From Event Logs to Managerial Decision Intelligence: A Business Data Analytics Framework for Automated Workflow Discovery

Mengqiu Zhang[1], Fei Sun[1], Jinping Mao[1,*]

[1] School of Smart Finance and Business, Hefei College of Finance & Economics, Hefei, China 230601

* Email: 820813179@qq.com (Corresponding Author)

Received 29 January 2026 · Revised 15 March 2026 · Accepted 29 April 2026 · Published 30 September 2026

Abstract

Event logs generated by enterprise systems provide a detailed empirical record of how business workflows actually unfold, yet many organizations still translate this evidence into managerial decisions through fragmented dashboards, manually selected process-mining settings, and ad hoc interpretations of discovered models. This study develops a business data analytics framework that connects event-log preparation, automated workflow discovery, multi-criteria model evaluation, and managerial decision intelligence. The framework is inspired by recent advances in multi- objective process discovery, but it reframes the technical discovery problem as a broader management analytics task: selecting a process model not only for statistical quality, but also for interpretability, governance relevance, and operational decision value. An illustrative analysis using four representative enterprise event logs demonstrates how automated configuration search improves the balance among fitness, precision, generalization, simplicity, and managerial usefulness when compared with default and expert-manual discovery strategies. The results indicate that the automated decision-intelligence setting raises the average decision usefulness score from 0.7l3 to 0.858 while preserving acceptable runtime performance through parallel evaluation. The study contributes to business data analytics by offering a structured bridge from process evidence to management action, showing how Pareto-based discovery and preference- aware model selection support bottleneck identification, compliance monitoring, resource allocation, and workflow redesign. Practical implications are provided for managers seeking to convert operational event data into auditable, repeatable, and strategically meaningful workflow intelligence.

Keywords: Event logs; business data analytics; Automated workflow discovery; Process mining; Multi-criteria model selection; Decision intelligence; Workflow analytics; managerial interpretation

From Event Logs to Managerial Decision Intelligence: A Business Data Analytics Framework for Automated Workflow Discovery

1. Introduction

Enterprise work is increasingly recorded before it is understood. Order management systems, customer relationship management platforms, enterprise resource planning modules, warehouse systems, service desks, and robotic process automation platforms continuously produce event records that identify cases, activities, timestamps, users, systems, and outcomes. These records are not merely technical exhaust. They constitute a process-level evidence base that reveals waiting time, rework, handoff friction, exception patterns, and compliance deviations. The central managerial challenge is therefore no longer the absence of data. It is difficult to transform dense event streams into workflow knowledge that executives, process owners, and operational analysts are willing to trust and use. This positioning follows the broader view that business intelligence and analytics turn enterprise data into organizational impact (Chen et al.,2012).

Process mining has become a leading approach for converting event logs into discovered process models, conformance diagnostics, and performance insights. In principle, the value proposition is direct: organizations observe what happened, discover how work actually flowed, compare this behavior against expected procedures, and redesign the process where evidence indicates loss of efficiency or control. In practice, however, the path from an event log to a managerial decision remains fragile. The discovered model depends heavily on log preprocessing, activity filtering, lifecycle interpretation, discovery algorithm selection, and hyperparameter settings. Different configurations applied to the same log frequently yield different models, each emphasizing different aspects of the process. One model may fit the log very well but be too complex for managerial interpretation. Another may be simple and persuasive, but may omit important exceptional behavior that matters for compliance or service recovery. Process mining is commonly defined as a bridge between event data and process knowledge (van der Aalst,2012).

This configuration sensitivity creates an important gap in business data analytics. Workflow discovery is often treated as a technical modeling problem, whereas managers require an auditable decision procedure. A process owner deciding whether to automate approval, reallocate employees, redesign case routing, or revise service-level agreements needs to know why a specific model was selected and how that selection reflects the organization's priorities. An analyst may prefer fitness and precision when the task is forensic investigation. A manager may prefer simplicity and interpretability when the task is communication to a steering committee. A risk officer may prefer the inclusion of rare exceptions when the task is internal control. Thus, workflow discovery becomes a multi-criteria decision problem rather than a single-metric optimization problem. The importance of this problem is reinforced by evidence that the main quality dimensions of process discovery must be balanced rather than optimized in isolation (Buijs et al.,2014).

Recent software-oriented advances in automated process discovery show that the search over preprocessing strategies, algorithms, and hyperparameters may be framed as a multi-objective optimization task. This direction is highly relevant to management analytics because it replaces manual trial and error with a structured search for candidate models that represent different quality trade-offs. Nevertheless, technical automation alone is insufficient. A discovered model becomes decision intelligence only when it is connected to business semantics, decision roles, governance requirements, and managerial actions. The present study develops this connection by proposing a business data analytics framework that extends automated workflow discovery into a broader decision-intelligence architecture. The multi-objective logic is consistent with evolutionary search methods designed to approximate efficient trade-off frontiers (Deb et al.,2002).

The article titled From Event Logs to Managerial Decision Intelligence addresses three research questions. First, how should an enterprise transform raw event logs into a controlled analytical object suitable for automated workflow discovery? Second, how should multiple discovered process models be evaluated and selected when quality criteria conflict with one another? Third, how should the selected model be translated into managerial recommendations that remain traceable to the underlying log evidence? These questions are answered through a framework that integrates event evidence, discovery analytics, multi-criteria model selection, and decision interpretation. Treating these questions as decision questions is consistent with data- science research that links modeling to data-driven action (Provost and Fawcett,2013).

Figure 1 summarizes the conceptual architecture of the proposed framework. Unlike a conventional process- mining pipeline that terminates with a model, the proposed architecture treats a discovered workflow as an intermediate analytical artifact. The top layer is managerial decision intelligence: an evidence-based interpretation of what should be monitored, improved, automated, or governed. The middle layers concern model selection and automated discovery, while the bottom layer anchors the analysis in event evidence. The figure intentionally avoids a directional arrow layout to emphasize that the layers represent governance responsibilities as much as computational steps.

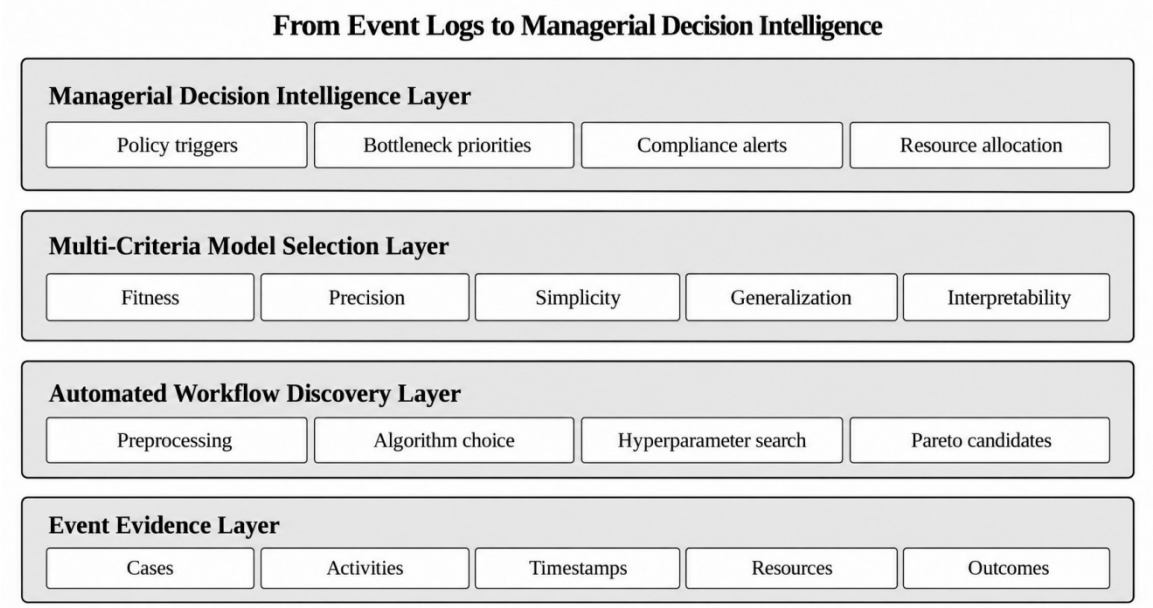


Figure 1. Business Data Analytics Framework for Workflow Discovery and Decision Intelligence

The framework contributes to the literature and practice of business data analytics in five ways. First, it reframes automated process discovery as a managerial decision-intelligence task rather than a purely technical search problem. Second, it introduces an explicit distinction between model quality and decision usefulness, recognizing that a technically strong model may still fail to support business action if it lacks interpretability or governance relevance. Third, it outlines an end-to-end analytical procedure from log ingestion to decision translation. Fourth, it provides an illustrative numerical analysis demonstrating how automated configuration search compares with default and expert-manual discovery strategies. Fifth, it identifies practical governance controls for repeatability, auditability, and responsible interpretation of discovered workflow models. This contribution also aligns with research that places process mining across organizational, analytical, and technological levels (vom Brocke et al.,2021).

The remainder of the article is organized as follows. Section 2 reviews the literature on process mining, automated workflow discovery, multi-criteria analytics, and managerial decision intelligence. Section 3 develops the proposed framework and defines its core components. Section 4 explains the

analytical procedure and the illustrative data context. Section 5 presents data analysis results. Section 6 discusses theoretical and managerial implications. Section 7 identifies limitations and future research directions. Section 8 concludes the article.

2. Literature Review

2.1 Process Mining and Event Log Analytics

Process mining emerged at the intersection of data mining and business process management. Its distinguishing feature is the use of event data to reveal, verify, and improve operational processes. Early work on workflow mining demonstrated that process models could be discovered from event logs rather than designed solely from interviews or procedural documentation. Subsequent research expanded the field into three closely related domains: discovery, conformance checking, and enhancement. Discovery derives a process model from observed behavior. Conformance checking compares observed behavior with a reference model. Enhancement uses event evidence to enrich or improve existing models through performance, organizational, or decision-related information. The original workflow-mining literature established that process models could be derived directly from event logs rather than only from interviews or documentation (van der Aalst et al.,2004).

