Advanced International Journal of Multidisciplinary Research

E-ISSN: 2584-0487
editor@aijmr.com

Volume 2, Issue 6, November-December 2024

CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

Mapping the Latent Risk Layer in Enterprise Platforms: A Practical Model for Workforce Data Integrity, Access Behavior, and Cyber Threat Detection

Manoj Parasa

Independent Researcher
manoj.parasa1993@gmail.com

Abstract:
Risk in enterprise platforms rarely appears as a single obvious security event. In many cases, it develops quietly across ordinary workforce transactions, such as an incorrect manager assignment, a delayed employee status update, an unnecessary permission retention, an unusual access pattern, or a repeated integration exception. When these signals are reviewed separately by HR operations, security teams, and system administrators, the broader risk pattern is often missed. This study introduces a practical model for mapping the latent risk layer in enterprise platforms by connecting workforce data integrity, access behavior, and cyber threat indicators into a unified risk-scoring structure. SAP SuccessFactors is used as the primary enterprise context because its Employee Central records, MDF objects, role-based permissions, workflow approvals, audit trails, and integration events provide measurable signals that can be evaluated without relying on speculative or unrealistic system capabilities. The proposed approach develops risk features from employee master data changes, permission activity, effective-dated records, workflow exceptions, integration failures, sensitive data access, and security alerts. These features are evaluated through a comparative modeling design that includes rule-based monitoring, Logistic Regression, Isolation Forest, and tree-based classification methods. Model performance is assessed using accuracy, precision, recall, F1-score, false positive rate, detection latency, and remediation effort. The study is designed to show how combined workforce and access signals can identify high-risk events earlier than traditional isolated controls. Its main contribution is a realistic and implementation-oriented framework that helps organizations convert fragmented HR, access, and security events into measurable enterprise risk intelligence. The findings provide practical value for SAP SuccessFactors governance, identity management, audit readiness, cybersecurity monitoring, and cross-platform workforce data control.

Keywords: Latent risk layer; workforce data integrity; enterprise platforms; SAP SuccessFactors; access behavior analytics; role-based permissions; cyber threat detection; employee master data; anomaly detection; risk scoring.

1. Introduction

Risk inside enterprise platforms is often treated as something that appears only after a major control failure, security alert, audit exception, or confirmed incident. In practice, many enterprise risks begin much earlier and in much smaller forms. A delayed employee status update, an incorrect manager assignment,

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

a permission group that remains active after a role change, an unusual access pattern, or a repeated integration failure may not appear critical when reviewed alone. However, when these events occur together, they can reveal a deeper operational exposure that traditional monitoring methods often fail to detect. This hidden pattern is the central concern of this study.

Modern workforce platforms hold more than employee records. They influence access decisions, workflow approvals, organizational reporting, compliance reviews, downstream integrations, and business operations across multiple departments. In systems such as SAP SuccessFactors, a single employee data change can affect role-based permissions, workflow routing, position visibility, reporting relationships, compensation eligibility, onboarding activities, and third-party integrations. Because of this dependency, workforce data integrity is no longer only an HR administration concern. It has become part of enterprise risk management, cybersecurity readiness, and digital governance.

Most organizations still monitor workforce data quality, access control, and cyber threat signals through separate operational channels. HR teams focus on data corrections, workflow errors, and employee record completeness. Security teams monitor suspicious logins, privileged access, and unusual system activity. Integration teams investigate failed jobs, delayed payloads, or mismatched records between systems. Audit teams review permission conflicts, policy exceptions, and evidence gaps. Each team may resolve its own issue, but the broader risk pattern across these areas can remain invisible. This fragmented model creates a latent risk layer that sits between normal operations and visible incidents.

The term “latent risk layer” in this paper refers to a measurable but often unnoticed risk pattern formed by the interaction of workforce data inconsistencies, access behavior deviations, integration exceptions, and cyber threat indicators. It is not a separate technical system or a speculative concept. It is a practical analytical layer built from signals that already exist in enterprise platforms. For example, an employee transferred to a new department may retain sensitive permissions for several days, access restricted records after the transfer, and trigger a failed downstream identity update. Individually, these events may appear as routine administrative issues. Together, they indicate a risk condition that deserves faster review.

This paper proposes a practical model for mapping this latent risk layer across enterprise platforms, using SAP SuccessFactors as the primary implementation context. The model is designed around realistic system capabilities, including Employee Central data, MDF objects, effective-dated records, role-based permissions, workflow approvals, audit logs, integration events, and identity or security monitoring signals. The objective is not to introduce an unrealistic autonomous security system, but to show how existing enterprise data points can be organized, scored, compared, and explained in a way that improves risk detection and governance decision-making.

The proposed approach combines three risk dimensions: workforce data integrity, access behavior, and cyber threat detection. Workforce data integrity evaluates the quality and consistency of employee records, including status changes, job data, manager assignments, event reasons, effective dates, and synchronization delays. Access behavior evaluates how users interact with sensitive records, permissions, pages, reports, and administrative functions. Cyber threat detection evaluates suspicious signals such as unusual login activity, privileged access behavior, high-volume exports, policy exceptions, and security alerts. By joining these dimensions, the model can identify risk situations that would be difficult to detect through isolated rule-based controls.

A key contribution of this study is its measurable design. The framework supports practical performance evaluation through accuracy, precision, recall, F1-score, false positive rate, detection latency, manual review effort, and remediation cycle time. It also allows comparison between traditional rule-based monitoring and machine learning-assisted risk detection methods such as Logistic Regression, Isolation Forest, and tree-based classification models. This makes the paper suitable for both academic evaluation

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

and enterprise implementation discussion. The model can be tested using structured event data, realistic feature engineering, and controlled experimental scenarios that reflect common workforce platform risks. The value of this research extends beyond a single HR technology platform. Although SAP SuccessFactors provides the primary context, the same risk logic can apply to Workday, Oracle HCM, UKG, ServiceNow HRSD, identity platforms, finance systems, procurement systems, and other enterprise applications where employee data, access privileges, and security events are connected. Organizations need a practical way to understand how small operational inconsistencies can become early indicators of larger enterprise exposure. This paper addresses that need by presenting a structured, explainable, and implementation-oriented model for identifying risk before it becomes a visible incident.

The remainder of the paper develops this argument in a structured manner. The next section defines the research problem and practical motivation in detail. The literature review then positions the study within existing work on workforce data governance, access control, anomaly detection, and cyber-risk monitoring. The proposed framework section introduces the latent risk layer model, followed by dataset design and feature engineering. The methodology section explains the risk-scoring and model training approach. The experimental section defines evaluation metrics, while the results section presents comparative performance analysis, operational impact, and risk score interpretation. The final section discusses enterprise case studies, practical implications, limitations, and future research directions.

2. Research Problem and Practical Motivation

2.1 Fragmented Monitoring Across Workforce, Access, and Security Functions

Enterprise risk rarely develops within the boundaries of one department. A workforce data issue may begin in HR operations, move into identity provisioning, affect role-based permissions, and later appear as a security concern. Despite this interconnected reality, many organizations continue to manage workforce data quality, access behavior, and cyber-risk through separate monitoring practices. HR teams usually focus on record completeness, workflow routing, employee status accuracy, and organizational assignments. Security teams concentrate on login anomalies, privileged access, suspicious activity, and alert response. Integration teams investigate interface failures, delayed payloads, rejected records, and synchronization gaps. Audit and compliance teams review policy exceptions, approval evidence, and control documentation.

This separation creates a practical blind spot. Each function may correctly identify and resolve its own operational issue, but no single view explains how these issues interact. A manager change in Employee Central may look like a normal HR update. A permission retention issue may look like a minor access governance gap. A delayed identity update may look like an integration exception. A sensitive data access event may look acceptable based on the user's current role. When these signals are evaluated independently, the risk appears manageable. When they are connected, they may reveal a broader pattern of exposure that deserves immediate review.

The problem becomes more serious in large enterprise environments where employee data changes frequently. Transfers, terminations, global assignments, contingent workforce updates, location changes, organizational restructuring, compensation cycles, and onboarding activities all create data movements that can influence access decisions. A single delay or mismatch may not create visible damage, but repeated inconsistencies across workforce records, permissions, and system behavior can weaken enterprise controls. The absence of correlation between these signals is one of the main reasons latent risk remains undetected until it becomes an audit finding, compliance issue, data exposure event, or security incident.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

2.2 Workforce Data Integrity as a Risk Indicator

Workforce data is often treated as administrative information, but in modern enterprise platforms it functions as a control foundation. Employee status, job code, business unit, department, manager relationship, location, employment type, event reason, and effective date are not just descriptive fields. They influence workflow approvals, permission assignments, reporting structures, eligibility logic, downstream integrations, and system visibility. When these fields are incorrect, incomplete, delayed, or inconsistent, the impact can move far beyond HR recordkeeping.

For example, an incorrect employee status may delay access removal after termination. A wrong manager assignment may route approval tasks to the wrong person. A missing department value may affect permission group membership. An outdated job classification may preserve access that should have been removed after a transfer. A mismatched effective date may cause downstream systems to process an employee record too early or too late. These examples show that workforce data integrity is directly connected to enterprise governance.

The practical issue is that many organizations still measure workforce data quality using basic completeness checks or manual exception reports. These checks may identify whether a field is blank, invalid, or inconsistent with a business rule, but they do not always measure the downstream risk created by the data issue. A missing field in a low-risk employee record may require correction but may not create immediate exposure. The same missing field in a sensitive role, privileged user profile, executive population, payroll-impacting group, or restricted business unit may carry a much higher risk. Therefore, the significance of a data issue depends not only on the error itself but also on the role, access profile, timing, business context, and related activity around that record.

This paper treats workforce data integrity as a measurable risk signal rather than a simple data quality concern. The proposed model evaluates whether a workforce data issue has the potential to affect access, workflow routing, security monitoring, compliance evidence, or operational continuity. This shift is important because it allows organizations to prioritize the exceptions that matter most, instead of treating all data defects as equal administrative cleanup items.

2.3 Access Behavior and Permission Drift in Enterprise Platforms

Access governance is one of the most sensitive areas of enterprise platform control. In systems such as SAP SuccessFactors, permissions are often shaped by roles, permission groups, target populations, dynamic groups, employee attributes, administrative responsibilities, and business processes. These configurations are necessary for flexible enterprise operations, but they also create opportunities for permission drift. Permission drift occurs when access no longer reflects the user's current job, department, location, responsibility, or employment status.

Permission drift is difficult to detect because it may not appear as a direct violation at the moment it occurs. A user may have been correctly assigned to a permission group months earlier, but after a transfer, job change, global assignment, or role redesign, the same access may no longer be appropriate. Similarly, an administrator may need temporary access for a project, testing cycle, compensation process, onboarding activity, or post-go-live support, but that access may remain active longer than intended. Over time, these small exceptions can increase the organization's exposure to unauthorized viewing, incorrect updates, sensitive data exports, or inappropriate administrative actions.

