Information Systems 140 (2026) 102715

Contents lists available at ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/is

Domain experts in the loop: Leveraging generative artificial intelligence for interactive data validation in process mining☆

Julian Armin Dormehl [a,b,c,*], Robert Andrews [d], Wolfgang Kratsch [e,f,g], Maximilian Röglinger [a,b,c], Moe Thandar Wynn [d], Felix Zetzsche [a,b,c]

[a] FIM Research Center for Information Management, Wittelsbacherring 10, Bayreuth, 95444, Bavaria, Germany
[b] University of Bayreuth, Wittelsbacherring 10, Bayreuth, 95444, Bavaria, Germany
[c] Fraunhofer FIT, Wittelsbacherring 10, Bayreuth, 95444, Bavaria, Germany
[d] Queensland University of Technology, 2 George St, Brisbane City, 4000, Queensland, Australia
[e] FIM Research Center for Information Management, Alter Postweg 101, Augsburg, 86159, Bavaria, Germany
[f] Augsburg Technical University of Applied Sciences, Alter Postweg 101, Augsburg, 86159, Bavaria, Germany
[g] Fraunhofer FIT, Alter Postweg 101, Augsburg, 86159, Bavaria, Germany

ARTICLE INFO

Keywords:
Process mining
Event data quality
Data validation
Generative artificial intelligence
Domain knowledge
Design science research
Fitness for purpose

ABSTRACT

Process mining analyzes process execution data to derive insights that support operational process improvement. However, event logs often suffer from poor data quality, typically resulting from process deficiencies, which can lead to inaccurate or misleading insights. To mitigate this risk, domain experts and process analysts engage in data validation during event data preparation to assess whether an event log is fit for its intended analytical purpose. Yet, current practices often fail to sufficiently align event logs with their analytical objectives, commonly formalized as analysis questions. This misalignment impedes the detection of data quality issues, which frequently vary across application domains and analytical contexts. Generative artificial intelligence offers promising capabilities in this regard, including adaptability to diverse contexts, the ability to interpret complex data, and the generation of context-aware recommendations. To leverage this potential, we adopt the Design Science Research paradigm to iteratively develop Artificial Intelligence-Assisted Data Validation For Domain Experts (AID4DE) that integrates domain knowledge — rooted in experts' practical engagement with operational processes — with generative artificial intelligence support to facilitate interaction with complex event log data. We instantiate AID4DE as an open-source software prototype and evaluate it through a three-phase approach: a competing artifact analysis, 14 semi-structured expert interviews, and a user study involving 18 information systems researchers. Our results show that AID4DE is both applicable and effective in supporting domain experts in data validation, enabling the systematic externalization of domain knowledge and rigorous assessment of event log's fitness for purpose.

1. Introduction

Effective business processes are fundamental to organizational operations, as they determine how value is created [1,2]. The field of Business Process Management (BPM) provides methods, techniques, and tools to model, analyze, automate, and monitor these processes [1–3]. Within this context, Process Mining (PM) has emerged as a data-driven approach that extracts insights from event logs generated by information systems (IS) [1,3]. Unlike model-centric BPM, PM offers empirical evidence of actual process executions, supporting decision-making based on observed behavior [3,4].

The quality of input data, particularly event logs, critically affects the validity and usefulness of PM insights [4–6]. Event logs often contain irrelevant, imprecise, or missing data due to limitations in logging practices [7]. Using low-quality logs can distort analysis and lead to incorrect conclusions [5–9]. Consequently, event data preparation, which identifies and corrects such issues, is essential but resource-intensive, often accounting for up to 80% of PM project effort [5,6,10,11]. This effort varies across domains and depends on analytical objectives, frequently requiring context-sensitive approaches, stakeholder consultation, and manual data manipulation [5,8,11–13]. To reduce this burden, methods based on Artificial Intelligence (AI) are increasingly

☆ This article is part of a Special issue entitled: 'AI-Enhanced BPM' published in Information Systems.

* Corresponding author.
E-mail address: julian.dormehl@fit.fraunhofer.de (J.A. Dormehl).

https://doi.org/10.1016/j.is.2026.102715
Received 30 June 2025; Received in revised form 26 January 2026; Accepted 3 March 2026
Available online 4 March 2026

applied to detect and correct data quality issues adaptively, leveraging data-driven learning mechanisms that surpass traditional rule-based techniques [14,15].

Data validation is the final step of event data preparation and is typically performed collaboratively by process analysts and domain experts. This collaboration mitigates the risk of data quality issues compromising PM outcomes by enabling a more context-aware evaluation of data quality [5,9,16,17]. Analysts contribute expertise in data preparation and PM tools [18,19], while domain experts provide practical knowledge of process execution that extends beyond the recorded data [8,20,21]. Together, they ensure that event logs reflect real-world processes and identify areas requiring further refinement [9,16].

Despite its importance, data validation in PM remains underexplored [19]. Logs can be assessed along two dimensions: *fitness for PM* and *fitness for purpose*. The former evaluates technical and structural suitability for analysis [22,23], while the latter assesses alignment with analytical objectives and relevant process aspects [8]. Existing approaches predominantly focus on fitness for PM, leaving fitness for purpose challenging due to difficulties in translating tacit domain knowledge into actionable assessments and bridging communication gaps between domain experts and analysts [18,19,24]. Several methods support expert interaction with event logs [25–27], and others detect and correct log issues [28], often leveraging AI [29]. However, systematic evaluations of log alignment with project objectives, formalized as analysis questions (AQs), are still lacking [5,13,30]. As a result, undetected data quality issues can compromise PM outcomes [5,10,11,28,31]. This study therefore focuses on fitness for purpose, aiming to enable domain experts to evaluate event logs against analytical objectives and improve the reliability and relevance of PM analyses.

To support this evaluation, we propose leveraging Generative AI (GenAI), which can interpret complex patterns, generate comprehensible representations, and provide context-aware recommendations [32]. Our research question (RQ) is: *How can GenAI support domain experts in validating the fitness for purpose of event logs for PM?*

We adopt the Design Science Research (DSR) paradigm [33] and conceptualize **AI**-Assisted **D**ata Validation **For D**omain **E**xperts (AID4DE), a system designed to support domain experts in aligning event logs with their intended analytical purpose. AID4DE integrates visual and textual representations of event log characteristics, enables interactive engagement, and offers context-sensitive guidance. Its objective is to improve data quality by enabling the identification of discrepancies between the event log and the underlying AQ, serving as a first step toward their resolution. The system is instantiated as an open-source prototype and evaluated following the Framework for Evaluation in Design Science (FEDS) [34]. Evaluation comprises three stages: (1) benchmarking and refinement of the design specification through seven semi-structured interviews with domain experts in the PM field, (2) assessment of the conceptual model via seven additional expert interviews, and (3) a simulated data validation session with 18 IS researchers.

This paper makes two primary contributions. First, it introduces an artifact supporting assessment of event log's fitness for purpose — an underexplored aspect of PM data validation. Second, it facilitates the externalization of tacit domain knowledge, improving data quality and supporting subsequent PM analyses. To our knowledge, this is the first approach combining domain expertise with GenAI for this purpose.

The remainder of the paper is organized as follows. Section 2 reviews relevant literature. Section 3 outlines the research approach. Section 4 presents AID4DE, followed by Section 5, which reports its evaluation. Section 6 discusses implications, and Section 7 concludes the paper.

2. Background and related work

2.1. Background

PM encompasses a set of data-driven analysis techniques that support BPM by leveraging event logs to gain insights into the actual execution of business processes [1,3,4,35]. PM projects typically follow a lifecycle, such as the PM^2 methodology proposed by van Eck et al. [35], which begins with planning and data extraction. In this initial phase, the project scope is defined, an AQ is formulated, and relevant event data is extracted from source systems [10,35,36]. Subsequently, projects iterate through event data preparation, mining and analysis, and result evaluation [35]. These iterations apply PM techniques to address the AQ, for example by discovering process models, assessing conformance, or identifying improvement opportunities through advanced analytics [35]. Once the AQ is sufficiently addressed, findings are consolidated to support implementation, enable continuous monitoring, and inform evidence-based decision-making [3,35].

During the planning stage, the processes to be analyzed are identified, and an AQ is formulated in alignment with the overarching objectives of the PM project [35]. Typical motivations include improving process performance or verifying compliance with organizational rules [35,37]. AQs may be descriptive, comparative, explanatory, or suggestive and can focus on various process dimensions, such as quality, time, or cost [35,38]. As a critical input for subsequent stages, AQs guide both the extraction of relevant event data and the selection of suitable mining and analysis techniques [35–38]. Defining a clear and feasible AQ is therefore essential for effective event data preparation and analysis.

Before PM techniques can be applied, raw event data must be transformed into an event log representing process executions as ordered sequences of events [4,7]. Each event refers to a specific activity and is associated with a case, i.e., an instance of the process [4,7]. Event logs are commonly stored in the case-centric eXtensible Event Stream (XES) format [39,40], which requires (1) a case identifier (ID), (2) an activity reference, and (3) a temporal ordering of events, typically based on timestamps [39]. In addition to these core components, event logs may include contextual attributes, such as the resource executing an activity, which enhance log richness and support multi-perspective analysis [39]. As an alternative, the Object-Centric Event Log (OCEL) format captures interactions between events and multiple objects, enabling more flexible representations without relying on a predefined case notion [41,42].

Beyond structural correctness, event logs must also be of sufficient quality, as analytical outcomes are directly influenced by the reliability of the underlying data [4–6,10,43]. Poor data quality can compromise PM insights, as emphasized by the *"garbage in, garbage out"* principle. Accordingly, event data preparation aims to transform raw data into event logs of sufficient quality that are both structurally sound and appropriate for addressing the defined AQ [5,10,35,39]. The first dimension, *fitness for PM*, captures the extent to which an event log fulfills the structural prerequisites required by PM algorithms [5,22,39]. The second, *fitness for purpose*, concerns the degree to which an event log accurately reflects relevant aspects of the underlying business processes and supports the analytical objectives formalized as an AQ [5,8,13,30].

Ensuring that event logs are both *fit for PM* and *fit for purpose* requires a range of tasks during event data preparation, typically executed by data engineers and process analysts. Data engineers are responsible for extracting event data from source systems and ensuring technical completeness, while process analysts guide preparation and perform subsequent analysis using PM techniques [1]. Common tasks include log filtering, abstraction, and transformation [44]. Log filtering reduces complexity by removing irrelevant traces or events and facilitates the detection of noise and data quality issues through attribute- or variance-based methods [5,35,45]. Log abstraction aggregates events to an appropriate level of granularity, aligning event detail with analytical objectives [39,46]. Log transformation modifies event data to enhance its *fitness for PM*. Due to their diversity and complexity, these activities are often time-consuming and require repeated interaction with domain experts [5,8,11,13].

The involvement of domain experts is essential to ensure that event logs accurately reflect real-world process executions and support the defined analytical objectives [1,8,35]. Their practical experience provides contextual knowledge that complements standardized preparation procedures, enabling a nuanced understanding of actual process behavior and serving as a reference against recorded event data [7,8,20,21,24,47,48]. Such knowledge may be explicit, having been externalized through communication or documentation, or tacit, residing in internalized cognitive processes not yet formally articulated [18,49]. Incorporating domain knowledge supports accurate interpretation of the AQ and enables the identification of data quality issues by contrasting expert expectations with event data, thereby reducing the risk of biased analyses and misguided improvement decisions [8,10,11,20,21,36,50].

Data validation usually represents the final step of event data preparation and is a domain knowledge-intensive activity aimed at determining whether an event log requires further refinement before applying PM analyses [9,16]. To support this decision, process analysts and domain experts jointly assess whether the event log is valid for its intended analytical purpose and adequately reflects actual process behavior [5,9,16,17]. Analysts contribute expertise in data quality and PM requirements, while domain experts provide the contextual understanding needed to interpret recorded data [11,51]. This collaboration establishes trust in the event log's quality and analytical relevance, forming a robust foundation for subsequent mining and analysis [9,16].

The execution of data validation generally involves four iterative tasks. First, the process analyst introduces the domain expert to PM principles and outlines the objectives of the validation effort. To support contextual understanding, raw event data is presented along with explanations of its provenance and structure, and the overarching AQ is discussed [9]. Second, overview statistics are generated to help the domain expert develop an initial understanding of the event log [9]. Depending on the technology stack, data validation dashboards may provide automatically generated visual representations [52]. Third, the analyst and domain expert collaboratively explore the data in greater depth by applying filters or examining subsets to identify emerging patterns [51]. When anomalies are observed, such as atypical process paths, unusual case durations, or violations of known business rules, specific instances are examined to uncover potential data quality issues and their causes [9,16]. Fourth, insights regarding data quality are systematically documented [51]. If the event log proves inadequate to address the AQ, validation outcomes inform further iterations of event data preparation [5,53].

The emergence of GenAI offers new opportunities to support data validation due to its ability to learn from data, generalize to new contexts, and operate effectively in data-intensive environments. GenAI, a subfield of AI grounded in Machine Learning, involves systems that learn from data rather than relying on explicitly programmed rules [54,55]. More specifically, GenAI is largely enabled by Deep Learning, which employs multi-layered neural networks to model complex data structures [56]. By learning patterns from large-scale datasets, GenAI systems can autonomously generate new content, such as text, images, or code, thereby supporting the automation and augmentation of knowledge-intensive tasks [32,47,57]. The technical foundation of GenAI primarily rests on the transformer architecture underpinning most Large Language Models (LLMs), which relies on self-attention mechanisms and embeddings to capture context, semantics, and long-range dependencies [32,58]. Through extensive pretraining, LLMs achieve strong generalization across understanding and generation tasks, positioning them as promising tools for data validation [47,59].

2.2. Related work

Despite its importance for ensuring valid analytical outcomes, no standardized framework for data validation in PM currently exists. Outside PM, van der Loo and de Jonge [17] propose a validation approach implemented in the R package 'validate', enabling analysts to specify and verify rules based on domain knowledge. In relational databases, similar mechanisms are realized through integrity constraints, such as 'CHECK' constraints, which enforce consistency at the schema level [60]. Within PM, Andrews et al. [53] propose a validation approach based on the Cross-Industry Standard Process for Data Mining, iteratively combining data quality assessment, preliminary PM, and evaluation to detect remaining issues. Complementarily, Martin et al. [30] introduce an interactive preparation framework that conceptualizes validation as a bidirectional process integrating data-centric assessment with model-driven discovery.