Event-log analytics differs from conventional transaction reporting because it treats time-ordered sequences as the primary analytical unit. The same event set may indicate different managerial problems depending on how cases are defined and how activities are grouped. For instance, an invoice approval process may appear efficient when measured by average cycle time but may still show excessive rework in high-value cases. A customer-service process may satisfy average response targets while concealing a small group of cases that repeatedly transfer between teams. Process mining therefore contributes a behavioral view that complements traditional business intelligence dashboards. This sequence-oriented perspective complements predictive analytics research in information systems, where empirical models are evaluated by their usefulness for prediction and explanation (Shmueli and Koppius,2011).

The quality of the event log is a central precondition for reliable analysis. Event logs often contain incomplete timestamps, duplicated events, inconsistent activity labels, missing case identifiers, atypical lifecycle states, and resource names that have been anonymized or aggregated. These imperfections are not minor technical issues. They affect the discovered model and the conclusions drawn from it. A log with excessive activity granularity may create a spaghetti-like model that is difficult to interpret. A log that removes rare exceptions too aggressively may understate operational risk. A log that merges different product or customer segments may produce a model that fits no segment particularly well. Event-log imperfections such as missing timestamps, duplicated entries, and ambiguous activity labels are known to distort process-mining results (Suriadi et al.,2017).

Business data analytics therefore requires a disciplined log-preparation stage. This stage includes case definition, activity normalization, timestamp verification, outlier analysis, and segmentation. The preparation stage should not be hidden from managerial users, because its choices shape what counts as process behavior. In the proposed framework, log preparation is treated as a governance decision rather than as a purely technical cleaning activity. Analysts document each transformation and preserve the ability to reproduce the input log used for discovery. A review of event-log preprocessing likewise shows that data preparation is a central determinant of discovery reliability (Marin-Castro and Tello-Leal,2021).

2.2 Automated Workflow Discovery and Quality Metrics

Workflow discovery algorithms differ in their assumptions, output structures, and tolerance for noise. Inductive approaches often produce sound block-structured models, while heuristic approaches emphasize frequency-based dependencies and are useful when event logs include noise or informal behavior. Split-

based approaches frequently seek a balance between model accuracy and simplicity. Declarative discovery emphasizes constraints rather than procedural flows and may fit processes where flexibility is expected. No single algorithm dominates across all event logs, industries, or managerial goals. The appropriate discovery strategy depends on the structure of the log and the intended use of the model. The relevance of inductive and block-structured discovery follows from work showing how structured models can be discovered from event logs (Leemans et al.,2013). Model readability also matters because process-modeling guidelines associate understandable structure with better human interpretation (Mendling et al.,2010).

The algorithmic choice is only one part of the configuration problem. Preprocessing decisions, filtering thresholds, activity aggregation, variant selection, and hyperparameter values strongly influence the resulting model. Manual tuning remains common in organizational projects because analysts experiment with several settings until a model appears plausible. This practice is understandable, but it reduces repeatability and makes the analytical path difficult to defend. A process model selected through informal trial and error may support a persuasive presentation, yet the organization may not be able to reproduce why that model was selected rather than an alternative. This dependence on preprocessing is consistent with work on quality- informed event-log generation for process mining (Andrews et al.,2020).

Automated workflow discovery addresses this problem by searching over a configuration space. Each candidate configuration specifies a combination of preprocessing choices, discovery algorithm, algorithm variant, and parameter values. The candidate is evaluated using a set of quality metrics. Fitness measures the extent to which the model reproduces observed log behavior. Precision penalizes models that allow too much behavior not seen in the log. Generalization measures whether the model avoids excessive overfitting to specific traces. Simplicity captures the structural understandability of the model. Other metrics may include structural complexity, soundness, runtime, robustness, and domain-specific adequacy. Open process-mining libraries further lower the barrier to evaluating alternative log, model, and metric configurations (Berti, van Zelst, and Schuster,2023).

These metrics frequently conflict. A model with high fitness may be complex and low in simplicity. A highly precise model may exclude rare but meaningful exceptions. A simple model may communicate well but may hide important rework. This creates a natural multi-objective optimization problem. Instead of searching for one universally best model, automated discovery returns a set of non-dominated models representing different trade-offs. The managerial task then becomes selecting a model whose trade-off profile matches the decision context. This tension between fit and behavioral restriction is closely related to alignment-based conformance analysis (Adriansyah et al.,2011).

2.3 Multi-Criteria Decision Intelligence in Business Analytics

Multi-criteria decision analysis has long argued that decisions involving conflicting objectives should make preferences explicit. This principle is highly relevant to workflow discovery. A compliance investigation, a lean improvement project, an automation feasibility study, and a customer-experience redesign initiative do not require the same process model. A compliance investigation may place greater weight on exceptional paths and precision. A lean improvement project may place greater weight on bottlenecks and cycle-time clarity. An automation study may require a model that exposes repeatable, high-volume activities. A customer-experience project may privilege case-level pathways that explain delays for specific customer groups. The need to make preferences explicit is compatible with classic multi-criteria decision theory, especially when alternatives cannot be ranked by a single criterion (Saaty, 1987). More recent surveys of multi-criteria decision methods similarly emphasize explicit trade-off handling in complex managerial settings (Mardani et al.,2015).

Decision intelligence extends business intelligence by linking data, models, actions, and organizational objectives. It asks not only what the data show, but what decision should be made, by whom, with what evidence, and under what governance constraints. In the context of event logs, decision

intelligence requires a translation layer. The discovered model must be mapped to managerial questions such as which rework loops are economically material, where automation is justifiable, which teams are overloaded, which exceptions deserve policy revision, and which variants indicate operational risk. The decision-intelligence framing is also consistent with prescriptive analytics, where predictions become useful only when connected to decisions (Bertsimas and Kallus,2020).

The proposed framework uses multi-criteria model selection as the interface between technical discovery and managerial decision intelligence. Analysts first generate candidate models through automated discovery. Then they evaluate candidates using technical metrics and managerial criteria. The final model is selected using an explicit preference profile, such as efficiency-oriented, compliance-oriented, customer-oriented, or automation-oriented. The selected model is then interpreted through business rules and performance measures. This structure reduces the risk that a model is selected simply because it looks familiar or visually attractive. Preference-aware selection also benefits from explanation techniques that clarify why a model or algorithmic output should be trusted (Ribeiro et al.,2016).

2.4 Governance, Data Quality, and Managerial Interpretation

Governance is essential because process models influence organizational decisions. A discovered workflow may justify resource changes, control redesign, system modification, outsourcing, or automation investment. These decisions affect employees, customers, suppliers, and compliance obligations. Therefore, the analytical procedure should be transparent, reproducible, and documented. Governance includes input-data lineage, version control for log transformations, documented metric choices, recorded preference weights, and validation by domain experts. Such governance requirements are consistent with data-governance research that assigns decision rights and accountability for enterprise data assets (Khatri and Brown,2010).

Managerial interpretation is also a socio-technical activity. A model that is technically strong may be rejected if it contradicts how process owners understand their work, while a model that confirms existing beliefs may be accepted too quickly. The analyst's responsibility is to use event evidence to challenge assumptions without treating the model as an unquestionable truth. Domain experts should review whether discovered activities correspond to meaningful operational steps, whether rare paths represent errors or legitimate exceptions, and whether suggested interventions are feasible within the organization's constraints. The role of domain experts reflects the view that data quality is partly defined by fitness for consumer use rather than by technical accuracy alone (Wang and Strong, 1996).

The literature therefore points to an integrated problem. Process mining supplies methods for discovery and conformance. Automated search supplies a way to handle configuration complexity. Multi-criteria decision analysis supplies a way to select among competing models. Decision intelligence supplies the managerial framing that connects evidence to action. The proposed framework combines these perspectives into a single business data analytics procedure. This integration is consistent with business process management research that treats modeling, analysis, improvement, and governance as connected lifecycle activities (Dumas et al.,2018).

3. Business Data Analytics Framework for Automated Workflow Discovery

3.1 Problem Definition and Design Principles

The framework begins from a practical problem faced by process owners: an enterprise possesses event logs for a workflow, but the organization lacks a consistent method for selecting the process model that should guide operational decisions. The objective is not merely to maximize a technical metric. The objective is to produce a decision-ready representation of workflow behavior that supports improvement, control, and redesign. The framework is designed around five principles: evidence anchoring, configuration

transparency, multi-criteria evaluation, preference-aware selection, and decision translation. The framework therefore follows a design-science logic in which an artifact is constructed to solve a relevant organizational problem (Hevner et al.,2004).

Evidence anchoring means that each analytical claim should trace back to the event log and its preprocessing history. Configuration transparency means that the selected discovery pipeline should be recorded in enough detail for reproduction. Multi-criteria evaluation means that no single quality metric is treated as sufficient. Preference-aware selection means that the organization explicitly states whether the current decision context favors accuracy, simplicity, compliance sensitivity, or automation potential. Decision translation means that model outputs are converted into managerial statements about bottlenecks, variants, exceptions, resources, and policy changes. Documenting transformations also supports systematic assessment of data-quality dimensions such as completeness, consistency, and timeliness (Pipino et al.,2002).

Table 1. Core Components of the Proposed Business Data Analytics Framework

Component	Analytical content	Managerial role	Illustrative indicators
Event evidence layer	Raw and curated event logs, case identifiers, activities, timestamps, resources, and outcomes.	Defines the empirical boundary of the workflow under analysis.	Case completeness, timestamp validity, variant coverage
Discovery configuration layer	Candidate preprocessing rules, algorithms, variants, and hyperparameters.	Controls the model search space and determines reproducibility.	Search breadth, parameter traceability, algorithm diversity
Quality evaluation layer	Metrics such as fitness, precision, generalization, simplicity, runtime, and structural clarity.	Quantifies trade-offs among alternative models.	Metric balance, non-dominated candidates, stability
Preference selection layer	Managerial weights assigned to technical and business criteria.	Links model selection to the decision context.	Documented weights, scenario fit, stakeholder approval
Decision intelligence layer	Interpretive outputs such as bottleneck priorities, compliance warnings, automation opportunities, and redesign suggestions.	Turns discovered workflows into managerial action.	Action relevance, auditability, economic significance

Table 1 shows that automated workflow discovery is positioned inside a larger analytics architecture. The discovery algorithm is important, but it is not the only managerial concern. The quality of the event evidence, the search design, the selected evaluation metrics, and the decision context all shape the final interpretation. This broader view is particularly important for organizations that seek to institutionalize process mining rather than treat it as a one-off diagnostic exercise.

