledge required during the pre-analysis stage of a process mining project and identifying mechanisms that facilitate knowledge exchange, along with challenges that hinder effective KM. Section "Discussion" contextualizes these findings and discusses their implications, while acknowledging the limitations. Finally, Section "Conclusion" concludes the paper by summarizing our contributions and proposing future research directions.

Background and research questions

Knowledge is critical to organizational success, serving as a foundation for effective decision-making, innovation, and operational efficiency. It has been defined in numerous ways, with Zagzebski (2017) describing knowledge as a form of *"cognitive contact with reality arising out of acts of intellectual virtue"*. Another prominent definition posits knowledge as *"justified true belief"* (Nonaka and Takeuchi 1995). Metaphors such as knowledge being fluid and dynamic, circulating from its origin to areas of application, further illustrate its transformative potential within organizations (Nissen 2005; Bolisani and Bratianu 2018).

Process mining initiatives rely on collaboration among diverse experts, each contributing domain-specific and technical knowledge to analyze and improve business processes (Suriadi et al. 2013). The PM^2 methodology highlights the importance of forming a cross-functional project team early in the planning phase of a process mining initiative (Eck et al. 2015). Central to this collaboration are *process owners* (i.e., business experts), who possess deep insights into actual process execution, and *process analysts*, who apply process mining techniques to uncover actionable insights. Supporting them are *information system (IS) experts*, whose understanding of IT infrastructure is critical for identifying data sources and resolving data quality issues. Effective collaboration across these roles ensures that analyses are technically valid and contextually relevant.

The importance of expert involvement is echoed in success factor models for process mining, such as that of Mamudu et al. (2022), which highlight the interplay between management, external stakeholders, subject matter experts, and user groups. Yet while expert involvement is widely acknowledged, how expert knowledge is exchanged and applied—particularly during the early, pre-analysis stage—remains poorly understood. A multidisciplinary team ensures that a broad range of expertise is available to support data selection, event log creation, and process analysis (Suriadi et al. 2013; Mamudu et al. 2022; Brocke et al. 2021).

The pre-analysis stage involves crucial decisions, including identifying relevant systems, selecting appropriate data sources, building event logs, and aligning analysis goals (Pradhan et al. 2025). These decisions often depend on knowledge that is tacit, distributed across roles, and highly contextual (Brock et al. 2024). For example, process owners may be aware of undocumented process variants or workarounds, while IS experts may know which data fields are unreliable or inconsistently logged. Process analysts must gather, interpret, and apply this fragmented knowledge to construct event logs that faithfully reflect the process under investigation. This process involves activities such as knowledge acquisition, creation, application, storage, and transfer (Shongwe 2016).

In complex and dynamic settings like process mining pre-analysis, knowledge exchange is not a simple transfer of information. It is a two-way, collaborative process in which both the sender and the recipient shape the effectiveness of communication (Basque et al. 2004; Deken et al. 2012; Goodbrand et al. 2021). To conceptualize why such collaboration can be difficult, this study draws on *knowledge boundary theory* (Carlile 2004). This perspective explains that when experts from different domains collaborate, they encounter distinct types of boundaries—syntactic, semantic, and pragmatic—that arise from differences in meaning, context, and interests. These boundaries require varying levels of knowledge management effort, from simple information transfer to translation and, in more complex situations, transformation of knowledge. This theoretical view is relevant to process mining pre-analysis, where business, IS, and analytical experts must integrate highly specialized knowledge to build valid event logs and align analysis objectives.

Nevertheless, knowledge exchange faces several challenges. While traditional models viewed knowledge transfer as a straightforward exchange of information (Shannon 1948), more recent perspectives recognize it as a complex process involving social interaction, mutual interpretation, and sense-making (Oliver 2001). Barriers may arise at the individual level—such as reluctance to share knowledge due to fears of losing power or concerns about effort—and at the social level, including trust deficits and unclear recognition of expertise (Disterer 2001). Moreover, contextual factors like the relationship between sender and receiver, the characteristics of the exchange setting, and broader organizational dynamics can further complicate the process (Eppler 2007).

Although KM has been extensively studied, its specific application to process mining remains underexplored (Brock et al. 2024; Dixit et al. 2017; Eichele et al. 2023). In particular, there is a lack of research on how knowledge is structured, acquired, and shared during process mining initiatives—especially in the pre-analysis stage. This stage is pivotal, as it brings together diverse expert insights that influence the validity of event logs and the effectiveness of subsequent analyses (Pradhan et al. 2025). A better understanding of how knowledge is managed during this phase can help identify knowledge gaps, reduce miscommunication, and ensure that critical expertise is appropriately integrated. These improvements can, in turn, enhance collaboration, data selection, and analytical outcomes—key success factors in process mining initiatives.

To address this gap, we drew on prior work in the KM domain. Specifically, we reviewed eight studies (Contandriopoulos et al. 2010; Battistella et al. 2016; Prihodova et al. 2019; Van Eerd 2019; Fazey et al. 2013; Mitton et al. 2007; Nidhra et al. 2013; Wit-de Vries et al. 2019) that examined knowledge-related factors in organizational and IS contexts. From this body of work, we synthesized the dimensions most consistently highlighted across studies and most relevant to the pre-analysis stage of process mining. On this basis, we selected three core dimensions that were both central in the literature and feasible to examine systematically within the scope of this study. These three dimensions provide the conceptual basis for our research questions. The three core dimensions are as follows: *i)* **Knowledge Units** (the specific content that must be exchanged (Contandriopoulos et al. 2010; Battistella et al. 2016; Prihodova et al. 2019; Van Eerd 2019)), *ii)* **Mechanisms of Knowledge Exchange** (the methods, tools, and processes used for exchanging knowledge (Battistella et al. 2016; Prihodova et al. 2019; Van Eerd 2019; Fazey et al. 2013)), *iii)* **Callenges in Knowledge Exchange** (the barriers that hinder effective knowledge exchange (Mitton et al. 2007; Nidhra et al. 2013; Wit-de Vries et al. 2019)). These dimensions guide the formulation of the following research questions:

RQ1: What knowledge do experts require and utilize during the process mining pre-analysis stage?

RQ2: How do experts share knowledge during the process mining pre-analysis stage?

RQ3: Which challenges affect knowledge exchange during the process mining pre-analysis stage?

Research methodology

The objective of this study is to explore the knowledge required by experts in the pre-analysis stage of process mining, the mechanisms for knowledge exchange, and the challenges they encounter. We employed a qualitative, interview-based research design, using semi-structured interviews, which ensures conceptual consistency while allowing open-ended exploration (Adams 2015).

As part of the design phase, we identified relevant expert profiles based on a case study review and the outcomes of an international brainstorming seminar. These preparatory steps provided a deductive frame of reference that guided the sampling of interviewees. In contrast, the analysis of the interview data followed the Gioia approach in an inductive manner, allowing first-order concepts, second-order themes, and aggregate dimensions to emerge directly from the transcripts. This mixed strategy reflects a deliberate methodological choice: deductive elements ensured focus and alignment with prior work, while inductive coding captured novel insights from practitioners. The subsequent sections detail the specific analytic steps undertaken.

Identifying the expert profiles

Process mining initiatives rely on experts with diverse skill sets who collaborate within varying organizational settings (Eck et al. 2015; Suriadi et al. 2013). To examine expert involvement during the pre-analysis stage, we reviewed case studies published by the *Task Force on Process Mining (TF-PM)*[1] and those presented in a textbook (Reinkemeyer 2020). This was complemented by insights from an academic seminar[2] hosted by the *Scientific Research Community on Process Mining*, which two authors attended. The seminar focused on the challenges of integrating domain expertise into process mining efforts. These sources collectively informed our identification of key expert profiles and team structures, with particular attention to success factors such as team configuration (Mamudu et al. 2022).

As detailed in Sect. 4.1, the analysis identified three key expert profiles involved in the pre-analysis phase: *process owners*, *IS experts*, and *process analysts*.

Interview guidelines and participants

To maintain consistency across interviews, we developed an interview guideline based on Taherdoost (2022). The guideline included a research preamble, interview objectives, and a standardized form (Appendix A). We also conducted two pilot interviews, which helped refine the interview questions and eliminate unnecessary ones. During the study, the questions served as guiding prompts rather than a fixed sequence. Appendix B provides the questions used in the interviews.

Prior to data collection, institutional ethics approval was obtained. The submission to receive the approval included protocols, outreach content, and supporting materials, all prepared according to the institutional template.

Data was collected from five companies operating in different industries (see Table 1). Companies A, B, and E maintained in-house process mining teams, while Companies C and D offered process mining as part of broader consulting services.

[1] https://www.tf-pm.org/resources/casestudy (https://perma.cc/8SZW-Z9BF)

[2] https://srcprocessmining.com/activities/brainstorm-seminars/brainstorm-seminar-2-the-domain-expert-in-the-loop/ https://perma.cc/SW2B-2CMG)

Table 1 List of companies and their industries

Company	Industry	Country
A	Medical Systems	Japan
B	Banking	The Netherlands
C	Consulting	The Netherlands, Germany
D	Consulting	Belgium
E	Food and Beverages	The Netherlands

Table 2 List of interviewees

Company	Department	Role	Expert Domain			Experience
			Process Owner	IS Expert	Process Analyst	
A	Operational Excellence	Transformation Manager	✓	✓	✓	4 years
A	Performance Analytics & Control	EMEA Chief of Department	✓			1.5 years
B	Process Management	Process Manager	✓		✓	1 year
B	Financial & Economic Crime	Analytics Translator	✓	✓	✓	1 year
B	IT	Process Mining Analyst	✓	✓	✓	7 years
B	Special Asset Management	Business Analyst	✓	✓		3 years
B	Chief Data Analytics Office	Product Owner		✓	✓	16 years
C	Consulting	Data Scientist			✓	3 years
C	Financial, Accountancy, & Advisory Services	Senior Consultant	✓		✓	2 years
C	Financial, Accountancy, & Advisory Services	Senior Manager	✓		✓	7 years
C	Financial, Accountancy, & Advisory Services	Senior Consultant	✓	✓	✓	3 years
D	Risk, Control & Compliance	Manager	✓	✓	✓	6 years
D	Risk, Control & Compliance	Senior Associate		✓	✓	3 years
E	Internal Audit	Manager - IT Audit	✓		✓	3 years
E	Internal Audit	Manager - Data Analytics	✓	✓		3 years

Participants were recruited using purposive sampling (Suri 2011), ensuring that each held at least one expert profile identified in the case study review and academic seminar (see Section "Identifying the Expert Profiles"). All participants were briefed on the study's objectives, informed about audio recording, and asked to sign an informed consent form. For those who declined recording, detailed notes were taken by the researchers. Interviews were conducted via Google Meet with automatic transcription, which was subsequently verified and corrected for any inaccuracies manually.

Out of the 15 interviews, four participants declined audio recording. For these interviews, detailed notes were taken during the sessions and immediately expanded and organized after each interview. These notes captured key points and contextual observations. They were subsequently coded line-by-line using the same coding framework as the recorded and transcribed interviews. Although direct quotations could not be extracted from these sessions, their content contributed to the same codes and themes, ensuring that their perspectives were represented in the thematic analysis.

Table 2 offers an overview of the interviewees who contributed to the study. A single individual may encompass multiple expert profiles. Specifically, five interviewees identified themselves as belonging to all three expert profiles. Additionally, four interviewees identified with both the process owner and process analyst profiles. Two interviewees each reported expertise in both business process and IS, as well as IS and process mining. Furthermore, one interviewee each identified as possessing a process analyst profile and a process owner profile.

Regarding the highest academic degree attained by the interviewees, all participants possess a master's degree. Their professional experience spans 1 to 16 years, with an average of 4.1 years, a median of 3 years, and a standard deviation of 3.9 years.

Coding

Thematic coding is applied on the interview transcripts (Gibbs 2007), which is a widely recognized qualitative research method that identifies recurring patterns and key concepts within textual data. To ensure a grounded and unbiased analysis, an inductive open coding approach was used (Chandra and Shang 2019), applied separately for each research question. While codes were derived directly from participant responses without relying on a predefined codebook, the research questions served as a broad lens for structuring the analysis. This process allowed for a flexible yet focused exploration of themes emerging from participant responses.

Initially, logical groups of sentences were coded. During this phase, similar or overlapping codes were merged to refine the organization of concepts, reducing redundancy while preserving the richness of the data. We used *Atlas.TI*, a qualitative data analysis software, to manage and categorize the codes.

Following this, we applied the Gioia methodology (Gioia et al. 2013), a structured approach that enhances the rigor of qualitative research by systematically consolidating first-order codes into higher-order groups. Specifically, we used the steps proposed by Braun and Clarke (2006), which includes (i) *familiarizing with the data*, (ii) *generating initial codes*, (iii) *searching for themes*, (iv) *reviewing themes*, and (v) *defining and naming themes*. This method ensures that emergent groups reflect a structured relationship among codes, forming a coherent hierarchy of categories and subcategories (Noble and Mitchell 2016). By employing this approach, we were able to move beyond surface-level patterns and uncover more profound insights.

Appendix C provides the coding data structures for the three main topics of study—knowledge units, knowledge exchange mechanisms, and knowledge exchange challenges.

Results

This section presents the findings of our study. Section "Expert Profiles" presents the expert profiles identified through case study review and brainstorming seminar. Section "Knowledge Units" defines essential knowledge required, addressing RQ1, while Section "Mechanisms of Knowledge Exchange" examines KM mechanisms, addressing RQ2. Section "Challenges in Knowledge Exchange" highlights key challenges (RQ3).

Expert profiles

The analysis of case studies and seminar insights helped identify three key expert profiles in the pre-analysis phase of process mining initiatives: *process owners*, *IS experts*, and *process analysts*. Each role contributes unique expertise to ensure effective data preparation and process understanding:

- **Process owners** offer contextualized process knowledge and help interpret process mining results within the domain.
- **IS experts** support the identification, extraction, and exchange of relevant event data.
- **Process analysts** apply process mining techniques to uncover inefficiencies and improvement opportunities.

These roles are not always strictly separated; in several cases, a single individual fulfilled multiple responsibilities, such as a process owner also acting as a process analyst.