In addition to the lack of an established framework, current data validation practice in PM faces several challenges [19]. Effective validation requires analysts and domain experts to understand the event log within its business context [4]. However, asymmetries in expertise often hinder collaboration, complicating assessments of whether event data accurately reflects real process executions [18,19]. This disconnect increases the risk of misaligned interpretations and misleading analytical outcomes. LLMs offer a promising means to bridge this gap by reasoning about data without rigid rules, communicating findings in natural language, and adapting to new domains through prompting or fine-tuning [32,47,59]. In this context, Brützke et al. [61] demonstrate that GenAI can support users in PM analyses. Accordingly, GenAI-based support may enable domain experts to better assess whether event logs are semantically aligned with analytical objectives.

Although the application of GenAI for data validation in PM remains largely underexplored, initial studies demonstrate its potential within event data preparation. Nguyen et al. [62] use autoencoders to detect anomalies and repair missing attribute values, while Wu et al. [63] employ masked transformer-based models to restore missing activities. Schmid et al. [64] address identical timestamp errors using a Generative Adversarial Network. Moving beyond fully automated techniques, Zetzsche et al. [29] introduce a hybrid human-in-the-loop approach that combines rule-based methods with transformer architectures to address the elusive case data quality issue. Beyond GenAI, several approaches leverage domain knowledge to improve event data preparation [65], including gamification-based correction of activity labels [28], enrichment of event logs with unstructured text and expert feedback [50], and knowledge graph-based log construction and enrichment [66]. Additional heuristics informed by domain knowledge address missing events [67,68], incorrect timestamps [69], and deviating event sequences [70]. Despite these efforts, existing work primarily targets structural detection and repair and does not enable domain experts to evaluate the *fitness for purpose* of event logs.

3. Research design

To address the formulated RQ, we adopt the established DSR paradigm, which focuses on creating purposeful artifacts that solve real-world problems and contribute to design knowledge [71–73]. This study addresses two main challenges: (1) asymmetries in expertise between domain experts and process analysts during data validation, and (2) the limited consideration of event log fitness for purpose in current validation practices. To bridge these gaps, we propose a conceptual model for AI-assisted event data validation that supports domain experts. Following the notion of exaptation [71], existing AI technologies are repurposed for a new application context. The conceptual model represents the core solution artifact and is instantiated through a software prototype, aligning with DSR's emphasis on artifact realization [72,74]. Given its socio-technical nature, development proceeded through several iterative cycles [75], making DSR an appropriate framework due to its iterative and problem-solving orientation [72,73].

We follow the six-step DSR methodology proposed by Peffers et al. [33], illustrated in Fig. 1. Section 1 introduced and motivated the research problem (Step 1), while Section 2 reviewed related work to refine the research challenge. Drawing on data validation practice and

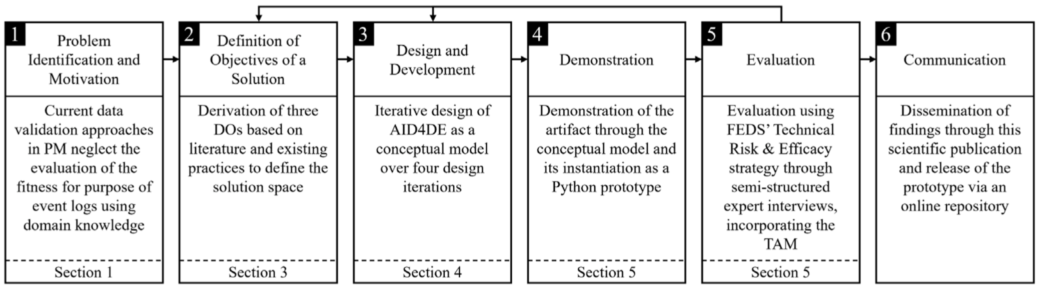


Fig. 1. Overview of the research activities based on the DSR methodology by Peffers et al. [33].

literature on PM and visual analytics (VA), we derived three design objectives (DOs) to delineate the solution space and guide artifact development (Step 2) [72]. The artifact was designed iteratively (Step 3), specified in Section 4, and instantiated as a software prototype (Step 4). Design and evaluation activities (Step 5) were closely integrated using the FEDS framework [34] to ensure practical relevance. To the best of our knowledge, this is the first study to address AI-assisted data validation tailored for domain experts. Finally, this paper communicates the results (Step 6) and makes the prototype's source code publicly available to support transparency and reproducibility.

3.1. Design objectives

The DOs guide artifact development and are derived from the research problem and a systematic literature review [33]. A central challenge in PM initiatives is eliciting and utilizing domain experts' tacit knowledge [4,8]. Accumulated through experience, this knowledge enables experts to contextualize and interpret event data [18, 24,47,76]. To support this process, the artifact provides meaningful visual representations of event logs, enhancing cognitive processing and insight generation [49,77]. As visualization effectiveness depends on design quality and user familiarity [78], combining visual elements with natural language annotations helps reduce cognitive load and improve interpretability [76,78]. Accordingly, we define the first DO:

DO1 (visualize and describe): *AI-assisted event data validation should provide meaningful visualizations of the event log and its process characteristics, alongside clear natural language explanations of relevant terminology.*

Data validation often involves analyzing samples of event logs to develop a detailed understanding of their characteristics [9,51,77,79]. To support multi-perspective exploration, interactive mechanisms are essential. Following the VA principle *"overview first, zoom and filter, then details on demand"* [80], the artifact enables experts to drill down into complex data while accessing supplementary details as needed. This leads to DO2:

DO2 (interact): *AI-assisted event data validation should enable interaction with both the event log and its visual representations.*

Domain experts may lack clarity on how to perform validation, which activities are involved, or what constitutes a satisfactory outcome [81]. Providing context-sensitive guidance throughout the process supports effective validation and mitigates potential obstacles [82]. Such guidance aligns with the prevailing AQ, helping experts assess the event log in relation to its intended analytical purpose [5,30]. Leveraging intrinsic event log characteristics, including anomalies, the artifact supports semantic evaluation, leading to DO3:

DO3 (guide): *AI-assisted event data validation should provide context-sensitive guidance based on the AQ and proactively highlight potential data quality issues.*

3.2. Design and development

The conceptual model constitutes the central solution artifact, synthesizing insights from iterative build and evaluation cycles [71]. It captures the interactive relationship between domain experts and the system underlying AID4DE [83]. Conceptual modeling is well suited for this purpose, as it represents application functionality through entities, attributes, operations, and relationships while directly linking these elements to user tasks [83].

We structured the model design using the six-step methodology for empirically grounded reference architectures by Galster and Avgeriou [84]. Initially (Step a), we developed an industry-agnostic model to ensure broad applicability. Drawing on conceptual models in VA [49,81, 85], we adapted existing structures to inform the design (Step b). Next (Step c), insights from data validation literature (Section 2.1) and the derived DOs guided construction of the AID4DE model. Empirical input was obtained through expert interviews with academics and practitioners (Sections 5.1 and 5.2). We modeled expert–system interactions as actions and states (Step d) and incorporated variation points to enable reuse and customization (Step e). The resulting conceptual model is presented in Section 4, with its evaluation reported in Section 5 (Step f).

Guided by the DOs, we developed the prototype through several iterative cycles, summarized in Table A.6 in Appendix A. The first iteration explored the potential of GenAI to support domain experts during data validation, examining whether LLMs can bridge the gap between complex event logs and non-technical users. Given the limited context windows of LLMs [86], the second iteration focused on mitigating this constraint by generating visual representations that extract representational semantics from the event log [86,87]. These visualizations improved data accessibility for domain experts while enhancing the LLM's contextual understanding. Subsequent iterations refined the GenAI-based support: iteration 3 addressed DO1 and DO2 by extending visualizations, analytical perspectives, and interaction mechanisms, forming the basis for DO3 in iteration 4, where guidance features informed by prior literature were implemented. During instantiation, it became evident that combining representational semantics from the event log with the specified AQ provides essential input for the GenAI-based guidance mechanism.

3.3. Demonstration and evaluation

During the iterative build-and-evaluate cycles, we demonstrated the artifact from the initial design specification to the conceptual model and finally to its Python-based prototype.[1] These demonstrations illustrate the artifact's utility and suitability for addressing the identified challenges [33]. To evaluate feasibility, we implemented the prototype using a ticket management event log from an Italian software company [88].

[1] Link to the AID4DE source code: https://github.com/jul-dor/AID4DE

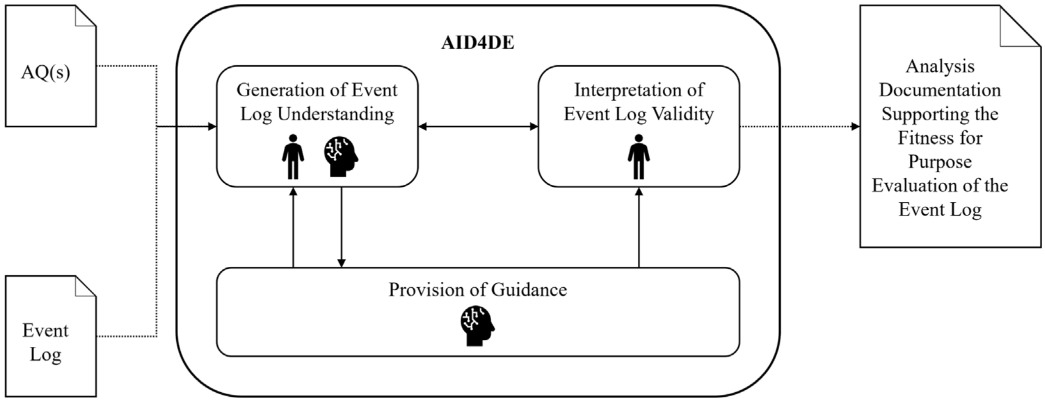


Fig. 2. Overview of AID4DE.

The primary evaluation objective was to identify opportunities to improve the conceptual model early, ensuring effectiveness and reliability in practical scenarios. The greatest uncertainty concerned the LLM component, whose behavior could not be fully anticipated during design. We therefore adopted the Technical Risk and Efficacy evaluation strategy [34], structured into three episodes.

Episode 1: Artificial formative evaluation. Conducted in a controlled setting to refine the artifact during development. Semi-structured interviews with seven academic and industry experts focused on the design specification [89], namely leveraging GenAI to support domain experts in assessing event log fitness for purpose. A comparative analysis of existing artifacts further informed design refinements (see Section 5.1).

Episode 2: Artificial summative evaluation. Assessed the artifact's effectiveness and suitability for real-world scenarios [34]. Seven additional experts evaluated the conceptual model through semi-structured interviews, focusing on completeness, consistency, and robustness for supporting domain expert data validation (see Section 5.2).

Episode 3: Naturalistic summative evaluation. Involved a live demonstration of the prototype using a publicly available help desk dataset [88] for an audience of 18 IS researchers. Participants subsequently completed a survey with six statements assessing perceived usefulness and ease of use based on the Technology Acceptance Model (TAM) [90], providing insights into behavioral intentions toward adopting the AID4DE system (see Section 5.3).

4. Results

4.1. Conceptual model for Artificial-Intelligence-assisted event data validation for domain experts

Fig. 2 depicts the key operational components and their interrelationships within AID4DE. The model relies on two primary inputs: AQ(s), defined during the planning phase of a PM project, and an event log capturing process execution data. The validation process comprises three interrelated components: (1) generation of event log understanding, (2) interpretation of event log validity, and (3) provision of guidance. Their tight integration enables iterative validation cycles and shared responsibilities between domain experts and the system underlying AID4DE. The combined execution of these components produces an analysis documentation supporting the evaluation of the event log's fitness for purpose.

Developing an understanding of the event log is fundamental to AI-assisted data validation. To this end, fulfillment of DO1 (visualize and describe) and DO2 (interact) is required. DO1 is addressed by providing visual representations from multiple analytical perspectives, complemented by natural-language explanations that clarify both visual content and terminology. DO2 is achieved by enabling domain experts to manipulate the underlying event log data and corresponding visualizations, for example through filtering, aggregation, or clustering, while supporting the documentation of insights during exploration. In parallel, the system underlying AID4DE extracts representational semantics, i.e., core information embedded in the event log, to strengthen comprehension and support expert interpretation.

This understanding directly informs the guidance component, represented by DO3 (guide). DO3 is achieved by offering context-aware support that assists domain experts in executing data validation tasks while ensuring alignment with the defined AQ. By integrating event log understanding with guidance provision, the system supports expert-driven interpretation of event log validity and promotes identification of discrepancies between the prevailing event log and the AQs.

Because event log data must be validated from multiple perspectives, interpretation of validity typically triggers further iterations of the validation cycle. When no additional insights emerge, the system generates an analysis report summarizing validation activities and findings, facilitating evaluation of the event log's fitness for its intended purpose.

The following subsections introduce the core components of AID4DE, each forming part of the conceptual model. To ensure clarity and interpretative consistency, three foundational assumptions underpin the exposition. First, following Federico et al. [49], the model distinguishes between two conceptual spaces: the perceptual space, representing the cognitive activities and interactions of the domain expert, and the computational space, encompassing operations performed by the system underlying AID4DE. Second, as a solution artifact, the conceptual model provides an abstract representation of the system architecture, explicitly highlighting expert–system interaction. Third, the model assumes an analysis-ready setting in which the system has already been provided with both the event log and the corresponding AQs. The complete composition of all subcomponents is illustrated in Fig. B.8 in Appendix B.

4.1.1. Generation of event log understanding

The generation of event log understanding constitutes the foundational component of AID4DE. As illustrated in Fig. 3, it comprises two interrelated processes: generation of event log understanding for the domain expert (U1) and for the system (U2). Together, these processes establish a comprehensive understanding of the event log. In all conceptual model illustrations, circles denote components where input or output is accumulated, squares indicate processing components, and arrows represent logical flow. The left side depicts the computational space (system), while the right side represents the perceptual space (domain expert).