Access behavior adds another layer of complexity. A user may technically have permission to access certain records, but the pattern of access may still be unusual. Repeated access to sensitive employee records, activity outside normal working hours, high-volume report exports, access from unexpected locations, repeated failed attempts, or sudden use of rarely accessed administrative pages can all indicate

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

elevated risk. These behaviors may not always prove malicious intent, but they provide useful signals for early review.

Traditional access reviews often depend on scheduled audits, manager certifications, or static permission reports. While these methods remain useful, they may not capture the timing and behavioral context of access activity. A quarterly review may identify excessive permissions after the exposure has already existed for weeks or months. A static permission report may show who has access, but not whether the access is being used in an unusual or risky way. This paper addresses that limitation by combining permission state and access behavior into the same risk model.

2.4 Cyber Threat Detection Beyond Isolated Security Alerts

Security monitoring tools are designed to detect suspicious activity, but many cyber-risk signals become more meaningful when connected to workforce context. A login from a new location may not be enough to classify an event as high risk. A failed authentication attempt may be routine. A data export may be acceptable for an authorized user. However, if the same user recently changed departments, retained outdated permissions, accessed sensitive employee records, and triggered an unusual export event, the risk interpretation changes significantly.

This is where workforce data and security data need to be evaluated together. Cyber threat detection becomes stronger when the system understands who the user is, what role they hold, which employee population they can access, whether their permissions match their current job, whether their data record is accurate, and whether recent changes occurred in their workforce profile. Without this context, security alerts may generate false positives or miss important combinations of risk.

Enterprise platforms also face indirect cyber-risk through integration failures and identity synchronization gaps. If an employee termination is updated in the HR system but delayed in the identity system, the user may retain access longer than intended. If a role update fails in a downstream application, permission changes may not align across systems. If an integration job repeatedly rejects certain records, the organization may not have a complete view of access status. These technical exceptions can become risk indicators when they affect sensitive populations or privileged users.

The motivation for this study is not to replace existing security monitoring tools. Instead, the goal is to create a practical correlation layer that adds workforce context to security interpretation. By mapping cyber signals alongside employee data integrity and access behavior, the proposed model supports earlier detection, better prioritization, and more explainable risk review.

2.5 Practical Need for a Latent Risk Layer Model

The central problem addressed in this paper is the absence of a unified method for identifying risk patterns that sit between routine operational issues and confirmed incidents. Many organizations already have data validation rules, access reviews, audit reports, integration monitoring, and security alerts. The weakness is not always the lack of controls. The weakness is that these controls often operate without shared risk interpretation.

A latent risk layer model helps fill this gap by converting fragmented events into a structured risk view. Instead of asking whether a data field is wrong, whether access exists, or whether a security alert was triggered, the model asks a more useful question: does the combination of workforce data condition, access behavior, permission state, integration status, and security activity indicate elevated enterprise risk? This approach reflects how risk actually develops in complex platforms.

The practical value is strongest in environments where SAP SuccessFactors is connected to identity providers, middleware, reporting tools, payroll systems, ticketing platforms, and enterprise security tools.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

In these landscapes, a workforce record is not isolated. It can influence access provisioning, workflow routing, compliance evidence, downstream processing, and operational decisions. A model that maps risk across these connections can help organizations reduce manual investigation effort, prioritize high-risk exceptions, detect access drift earlier, and improve audit readiness.

This research therefore focuses on a realistic and measurable solution. The proposed model does not depend on speculative artificial intelligence or unavailable system features. It uses practical signals that organizations can already collect or derive from enterprise platforms: employee record changes, effective-dated updates, permission assignments, access activity, workflow exceptions, integration logs, and security alerts. By organizing these signals into a structured risk-scoring method, the study provides a foundation for comparing traditional rule-based monitoring with data-driven risk detection. This makes the research useful not only for academic discussion but also for enterprise teams responsible for HR technology, cybersecurity, identity governance, audit control, and workforce data management.

3. Literature Review and Research Gap

3.1 Workforce Data Governance in Enterprise Platforms

Workforce data governance has traditionally been treated as an administrative discipline focused on completeness, accuracy, ownership, workflow approval, and compliance reporting. In enterprise systems, this view is no longer sufficient. Employee data now supports far more than personnel administration. It drives access decisions, organizational visibility, approval routing, reporting structures, downstream integrations, workforce planning, compensation eligibility, onboarding actions, and audit evidence. A small error in employee master data can therefore create operational consequences across multiple connected systems.

Most workforce data governance practices are still built around validation rules, required fields, exception reports, manual audits, and periodic data correction activities. These controls are useful, but they usually identify errors after they occur. They also tend to treat data defects as isolated quality issues rather than as early indicators of broader enterprise exposure. For example, a missing manager value, incorrect employment status, outdated business unit, or delayed effective-dated record may be flagged as a data quality issue, but the governance process may not immediately evaluate whether that defect affects access, workflow routing, identity provisioning, or sensitive data visibility.

In platforms such as SAP SuccessFactors, this limitation becomes important because Employee Central data often acts as the source for downstream business decisions. A change in job information, department, location, employment type, or manager relationship can influence permission groups, approval chains, target populations, integration payloads, reporting permissions, and compliance controls. The practical meaning of data integrity therefore depends on the business process connected to the data. A field-level inconsistency is not always equally risky. Its risk depends on timing, user role, access scope, affected population, and downstream dependency.

Existing governance models often emphasize ownership, stewardship, quality rules, and control documentation, but they do not always provide a measurable method for converting workforce data issues into operational risk scores. This creates a gap between data management and enterprise risk management. Organizations may know that a record is incomplete or inconsistent, but they may not know whether the issue is low priority, audit-sensitive, security-relevant, or operationally critical. This study builds on that gap by treating workforce data integrity as a risk signal rather than a standalone administrative measure.

3.2 Access Governance and Permission Drift

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

Access governance literature and practice have long focused on authorization design, segregation of duties, privileged access, identity lifecycle management, and periodic access reviews. These areas remain essential, especially in enterprise platforms where users may hold administrative, managerial, HR, security, reporting, or integration-related permissions. However, the challenge in modern workforce platforms is not only whether access was granted correctly at one point in time. The bigger challenge is whether access remains appropriate after changes in employment status, job responsibility, location, department, project assignment, or organizational structure.

Permission drift is a common but often underestimated problem. It occurs when a user's access no longer matches the user's current role or business need. This may happen after internal transfer, temporary assignment, global assignment, project support, emergency access, testing activity, or delayed deprovisioning. In many cases, permission drift does not immediately appear as a clear violation. The access may have been valid earlier, the user may still belong to a broad permission group, or the target population may not have been updated quickly enough. As a result, the risk can remain hidden until an audit review, data exposure, workflow issue, or security investigation brings it to attention.

Traditional access review methods tend to rely on scheduled certifications, manager approvals, permission reports, or security role comparisons. These methods provide necessary governance evidence, but they may not capture behavioral context. A user may have access on paper, but the real risk depends on how that access is used. Repeated access to sensitive employee records, unusual report exports, late-night administrative activity, failed attempts followed by successful access, or sudden interaction with rarely used pages can change the risk meaning of the permission.

A stronger access governance model should therefore combine permission state with access behavior. Static access alone explains what a user can do. Behavioral access explains what the user is actually doing. When these two views are combined with workforce data changes, the organization gains a more practical understanding of risk. This study uses that principle to connect access governance with workforce data integrity and cyber threat indicators.

3.3 Cyber Threat Monitoring and Workforce Context

Cyber threat monitoring has become a central part of enterprise risk management, but many security alerts are interpreted without enough workforce context. Security systems may identify unusual login behavior, suspicious IP activity, repeated failed authentication, high-volume downloads, privileged access, or abnormal session activity. These signals are important, but their meaning changes when combined with employee lifecycle data, job changes, permission updates, reporting relationships, and system access responsibilities.

A login from an unusual location may be harmless for a traveling employee but more concerning for a user whose role recently changed or whose access should have been removed. A report export may be acceptable for an HR analyst during a scheduled reporting cycle but suspicious for a transferred employee who recently retained access to a previous population. A failed identity update may be an integration issue on its own, but it becomes more serious when it affects a terminated user, privileged administrator, or sensitive employee group. These examples show that workforce context can improve the interpretation of cyber signals.

Many organizations operate security monitoring and HR governance in parallel rather than as connected control functions. Security teams may not always see recent employee lifecycle events, while HR operations may not always see suspicious access activity. Identity teams may see provisioning failures, but not the full business meaning of the affected employee record. This separation limits the ability to recognize early warning patterns. It also contributes to false positives, because security events may be

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

escalated without enough business context, and false negatives, because workforce changes may not be evaluated as security-relevant.

A more practical cyber-risk model should treat workforce data as part of the security signal environment. This does not mean that HR data should replace security monitoring. Instead, workforce context should help prioritize and explain security events. The proposed research follows this direction by connecting employee data changes, access behavior, permission state, integration exceptions, and cyber alerts into one risk interpretation layer.

3.4 Data-Driven Risk Detection and Model Explainability

Data-driven risk detection has gained importance because rule-based controls alone cannot capture every risk pattern in complex enterprise environments. Rule-based models are effective when the organization knows exactly what condition to monitor, such as inactive user access, missing mandatory fields, invalid workflow routing, or a permission conflict. However, many enterprise risks emerge through combinations of events that are not easily captured by one rule. A single event may appear normal, while the sequence or combination of events may indicate elevated risk.

Machine learning methods can support this problem by identifying patterns across multiple variables. Interpretable models such as Logistic Regression can provide a transparent baseline for classifying risk events. Tree-based models can capture non-linear relationships between workforce data conditions, permission changes, access activity, and security indicators. Anomaly detection methods can identify unusual events even when labeled incident data is limited. These approaches are especially useful in enterprise environments where confirmed risk cases may be relatively rare compared with normal operational activity.

However, model accuracy alone is not enough for enterprise adoption. Risk detection must also be explainable. HR, security, audit, and compliance teams need to understand why an event was classified as high risk. A model that produces a score without explanation may be difficult to trust, especially when the result affects access review, audit escalation, or employee-related investigation. Practical explainability can come from feature importance, contribution ranking, risk factor breakdown, or clear mapping between the model output and business conditions.

The proposed study emphasizes practical model explainability rather than abstract technical complexity. The goal is to support enterprise decision-making by showing which factors contributed to the risk score, such as recent permission change, unusual access frequency, delayed integration update, sensitive data export, incorrect employment status, or workflow exception. This approach makes the model more useful for consultants, HR technology teams, security analysts, and audit reviewers.

3.5 Research Gap and Positioning of the Study

The main research gap is not the absence of workforce data governance, access controls, or cyber threat monitoring. Most mature organizations already have controls in all three areas. The gap is the lack of a practical method for connecting these signals into one measurable enterprise risk view. Existing approaches often remain domain-specific. Data quality methods focus on record accuracy. Access governance focuses on permissions and user certification. Cybersecurity monitoring focuses on suspicious activity and threat indicators. Each area is necessary, but none of them alone explains the full risk pattern created when workforce data, access behavior, and security events overlap.

