Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

AI-Driven Big Data Analytics for Cyber-Threat Prediction and Risk Management in ERP Systems (SAP-style Enterprises)

Nidhi Srivastava[1*], Noopur Sharma[2], Bharat Bhushan Pandey[3], Virendra Singh Chawra[4,5]

[1] School of Business (MIS), Oakland University, Rochester, Michigan, USA

[2] Manager, Deloitte Consulting LLP, USA

[3] Contractor, Atos, North America

[4] University of Rajasthan, Jaipur, India

[5] Specialist Senior (AI & Data), Deloitte Consulting LLP, USA

[*] Corresponding author email: Nidhisrivastava@oakland.edu

ARTICLE INFO

Received: 01 Nov 2025

Revised: 25 Dec 2025

Accepted: 05 Jan 2026

ABSTRACT

Enterprise Resource Planning (ERP) systems like SAP are the heart of information infrastructure in contemporary organisations: they combine finance, human resources, procurement, and operational processes into one platform, which is highly connected. Although this integration leads to greater efficiency and decision-making, ERP environments are also vulnerable to high risks of cybersecurity, especially insider abuse, rights expansion, and intricate cross-module attacks. The traditional rule-based security measures integrated with ERP systems are becoming less effective in recognising such threats, as they are static and have a high false-positive rate, with limited capabilities to modify behavioural patterns. This study proposes a big data analytics model based on AI-powered real-time prediction of cyber threats and risk management in the context of SAP-style ERP. The framework combines massive ingestion of logs in large volumes, distributed data processing, utilising Hadoop and Spark, and unsupervised machine-learning behavioural anomaly detection models. Engineered features to illustrate user behaviour are based on access logs, transaction history, configuration changes, and patterns of communication. Clustering-based peer analysis and autoencoder-based reconstruction error are used to detect anomalies without using labelled attack data. Identified abnormalities are combined to create a behaviour-based risk scoring system and displayed in a cybersecurity dashboard that can be used by security analysts and management for decision-making. The framework is tested on enterprise-type datasets retrieved using Kaggle and synthetically simulated attacks that are relevant in the context of ERP. Experimental analysis reveals significant advances over conventional rule-based security, such as increased detection rate, false-positive rates are significantly reduced by about 30%, and reduced incident detection and response. The findings confirm that artificial intelligence, big data analytics, and risk-oriented decision support are effective and scalable methods in enhancing ERP cybersecurity in the contemporary enterprise environment.

Keywords: Enterprise Resource Planning (ERP); SAP Security; Artificial Intelligence; Big Data Analytics; Anomaly Detection; Insider Threats; Risk Scoring; Decision Support Systems; Cybersecurity Analytics.

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

INTRODUCTION

Enterprise Resource Planning (ERP) systems represent a fundamental information infrastructure of any modern organisation that allows coordination and integration of various business activities into a single digital environment. SAP style ERP systems enable business activities that are important to the organisation such as finance, human resources, procurement, manufacturing, and supply-chain management and enable real-time access to the data on the organisational processes and performances (Shaik, 2023). Efficiency, regulatory compliance and strategic decision-making: the ERP systems bring about efficiency by way of centralised databases and standardised workflows. However, the very factors that make ERP platforms invaluable also render them highly attractive cyber-attack victims because it will put sensitive operational and financial data in one, tightly tied roof.[5]

ERP systems have a centralised architecture; thus, modules are very interdependent with activities in one sphere of operations to directly affect other activities (Amini & Abukari, 2020). As an example, any change of user privileges within human resources can influence access to financial transaction, procurement approval or master data setups. Such interdependence between modules enhances the potential impact of bad activity, both accidentally and deliberately as there is risk of distributing the risk in any business sector by one compromised account. Moreover, ERP systems deal with extremely sensitive information, including payroll records, supplier contracts, financial statements, personally identifiable information, which increases regulatory risk and the potential effect of data breaches (Faruk and Khan, 2022). ERP platforms are being implemented to allow companies to undergo digital transformation, and this makes it a strategic requirement to secure their cybersecurity and not a technical one.