Event log understanding for both the domain expert and the system is driven by the expert's tacit knowledge (K^{τ}) and explicit knowledge (K^{ϵ}). These complementary knowledge forms inform exploration

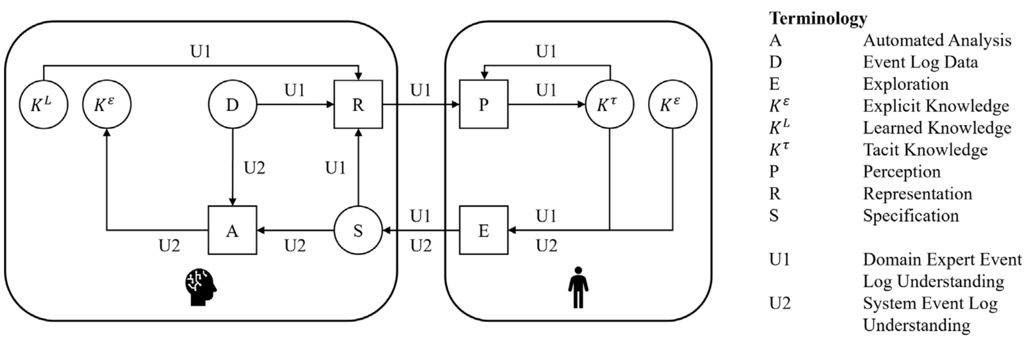


Fig. 3. Generation of event log understanding in AID4DE.

(E) of the event log, encompassing expert interactions with the data. Exploration techniques include filtering, aggregation, and clustering. Filtering narrows the focus to specific log segments, supporting targeted validation [35]. Aggregation derives higher-level metrics to provide condensed overviews [7], while clustering identifies structurally or behaviorally similar groups [45]. The selected exploration method is translated into a specification (S), a machine-readable representation — typically executable code — defining the intended data transformation or analysis. From this point, the generation of understanding diverges into distinct pathways for the domain expert and the system.

In line with DO1, event log understanding for the domain expert (U1) is operationalized through representations (R) derived from the current specification (S) and event log data (D). These include visualizations, tabular views, and natural language descriptions [79,91,92]. Visualizations span chart-based, network-based, matrix-based, hierarchy-based, and timeline-based formats, encompassing both general-purpose and process-specific techniques [30,79,91,92]. For example, chart-based visualizations commonly used in PM include bar charts, line graphs, and heatmaps, while network-based visualizations may depict process models such as Business Process Model and Notation (BPMN) diagrams [79]. These representations encode facets such as temporal patterns, spatial distributions, event relationships, and attribute-specific characteristics [79].

The representations trigger iterative cognitive and perceptual processes (P) within the domain expert, generating insights and fostering the development of new tacit knowledge (K^τ) [49,93]. Effective comprehension requires alignment between visualization choices, encoded data facets, and domain-specific characteristics [91].

The system complements these processes by providing supplementary explanations derived from learned knowledge (K^L). Learned knowledge denotes domain-specific and task-related knowledge acquired through pretraining and, where applicable, fine-tuning of the underlying LLM. In this study, no fine-tuning was performed; all learned knowledge originates from pretraining. Unlike explicit knowledge extracted from the analyzed event log, learned knowledge is not dataset-specific but encodes general regularities, heuristics, and interpretive conventions associated with event logs and process-oriented data. Examples include typical data-validation heuristics (e.g., detecting missing or inconsistent events), commonly observed activity patterns (e.g., frequently observed sequences of activities), and standard interpretations of event sequences (e.g., loops, optional activities, or parallel executions). In practice, this knowledge conditions the system's natural language explanations and guidance based on the current specification and extracted representational semantics, enabling contextualization of event log representations and directing expert attention to relevant aspects, thereby reducing cognitive effort [49]. By providing this support, the system allows domain experts to focus on higher-level reasoning rather than low-level operational details.

The tacit knowledge thus acquired guides further exploration of the event log, leading to adjustments in the specification and representations, as well as renewed cognitive and perceptual processes. This feedback loop supports a closed cycle of cumulative knowledge acquisition regarding the event log data [49].

From the system's perspective (U2), event log understanding is generated through automatic analysis (A), which extracts representational semantics from the event log based on the current specification (S). This produces a condensed summary of key log features that adapts as the specification evolves. The extracted semantics integrate into the system's explicit knowledge (K^ε), stored as contextual input within the LLM embedding space, enabling the system to maintain an updated internal understanding of the event log.

Overall, AID4DE supports the development of a shared understanding of the event log between the domain expert and the system, emerging from interactive exploration and interpretation of event log data.

4.1.2. Interpretation of event log validity

Within AID4DE, the interpretation of event log validity builds on the previously generated event log understanding. As illustrated in Fig. 4, this component comprises two processes: the externalization of domain knowledge (I1) and the extraction of exploration activities (I2). Together, these processes enable domain experts to assess the validity of the event log.

Both the externalization of domain knowledge and the extraction of exploration activities rely on tacit and explicit knowledge. Tacit knowledge (K^τ) stems from prior exploration cycles (U1) and accumulated experience in executing business processes, while explicit knowledge (K^ε) refers to formalized business rules and domain conventions. Integrating newly acquired tacit knowledge with existing tacit and explicit knowledge may lead to differing perceptions of event log validity. These perceptions inform an assessment of the log's fitness for purpose, i.e., the extent to which it accurately represents the underlying business process and supports the analytical objectives defined during planning (cf. Section 2.1).

The externalization of domain knowledge (I1) constitutes a key mechanism for interpreting event log validity. Through externalization (X), tacit and explicit knowledge regarding event log validity is transferred from the domain expert to the system, either qualitatively (e.g., free-text input) or quantitatively (e.g., Likert scales). Externalized knowledge may capture, for instance, the expert's judgment of the event log's fitness for purpose. As externalization occurs alongside ongoing exploration, its content evolves throughout the validation process. Within the system, this knowledge is stored as contextual input in the LLM embeddings, thereby becoming part of the system's explicit knowledge (K^ε).

During event log exploration and externalization, domain experts may develop differing perceptions of the event log's suitability for the intended analysis. Specifically, they may assess how well the log satisfies requirements derived from the analytical objectives. These requirements can be evaluated along several dimensions characterizing fitness for purpose, including:

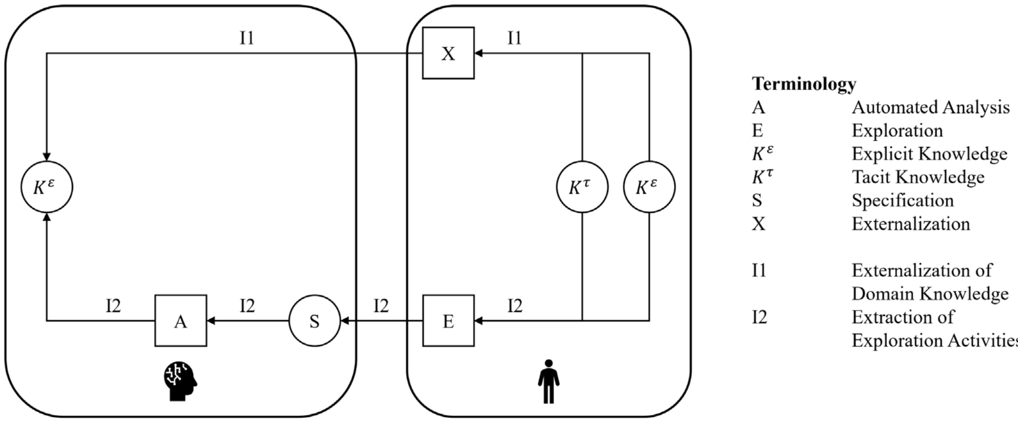


Fig. 4. Conceptual structure of the event log validity interpretation component.

- **Variability**, which reflects the degree of variation within the log, such as structural and temporal differences [7];
- **Volume**, referring to the total number of events and associated attributes in the log [4,7];
- **Granularity**, describing the level of detail captured, such as timestamp precision or the specificity of activity labels [4,39,46,94];
- **Temporal coverage**, relating to time-related characteristics, including overall duration as well as case- and activity-level durations [78,79];
- **Contextual completeness**, indicating whether the log contains all necessary contextual information to support meaningful analysis [5,6,39];
- **Behavioral validity**, referring to how accurately the recorded events reflect the actual behavior of the real-world process [5,7].

The extraction of exploration activities (I2) provides an additional mechanism for interpreting event log validity. As shown in the system-side generation of event log understanding (U2), the domain expert's tacit and explicit knowledge informs exploration (E), which guides the derivation of updated specifications (S). The automatic analysis component (A) captures these exploration activities and stores the resulting explicit knowledge (K^{ε}) as contextual input within the LLM embedding space. This enables the system to model the expert's exploration behavior, revealing both considered and omitted analytical perspectives. Since omitted perspectives may negatively affect validity depending on the analytical objective, exploration activities constitute a crucial input for the guidance component discussed next.

4.1.3. *Provision of guidance*

The provision of guidance by AID4DE supports domain experts in both generating an understanding of the event log and interpreting its validity. Guidance is particularly important because domain experts are often only partially responsible for data validation and may lack full awareness of its overarching goals or required steps [81]. As illustrated in Fig. 5, the guidance mechanism comprises three components: two inputs (G1 and G2) and an output component (G3) that captures how the system actively supports validation activities.

The primary input for system-driven guidance (G1) is the explicit knowledge (K^{ε}) stored in the embedding space of the underlying LLM. This knowledge is derived from event log understanding (U1), externalized domain knowledge (I1), and extracted exploration activities (I2). It is complemented by learned knowledge (K^{L}), acquired during pretraining or, where applicable, fine-tuning of the LLM, and may include general insights into the application domain and common data validation patterns.

The input for AQ-based guidance (G2) consists of the set of AQs defined for the PM project. As discussed in Section 2.1, these questions shape event data preparation and subsequent analyses. To provide context-aware guidance, the system must be informed about the prevailing AQs, which may reflect diverse analytical perspectives or their combinations [37,38,91]. For example, Klinkmüller et al. [91] distinguish perspectives such as discovery, conformance checking, time analysis, organizational analysis, case analysis, prediction, drift detection, and familiarization. Milani et al. [38] further map AQs to these analytical purposes, illustrating questions such as: for discovery, *"How are cases executed?"*; for drift detection, *"How has process behavior changed over time?"*; or for prediction, *"What is the remaining time of an ongoing case?"*.

Because each AQ imposes different requirements on event log data, integrating AQs into the validation process enables the system to refine its contextual understanding and deliver more specific, relevant guidance to domain experts.

The output of guidance (G3) can take various forms with differing levels of influence on the domain expert. *Orienting guidance* provides cues (C) about characteristics of the prevailing event log, for example through natural language explanations that complement existing representations, shaping the expert's perception (P) [81]. *Directing guidance* suggests potential next steps by presenting options (O), such as specific exploration activities (E) that reveal alternative analytical perspectives [81]. Finally, *prescribing guidance* bypasses human input by directly modifying the specification (S) based on the system's assessment of promising next steps, for instance by generating targeted visualizations to address underexplored perspectives.

Overall, system guidance emerges from the interaction of orienting, directing, and prescribing mechanisms. It continuously adapts based on explicit domain knowledge, learned knowledge, and the prevailing AQs. As explicit knowledge evolves through ongoing expert interaction, guidance is progressively refined to better support event data validation.

4.2. *Instantiation as a software prototype*

To demonstrate the applicability of the conceptual model, we instantiated the underlying system of AID4DE as a software prototype operationalizing its core components. The prototype comprises multiple interface pages, each offering a distinct analytical perspective on the uploaded event log. After uploading an event log file, fit for PM, and specifying the relevant AQs, domain experts can initiate an exploration process to assess event log validity and quality.

Event log exploration is organized across three interface pages, each supporting a complementary analytical view. The initial data exploration page provides predefined visualizations for an overview of the log. The initial process exploration page presents process-oriented representations for detailed examination of process behavior. The interactive event log exploration page dynamically recommends additional

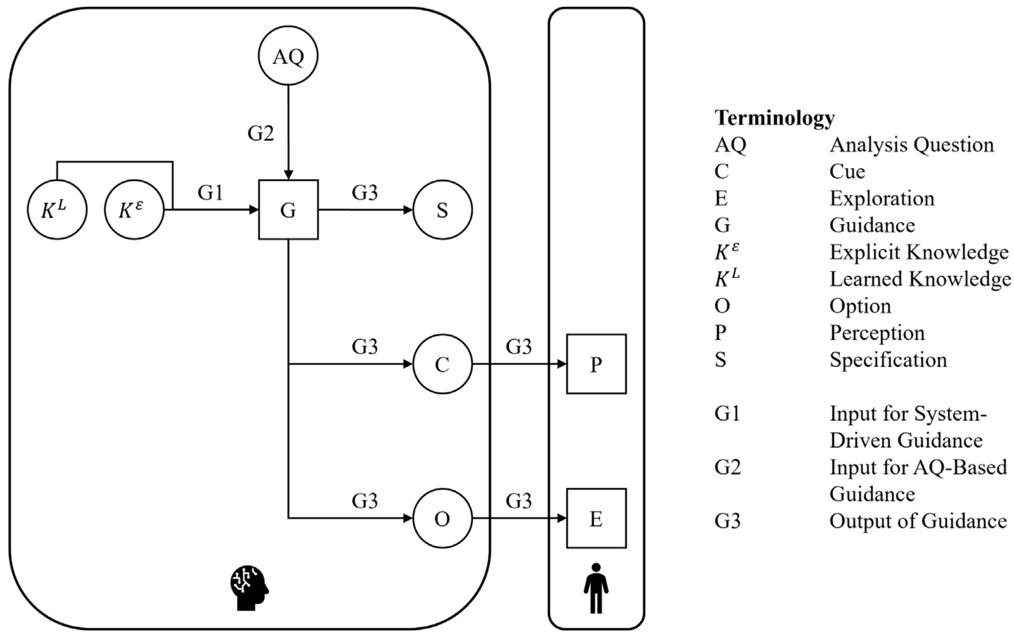


Fig. 5. Conceptual model of guidance provision in AID4DE.

visualizations based on the uploaded AQs. After completing the analysis, experts can export a summary report of validation activities and findings. The following subsections detail how these pages implement the conceptual model's components.