The framework also separates three types of value. Technical value concerns whether the discovered model represents the event log accurately. Cognitive value concerns whether managers understand and trust the model. Strategic value concerns whether the model supports a specific operational decision. An automated search procedure may improve technical value by exploring more configurations. A preference-aware selection procedure improves cognitive and strategic value by linking the final model to the decision context. The separation of technical, cognitive, and strategic value is compatible with management analytics as an interdisciplinary field connecting analytical methods with managerial action.

3.2 Event Log Preparation and Business Semantics

The first operational stage is the transformation of enterprise data into an analyzable event log. This stage begins by defining the case notion. In order-to-cash processes, a case may represent a sales order, a customer order line, or an invoice. In procurement processes, a case may represent a purchase requisition, a purchase order, or a supplier invoice. Each choice changes the meaning of the discovered workflow. The framework requires analysts to document the case definition and explain why it matches the managerial question. Tool- supported analysis depends on reliable interchange among event logs, discovery plugins, and process-model representations (van Dongen et al.,2005).

Activity labels are then normalized. Enterprise systems frequently use technical codes or localized labels that are difficult for managers to interpret. Analysts should map these codes into meaningful business activities without collapsing activities that have different control implications. Timestamp validation follows. Missing or reversed timestamps may create impossible event sequences, while duplicated records may inflate activity frequency. Resource fields require similar attention because employees, teams, or systems may be coded inconsistently across systems. The final curated log should remain connected to the original extraction through a data lineage record. The emphasis on validation also reflects the organizational cost of poor data quality in ordinary enterprise work (Redman, 1998).

Business semantics are added through attributes that support decision interpretation. These attributes may include customer segment, product family, order value, supplier risk tier, service priority, region, channel, exception reason, or approval authority. Segment-level discovery is often more useful than aggregate discovery because the same workflow may behave differently for high-value orders, rush cases, regulated products, or strategic suppliers. The framework therefore treats segmentation as a hypothesis about how management should interpret workflow behavior. Segment-level discovery is increasingly important because object-centric process mining shows that real processes may involve several interacting business objects (van der Aalst,2023).

3.3 Discovery Configuration Space

After the event log is prepared, the framework constructs a configuration space. This space includes preprocessing alternatives, discovery algorithms, and hyperparameters. Preprocessing alternatives may include variant filtering, infrequent-activity filtering, event abstraction, lifecycle selection, resource grouping, and noise treatment. Algorithm alternatives may include inductive, heuristic, split-based, and declarative discovery families. Hyperparameters may govern dependency thresholds, concurrency detection, noise thresholds, structural simplification, and loop handling. The inclusion of object-centric alternatives is relevant where process outcomes depend on interacting objects rather than a single case identifier (Gherissi et al.,2023).

The configuration space should be broad enough to capture meaningful alternatives but bounded enough to be computationally feasible. A common error is to include every possible setting without considering the managerial question. For example, a compliance-focused study should not remove rare pathways too aggressively because those pathways may be the most important evidence. An automation-focused study may reasonably prioritize frequent paths and stable routines. The framework therefore recommends that analysts define the configuration space with explicit design rules tied to the decision context. This bounded-search principle is especially important when automated discovery is connected to robotic-process-automation task segmentation (Agostinelli et al.,2021).

Automated search evaluates candidate configurations and returns a set of strong alternatives. The output is not a single result but a portfolio of process models. Each model represents a different trade-off among metrics. This portfolio structure matters because it creates an opportunity for managerial preference to enter the analysis transparently. Instead of disguising judgment as technical objectivity, the framework

records how business priorities influence the final model selection. A portfolio of candidate models can also support later resource-allocation decisions when predictive monitoring is used operationally (Park and Song,2023).

3.4 Multi-Criteria Model Evaluation

The evaluation layer assigns quality values to each candidate model. Fitness, precision, generalization, and simplicity form the technical base. However, the proposed framework adds managerial criteria: interpretability, diagnostic relevance, actionability, and governance sensitivity. Interpretability refers to whether managers are able to explain the model without excessive technical mediation. Diagnostic relevance refers to whether the model exposes the operational issue under investigation. Actionability refers to whether the model supports a specific intervention. Governance sensitivity refers to whether the model preserves paths that matter for audit, compliance, or risk management. Outcome-oriented predictive monitoring research similarly shows that evaluation choices affect the practical value of process predictions (Teinemaa et al.,2019).

Because some managerial criteria are qualitative, the framework uses structured scoring rather than pretending that all business value is purely numerical. Domain experts may rate interpretability and actionability using standardized rubrics. Analysts should report these scores alongside technical metrics, not in place of them. A model with excellent technical metrics but weak actionability may be useful for research comparison but less useful for management. Conversely, a simple model with moderate fitness may be acceptable for a strategic communication task if its limitations are clearly reported. Toolkits for predictive process monitoring further demonstrate the need to compare models, features, and explanations within a controlled pipeline (Rizzi et al.,2025).

The final evaluation produces a Pareto-oriented model portfolio. A candidate is non-dominated when no other candidate performs better on all relevant criteria. The portfolio is then filtered through preference profiles. For example, an efficiency profile may assign greater weight to bottleneck visibility and model simplicity. A compliance profile may assign greater weight to precision and rare-path retention. A digital automation profile may prioritize frequent-path clarity and structural regularity. This process makes the managerial logic of model selection explicit. This portfolio view is also reinforced by explainable predictive process analytics, where users need reasons for model recommendations, not only scores (Galanti et al.,2023a).

3.5 Managerial Decision Intelligence Layer

The decision intelligence layer converts the selected model into managerial outputs. These outputs include process maps, variant rankings, bottleneck diagnostics, rework-loop analysis, handoff analysis, conformance exceptions, and intervention suggestions. The key difference from a conventional process-mining report is that each output is linked to an explicit decision. A bottleneck indicator should be connected to a resource allocation question. A rework loop should be connected to a quality or policy question. An exceptional path should be connected to a compliance or training question. The emphasis on bottlenecks, variants, and redesign levers is consistent with the process-redesign literature on improvement heuristics (Reijers and Mansar,2005).

The framework recommends four decision translation templates. The first is the bottleneck template, which identifies where waiting time accumulates and what managerial lever may reduce it. The second is the variant-control template, which identifies high-frequency variants and evaluates whether they should be standardized, automated, or preserved for flexibility. The third is the compliance template, which identifies deviations from required sequences and differentiates legitimate exceptions from control failures. The fourth is the automation template, which identifies stable, high-volume, low-exception activity patterns

suitable for digital automation or workflow-rule implementation. The compliance template also draws on auditing research where event logs serve as analytical evidence for control assessment (Jans et al.,2014).

Each template requires managerial validation. Event logs show behavior but not always intent. A repeated approval loop may indicate poor system design, but it may also represent prudent control for high-risk cases. A rare path may indicate a violation, but it may also represent a legitimate emergency procedure. The decision intelligence layer therefore combines model evidence with domain interpretation. The goal is not to replace managerial judgment, but to make judgment evidence-aware and reproducible. Managerial validation is especially important because process-mining transparency can change how organizations perceive and debate their own work (Eggers et al.,2021).

4. Research Design and Analytical Procedure

4.1 Illustrative Data Context

To demonstrate the proposed framework, this study uses an illustrative analytical design built around four representative enterprise event-log contexts: procure-to-pay, order-to-cash, IT incident management, and HR onboarding. These processes were selected because they differ in frequency, regulatory relevance, cycle-time structure, and managerial interpretation. The procure-to-pay process emphasizes supplier coordination, approval control, and invoice matching. The order-to-cash process emphasizes revenue flow, customer service, and fulfillment timeliness. The IT incident process emphasizes service responsiveness and escalation behavior. The HR onboarding process emphasizes cross-functional coordination and compliance with administrative milestones. Healthcare process analysis illustrates why different operational contexts require different discovery and interpretation choices (Rebuge and Ferreira,2012). Cross-organizational process analysis similarly shows that workflows may span firms, messages, and performance measures (Engel et al.,2016).

The analysis does not claim that the numerical values represent a particular organization. Instead, the values are structured to resemble common enterprise event-log profiles and to illustrate how the framework operates. The data include case counts, event counts, number of activities, variant counts, average cycle time, and estimated noise rate. Each log is treated as a separate decision context, and model selection is performed according to a workflow goal. This design allows the paper to show how the same discovery framework adapts to different managerial questions.

Table 2. Event-Log Contexts Used in the Illustrative Analysis

Process context	Cases	Events	Activities	Variants	Avg. cycle time	Primary managerial question
Procure-to-pay	15,200	128,940	18	3, 160	8.4 days	Supplier approval and invoice matching
Order-to-cash	22,500	180,300	16	1,870	5.2 days	Fulfillment delay and customer promise reliability
IT incident management	31, 100	210,440	14	2,940	1.6 days	Escalation control and service recovery
HR onboarding	7,800	62,400	12	960	11.5 days	Administrative compliance and employee readiness

Table 2 highlights the diversity of workflow conditions that an automated discovery framework must address. The IT incident log has the largest number of cases and the shortest average cycle time, suggesting a high- volume service environment where escalation and rapid resolution are managerial priorities. The HR onboarding log has fewer cases but a longer cycle time, indicating a process in which coordination quality and missing milestones may matter more than sheer volume. The procure-to-pay log has high variant complexity, making it an appropriate setting for variant control and exception analysis.