Moreover, interview insights revealed that additional stakeholders—such as product managers, key users, and super users—although labeled differently across organizations, ultimately mapped onto the same three core expert profiles. This suggests that while role titles may vary, the underlying responsibilities consistently align with the three roles we distinguish.

The profiles were consistently mentioned across reviewed case studies. In the 39 TF-PM case studies published by November 2023, 34 referenced expert involvement. Process owners and analysts were cited in 24 and 26 cases, respectively, whereas IS experts appeared in only two. In the case studies in Reinkemeyer's textbook (Reinkemeyer 2020), process owners were mentioned in nine out of twelve, while six involved process analysts.

Knowledge units

This section outlines the units of knowledge required during the pre-analysis stage of process mining, thereby addressing RQ1. We categorize this knowledge into seven groups, detailed in Sections "Best Practices" to "System Knowledge", with Fig. 1 and Table 3 summarizing the required knowledge units. Additionally, Fig. 2 presents the

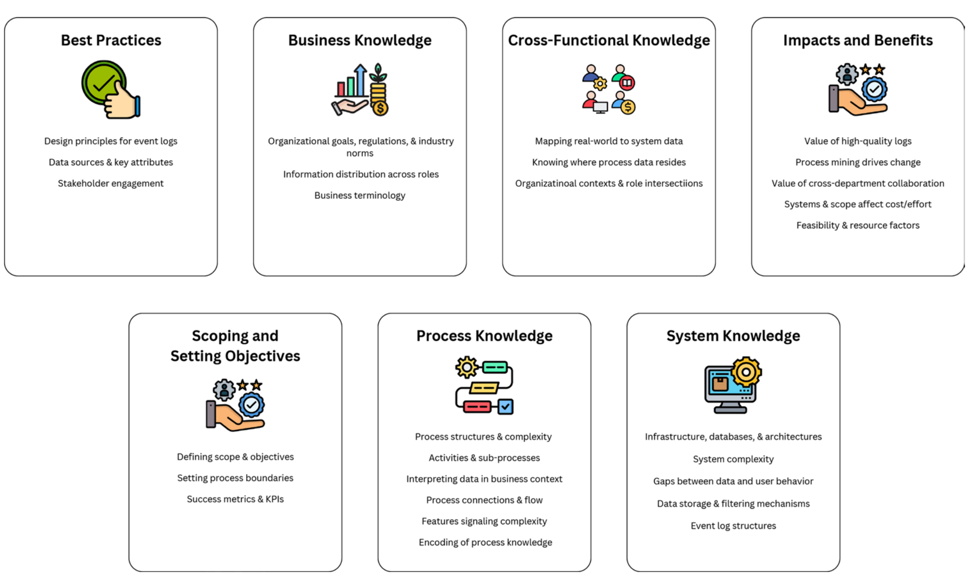

Fig. 1 Identified required units of knowledge for the pre-analysis stage of process mining

Table 3 Required knowledge units and their descriptions

Required Knowledge Unit	Description
Best Practices	Best practices refer to knowledge units that provide actionable guidance for how pre-analysis should be carried out, as opposed to contextual factors (e.g., feasibility, resource constraints) or entry conditions (e.g., data availability). These practices ensure structured event log-building, effective stakehgagement, proper data selection, and a balance between formal and informal knowledge.
Business Knowledge	Business knowledge ensures process mining aligns with organizational goals, regulatory requirements, and industry trends. It helps frame analyses within the right strategic context, ensures event logs capture relevant activities, and facilitates collaboration through a shared understanding of business terminology and process logic.
Cross-Functional Knowledge	Cross-functional knowledge bridges business expertise, IS, and process mining. It includes translational knowledge to map real-world processes to system data, understanding the data sources and system constraints, and grasping the organizational context. A basic understanding across disciplines ensures alignment in data selection and process definition.
Knowledge of Impacts and Benefits	Understanding impacts and benefits of process mining helps align expectations, assess feasibility, and optimize resource allocation. It emphasizes event log quality, cross-departmental collaboration, and work practice changes driven by process insights. Awareness of system constraints, project costs, and feasibility ensures realistic planning, while training and stakeholder engagement support adoption and long-term success.
Knowledge of Scoping and Setting Objectives	Scoping and objective-setting ensure process mining initiatives are aligned with organizational needs. This involves defining clear scope, establishing goals, and setting process boundaries to distinguish standard flows from exceptions. Understanding success metrics and KPIs ensures measurable outcomes, while validation and data interpretation enhance accuracy and reliability.
Process Knowledge	Process knowledge ensures accurate interpretation of workflows, activities, and subprocesses. It involves understanding process structures, complexities, and connections to identify gaps and inefficiencies. This knowledge helps assess bottlenecks and variations, ensuring process mining focuses on the most impactful areas.
System Knowledge	System knowledge is crucial for data extraction, structuring, and event log-building. It includes expertise in databases, system architectures, and data storage to ensure accurate retrieval. Understanding system complexity and real-world user behavior helps align technical data with business processes. Mastery of event log structures and filtering mechanisms ensures meaningful and effective process mining analysis.

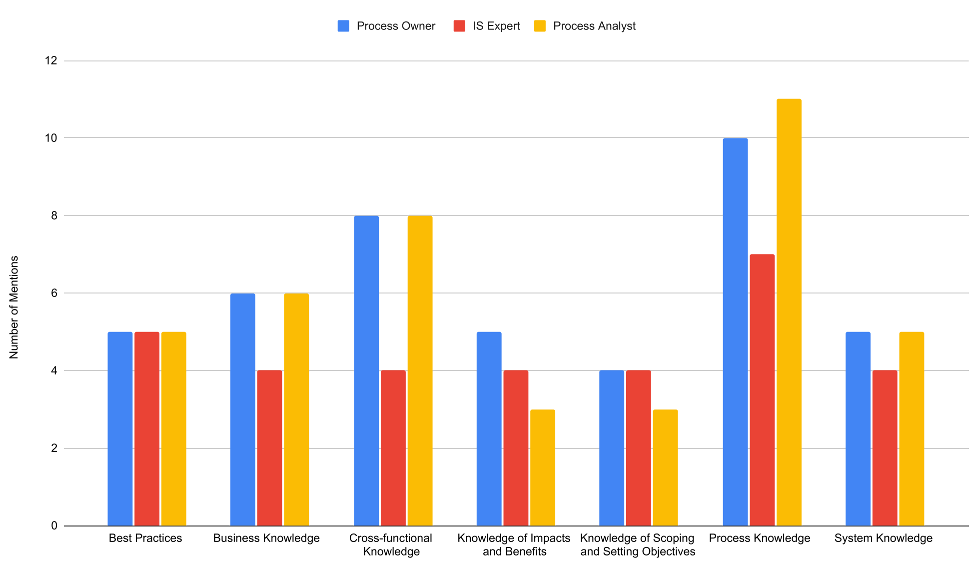

Fig. 2 Distribution of expert profiles mentioning each unit of knowledge required in process mining pre-analysis

number of expert profiles that mentioned each unit, showing how many individuals from each profile referred to these knowledge areas.

Best practices

Several interviewees highlighted some best practices that experts need to be aware of during the pre-analysis stage. Such elements could be considered as foundational knowledge units that experts must possess.

A core area of focus concerned the **principles and practices for designing event logs**. One such knowledge unit is the criteria for high-quality event logs. This includes understanding which attributes are necessary to ensure logs are suitable for analysis. Interviewees emphasized the importance of elements such as complete and unique case identifiers, consistent and correctly formatted timestamps, and clearly defined event types. As one expert explained: *"And to have a set of criteria say ... I want to have best practices for good event log design."* Although these criteria might also be interpreted as prerequisites, interviewees consistently framed them as design-oriented guidance—emphasizing how to construct event logs well, rather than conditions for mere participation. This knowledge allows process analysts to assess whether available logs meet the basic requirements before conducting deeper analysis.

Equally important is knowledge of the **types of data sources and key attributes** typically required in event logs. IS experts must understand which systems are likely to store relevant process information and how to identify useful fields within those systems. As one participant noted: *"It is probably a good practice to limit what you encode there, but you probably need to do some basic relationships and filtering there."* This reflects the need for conceptual knowledge about how data is structured and what attributes—such as case ID, activity name, and timestamp—are critical for log-building.

Beyond the design of event logs, interviewees also stressed the importance of micro-level knowledge about **how to engage stakeholders effectively** during pre-analysis. This type of knowledge concerns interactional practices—how to ask the right questions, interpret partial perspectives, and adapt communication strategies to different stakeholder groups. Practitioners must know who to involve, what to ask, and how to interpret partial or biased perspectives. For example, process owners often possess valuable contextual knowledge about real-world process deviations, while IT staff provide insight into how process data is recorded. As one interviewee reflected: *"So, how do I engage with my stakeholders? And what questions should I ask? What should I keep in mind when discussing this with business users? I should know that they tell me a limited story."* While stakeholder engagement involves interpersonal skill, the unit of knowledge here lies in tactical awareness of how to approach, probe, and make sense of what each group can offer during pre-analysis.

In line with the coding structure presented in Appendix C, these elements are grouped under "best practices" because interviewees consistently described them as actionable, practice-oriented guidance for conducting pre-analysis, rather than as prerequisites or feasibility factors.

Business knowledge

Business knowledge is essential during the pre-analysis stage for aligning process mining efforts with organizational realities. Interviewees highlighted the importance of understanding **organizational goals**, **regulatory requirements**, and **industry practices**. This includes awareness of compliance constraints—such as data privacy or auditability—and how industry norms shape the structure and availability of process data. As one interviewee noted: *"... businesses are constantly evolving, a lot of regulatory changes ... they should not be completely oblivious to those changes."* Such knowledge ensures that pre-analysis remains grounded in both strategic priorities and external constraints.

In addition to understanding the broader business environment, interviewees emphasized the importance of meso-level knowledge about **how process-relevant information is distributed across stakeholder roles**. This unit of knowledge involves recognizing who typically holds strategic, operational, or technical insights relevant to the process under study. Interviewees emphasized that such knowledge is not evenly distributed and that understanding this distribution is critical for sourcing the right information. Knowing how to map specific types of knowledge to specific stakeholder roles enables practitioners to select the right participants and avoid blind spots during pre-analysis.

Finally, process analysts must be able to interpret the language and structure of business data itself. Familiarity with **business terminology and the way business processes manifest themselves in organizational data** is key. Interviewees pointed out that process analysts must be able to recognize how common business concepts appear in datasets. One participant noted: *"... if you do not know the context, it is quite difficult to perceive it."* This underscores the value of understanding domain-specific language and data conventions in order to correctly interpret the meaning behind raw event data.

Cross-functional knowledge

Possessing cross-functional knowledge across business, IS, and process mining is essential during pre-analysis. Experts need a foundational understanding of core concepts from each of these domains to recognize how processes are executed, recorded, and analyzed. As one interviewee noted: *"And process mining has a little bit of knowledge of everything. But that is at least my role a bit. I know little about business and IT, but not enough."* This highlights the necessity of broad, integrative knowledge to situate process mining efforts effectively within organizational contexts.

A particularly important aspect of this cross-functional understanding involves knowing **how real-world business activities are reflected in system data structures**. Experts must be able to trace how operational work translates into system-generated logs, and how different applications capture process-related information. As one interviewee explained: *"They own a little bit of the data model and should be able to translate what is happening in the real world into a language that reflects what is happening in the system."*

Another critical knowledge area is understanding **where the relevant process data is stored and what types of information are typically needed**. Practitioners must understand both business data requirements and the system constraints that affect data availability. As one participant shared: *"I have the domain and system knowledge on what data would be needed and where we could find it for the applications we have in our department."*

In practice, this requires more than technical know-how; it calls for the ability to navigate between system logic and business meaning. Interviewees described how

understanding the organizational context, including how business applications function, the definitions of system fields and events, and their relationship to real-world processes, is crucial. As one expert noted: *"Understanding the application you are working with and the definitions of everything you see. On the other hand, you need to understand the process you are looking at—the actual business process."*

Finally, interviewees pointed to the need for macro-level cross-functional knowledge of **how roles and responsibilities intersect across domains**, particularly between business and IT. This goes beyond knowing individuals who possess the knowledge and how to extract it to understanding how responsibilities, decision rights, and constraints are distributed across organizational units and how these elements interact. Practitioners must recognize who is involved, what expertise they bring, and how their decisions and constraints interact across organizational boundaries. This relational knowledge supports coordination and ensures that the full spectrum of perspectives—business intent, operational constraints, and system implementation—is incorporated into the pre-analysis stage.

Knowledge of impacts and benefits

Another knowledge unit that is important during the pre-analysis stage is the understanding of potential impacts and benefits of process mining. For example, practitioners must recognize **how high-quality event logs can simplify data preparation and enable more efficient analysis**. One interviewee explained the benefit: *"If you have an excellent event log, you can throw away all the snapshots and simply aggregate on the process step or technical definition you want to look at."* This highlights the importance of prioritizing event log quality early to maximize the value of process mining initiatives.

However, process mining's influence often extends beyond analysis. Practitioners should also be **aware that process mining can also lead to organizational or technical change**. Recognizing the broader impact on the organizational level can lead to realistic expectations during early discussions. As one participant observed: *"At some point, you might achieve that people change how they work or that something changes in the source system."*

An important unit of knowledge is the strategic-level awareness that **cross-departmental collaboration enhances process mining outcomes**. Engaging stakeholders from different parts of the organization helps generate relevant analytical questions and uncover inefficiencies that might otherwise go unnoticed, ensuring that pre-analysis reflects real-world operational needs. As one interviewee noted: *"Involving people outside our department helps identify improvement areas, such as reducing rework or optimizing processes."* This form of knowledge is not procedural or structural, but rather reflective: it concerns understanding the business value of inclusive collaboration and recognizing how diverse perspectives contribute to effective pre-analysis stage.

Understanding must also be grounded in technical and operational realities. Recognizing that the **underlying systems and the project scope significantly impact costs and efforts** in process mining initiatives is also vital. One interviewee emphasized, *"Understanding the underlying systems is crucial from the outset, as it significantly impacts costs and efforts. These initial discussions are essential as they define the scope of effort and* *lead directly to the decision to proceed."* This knowledge ensures that feasibility assessments are realistic and grounded in system constraints, avoiding costly misalignments later.