4.2.1. Initial data exploration

On the data exploration page, domain experts develop an initial understanding of the event log (U1) through automatically generated visualizations. In parallel, the system underlying AID4DE forms its internal representation (U2) by extracting representational semantics from these visualizations. To support interpretation, a chatbot provides contextual cues and explanations, guiding the expert's cognitive processing (U1, G3).

Uploaded event logs are converted into tabular form using the `pandas` library and standardized via `pm4py` utilities to include essential columns (Case ID, Activity, Timestamp, Resource). Descriptive summaries, including activity frequencies, case variants, temporal patterns, and resource associations, are computed and stored in structured session objects that constitute the system's internal knowledge base.

Visualizations are generated using the `pm4py.vis`, `pm4py.stats`, and `pm4py.filtering` modules [95] and rendered with `matplotlib` [96]. They cover summary statistics, frequency and distribution, temporal, performance, and, if available, resource analysis, using histograms, bar charts, heatmaps, and line charts.

Each visualization includes a free-text input field for expert observations regarding data validity. Experts can record whether observed patterns align with real-world processes, noting discrepancies or anomalies (I1). For example, an expert might note: "Activity B occurs more frequently than expected, suggesting a logging error or exceptional variant". These annotations externalize tacit knowledge for assessing the event log's fitness for purpose. Future versions may incorporate context-specific guiding questions for each visualization type.

The chatbot, powered by GPT-3.5 Turbo via *Azure OpenAI* [97], responds to expert queries during exploration. Structured summaries of event log features, such as activity frequencies, variants, and durations, are embedded into the prompt as explicit knowledge to ground explanations in actual data, ensuring interpretability, efficiency, and privacy. An example of such a textual representation is shown below:

```
Top activities: A(50%), B(30%), C(20%); Most frequent variants: [A-B-C (25%), A-C-B (15%)]; Average case duration: 5.2 h; Resource-task associations: R1(A), R2(B)...
```

4.2.2. Initial process exploration

The process exploration page enables domain experts to analyze process behavior captured in the event log (U1). Semantic information extracted from process representations is stored as structured, machine-readable summaries (U2), including start and end activity frequencies, case variants, and common sequences. These summaries support subsequent analyses and contextualized LLM-based explanations, while expert annotations externalize domain knowledge (I1).

Experts can interactively adjust a variant coverage threshold (default: 80%), defining the subset of cases used for process-centric visualizations. These include a BPMN model with natural language explanations, start and end activity statistics, and both DECLARE [98] and footprint models. Interactive filtering and model generation rely on the `pm4py.filtering` and `pm4py.discovery` modules [95]. The LLM interprets structured process summaries to provide data-grounded insights rather than generic model descriptions.

4.2.3. Interactive event log exploration

The interactive event log exploration page presents system-generated analytical questions that introduce new perspectives on the event log (G3). Generated visualizations support understanding for both the domain expert (U1) and the system (U2). Experts externalize knowledge about event log validity via dedicated input fields (I1), while the system records interaction data as explicit knowledge (I2), including selected questions, thresholds, and generated visualizations, to refine adaptive recommendations and focus on unexplored aspects.

Analytical questions are proposed based on the uploaded AQs (G2) and event log characteristics, targeting analytical perspectives not yet explored. When a question is selected, the system generates the corresponding visualization, enabling continued validation until no further questions remain. This iterative process enriches the shared knowledge base with structured summaries and tacit expert insights.

Table 1
Overview of related artifacts and fulfillment of identified DOs.

Artifact	DO1 (visualize and describe)	DO2 (interact)	DO3 (guide)
Barbieri et al. [26]	◒	◒	○
Case Group Explorer [99]	◒	●	○
InterLog [100]	◒	●	○
Log to Model Explorer [45]	◒	●	○
ProcessExplorer [101]	◒	●	◒
P2T [102]	◒	◒	○
RDB2Log [22]	◒	◒	◒
VisProc [103]	◒	●	○
VIT-PLA [104]	◒	●	○
AID4DE	●	●	●

5. Evaluation

5.1. Artificial formative evaluation

We began the evaluation with an artificial formative assessment to explore the problem space, investigate potential solutions, and iteratively refine the design based on early feedback. This phase comprised two activities. First, we conducted a competing artifact analysis to compare the proposed design specification with related solutions from the literature and highlight its added value [34]. Second, we carried out seven semi-structured interviews with academic and industry experts to assess the initial design specification.

5.1.1. Competing artifact analysis

Nine artifacts from the literature were selected as competing artifacts. Table 1 summarizes the extent to which these artifacts address our DOs, using ideographs to indicate levels of support: ● (full), ◒ (partial), and ○ (none). Full support indicates alignment in objectives and fulfillment of the respective DO; partial support reflects limited alignment due to differing focus or incomplete coverage; no support indicates the DO is not addressed. Detailed justifications are provided in Table C.7 in Appendix C.

The competing artifacts address only specific aspects of the research problem. For DO1, most provide either natural language descriptions [26,102], visual representations [100,101,103,104], or tabular formats [22]. Only Case Group Explorer [99] combines visualization with textual descriptions; however, its language targets process analysts rather than domain experts. Effectively supporting domain experts requires an integrated combination of clear visualizations, natural language explanations, and terminology clarification.

For DO2, several artifacts enable interactive exploration of event logs, including filtering infrequent behavior [45,100], clustering traces [45,99,101,104], and applying diverse filtering methods [100,103]. In contrast, artifacts relying solely on textual descriptions [26,102] substantially limit interaction and exploratory potential.

Regarding DO3, most artifacts provide limited guidance for data validation. Notable exceptions include RDB2Log, which highlights data quality issues through color-coding [22], and ProcessExplorer, which offers dynamic recommendations based on selected event log subsets [101].

While these artifacts offer valuable functionalities, none fully address all DOs. In contrast, AID4DE comprehensively covers each DO by supporting visualization with textual descriptions (DO1), interactive exploration (DO2), and targeted guidance throughout the validation process (DO3). Its guidance capabilities are particularly novel, providing contextualized support based on AQs, event log characteristics, and the domain expert's prior interactions.

Table 2
Participants involved in the first semi-structured interview round.

ID	Sector	Role	PM experience (years)	Country
1	Industry	Vice President Value Engineering	8	USA
2	Research	Professor	5	Switzerland
3	Industry	Senior Consultant Data Analytics	5	Germany
4	Research	Doctoral Candidate	4	Germany
5	Industry	PM Consultant	2	Germany
6	Industry	Business Excellence Advisor Board Office	4	Germany
7	Research	Doctoral Candidate	4	Germany

5.1.2. Semi-structured interviews on the design specification

In parallel with the competing artifact analysis, the design specification was evaluated through seven semi-structured expert interviews. Participants were selected for their expertise in PM and BPM, representing both academic and industrial perspectives (cf. Table 2). All participants routinely handled process data, including event-log extraction, cleaning, and validation. On average, they had 4.6 years of PM experience, and interviews lasted approximately 54 min.

The interviews comprised two parts. The first began with an introduction to the research problem, highlighting the expertise asymmetry between domain experts and process analysts during data validation and the limited attention given to assessing event logs' fitness for purpose. Interviewees then reflected on this research gap, discussed comparable challenges from their experience, and rated the relevance of the formulated RQ.

In the second part, the proposed DOs and resulting design specification were demonstrated using illustrative examples. Interviewees provided general feedback and evaluated the specification using an end-labeled unipolar seven-point Likert scale [105]. Following Sonnenberg and vom Brocke [89], four criteria were assessed: *novelty*, *understandability*, *completeness*, and *applicability*. Novelty examined whether the design addressed a unique problem and advanced current knowledge; understandability, whether it was clearly communicated and accessible; completeness, whether it sufficiently covered the research problem; and applicability, its potential for real-world implementation. Each criterion was briefly defined and accompanied by a guiding question to ensure consistent interpretation. Participants justified their ratings qualitatively and provided additional feedback beyond the predefined criteria. Quantitative results are shown in Fig. 6.

Interviewees, labeled 'I' followed by their IDs (cf. Table 2), acknowledged the importance of the data validation challenges and the associated research gap. I1 emphasized that recurring asymmetries between domain experts and technical stakeholders complicate validation efforts. I5 noted that assessing an event log's fitness for purpose is often neglected, stating that "it can't be proceeded with the next steps within a PM project without properly validating the data". I3 highlighted the critical role of the information technology (IT) department, which typically owns and understands the data. Overall, the RQ was rated as highly important. More critical reflections included I2's observation that "GenAI is not a direct consequence of the problem" and I4's remark that "GenAI brings new potentials, but also limits the solution space". These insights informed refinements to the research motivation in Section 1.

Interviewees generally found the design specification highly comprehensible. I6 suggested clarifying the specific value contribution of the GenAI component and providing a more concrete depiction of expected outputs, while also emphasizing greater focus on semantic evaluation of the event log. I7 noted that although the abstract level supports overall understanding, deeper insight into component interactions requires a more integrated perspective, recommending illustrations of component interrelations and their integration into standard

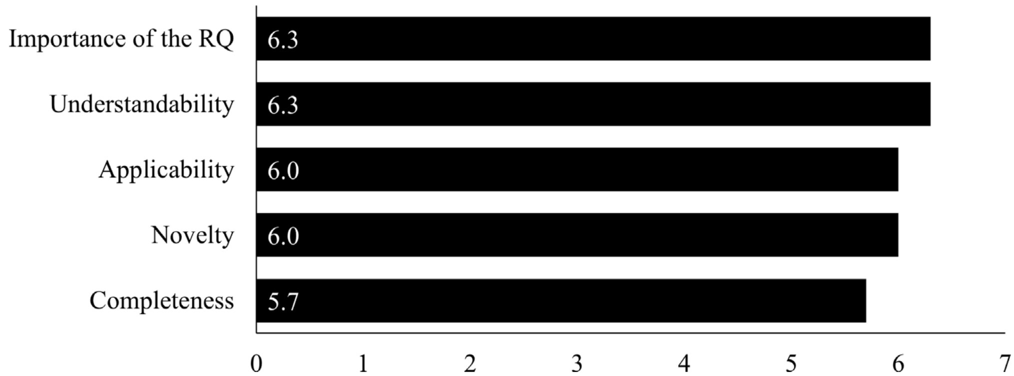


Fig. 6. Average Likert scale ratings for artificial formative evaluation criteria.

data validation steps. While this feedback did not prompt immediate revisions, it informed the more detailed exposition in Section 4.

Interviewees generally perceived the design specification as innovative. I1 noted that while it does not introduce new technology, it combines existing concepts in novel ways, particularly through the active involvement of domain experts in interpreting event log validity. I5 emphasized that visual representations enhance understanding and highlighted the GenAI-enabled conversational interface as a distinctive feature, "building a bridge between complex event log data and domain experts". I7 agreed, while noting that VA-based data validation approaches are not entirely new. Overall, both the competing artifact analysis and interview feedback confirmed the contribution of a novel solution to data validation.

Interviewees highlighted both human and technical factors influencing applicability. I7 observed that "each modern company can implement the approach" but questioned domain experts' readiness to adopt it. I1, I2, and I3 emphasized the importance of human engagement, including motivation, workload, and the extent to which domain experts' perspectives are reflected in the event log, all of which affect trust and analysis validity. Technical considerations included context sensitivity, with I2 stressing alignment between visualizations, AQs, and event data characteristics, and I6 warning that GenAI hallucinations may undermine reliability. Overall, applicability largely depends on domain expert engagement, a factor beyond full system control. These insights informed refinements to enhance accessibility and practical relevance (Section 4).

Interviewees generally considered the design specification sufficiently complete for supporting domain experts in validating event log data for PM, while identifying opportunities for improvement. I5 recommended including summary statistics and visualizations but cautioned that generating multiple in-depth explanations can be time-consuming and subject to diminishing returns. I1 suggested a key performance indicator (KPI) to quantify validation progress, potentially expressed as a percentage. I5 also advocated system-generated explanations clarifying the need for dedicated tools and specific validation actions, while I6 emphasized explainability, including translating PM terminology and clarifying the AI component's workings. Several suggestions were incorporated into the final specification, such as technical term translation and detailed activity descriptions, while others (e.g., validation progress KPI and enhanced AI explainability) remain promising avenues for future research.

5.2. *Artificial summative evaluation*

In the second evaluation phase, we conducted a summative assessment of the artifact in an artificial context [34] prior to real-world application. During this phase, the design specification was refined and the conceptual model finalized. The model was presented to seven additional research and industry experts selected for their PM and BPM experience (cf. Table 3). Participants had an average of 6.4 years of PM experience, and each interview lasted approximately 57 min.

Table 3
Participants in the second semi-structured interview round.

ID	Sector	Role	PM experience (years)	Country
8	Industry	Center of Excellence Lead PM	10	Germany
9	Industry	Head of Process AI	6	Germany
10	Industry	Chief Executive Officer of PM Consultancy	9	Germany
11	Industry	Head of Business Process Excellence	7	Liechtenstein
12	Research	Doctoral Candidate	6	Germany
13	Research	Doctoral Candidate	2	Germany
14	Research	Professor	5	Netherlands

The interviews followed a two-part structure. First, participants were introduced to the research problem and RQ, using materials similar to the initial interview round, with a minor refinement of the research gap based on earlier insights. In the second part, the revised DOs, conceptual model, and its components were presented. The model was introduced top-down, starting with its overall structure and terminology, followed by detailed discussion of individual components — generation of event log understanding, interpretation of event log validity, and guidance of domain experts. The complete model was then revisited, color-coded to illustrate component integration.

Following the introduction, interviewees evaluated the conceptual model using end-labeled unipolar seven-point Likert scales [105]. Five established DSR criteria were assessed [89]: *completeness* (extent to which the model includes all necessary concepts and relationships), *fidelity with the real world* (accuracy in reflecting actual practices), *internal consistency* (logical coherence), *level of detail* (appropriate abstraction), and *robustness* (utility under minor contextual variations). Each criterion was introduced with a brief definition and guiding question. Participants provided quantitative ratings and qualitative justifications, with results shown in Fig. 7. Interviews concluded with an open-ended question on the model's applicability and usefulness in practice, summarized in Table 4.