The four logs also illustrate why default discovery settings are rarely sufficient. A default setting may overgeneralize the procure-to-pay process and fail to expose approval exceptions. It may underrepresent escalation behavior in IT incidents. It may generate a model for HR onboarding that is accurate but too detailed for policy owners. Automated configuration search provides a more disciplined way to handle these differences by evaluating multiple alternatives under process-specific decision priorities. Similar logic has been applied in manufacturing settings where process-mining diagnostics support productivity improvement (Lorenz et al.,2021).

4.2 Candidate Discovery Strategies

The analysis compares three discovery strategies. The first strategy is the default setting, which applies a standard discovery algorithm with minimal tuning and conventional preprocessing. This strategy approximates the starting point used in many exploratory process-mining projects. The second strategy is expert-manual tuning, in which an analyst adjusts filters and algorithm parameters based on domain knowledge and visual inspection. This strategy reflects common practice in consulting and internal analytics teams. The third strategy is automated managerial decision intelligence, abbreviated as automated MDI. This strategy searches over a configuration space and selects a model through a preference-aware evaluation procedure. Comparing strategies also requires a structured project method so that discovery, evaluation, and interpretation remain reproducible (Aguirre et al.,2017).

For each process context, candidate models are evaluated using fitness, precision, generalization, simplicity, and decision usefulness. Decision usefulness is a composite managerial assessment based on interpretability, diagnostic relevance, and actionability. Runtime is also reported because computational cost affects whether the approach is practical for recurring organizational use. Automated MDI is expected to improve the quality balance and decision usefulness, but it may require more runtime than default or manual strategies because it evaluates multiple candidates. Decision usefulness is included because analytics capabilities produce value when they are connected to dynamic organizational capabilities (Mikalef et al.,2020).

4.3 Decision Scenarios and Preference Profiles

Preference profiles translate business intent into model selection criteria. Four profiles were defined for the illustrative analysis. The efficiency profile emphasizes bottleneck visibility, simplicity, and cycle-time interpretation. It is applied to order-to-cash because managers are concerned with promise reliability and fulfillment delay. The compliance profile emphasizes precision, exception retention, and governance sensitivity. It is applied to HR onboarding because administrative steps and required milestones must be monitored. The supplier-control profile emphasizes variant visibility, approval loops, and rework. It is applied to procure-to-pay. The service-recovery profile emphasizes escalation paths, queue transitions, and high-volume actionability. It is applied to IT incident management. Preference profiles also reflect the strategic alignment logic used in studies of analytics capability and firm performance (Wamba et al.,2017).

These profiles show that the model selection step is not value-neutral. A manager seeking standardization may prefer a model that compresses low-frequency variation. A risk officer may reject that same model if it removes important exceptions. A digital transformation team may prefer a model that highlights stable routines even if it provides less detail about edge cases. Preference profiles allow these

differences to be stated directly and recorded as part of the analytical procedure. This supports the argument that model selection should be tied to business value rather than treated as an isolated technical exercise (Grover et al.,2018).

4.4 Analytical Workflow and Validation

The analytical procedure follows six steps. First, the raw event log is profiled for case completeness, timestamp validity, activity frequency, and variant concentration. Second, the log is transformed into a curated event set with documented preprocessing decisions. Third, the candidate configuration space is defined. Fourth, candidate process models are generated and evaluated against technical and managerial metrics. Fifth, non-dominated candidates are ranked using the relevant preference profile. Sixth, the selected model is translated into managerial recommendations and validated by domain experts. Stability checks are important because sampling and evaluation choices can preserve or distort monitoring performance (Fani Sani et al.,2023).

Validation occurs at two levels. Technical validation examines whether the selected model achieves acceptable metric values and whether the results remain stable under modest changes in filtering thresholds. Managerial validation examines whether the model explains recognizable process behavior and whether proposed actions are feasible. The two forms of validation are complementary. A technically strong model that managers cannot interpret has limited business value. A managerially intuitive model that lacks event-log support may reinforce existing myths. Figure 2 presents the analytical workflow used in the study. Technical validation can also identify when discovered or reference models require repair to better reflect observed behavior (Fahland and van der Aalst,2015).

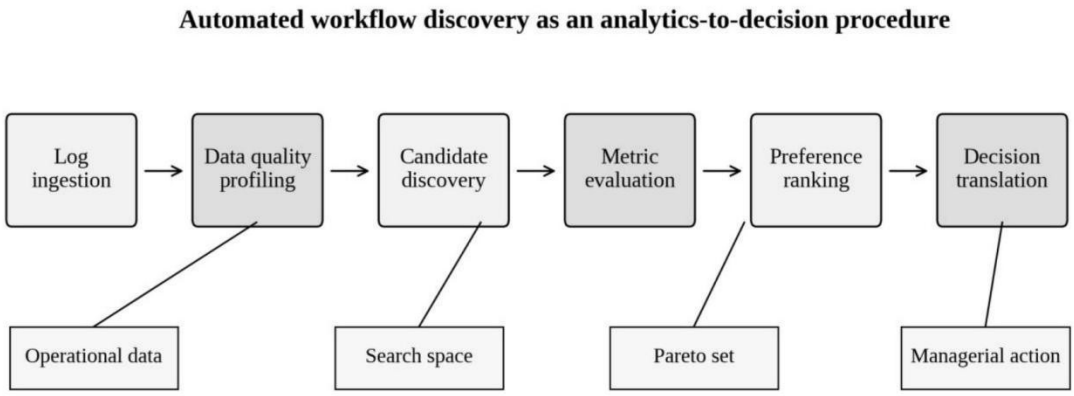


Figure 2. Analytical Workflow for Discovery, Evaluation, and Decision Translation

The workflow shown in Figure 2 emphasizes the repeated interaction between evidence, computation, and decision logic. Event data enter the analysis through log ingestion and quality profiling. Automated discovery explores the configuration space and generates candidate models. Metric evaluation makes model trade-offs visible. Preference ranking selects a model for the current decision context. Decision translation converts the model into managerial statements. Each stage leaves an audit trail, which supports repeatability and organizational learning.

5. Results and Data Analysis

5.1 Event-Log Diagnostic Findings

The initial diagnostic analysis indicates that each log presents a different discovery challenge. The procure- to-pay process shows a high number of variants relative to its case count. Approximately 22% of cases fall into low-frequency variants, suggesting heterogeneous supplier behavior, repeated approval loops, or exception handling. The order-to-cash process shows a more concentrated variant structure, with the ten most common variants covering a larger share of cases. This pattern supports an efficiency-oriented analysis focused on cycle-time bottlenecks rather than broad exception discovery. Such differences are familiar in healthcare and service contexts, where process-mining applications show strong variation in data structure and analytical purpose (Erdogan and Tarhan,2018).

The IT incident process shows high case volume and multiple escalation paths. Variant diversity is partly caused by differences in incident severity, support group assignment, and customer response time. The HR onboarding process shows the lowest event count but the longest average cycle time. Its main issue is not model scale but the identification of delayed administrative milestones. These diagnostics support the use of distinct preference profiles rather than a single evaluation template across all logs. This finding is also relevant for predictive monitoring because sequence models learn escalation and timing patterns from running cases (Tax et al.,2017).

Across all logs, the most common data-quality issues are activity naming inconsistency, duplicate status events, and incomplete resource identifiers. The framework treats these issues as part of the business analytics problem. When activity naming inconsistency is high, model complexity may reflect data semantics rather than real workflow complexity. When resource identifiers are incomplete, handoff analysis becomes less reliable. When duplicate status events remain in the log, bottleneck estimates may be distorted. Therefore, the diagnostic stage is necessary before any automated search is performed. Data-quality treatment is particularly important for object-centric process predictive analytics, where several object types may interact in a single event record (Galanti et al.,2023b).

5.2 Comparative Model Quality

Table 3 compares the three discovery strategies across the four process contexts. The automated MDI strategy produces the highest decision usefulness score in all four contexts. The improvement is not based on maximizing one metric at the expense of all others. Instead, the automated strategy improves the overall balance among fitness, precision, generalization, and simplicity. This balance is especially important for managerial interpretation because process owners require a model that is both empirically credible and readable.

In the procure-to-pay process, automated MDI increases precision from 0.61 under default settings to 0.76 while also improving simplicity from 0.68 to 0.78. This matters because the process contains many variants; a model that allows too much behavior would reduce the value of variant-control recommendations. In order- to-cash, the automated strategy improves decision usefulness from 0.73 to 0.87, primarily because the selected model exposes the main fulfillment-delay loop without overwhelming users with rare paths. In IT incidents, the automated strategy strengthens escalation visibility while maintaining high fitness. In HR onboarding, the selected model improves compliance interpretation by preserving delayed milestone paths while simplifying routine administrative sequences. The governance relevance of such results increases when digital infrastructure creates auditable transaction trails across process participants (Mendling et al.,2018). The auditing relevance of immutable and verifiable operational records is also reflected in recent work on blockchain-enabled internal auditing.

Table 3. Comparative Discovery Outcomes across Process Contexts

Process	Strategy	Fitness	Precision	Generalization	Simplicity	Decision usefulness
Procure-to-pay	Default settings	0.89	0.61	0.74	0.68	0.713
Procure-to-pay	Expert manual	0.92	0.69	0.77	0.72	0.779
Procure-to-pay	Automated MDI	0.95	0.76	0.81	0.78	0.846
Order-to-cash	Default settings	0.88	0.64	0.76	0.71	0.730
Order-to-cash	Expert manual	0.91	0.70	0.79	0.75	0.801
Order-to-cash	Automated MDI	0.94	0.78	0.83	0.80	0.870
IT incident	Default settings	0.91	0.59	0.72	0.66	0.699
IT incident	Expert manual	0.93	0.68	0.75	0.71	0.772
IT incident	Automated MDI	0.96	0.74	0.80	0.77	0.854
HR onboarding	Default settings	0.86	0.66	0.74	0.70	0.711
HR onboarding	Expert manual	0.89	0.72	0.78	0.74	0.773
HR onboarding	Automated MDI	0.93	0.79	0.82	0.81	0.862

The results in Table 3 support the argument that automated discovery should be interpreted as a model- selection procedure rather than as a purely technical convenience. Manual expertise improves over defaults in every case, which confirms the value of domain knowledge. However, automated MDI achieves further gains because it searches a broader configuration space and evaluates candidates consistently against multiple criteria. The difference is particularly visible in precision and decision usefulness, two dimensions that influence whether managers trust the model as a basis for action. Scalability is important because large event logs require discovery and conformance methods that remain computationally feasible (Leemans et al.,2018).