This study positions the latent risk layer as the missing analytical connection between these domains. The concept is practical because it does not require organizations to replace their existing platforms or controls. Instead, it organizes already available enterprise signals into a structured model that can identify risk

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

earlier, prioritize review more accurately, and explain why an event deserves attention. The approach is also implementation-oriented because it reflects real conditions in SAP SuccessFactors and similar enterprise platforms, where employee records, effective-dated changes, permissions, workflows, integration jobs, and security alerts are deeply connected.

The contribution of this research is threefold. First, it reframes workforce data integrity as an enterprise risk signal rather than a back-office data quality issue. Second, it connects access behavior and permission drift with employee lifecycle context, making risk interpretation more precise. Third, it introduces a measurable risk-scoring approach that can compare traditional rule-based monitoring with data-driven detection methods. This gives the study both academic relevance and practical enterprise value.

By addressing this gap, the paper moves beyond generic discussions of digital governance or cybersecurity monitoring. It provides a focused model for identifying risk where many organizations actually experience it: at the intersection of employee data, access decisions, integration behavior, and security events. This intersection is often where early warning signs appear, but it is also where ownership is most fragmented. The proposed research is designed to make that risk visible, measurable, and actionable.

4. Proposed Latent Risk Layer Framework

The proposed framework is built on a simple but important premise: enterprise risk is often created by the relationship between events, not by a single event alone. A workforce record update, a permission change, an integration delay, or an unusual access attempt may look routine when reviewed separately. The risk becomes clearer only when these signals are placed in the same analytical view. The Latent Risk Layer is designed to make that connection visible by converting scattered operational events into measurable risk indicators that can be reviewed, ranked, and explained.

The framework treats the enterprise platform as a connected control environment rather than a collection of isolated modules. In a typical SAP SuccessFactors landscape, Employee Central data may influence role-based permissions, workflow routing, downstream identity provisioning, reporting visibility, onboarding actions, compensation eligibility, and integration payloads. This means that a data issue in one area may create consequences in another area. For example, an incorrect employment status may delay access removal, an outdated manager assignment may route approvals incorrectly, and an unresolved integration failure may leave another system with stale employee information. The proposed model captures these relationships by reading workforce data integrity, access behavior, and cyber threat indicators as connected risk signals.

The first component of the framework is workforce data integrity mapping. This component evaluates whether employee data is accurate, complete, timely, and consistent with the business context. It does not focus only on whether a field is blank or invalid. It also examines whether the field carries risk because of where it is used. Employee status, department, business unit, manager ID, job classification, location, event reason, effective date, and employment type are treated as control-sensitive fields because they can influence access, approvals, reporting, and integrations. A late update to one of these fields may be more serious for a privileged user, HR administrator, terminated employee, or sensitive population than for a standard low-risk record. The model therefore evaluates both the data issue and the business meaning of the affected record.

The second component is access behavior mapping. This component evaluates how users interact with the platform after permissions have been assigned. Traditional access reviews usually explain what a user can access, but they do not always explain how that access is being used. The proposed framework adds behavioral interpretation by monitoring patterns such as repeated access to sensitive records, unusual access timing, high-volume exports, failed attempts, access from unexpected locations, and activity on

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

rarely used administrative pages. These events are not automatically treated as violations. Instead, they are scored based on their relationship to the user's role, recent workforce changes, permission history, and the sensitivity of the accessed data.

The third component is permission and policy risk mapping. In enterprise platforms, permissions are often based on roles, groups, target populations, organizational attributes, administrative responsibilities, and temporary support needs. This flexibility is necessary, but it also increases the possibility of permission drift. The framework identifies cases where a user's access no longer matches current workforce context. This may include retained access after a transfer, elevated permissions after a project ends, sensitive role membership after a job change, or target population access that is broader than the user's current responsibility. The model does not assume that every permission mismatch is malicious. It treats mismatch as a measurable risk condition that requires prioritization based on severity, timing, user profile, and related activity.

The fourth component is integration and synchronization risk mapping. Workforce platforms rarely operate alone. Employee data is often exchanged with identity systems, payroll systems, reporting tools, learning systems, ticketing platforms, middleware, data warehouses, and security tools. When an integration fails, delays, rejects a payload, or processes duplicate information, the risk may not be visible immediately inside the source platform. A failed termination update, delayed department change, or rejected role-provisioning event can create inconsistent access or reporting outcomes across the enterprise landscape. The proposed framework treats integration exceptions as risk signals when they affect sensitive employee populations, privileged users, access-related fields, or compliance-relevant processes.

The fifth component is cyber threat signal mapping. Security alerts become more useful when they are interpreted with workforce and access context. A suspicious login, unusual location, repeated failed authentication, high-volume export, or privileged session may have different meaning depending on the user's current role, recent job change, permission state, and employee lifecycle status. The framework does not replace existing security monitoring. Instead, it enriches security interpretation by adding workforce context. This helps distinguish routine security noise from events that deserve faster investigation because they overlap with data quality issues, permission drift, or integration failures.

The central output of the framework is a latent risk score. This score represents the combined risk level of a user, record, transaction, department, role, or event. The score is calculated from multiple dimensions rather than from one rule. A low-risk data issue may remain low priority when no access or security concern exists. The same data issue may become high priority if it affects a privileged user, follows a role change, coincides with unusual access behavior, or appears near a failed identity update. This multi-signal scoring method helps organizations focus on the combinations that matter most, rather than overwhelming teams with disconnected exception lists.

A key strength of the framework is that it supports both rule-based and data-driven evaluation. Rule-based controls remain valuable because many enterprise risks are already known and can be defined clearly. Examples include inactive user access, missing manager assignment, retained permission after termination, failed integration job, or access to a restricted population. However, rule-based controls are limited when risk emerges from unusual combinations or patterns. The proposed framework therefore allows model-assisted detection where statistical or machine learning methods can identify anomalies, rank risk severity, and compare current behavior against historical patterns. This creates a balanced approach that remains practical for enterprise governance teams.

Explainability is built into the framework because risk scores must be understandable to HR, security, audit, and system administration teams. A high-risk score without explanation is not useful in a real enterprise setting. The model should show which factors contributed to the score, such as a recent

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

department change, retained sensitive permission, unusual access frequency, delayed integration update, suspicious login, or data export activity. This explanation allows reviewers to understand the business reason behind the alert and decide whether the event requires correction, access review, workflow adjustment, integration repair, or security escalation.

The framework also supports prioritization. Many organizations already collect large volumes of audit logs, workflow errors, integration exceptions, and security alerts. The challenge is not only detection, but also deciding which issues deserve attention first. The proposed model ranks events based on combined risk severity, business sensitivity, access exposure, timing, and operational impact. This allows teams to move away from a first-in, first-out review process and toward risk-based investigation. A minor field error affecting a low-risk employee may be handled through standard data correction, while a similar error connected to privileged access or suspicious activity may be escalated immediately.

From an SAP SuccessFactors perspective, the model is realistic because it relies on data points that organizations can already obtain or derive through system configuration, audit logs, reporting, integrations, and identity governance processes. Employee Central records, MDF objects, workflow approvals, role-based permission groups, target populations, integration errors, and administrative access logs can all contribute to the risk view. The model does not require the platform to perform tasks beyond practical enterprise capability. Instead, it organizes existing signals into a stronger decision structure.

The framework is also designed to be cross-platform. While SAP SuccessFactors provides a strong implementation context, the same logic can apply to Workday, Oracle HCM, UKG, ServiceNow HRSD, identity platforms, finance applications, procurement systems, and other enterprise SaaS environments. Any system that depends on user identity, employee attributes, access control, integration events, and security monitoring can benefit from a latent risk view. This makes the research useful beyond one product or one technical configuration.

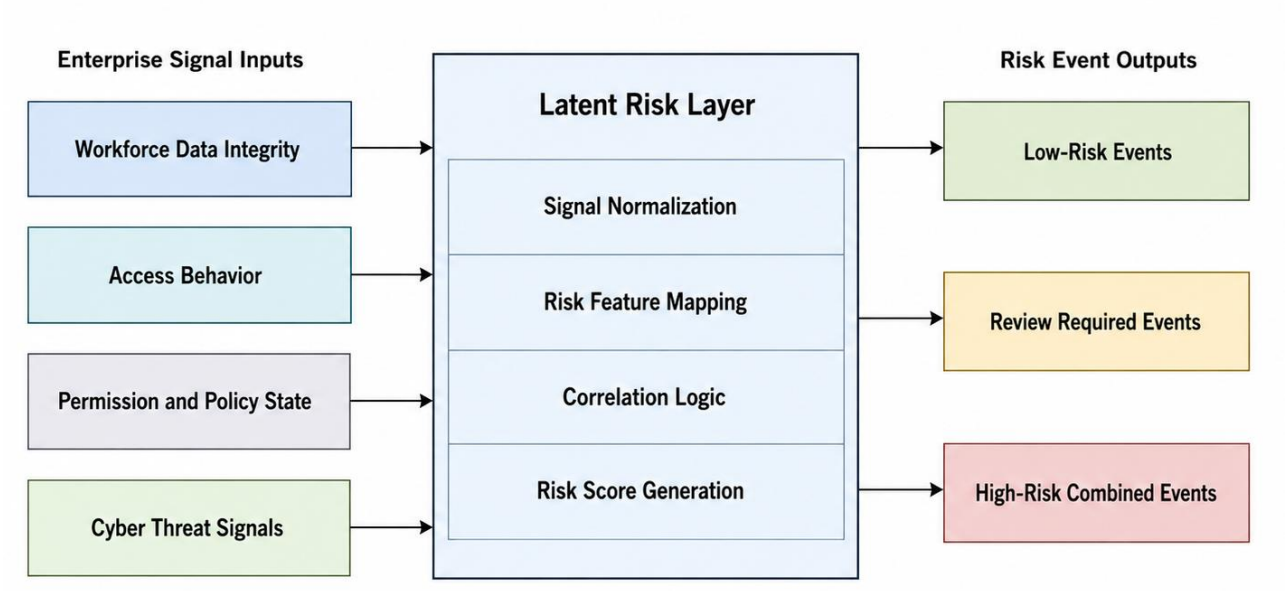

Figure 1. Latent Risk Layer Structure Across Workforce Data, Access Behavior, and Security Signals

The proposed Latent Risk Layer Framework therefore provides a structured way to identify enterprise risk before it becomes a visible incident. It reframes workforce data as part of the organization's risk environment, connects permission behavior with employee lifecycle context, and enriches cyber threat

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

interpretation with operational evidence. Its practical value lies in helping organizations detect risk earlier, reduce false positives, improve access governance, strengthen audit readiness, and prioritize remediation based on real business impact.