Finally, practitioners must have **knowledge of key factors influencing feasibility and resource requirements**. Early understanding of system capabilities, budget limitations, and required efforts is essential for realistic project planning. As one participant stressed: *"Understanding the scope means determining feasibility, execution potential, budget, and required efforts."*

Knowledge of scoping and setting objectives

Knowledge of **how to define project scope and set objectives** is fundamental during the pre-analysis stage of process mining. Practitioners must understand what constitutes a clear project scope and how objectives should align with stakeholder expectations and organizational priorities. As one interviewee noted: *"We define the scope and understand two main things: scope and objectives ... We need to clarify the objectives."* This highlights the importance of grounding process mining efforts in a shared understanding of purpose.

Beyond overall goals, process analysts must also establish clear analytical boundaries for their investigation. Understanding **how to define process boundaries**, including identifying appropriate start and endpoints, is also crucial. Correctly setting process boundaries ensures that analysis captures the full flow of activities and distinguishes standard operations from exceptions. As one interviewee explained: *"... this is the real beginning, and this is the real end. How we can define those points where people would fall out of the flow ..."*

Another critical knowledge area concerns **success metrics and KPIs relevant to process performance evaluation** in the context of process mining. Process analysts must be familiar with metrics such as first-time-right percentages or acceptable deviation rates, which provide concrete criteria for assessing process behavior. As one participant noted: *"... we often look at the "first time right" percentage ... what does a process execution look like when it is right and which deviations are allowed."*

Process knowledge

Process knowledge is essential during the pre-analysis stage, enabling experts to accurately interpret workflows, assess data completeness, and ensure meaningful analysis. A critical unit of knowledge is an **understanding of business processes and their complexities**, including the typical steps involved in processes such as purchase-to-pay and order-to-cash. As one interviewee noted: *"Using process mining requires business experience and process expertise ... You need to understand the steps in processes like purchase-to-pay or order-to-cash ..."* This highlights the necessity of domain-specific knowledge to prepare reliable event logs.

Experts must also possess **knowledge of process activities and sub-processes**, including the ability to recognize standard activities and understand their meaning within business workflows. One interviewee emphasized: *"The focus then shifts to process knowledge ... a person needs to understand what each activity means."*

Some interviewees also stressed the importance of **interpreting process data in a way that aligns with business meaning and stakeholder expectations**, as it was essential to make analysis useful. As one participant explained:*"Data, on its own, doesn't mean anything unless you give it meaning. Information is very different from raw data, and interpreting data is something that now helps me manage customer expectations."*

Additionally, **understanding how process steps connect** is crucial for identifying gaps or missing events in event logs. Recognizing typical flow patterns supports the validation of process completeness. As one expert remarked: *"Sometimes I come across an event log and say, "This is something that we should look at because this is very interesting." But the business has not thought of that yet."*

Another important unit of knowledge is **awareness of structural features that indicate process complexity**. Practitioners must know how to recognize whether a process is straightforward or involves many variations and activities, which impacts data preparation. As one interviewee explained: *"On different points, we need to have a general understanding of the process to say if it is simple or complex, or if it will involve many or not many activities."*

Finally, interviewees also drew attention to **how process knowledge is encoded**, both formally and informally. Formal documentation—such as standard operating procedures—may not always reflect the actual process as executed. Tacit or informal knowledge, often held by frontline workers, can help explain process deviations and contextual nuances. One expert remarked: *"And the knowledge was codified and made explicit ... But at the end of the day, a lot of this is informal knowledge that people acquire, and even the coded knowledge is imperfect ..."* This awareness supports a more holistic understanding of the process context and reduces the risk of misinterpreting the data.

System knowledge

System knowledge is a critical requirement during the pre-analysis stage, enabling experts—particularly IS experts—to correctly interpret, extract, and structure data for process mining. A foundational unit of knowledge is a **deep understanding of technical infrastructures, databases, and system architectures**, which ensures that relevant data can be retrieved and transformed into usable event logs. As one interviewee emphasized: *"Usually, the system expert is the IT team of the clients. The client is never involved at that stage."* Another interviewee reflected: *"I gained some experience as a business analyst writing down requirements for IT solutions, and I also became a programmer as a hobby ... So, I knew something about databases."*

Another important knowledge area is an **understanding of IS complexity**, including factors such as the degree of customization, the number of integrated systems, and the familiarity of the system environment. This knowledge allows IS experts to assess potential challenges in data extraction. As one interviewee explained: *"Understanding the complexity of systems is important ..."*

Experts must also have **awareness that system data may not accurately reflect real-world user behavior**. Recognizing that users often interact with systems in unexpected ways helps IS experts interpret technical logs more accurately. As one participant noted: *"What you see in a source system does not necessarily reflect what people think they do. People use things differently than you might expect."*

A key unit of technical knowledge concerns **understanding where relevant data is stored (tables and fields) and how data filtering mechanisms operate**. This includes identifying which tables and fields contain process-relevant information and applying appropriate filters during extraction. As one interviewee stated: *"If we talk about process mining in terms of event log creation, you need to understand the data, tables, and fields you are working with. This system knowledge is crucial before creating an event log and transforming data."*

Finally, IS experts must possess **knowledge of event log structures**, understanding how logs should be designed to accurately represent process steps for analysis. Familiarity with necessary attributes such as case IDs, activity names, and timestamps is essential for building effective event logs. As one participant emphasized: *"What helps me greatly is that I understand event log structures—the actual fuel that powers all process mining algorithms."*

Mechanisms of knowledge exchange

To address RQ2, this subsection examines the practical mechanisms through which knowledge is exchanged in collaborative settings. Our analysis reveals four distinct yet interrelated mechanisms that structure how knowledge flows within and across teams.

Figure 3 provides an overview of these mechanisms. *Planned structured interactions*, *Responsive and ad-hoc interactions*, and *Informal ongoing exchanges* are the active knowledge exchange mechanisms, whereas *Knowledge retention practices* act as a supporting layer facilitating all three active mechanisms. Table 4 offers detailed descriptions of these mechanisms. Together, they highlight the dynamic interplay between formal structures, responsive coordination, informal routines, and documentation practices in shaping knowledge exchange.

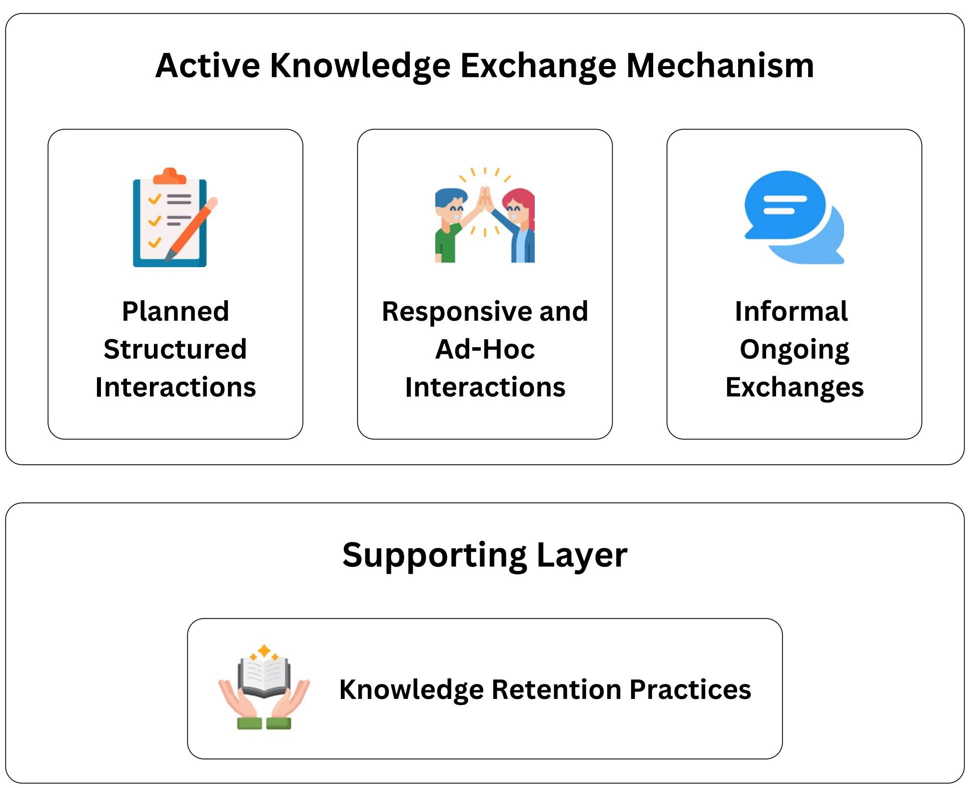

Fig. 3 Overview of the mechanisms of knowledge exchange

Table 4 Mechanisms of knowledge exchange and their descriptions

Mechanism	Description
Planned Structured Interactions	Formal, scheduled sessions such as kickoff meetings, workshops, and progress reviews that facilitate deliberate knowledge transfer across expert profiles and align business, technical, and analytical perspectives.
Responsive and Ad-Hoc Interactions	Unplanned, need-driven exchanges that arise in response to emerging issues, involving quick clarification, spontaneous workshops, or iterative follow-ups across roles and departments.
Informal Ongoing Exchanges	Lightweight, continuous knowledge flows occurring outside formal structures, often through chats, quick conversations, or digital messaging, enabling fast clarification and shared process understanding.
Knowledge Retention Practices	Documentation-based practices that capture and preserve shared knowledge for future reference, enabling continuity and supporting asynchronous knowledge exchange across evolving teams.

Planned structured interactions

Planned structured interactions form a foundational mechanism for knowledge exchange in process mining initiatives. These include **kickoff meetings**, **alignment workshops**, and **scheduled progress reviews**, all designed to transfer critical business, operational, and technical knowledge between expert profiles.

The kickoff meeting is the initial forum where process owners share the business case, define project objectives, and outline key performance metrics. This exchange can be viewed through the lens of Carlile's knowledge boundary framework (Carlile 2004), as it marks a transition from addressing a syntactic boundary—where common terminology and data are shared—to engaging a semantic boundary, where participants begin interpreting and contextualizing that information. In this stage, knowledge is not only transferred but also translated to ensure process analysts can meaningfully interpret event data. Operational insights are further conveyed through discussions of personas representing actual process participants, ensuring a grounded understanding of daily execution realities. During these sessions, the project is scoped, effort is estimated, and expectations around performance metrics are clarified before deeper analysis begins. These discussions often include early reflections on system complexity, available data, and likely resource demands. As one expert explained: *"Understanding the scope means being able to determine if it is feasible or not … there's also the budget question and understanding the efforts required."* Before these sessions, a structured intake process is often used to gather preliminary information and align on goals. Practitioners may rely on checklists or standardized forms to ensure essential information is captured. Such mechanisms can be seen as supporting the development of a syntactic boundary, facilitating consistent and reliable knowledge transfer across stakeholders.

Collaborative discussions and workshops are commonly used to synthesize diverse viewpoints. As one participant noted, *"We sat with three to four people from different lines of services to get a general understanding of the process."* Such sessions refine the collective understanding of the process landscape and surface nuances across organizational units, thereby fostering shared meaning and alignment across the semantic boundary.

Further structured workshops focus on translating business processes into technical system knowledge. IS experts often act as intermediaries during these sessions, ensuring accurate representation of business requirements within event logs. One participant emphasized, *"We definitely need the system expert in between to do this translation."*

Another interviewee remarked on the challenge of misalignment, stating, *"We have a large disconnect between what the business calls their process steps and what the developers call their actions."* Such misalignments reflect the shift from managing a semantic boundary—where meanings must be negotiated—to engaging a pragmatic boundary, where differing objectives and constraints require the transformation of knowledge itself. Structured discussions and formal business-to-system mappings help reconcile these perspectives, enabling consistent terminology, shared accountability, and traceability across stakeholders.

Regular progress meetings, often held digitally, further support structured knowledge alignment. Organizations use platforms such as Microsoft Teams or Google Meet to conduct bi-weekly or monthly syncs, ensuring ongoing collaboration and clarification of emerging issues.

These structured interactions are often supported by a **hub-and-spoke organizational model**. Centralized process mining teams (the hub) typically handle documentation practices, training, capability demonstrations, and infrastructure management, enabling consistent knowledge-sharing across departments. Local teams (the spokes) initiate and conduct analyses, validate findings with operational insights, and translate business needs into technical action. As one participant noted, *"It is best to organize this analysis (process mining pre-analysis) as close to their operations as possible to validate what they see and exchange insights into action promptly."*

By implementing kickoff sessions, translation workshops, and regular progress reviews, organizations create deliberate, role-spanning environments for knowledge exchange that are critical for accurate and actionable process mining outcomes.

Responsive and ad-hoc interactions

Responsive and ad-hoc interactions are another critical mechanism for knowledge exchange, enabling teams to quickly address emerging challenges without relying on formal meeting cycles. These spontaneous exchanges are triggered by specific needs, data discrepancies, or gaps in understanding encountered during project execution.

Teams frequently engage in **on-demand discussions** to resolve ambiguities. One interviewee shared, *"We went to three people in a row for certain questions until we found someone who could answer them."* This highlights how knowledge is sought dynamically across organizational lines to maintain project momentum.

Ad-hoc workshops and collaborative refinement sessions also play a key role in addressing evolving needs, particularly when bridging business and technical perspectives. As one participant explained, *"We create a workshop with developers saying, 'These are the technical logging steps we want to see,' and then align that with what the business is looking for."* Other interviewee described continuous co-creation: *"We do all the heavy lifting … then have some touchpoints along the way where we ask, 'Is this the right direction?'"* Such interactions exemplify ongoing negotiation of meaning and knowledge translation across semantic and pragmatic boundaries (Carlile 2004), where teams must not only interpret but adapt knowledge to reconcile differing viewpoints.

Once initial models are created, teams enter a phase of **validation and feedback loops**. Several interviewees emphasized that findings are never accepted at face value; they are iteratively reviewed with business stakeholders. As one put it: *"After this, we conduct a first data validation meeting with business experts … we go through the* *environments."* This iterative structure creates space for realignment based on user expertise and avoids premature assumptions about what the data truly represents. *"Here's the process—please check that and let me know whether you see any discrepancies or deviations,"* one participant explained.