Figure 3 aggregates the results across process contexts. The default strategy provides the shortest runtime, but it delivers lower quality and lower decision usefulness. Expert-manual tuning improves results but depends on analyst experience and remains difficult to reproduce. Automated MDI delivers the highest mean quality score and the highest decision usefulness score. Its runtime is higher, yet the relative increase remains acceptable for strategic process analysis, particularly when parallel evaluation is available and the analysis is not performed as a real-time transaction.

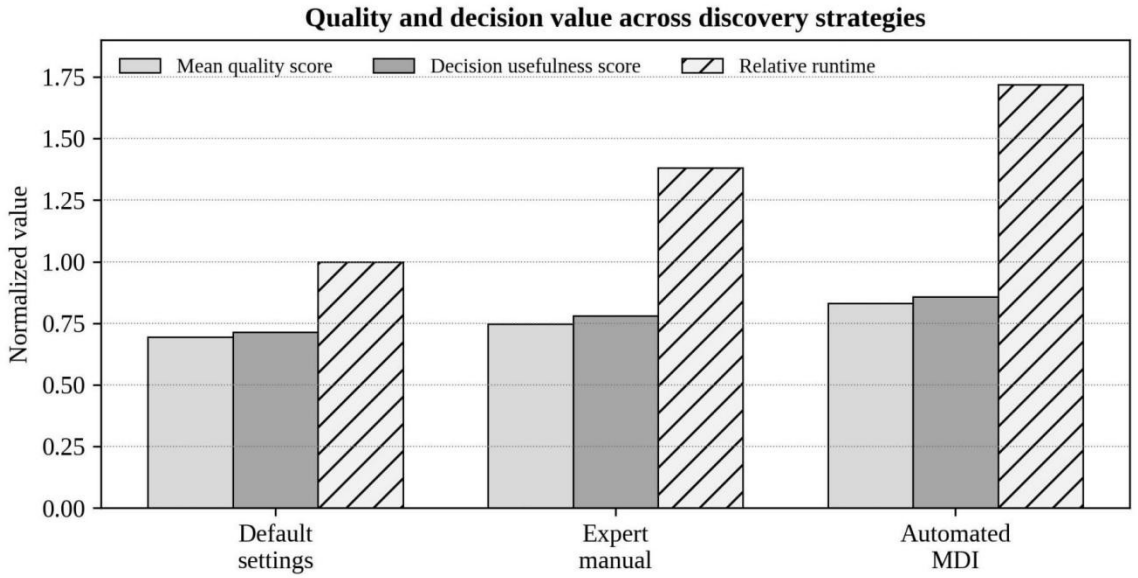


Figure 3. Discovery Strategy Comparison across Quality, Decision Usefulness, and Runtime

Figure 3 also reveals a practical trade-off. Automated MDI improves decision usefulness, but it requires greater computational effort. This trade-off is acceptable when the workflow under analysis affects cost, compliance, customer experience, or automation investment. It may be less appropriate for quick exploratory analysis or small logs with low managerial stakes. The implication is not that every process-mining task requires automated search. Rather, automated search is most valuable when model choice materially affects the decision being made. Incremental monitoring research suggests that computational effort can be reduced when models update efficiently as new cases arrive (Pauwels and Calders,2021).

5.3 Runtime and Scalability

Runtime analysis shows that the cost of automated discovery varies with event count, variant complexity, selected algorithms, and metric computation method. Alignment-based conformance metrics are often more computationally demanding than replay-based alternatives, while highly variable logs require more candidate evaluations before stable trade-offs emerge. In the illustrative analysis, automated MDI requires between 58 and 91 minutes depending on the log and preference profile. Default settings require between 24 and 35 minutes, while expert-manual tuning requires between 38 and 61 minutes, excluding undocumented experimentation time. Deep-learning approaches for predictive process analytics also demonstrate why metric computation can become expensive on large event logs (Pasquadibisceglie et al.,2019).

The runtime comparison should be interpreted carefully. Manual tuning appears faster when only the final execution time is recorded, but it often excludes the analyst's exploratory iterations. In practice, analysts may spend hours comparing visual outputs and adjusting thresholds. Automated search records these iterations as part ofa reproducible procedure. The managerial value of the approach therefore lies not only in better results, but also in transparent documentation of the search process. Collaborative predictive monitoring further shows that hidden coordination costs often matter beyond the final model execution time (Delgado et al.,2025).

Parallel evaluation reduces runtime pressure by distributing candidate configuration assessment across multiple worker threads or compute nodes. The greatest benefits appear when candidate evaluations are independent and log loading is handled efficiently. However, parallelism does not remove the need for intelligent search design. A poorly bounded configuration space remains costly even with parallel

execution. The framework therefore recommends a two-stage procedure: a broad preliminary scan to identify promising configuration regions, followed by a focused search under the selected preference profile. Recent work on event-log extraction with large language models suggests that upstream extraction and preparation may also become part of scalable process analytics (Stein Dani et al.,2025).

5.4 Managerial Interpretation of Discovered Models

The selected models provide distinct managerial insights in the four contexts. For procure-to-pay, the automated MDI model identifies a recurring loop between purchase-order correction and approval re-submission. The loop is not the most frequent path, but it accounts for a substantial share of cycle-time delay in high-value cases. The recommended action is to revise supplier master-data validation and introduce pre- approval checks for high-risk purchase orders. For order-to-cash, the selected model highlights a fulfillment- delay path involving inventory confirmation and shipment rescheduling. The recommended action is to connect promised delivery dates with real-time inventory availability. Rework-loop findings are actionable when linked to redesign heuristics rather than reported only as visual patterns (Reijers and Mansar,2005).

For IT incident management, the selected model reveals that escalations from first-level support to specialist teams are concentrated in a small set of incident categories. The action is to redesign the triage rule and provide targeted knowledge-base prompts for first-level agents. For HR onboarding, the model identifies a delay between contract confirmation and access provisioning. The action is to trigger automated reminders when access provisioning does not occur within the required time window. These examples show how a discovered process model becomes managerial intelligence when it is translated into operational levers. Escalation analysis can also feed predictive monitoring designs that anticipate future case behavior (Márquez- Chamorro et al.,2018).

Table 4. Managerial Decision Interpretation Derived from Automated Workflow Discovery

Process	Model-based finding	Business significance	Recommended managerial action	Primary owner
Procure-to-pay	Correction and approval re-submission loop in high-value cases	Medium frequency; high delay impact	Introduce supplier-data prechecks and approval-rule clarification	Procurement operations manager
Order-to-cash	Inventory confirmation delay before shipment scheduling	High frequency; moderate customer impact	Link order promising to inventory visibility and exception alerts	Revenue operations manager
IT incident management	Escalation concentration in selected incident categories	High frequency; high service impact	Revise triage rules and strengthen first-level knowledge prompts	Service desk manager
HR onboarding	Access provisioning delay after contract confirmation	Lower frequency; high compliance impact	Automate milestone reminders and exception escalation	HR operations manager

Table 4 demonstrates why decision translation is necessary. Technical process metrics indicate that the model is reliable enough to interpret, but they do not specify what managers should do. The decision interpretation step connects structural patterns to accountable owners and action categories. This connection is critical for organizational adoption because process-mining outputs often fail when they remain at the level of diagrams rather than decision responsibilities.

Figure 4 further summarizes the action intensity generated by the selected models. The highest priority values appear in the areas where model evidence and managerial relevance jointly indicate action.

For example, procure-to-pay has a high bottleneck focus index because approval loops produce measurable delay, while IT incident management has a high compliance watch index because escalation paths affect service obligations. HR onboarding has high manager-review and compliance-watch scores because the process involves cross-functional accountability and employee readiness.

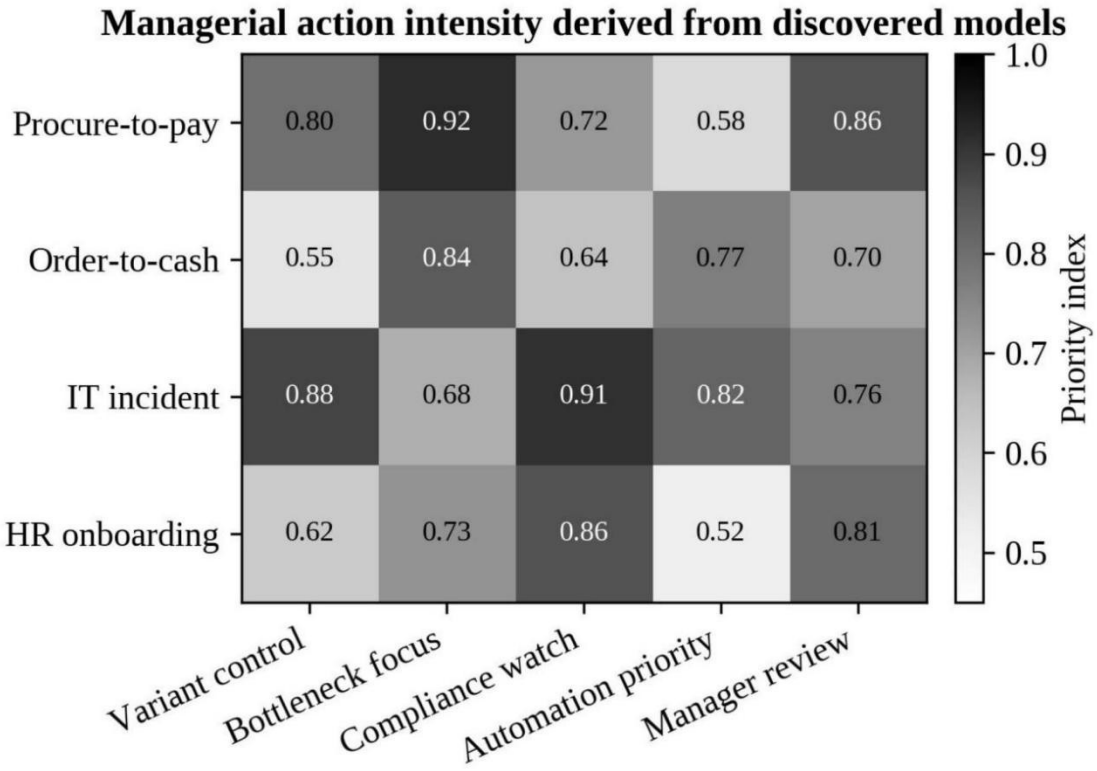


Figure 4. Managerial Action Priorities across Processes and Decision Dimensions

The matrix in Figure 4 should not be read as an automated decision rule. It is a structured decision aid. The values prioritize where management attention should be directed, but human validation remains necessary. A high automation priority may still be rejected if the process requires legal judgment or human empathy. A high compliance-watch value may require policy revision rather than system control. The value of the framework is that it creates a transparent starting point for such discussions.