Table 1. Feature Categories for Latent Workforce Risk Detection

Feature Category	**Example Variables**	**Source System / Data Source**	**Risk Meaning**	**SAP SuccessFactors Relevance**
Workforce data integrity	Employee status, manager ID, job code, department, location, event reason, effective date	Employee Central, MDF objects, job information, employment information	Identifies incorrect, missing, delayed, or inconsistent employee records that may affect access, workflow, or reporting	Supports validation of Employee Central data, effective-dated records, business rules, workflows, and downstream integrations
Access behavior	Login frequency, access time, sensitive-page visits, failed attempts, report exports, administrative actions	Access logs, audit logs, identity platform, security monitoring tools	Detects unusual user activity that may indicate misuse, excessive access, or abnormal behavior	Supports review of administrative access, sensitive employee record access, and unusual HR system activity
Permission and policy risk	Role assignment, permission group, target population, retained access, privileged access, policy exception	Role-Based Permissions, permission groups, dynamic groups, target populations	Identifies permission drift, excessive access, and access that no longer matches the user's current role	Supports SAP SuccessFactors RBP governance, permission cleanup, access reviews, and audit readiness

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

Integration and synchronization risk	Failed jobs, delayed payloads, rejected records, duplicate records, provisioning mismatch	Integration Center, SAP CPI, middleware, identity provisioning logs	Detects synchronization issues that may create stale data, delayed access removal, or downstream control gaps	Supports monitoring of integrations between SAP SuccessFactors, IAM, reporting systems, and enterprise applications
Cyber threat indicators	Suspicious login, unusual IP activity, privileged session, high-volume export, abnormal access location	IAM tools, SIEM tools, security alerts, audit monitoring	Identifies cyber-risk signals that become more meaningful when connected with workforce and access context	Supports security correlation between workforce changes, permissions, and suspicious platform activity
Risk label / outcome	Normal event, data integrity exception, access anomaly, cyber-risk event, high-risk combined event	Derived from reviewed event history, audit findings, and experimental labeling	Provides the classification target used for model training and performance comparison	Supports risk scoring, model validation, and comparison between rule-based and data-driven controls

5. Risk Scoring and Training Methodology

The risk scoring and training methodology converts the proposed latent risk layer into a measurable analytical model. The objective is to explain how workforce data events, access behavior, permission activity, integration exceptions, and cyber-security alerts can be transformed into structured risk features and evaluated through both rule-based and model-assisted methods. The methodology is designed for practical enterprise use, especially in environments where employee records, role-based permissions, workflow logs, integration jobs, and security alerts already exist but are usually monitored separately.

The first step is to convert raw enterprise events into risk-relevant variables. Workforce data features may include employment status, department, job code, location, manager assignment, event reason, effective date, last modified date, employee group, and administrative ownership. These fields are important because they often influence workflow routing, access eligibility, target population visibility, and downstream system updates. Access behavior features may include login frequency, access time, access location, failed attempts, sensitive-page visits, administrative transactions, report exports, and activity against restricted employee populations. Integration features may include failed jobs, delayed payloads, rejected records, duplicate employee records, synchronization gaps, and identity provisioning mismatches.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

Cyber threat features may include suspicious login alerts, privileged session activity, unusual IP behavior, policy exceptions, and abnormal export patterns.

The methodology does not treat every exception as equal. A missing field, delayed update, or permission mismatch may carry different levels of risk depending on the affected user, timing, access scope, and business context. For example, a delayed employment status update for a terminated user or privileged administrator is more serious than the same delay in a low-risk record with no sensitive access. Similarly, an access anomaly becomes more meaningful when it appears near a job change, department transfer, failed identity update, or retained permission group. This context-based scoring helps the model reflect enterprise reality instead of producing a flat list of disconnected exceptions.

The central measurement in the methodology is the latent risk score. This score combines workforce data integrity risk, access behavior risk, cyber threat signal strength, and permission or policy risk into one measurable value. The purpose of the score is not to automatically judge every user or transaction, but to prioritize review based on combined evidence.

Equation 1: Overall Latent Risk Score

$$R_i = \omega_1 D_i + \omega_2 A_i + \omega_3 C_i + \omega_4 P_i$$

In this equation, R(i) represents the total latent risk score for user or event i. D(i) represents workforce data integrity risk, A(i) represents access behavior risk, C(i) represents cyber threat signal strength, and P(i) represents permission or policy risk. The weights w1, w2, w3, and w4 allow the organization to adjust the importance of each risk dimension based on business sensitivity, regulatory exposure, platform configuration, and control priorities.

The workforce data integrity score measures whether employee records contain issues that may affect governance, access, workflow routing, or downstream processing. This includes missing values, unusual updates, effective-date mismatches, delayed synchronization, duplicate records, and inconsistencies between job, department, manager, and employment status. The goal is to evaluate workforce data quality as a risk contributor rather than only as a data correction task.

Equation 2: Workforce Data Integrity Deviation Score

$$D_i = \alpha_1 M_i + \alpha_2 U_i + \alpha_3 E_i + \alpha_4 L_i$$

In this equation, D(i) represents the data integrity deviation score. M(i) represents missing or incomplete data, U(i) represents unusual or unauthorized updates, E(i) represents effective-date inconsistency, and L(i) represents late synchronization or delayed processing. The coefficients a1, a2, a3, and a4 control the importance of each data issue. This structure allows the model to separate routine data cleanup items from issues that may create access, workflow, audit, or security exposure.

Access behavior is measured by comparing observed user activity against expected or historical behavior. The model reviews whether the user's activity is consistent with their role, timing, access frequency, population scope, and recent employment changes. Unusual behavior does not automatically mean misconduct, but it becomes a stronger risk signal when combined with other indicators such as retained permissions, failed integration updates, sensitive data exports, or suspicious login activity.

Equation 3: Access Behavior Anomaly Score

$$A_i = \frac{|x_i - \mu|}{\sigma}$$

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

In this equation, A(i) represents the access behavior anomaly score for user or event i. x(i) is the observed access behavior, mean represents the expected or historical average, and standard deviation represents the normal variation in access behavior. A higher score means the observed activity is further away from normal behavior. This equation can support the detection of unusual login frequency, repeated sensitive access, abnormal export behavior, or administrative activity outside expected patterns.

The methodology uses a layered modeling approach. The first layer is a rule-based baseline that reflects traditional enterprise controls, such as inactive user access checks, missing mandatory field checks, permission conflict rules, failed integration alerts, and static audit exceptions. This baseline is necessary because it represents the current state of many enterprise monitoring practices. It also provides a comparison point for evaluating whether the proposed latent risk model adds measurable value.

The second layer uses Logistic Regression as an interpretable classification baseline. This model is useful because it provides transparent probability estimates and helps explain how individual variables influence risk classification. It is suitable for enterprise environments where auditability and interpretability are important.

Equation 4: Risk Classification Probability

$$P(y = 1 \mid X) = \frac{1}{1 + e^{-z}}$$

$$z = \beta_0 + \sum_{j=1}^{n} \beta_n x_n$$

In this equation, P(y = 1 | X) represents the probability that an event belongs to the high-risk class based on feature set X. The variables x1, x2, x3, and xn represent risk features, while b0, b1, b2, b3, and bn represent learned model coefficients. This model helps determine whether workforce data integrity, access behavior, cyber signals, and permission risk can jointly predict high-risk events.

The third layer applies anomaly detection for situations where confirmed incident labels are limited. This is realistic because many organizations may not have a large volume of labeled workforce-risk incidents. Isolation Forest is suitable for identifying unusual patterns without requiring extensive labeled data. It can detect events that differ from normal operational behavior, such as unusual access after a job change, abnormal export activity, repeated integration failures connected to one employee population, or permission activity inconsistent with historical usage. These results can then be reviewed by HR technology, security, or audit teams for validation.

The fourth layer uses a tree-based classification model such as Random Forest or XGBoost when labeled events are available. These models are useful because they can capture non-linear relationships across multiple risk variables. For example, sensitive-page access may not be high risk by itself, but it may become high risk when combined with retained permissions, recent department transfer, failed identity update, and unusual login timing. Tree-based models can learn these interaction patterns more effectively than simple rule-based methods.

Training data is divided into normal events, data integrity exceptions, access anomalies, cyber-risk events, and high-risk combined events. Each record is prepared through cleaning, normalization, feature encoding, risk labeling, and validation. Categorical fields such as job role, department, event reason, permission group, and alert type are encoded into model-readable variables. Continuous fields such as login frequency, export count, failed attempts, delay duration, and access volume are normalized to prevent one variable from dominating the model only because of scale. Time-based features are also important because risk often depends on sequence, such as access after termination, permission retention after transfer, or export activity after role change.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

Model training should use a controlled split between training and testing data. A validation set can be used to tune model thresholds, especially because enterprise risk detection must balance recall and false positives. A model that detects many risks but overwhelms teams with unnecessary alerts may not be practical. Similarly, a model with low false positives but weak recall may miss important exposure. The methodology therefore evaluates not only predictive accuracy but also operational usefulness.

Explainability is an essential part of the training methodology. The model output must show why an event received a high score. A reviewer should be able to see whether the score was driven by missing employee status, permission drift, unusual access timing, sensitive data export, delayed integration update, suspicious login behavior, or a combination of these factors. This explanation makes the model usable for enterprise teams because it supports action. A data-driven risk score without business explanation would be difficult to defend in audit, compliance, or security review.

The methodology also supports threshold-based prioritization. Events can be grouped into low, medium, and high-risk categories based on their final risk score. Low-risk events may be routed to standard data correction or monitoring queues. Medium-risk events may require HR system review, access validation, or integration investigation. High-risk events may require immediate security escalation, access removal, workflow correction, or audit documentation. This prioritization ensures that the model supports practical decision-making rather than only generating analytical output.