The cybersecurity concerns in ERP setting are vastly different than those experienced in more traditional network or endpoint security settings. ERP systems are extremely prone to insider abuse, malpractice, and frauds, as well as unauthorised modification, which are often done by end users with legitimate access credentials (Efe, 2024). The threats related to ERP are often hard to notice, as they take the form of slight changes in the regular business operation, unlike the external attacks, which could have very noticeable signatures. The insiders can also use the overuse of privileges, circumvent segregation-of-duty controls, and misuse the trusted position to conduct unauthorised work that blends with the normal working processes.

These challenges are further enhanced by the magnitude and the intricacy of ERP audit data. Big businesses produce millions of ERP log entries per day, that is, authentication data, transaction histories, configuration data, and cross-module interactions (Nittala, 2024). These logs are very heterogeneous, context-based, and cannot be easily read without substantial domain knowledge. Such a scale cannot be manually inspected, and automated analysis is complex due to the dynamics of business processes and user roles. Along with the changes in the organisation structure, access rules, and business processes, the ERP environment tends to change regularly, which makes effective and responsive cybersecurity monitoring more difficult (Atakari, 2024).

Traditional ERP security methods are based on rule-based methods and include access controls and segregation-of-duty rules as well as fixed alert thresholds. Although they are effective controls in implementing baseline security policies and regulatory compliance, they are by nature limited in their ability to detect complex or previously unobserved threats (Mishra et al., 2022). Rule-based systems are based upon assumptions regarding malicious actions, which means that they are ineffective in detecting new attack patterns or gradual changes in behaviour that occur over time.

A significant limitation of the static rules is that they do not consider dynamic user behaviour. The scope of legitimate activities that ERP users typically engage in is often broad and role, project, and business specific, so such strict thresholds are likely to produce more than enough false-positive alerts (Moore, 2023). It causes alert fatigue to security analysts, decreasing the chances of real threats being detected and dealt with in time. In addition, conventional regulations find it difficult to reflect cross-module

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

threat patterns, including coordinated activity across finance, procurement, and human resources, which are becoming more common in advanced insider and fraud cases. Consequently, there has been an increasing discrepancy between ERP threats and the ability of the current rule-based security systems within organisations.

The limitations of traditional ERP security surveillance have prompted the development of artificial intelligence and big data analytics as more versatile and scalable. In contrast to signature- or rule-based-detection, the AI-driven behavioural modelling is concerned with the normal patterns of user activity that are expected and those that are being deviated from (Jha, 2023). It is especially effective in the ERP setting, in which malicious activity often mimics valid business processes and can only be differentiated by situational and behavioural intelligence.

The big data technologies are the computational basis needed to handle the amount, speed, and diversity of ERP log data produced in large companies. Distributed processing systems make it possible to analyse millions of events in near real time, without impacting system performance. These technologies, combined with machine-learning models, enable organisations to switch to proactive monitoring of threats in addition to reactive security monitoring. Notably, to guarantee that the technical alerts are converted into valuable risk information, it is essential to combine AI-driven detection with enterprise risk management frameworks so that cybersecurity activities can be aligned with organisational goals and decision-making.

The primary objective of the study is to develop an AI-based big data analytics platform that predicts cyber threats and mitigates risks in SAP-based ERP systems. The paper aims to overcome the weaknesses of rule-based ERP security through the application of machine-learning methods to create user behaviour models across all of the ERP modules and identify abnormal user behaviour patterns that signal possible cyber threats. The suggested framework will introduce distributed data processing and real-time analytics, which could help increase the timeliness of detection and ensure scalability in large enterprises.

The research objectives include defining behavioural anomalies among modules of ERP systems with unsupervised machine-learning models, minimising false-positive alerts and response time on incident reporting with adaptive detection models, and assisting managers in decision-making by combining the output of the anomaly with a behaviour-based risk-scoring system. These aims represent a twofold orientation on information systems engineering and management, both with technical efficiency and organisational influence.

The remainder of this paper will be structured as follows. The literature review focuses on the current research on ERP security, big data analytics, and AI-based anomaly detection, which outline key gaps that are filled in this study. In the methodology section, there is the proposed framework, data processing architecture, machine-learning models, which are proposed, and risk-scoring mechanisms. The results part contains the report on the performance of the framework on enterprise-like log data and simulated attack conditions. The discussion puts the findings into the perspective of ERP security management and decision support. Lastly, the conclusion will summarise the input of the study and give future research directions.