Interviewees, labeled 'I' followed by their IDs (cf. Table 3), generally agreed on the conceptual model's completeness. I8, I11, I12, and I13 explicitly affirmed this assessment. I8 highlighted the role of visualizations in enabling closed-loop communication between domain experts and the event log, noting that similar approaches exist in their

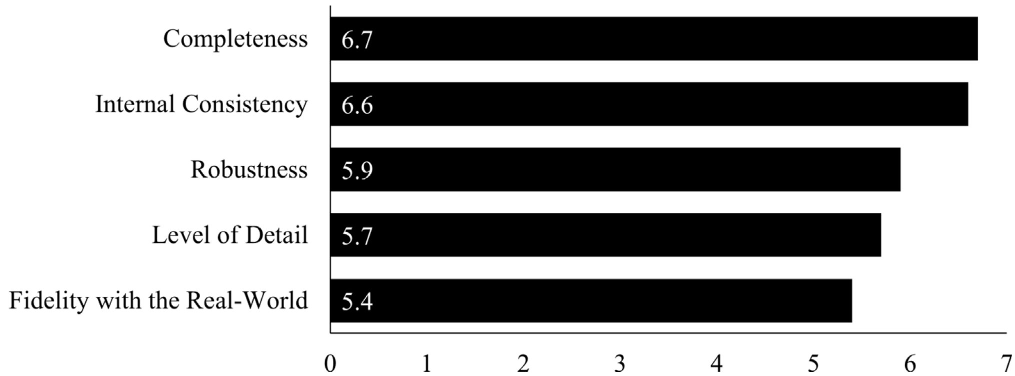


Fig. 7. Average Likert scale ratings for the conceptual model based on the artificial summative evaluation criteria.

organization but are currently manual and lack tool support. I10 suggested making the integration of common knowledge from GenAI more explicit, which we addressed by embedding learned knowledge in the system underlying AID4DE. I9 recommended linking the model directly to a PM environment, a valuable but out-of-scope implementation consideration.

Interviewees emphasized the conceptual model's internal consistency. I8, I9, I10, and I11 reported no noticeable inconsistencies. I12 noted that consistency was not immediately evident in the top-down presentation but became clear after individual components were explained. I13 emphasized that the AQ originates from the domain expert interacting with the system underlying AID4DE, appearing in the model only once the analysis-ready stage begins; this aligns with our design intent and was incorporated into the final model. I13 also noted the absence of explicit start and end points for completing the iterative loop. As discussed in Section 5.1.2, including a validation progress KPI remains an area for future research.

Interviewees generally regarded the conceptual model as robust. I8, I10, and I14 noted that it maintains integrity under minor environmental changes. I10 cautioned that future revisions may be required as GenAI evolves, particularly with emerging paradigms such as agentic AI, which are beyond the scope of this study. I11 emphasized that iterative refinement supported by the closed-loop design enhances robustness. I9 suggested potential generalization beyond event logs to conventional datasets. Considering hypothetical scenarios, I12 observed that parallel validation by multiple domain experts may introduce variation, while I13 highlighted that erroneous expert input naturally leads to erroneous output, a limitation explicitly acknowledged in our approach.

Regarding the level of detail, interviewees generally agreed that the model's abstraction is appropriate, while suggesting selective refinements. I8, I12, and I13 considered the current level sufficient, with I13 noting that component relationships are clearly conveyed. I9, I10, I11, and I14 suggested modeling additional interaction possibilities, integrating knowledge of the underlying LLM more explicitly, and representing iterative validation steps. In response, the final model summarizes key interactions within the exploration activity and represents knowledge of the underlying LLM as learned knowledge in a separate processing component. Individual iterations are not depicted, as the closed-loop design better captures the dynamic nature of data validation.

Interviewees generally agreed that the conceptual model is applicable in practice, while noting its potential to alter existing validation procedures. I11 and I14 observed that typical data validation involves multiple iterative stakeholder interactions, whereas AID4DE follows a more sequential structure. I11 added that deviations from expected process behavior often require additional consultation before validation proceeds. I12 emphasized that effectiveness depends on domain expert motivation and engagement. I9 noted that validation may be infeasible in contexts with missing data or unavailable expertise. In contrast, I8 and I13 expressed strong confidence in the model's practical utility. I13 further highlighted that AQs are typically derived during planning phases, an assumption reflected in the model.

Table 4
Positive and negative comments (with participant references) on the conceptual model.

Positive comments	Negative comments
• Closing the knowledge gap between domain experts and process analysts (I9, I10, I13, I14) • Identification of data quality issues (I12) • Extraction of domain knowledge (I8, I10) • Enhancing trust in the data (I8) • Appreciation of domain experts' knowledge (I11) • Learning opportunity for domain experts to do data validation (I11, I12)	• Success dependency on the domain expert (I8, I12) • Neglect of the process analyst (I13) • Increased time investment for data validation (I14)

Responses to the open-ended question on applicability and usefulness were predominantly positive (cf. Table 4). Interviewees highlighted the model's potential to enhance data validation by bridging the knowledge gap between domain experts and process analysts. They also emphasized its value for future initiatives, noting that it recognizes domain experts' knowledge and provides a structured, systematic approach to validation.

Interviewees identified several limitations of the conceptual model, including its reliance on domain experts, potential increases in validation time, and the absence of the process analyst. We acknowledge these concerns but argue that focusing on domain experts enables the extraction of valuable tacit knowledge, improving event log quality and the validity of subsequent insights. Although this approach may require additional time, the anticipated benefits justify the investment, particularly given that non-automated data validation inherently requires active domain expert engagement.

5.3. Naturalistic summative evaluation

The final evaluation phase was a naturalistic summative evaluation, aimed at demonstrating the artifact's proof of value, defined as its ability to generate value across multiple domains [106]. We conducted

Table 5
Summary of participants' ratings of statements on perceived usefulness (S1–S6) and perceived ease of use (S7–S12) of the prototype.

Statement	Mean	Median	Std
Perceived Usefulness (S1–S6)			
S1: Using the prototype in my job would enable me to accomplish data validation tasks more quickly.	5.83	6.0	1.25
S2: Using the prototype would improve my data validation performance.	5.56	6.0	1.38
S3: Using the prototype for data validation would increase my productivity.	5.89	6.0	1.18
S4: Using the prototype would enhance my effectiveness in data validation.	5.56	6.0	1.25
S5: Using the prototype would make it easier to do data validation.	6.00	6.0	1.28
S6: I would find the prototype useful for data validation.	5.72	6.0	1.36
Perceived Ease of Use (S7–S12)			
S7: Learning to operate the prototype would be easy for me.	6.33	7.0	0.91
S8: I would find it easy to get the prototype to do what I want it to do.	5.28	5.0	1.32
S9: My interaction with the prototype would be clear and understandable.	5.72	6.0	1.07
S10: I would find the prototype to be flexible to interact with.	5.28	5.0	1.13
S11: It would be easy for me to become skillful at using the prototype.	6.28	6.0	0.75
S12: I would find the prototype easy to use.	6.17	6.0	0.99

a live demonstration of the conceptual model, instantiated as a software prototype, for 18 IS researchers previously introduced to the research gap and RQ. Although participants were not specialized in PM or BPM, their extensive experience in data-driven IS research and industry collaboration enabled informed assessment of the prototype's conceptual soundness, analytical capabilities, and usability on real-world data.

For the demonstration, we used a publicly available XES event log from an Italian software company's help desk [88], comprising 21,348 events, 4,580 cases, and 14 activities. The guiding AQ was: "Where do inefficiencies occur within the ticket management process?" Participants assumed the role of domain experts responsible for processing IT support tickets. While some had prior experience in this domain, others approached the task from a broader IS perspective. Participants were encouraged to propose interactions with the prototype, fostering collaborative exploration and interpretation. Qualitative and quantitative feedback was collected through a user study survey following the session.

The live demonstration lasted approximately 35 min and concluded with a discussion of how the prototype addresses the RQ. Participants actively explored inefficiencies and proposed interactions with the system. For example, when examining a bar chart of monthly events, they observed a drop in August and hypothesized missing data rather than reduced activity, annotating the visualization accordingly. Using the chatbot, participants queried potential explanations related to the AQ, which suggested seasonal factors such as vacation periods. This interaction highlighted the potential benefits of integrating an organizationally trained LLM into AID4DE to provide context-aware support during complex validation scenarios.

During the demonstration, participants identified several notable patterns in the event log:

1. Events occurred on Saturdays and Sundays, albeit less frequently than on weekdays, suggesting potential anomalies or inconsistencies in process logging.
2. Approximately 3% of tickets were opened but never closed, possibly reflecting long-running cases with unclear ownership or tickets that were closed but not properly logged.
3. Two labels, "Closed" and "RESOLVED", were used for resolved tickets, representing a *synonymous label* imperfection pattern [5].
4. One resource, "Value 2", exhibited substantially higher capacity utilization than others. While potentially concerning for IT ticket handling, participants noted that this pattern reflects analytical insights rather than data validation issues.

All identified anomalies were documented using free-text annotation fields linked to the corresponding visualizations. These annotations, together with the analytical context, provide a traceable record of the validation process and the interaction between domain experts and the AID4DE system. This documentation supports subsequent conventional data validation sessions between process analysts and domain experts, facilitating assessment of whether observed patterns indicate data quality issues. The collaborative use of AID4DE thus served as an initial step toward comprehensive event log understanding and activation of domain knowledge for validation.

After the demonstration, participants completed a three-part survey. The first part collected background information on experience in IS, BPM, and PM. The second part applied the TAM, which identifies perceived usefulness and perceived ease of use as key determinants of system acceptance [90]. Perceived usefulness refers to the "degree to which a person believes that using a system would enhance job performance", and perceived ease of use refers to the "degree to which a person believes that using a system would be free of effort" [90]. Participants rated 12 statements on a seven-point Likert scale (1 = "I fully disagree" to 7 = "I fully agree"), with S1–S6 addressing usefulness and S7–S12 addressing ease of use. Table 5 reports the mean, median, and standard deviation (Std) for each statement. The third part consisted of two open-ended questions soliciting qualitative feedback on strengths and areas for improvement.

The quantitative TAM results indicate that participants evaluated the AID4DE prototype favorably in terms of both perceived usefulness and ease of use. The mean score for ease of use (5.84) was slightly higher than that for usefulness (5.76). Median values were 5.94 and 5.78, respectively. Stds were higher for ease of use (Std = 0.49) than for usefulness (Std = 0.18), indicating greater variability in perceptions of ease of use.

For perceived usefulness, S5 achieved the highest mean score (6.0), indicating strong agreement that the prototype facilitates data validation, while S2 and S4 showed slightly lower means (5.56). For perceived ease of use, S7 scored highest (6.33) with low variability, suggesting participants consistently found the prototype easy to learn. In contrast, S8 and S10 received the lowest scores (5.28), indicating limitations in interaction flexibility. Overall, ease of use exhibited greater variability across statements than usefulness, explaining its higher aggregate Std.

Qualitative feedback highlighted several strengths of the prototype. Participants noted its effectiveness in supporting domain experts in developing foundational data-related skills. They emphasized its suitability for users with varying expertise, citing interpretable visualizations, a clear structure, and minimal prior knowledge requirements. Participants also valued the presentation of multiple analytical perspectives, which fostered a comprehensive understanding of the event log.

Participants suggested enhancing guidance and explainability for domain experts. They emphasized the need for clearer terminology and more detailed explanations, particularly for users without a background in process science. For example, participants with expertise in Data Science and Applied AI requested clarification of terms such as *cases* and *variants*. Improved guidance could be achieved by fine-tuning the underlying LLM on domain-specific data and by providing explicit recommendations for next steps. Proactively posing context-aware questions related to specific visualizations was also suggested as a means to increase system support.

Participants also noted the need to improve system performance. Although performance optimization was not a primary focus of the prototype, noticeable latency occurred during backend LLM interactions, such as generating process model descriptions or responding to chatbot queries. These delays affected the demonstration experience. Addressing such performance issues will be essential for real-world deployment of AID4DE, ensuring a smooth and reliable user experience while acknowledging inherent limitations of current LLMs.

6. Discussion

This paper conceptualizes the interaction between domain experts and the system underlying AID4DE to facilitate validation of event logs with respect to their intended analytical purpose. We leverage GenAI to translate complex event log data into comprehensible visualizations, enable interactive exploration, and provide context-sensitive guidance throughout the validation process. The conceptual model was evaluated through 14 semi-structured interviews, instantiated as a software prototype, and subsequently assessed in a user study with 18 IS researchers. Overall, the results indicate that the model effectively supports domain experts in assessing the fitness for purpose of event logs.

This study makes three key theoretical contributions. First, it provides a foundation for evaluating the fitness for purpose of event logs through expert validation, extending prior work [49,81,85] by explicitly conceptualizing knowledge generation processes across both human and system spaces. Second, we introduce an extraction mechanism that transforms raw event log data into representational semantics via visualizations, addressing LLM context window limitations while integrating seamlessly into the overall model. Third, the artifact incorporates a guidance mechanism for validation that can be adapted to other PM tasks requiring expert input. This mechanism combines event log characteristics, particularly representational semantics, with the stated AQ to direct meaningful validation actions, offering a transferable blueprint for GenAI-powered systems.

We identify four primary practical contributions. First, AID4DE enhances domain experts' understanding of event logs and their analytical validity, thereby supporting more effective collaboration with process analysts. Second, it facilitates the externalization of tacit domain knowledge through visualizations and context-sensitive guidance, generating richer insights for other PM stakeholders. Third, the artifact improves usability and understanding of PM, addressing challenges highlighted in the PM Manifesto [4]. Fourth, AID4DE demonstrates the potential of LLM-based support in organizational settings, while also highlighting risks arising from incomplete or inaccurate organizational data that may affect validation reliability.