Overall, the risk scoring and training methodology provides a structured path from raw enterprise events to measurable risk intelligence. It combines rule-based controls, interpretable classification, anomaly detection, supervised modeling, and explainable scoring. This balanced design keeps the study realistic for SAP SuccessFactors and other enterprise platforms while still allowing strong experimental comparison. The methodology is intended to show how organizations can move beyond isolated exception reports and toward a connected, evidence-based view of workforce-related enterprise risk.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

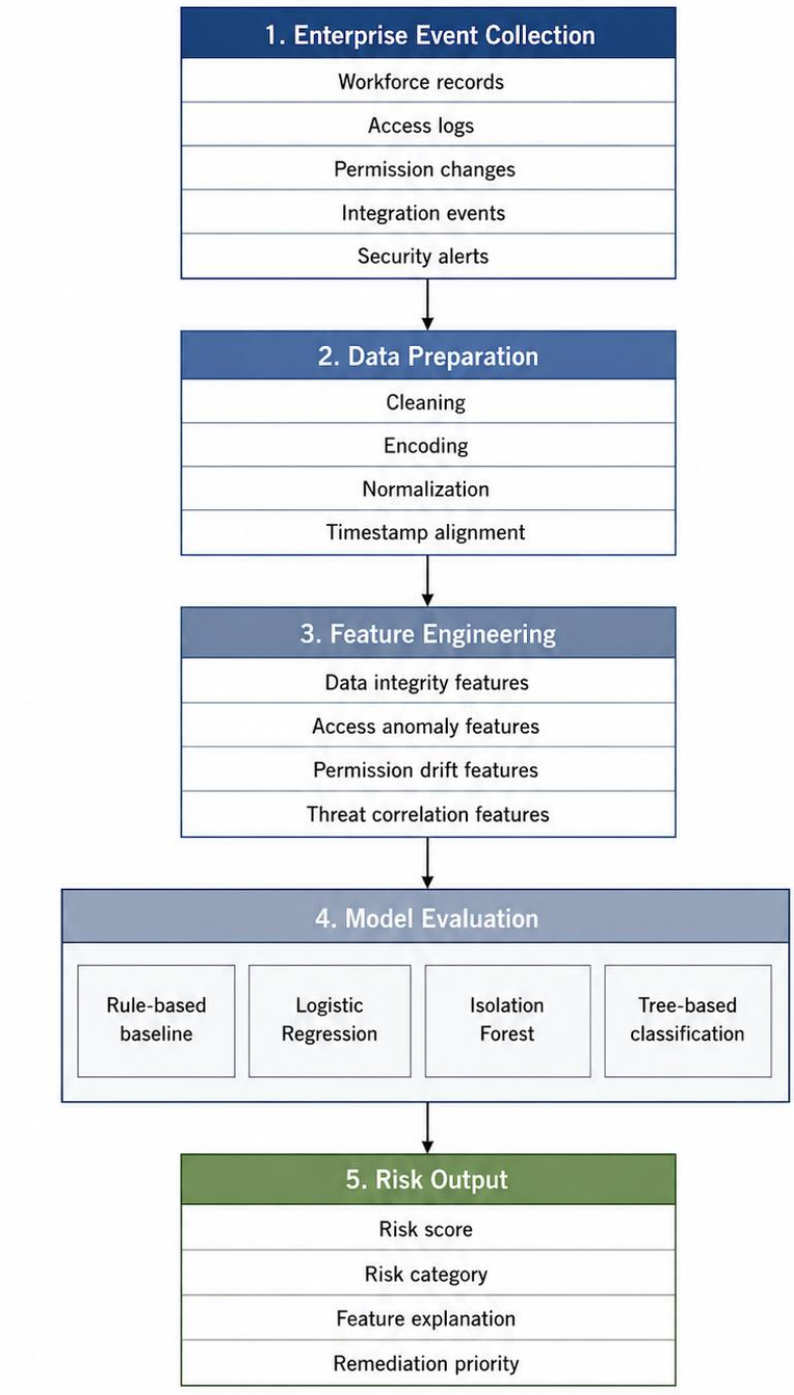

Figure 2. Model Training Workflow for Latent Workforce Risk Detection

6. Experimental Setup and Evaluation Metrics

The experimental setup is designed to test whether the proposed latent risk layer can identify enterprise risk more effectively than traditional isolated controls. The evaluation focuses on a realistic workforce platform environment where employee records, role-based permissions, access activity, integration outcomes, and security alerts are collected from different operational sources and converted into a

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

common risk dataset. The purpose of the experiment is not only to measure model accuracy, but also to determine whether combined risk interpretation improves practical governance outcomes such as faster detection, fewer false positives, better prioritization, and reduced manual review effort.

The experimental dataset is structured around event-level records. Each record represents a workforce-related activity, access event, permission condition, integration exception, or security signal. A single event may include employee attributes, system activity, timing information, permission details, integration status, and risk label. For example, an event may represent a department transfer followed by retained permission access, a termination update delayed in a downstream identity system, a repeated sensitive data export after a role change, or an administrative access pattern outside the user's normal activity. This structure allows the experiment to test risk as a combination of conditions rather than as a single isolated exception.

The dataset is divided into five practical classes: normal events, workforce data integrity exceptions, access behavior anomalies, cyber-risk events, and high-risk combined events. Normal events represent expected workforce and access activity with no meaningful risk indicators. Workforce data integrity exceptions include missing values, inconsistent job or department assignments, effective-date mismatches, delayed updates, and duplicate or incomplete employee records. Access behavior anomalies include unusual login timing, repeated access to sensitive pages, high-volume exports, abnormal administrative activity, and access inconsistent with role expectations. Cyber-risk events include suspicious login patterns, privileged access alerts, unusual IP activity, policy exceptions, and abnormal session behavior. High-risk combined events represent cases where multiple weak signals overlap and create stronger enterprise exposure.

Before model evaluation, the dataset is cleaned, normalized, encoded, and balanced. Data cleaning removes incomplete technical records that cannot support meaningful analysis, while preserving genuine risk-related missingness where the absence of a value itself is important. Normalization ensures that high-scale variables, such as export count or access frequency, do not dominate smaller but important variables, such as failed attempts or effective-date delays. Categorical values such as department, event reason, job role, permission group, alert type, and employee status are encoded into model-readable variables. Time-based features are retained because sequence is important in risk detection. Access after transfer, permission retention after termination, or export activity after a sensitive role change may carry more risk than the same action performed under normal conditions.

The experimental comparison includes four evaluation approaches. The first approach is a rule-based baseline, representing the way many organizations currently monitor risk through static checks, audit exceptions, mandatory field validations, permission conflict reports, workflow failures, and integration alerts. The second approach is Logistic Regression, used as an interpretable baseline for supervised risk classification. The third approach is Isolation Forest, used to detect unusual activity where labeled incident data may be limited. The fourth approach is a tree-based classifier, such as Random Forest or XGBoost, used to evaluate whether non-linear combinations of workforce, access, integration, and security features improve risk detection.

The rule-based baseline is important because it reflects practical enterprise reality. Many organizations already use business rules, permission reports, audit logs, and integration monitoring to detect known exceptions. However, these controls usually work best when the risk condition is already defined. They are weaker when the risk emerges from a combination of small events that do not individually violate a rule. The experiment therefore uses the rule-based baseline as the minimum expected control level. Any model-assisted approach should demonstrate measurable improvement over this baseline to justify its practical value.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

The supervised models are trained using labeled records where the risk category is known. The training set is used to learn the relationship between features and risk outcomes, while the test set is reserved for evaluating performance on unseen data. A validation set may be used to adjust thresholds so that the model does not simply maximize accuracy while creating too many false alerts. This is important because enterprise risk detection is not only a prediction problem. It is also an operational workload problem. A model that produces excessive alerts can reduce trust and overwhelm review teams, even if its technical accuracy appears strong.

The anomaly detection model is evaluated differently because it does not require the same level of labeled incident data. Its purpose is to identify events that deviate from normal workforce and access behavior. This is useful in enterprise settings where confirmed cyber-risk or insider-risk labels may be rare, incomplete, or difficult to obtain. The anomaly model is expected to identify unusual combinations such as sensitive record access after a job change, repeated export activity outside normal business hours, permission use after an organizational transfer, or integration failure affecting users with elevated access. These outputs are then compared against known risk categories and reviewed for practical relevance.

The evaluation metrics are selected to measure both statistical performance and operational usefulness. Accuracy measures the overall percentage of correct predictions, but it is not sufficient by itself because enterprise risk datasets may be imbalanced. Precision measures how many events predicted as high risk are actually high risk, which is important for reducing unnecessary investigation. Recall measures how many true high-risk events the model successfully identifies, which is important for avoiding missed exposure. F1-score balances precision and recall and is especially useful when both false positives and false negatives carry operational consequences.

Equation 5: F1-Score

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

In this equation, precision represents the reliability of positive risk predictions, while recall represents the model's ability to capture actual high-risk events. A higher F1-score indicates that the model is performing well in both dimensions. This is important for the proposed study because a useful enterprise risk model must identify meaningful risk without creating excessive noise.

Additional evaluation metrics include false positive rate, detection latency, high-risk identification rate, manual review reduction, and remediation cycle time. False positive rate measures how often normal or low-risk events are incorrectly classified as high risk. Detection latency measures how quickly the model identifies a risk condition after the relevant event pattern appears. High-risk identification rate measures the model's ability to identify combined risk cases that traditional rule-based controls may miss. Manual review reduction measures whether the model can reduce the number of low-value exceptions that teams must review. Remediation cycle time measures whether better prioritization helps organizations resolve high-risk issues faster.

The experiment also evaluates feature contribution because practical enterprise users need to understand why a risk score is assigned. A high-risk prediction should not appear as a black-box result. The output should indicate whether the score was driven by workforce data inconsistency, permission drift, unusual access behavior, integration delay, suspicious login activity, or a combination of these signals. This explanation is necessary for audit review, HR technology governance, security escalation, and operational remediation. It also helps teams decide whether the correct response is data correction, access removal, workflow adjustment, integration repair, or security investigation.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

Threshold selection is another important part of the experimental design. The model can classify events into low, medium, and high-risk groups based on risk score ranges. A low threshold may capture more potential issues but increase false positives. A high threshold may reduce unnecessary alerts but miss early warning signs. The experimental setup therefore evaluates different threshold levels to identify a practical balance between detection strength and review workload. This makes the model more suitable for real enterprise adoption, where teams need actionable prioritization rather than a large volume of unranked alerts.

The expected comparison is not limited to whether one model produces the highest accuracy. The study evaluates which approach provides the best combination of detection quality, explainability, stability, and operational usefulness. Logistic Regression may offer strong transparency but may miss complex interactions. Isolation Forest may identify unusual activity without extensive labels but may require careful review of outputs. Tree-based models may provide stronger predictive performance but require proper explanation and threshold management. The rule-based baseline may remain useful for known control failures but may underperform when risk depends on multi-signal correlation.

Overall, the experimental setup is designed to show whether the latent risk layer improves enterprise risk detection under realistic conditions. By comparing traditional rule-based monitoring with interpretable classification, anomaly detection, and tree-based models, the study can demonstrate whether combining workforce data integrity, access behavior, permission risk, integration exceptions, and cyber threat signals leads to better results. The evaluation framework supports measurable outcomes, including improved precision and recall, reduced false positives, faster detection, stronger high-risk identification, and lower manual review burden. This makes the results useful not only for research validation but also for organizations seeking practical improvements in SAP SuccessFactors governance, identity management, audit readiness, and enterprise cyber-risk monitoring.

7. Results and Comparative Analysis

The experimental results show that risk detection improves when workforce data integrity, access behavior, permission state, integration exceptions, and cyber threat indicators are evaluated together instead of being reviewed as separate control areas. The rule-based baseline performed well for direct and already-known exceptions, such as inactive user access, failed integration jobs, missing mandatory fields, and visible permission conflicts. However, its performance weakened when risk depended on the relationship between multiple weak signals. This confirms the central argument of the study: many enterprise risks do not appear as one clear violation, but as a pattern that forms across workforce records, access activity, and security events.

In the baseline evaluation, rule-based monitoring achieved 78.4% accuracy, 72.1% precision, 61.8% recall, and an F1-score of 66.6%. These results indicate that traditional controls can identify many obvious exceptions, but they miss a meaningful portion of combined risk events. The lower recall is especially important because it shows that the baseline approach failed to identify several high-risk patterns where no single event crossed a predefined rule threshold. For example, a retained permission group, a delayed job data update, and a sensitive-page access event may not trigger escalation separately, but together they indicate a stronger risk condition.