6. Discussion

6.1 From Automated Discovery to Decision Intelligenc

The findings support the central argument of this article: the value of automated workflow discovery lies not only in producing better process models, but in creating a repeatable pathway from operational evidence to managerial decisions. Default discovery settings provide a quick view, but they may obscure important trade- offs. Expert-manual tuning improves results, yet it depends heavily on individual judgment and often lacks a complete audit trail. Automated MDI strengthens the analytical procedure by searching more systematically, recording model trade-offs, and linking selection to explicit preferences. This finding is consistent with organizational research identifying both opportunities and challenges when process mining is institutionalized (Martin et al.,2021).

This does not imply that technical metrics should be abandoned. Fitness, precision, generalization, and simplicity remain essential. However, they should be interpreted as inputs into a larger decision

process. A model with high fitness but low precision may overgeneralize the process. A model with high precision but excessive complexity may be difficult to use in a management meeting. A simple model may be useful for communication but inadequate for compliance monitoring. Decision intelligence emerges when these trade- offs are made visible and tied to action. Predictive process monitoring research reaches a similar conclusion by separating algorithmic performance from deployment and interpretation challenges (Ceravolo et al.,2024).

The proposed framework also shifts attention from diagram production to decision accountability. Many process-mining projects end with a set of visual models and performance dashboards. The framework extends this by asking who owns the issue, what decision is required, which evidence supports the decision, and how the decision will be monitored after intervention. This accountability orientation is particularly important for enterprise analytics programs that seek measurable operational impact rather than isolated analytical outputs. Accountability also matters when proactive conformance checking turns process evidence into early warnings about deviations (Grohs et al.,2025).

6.2 Managerial Implications

For managers, the first implication is that event logs should be treated as strategic analytical assets. Organizations invest heavily in enterprise systems, but the behavioral evidence contained in those systems is often underused. Process owners should define which workflows require recurring event-log analysis and which decisions will be supported by discovery results. Candidate areas include procurement exceptions, order fulfillment delays, service escalations, claims processing, customer onboarding, invoice disputes, and compliance reviews. This implication is consistent with management analytics research that views analytical capabilities as part of modern decision-making practice. The managerial value of automated discovery also depends on how organizations understand the evolution, models, and applications of artificial intelligence. Analytics can inform routine managerial action when organisational processes incorporate the evidence it produces (Davenport & Harris, 2007).

The second implication concerns preference transparency. Managers should not delegate all model-selection decisions to analysts without stating the business objective. An analyst may optimize for technical accuracy, while the decision context may require interpretability or exception sensitivity. The framework provides a mechanism for making such preferences explicit. Before discovery begins, stakeholders should identify whether the analysis is designed for efficiency improvement, compliance monitoring, automation selection, customer-experience redesign, or resource planning. Transparency becomes more important as object-centric process mining introduces richer representations and a larger design space (Berti, Montali, and van der Aalst,2023). The same transparency principle becomes more salient as AI-augmented BPM systems make process execution more proactive and context-sensitive (Dumas et al.,2022).

The third implication concerns organizational learning. Automated discovery produces a record of candidate models and trade-offs. This record is valuable beyond a single project because it allows managers and analysts to learn which configurations tend to work for particular process types. Over time, organizations may build libraries of preference profiles, preprocessing rules, and validated discovery settings. Such libraries reduce project start-up time and improve consistency across business units. Organizational learning from process data is also relevant to logistics and supply-chain analytics, where large operational datasets support performance improvement (Wamba et al.,2018).

The fourth implication concerns governance. Event-log analysis may reveal employee behavior, control weaknesses, or sensitive customer information. Managers should establish rules for data access, anonymization, interpretation, and communication. Process-mining outputs should not be used to blame individuals without contextual validation. Instead, they should be used to improve systems, clarify policies, and reduce avoidable friction. Responsible interpretation increases the legitimacy of analytics in the organization. Privacy-preserving process mining is therefore essential when event logs include personal or

sensitive operational attributes (Mannhardt et al.,2019). When event logs are collected from IoT-enabled operations, privacy and cybersecurity controls become part of the governance architecture.

6.3 Implications for Business Data Analytics Research

For research, the framework contributes by positioning automated workflow discovery within business data analytics rather than only within process-mining software engineering. This positioning opens several research opportunities. First, researchers may examine how managerial preference profiles affect model selection and subsequent organizational action. Second, studies may compare the effectiveness of different decision translation templates across industries. Third, researchers may investigate how model interpretability influences stakeholder trust and adoption. Fourth, future work may connect automated discovery with causal inference to evaluate whether workflow interventions actually improve performance. This positioning also draws on the view that process mining is a data-science discipline with managerial and operational consequences (van der Aalst,2016). Bibliometric work on management analytics further supports the need to connect analytical techniques with managerial domains.

The framework also encourages more careful treatment of model usefulness. Process-mining benchmarks often compare algorithms through technical metrics. Such benchmarks remain necessary, but business analytics research should also examine whether selected models support decisions. A model may be technically superior in a benchmark but less valuable for a manager facing time pressure, governance constraints, or communication needs. Decision usefulness should therefore be studied as a construct with measurable dimensions such as interpretability, diagnostic relevance, actionability, and governance fit. Benchmarking also needs to consider predictive monitoring designs that cluster traces and evaluate future behavior (Di Francescomarino et al.,2019).

Finally, the framework suggests that automated discovery should be integrated with broader enterprise analytics architectures. Event-log analytics is connected to data warehousing, master-data governance, process automation, risk management, and performance management. A model selected through automated workflow discovery may feed a dashboard, trigger an alert, support a business case for automation, or guide process redesign workshops. Research that treats process mining as part of a larger decision system is likely to generate stronger managerial insights than research that examines discovery algorithms in isolation. This integration requires analytics capabilities that combine data, technology, managerial skills, and organizational routines (Gupta and George,2016). It also reflects Industry 4.0 research in which data-intensive operational systems become the basis for intelligent organizational coordination.

7. Limitations and Future Research

This study has several limitations. First, the numerical analysis is illustrative. It is designed to demonstrate the logic of the proposed framework rather than to estimate universal performance gains. Future research should apply the framework to real organizational event logs across industries and compare decision outcomes before and after implementation. Second, the decision usefulness score is structured but partly judgment-based. Future studies should validate measurement scales for interpretability, diagnostic relevance, actionability, and governance sensitivity. Third, the framework assumes that preference profiles are explicitly defined. In practice, stakeholders may disagree about priorities, and those disagreements may change over time. A limitation of illustrative analysis is that sector-specific validation may change the interpretation of discovered models, as shown in healthcare process-mining studies (Rojas et al.,2016). The same cautionapplies when workflow analysis is used to justify robotic-process-automation implementation (Aguirre and Rodriguez,2017). Additional empirical work should test process-mining criteria for assessing whether tasks are suitable for automation (Shafik Salah Elsayed and Kassem,2022). Such testing should also consider whether analytics capability and strategic alignment jointly shape performance outcomes (Akter et al.,2016).

Fourth, the framework focuses on discovery and model selection rather than full conformance and prediction. Future work should integrate conformance checking, predictive process monitoring, and prescriptive analytics into the decision intelligence layer. For example, a selected model might identify a bottleneck, a predictive model might estimate which cases are likely to experience that bottleneck, and a prescriptive model might recommend resource reallocation. Fifth, computational efficiency remains a practical concern. Automated search may become expensive for very large event logs, high-dimensional configuration spaces, or complex conformance metrics. Adaptive search strategies, sampling methods, and cloud-based parallelization deserve additional investigation. Future work on object-centric predictive monitoring may improve the treatment of interacting cases and complex enterprise processes (Smit et al.,2024). Practical extensions should also evaluate how predictive monitoring is embedded in operational platforms such as Apromore (Verenich et al.,2018). Future work may also examine object-centric Petri nets as a formal basis for complex workflow representation (van der Aalst and Berti,2020). Comparative work on object-centric and classical event logs suggests that future studies should test whether the same decision-usefulness patterns hold across data representations (Fioretto and Masciari,2025).

Sixth, managerial interpretation may be influenced by organizational politics. A process model that reveals inefficiency in one department may be contested by stakeholders who fear blame or resource loss. The framework recommends governance and domain validation, but future research should examine the organizational conditions under which event-log evidence leads to constructive process improvement rather than defensive behavior. Finally, ethical considerations need further attention. Event logs may contain information about individual employees or customers. Responsible analytics should protect privacy, avoid simplistic performance judgments, and focus on systemic improvement. These political and organizational effects may become stronger when process evidence feeds automation programs and role redesign (Syed et al.,2020). A recent review of data-driven robotic process automation also points to the need for better links between process evidence and automation decisions (El-Gharib and Amyot,2022). Automation politics should also be evaluated against the broader evolution of robotic process automation in information systems (van der Aalst et al.,2018). These issues should be studied as socio-technical change rather than as narrow software deployment (Doolin et al.,2025).