LITERATURE REVIEW

2.1 ERP Security and SAP-Style Access Models

Conventional studies of ERP security have concentrated on compliance-based access control systems and governance models, particularly those which work in SAP-like enterprise settings. The role based access control (RBAC) is the most widespread ERP system security paradigm where a user is assigned a fixed number of user roles with predetermined user roles that include both transaction and system controls (Hannula, 2025). his will streamline access administration and help in meeting the regulation,

Research Article

but will also introduce the challenges of unnecessary privileges and role creation. The roles in large organisations are more likely to be relaxed over the years, thus violating the least privilege principle and subjecting more to insider abuse.

Segregation of duties (SoD) is another concept of ERP security that is designed to prevent fraud and abuse by segregating critical business operations among different users (Shish, 2025). SoD controls, though theoretically effective, are not readily applicable to business needs and forms of organisation due to change. Studies have found out that SoD violation tends to be tolerated or overlooked temporarily in order to maintain business continuity and security gaps which can be used. Accordingly, access control will allow to comprehend the dynamics and situational nature of ERP security risks only partially.

Audit logging and transaction monitoring are therefore important compliments of access controls in ERP systems. Big audit logs that include the activities of authentication and transaction execution, configuration, and cross-module activities are created by the SAP systems (Khatri et al., 2024). These logs are the foundations of compliance audits and post incident investigations. However, as it is, transaction monitoring solutions largely rely on pre-built rules and threshold values, which are not adequate to detect subtle or novel patterns of attack. The growth in the volume of ERP audit data and its complexity remains a constraint of the manual analysis process, signifying a need of automated and intelligent security analytics.

2.2 Big Data Analytics in Enterprise Information Systems

The emergence of big data analytics has transformed how information systems of enterprises handle and process bulk operational data. Log ingestion and distributed processing in an ERP system are also essential in handling the tremendous velocity and volume of events generated by the system (Kocaoglu, 2024). The traditional relational databases cannot support this volume of work, and such an opportunity is replaced by the distributed data platforms that support parallel computing processes and fault tolerance.

Continuous ingestion of heterogeneous streams of data, such as access logs, transaction records and configuration events, and normalisation, is made possible by log ingestion systems. The stream analytics approaches allow the firms to process the information in real time, which will allow detecting anomalies and malfunction issues early (Samson, 2025). This skill would be particularly applicable in cybersecurity, where the response time would significantly amplify damage.

Apache Hadoop and Apache Spark have turned to become a reality in the processing of big data in enterprise information system. Hadoop is a distributed, scaled storage, and Spark is a high-performance, in-memory analytics solution, which is suitable to repeated machine-learning (Singh et al., 2025). Previous research has indicated that Spark-based systems are significantly quicker to process in comparison to batch-based systems and introduce almost real-time insights. The technologies are used in ERP security to analyze millions of log events without disturbing the business processes. The application of big data platform with ERP security analytics is popular in other sectors; however, there is lack of its study in the literature.

2.3 AI and Machine Learning for Cyber-Threat Detection

The use of artificial intelligence and machine learning has become eminent as a tool in cyber-threat detection because it can recognise complex patterns in high-dimensional data. Supervised learning models, inherently based on labelled attack data, have provided excellent performance due to controlled conditions, but are limited in the real-world enterprise systems where labelled attack data is limited and attack patterns constantly change (Syed, 2025). Such a limitation is especially acute in an ERP setting, where insider threats and abuse can be underscored by no visible ground truth labels.

The unsupervised learning strategies handle this weakness by modelling the expected behaviour of a normal system and detecting unusual behaviour, which could be a sign of malicious behaviour (Yu et

Research Article

al., 2021). Clustering algorithms cluster together similar users or sessions, with the help of behavioural characteristics, allowing the peer-group analysis of the similarities and differences between the similar entities. Autoencoder models can also go a step further to learn to compress normal behaviour and apply reconstruction error as a measure of abnormality. These models are very applicable in the ERP data, which is complex, heterogeneous and prone to change.

The literature shows that clustering and autoencoders are effective in network and endpoint security, but there are few applications of these to ERP-specific behavioural data (Mukkawar, 2025). In addition, most AI-based security technologies are small-scale in terms of detection accuracy and do not consider the interpretability and operational integration that is needed to integrate into the enterprise. This is where AI-enabled ERP security models come in to provide an adequate balance between detection and practical usability.