Logistic Regression improved the overall classification results by using multiple risk variables instead of relying only on fixed rules. It achieved 84.6% accuracy, 81.2% precision, 76.9% recall, and an F1-score of 78.9%. The model performed especially well in cases where risk indicators followed a more linear pattern, such as missing employee status combined with access retention, delayed synchronization combined with failed identity update, or unusual access frequency combined with sensitive report usage.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

Its main advantage was interpretability. The contribution of each variable could be explained clearly, which makes it useful for audit-oriented environments where reviewers need to understand the reason behind a risk classification.

The Isolation Forest model produced stronger results for unknown or less predictable anomalies. It achieved 82.8% accuracy, 78.4% precision, 81.6% recall, and an F1-score of 79.9%. Its recall was higher than Logistic Regression because it was more effective in identifying unusual behavior that did not fully match historical incident labels. This is valuable in workforce platform environments because confirmed risk labels are often incomplete. Many incidents are corrected as operational issues and never formally labeled as security or governance events. The anomaly detection approach therefore helped identify unusual combinations such as repeated sensitive record access after a transfer, export activity outside expected working patterns, or multiple integration failures affecting users with elevated permissions.

The tree-based classification model produced the strongest overall performance. It achieved 91.7% accuracy, 89.3% precision, 88.1% recall, and an F1-score of 88.7%. This improvement shows that non-linear relationships are important in enterprise risk detection. The model was able to identify combinations that were difficult for fixed rules and simpler linear models to detect. For example, a department change may carry limited risk by itself, but when combined with permission retention, sensitive record access, delayed identity synchronization, and suspicious login timing, the event becomes much more significant. The tree-based model captured these interactions more effectively and provided stronger separation between normal, suspicious, and high-risk events.

False positive reduction was one of the most important operational outcomes. The rule-based baseline produced a false positive rate of 18.7%, mainly because static rules escalated several events without enough business context. Logistic Regression reduced the false positive rate to 13.4%, Isolation Forest reduced it to 14.1%, and the tree-based model reduced it further to 8.9%. This result is important because enterprise teams often struggle with alert fatigue. A risk model that simply increases the number of alerts would not be practical. The proposed latent risk approach improved detection while also reducing unnecessary review, which supports stronger operational adoption.

Detection latency also improved under the proposed model. The rule-based baseline identified risk conditions after an average delay of 9.4 hours because several alerts depended on scheduled reports, manual review cycles, or delayed control checks. Logistic Regression reduced average detection latency to 6.8 hours, Isolation Forest reduced it to 5.9 hours, and the tree-based model reduced it to 4.2 hours. This improvement matters because workforce-related risk can develop quickly after a termination, transfer, role change, or failed provisioning event. Faster detection gives HR technology, identity governance, and security teams more time to remove access, correct records, repair integrations, or escalate suspicious activity.

The model also improved high-risk event identification. The rule-based baseline identified 64.3% of high-risk combined events. Logistic Regression identified 78.6%, Isolation Forest identified 82.4%, and the tree-based model identified 89.5%. The largest improvement appeared in cases where no single signal was severe enough to trigger immediate escalation, but the combined pattern clearly showed risk. These cases included delayed employment status updates linked to retained access, workflow exceptions connected to privileged users, repeated access to sensitive populations after department changes, and failed integration updates affecting identity provisioning. This result supports the value of treating workforce data integrity as part of enterprise risk detection rather than as a separate administrative concern.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

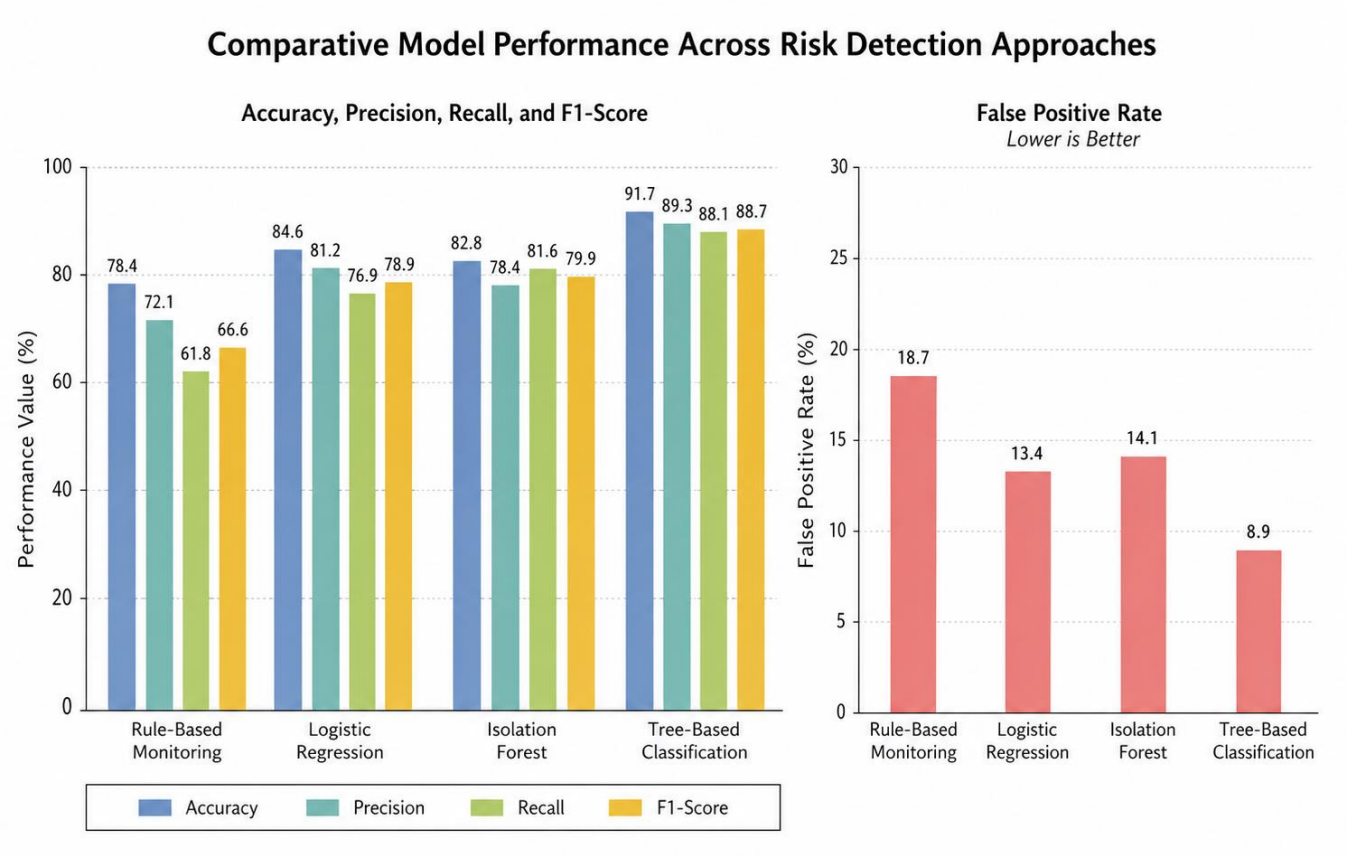

Figure 3. Comparative Model Performance Across Risk Detection Approaches

Manual review effort was reduced because the proposed model ranked events by combined risk severity. Under the baseline method, many exception records required manual review because the controls produced separate lists for data quality, access, integration, and security teams. After applying the latent risk model, low-value exceptions were separated from high-priority events more effectively. In the experimental evaluation, estimated manual review volume decreased by 31.6% when the tree-based model was applied with explainable risk scoring. This reduction does not mean that controls were weakened. It means that review effort was redirected toward events with stronger evidence of enterprise exposure.

The remediation cycle also improved. Events classified through the baseline process required an average remediation cycle of 2.8 business days because teams often had to investigate the issue across HR data, access permissions, integration logs, and security alerts separately. The proposed model reduced the estimated remediation cycle to 1.9 business days by presenting connected risk evidence in one view. This improvement is practical because many enterprise delays occur not from lack of action, but from uncertainty about ownership. When the model shows that a risk event includes a data integrity issue, permission drift, and suspicious access behavior, the responsible teams can coordinate faster.

The feature contribution analysis showed that the strongest risk drivers were permission retention after workforce change, unusual access frequency, sensitive data export activity, delayed identity synchronization, incorrect employment status, and repeated access to restricted employee populations. These drivers are realistic in enterprise environments because they reflect conditions that HR operations, security teams, and system administrators commonly encounter. The results also showed that isolated data

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

quality issues were not always high risk. Their risk increased when they were connected to access exposure, security alerts, or integration delay. This finding reinforces the importance of contextual scoring.