Underlying these exchanges are **adaptable communication modalities** that support responsiveness. Organizations use a mix of channels—digital meetings, quick check-ins, emails—depending on the urgency or nature of the issue. *"Now post-COVID, it's physical or digital, doesn't matter … but definitely all meetings."* These responsive, ad-hoc interactions form the connective tissue between structured phases of the project. They enable continuous interpretation, transformation, and alignment of knowledge across boundaries, sustaining collaboration where project complexity exceeds what can be resolved in formalized sessions.

Informal ongoing exchanges

Informal ongoing exchanges represent a third, essential mechanism for sustaining knowledge flow throughout process mining initiatives. Unlike planned or ad-hoc sessions, these exchanges occur casually and continuously, often outside structured project frameworks.

One core function of these exchanges is to **cut through unnecessary back-and-forth between siloed teams**. As one participant explained: *"Active communication—we really bring the teams together and get rid of these back and forths in Teams and emails, because we saw it leading to nothing."* Rather than routing through layers of approvals, experts often prefer direct, informal touchpoints to clarify ambiguity or accelerate progress. These touchpoints also support **lightweight collaboration for insight generation**. In cases where a shared process framework already exists, practitioners often forgo new workshops entirely. As one interviewee noted: *"If the framework is more or less aligned, I don't think you need a completely new workshop. Often it's just a simple call to the analyst or domain expert: 'I saw something interesting but think I'm missing something—do you agree?'"* Such low-friction conversations help surface overlooked perspectives and correct misinterpretations on the fly.

In addition, informal exchanges play a role in **collaborative troubleshooting**. Issues with data structures or unexpected behavior in systems are often addressed through impromptu dialogue. As one participant recalled: *"I just went over to the team and said, 'Hey, this column is causing issues; can you explain how to handle it?'"* These low-barrier interactions accelerate issue resolution and prevent bottlenecks in daily operations.

Informal exchanges—whether over Teams, hallway conversations, or a shared Slack channel—provide agility, continuity, and immediacy. They enable rapid clarification, quiet course correction, and cumulative insight building that formal processes alone cannot sustain. Viewed through Carlile's knowledge boundary lens (Carlile 2004), these exchanges represent the ongoing negotiation and transformation of knowledge across pragmatic boundaries, where solutions emerge through mutual adjustment rather than formal coordination.

Knowledge retention practices

While the primary focus is on active knowledge exchange during pre-analysis, organizations also implement structured knowledge retention practices to support continuity across process mining initiatives, especially as teams change or projects scale. A key mechanism for retention is the **use of digital platforms** like Jira and Confluence to track progress, store insights, and maintain project history. One participant explained, *"We have epics and user stories in Jira, where we document what has been done, maintain a continuous flow of comments and updates, and try to document what we learn in Confluence."* These tools serve both operational needs—like task tracking—and strategic ones, such as knowledge preservation.

In addition to structured systems, teams also rely on **real-time note-taking during virtual or in-person meetings**. As another interviewee noted: *"Most of the time, we held Google Meet meetings or met in the office, where we took extensive notes during our initial conversations. This is how we retained most of our knowledge."* Such practices ensure that tacit knowledge exchanged in spontaneous discussions is not lost over time.

Retention practices also intersect with **compliance concerns**, particularly around sensitive data. Some teams maintain documentation on internal data-sharing norms and storage locations. One participant explained: *"We have documentation on how we store and where we store data. And we have guidelines on how to share data with other parties outside the department—because of GDPR and all that."*

Together, these mechanisms—ranging from live note-taking to structured digital repositories—enable knowledge to outlast personnel turnover, support auditability, and ensure that insights from pre-analysis remain actionable throughout the process mining lifecycle. By balancing lightweight communication with structured documentation practices, organizations foster a collaborative environment where expertise is readily accessible, and process mining projects can adapt fluidly to changing needs.

To further clarify the distinct yet interdependent roles of these mechanisms, we developed Table 5, which summarizes them across multiple characteristics grounded in our empirical data. The dimensions—interaction format, temporal orientation, initiation source, functional contribution, participation load, and documentation practice—were not adopted from existing frameworks but were inductively derived from our analysis, thereby constituting an original conceptual lens.

As Table 5 shows, no single mechanism supports all dimensions of effective knowledge exchange. Planned structured interactions provide coordination and terminology alignment but require preparation and are typically top-down. Responsive and ad-hoc interactions offer speed and relevance, adapting quickly to project needs, while informal exchanges lower communication barriers and promote agile clarification. Knowledge retention practices are foundational for ensuring long-term continuity and reuse, particularly in asynchronous or evolving team settings. Their structured nature complements the dynamic forms of interaction that occur throughout a project. This comparative view reinforces our central claim: process mining initiatives require a blend of structured, emergent, synchronous, and asynchronous mechanisms to effectively coordinate expertise and sustain organizational learning.

Challenges in knowledge exchange

This section examines the challenges in knowledge exchange in the pre-analysis stage of process mining, addressing RQ3. Based on the interviews, we identified 12 challenges and categorized them into four groups, detailed in Sections 4.4.1 to 4.4.4. They are arranged in the order of the number of interviewees mentioning them. Table 6 provides

Table 5 Comparison of knowledge exchange mechanisms across key characteristics observed during the pre-analysis stage of process mining. ✓ = yes, X = No

Dimension	Characteristic	Planned	Responsive and Ad-Hoc	Informal	Retention
Form & Mode	Structured format	✓	Partially	X	✓
	Synchronous (real-time)	✓	✓	Often	X
	Asynchronous	X	Sometimes	✓	✓
	Captures contextual detail	✓	✓	Occasionally	✓
	Continuous engagement	Milestone-based	Episodic	✓	✓
Trigger & Timing	Proactive (scheduled)	✓	X	Sometimes	✓
	Reactive (need-based)	Sometimes ✓	✓	✓	✓
Initiation Source	Organization-initiated	✓	Sometimes	X	✓
	Team/individual-initiated	X	✓	✓	✓
	Requires cross-role coordination	✓	✓	Sometimes	✓
Functional Role	Cross-functional alignment	✓	✓	Occasionally	✓
	Long-term reuse	Limited (depends on notes)	X	X	✓
	Supports model refinement	✓	✓	X	Background only
	Terminology alignment	✓	✓	X	✓
Participation Load	Low barrier to participation	X	✓	✓	For retrieval
	Requires preparation/agenda	✓	Sometimes	X	✓
Documentation	Traceable output	If minutes exist	Sometimes	X	✓
	Feeds future projects	Limited	X	X	✓

Table 6 Challenges in knowledge management and their descriptions

Challenge	Description
Stakeholder Misalignment and Role Ambiguity	Challenges arising from unclear responsibilities, misaligned perspectives, and fragmented ownership of process knowledge among business, IT, and analytical roles.
Gaps in Communication and Understanding	Challenges stemming from inconsistent definitions, vague requirements, and misunderstandings that hinder the formation of a shared understanding during pre-analysis.
Missing or Inaccessible Knowledge	Challenges related to undocumented, tacit, or siloed knowledge that cannot be accessed or retained effectively during early-stage analysis.
System Complexity and Fragmentation	Challenges caused by technical environments that are decentralized, poorly documented, or distributed across incompatible systems, limiting visibility and knowledge integration.

the description of the identified challenges, whereas Table 7 groups them according to the companies in which they were observed.

Because the number of interviewees differed between consulting (6) and non-consulting (9) firms, raw counts would over-represent the latter. Therefore, the number of mentions per theme was normalized by group size (i.e., total mentions divided by the number of interviewees in each group). As shown in Table 8, non-consulting firms reported proportionally more challenges related to stakeholder misalignment and communication, whereas consulting firms exhibited higher relative frequencies of missing knowledge and system complexity.

Table 7 Challenges reported by interviewees across companies, grouped by thematic category. The *companies* column indicates, in brackets, the number of interviewees from each company who mentioned the challenge

Challenge Themes	Challenges	Companies (Consulting/Non-Consulting)
Stakeholder Misalignment and Role Ambiguity	Business-Technical Translation	Consulting: C (3) Non-consulting: A (1), B (2)
	Stakeholder Misalignment	Consulting: D (1) Non-consulting: B (2)
	Difficulty Identifying Stakeholders with Relevant Process Knowledge	Non-consulting: B (1), E (1)
	Fragmented Process and System Knowledge	Non-consulting: A (1), B (3), E (2)
Gaps in Communication and Understanding	Ambiguous Terminologies and Definitions	Consulting: C (1), D (1) Non-consulting: A (1), B (2)
	Scope and Expectation Mismatch	Consulting: C (1), D (1) Non-consulting: A (2), B (3)
	Unclear Requirements	Non-consulting: A (1), B (1), E (1)
Missing or Inaccessible Knowledge	Lack of Documentation and Knowledge Retention Policy	Consulting: D (2) Non-consulting: A (1), E (1)
	Lack of Knowledgeable Personnel	Consulting: D (1) Non-consulting: A (1), B (1)
	Lack of Structured Knowledge Sharing	Consulting: D (2) Non-consulting: A (1)
System Complexity and Fragmentation	Legacy and Unfamiliar Systems	Consulting: C (4), D (1) Non-consulting: B (1)
	Fragmented Systems	Non-consulting: B (2), E (2)

Table 8 Normalized distribution of reported challenges across consulting and non-consulting companies. Counts indicate the number of interviewees mentioning each challenge theme. "Per interviewee" values are normalized by group size (6 consulting, 9 non-consulting)

Challenge Theme	Consulting (mentions)	Consulting (per int.)	Non-Consulting (mentions)	Non-Consulting (per int.)
Stakeholder Misalignment and Role Ambiguity	4	0.67	13	1.44
Gaps in Communication and Understanding	4	0.67	11	1.22
Missing or Inaccessible Knowledge	5	0.83	5	0.56
System Complexity and Fragmentation	5	0.83	5	0.56

Stakeholder misalignment and role ambiguity

Business-technical translation (7 interviewees): During the pre-analysis stage of process mining, knowledge exchange is hindered by the disconnect between business terminology and its implementation in IS. While closely related to terminological ambiguity, this challenge specifically concerns the translation of business concepts into technical structures and actions, which requires intensive back-and-forth communication between business stakeholders and system experts. One expert noted, *"We have a very large disconnect between what the business is naming their process steps and what the application developers are actually calling their actions within the system."*

The challenge is compounded when technical implementations differ from business expectations; as one participant explained, *"The business can say, "I need step A," but if the tech guys implemented step A as Z, Y, and F, then you have no clue what you are looking for."* Similarly, business users may see terminology different from what is stored in the database on their screens, creating further confusion during data transformation or event log creation. This misalignment complicates the mapping of business process steps to technical data structures, which is critical for identifying relevant data sources and building accurate event logs during pre-analysis.

Moreover, bridging this divide often necessitates collaboration between business experts and system experts, as it is rare to find individuals who possess deep knowledge of both domains. Thus, this challenge directly reflects a difficulty of knowledge exchange: unless knowledge from all sides is effectively aligned, the pre-analysis stage risks producing incomplete or misleading results.

Fragmented process and system knowledge (5 interviewees): Knowledge exchange in the pre-analysis stage is often limited by gaps in stakeholders' understanding of the process to be analyzed and the IS involved. One expert observed, *"It is very difficult and almost impossible to find a business user with the entire process view,"* as most users possess deep but narrow expertise limited to their specific segment. Because of this fragmentation, knowledge must be exchanged across multiple individuals to piece together a complete view, yet such exchange is often incomplete or inconsistent. This fragmented knowledge makes it challenging to build a shared understanding across teams.

In addition, stakeholders may lack awareness that event logs can be designed for specific analysis purposes. As one interviewee noted, *"The largest gap I see is the realization that we can create event logs specifically for this purpose."* Without this foundational knowledge, relevant data cannot be easily identified or scoped during pre-analysis, since critical information is not effectively communicated between technical and business participants.

In some cases, analysts are assigned to projects in unfamiliar domains, lacking even basic understanding of the process itself. This absence of domain knowledge further obstructs the interpretation and structuring of event data, and hinders productive exchange with domain experts, who must fill in the missing context for analysis to progress.

Stakeholder misalignment (3 interviewees): Knowledge exchange in the pre-analysis stage is hindered when stakeholders are misaligned in their priorities, level of engagement, or understanding of each other's roles. One interviewee noted, *"It is quite a challenge to align those completely different stakeholders in such a way that they prioritize and get the value they each want to see."* When stakeholders do not share a common understanding of the project goals—or of each other's responsibilities, information needs, and expected outputs—they are less likely to participate actively in knowledge sharing, which directly undermines the consolidation of the knowledge needed for pre-analysis.

This misalignment is particularly evident between IT teams and business units. IT departments may not recognize the need for creating specific logs or sharing detailed data, assuming existing datasets are sufficient. One participant explained, *"IT does not really benefit—or at least that is what they think—from creating that."* As a result, essential knowledge about available data and its suitability for analysis is not exchanged. This gap in understanding and perceived value limits the availability of information needed to support effective pre-analysis. In parallel, limited cooperation from clients—such as repeated delays or resistance in providing needed information—can exacerbate knowledge gaps and stall pre-analysis work.

Difficulty locating stakeholders with relevant knowledge (2 interviewees): Even when process knowledge exists, knowledge exchange is often hindered by the difficulty of identifying who holds it. In large organizations, responsibilities are distributed across departments and locations, creating uncertainty about who is responsible for which part of the process. One expert explained, *"We spent a huge amount of time trying to find the business owner—who is in the end responsible for the process and who can change the process."* When roles are unclear, knowledge remains siloed and difficult to access.

This issue is compounded by frequent role changes and organizational restructuring. One participant noted, *"I tried to make a stakeholder overview … and stopped because it changed so much over time. People left, groups dissolved."* As a result, teams often have to re-identify key contacts, which delays the gathering of required information during pre-analysis and creates repeated bottlenecks. In addition, experts reported difficulty in identifying cross-functional stakeholders who understand both business processes and the technical aspects required for pre-analysis, further fragmenting the available knowledge. This challenge is compounded by the scarcity of individuals with end-to-end process expertise, as most stakeholders have knowledge limited to their own segments—undermining both the completeness and the effectiveness of knowledge exchanged during pre-analysis. These difficulties stem not from the content of stakeholders' knowledge, but from the organizational complexity involved in finding and accessing the right individuals in time, which ultimately restricts the flow of knowledge essential for pre-analysis.