Several limitations suggest directions for future research. First, as a DSR project, the results are context-specific and may not generalize across different organizational settings, domains, or event log types. Second, validation quality depends on the domain expert's tacit knowledge, which may vary in articulation and cognitive framing. While automated approaches could complement expert input, they may also limit the depth of knowledge extraction. Third, the artifact does not explicitly balance validation certainty and time investment; future work could introduce decision-support metrics, such as marginal rates of return, to guide validation effort. Fourth, the use of GenAI introduces risks related to biased or incorrect outputs and sensitivity to prompt formulation, underscoring the need for systematic evaluation and reproducibility practices. Trust in AI-generated outputs was not empirically assessed and remains an important avenue for future research. Fifth, the approach assumes XES event logs that are fit for PM and the availability of at least one AQ; future studies could extend the approach to alternative formats, such as OCEL. Sixth, the performance of the LLM in providing guidance and interaction was not systematically evaluated, warranting further investigation into handling domain-specific cues and explanations. Finally, the study does not yet offer a quantitative framework for assessing an event log's fitness for purpose; future research may develop scoring models and context-sensitive refinement strategies to complement expert judgment.

7. Conclusion

Event logs used in PM frequently suffer from data quality issues that can lead to inaccurate insights and suboptimal process improvements if not addressed during data preparation. Data validation is typically performed collaboratively by domain experts and process analysts to assess whether an event log is suitable for its intended analytical purpose; however, this alignment is often insufficiently examined in current practice. To address this gap, we investigated how GenAI can support domain experts in evaluating the fitness for purpose of event logs and identifying discrepancies between the prevailing event log and the underlying AQs. We developed AID4DE, a conceptual and instantiable artifact that enables domain experts to interact with complex event log data and generate structured knowledge about its validity. Following the DSR paradigm, AID4DE was iteratively designed and evaluated through a competing artifact analysis, 14 semi-structured expert interviews, and a user study with 18 IS researchers. The results indicate that AID4DE enhances data validation by improving experts' understanding of event log quality relative to its analytical purpose. Moreover, the externalized knowledge produced during validation can support subsequent data preparation activities, such as identifying and repairing data quality issues.

To our knowledge, this study is the first to explicitly conceptualize the knowledge generation processes involved in evaluating the fitness for purpose of event logs during data validation. Its primary contribution lies in offering an alternative to conventional validation practices by enabling domain experts to develop an understanding of event logs through visual representations, from which representational semantics are extracted to guide interpretation and further analysis. The conceptual model is instantiated as a software prototype that operationalizes these mechanisms and supports experts in carrying out validation tasks in a structured and interactive manner.

Table A.6
Overview of development iterations, evaluation outcomes, and key insights for the system underlying AID4DE.

Iteration	Development activities	Evaluation results	Key insights
1	• Reviewed literature on various data exploration tools and GenAI applications in BPM and PM	• Identified context window size constraints of LLMs as a limitation for directly uploading event logs • LLM applications can transform complex contexts into natural language representations	• Pre-processing of event logs is required to make them suitable for consumption by LLMs
2	• Reviewed literature from the VA domain • Developed an initial prototype visualizing event log data and extracting representational semantics for LLM input	• Extracting and transferring representational semantics enables the LLM to understand event log data • LLMs can summarize event log characteristics based on the provided semantic data	• GenAI can bridge the gap between complex event data and understandable representations for domain experts
3	• Incorporated additional visualizations in the prototype to provide more representational semantic data for the LLM • Enabled interactive features, such as filtering and annotations, to explore the event log	• Interactive functionalities facilitate deeper analytical insights into the event log • Enriching the LLM with more semantic data improves its comprehension of the event log	• Interactive features combined with enriched semantic data enhance both domain experts' and LLM's understanding of the event log
4	• Implemented GenAI-based guidance aligned with the stated AQ and stored semantic data • Conducted extensive prompt engineering to refine guidance quality	• Representational semantic data and AQ are essential for delivering effective guidance in data validation • Prompt engineering enhances the specificity and relevance of guidance provided	• GenAI can effectively support the interpretation of event log validity, forming a basis for assessing their fitness for purpose

While the proposed model and prototype effectively support the evaluation of event logs, several limitations remain. In particular, the approach depends on domain experts' ability to externalize relevant knowledge and engage in exploratory analysis, which may vary across individuals and contexts. Future research could investigate automated and semi-automated validation techniques, including agentic AI, to complement expert-driven processes and further scale validation efforts while preserving the benefits of human judgment.

CRediT authorship contribution statement

Julian Armin Dormehl: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Robert Andrews:** Writing – review & editing, Conceptualization. **Wolfgang Kratsch:** Writing – review & editing, Supervision. **Maximilian Röglinger:** Writing – review & editing, Supervision, Methodology, Funding acquisition. **Moe Thandar Wynn:** Writing – review & editing, Supervision, Conceptualization. **Felix Zetzsche:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Formal analysis, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT (developed by OpenAI) in order to improve the clarity of expression and to check grammar and spelling. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We gratefully acknowledge the StMWi (Bavarian Ministry of Economic Affairs, Regional Development and Energy) for their support of the project "QUAPRO (DIK-2209-0017//DIK0440/02)" that made this paper possible. This study received ethical approval from QUT HREC — approval number 9581.

Appendix A. Development iterations and insights

Table A.6 summarizes the key development iterations of the system underlying AID4DE, together with the associated evaluation outcomes and resulting insights.

Appendix B. Detailed architecture of the conceptual model

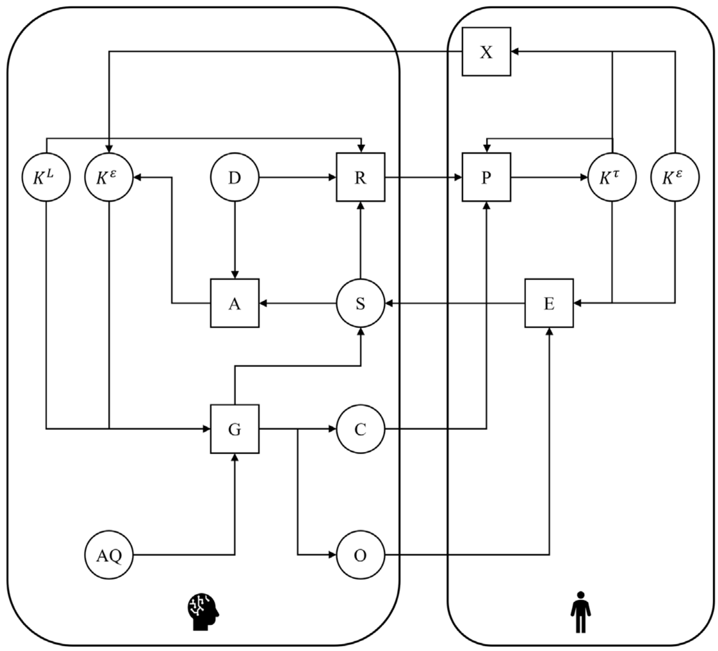
Fig. B.8 illustrates the detailed architecture of the AID4DE conceptual model.

Appendix C. Detailed competing artifact analysis

Table C.7 provides a detailed comparison of existing artifacts with respect to the three DOs defined in this study.

Data availability

The source code used in this study is publicly available on GitHub. The repository link is provided in the manuscript.



Terminology

A	Automated Analysis
AQ	Analysis Question
C	Cue
D	Event Log Data
E	Exploration
G	Guidance
K^ε	Explicit Knowledge
K^L	Learned Knowledge
K^τ	Tacit Knowledge
O	Option
P	Perception
R	Representation
S	Specification
X	Externalization

Fig. B.8. Conceptual model for AID4DE.

Table C.7
Detailed comparative analysis of existing artifacts with respect to the defined DOs.

Artifact	DO1	DO2	DO3
Barbieri et al. [26]	◒ Provides natural language answers but lacks visualizations	◒ Supports only natural language interaction with event logs	○ Does not provide user guidance
Case Group Explorer [99]	◒ Provides visualizations and explanations but lacks terminology clarification	● Supports interaction via clustering of traces into groups and linked visualizations	○ Does not provide user guidance
InterLog [100]	◒ Provides control-flow-focused visualizations but lacks natural language explanations	● Supports interaction through filtering functions	○ Does not provide user guidance
Log to Model Explorer [45]	◒ Provides control-flow-focused visualizations but lacks natural language explanations	● Supports interaction through clustering and filtering functions	○ Does not provide user guidance
ProcessExplorer [101]	◒ Provides visualizations but lacks natural language explanations	● Supports interaction through filtering functions	◒ Provides dynamic subset and insight recommendations
P2T [102]	◒ Provides natural language answers but lacks visualizations	◒ Supports only natural language interaction with event logs	○ Does not provide user guidance
RDB2Log [22]	◒ Provides tabular overviews but lacks natural language explanations	◒ Offers limited interaction by allowing selection of specific event log attributes	◒ Provides color-coding to indicate data quality of event attributes
VisProc [103]	◒ Enables visualizations but lacks natural language explanations	● Supports interaction through clustering functions	○ Does not provide user guidance
VIT-PLA [104]	◒ Provides visualizations but lacks natural language explanations	● Supports interaction through clustering and aggregation functions	○ Does not provide user guidance

References

[1] M. Dumas, M. La Rosa, J. Mendling, H.A. Reijers, Fundamentals of Business Process Management, 2nd ed., Springer, Berlin, Heidelberg, 2018, http://dx.doi.org/10.1007/978-3-662-56509-4.

[2] M. Weske, Business Process Management: Concepts, Languages, Architectures, 4th ed., Springer, Berlin, Heidelberg, 2024, http://dx.doi.org/10.1007/978-3-662-69518-0.

[3] W. van der Aalst, Process Mining: Data Science in Action, 2nd ed., Springer, Berlin, Heidelberg, 2016, http://dx.doi.org/10.1007/978-3-662-49851-4.

[4] W. van der Aalst, A. Adriansyah, A.K.A. de Medeiros, F. Arcieri, T. Baier, T. Blickle, et al., Process Mining Manifesto, in: F. Daniel, K. Barkaoui, S. Dustdar (Eds.), Business Process Management Workshops, in: Lecture Notes in Business Information Processing, Springer, Berlin, Heidelberg, 2012, pp. 169–194, http://dx.doi.org/10.1007/978-3-642-28108-2_19.

[5] S. Suriadi, R. Andrews, A.H.M. ter Hofstede, M.T. Wynn, Event log imperfection