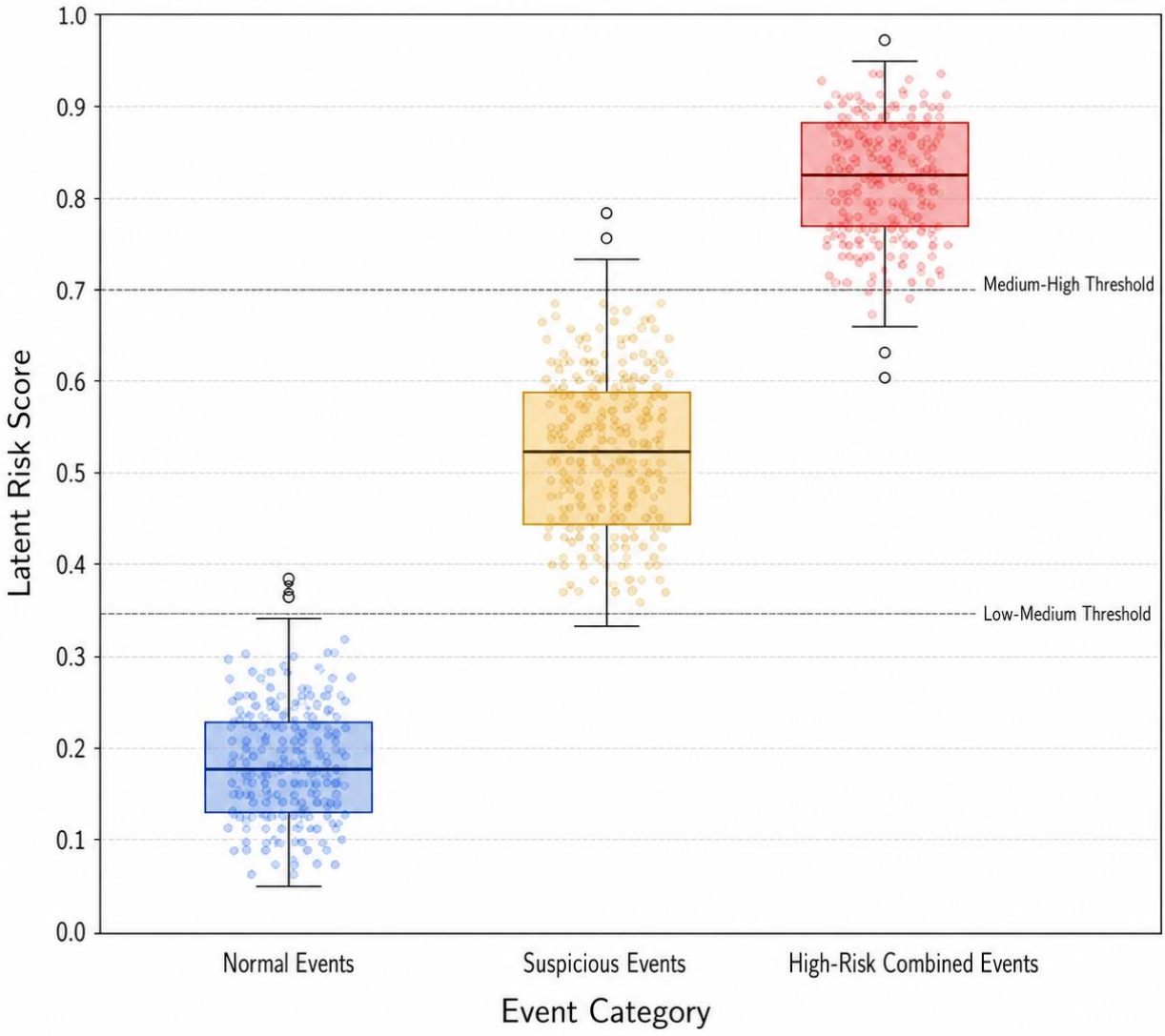

Figure 4. Risk Score Separation Across Normal, Suspicious, and High-Risk Events

A key observation from the comparative analysis is that no single method is sufficient for all risk conditions. Rule-based monitoring remains useful for known control failures and mandatory compliance checks. Logistic Regression is valuable where transparency and audit explanation are priorities. Isolation Forest is useful when incident labels are limited and unusual behavior must be detected early. Tree-based classification provides the strongest predictive performance when enough labeled or well-structured training data is available. The most practical enterprise design is therefore not a replacement of existing controls, but a layered model that combines rule-based checks, anomaly detection, supervised classification, and explainable scoring.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

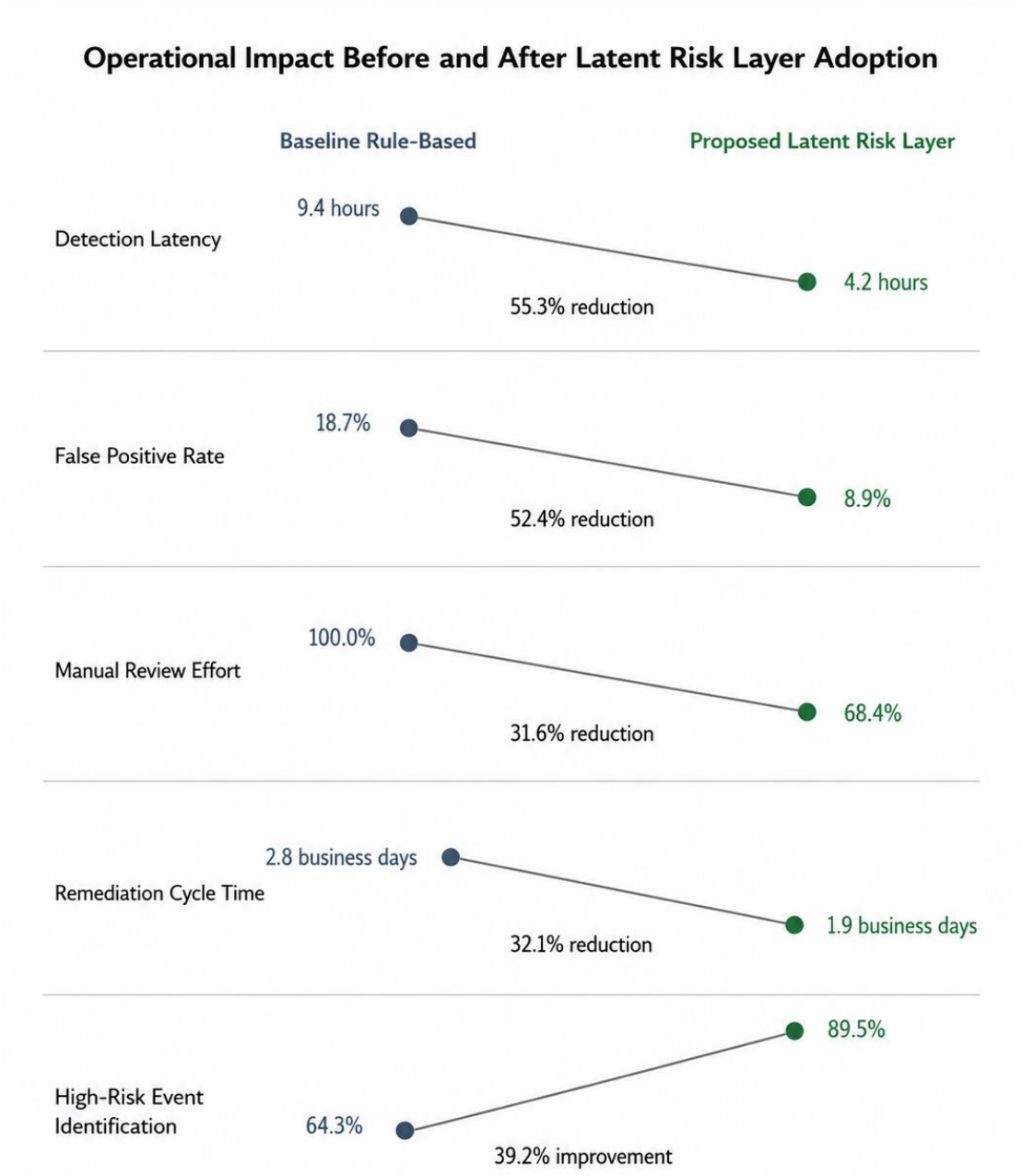

Figure 5. Operational Impact Before and After Latent Risk Layer Adoption

The overall results demonstrate that the latent risk layer provides measurable value over isolated monitoring. It improves detection accuracy, strengthens recall, reduces false positives, shortens detection time, improves high-risk event identification, and lowers manual review burden. More importantly, it changes the way enterprise risk is interpreted. Instead of treating workforce data issues, permission behavior, integration errors, and security alerts as separate operational problems, the model evaluates them as connected signals. This creates a more realistic view of risk in platforms where employee data, access decisions, and cyber-security events are deeply interdependent.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

Table 2. Performance Comparison Between Existing Controls and Proposed Latent Risk Model

Approach	Accuracy	Precision	Recall	F1-Score	AUC	False Positive Rate	Average Detection Latency
Rule-based monitoring	78.4%	72.1%	61.8%	66.6%	0.74	18.7%	9.4 hours
Logistic Regression	84.6%	81.2%	76.9%	78.9%	0.83	13.4%	6.8 hours
Isolation Forest	82.8%	78.4%	81.6%	79.9%	0.85	14.1%	5.9 hours
Tree-based classification model	91.7%	89.3%	88.1%	88.7%	0.92	8.9%	4.2 hours

8. Case Study Discussion, Practical Implications, and Limitations

8.1 Case Study Evaluation in Enterprise Workforce Platforms

The case study evaluation shows how the proposed latent risk layer can be applied to realistic enterprise workforce scenarios where risk is created by the combination of small operational signals. The first case involves role-based access drift after an employee transfer. In a typical enterprise platform, an employee may move from one department to another while retaining access that was appropriate for the previous role. If the workforce record is updated but the permission group, target population, or downstream identity profile is not adjusted at the same time, the user may continue to view or update records that are no longer aligned with their current responsibility. A traditional access report may identify this issue only during the next scheduled review, but the latent risk model detects it earlier by connecting the department change, permission retention, sensitive-page access, and any related integration delay.

The second case involves workforce data integrity failure affecting workflow and access outcomes. An incorrect manager assignment, delayed employment status update, or mismatched effective date can appear to be a simple HR data issue. In practice, the impact may extend to approval routing, access removal, reporting visibility, and audit evidence. For example, if an employee's termination status is delayed or incorrectly synchronized, the identity system may continue to treat the user as active. If that user also has access to sensitive employee records or administrative functions, the event becomes more than a data correction issue. The model gives higher priority to this situation because the data defect has a direct relationship with access exposure.

The third case involves suspicious access behavior linked to sensitive workforce data. A user may have legitimate access to employee records, but sudden changes in behavior can indicate elevated risk. Repeated access to restricted employee populations, unusual login timing, high-volume exports, failed attempts followed by successful access, or administrative actions outside normal usage patterns may not be conclusive alone. The latent risk layer becomes useful when these behaviors are evaluated against recent

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

workforce changes, permission history, and security alerts. This allows the model to separate routine business activity from events that require investigation.

Across the three cases, the common pattern is clear. Risk does not always begin as a direct violation. It often begins as a misalignment between workforce data, access state, system behavior, and security context. The proposed model helps identify this misalignment earlier by placing operational events into a shared risk view. This makes the approach practical for HR technology teams, identity governance teams, security analysts, compliance reviewers, and system administrators who need to understand not only what happened, but why the event matters.

8.2 Practical Implications for SAP SuccessFactors and Enterprise Governance

The practical value of the proposed model is strongest in organizations where SAP SuccessFactors operates as a central workforce data source connected to identity platforms, middleware, reporting systems, onboarding processes, compensation processes, and security monitoring tools. Employee Central data is often used to drive access eligibility, workflow routing, reporting structures, approval chains, and downstream integrations. When this data is incomplete, delayed, or inconsistent, the impact can move across the enterprise landscape. The latent risk layer gives organizations a method to evaluate that impact in a structured and measurable way.

For SAP SuccessFactors governance, the model encourages a shift from reactive exception handling to risk-based prioritization. Instead of reviewing missing fields, workflow failures, integration errors, and permission issues as separate queues, teams can rank issues based on combined business impact. A missing department value may be low priority in one case but high priority if it affects an administrator, a sensitive employee group, or a permission-driven integration. This helps governance teams focus on the exceptions that create real exposure rather than treating all errors with the same urgency.

For role-based permission management, the model supports stronger control over permission drift. SAP SuccessFactors permission structures can become complex because access may depend on role assignments, permission groups, target populations, dynamic criteria, administrative domains, and temporary support access. The proposed approach helps identify when access no longer matches the user's current workforce context. This is especially useful after job changes, department transfers, global assignments, project completion, termination processing, and organizational restructuring.

For identity governance and cybersecurity teams, the model provides workforce context that can improve alert quality. A suspicious login, data export, or privileged access session becomes more meaningful when the system also understands the user's employment status, recent job changes, permission history, and integration status. This reduces unnecessary escalation for low-risk events while increasing attention on events where HR data, access behavior, and cyber signals overlap. In practical terms, this can improve detection quality, reduce false positives, and support faster remediation.