Gaps in communication and understanding

Scope and expectation mismatch (7 interviewees): Knowledge exchange during the pre-analysis stage can be impeded by misaligned expectations between business and technical stakeholders. One interviewee noted, *"The main challenge we face is scope and expectations … Are the objectives achieved?"* When stakeholders initiate projects with vague or overly ambitious goals, such as, *"Let us process mine all the processes you can think of for a huge department,"* the lack of specificity creates gaps in shared understanding and obstructs effective exchange about priorities and feasibility.

These misunderstandings often stem from a limited or inaccurate perception of what process mining entails. As one participant explained, *"The business manager often initiates the request, thinking they are asking for process mining, but usually, they are just asking for a dashboard."* Such mismatches in terminology and expectations weaken knowledge exchange during pre-analysis discussions, as the two sides are not actually talking about the same thing. When process mining is treated as a generic solution without clearly defined scope or data needs, it becomes difficult to align priorities and facilitate effective knowledge exchange during project scoping.

Several experts also highlighted that stakeholders frequently underestimate the complexity of analysis during initial discussions, assuming it will be easy to solve without accounting for the data relationships and constraints. As one explained, *"the main risk in scoping lies in the lack of knowledge or incorrect assessments, especially when the goal is to evaluate feasibility and required effort."* Ambiguity also arises when trying to define what exactly should be analyzed—for instance, differing interpretations of key objects or user definitions can cause confusion when attempting to filter or generalize the process scope.

Moreover, stakeholders may be impressed by visual dashboards or demo outcomes and assume their own data is equally prepared. But, as one expert warned, *"the event logs we use were not created for this type of analysis—they just happen to be available."* This creates a mismatch between expectations and data readiness, leading to breakdowns in the exchange of realistic requirements and constraints during pre-analysis.

Ambiguous terminologies and definitions (5 interviewees): During the pre-analysis stage of process mining, knowledge exchange is impeded when key concepts such as time, events, and activities are interpreted differently by stakeholders. This ambiguity leads to fragmented understanding and complicates the identification and interpretation of process data. One expert illustrated the issue by asking, *"What do you mean by "active?" Because you could start a repair that takes two hours, but in between, you take a break or pause and then restart."*

Timestamps are a frequent source of confusion, as their meanings vary depending on the system context. As one interviewee explained, *"An approval timestamp could signify when a workflow was created, assigned, implemented, or closed."* Such ambiguity complicates the alignment of data interpretations between teams and affects event log-building, since each group exchanges knowledge based on different assumptions. Similarly, the lack of consistent definitions for key performance indicators can create breakdowns in shared understanding. One participant noted, *"When the business asked me, "Do you have any insights on first time right?" I said, "Yeah, fine, but what is the definition of first-time-right?"*

While this challenge overlaps with *Business-Technical Translation*, it is distinct in that it concerns ambiguity within or between stakeholder groups about the meaning of commonly used process terms, rather than cross-domain misalignment between business concepts and technical implementations.

Unclear requirements (3 interviewees): Knowledge exchange in the pre-analysis stage is hindered when project requirements are vague or poorly defined. One interviewee remarked, *"The biggest challenge is that people do not know what they want. Either they want it all and still have no clue what they want."* When objectives are unclear, stakeholders struggle to articulate and exchange the knowledge needed to define scope, identify relevant data, and align expectations.

Unclear requirements also lead to inefficient feedback loops. As another participant explained, *"There was always some back and forth where I just said, "No, it is not what I wanted to see." But that leads to nothing."* Without mechanisms to build shared understanding between technical and business teams, critical knowledge remains siloed—preventing effective exchange, weakening collaboration, and delaying the pre-analysis phase.

Missing or inaccessible knowledge

Lack of documentation and knowledge retention policy (4 interviewees):

Knowledge exchange during the pre-analysis stage is hindered by the absence of accessible documentation and institutional knowledge. One interviewee noted, *"No central knowledge database exists, and all operating companies work differently. We build a good understanding of a process at one company, which is completely different at the next."* Without reusable documentation, teams face inconsistencies in understanding prior processes, forcing knowledge exchange to occur informally and inconsistently across projects, and making it difficult to retain lessons learned over time.

Undocumented expertise further complicates collaboration. As one participant explained, *"The knowledge is encoded … and that is extremely difficult to manage and* *find later."* When key individuals leave, undocumented insights are lost, creating knowledge gaps that disrupt pre-analysis activities. Process analysts reported struggling to validate process transitions or understand system behavior without reference materials, which hinders systematic exchange and instead leads to reliance on business users and trial-and-error approaches.

In some cases, teams also lack access to formal training resources for key systems or tools. As one expert reported, they had to rely heavily on self-directed research due to the absence of structured onboarding or documentation, leaving knowledge exchange dependent on individual efforts instead of shared, structured organizational resources.

Lack of knowledgeable personnel (3 interviewees): Pre-analysis is hindered when organizations lack personnel with the specific expertise required to support knowledge exchange. One expert explained, *"We do not have a process mining analyst. We have a developer supporting us with reviews and analysis, but there is no certain position for that."* In such cases, teams rely on individuals who are already handling multiple responsibilities, limiting the time and focus available for knowledge sharing.

A related issue is the scarcity of specialized IS experts familiar with system logging or development processes. As one participant noted, *"Such an expert is really, really rare to have, especially in the applications that are still under development."* Without IS experts, it becomes difficult to access or exchange the technical knowledge required to identify relevant data during pre-analysis, leaving critical expertise unavailable to the teams that depend on it.

Knowledge loss due to personnel turnover can further disrupt early-stage activities. One interviewee recalled how the departure of a key team member led to delays because critical insights had not been documented or transferred, demonstrating how the absence of knowledgeable personnel directly interrupts the exchange of information needed for pre-analysis.

Lack of structured knowledge sharing (3 interviewees): In contrast to long-term knowledge loss, this theme captures the absence of structured mechanisms for real-time knowledge coordination during active projects. Knowledge exchange during the pre-analysis stage of process mining is often hindered by the absence of formal mechanisms for capturing and centralizing critical information. One interviewee described the reliance on informal communication methods such as phone calls and ad hoc discussions, noting, *"Phone calls, questions, and answers ... where do we write it down? At the end of the day, what we learned we code."* This fragmented approach leads to inconsistent understanding across teams and complicates retrieval of shared insights during pre-analysis.

In addition, gathering accurate input requires coordination among multiple experts, which can be time-consuming and error-prone. One participant noted, *"The main challenge is ensuring accurate and timely centralized knowledge because certain aspects can be time-consuming."* Another added, *"... gathering the correct information was sometimes challenging,"* emphasizing the difficulty of consolidating knowledge when defining project scope and identifying relevant data sources.

These challenges are further compounded by the imperfect and often orally transmitted nature of process expert knowledge. As one expert reflected, knowledge is frequently passed through conventions and workarounds—*"a huge knowledge transfer program"* becomes necessary when shifting from decentralized teams to shared services, showing how the absence of structured mechanisms forces knowledge exchange to remain ad hoc and inconsistent during pre-analysis.

System complexity and fragmentation

Legacy and unfamiliar systems (6 interviewees): During the pre-analysis stage, knowledge exchange is often obstructed by legacy or unfamiliar systems that lack clear documentation. Teams struggle to locate and interpret relevant data sources, which hampers their ability to share accurate insights to others—especially when dealing with outdated architectures. As one interviewee noted, *"A simple payments process might involve 50 different systems with outdated programming languages."*

System-specific nuances and undocumented configurations further complicate efforts to access necessary knowledge. Experts often must work without database schemas or reliable references. As one participant described, *"Is it the standard system, off-the-shelf, or a completely unknown system—a very niche solution that no one has ever heard of"* In such environments, knowledge tends to be siloed among a small number of individuals, creating bottlenecks for knowledge exchange during pre-analysis when technical understanding is needed to identify and extract relevant data.

These challenges are sometimes exacerbated by organizational resistance to change, where entrenched legacy systems persist not only due to technical debt but also due to cultural norms or management structures, further restricting the flow of knowledge needed to support effective pre-analysis.

Fragmented systems (4 interviewees): Knowledge exchange in the pre-analysis stage is hindered when organizations operate across fragmented systems. Different tools for workflow management, customer relationship management, and data analysis often result in siloed information, which prevents stakeholders from easily sharing or consolidating their knowledge into a unified process view. One interviewee explained, *"We use different vendors for every layer ... we spread our risks to avoid vendor lock-in, but then we end up heavily dependent on Microsoft, for instance, with their Azure platform."*

The issue is further complicated when decentralized ERP systems are involved. As another participant noted, *"We do not do process mining on one system ... we have different systems, logs, data types."* This system fragmentation makes it difficult to align data and coordinate knowledge across departments, ultimately restricting the exchange of critical insights needed to collaboratively prepare for process mining activities during pre-analysis.

Figure 4 presents the challenges categorized by type, ordered by the number of interviewees who explicitly mentioned each challenge. Notably, the challenges varied among interviewees, with no single challenge being explicitly cited by the majority.

Discussion

This section reflects on the study's findings, answers the research questions, and discusses the implications and limitations. Our study focused on the critical knowledge required for successful process mining pre-analysis, the mechanisms through which this knowledge is exchanged, and the challenges involved.

Table 9 provides a high-level overview of the findings of this study. The findings highlight the contributions of process owners, IS experts, and process analysts—each bringing unique expertise to this crucial stage. Process owners define objectives and KPIs, IS

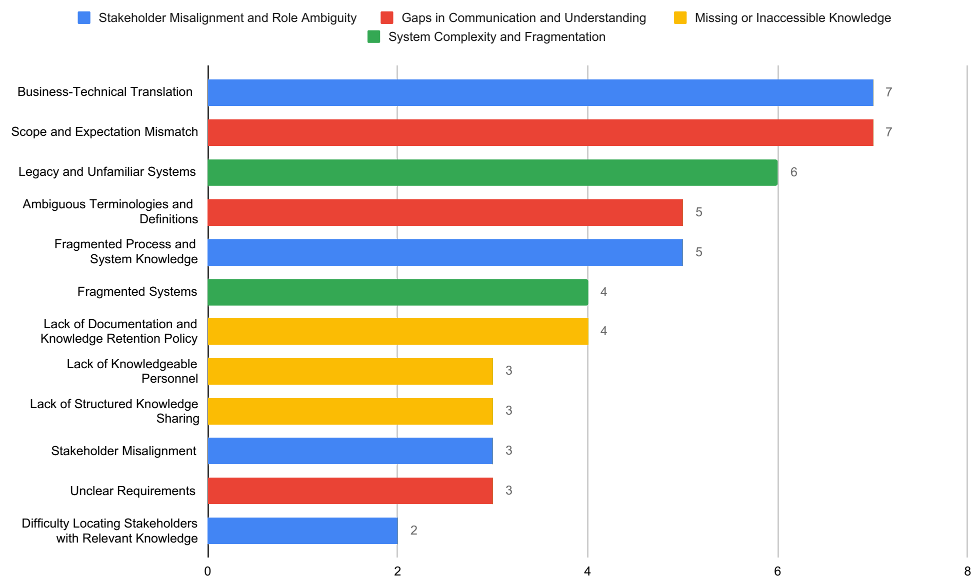

Fig. 4 Challenges identified from interview responses, categorized by their type

experts handle data extraction and technical requirements, and process analysts apply analytical tools to derive insights (Eck et al. 2015; Mamudu et al. 2022). Understanding how these experts interact and share knowledge is essential for effective process mining implementation.

Knowledge required in process mining pre-analysis (RQ1)

RQ1 asked what types of knowledge are required and utilized by different expert profiles during the pre-analysis stage of process mining. Our findings identify *process knowledge*—defined as an understanding of the activities, goals, metrics, and boundaries of the business process under analysis—as one of the key types of knowledge involved. This type of knowledge emerged as the most frequently mentioned requirement, cited by twelve interviewees. It shapes nearly all aspects of pre-analysis, from data extraction and event log-building to the interpretation of process metrics. For instance, when building event logs, knowing which activities constitute a particular process is essential for producing meaningful analyses. In one case, analysts needed process knowledge to connect Salesforce case handling with the wider order-to-cash process. They traced how customer emails created cases in Salesforce, how these were queued and handled, and how they eventually triggered order entry in SAP. Recognizing this sequence, and that it formed a subset of the broader order-to-cash process, was critical to constructing a meaningful event log.

Interestingly, the *scoping and objective setting* was the least cited knowledge unit, mentioned by only four participants. However, three of these held hybrid roles encompassing all three expert profiles, and the fourth combined process owner and IS expertise. This suggests that individuals with broader cross-domain exposure are more attuned to the strategic importance of clearly defined goals in the pre-analysis phase. While not frequently mentioned, this knowledge unit may be underrecognized rather than unimportant. For example, in Financial Economic Crime (FEC) process, the business requested insights on the KPI “first time right.” Analysts first had to clarify the scope of what “first time right” meant—whether avoiding rework across all steps, or simply reaching a final

Table 9 Overview of the findings of the study

Study Dimensions	Categories	Description
Knowledge Units	Best Practices	Actionable guidance for structured log-building, stakeholder engagement, and effective pre-analysis.
	Business Knowledge	Ensures alignment of process mining pre-analysis with organizational goals, regulations, industry context, and shared terminology.
	Cross-Functional Knowledge	Bridges business, IS, and process mining through translational knowledge and system-context awareness.
	Knowledge of Impacts and Benefits	Clarifies feasibility, costs, collaboration, and adoption by understanding outcomes and constraints.
	Knowledge of Scoping and Setting Objectives	Defines clear scope, goals, KPIs, and boundaries to ensure reliable pre-analysis.
	Process Knowledge	Provides deep understanding of workflows, variations, and bottlenecks.
	System Knowledge	Delivers expertise in data structure, data extraction, and system behavior for pre-analysis.
Mechanisms of Knowledge Exchange	Planned Structured Interactions	Formal, scheduled sessions enabling deliberate, cross-expert knowledge alignment.
	Responsive and Ad-Hoc Interactions	Unplanned, issue-driven exchanges that provide rapid clarification and adaptive collaboration.
	Informal Ongoing Exchanges	Lightweight, continuous conversations fostering shared understanding.
	Knowledge Retention Policies	Documentation and preservation methods ensuring continuity and supporting asynchronous knowledge sharing.
Challenges in Knowledge Exchange	Stakeholder Misalignment and Role Ambiguity	Unclear responsibilities and fragmented ownership create misaligned perspectives across roles.
	Gaps in Communication and Understanding	Inconsistent definitions and vague requirements obstruct shared understanding in pre-analysis.
	Missing or Inaccessible Knowledge	Tacit, siloed, or undocumented knowledge remains unavailable for effective early-stage analysis.
	System Complexity and Fragmentation	Decentralized, incompatible, or poorly documented systems hinder visibility and integration.

approval without rejection. These discussions illustrate how scoping involves negotiating shared definitions of key objectives before analysis can proceed.