8. Conclusion

This article developed a business data analytics framework for transforming event logs into managerial decision intelligence through automated workflow discovery. The framework addresses a common problem in enterprise analytics: organizations possess detailed operational event data, but the translation from logs to trusted managerial action remains inconsistent. By integrating event-log preparation, automated configuration search, multi-criteria model evaluation, preference-aware selection, and decision translation, the framework creates a more transparent and reproducible pathway from process evidence to management decisions.

The illustrative analysis across procure-to-pay, order-to-cash, IT incident management, and HR onboarding processes shows that automated managerial decision intelligence improves the balance of model quality and decision usefulness when compared with default settings and expert-manual tuning. The results also show that the selected model should be interpreted within a specific decision context. A model appropriate for compliance monitoring may not be the best model for automation planning. A model designed for executive communication may not satisfy the needs of forensic investigation. The framework makes these differences explicit by treating model selection as a multi-criteria decision task.

The broader contribution of the study is to reposition workflow discovery as a managerial intelligence process. Event logs provide evidence, automated discovery provides candidate explanations, multi-criteria evaluation exposes trade-offs, and decision translation connects analytical outputs to accountable action. Organizations that adopt this approach are better positioned to use process data for bottleneck reduction,

compliance monitoring, automation prioritization, service improvement, and workflow redesign. Future research should validate the framework with real enterprise deployments and develop richer methods for measuring the business value of discovered process models.

Acknowledgement

The authors thank the researchers whose published work provided the theoretical and methodological foundations for this study.

Author Contributions

Conceptualization, M.Q.Z. and J.P.M.; Methodology, M.Q.Z. and J.P.M.; Formal Analysis, M.Q.Z. and J.P.M.; Investigation, M.Q.Z. and F.S.; Data Curation, M.Q.Z. and F.S.; Software, M.Q.Z. and F.S.; Validation, M.Q.Z. and J.P.M. and F.S.; Resources, F.S.; Visualization, M.Q.Z. and F.S.; Writing – Original Draft, M.Q.Z.; Writing – Review & Editing, M.Q.Z. and J.P.M. and F.S.; Supervision, J.P.M.; Project Administration, M.Q.Z. and J.P.M. We have read and approved the final manuscript and take responsibility for the work.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Ethical Approval and Consent

Not applicable. This study used an illustrative numerical analysis and did not involve human participants, identifiable personal data, or animal subjects.

Data Availability Statement

The illustrative event-log characteristics and aggregate analytical results are reported in Tables 2–4. The analytical design and evaluation procedures are described in the manuscript.

Declaration of Generative AI Use

Generative AI was used to assist with manuscript editing, formatting, and the preparation of disclosure statements. The authors take full responsibility for the accuracy, integrity, and originality of the manuscript.

References

Adriansyah, A., van Dongen, B. F., & van der Aalst, W. M. P. (2011). Conformance checking using cost-based fitness analysis. In Proceedings of the 15th IEEE International Enterprise Distributed Object Computing Conference (pp. 55–64). IEEE. https://doi.org/10.1109/EDOC.2011.12

Agostinelli, S., Marrella, A., & Mecella, M. (2021). Exploring the challenge of automated segmentation in robotic process automation. In Research Challenges in Information Science (pp. 38–54). Springer. https://doi.org/10.1007/978-3-030-75018-3_3

Aguirre, S., Parra, C., & Sepulveda, M. (2017). Methodological proposal for process mining projects. International Journal of Business Process Integration and Management, 8(2), 102–113. https://doi.org/10.1504/IJBPIM.2017.083793

Aguirre, S., & Rodriguez, A. (2017). Automation of a business process using robotic process automation (RPA): A case study. In Applied Computer Sciences in Engineering (pp. 65–71). Springer. https://doi.org/10.1007/978-3-319-66963-2_7

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? International Journal of Production Economics, 182, 113–131. https://doi.org/10.1016/j.ijpe.2016.08.018

Andrews, R., van Dun, C. G. J., Wynn, M. T., Kratsch, W., Röglinger, M. K. E., & ter Hofstede, A. H. M. (2020). Quality-informed semi-automated event log generation for process mining. Decision Support Systems, 132, 113265. https://doi.org/10.1016/j.dss.2020.113265

Berti, A., Montali, M., & van der Aalst, W. M. P. (2023). Advancements and challenges in object-centric process mining: A systematic literature review. arXiv. https://doi.org/10.48550/arXiv.2311.08795

Berti, A., van Zelst, S. J., & Schuster, D. (2023). PM4Py: A process mining library for Python. Software Impacts, 17, 100556. https://doi.org/10.1016/j.simpa.2023.100556

Bertsimas, D., & Kallus, N. (2020). From predictive to prescriptive analytics. Management Science, 66(3), 1025–1044. https://doi.org/10.1287/mnsc.2018.3253

Buijs, J. C. A. M., van Dongen, B. F., & van der Aalst, W. M. P. (2014). Quality dimensions in process discovery: The importance of fitness, precision, generalization and simplicity. International Journal of Cooperative Information Systems, 23(1), 1440001. https://doi.org/10.1142/S0218843014400012

Ceravolo, P., Comuzzi, M., De Weerdt, J., Di Francescomarino, C., & Maggi, F. M. (2024). Predictive process monitoring: Concepts, challenges, and future research directions. Process Science, 1(1), 2. https://doi.org/10.1007/s44311-024-00002-4

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS Quarterly, 36(4), 1165–1188. https://doi.org/10.2307/41703503

Davenport, T. H., & Harris, J. G. (2007). Competing on analytics: The new science of winning. Harvard Business School Press.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), 182–197. https://doi.org/10.1109/4235.996017

Delgado, A., Calegari, D., Espino, C., & Ribero, N. (2025). Predictive process monitoring for collaborative business processes: Concepts and application. Discover Analytics, 3(1), Article 5. https://doi.org/10.1007/s44257-025-00031-8

Di Francescomarino, C., Dumas, M., Maggi, F. M., & Teinemaa, I. (2019). Clustering-based predictive process monitoring. IEEE Transactions on Services Computing, 12(6), 896–909. https://doi.org/10.1109/TSC.2016.2645153

Doolin, B., Techatassanasoontorn, A. A., Waizenegger, L., & Wallace-Carter, E. (2025). Theorising robotic process automation as socio-technical change: A process study. Australasian Journal of Information Systems, 29. https://doi.org/10.3127/ajis.v29.5451

Dumas, M., Fournier, F., Limonad, L., Marrella, A., Montali, M., Rehse, J.-R., Accorsi, R., Calvanese, D., De Giacomo, G., Fahland, D., Gal, A., La Rosa, M., Völzer, H., & Weber, I. (2022). AI-augmented business process management systems: A research manifesto. arXiv. https://doi.org/10.48550/arXiv.2201.12855

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2018). Fundamentals of Business Process Management. Springer. https://doi.org/10.1007/978-3-662-56509-4

Eggers, J., Hein, A., Böhm, M., & Krcmar, H. (2021). No longer out of sight, no longer out of mind? How organizations engage with process mining-induced transparency to achieve increased process awareness. Business & Information Systems Engineering, 63(5), 491–510. https://doi.org/10.1007/s12599-021-00715-x

El-Gharib, N. M., & Amyot, D. (2022). Robotic process automation using process mining: A systematic literature review. arXiv. https://doi.org/10.48550/arXiv.2204.00751

Engel, R., Krathu, W., Zapletal, M., Pichler, C., Bose, R. P. J. C., van der Aalst, W. M. P., Werthner, H., & Huemer, C. (2016). Analyzing inter-organizational business processes: Process mining and business performance analysis using electronic data interchange messages. Information Systems and e-Business Management, 14(3), 577–612. https://doi.org/10.1007/s10257-015-0295-2

Erdogan, T. G., & Tarhan, A. (2018). Systematic mapping of process mining studies in healthcare. IEEE Access, 6, 24543–24567. https://doi.org/10.1109/ACCESS.2018.2831244

Fahland, D., & van der Aalst, W. M. P. (2015). Model repair: Aligning process models to reality. Information Systems, 47, 220–243. https://doi.org/10.1016/j.is.2013.12.007

Fani Sani, M., Vazifehdoostirani, M., Park, G., Pegoraro, M., van Zelst, S. J., & van der Aalst, W. M. P. (2023). Performance-preserving event log sampling for predictive monitoring. Journal of Intelligent Information Systems, 61(1), 53–82. https://doi.org/10.1007/s10844-022-00775-9

Fioretto, S., & Masciari, E. (2025). A comparative analysis of predictive process monitoring: Object-centric versus classical event logs. Knowledge and Information Systems, 67, 7355–7398. https://doi.org/10.1007/s10115-025-02461-y

Galanti, R., de Leoni, M., Monaro, M., Navarin, N., Marazzi, A., Di Stasi, B., & Maldera, S. (2023a). An explainable decision support system for predictive process analytics. Engineering Applications of Artificial Intelligence, 120, 105904. https://doi.org/10.1016/j.engappai.2023.105904

Galanti, R., de Leoni, M., Navarin, N., & Marazzi, A. (2023b). Object-centric process predictive analytics. Expert Systems with Applications, 213, 119173. https://doi.org/10.1016/j.eswa.2022.119173

Gherissi, W., El Haddad, J., & Grigori, D. (2023). Object-centric predictive process monitoring. In Service-Oriented Computing - ICSOC 2022 Workshops (pp. 27–39). Springer. https://doi.org/10.1007/978-3-031-26507-5_3

Grohs, M., Pfeiffer, P., & Rehse, J.-R. (2025). Proactive conformance checking: An approach for predicting deviations in business processes. Information Systems, 127, 102461. https://doi.org/10.1016/j.is.2024.102461

Grover, V., Chiang, R. H. L., Liang, T. P., & Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. Journal of Management Information Systems, 35(2), 388–423. https://doi.org/10.1080/07421222.2018.1451951

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. Information & Management, 53(8), 1049–1064. https://doi.org/10.1016/j.im.2016.07.004

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS Quarterly, 28(1), 75–106. https://doi.org/10.2307/25148625

Jans, M., Alles, M. G., & Vasarhelyi, M. A. (2014). A field study on the use of process mining of event logs as an analytical procedure in auditing. The Accounting Review, 89(5), 1751–1773. https://doi.org/10.2308/accr-50807