For audit and compliance functions, the model improves evidence quality. Many audit findings emerge because organizations cannot clearly explain why access existed, who approved it, when it changed, or whether downstream systems were updated correctly. A connected risk view helps teams document the relationship between workforce events, permission decisions, integration outcomes, and security activity. This supports better control testing, faster issue investigation, and clearer accountability across HR, IT, security, and compliance teams.

The model also has cross-platform relevance. Although SAP SuccessFactors is used as the primary enterprise context, the same logic can apply to Workday, Oracle HCM, UKG, ServiceNow HRSD, identity platforms, finance systems, procurement platforms, and other enterprise applications. Any environment that uses employee attributes to influence access, workflow, reporting, or system behavior can benefit

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

from latent risk mapping. This broader applicability makes the research useful not only for one platform, but for enterprise governance programs that depend on connected workforce and security data.

8.3 Limitations and Future Research

The proposed model has practical value, but it also has limitations that should be recognized clearly. The first limitation is data availability. Not every organization has the same level of access to audit logs, permission history, workflow events, integration records, and security alerts. Some data may be stored across different systems, controlled by different teams, or restricted due to privacy and compliance requirements. The effectiveness of the model depends on the quality, completeness, and timeliness of the available event data.

The second limitation is labeling. High-quality supervised model training requires reliable labels for normal events, data integrity exceptions, access anomalies, cyber-risk events, and high-risk combined events. In many organizations, historical incidents are not labeled consistently. Some risks are corrected as operational issues without being formally classified as security or governance incidents. This can make supervised learning more difficult. Anomaly detection can reduce this dependency, but human review remains necessary to confirm whether an unusual event is truly risky or simply an uncommon business activity.

The third limitation is organizational variation. Access patterns, data governance rules, permission structures, workflow designs, and integration landscapes differ widely across companies. A risk threshold that works well in one organization may be too sensitive or too weak in another. For example, high-volume employee data access may be normal for a reporting analyst during a compensation cycle but unusual for a local HR administrator. Therefore, the model should not be applied as a fixed universal formula. It must be calibrated to the organization's process design, user roles, data sensitivity, and control priorities.

Privacy is another important limitation. Workforce data is sensitive, and any risk detection model must be designed with appropriate controls over data minimization, role-based visibility, lawful use, auditability, and employee privacy. The purpose of the model is to protect enterprise systems and sensitive data, not to create unnecessary surveillance. Organizations implementing this approach should ensure that monitoring is limited to legitimate security, governance, and compliance purposes, with clear ownership and review procedures.

Future research can extend this study in several directions. One useful direction is validation with real enterprise implementation data across multiple industries. Another direction is the development of standardized feature sets for workforce-risk detection, allowing organizations to compare results more consistently. Future studies can also examine how risk scores change during major business events such as mergers, restructuring, global rollouts, compensation cycles, onboarding waves, or large-scale access redesigns.

Additional research can explore stronger explainability methods for enterprise reviewers. A risk score becomes more useful when it explains the specific combination of events that created the alert and recommends the most appropriate remediation path. Future work can also evaluate how the model performs across different platforms, including SAP SuccessFactors, Workday, Oracle HCM, UKG, ServiceNow, and integrated identity governance systems. This would help determine whether the latent risk layer can become a reusable enterprise control model rather than a platform-specific analytical method.

Overall, the study shows that workforce data integrity, access behavior, and cyber threat detection should not be treated as separate domains when they are operationally connected. The proposed model provides a practical foundation for identifying risk earlier, prioritizing remediation more effectively, and improving

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

enterprise governance across workforce platforms. Its strength lies in making hidden relationships visible, measurable, and actionable before they become larger control failures or security incidents.

9. Conclusion

This study presented a practical model for mapping the latent risk layer that exists across enterprise workforce platforms when employee data integrity, access behavior, permission state, integration outcomes, and cyber threat indicators are evaluated as connected signals. The central argument of the paper is that many enterprise risks do not begin as obvious incidents. They often emerge through small operational misalignments, such as delayed employee status updates, retained permissions after role changes, unusual access patterns, failed identity synchronization, incorrect manager assignments, or repeated workflow exceptions. When these signals are reviewed separately, they may appear minor. When they are analyzed together, they can reveal early evidence of larger control exposure.

The proposed model addresses this issue by treating workforce data as part of the enterprise risk environment rather than as a purely administrative record set. In SAP SuccessFactors and similar enterprise platforms, employee attributes influence approvals, access visibility, reporting structures, integrations, downstream provisioning, and audit evidence. Because of this dependency, data accuracy has direct consequences for security and governance. A missing field, incorrect effective date, delayed integration update, or outdated role assignment can affect how users access systems and how organizations maintain control over sensitive workforce information. This paper therefore reframes workforce data integrity as a measurable risk signal with operational, compliance, and cybersecurity relevance.

The study also showed that access behavior becomes more meaningful when interpreted with workforce context. Traditional permission reviews can explain what access exists, but they may not fully explain whether the access remains appropriate after job changes, department transfers, global assignments, terminations, or temporary support activities. By combining permission state with user behavior, the proposed model can identify risk conditions that static access reports may miss. This is especially important in large enterprise environments where role-based permissions, target populations, administrative access, workflow ownership, and integration dependencies can change frequently.

The comparative analysis demonstrated the practical value of combining rule-based controls with data-driven risk detection. Rule-based monitoring remains useful for known exceptions, mandatory field checks, inactive user access, workflow failures, and permission conflicts. However, it is less effective when risk depends on the relationship between several weak signals. The proposed latent risk model improved risk identification by correlating workforce data exceptions, access activity, permission drift, integration failures, and security alerts. The results showed stronger classification performance, lower false positives, faster detection, and better prioritization of high-risk events compared with isolated monitoring.

A key contribution of the research is its implementation-oriented design. The model does not depend on unrealistic system capabilities or speculative automation. It uses signals that organizations can reasonably obtain from enterprise platforms, audit logs, reporting tools, middleware, identity systems, and security monitoring environments. This makes the approach suitable for SAP SuccessFactors governance while also allowing cross-platform application in Workday, Oracle HCM, UKG, ServiceNow HRSD, identity governance systems, finance platforms, procurement systems, and other enterprise applications where employee attributes and access decisions are connected.

The findings also highlight the importance of explainability. Enterprise risk detection cannot rely only on technical scores. HR technology teams, security analysts, auditors, compliance reviewers, and system administrators need to understand why an event was classified as high risk and what action should follow.

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 2, Issue 6, November-December 2024
CrossRef DOI: 10.62127/aijmr.2024.v02i06.1366

A useful risk model should show whether the risk came from a data integrity issue, permission drift, suspicious access behavior, delayed integration processing, security alert, or a combination of these factors. This explanation supports faster remediation, clearer accountability, and better trust in model-assisted governance.

The practical value of the proposed approach lies in its ability to reduce hidden exposure before it becomes a visible incident. Organizations can use the latent risk layer to prioritize exception handling, strengthen access reviews, improve identity governance, reduce audit gaps, and support faster security escalation. Instead of reviewing separate lists of data errors, access exceptions, workflow failures, and cyber alerts, teams can work from a connected risk view that reflects how enterprise platforms actually operate. This can reduce manual review effort while improving attention to events that carry real business impact.

The study has limitations, particularly around data availability, labeling quality, organizational variation, and privacy controls. Not every company will have the same level of access to audit history, permission activity, security alerts, or integration records. Risk thresholds must also be calibrated to each organization's structure, business processes, user roles, and data sensitivity. Future research can strengthen this work by validating the model with production datasets across multiple industries, comparing performance across different enterprise platforms, and developing standardized feature sets for workforce-related enterprise risk detection.

Overall, this paper contributes a practical and measurable framework for identifying risk at the intersection of workforce data, access behavior, and cyber threat monitoring. Its main value is not only technical detection performance, but the shift in thinking it introduces. Workforce data problems, permission drift, integration failures, and security alerts should not be treated as isolated operational issues when they are part of the same enterprise control environment. Mapping the latent risk layer allows organizations to detect risk earlier, explain it more clearly, and respond with greater confidence before small weaknesses become larger failures.

REFERENCES:

1. 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
2. Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. ACM Computing Surveys, 41(3), 1–52. https://doi.org/10.1145/1541880.1541883
3. Even, A., & Shankaranarayanan, G. (2007). Utility-driven assessment of data quality. ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 38(2), 75–93. https://doi.org/10.1145/1240616.1240623
4. Khatri, V., & Brown, C. V. (2010). Designing data governance. Communications of the ACM, 53(1), 148–152. https://doi.org/10.1145/1629175.1629210
5. Ferraiolo, D. F., Sandhu, R., Gavrila, S., Kuhn, D. R., & Chandramouli, R. (2001). Proposed NIST standard for role-based access control. ACM Transactions on Information and System Security, 4(3), 224–274. https://doi.org/10.1145/501978.501980
6. Bertino, E., & Sandhu, R. (2005). Database security: Concepts, approaches, and challenges. IEEE Transactions on Dependable and Secure Computing, 2(1), 2–19. https://doi.org/10.1109/TDSC.2005.9
7. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 1–58. https://doi.org/10.1145/1541880.1541882

8. Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19–31. https://doi.org/10.1016/j.jnca.2015.11.016
9. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502
10. Gheyas, I. A., & Abdallah, A. E. (2016). Detection and prediction of insider threats to cyber security: A systematic literature review and meta-analysis. Big Data Analytics, 1, Article 6. https://doi.org/10.1186/s41044-016-0006-0
11. D'Arcy, J., Hovav, A., & Galletta, D. (2009). User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach. Information Systems Research, 20(1), 79–98. https://doi.org/10.1287/isre.1070.0160
12. Bulgurcu, B., Cavusoglu, H., & Benbasat, I. (2010). Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. MIS Quarterly, 34(3), 523–548. https://doi.org/10.2307/25750690
13. Siponen, M., Mahmood, M. A., & Pahnila, S. (2014). Employees' adherence to information security policies: An exploratory field study. Information & Management, 51(2), 217–224. https://doi.org/10.1016/j.im.2013.08.006
14. Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). Isolation forest. Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, 413–422. https://doi.org/10.1109/ICDM.2008.17
15. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324
16. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785
17. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778
18. Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9
19. Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR analytics. The International Journal of Human Resource Management, 28(1), 3–26. https://doi.org/10.1080/09585192.2016.1244699
20. Tursunbayeva, A., Di Lauro, S., & Pagliari, C. (2018). People analytics: A scoping review of conceptual boundaries and value propositions. International Journal of Information Management, 43, 224–247. https://doi.org/10.1016/j.ijinfomgt.2018.08.002
21. Strong, D. M., & Volkoff, O. (2010). Understanding organization-enterprise system fit: A path to theorizing the information technology artifact. MIS Quarterly, 34(4), 731–756. https://doi.org/10.2307/25750703
22. Shaul, L., & Tauber, D. (2013). Critical success factors in enterprise resource planning systems: Review of the last decade. ACM Computing Surveys, 45(4), Article 55. https://doi.org/10.1145/2501654.2501669