patterns for process mining: Towards a systematic approach to cleaning event logs, Inf. Syst. 64 (2017) 132–150, http://dx.doi.org/10.1016/j.is.2016.07.011.
[6] M.T. Wynn, S. Sadiq, Responsible Process Mining - A Data Quality Perspective, in: T. Hildebrandt, B.F. van Dongen, M. Röglinger, J. Mendling (Eds.), Business Process Management, in: Lecture Notes in Computer Science, Springer, Cham, 2019, pp. 10–15, http://dx.doi.org/10.1007/978-3-030-26619-6_2.
[7] R.P.J.C. Bose, R.S. Mans, W.M.P. van der Aalst, Wanna improve process mining results? in: 2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM), 2013, pp. 127–134, http://dx.doi.org/10.1109/CIDM.2013.6597227.
[8] I. Beerepoot, C. Di Ciccio, H.A. Reijers, S. Rinderle-Ma, W. Bandara, A. Burattin, et al., The biggest business process management problems to solve before we die, Comput. Ind. 146 (2023) 103837, http://dx.doi.org/10.1016/j.compind.2022.103837.
[9] Fluxicon, Detect and Fix Data Quality Problems, 2024, https://fluxicon.com/book/read/dataquality/. (Accessed 27 May 2025).
[10] F. Emamjome, R. Andrews, A.H.M. ter Hofstede, A Case Study Lens on Process Mining in Practice, in: H. Panetto, C. Debruyne, M. Hepp, D. Lewis, C.A. Ardagna, R. Meersman (Eds.), on the Move To Meaningful Internet Systems: OTM 2019 Conferences, in: Lecture Notes in Computer Science, Springer, Cham, 2019, pp. 127–145, http://dx.doi.org/10.1007/978-3-030-33246-4_8.
[11] M.T. Wynn, J. Lebherz, W.M.P. van der Aalst, R. Accorsi, C. Di Ciccio, L. Jayarathna, H.M.W. Verbeek, Rethinking the Input for Process Mining: Insights from the XES Survey and Workshop, in: J. Munoz-Gama, X. Lu (Eds.), Process Mining Workshops, in: Lecture Notes in Business Information Processing, Springer, Cham, 2022, pp. 3–16, http://dx.doi.org/10.1007/978-3-030-98581-3_1.
[12] S. Franzoi, S. Hartl, T. Grisold, H. van der Aa, J. Mendling, J. vom Brocke, Explaining process dynamics: a Process Mining Context Taxonomy for sense-making, Process. Sci. 2 (1) (2025) 2, http://dx.doi.org/10.1007/s44311-025-00008-6.
[13] F. Zerbato, J.J. Koorn, I. Beerepoot, B. Weber, H.A. Reijers, On the Origin of Questions in Process Mining Projects, in: J.P.A. Almeida, D. Karastoyanova, G. Guizzardi, M. Montali, F.M. Maggi, C.M. Fonseca (Eds.), Enterprise Design, Operations, and Computing, in: Lecture Notes in Computer Science, Springer, Cham, 2022, pp. 165–181, http://dx.doi.org/10.1007/978-3-031-17604-3_10.
[14] M. Dumas, F. Fournier, L. Limonad, A. Marrella, M. Montali, J.-R. Rehse, et al., AI-augmented Business Process Management Systems: A Research Manifesto, ACM Trans. Manag. Inf. Syst. 14 (1) (2023) 1–19, http://dx.doi.org/10.1145/3576047.
[15] S. Weinzierl, S. Zilker, S. Dunzer, M. Matzner, Machine learning in business process management: A systematic literature review, Expert Syst. Appl. 253 (2024) 124181, http://dx.doi.org/10.1016/j.eswa.2024.124181.
[16] A. del Río Ortega, I. Beerepoot, H. van der Aa, J. Evermann, Process Mining Without Perfect Data? Anne Rozinat Says Yes! A Practitioner's View on Event Log Quality, Bus. Inf. Syst. Eng. 67 (3) (2025) 385–387, http://dx.doi.org/10.1007/s12599-025-00938-2.
[17] M.P.J. van der Loo, E. de Jonge, Data Validation Infrastructure for R, J. Stat. Softw. 97 (10) (2021) 1–31, http://dx.doi.org/10.18637/jss.v097.i10.
[18] O. Marjanovic, G. Patmore, N. Balnave, Visual Analytics: Transferring, Translating and Transforming Knowledge from Analytics Experts to Non-technical Domain Experts in Multidisciplinary Teams, Inf. Syst. Front. 25 (4) (2023) 1571–1588, http://dx.doi.org/10.1007/s10796-022-10310-4.
[19] L. Zimmermann, F. Zerbato, B. Weber, What makes life for process mining analysts difficult? A reflection of challenges, Softw. Syst. Model. 23 (6) (2024) 1345–1373, http://dx.doi.org/10.1007/s10270-023-01134-0.
[20] A. Mamudu, W. Bandara, M.T. Wynn, S.J.J. Leemans, Process Mining Success Factors and Their Interrelationships, Bus. Inf. Syst. Eng. 67 (3) (2025) 389–408, http://dx.doi.org/10.1007/s12599-024-00860-z.
[21] F. Zerbato, P. Soffer, B. Weber, Initial Insights into Exploratory Process Mining Practices, in: A. Polyvyanyy, M.T. Wynn, A. Van Looy, M. Reichert (Eds.), Business Process Management Forum, in: Lecture Notes in Business Information Processing, Springer, Cham, 2021, pp. 145–161, http://dx.doi.org/10.1007/978-3-030-85440-9_9.
[22] R. Andrews, C.G.J. van Dun, M.T. Wynn, W. Kratsch, M.K.E. Röglinger, A.H.M. ter Hofstede, Quality-informed semi-automated event log generation for process mining, Decis. Support Syst. 132 (2020) 113265, http://dx.doi.org/10.1016/j.dss.2020.113265.
[23] D.A. Fischer, K. Goel, R. Andrews, C.G.J. van Dun, M.T. Wynn, M. Röglinger, Towards interactive event log forensics: Detecting and quantifying timestamp imperfections, Inf. Syst. 109 (2022) 102039, http://dx.doi.org/10.1016/j.is.2022.102039.
[24] H. van der Voort, S. van Bulderen, S. Cunningham, M. Janssen, Data science as knowledge creation a framework for synergies between data analysts and domain professionals, Technol. Forecast. Soc. Change 173 (2021) 121160, http://dx.doi.org/10.1016/j.techfore.2021.121160.
[25] L. Barbieri, E. Madeira, K. Stroeh, W. van der Aalst, A natural language querying interface for process mining, J. Intell. Inf. Syst. 61 (1) (2023) 113–142, http://dx.doi.org/10.1007/s10844-022-00759-9.
[26] L. Barbieri, K. Stroeh, E.R.M. Madeira, W.M.P. van der Aalst, An LLM-Based Q&A Natural Language Interface to Process Mining, in: A. Delgado, T. Slaats (Eds.), Process Mining Workshops, in: Lecture Notes in Business Information Processing, Springer, Cham, 2025, pp. 5–17, http://dx.doi.org/10.1007/978-3-031-82225-4_1.
[27] X. Han, L. Hu, J. Sen, Y. Dang, B. Gao, V. Isahagian, et al., Bootstrapping Natural Language Querying on Process Automation Data, in: 2020 IEEE International Conference on Services Computing (SCC), 2020, pp. 170–177, http://dx.doi.org/10.1109/SCC49832.2020.00030.
[28] S. Sadeghianasl, A.H.M. ter Hofstede, M.T. Wynn, S. Türkay, Humans-in-the-loop: Gamifying activity label repair in process event logs, Eng. Appl. Artif. Intell. 132 (2024) 107875, http://dx.doi.org/10.1016/j.engappai.2024.107875.
[29] F. Zetzsche, R. Andrews, A.H.M. ter Hofstede, M. Röglinger, S.J. Schmid, M.T. Wynn, Case ID Revealed HERE: Hybrid Elusive Case Repair Method for Transformer-Driven Business Process Event Log Enhancement, Bus. Inf. Syst. Eng. 67 (3) (2025) 311–337, http://dx.doi.org/10.1007/s12599-025-00935-5.
[30] N. Martin, A. Martinez-Millana, B. Valdivieso, C. Fernández-Llatas, Interactive Data Cleaning for Process Mining: A Case Study of an Outpatient Clinic's Appointment System, in: C. Di Francescomarino, R. Dijkman, U. Zdun (Eds.), Business Process Management Workshops, in: Lecture Notes in Business Information Processing, Springer, Cham, 2019, pp. 532–544, http://dx.doi.org/10.1007/978-3-030-37453-2_43.
[31] R. Andrews, S. Suriadi, C. Ouyang, E. Poppe, Towards Event Log Querying for Data Quality: Let's Start with Detecting Log Imperfections, in: H. Panetto, C. Debruyne, H.A. Proper, C.A. Ardagna, D. Roman, R. Meersman (Eds.), On the Move To Meaningful Internet Systems. OTM 2018 Conferences, in: Lecture Notes in Computer Science, Springer, Cham, 2018, pp. 116–134, http://dx.doi.org/10.1007/978-3-030-02610-3_7.
[32] S. Feuerriegel, J. Hartmann, C. Janiesch, P. Zschech, Generative AI, Bus. Inf. Syst. Eng. 66 (1) (2024) 111–126, http://dx.doi.org/10.1007/s12599-023-00834-7.
[33] K. Peffers, T. Tuunanen, M.A. Rothenberger, S. Chatterjee, A Design Science Research Methodology for Information Systems Research, J. Manag. Inf. Syst. 24 (3) (2007) 45–77, http://dx.doi.org/10.2753/MIS0742-1222240302.
[34] J. Venable, J. Pries-Heje, R. Baskerville, FEDS: a Framework for Evaluation in Design Science Research, Eur. J. Inf. Syst. 25 (1) (2016) 77–89, http://dx.doi.org/10.1057/ejis.2014.36.
[35] M.L. van Eck, X. Lu, S.J.J. Leemans, W.M.P. van der Aalst, PM^2: A Process Mining Project Methodology, in: J. Zdravkovic, M. Kirikova, P. Johannesson (Eds.), Advanced Information Systems Engineering, in: Lecture Notes in Computer Science, Springer, Cham, 2015, pp. 297–313, http://dx.doi.org/10.1007/978-3-319-19069-3_19.
[36] V. Stein Dani, H. Leopold, J.M.E.M. van der Werf, X. Lu, I. Beerepoot, J.J. Koorn, H.A. Reijers, Towards Understanding the Role of the Human in Event Log Extraction, in: A. Marrella, B. Weber (Eds.), Business Process Management Workshops, in: Lecture Notes in Business Information Processing, Springer, Cham, 2022, pp. 86–98, http://dx.doi.org/10.1007/978-3-030-94343-1_7.
[37] L. Reinkemeyer, Process Mining in Action: Principles, Use Cases and Outlook, Springer, Cham, 2020, http://dx.doi.org/10.1007/978-3-030-40172-6.
[38] F. Milani, K. Lashkevich, F.M. Maggi, C. Di Francescomarino, Process Mining: A Guide for Practitioners, in: R. Guizzardi, J. Ralyté, X. Franch (Eds.), Research Challenges in Information Science, in: Lecture Notes in Business Information Processing, Springer, Cham, 2022, pp. 265–282, http://dx.doi.org/10.1007/978-3-031-05760-1_16.
[39] J. De Weerdt, M.T. Wynn, Foundations of Process Event Data, in: W.M.P. van der Aalst, J. Carmona (Eds.), Process Mining Handbook, in: Lecture Notes in Business Information Processing, Springer, Cham, 2022, pp. 193–211, http://dx.doi.org/10.1007/978-3-031-08848-3_6.
[40] Institute of Electrical and Electronics Engineers, IEEE Standard for eXtensible Event Stream (XES) for Achieving Interoperability in Event Logs and Event Streams, 2016, http://dx.doi.org/10.1109/IEEESTD.2016.7740858, IEEE 1849-2016.
[41] A. Berti, I. Koren, J.N. Adams, G. Park, B. Knopp, N. Graves, et al., OCEL (Object-Centric Event Log) 2.0 Specification, 2024, arXiv:2403.01975.
[42] A.F. Ghahfarokhi, G. Park, A. Berti, W.M.P. van der Aalst, OCEL: A Standard for Object-Centric Event Logs, in: L. Bellatreche, M. Dumas, P. Karras, R. Matulevičius, A. Awad, M. Weidlich, et al. (Eds.), New Trends in Database and Information Systems, in: Communications in Computer and Information Science, Springer, Cham, 2021, pp. 169–175, http://dx.doi.org/10.1007/978-3-030-85082-1_16.
[43] S.K. Pradhan, M. Jans, N. Martin, Getting the Data in Shape for Your Process Mining Analysis: An In-Depth Analysis of the Pre-Analysis Stage, ACM Comput. Surv. 57 (6) (2025) 1–37, http://dx.doi.org/10.1145/3712587.
[44] Y. Liu, V. Stein Dani, I. Beerepoot, X. Lu, Turning Logs into Lumber: Preprocessing Tasks in Process Mining, in: J. De Smedt, P. Soffer (Eds.), Process Mining Workshops, in: Lecture Notes in Business Information Processing, Springer, Cham, 2024, pp. 98–109, http://dx.doi.org/10.1007/978-3-031-56107-8_8.
[45] X. Lu, D. Fahland, W.M.P. van der Aalst, Interactively Exploring Logs and Mining Models with Clustering, Filtering, and Relabeling, in: L. Azevedo, C. Cabanillas (Eds.), BPM Demo Track 2016, 2016, pp. 44–49, URL: https://ceur-ws.org/Vol-1789/bpm-demo-2016-paper9.pdf.

[46] S.J. van Zelst, F. Mannhardt, M. de Leoni, A. Koschmider, Event abstraction in process mining: literature review and taxonomy, Granul. Comput. 6 (3) (2021) 719–736, http://dx.doi.org/10.1007/s41066-020-00226-2.

[47] A. Beheshti, J. Yang, Q.Z. Sheng, B. Benatallah, F. Casati, S. Dustdar, et al., ProcessGPT: Transforming Business Process Management with Generative Artificial Intelligence, in: C. Ardagna, B. Benatallah, H. Bian, C.K. Chang, R.N. Chang, J. Fan, et al. (Eds.), 2023 IEEE International Conference on Web Services (ICWS), 2023, pp. 731–739, http://dx.doi.org/10.1109/ICWS60048.2023.00099.

[48] D. Schuster, S.J. van Zelst, W.M.P. van der Aalst, Utilizing domain knowledge in data-driven process discovery: A literature review, Comput. Ind. 137 (2022) 103612, http://dx.doi.org/10.1016/j.compind.2022.103612.

[49] P. Federico, M. Wagner, A. Rind, A. Amor-Amorós, S. Miksch, W. Aigner, The Role of Explicit Knowledge: A Conceptual Model of Knowledge-Assisted Visual Analytics, in: B. Fisher, S. Liu, T. Schreck (Eds.), 2017 IEEE Conference on Visual Analytics Science and Technology (VAST), 2017, pp. 92–103, http://dx.doi.org/10.1109/VAST.2017.8585498.

[50] D.T. Kapugama Geeganage, M.T. Wynn, A.H.M. ter Hofstede, Text2EL+: Expert Guided Event Log Enrichment Using Unstructured Text, ACM J. Data Inf. Qual. 16 (1) (2024) 1–28, http://dx.doi.org/10.1145/3640018.

[51] F. Zerbato, P. Soffer, B. Weber, Process Mining Practices: Evidence from Interviews, in: C. Di Ciccio, R. Dijkman, A. del Río Ortega, S. Rinderle-Ma (Eds.), Business Process Management, in: Lecture Notes in Computer Science, Springer, Cham, 2022, pp. 268–285, http://dx.doi.org/10.1007/978-3-031-16103-2_19.

[52] Celonis, Validating data with Data Explorer, 2025, https://docs.celonis.com/en/validating-data-with-data-explorer.html. (Accessed 07 June 2025).

[53] R. Andrews, M.T. Wynn, K. Vallmuur, A.H.M. ter Hofstede, E. Bosley, M. Elcock, S. Rashford, Leveraging Data Quality to Better Prepare for Process Mining: An Approach Illustrated Through Analysing Road Trauma Pre-Hospital Retrieval and Transport Processes in Queensland, Int. J. Environ. Res. Public Health 16 (7) (2019) 1138, http://dx.doi.org/10.3390/ijerph16071138.

[54] E. Brynjolfsson, T. Mitchell, What can machine learning do? Workforce implications, Science 358 (6370) (2017) 1530–1534, http://dx.doi.org/10.1126/science.aap8062.

[55] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed., Pearson, Harlow, 2021.

[56] C. Janiesch, P. Zschech, K. Heinrich, Machine learning and deep learning, Electron. Mark. 31 (3) (2021) 685–695, http://dx.doi.org/10.1007/s12525-021-00475-2.

[57] A. Berti, H. Kourani, H. Häfke, C.-Y. Li, D. Schuster, Evaluating Large Language Models in Process Mining: Capabilities, Benchmarks, and Evaluation Strategies, in: H. van der Aa, D. Bork, R. Schmidt, A. Sturm (Eds.), Enterprise, Business-Process and Information Systems Modeling, in: Lecture Notes in Business Information Processing, Springer, Cham, 2024, pp. 13–21, http://dx.doi.org/10.1007/978-3-031-61007-3_2.

[58] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, et al., Attention Is All You Need, 2017, arXiv:1706.03762.

[59] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, et al., A Survey on Evaluation of Large Language Models, ACM Trans. Intell. Syst. Technol. 15 (3) (2024) 1–45, http://dx.doi.org/10.1145/3641289.

[60] E.F. Codd, A relational model of data for large shared data banks, Commun. ACM 13 (6) (1970) 377–387, http://dx.doi.org/10.1145/362384.362685.

[61] P. Brützke, R. Killewald, S. Franzoi, J. vom Brocke, AI-Assisted Process Mining for Context-Sensitive Analysis Support, in: ECIS 2025 Proceedings, (7) 2025, URL: https://aisel.aisnet.org/ecis2025/bpm/bpm/7.