To further contextualize these findings, we map the identified knowledge areas to established typologies of knowledge. Drawing on De Jong et al. (Jong and Ferguson-Hessler 1996), we interpret knowledge of event log structures and process models as *conceptual knowledge*, while *situational knowledge* includes knowing which data is needed and where it resides. Knowledge of how knowledge is shared and coordinated within teams reflect *procedural knowledge* and the ability to define objectives and scope aligns with *strategic knowledge*.

We also distinguish between *explicit* and *tacit* knowledge (Smith 2001; Gorman 2002). Explicit knowledge—such as documented best practices, structured models, and system documentation—is easily codified and transferred. In contrast, tacit knowledge resides with individuals and is often transmitted through social interaction. This distinction is particularly relevant in process mining pre-analysis, where much of the critical knowledge (e.g., exceptions to process norms, system idiosyncrasies) might often remain undocumented. The reliance on tacit knowledge highlights the need for deliberate knowledge-sharing mechanisms and emphasizes the vulnerability of organizations where knowledge is concentrated in individuals.

Our findings further show that while deep expertise in one domain (e.g., IT, business, or analytics) remains vital, professionals benefit from a working knowledge of adjacent areas. This hybrid understanding enhances collaboration and enables more informed decision-making across expert boundaries. As one participant put it, becoming effective in process mining required wearing different "hats": first the system hat to understand tables and fields, then the business hat to grasp procurement processes, and finally the process mining hat to build and analyze event logs. The ability to switch between these hats illustrates how working knowledge of adjacent domains equips professionals to connect perspectives and produce more meaningful outcomes.

However, most knowledge in organizations remains siloed, with limited formal mechanisms for integration. The dominant mode of knowledge exchange remains verbal, supplemented by visual tools such as process maps and entity-relationship diagrams. According to Bhatt's framework (Bhatt 2002), knowledge is still largely held by individuals rather than systematically captured at the organizational level. This signals an opportunity for organizations to adopt structured knowledge retention practices that promote institutional memory and reduce reliance on ad hoc knowledge sharing.

Mechanisms of knowledge exchange (RQ2)

To answer RQ2, we investigated how knowledge is exchanged among expert profiles during the pre-analysis stage of process mining. Our findings reveal four complementary mechanisms: *i)* planned structured interactions, *ii)* responsive and ad-hoc interactions, *iii)* informal ongoing exchanges, and *iv)* knowledge retention practices. Together, these mechanisms support the collaborative translation of real-world processes into structured, analyzable event logs.

Planned structured interactions—such as kickoff meetings, alignment workshops, and progress reviews, form the backbone of early-stage knowledge exchange. They enable expert profiles to clarify project goals, share domain-specific insights, and define expectations. These sessions are particularly effective in orchestrating communication between process owners, IS experts, and process analysts, each of whom brings distinct but interdependent perspectives. Regular meetings and structured business-to-system translation workshops help bridge terminology gaps and ensure that process models reflect both business reality and technical feasibility. In one case, this structure was supported by an intake form with a predefined set of questions for requestors. The form forced early clarification of goals and boundaries—for example, distinguishing between what process mining can address (e.g., identifying where users drop out of an onboarding process) and what falls outside its scope (e.g., demographic profiling or reasons behind user behavior). This ensured that expectations were aligned before analysis began.

Responsive and ad-hoc interactions such as, spontaneous workshops, targeted follow-up sessions, and impromptu consultations, complement the planned interactions by addressing emerging challenges in real-time. These interactions, which are often informal and immediate, tend to be triggered by data discrepancies, process ambiguities, or evolving stakeholder needs. While unplanned, they play a critical role in sustaining momentum and resolving uncertainties that may not have been anticipated in structured meetings. For instance, in an HR project, an initial workshop provided a broad overview of the recruitment process, but new questions kept surfacing. This led to five or six shorter follow-up meetings focused on clarifying details directly in the Workday system.

Informal ongoing exchanges further enable fluid collaboration across expert roles. These interactions—often facilitated by messaging platforms and casual conversations—support continuous clarification, lightweight problem-solving, and quick feedback loops. Rather than being tied to formal milestones, they operate in the background of daily project work, lowering barriers to communication and supporting agile adaptation.

Although the mechanisms above emphasize active interaction, knowledge retention practices also play a pivotal role in sustaining knowledge exchange over time. Organizations use tools such as *Jira* and *Confluence* to document decisions, track progress, and capture lessons learned. These practices do not merely serve as repositories but function as enablers of future exchanges by ensuring that knowledge persists beyond the individuals initially involved.

A critical insight from our study is that knowledge exchange in process mining is not unidirectional. Experts frequently shift between roles as knowledge providers and recipients depending on the domain under discussion. For example, a process owner working on special asset management first used their domain knowledge to help others interpret the process model. Yet in the same project, they became the learner when an analyst explained how data access worked in Celonis. This demonstrates how expertise circulates across roles rather than flowing in a single direction.

Challenges in knowledge exchange (RQ3)

RQ3 explored barriers to effective knowledge exchange during the pre-analysis stage. Our findings reveal a complex set of interrelated challenges—spanning semantic, structural, organizational, and technical domains—that collectively hinder collaboration.

A key challenge emerges from the disconnect between business language and technical implementations. Business concepts are often not reflected directly in system structures, leading to misinterpretations and inefficiencies when identifying relevant data sources. This misalignment makes it a challenge to establish a cross-domain understanding. Moreover, this disconnect is not limited to terminology; it extends to differing assumptions about responsibilities, priorities, and project goals. Such a disconnect hampers efforts to establish a common frame of reference essential for effective process mining pre-analysis.

Another recurring obstacle involves ambiguous definitions of core process concepts such as activities, timestamps, or performance indicators. Stakeholders often operate with implicit or varying understandings of terms like "approval time" or "first-time-right," which undermines the precision required for effective exchange. This resonates with Eppler's framework of knowledge communication, in which semantic ambiguity and contextual variability lead to communication breakdowns (Eppler 2007). Without mechanisms to surface and resolve these differences, knowledge exchange remains inconsistent and prone to error.

Difficulties also stem from the challenge of locating individuals with the appropriate process knowledge, particularly in large or decentralized organizations. In such environments, knowledge is dispersed and held by individuals whose roles may be fluid or poorly defined. This creates delays and increases reliance on informal communication. Combined with frequent personnel changes, such conditions obstruct the continuity and completeness of knowledge exchange. In these situations, where formal structures and documentation are insufficient, knowledge exchange often depends on personal networks and informal recognition of expertise. These findings reinforce prior observations that trust, recognition of expertise, and stable social relationships are vital enablers of knowledge exchange (Oliver 2001; Disterer 2001).

Technical barriers further constrain exchange. Legacy and fragmented systems isolate information, limit data visibility, and require expert interpretation—yet those experts are often unavailable or overwhelmed. When systems are poorly documented, stakeholders are forced to rely on trial-and-error exploration or indirect knowledge sources. This significantly slows the early analysis phase and introduces uncertainty. The issue is not merely the absence of data, but the difficulty in transforming what is technically available into knowledge that is usable across stakeholder groups.

Finally, the lack of structured mechanisms for capturing and centralizing knowledge leads to inconsistent and ephemeral exchanges. Knowledge is often shared through ad hoc conversations, undocumented interactions, or temporary notes—none of which provide the persistence or transparency needed to support ongoing collaboration. Without deliberate structure for exchange—such as shared glossaries, stakeholder maps, or communication protocols—teams struggle to align on basic assumptions, coordinate effectively across roles, and maintain a shared understanding throughout the pre-analysis phase. This, in turn, undermines their ability to define scope, extract data, and construct useful event logs.

The distribution of challenges, as seen in Table 7, reveals distinct patterns between consulting and non-consulting firms. Consulting companies (C and D) are more frequently associated with issues of business–technical translation, ambiguous terminology, and unclear requirements. This underscores the communication barriers inherent in the intermediary role that consulting companies play between the client and process mining. In contrast, non-consulting firms (A, B, E) emphasize missing or inaccessible knowledge, such as insufficient documentation, lack of structured sharing, and weak retention practices—reflecting their dependence on long-term organizational memory. System fragmentation and stakeholder misalignment appear across both groups, indicating that these problems transcend organizational type and represent systemic challenges to process improvement initiatives.

Overall, these challenges indicate that knowledge exchange during pre-analysis is highly context-dependent and sensitive to organizational structure, project dynamics, and the evolving groups of actors. The diversity of challenges reported across interviews suggests that no single barrier dominates universally. Consequently, a one-size-fits-all approach to knowledge exchange is unlikely to succeed. Instead, process mining initiatives should incorporate adaptive exchange strategies that respond to specific stakeholder needs, system environments, and communication norms. Recognizing and actively managing the conditions under which knowledge is shared will be essential to improving the effectiveness of pre-analysis in process mining initiatives.

Implications

This study carries important implications for both practitioners and researchers engaged in process mining.

For practitioners, this study offers a structured repository of the types of knowledge required during pre-analysis. This repository can serve as a practical tool for competency development in organizations, thus enabling them to better prepare professionals who contribute to the early stages of process mining initiatives. It can support the design of training and certification programs by clarifying the specific knowledge units essential for conducting tasks such as data understanding, system mapping, and event log-building.

Moreover, organizations can also apply this repository to structure onboarding and mentoring processes, ensuring that new team members quickly assimilate the necessary domain, system, and process analysis knowledge in a coherent and efficient manner—thus complementing broader competency development efforts. Furthermore, by aligning roles and responsibilities based on these knowledge requirements, organizations can refine team structures to enhance collaboration and reduce friction during pre-analysis.

In addition, the study sheds light on knowledge exchange mechanisms and the challenges that hinder effective exchange. This information is equally valuable for practitioners, who can draw on the identified mechanisms to facilitate knowledge sharing during pre-analysis. They can also make informed decisions about which mechanisms are most appropriate in different circumstances, while anticipating potential barriers. By addressing these challenges proactively, organizations can establish mitigation measures before they disrupt real-world process mining initiatives.

For researchers, the study provides a foundation for theorizing about the interplay of knowledge domains in the pre-analysis stage. It invites exploration into how different expert profiles align and integrate their perspective to build a shared understanding of process contexts and system realities. The categorized knowledge units offer a baseline for developing models of collaborative knowledge formulation and exchange. These findings also open avenues for cross-sector studies comparing how pre-analysis knowledge needs vary by industry, system complexity, or regulatory environment. In addition, the study's examination of knowledge exchange mechanisms and challenges provides a basis for theorizing about how such mechanisms enable or constrain collaboration, and how challenges can be anticipated, mitigated, or turned into opportunities for organizational learning.

As prior research has not examined knowledge requirements in the process mining pre-analysis stage—or in process mining more broadly—this study provides a foundational contribution. This context is critical because process mining pre-analysis sits at the intersection of business, IS, and analytical expertise, where early misalignment can undermine entire initiatives. Existing KM and IS research does not account for this high dependency on event log-building and cross-domain coordination. Drawing on knowledge boundary theory (Carlile 2004), the study conceptualizes pre-analysis as a setting where experts from different domains must bridge differences in knowledge, terminologies, and objectives to align technical and business perspectives. By identifying the knowledge units required, the mechanisms that enable exchange, and the challenges that hinder it, the study applies and contextualizes knowledge boundary theory in a new analytic domain—showing how process, IS, and process mining experts navigate cross-domain differences during process mining pre-analysis. It contributes to process mining and knowledge management research by clarifying how these interactions shape collaboration outcomes and by highlighting pre-analysis as a critical determinant of analytic validity in process mining initiatives. Understanding these dynamics is essential for ensuring that organizations can reliably generate insights from process mining, making this stage both a practical success factor and a theoretically relevant context for studying collaborative knowledge work.

Limitations

While this study offers valuable insights, several limitations should be noted. First, the relatively small sample size of interviewees may not capture the full diversity of pre-analysis practices and may limit generalizability (Gerring 2004). Although participants represented multiple expert profiles within the same organization, broader studies across different industries and organizational sizes are needed to validate and refine our findings.

Second, the challenge of participant imbalance exists in our study. Most participants held hybrid profiles, and only one identified solely as a process owner, while none identified exclusively as IS experts. While this limited the presence of participants representing distinct profiles, we addressed it by asking not only about participants' own areas of expertise but also about the knowledge others sought from them. This approach allowed us to capture a broader understanding of knowledge exchange across profiles. Still, future research could further explore the fluidity and multidimensionality of these profiles by engaging participants with more narrowly defined specializations.

Third, since the study relies solely on interview responses, there is a risk of self-reporting bias. Participants with broad expertise may have overstated the value of their own knowledge domains or framed their contributions in an overly favorable light. To mitigate this, we probed for specific examples and asked about knowledge used by others. Nonetheless, the absence of data triangulation with other sources—such as existing literature on pre-analysis knowledge units, mechanisms of knowledge exchange, or related challenges—remains a limitation. Moreover, our initial exploration of the field revealed no prior literature on knowledge units for pre-analysis, not even within the broader domain of process mining. Observational studies or third-party evaluations would therefore provide stronger triangulation in future research.

Finally, four participants declined to be audio recorded, which limited the availability of verbatim quotations from their interviews. Detailed notes from these sessions were taken and coded using the same analytical framework as the transcribed interviews to ensure their perspectives were included in the thematic synthesis. The absence of direct quotations from these participants, however, may have constrained the representation of their voices in the reported findings.