Khatri, V., & Brown, C. V. (2010). Designing data governance. Communications of the ACM, 53(1), 148–152. https://doi.org/10.1145/1629175.1629210

Leemans, S. J. J., Fahland, D., & van der Aalst, W. M. P. (2013). Discovering block-structured process models from event logs: A constructive approach. In Application and Theory of Petri Nets and Concurrency (pp. 311–329). Springer. https://doi.org/10.1007/978-3-642-38697-8_17

Leemans, S. J. J., Fahland, D., & van der Aalst, W. M. P. (2018). Scalable process discovery and conformance checking. Software & Systems Modeling, 17(2), 599–631. https://doi.org/10.1007/s10270-016-0545-x

Lorenz, R., Senoner, J., Sihn, W., & Netland, T. (2021). Using process mining to improve productivity in make-to-stock manufacturing. International Journal of Production Research, 59(16), 4869–4880. https://doi.org/10.1080/00207543.2021.1906460

Mannhardt, F., Koschmider, A., Baracaldo, N., Weidlich, M., & Michael, J. (2019). Privacy-preserving process mining: Differential privacy for event logs. Business & Information Systems Engineering, 61(5), 595–614. https://doi.org/10.1007/s12599-019-00613-3

Mardani, A., Jusoh, A., Md Nor, K., Khalifah, Z., Zakwan, N., & Valipour, A. (2015). Multiple criteria decision-making techniques and their applications: A review of the literature from 2000 to 2014. Economic Research-Ekonomska Istrazivanja, 28(1), 516–571. https://doi.org/10.1080/1331677X.2015.1075139

Marin-Castro, H. M., & Tello-Leal, E. (2021). Event log preprocessing for process mining: A review. Applied Sciences, 11(22), 10556. https://doi.org/10.3390/app112210556

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

Martin, N., Fischer, D. A., Kerpedzhiev, G. D., Goel, K., Leemans, S. J. J., Röglinger, M., van der Aalst, W. M. P., Dumas, M., La Rosa, M., & Wynn, M. T. (2021). Opportunities and challenges for process mining in organizations: Results of a Delphi study. Business & Information Systems Engineering, 63(5), 511–527. https://doi.org/10.1007/s12599-021-00720-0

Mendling, J., Reijers, H. A., & van der Aalst, W. M. P. (2010). Seven process modeling guidelines (7PMG). Information and Software Technology, 52(2), 127–136. https://doi.org/10.1016/j.infsof.2009.08.004

Mendling, J., Weber, I., van der Aalst, W. M. P., vom Brocke, J., Cabanillas, C., Daniel, F., Debois, S., Di Ciccio, C., Dumas, M., Dustdar, S., Gal, A., Garcia-Banuelos, L., Governatori, G., Hull, R., La Rosa, M., Leopold, H., Leymann, F., Recker, J., Reichert, M., Reijers, H. A., Rinderle-Ma, S., Solti, A., Rosemann, M., Schulte, S., Singh, M. P., Slaats, T., Staples, M., Weber, B., Weidlich, M., Weske, M., Xu, X., & Zhu, L. (2018). Blockchains for business process management: Challenges and opportunities. ACM Transactions on Management Information Systems, 9(1), Article 4. https://doi.org/10.1145/3183367

Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. Information & Management, 57(2), 103169. https://doi.org/10.1016/j.im.2019.05.004

Park, G., & Song, M. (2023). Optimizing resource allocation based on predictive process monitoring. IEEE Access, 11, 38309–38323. https://doi.org/10.1109/ACCESS.2023.3267538

Pasquadibisceglie, V., Appice, A., Castellano, G., & Malerba, D. (2019). Using convolutional neural networks for predictive process analytics. In Proceedings of the International Conference on Process Mining (pp. 129–136). https://doi.org/10.1109/ICPM.2019.00028

Pauwels, S., & Calders, T. (2021). Incremental predictive process monitoring: The next activity case. In Business Process Management (pp. 123–140). Springer. https://doi.org/10.1007/978-3-030-85469-0_10

Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. Communications of the ACM, 45(4), 211–218. https://doi.org/10.1145/505248.506010

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. Big Data, 1(1), 51–59. https://doi.org/10.1089/big.2013.1508

Rebuge, A., & Ferreira, D. R. (2012). Business process analysis in healthcare environments: A methodology based on process mining. Information Systems, 37(2), 99–116. https://doi.org/10.1016/j.is.2011.01.003

Redman, T. C. (1998). The impact of poor data quality on the typical enterprise. Communications of the ACM, 41(2), 79–82. https://doi.org/10.1145/269012.269025

Reijers, H. A., & Mansar, S. L. (2005). Best practices in business process redesign: An overview and qualitative evaluation of successful redesign heuristics. Omega, 33(4), 283–306. https://doi.org/10.1016/j.omega.2004.04.012

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778

Rizzi, W., Di Francescomarino, C., Ghidini, C., & Maggi, F. M. (2025). Nirdizati: An advanced predictive process monitoring toolkit. Journal of Intelligent Information Systems, 63, 259–291. https://doi.org/10.1007/s10844-024-00890-9

Rojas, E., Munoz-Gama, J., Sepulveda, M., & Capurro, D. (2016). Process mining in healthcare: A literature review. Journal of Biomedical Informatics, 61, 224–236. https://doi.org/10.1016/j.jbi.2016.04.007

Saaty, T. L. (1987). The analytic hierarchy process: What it is and how it is used. Mathematical Modelling, 9(3–5), 161–176. https://doi.org/10.1016/0270-0255(87)90473-8

Shafik Salah Elsayed, N., & Kassem, G. (2022). Assessing process suitability for robotic process automation: A process mining approach. Wirtschaftsinformatik 2022 Proceedings, Student Track, Paper 18. https://aisel.aisnet.org/wi2022/student_track/student_track/18/

Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. MIS Quarterly, 35(3), 553–572. https://doi.org/10.2307/23042796

Smit, T. K., Reijers, H. A., & Lu, X. (2024). HOEG: A new approach for object-centric predictive process monitoring. In Advanced Information Systems Engineering (pp. 231–247). Springer. https://doi.org/10.1007/978-3-031-61057-8_14

Stein Dani, V., Dees, M., Leopold, H., Busch, K., Beerepoot, I., van der Werf, J. M., & Reijers, H. A. (2025). Event log extraction for process mining using large language models. In Cooperative Information Systems (pp. 56–72). Springer. https://doi.org/10.1007/978-3-031-81375-7_4

Suriadi, S., Andrews, R., ter Hofstede, A. H. M., & Wynn, M. T. (2017). Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs. Information Systems, 64, 132–150. https://doi.org/10.1016/j.is.2016.07.011

Syed, R., Suriadi, S., Adams, M., Bandara, W., Leemans, S. J. J., Ouyang, C., ter Hofstede, A. H. M., van de Weerd, I., Wynn, M. T., & Reijers, H. A. (2020). Robotic process automation: Contemporary themes and challenges. Computers in Industry, 115, 103162. https://doi.org/10.1016/j.compind.2019.103162

Tax, N., Verenich, I., La Rosa, M., & Dumas, M. (2017). Predictive business process monitoring with LSTM neural networks. In Advanced Information Systems Engineering (pp. 477–492). Springer. https://doi.org/10.1007/978-3-319-59536-8_30

Teinemaa, I., Dumas, M., La Rosa, M., & Maggi, F. M. (2019). Outcome-oriented predictive process monitoring: Review and benchmark. ACM Transactions on Knowledge Discovery from Data, 13(2), Article 17. https://doi.org/10.1145/3301300

van der Aalst, W. M. P. (2012). Process mining: Overview and opportunities. ACM Transactions on Management Information Systems, 3(2), Article 7. https://doi.org/10.1145/2229156.2229157

van der Aalst, W. M. P. (2016). Process Mining: Data Science in Action. Springer. https://doi.org/10.1007/978-3-662-49851-4

van der Aalst, W. M. P. (2023). Object-centric process mining: Unraveling the fabric of real processes. Mathematics, 11(12), 2691. https://doi.org/10.3390/math11122691

van der Aalst, W. M. P., & Berti, A. (2020). Discovering object-centric Petri nets. Fundamenta Informaticae, 175(1–4), 1–40. https://doi.org/10.3233/FI-2020-1946

van der Aalst, W. M. P., Bichler, M., & Heinzl, A. (2018). Robotic process automation. Business & Information Systems Engineering, 60(4), 269–272. https://doi.org/10.1007/s12599-018-0542-4

van der Aalst, W. M. P., Weijters, A. J. M. M., & Maruster, L. (2004). Workflow mining: Discovering process models from event logs. IEEE Transactions on Knowledge and Data Engineering, 16(9), 1128–1142. https://doi.org/10.1109/TKDE.2004.47

van Dongen, B. F., de Medeiros, A. K. A., Verbeek, H. M. W., Weijters, A. J. M. M., & van der Aalst, W. M. P. (2005). The ProM framework: A new era in process mining tool support. In Application and Theory of Petri Nets 2005 (pp. 444–454). Springer. https://doi.org/10.1007/11494744_25

Verenich, I., Mõškovski, S., Raboczi, S., Dumas, M., La Rosa, M., & Maggi, F. M. (2018). Predictive process monitoring in Apromore. In Information Systems in the Big Data Era (pp. 244–253). Springer. https://doi.org/10.1007/978-3-319-92901-9_21

vom Brocke, J., Jans, M., Mendling, J., & Reijers, H. A. (2021). A five-level framework for research on process mining. Business & Information Systems Engineering, 63, 483–490. https://doi.org/10.1007/s12599-021-00718-8

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. Journal of Business Research, 70, 356–365. https://doi.org/10.1016/j.jbusres.2016.08.009

Wamba, S. F., Gunasekaran, A., Papadopoulos, T., & Ngai, E. W. T. (2018). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. International Journal of Logistics Management, 29(2), 478–484. https://doi.org/10.1108/IJLM-02-2018-0026

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, 12(4), 5–33. https://doi.org/10.1080/07421222.1996.11518099