[62] H.T.C. Nguyen, S. Lee, J. Kim, J. Ko, M. Comuzzi, Autoencoders for improving quality of process event logs, Expert Syst. Appl. 131 (2019) 132–147, http://dx.doi.org/10.1016/j.eswa.2019.04.052.

[63] P. Wu, X. Fang, H. Fang, Z. Gong, D. Kan, An Event Log Repair Method Based on Masked Transformer Model, Appl. Artif. Intell. 38 (1) (2024) 2346059, http://dx.doi.org/10.1080/08839514.2024.2346059.

[64] S.J. Schmid, L. Moder, P. Hofmann, M. Röglinger, Everything at the proper time: Repairing identical timestamp errors in event logs with Generative Adversarial Networks, Inf. Syst. 118 (2023) 102246, http://dx.doi.org/10.1016/j.is.2023.102246.

[65] A.H.M. ter Hofstede, A. Koschmider, A. Marrella, R. Andrews, D.A. Fischer, S. Sadeghianasl, et al., Process-Data Quality: The True Frontier of Process Mining, ACM J. Data Inf. Qual. 15 (3) (2023) 1–21, http://dx.doi.org/10.1145/3613247.

[66] P. Filipp, R. Dorsch, A. Harth, EVErPREP: Towards an Event Knowledge Graph Enhanced Workflow Model for Event Log Preparation, in: A. Delgado, T. Slaats (Eds.), Process Mining Workshops, in: Lecture Notes in Business Information Processing, Springer, Cham, 2025, pp. 31–43, http://dx.doi.org/10.1007/978-3-031-82225-4_3.

[67] C. Di Francescomarino, C. Ghidini, S. Tessaris, I.V. Sandoval, Completing Workflow Traces Using Action Languages, in: J. Zdravkovic, M. Kirikova, P. Johannesson (Eds.), Advanced Information Systems Engineering, in: Lecture Notes in Computer Science, Springer, Cham, 2015, pp. 314–330, http://dx.doi.org/10.1007/978-3-319-19069-3_20.

[68] A. Rogge-Solti, R.S. Mans, W.M.P. van der Aalst, M. Weske, Repairing Event Logs Using Timed Process Models, in: Y.T. Demey, H. Panetto (Eds.), On the Move To Meaningful Internet Systems: OTM 2013 Workshops, in: Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2013, pp. 705–708, http://dx.doi.org/10.1007/978-3-642-41033-8_89.

[69] P.M. Dixit, S. Suriadi, R. Andrews, M.T. Wynn, A.H.M. ter Hofstede, J.C.A.M. Buijs, W.M.P. van der Aalst, Detection and Interactive Repair of Event Ordering Imperfection in Process Logs, in: J. Krogstie, H.A. Reijers (Eds.), Advanced Information Systems Engineering, in: Lecture Notes in Computer Science, Springer, Cham, 2018, pp. 274–290, http://dx.doi.org/10.1007/978-3-319-91563-0_17.

[70] W. Song, H.-A. Jacobsen, P. Zhang, Self-Healing Event Logs, IEEE Trans. Knowl. Data Eng. 33 (6) (2021) 2750–2763, http://dx.doi.org/10.1109/TKDE.2019.2956520.

[71] S. Gregor, A.R. Hevner, Positioning and Presenting Design Science Research for Maximum Impact, MIS Q. 37 (2) (2013) 337–355, http://dx.doi.org/10.25300/MISQ/2013/37.2.01.

[72] A.R. Hevner, S.T. March, J. Park, S. Ram, Design Science in Information Systems Research, MIS Q. 28 (1) (2004) 75–105, http://dx.doi.org/10.2307/25148625.

[73] T. Tuunanen, R. Winter, J. vom Brocke, Dealing with Complexity in Design Science Research: A Methodology Using Design Echelons, MIS Q. 48 (2) (2024) 427–458, http://dx.doi.org/10.25300/MISQ/2023/16700.

[74] S.T. March, G.F. Smith, Design and natural science research on information technology, Decis. Support Syst. 15 (4) (1995) 251–266, http://dx.doi.org/10.1016/0167-9236(94)00041-2.

[75] E. Mumford, A Socio-Technical Approach to Systems Design, Requir. Eng. 5 (2) (2000) 125–133, http://dx.doi.org/10.1007/PL00010345.

[76] C. Di Ciccio, S. Miksch, P. Soffer, B. Weber, G. Meroni, Human in the (Process) Mines (Dagstuhl Seminar 23271), Dagstuhl Rep. 13 (7) (2024) 1–33, http://dx.doi.org/10.4230/DagRep.13.7.1.

[77] T. Gschwandtner, Visual Analytics Meets Process Mining: Challenges and Opportunities, in: P. Ceravolo, S. Rinderle-Ma (Eds.), Data-Driven Process Discovery and Analysis, in: Lecture Notes in Business Information Processing, Springer, Cham, 2017, pp. 142–154, http://dx.doi.org/10.1007/978-3-319-53435-0_7.

[78] S. Miksch, C. Di Ciccio, P. Soffer, B. Weber, Visual Analytics Meets Process Mining: Challenges and Opportunities, IEEE Comput. Graph. Appl. 44 (6) (2024) 132–141, http://dx.doi.org/10.1109/MCG.2024.3456916.

[79] S. van den Elzen, M. Jans, N. Martin, F. Pieters, C. Tominski, M.-C. Villa-Uriol, S.J. van Zelst, Towards Multi-Faceted Visual Process Analytics, Inf. Syst. 133 (2025) 102560, http://dx.doi.org/10.1016/j.is.2025.102560.

[80] B. Shneiderman, The eyes have it: a task by data type taxonomy for information visualizations, in: Proceedings 1996 IEEE Symposium on Visual Languages, IEEE Computer Society Press, Los Alamitos, 1996, pp. 336–343, http://dx.doi.org/10.1109/VL.1996.545307.

[81] D. Ceneda, T. Gschwandtner, T. May, S. Miksch, H.-J. Schulz, M. Streit, C. Tominski, Characterizing Guidance in Visual Analytics, IEEE Trans. Vis. Comput. Graph. 23 (1) (2017) 111–120, http://dx.doi.org/10.1109/TVCG.2016.2598468.

[82] D. Ceneda, N. Andrienko, G. Andrienko, T. Gschwandtner, S. Miksch, N. Piccolotto, et al., Guide Me in Analysis: A Framework for Guidance Designers, Comput. Graph. Forum 39 (6) (2020) 269–288, http://dx.doi.org/10.1111/cgf.14017.

[83] J. Johnson, A. Henderson, Conceptual Models: Core to the Design of Interactive Applications, 2nd ed., Springer, Cham, 2024, http://dx.doi.org/10.1007/978-3-031-50852-3.

[84] M. Galster, P. Avgeriou, Empirically-grounded reference architectures, in: QoSA-ISARCS '11: Proceedings of the joint ACM SIGSOFT conference – QoSA and ACM SIGSOFT symposium – ISARCS on Quality of software architectures – QoSA and architecting critical systems – ISARCS, 2011, pp. 153–158, http://dx.doi.org/10.1145/2000259.2000285.

[85] J.J. van Wijk, The value of visualization, in: VIS 05. IEEE Visualization, 2005, 2005, pp. 79–86, http://dx.doi.org/10.1109/VISUAL.2005.1532781.

[86] A. Berti, D. Schuster, W.M.P. van der Aalst, Abstractions, Scenarios, and Prompt Definitions for Process Mining with LLMs: A Case Study, in: J. De Weerdt, L. Pufahl (Eds.), Business Process Management Workshops, in: Lecture Notes in Business Information Processing, Springer, Cham, 2024, pp. 427–439, http://dx.doi.org/10.1007/978-3-031-50974-2_32.

[87] F. Monti, F. Leotta, J. Mangler, M. Mecella, S. Rinderle-Ma, NL2ProcessOps: Towards LLM-Guided Code Generation for Process Execution, in: A. Marrella, M. Resinas, M. Jans, M. Rosemann (Eds.), Business Process Management Forum, in: Lecture Notes in Business Information Processing, Springer, Cham, 2024, pp. 127–143, http://dx.doi.org/10.1007/978-3-031-70418-5_8.

[88] M. Polato, Dataset belonging to the help desk log of an Italian Company [dataset], 2017, http://dx.doi.org/10.4121/UUID:0C60EDF1-6F83-4E75-9367-4C63B3E9D5BB, 4TU.Centre for Research Data.

[89] C. Sonnenberg, J. vom Brocke, Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research, in: K. Peffers, M. Rothenberger, B. Kuechler (Eds.), Design Science Research in Information Systems: Advances in Theory and Practice, in: Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2012, pp. 381–397, http://dx.doi.org/10.1007/978-3-642-29863-9_28.

[90] F.D. Davis, Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, MIS Q. 13 (3) (1989) 319–340, http://dx.doi.org/10.2307/249008.

[91] C. Klinkmüller, R. Müller, I. Weber, Mining Process Mining Practices: An Exploratory Characterization of Information Needs in Process Analytics, in: T. Hildebrandt, B.F. van Dongen, M. Röglinger, J. Mendling (Eds.), Business Process Management, in: Lecture Notes in Computer Science, Springer, Cham, 2019, pp. 322–337, http://dx.doi.org/10.1007/978-3-030-26619-6_21.

[92] A. Yeshchenko, J. Mendling, A survey of approaches for event sequence analysis and visualization, Inf. Syst. 120 (2024) 102283, http://dx.doi.org/10.1016/j.is.2023.102283.

[93] D. Sacha, A. Stoffel, F. Stoffel, B.C. Kwon, G. Ellis, D.A. Keim, Knowledge Generation Model for Visual Analytics, IEEE Trans. Vis. Comput. Graph. 20 (12) (2014) 1604–1613, http://dx.doi.org/10.1109/TVCG.2014.2346481.

[94] R.S. Mans, W.M.P. van der Aalst, R.J.B. Vanwersch, A.J. Moleman, Process Mining in Healthcare: Data Challenges When Answering Frequently Posed Questions, in: R. Lenz, S. Miksch, M. Peleg, M. Reichert, D. Riaño, A. Teije (Eds.), Process Support and Knowledge Representation in Health Care, in: Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2013, pp. 140–153, http://dx.doi.org/10.1007/978-3-642-36438-9_10.

[95] A. Berti, S. van Zelst, D. Schuster, PM4Py: A process mining library for Python, Softw. Impacts 17 (2023) 100556, http://dx.doi.org/10.1016/j.simpa.2023.100556.

[96] Matplotlib, Matplotlib: Visualization with Python, 2025, https://matplotlib.org/. (Accessed 29 June 2025).

[97] Microsoft, Azure OpenAI in Azure AI Foundry Models, 2025, https://learn.microsoft.com/en-us/azure/ai-services/openai/concepts/models. (Accessed 29 June 2025).

[98] C. Di Ciccio, M. Montali, Declarative Process Specifications: Reasoning, Discovery, Monitoring, in: W.M.P. van der Aalst, J. Carmona (Eds.), Process Mining Handbook, in: Lecture Notes in Business Information Processing, Springer, Cham, 2022, pp. 108–152, http://dx.doi.org/10.1007/978-3-031-08848-3_4.

[99] T. Fehrer, L. Moder, M. Röglinger, An interactive approach for group-based event log exploration, Inf. Syst. 134 (2025) 102575, http://dx.doi.org/10.1016/j.is.2025.102575.

[100] M. Vidgof, D. Djurica, S. Bala, J. Mendling, Interactive log-delta analysis using multi-range filtering, Softw. Syst. Model. 21 (3) (2022) 847–868, http://dx.doi.org/10.1007/s10270-021-00902-0.

[101] A. Seeliger, A. Sánchez Guinea, T. Nolle, M. Mühlhäuser, ProcessExplorer: Intelligent Process Mining Guidance, in: T. Hildebrandt, B.F. van Dongen, M. Röglinger, J. Mendling (Eds.), Business Process Management, in: Lecture Notes in Computer Science, Springer, Cham, 2019, pp. 216–231, http://dx.doi.org/10.1007/978-3-030-26619-6_15.

[102] Y. Fontenla-Seco, M. Lama, A. Bugarín, Process-To-Text: A Framework for the Quantitative Description of Processes in Natural Language, in: F. Heintz, M. Milano, B. O'Sullivan (Eds.), Trustworthy AI - Integrating Learning, Optimization and Reasoning, in: Lecture Notes in Computer Science, Springer, Cham, 2021, pp. 212–219, http://dx.doi.org/10.1007/978-3-030-73959-1_19.

[103] S. Fiol González, L. Schirmer, L.Q. Campagnolo, A.M.B. Rodrigues, G.G. Schardong, R. França, et al., Visual Filtering Tools and Analysis of Case Groups for Process Discovery, in: S. Hammoudi, M. Śmiałek, O. Camp, J. Filipe (Eds.), Enterprise Information Systems, in: Lecture Notes in Business Information Processing, Springer, Cham, 2019, pp. 303–323, http://dx.doi.org/10.1007/978-3-030-26169-6_15.

[104] S. Yang, X. Dong, Z. Moliang, X. Li, S. Chen, R. Webmann, et al., VIT-PLA: Visual Interactive Tool for Process Log Analysis, in: P. Chau, J. Vreeken, M. van Leeuwen, D. Shahaf, C. Faloutsos (Eds.), Proceedings of the ACM SIGKDD 2016 Full-day Workshop on Interactive Data Exploration and Analytics, 2016, pp. 130–137, URL: https://poloclub.gatech.edu/idea2016/papers/p130-yang.pdf.

[105] J.K. Höhne, D. Krebs, S.-M. Kühnel, Measurement properties of completely and end labeled unipolar and bipolar scales in Likert-type questions on income (in)equality, Soc. Sci. Res. 97 (2021) 102544, http://dx.doi.org/10.1016/j.ssresearch.2021.102544.

[106] J.F. Nunamaker Jr., R.O. Briggs, D.C. Derrick, G. Schwabe, The Last Research Mile: Achieving Both Rigor and Relevance in Information Systems Research, J. Manag. Inf. Syst. 32 (3) (2015) 10–47, http://dx.doi.org/10.1080/07421222.2015.1094961.