Nonetheless, this study provides a foundation for advancing research on expert knowledge in process mining—particularly in the pre-analysis stage—and highlights the need for ongoing empirical and theoretical exploration in this domain.

Conclusion

This study examined the expert knowledge required during the pre-analysis phase of process mining, focusing on three expert profiles—process owners, IS experts, and process analysts. Through qualitative case studies, we identified distinct types of knowledge and revealed how these are exchanged through both formal and informal mechanisms. We also surfaced persistent challenges related to stakeholder misalignment, gaps in communication, missing or inaccessible knowledge, and system complexity.

Future research can build on these findings in several ways. First, the often-overlooked but vital step of scoping and setting objectives highlights the need to explore how experts with multiple areas of expertise affect strategic alignment before analysis begins. Second, the dominance of tacit knowledge and informal exchange mechanisms highlights the value of studying how organizations can institutionalize these exchanges without losing their agility. This could involve exploring lightweight documentation methods, team rituals, or digital collaboration tools tailored for pre-analysis contexts.

Third, to support such institutionalization efforts, especially across diverse settings, future work should consider the contextual nature of knowledge exchange challenges.

For example, comparative studies across organizational maturity levels, industries, or system landscapes could help identify which knowledge exchange mechanisms and underlying practices generalize and which remain context-specific. In particular, the role of organizational culture in shaping trust, communication norms, and knowledge ownership warrants deeper examination. Additionally, frameworks that assess knowledge quality—not only in terms of accuracy but also usability across expert boundaries—could provide valuable insights for teams engaged in process mining.

Appendix A Standardized form for interviewee information

Table A1 Description of the fields of the standardized form of interviewee information

Field	Description
Interview ID	A unique identifier of an interview.
Date	The date of the interview.
Start time	The start time of the interview.
End time	The end time of the interview.
Duration	The total interview duration, calculated with the formula, *End time - Start time*.
Name	The name of the participant.
Industry	The industry type in which the participant is employed.
Company	The name of the company in which the participant is employed.
Department	The internal company department in which the participant is employed.
Role	The participant's role at the company where they are employed.
Functional expertise	A multiple-choice question with the following options: *process owner*, *IS expert*, *process analyst*
Academic background	The highest academic degree held by the participant.
Years of experience	The total number of years of work experience of the participant (irrespective of the duration of their current employment).
Consent for audio recording	A Boolean value with a yes/no option.
Consent for publication	A Boolean value with a yes/no option.

Appendix B Interview questions

1. Can you tell me about your day-to-day in terms of process mining, specifically the pre-analysis stage?
2. (for each role) What kind of knowledge or skills do you think are essential for you to perform your tasks effectively?
3. In your tasks, do you need information or knowledge from other parties? If so, what kind of information or knowledge do you require?
4. Does anyone else come to you asking for certain information or knowledge?
5. How do process mining initiatives start at your organization? Who decides to start a process mining initiative, and what is the motivation behind this decision? Once the decision is made, what are the actual steps taken to implement process mining? What does the collaboration between the stakeholders look like?
6. How are event logs built? How do you identify the relevant data? Please walk me through this process.
7. What are the results of your tasks and who do you hand it over to?
8. In terms of the challenges that you face, what are them?
9. Did you begin with both/all expert profiles simultaneously, or did you start with one and adopt the others later? What prompted you to adopt the other profiles, and how has your knowledge evolved over the years?

10. With the overall goal of the study in mind, do you think there is anything I may have overlooked or missed asking?

Appendix C Coding data structures

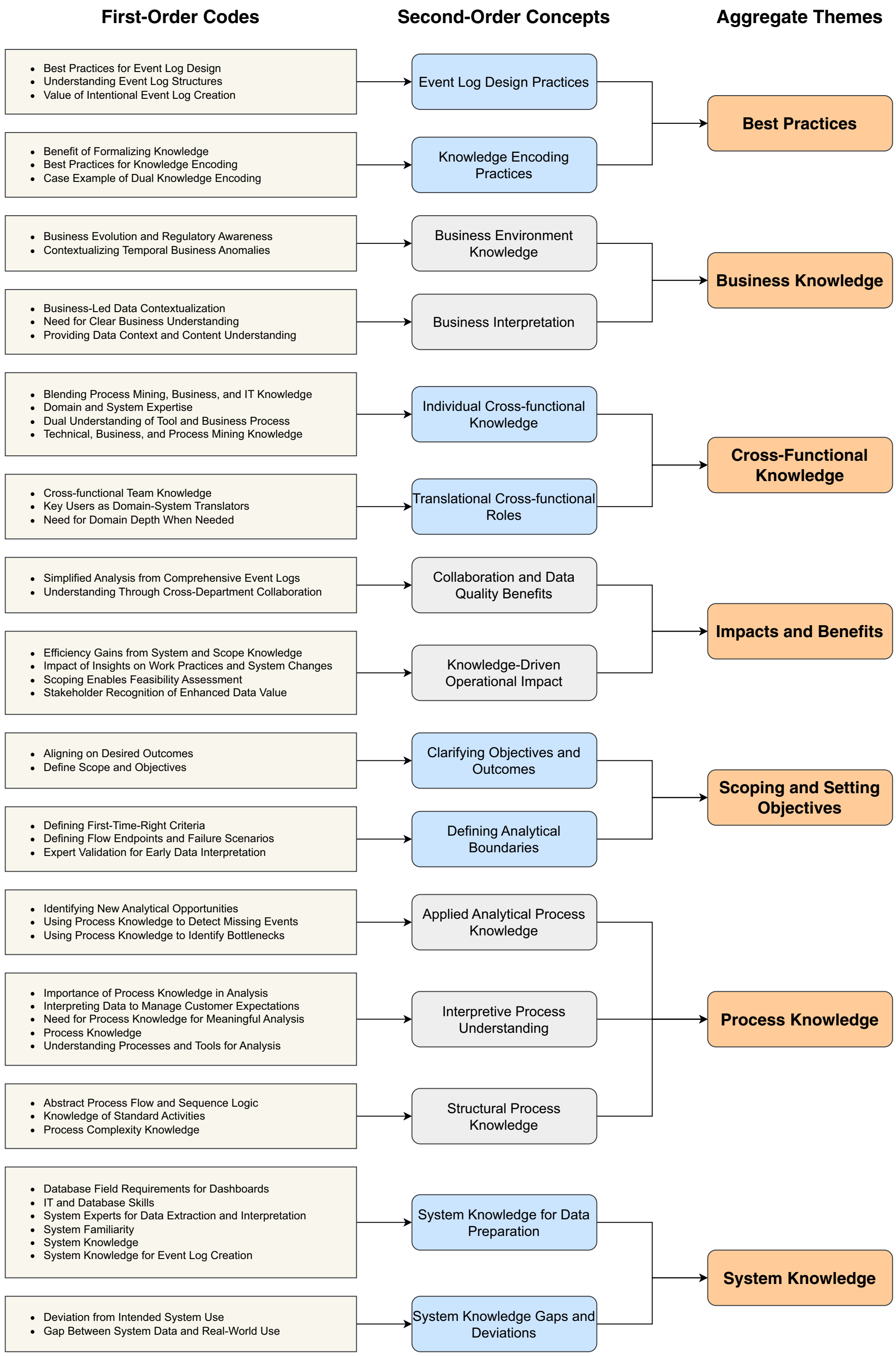

Fig. C1 Coding data structure for knowledge units

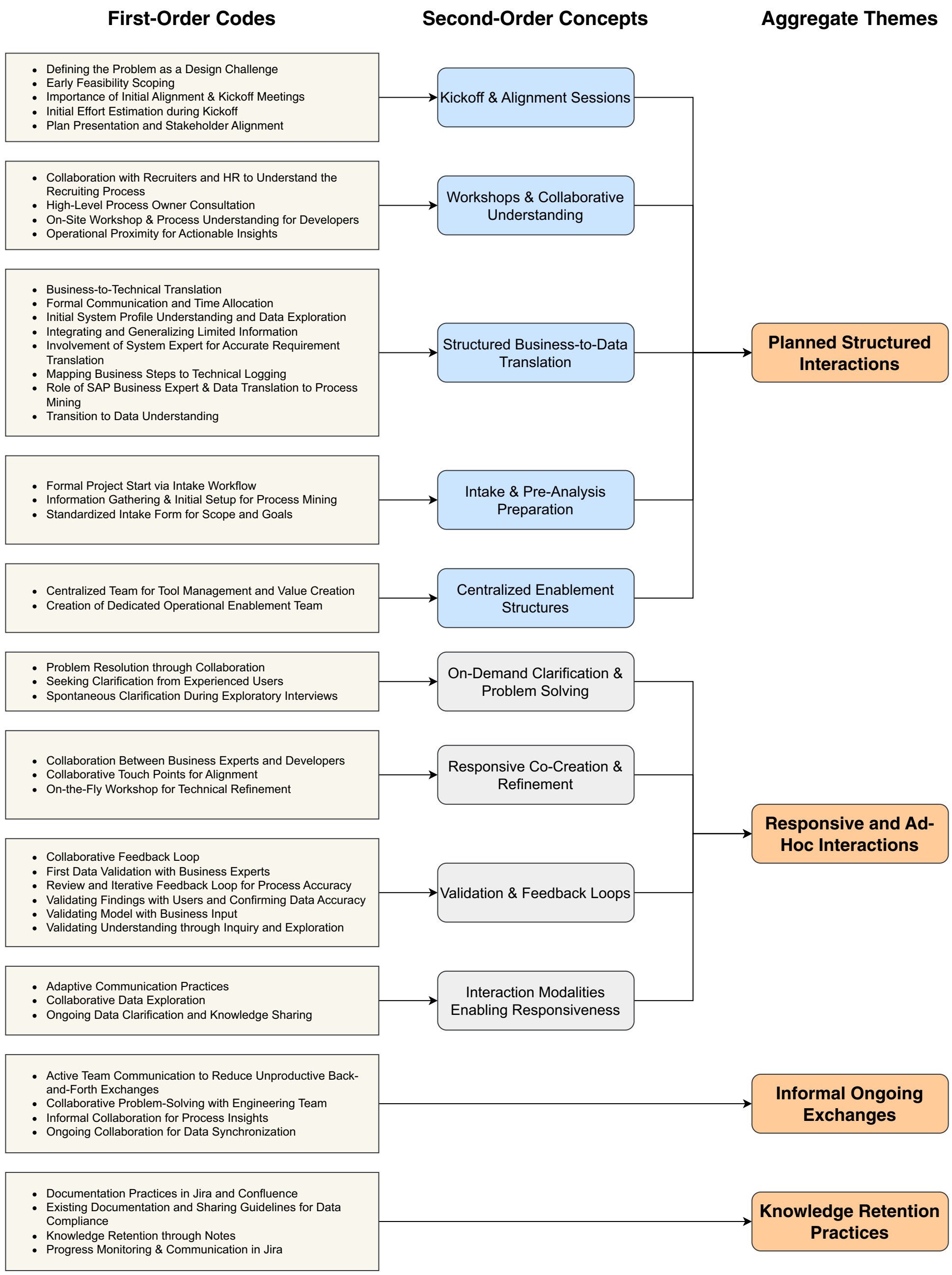

Fig. C2 Coding data structure for knowledge exchange mechanisms

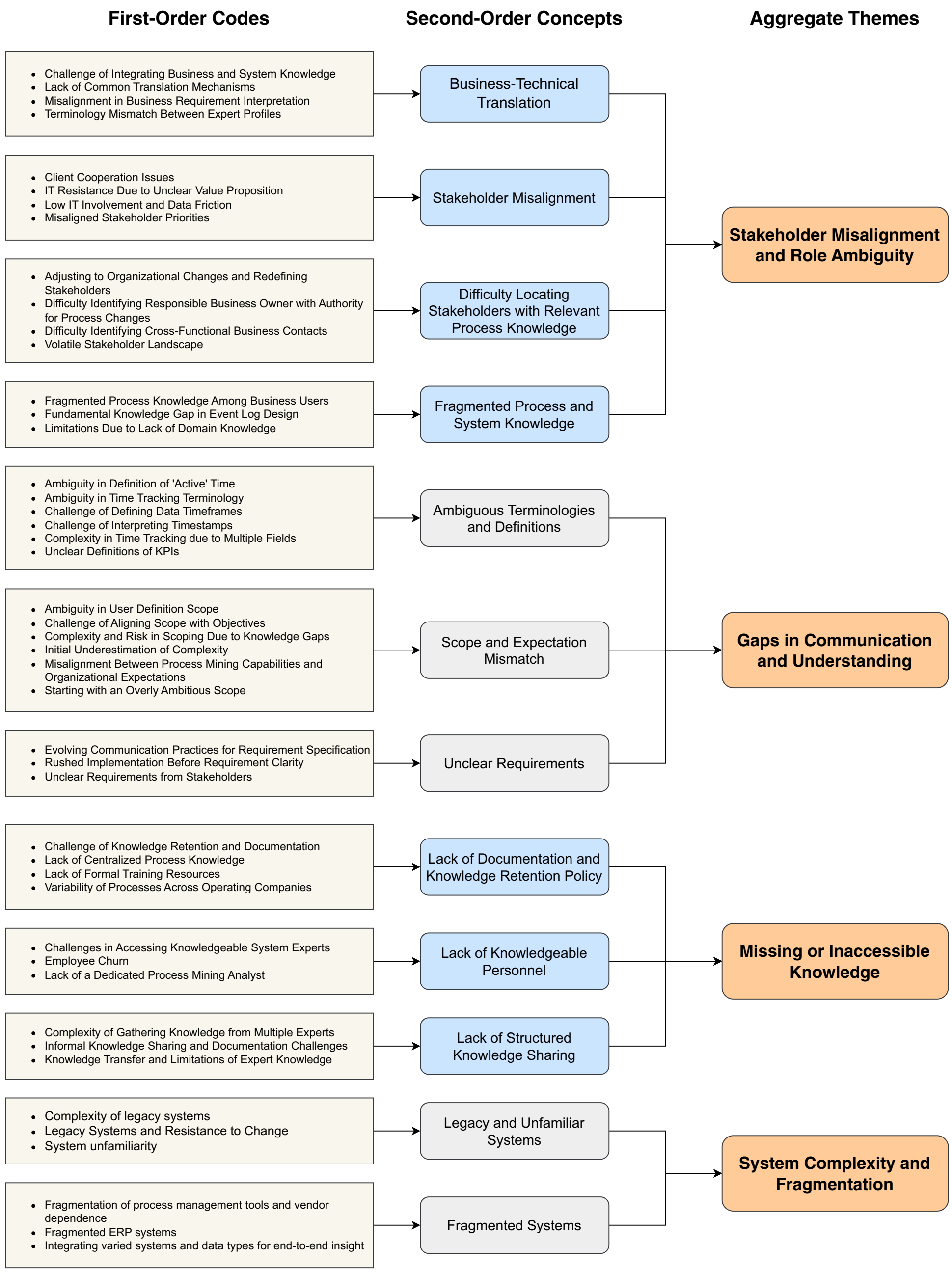

Fig. C3 Coding data structure for knowledge exchange challenges

Author contributions
S.P. wrote the main manuscript text and improved it based on the feedback from M.J. and N.M. All authors reviewed the manuscript.

Funding
This study was supported by the Special Research Fund (BOF) ofHasselt University under Grant Nos. BOF21OWB22 and BOF24TT02.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethical approval
Ethics approval was received from Hasselt University Sociaal-Maatschappelijke Ethische Commissie (SMEC).

Consent to participate
All participants provided their informed consent to take part in this study. Participation was entirely voluntary, and all individuals were informed about the purpose, procedures, and their rights, including the right to withdraw at any time without consequence.

Competing interests
The authors declare no competing interests.

Received: 5 June 2025 / Accepted: 25 November 2025
Published online: 09 December 2025

References

Aalst WMP (2022) Process mining: a 360 degree overview. In: Aalst WMP, Carmona J (eds) Process mining handbook. Springer, Cham, pp 3–34. https://doi.org/10.1007/978-3-031-08848-3-_1
Adams WC (2015) Conducting semi-structured interviews. In: Handbook of practical program evaluation. John Wiley & Sons, Ltd, pp 492–505. https://doi.org/10.1002/9781119171386.ch19
Al-Dowail M, Al-Hashedi A (2021) Stakeholders-driven process mining method for analyzing emergency department processes. In 2021 International Conference of Modern Trends in Information and Communication Technology Industry (MTICTI), pp 1–5). https://doi.org/10.1109/MTICTI53925.2021.9664783
Basque J, Imbeault C, Pudelko B, Léonard M (2004) Collaborative knowledge modeling between experts and novices: a strategy to support transfer of expertise in an organization. In: Canas AJ, Novak JD, Gonzalez FM (eds.) Proceedings of the First International Conference on Concept Mapping (CMC 2004), vol 1. Universidad Publica de Navarra, Pamplona, Espagne, 75–81). http://cmc.ihmc.us/cmc2004Proceedings/cmc2004%20-%20Vol%201.pdf
Battistella C, De Toni AF, Pillon R (2016) Inter-organisational technology/- knowledge transfer: a framework from critical literature review. The Journal of Technology Transfer 41(5):1195–1234. https://doi.org/10.1007/s10961-015-9418-7
Bhatt GD (2002) Management strategies for individual knowledge and organizational knowledge. Journal of Knowledge Management 6(1):31–39. Publisher: MCB UP Ltd https://doi.org/10.1108/13673270210417673.
Bolisani E, Bratianu C (2018) The elusive definition of knowledge. In: Bolisani E, Bratianu C (eds) Emergent knowledge strategies: strategic thinking in knowledge management. Knowledge management and organizational learning,. Springer, Cham, pp 1–22. https://doi.org/10.1007/978-3-319-60657-6_1
Braun V, Clarke V (2006) Using thematic analysis in psychology. Qualitative Research In Psychology 3(2):77–101. https://doi.org/10.1191/1478088706qp063oa
Brock J, Brennig K, Löhr B, Bartelheimer C, Enzberg S, Dumitrescu R (2024) Improving process mining maturity – from intentions to actions. Business Information Systems Engineering. https://doi.org/10.1007/s12599-024-00882-7
Brocke J, Jans M, Mendling J, Reijers (2021) H.A.: a five-level framework for research on process mining. Business Information Systems Engineering 63(5):483–490. https://doi.org/10.1007/s12599-021-00718-8
Carlile PR (2004) Transferring, translating, and transforming: an integrative framework for managing knowledge across boundaries. Organization Science 15(5):555–568. https://doi.org/10.1287/orsc.1040.0094
Chandra Y, Shang L (2019) Inductive coding. In: Chandra Y, Shang L (eds) Qualitative research using r: a systematic approach. Springer, Singapore, pp 91–106. https://doi.org/10.1007/978-981-13-3170-1_8
Contandriopoulos D, Lemire M, Denis J-L, Tremblay É (2010) Knowledge exchange processes in organizations and policy arenas: a narrative systematic review of the literature. The Milbank Quarterly 88(4):444–483. https://doi.org/10.1111/j.1468-0009.2010.00608.x
Davenport TH, De Long DW, Beers MC (1998) Successful knowledge management projects. Sloan Management Review 39(2):43–57). Num Pages: 15 Place: Cambridge, United States Publisher: Massachusetts Institute of Technology, Cambridge, MA
Deken F, Kleinsmann M, Aurisicchio M, Lauche K, Bracewell R (2012) Tapping into past design experiences: knowledge sharing and creation during novice–expert design consultations. Research In Engineering Design 23(3):203–218. https:
DeWeerdt J, Wynn MT (2022) Foundations of process event data. In: Aalst WMP, Carmona J (eds) Process mining handbook. Springer, Cham, pp 193–211. https://doi.org/10.1007/978-3-031-08848-3_6
Disterer G (2001) Individual and social barriers to knowledge transfer. In Proceedings of the 34th Annual Hawaii International Conference on System Sciences, pp 7). https://doi.org/10.1109/HICSS.2001.927138
Dixit PM, Buijs JCAM, Aalst WMP, Hompes BFA, Buurman J (2017) Using domain knowledge to enhance process mining results. In: Ceravolo P, Rinderle-Ma S (eds) Data-driven process discovery and analysis. Lecture notes in business information processing. Springer, Cham, pp 76–104. https://doi.org/10.1007/978-3-319-53435-0_4
Eck ML, Lu X, Leemans SJJ, Aalst WMP (2015) PM2: a process mining project methodology. In: Zdravkovic J, Kirikova M, Johannesson P (eds) Advanced information systems engineering. Lecture notes in computer science. Springer, Cham, pp 297–313. https://doi.org/10.1007/978-3-319-19069-3_19
Edvardsson IR, Durst S (2013) The benefits of knowledge management in small and medium-sized enterprises. Procedia - social and behavioral sciences 81:351–354. https://doi.org/10.1016/j.sbspro.2013.06.441. World Congress on Administrative and Political Sciences
Eichele S, Hinkelmann K, Spahic-Bogdanovic M (2023) Ontology-driven enhancement of process mining with domain knowledge. In Proceedings of the AAAI 2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAAI-MAKE 2023), vol 3433. CEUR Workshop Proceedings, San Francisco Airport, California, USA
Eppler MJ (2007) Knowledge communication problems between experts and decision makers: an overview and classification. Electronic Journal Of Knowledge Management 5(3):291–300291300). Number: 3
Fazey I, Evely AC, Reed MS, Stringer LC, Kruijsen J, White PCL, Newsham A, Jin L, Cortazzi M, Phillipson J, Blackstock K, Entwistle N, Sheate W, Armstrong F, Blackmore C, Fazey J, Ingram J, Gregson J, Lowe P, Morton S, Trevitt C (2013) Knowledge

exchange: a review and research agenda for environmental management. Environmental Conservation 40(1):19–36. https://doi.org/10.1017/S037689291200029X. . Publisher: Cambridge University Press
Galagan PA (1997) Smart companies. Training and Development 51(12):20–24). ERIC Number: EJ555022
Gerring J (2004) What is a case study and what is it good for? American Political Science Review 98(2):341–354. https://doi.org/10.1017/S0003055404001182
Gibbs GR (2007) Thematic coding and categorizing. In: Analyzing Qualitative Data. SAGE Publications, Ltd, London, UK, pp 38–56. https://doi.org/10.4135/9781849208574
Gioia DA, Corley KG, Hamilton AL (2013) Seeking qualitative rigor in inductive research: notes on the Gioia methodology. Organizational Research Methods 16(1):15–31. https://doi.org/10.1177/1094428112452151. Publisher: SAGE Publications Inc
Goodbrand PT, Deng C, Turner N, Uggerslev KL, Gordon J, Martin K, McClelland CR (2021) Exploring safety knowledge sharing among experienced and novice workers. Journal Of Safety Research 79:125–134. https://doi.org/10.1016/j.jsr.2021.08.013
Gorman ME (2002) Types of knowledge and their roles in technology transfer. The Journal of Technology Transfer 27(3):219–231. https://doi.org/10.1023/A:1015672119590
Jong T, Ferguson-Hessler MGM (1996) Types and qualities of knowledge. Educational Psychologist 31(2):105–113. https://doi.org/10.1207/s15326985ep3102_2
Mamudu A, Bandara W, Wynn MT, Leemans, Dijkman R (2022) S.J.J.: a process mining success factors Model. In: Di Ciccio C, Río Ortega A, Rinderle-Ma S (eds) Business process management. Lecture notes in computer science. Springer, Cham, pp 143–160. https://doi.org/10.1007/978-3-031-16103-2_12
Mitton C, Adair CE, Mckenzie E, Patten SB, Perry BW (2007) Knowledge transfer and exchange: review and synthesis of the literature. The Milbank Quarterly 85(4):729–768. https://doi.org/10.1111/j.1468-0009.2007.005.06.x
Nidhra S, Yanamadala M, Afzal W, Torkar R (2013) Knowledge transfer challenges and mitigation strategies in global software development—a systematic literature review and industrial validation. International Journal Of Information Management 33(2):333–355. https://doi.org/10.1016/j.ijinfomgt.2012.11.004
Nissen ME (2005) Harnessing knowledge dynamics: principled organizational knowing & learning: principled organizational knowing & learning. IGI Global
Noble H, Mitchell G (2016) What is grounded theory? Evidence-Based Nursing 19(2):34–35. https://doi.org/10.1136/eb-2016-102306. . Publisher: Royal College of Nursing Section: Research made simple
Nonaka I, Takeuchi H (1995) The knowledge-creating company: How Japanese companies create the dynamics of innovation. Oxford University Press. https://doi.org/10.1093/oso/9780195092691.001.0001
Oliver R (2001) Developing e-learning environments that support knowledge construction in higher education. Research Outputs Pre 2011
Pradhan SK, Jans M, Martin N (2025) Getting the data in shape for your process mining analysis: an in-depth analysis of the pre-analysis stage. ACM Comput. Surv. https://doi.org/10.1145/3712587
Prihodova L, Guerin S, Tunney C, Kernohan WG (2019) Key components of knowledge transfer and exchange in health services research: findings from a systematic scoping review. Journal Of Advanced Nursing 75(2):313–326. https://doi.org/10.1111/jan.13836
Randeree E (2006) Knowledge management: securing the future. Journal Of Knowledge Management 10(4):145–156
Reinkemeyer L (ed.) (2020) Process mining in action: principles, use cases and outlook. Springer, Cham. https://doi.org/10.1007/978-3-030-40172-6
Rowley J (1999) What is knowledge management? Library Management 20(8):416–420. https://doi.org/10.1108/01435129910291175. Publisher: MCB UP Ltd
Shannon CE (1948) A mathematical theory of communication. The Bell System Technical Journal 27(3):379–423. https://doi.org/10.1002/j.1538-7305.1948.t.b01338.x.
Shongwe MM (2016) An analysis of knowledge management lifecycle frameworks: towards a unified framework. Electronic Journal of Knowledge Management 14(3):140–153). Number: 3
Smith EA (2001) The role of tacit and explicit knowledge in the workplace. Journal of Knowledge Management 5(4):311–321. https://doi.org/10.1108/13673270110411733. Publisher: MCB UP Ltd
Stein Dani V, Leopold H, Werf JMEM, Lu X, Beerepoot I, Koorn JJ, Reijers HA (2022) Towards understanding the role of the human inEvent log extraction. In: Marrella A, Weber B (eds) Business process ManagementWorkshops. Lecture notes in business information processing. Springer, Cham, pp 86–98. https://doi.org/10.1007/978-3-030-94343-1_7
Suri H (2011) Purposeful sampling in qualitative research synthesis. Qualitative Research Journal 11(2):63–75. https://doi.org/10.3316/QRJ1102063. Publisher: Emerald Group Publishing Limited
Suriadi S, Wynn MT, Ouyang C, Hofstede AHM, Dijk NJ (2013) Understanding process behaviours in a large insurance company in Australia: a case study. In: Salinesi C, Norrie MC, Pastor Ó (eds) Advanced information systems engineering. Springer, Berlin, Heidelberg, pp 449–464. https://doi.org/10.1007/978-3-642-38709-8_29
Taherdoost H (2022) How to conduct an effective Interview; a Guide to interview design in research study, Rochester, NY. https://papers.ssrn.com/abst ract=4178687
Van Eerd D (2019) Knowledge transfer and exchange in health and safety: a rapid review. Policy And Practice In Health And Safety 17(1):54–77. https://doi.org/10.1080/14773996.2018.1508116. Publisher: Taylor & Francis
Wit-de Vries E, Dolfsma WA, Windt HJ, Gerkema MP (2019) Knowledge transfer in university–industry research partnerships: a review. The Journal Of Technology Transfer 44(4):1236–1255. https://doi.org/10.1007/s10961-018-9660-x
Zagzebski L (2017) What is knowledge? In: The Blackwell Guide to epistemology. John Wiley & Sons, Ltd, pp 92–116. https://doi.org/10.1002/9781405164863.ch3
Zimmermann L, Zerbato F, Weber B (2023) What makes life for process mining analysts difficult? A reflection of challenges. Software And Systems Modeling. https://doi.org/10.1007/s10270-023-01134-0

Publisher's Note