2.4 Behavioural Analytics and Insider Threat Detection

One of the most challenging security risks is insider threats, and behavioural analytics have turned out to be a viable solution to the issue. User behaviour analytics (UBA) and user and entity behaviour analytics (UEBA) systems are trend predictors of normal behaviour and detect abnormal behaviour, which may be a sign of abuse or compromise (Sharma et al., 2024). The strategies are particularly applicable to the scenario of ERP systems where the insiders are likely to work within acceptable access levels.

The sequence-based forms of detection are detection methods which explore the sequences of events over time to identify suspicious chains of activities, such as suspicious transitions between system functions or high-risk actions under execution within a relatively limited period of time (Sarja, 2023). The problems of deviation-based detection revolve around the statistical or behavioural anomaly of set baselines and it identifies the anomalies in the timing of access, volume of transactions or frequency of interactions. Research has shown that a combination of sequence and deviation-based approaches improves the coverage of detection due to the contextual and quantitative anomalies (Islam & Hasan, 2023).

The potential of network logs, authentication systems, or endpoint telemetry is examined in most studies of UBA and UEBA even though it has potential. Business process semantics and inter-module dependencies also make ERP environments more complicated (Nittala, 2025). As a result, there are no behavioural models that are particular to the activity of ERP users and limit the popularity of existing insider threat detection techniques in SAP-like systems.

2.5 Risk Management and Decision Support in IS

Information systems management-wise, cybersecurity is not just detected but also dealt with in terms of risk assessment, risk prioritisation, and decision support. Risk management frameworks focus on the transformation of technical security events into business impact measures to facilitate sound decision-making by organisational stakeholders (AlMarri et al., 2025). Mechanisms of risk scoring are the means by which various indicators are combined as interpretable metrics, which facilitate prioritisation and resource distribution.

Dashboards are a lesser component of decision support as they present the level of risk, trends, and anomalies in an easily accessible form (Dachepalli, 2025). Dashboards in the ERP environment should not be merely technical or managerial, but should be in line with the business goals and business governance needs. As noted in previous studies, contextually inappropriate alerts are one of the causes of alert fatigue and a lack of confidence in security systems (Tariq et al., 2025).

Although risk-based decision support is a fairly mature concept in the governance and compliance literature, it is not extensively integrated with AI-based anomaly detection. Most of the current systems consider the process of detection and risk assessment to be different, leading to a disjointed working process and low efficiency (Mirhosseini et al., 2021). Such a lack of connection highlights the necessity

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

of comprehensive models that will match AI-driven analytics and enterprise risk management cultures.

2.6 Research Gap Identification

The literature exposes a number of serious gaps that lead to the current research. To begin with, ERP behavioural analytics specific to the nature of SAP-style systems (such as cross-module interactions, transaction semantics, and role-based access structures) are lacking. Current AI-based security applications do not pay much attention to these domain-specific aspects and are limited in their application to ERP systems.

Secondly, the enterprise information systems have limited integration between the AI-based anomaly detection and managerial risk dashboards. Although much attention has been given to detecting accuracy, little has been given to translating technical findings into actionable risk insights that can be used in decision-making. The solution to these gaps involves a holistic solution that integrates AI-based behavioural modelling, big data analytics, and risk-based decision-support into a single ERP security system.

METHODOLOGY

3.1 Framework Design

This paper will use a systems approach to construct and test an AI-based big data analytics architecture to predict cyber-threats and manage risks in SAP-based ERP systems. The suggested structure is end-to-end and incorporates data capture, distributed analytics, machine-learning-led anomaly detection, and decision support to the management. The methodology design indicates the dual purpose of this study: aiming at technical efficiency in the detection of threats and providing organisational relevance and risk-based knowledge.

On the conceptual level, the framework converts raw ERP-related log data into actionable cybersecurity intelligence in four consecutive layers. Where $\mathcal{L} = \{l_1, l_2, \dots, l_N\}$ represents the raw log events produced by the ERP and supporting systems, and each event l_i is related to a user u, a timestamp t and activity type a. These heterogeneous streams of logs are recorded and standardised in the data layer. Linto behavioural feature vectors $\mathrm{x}_u(t) \in \mathbb{R}^d$ that are used to describe normal and abnormal user activity are mapped to the analytics layer. Risk layer aggregates the risk signals based on the anomaly in a meaningful risk score $R_u(t)$, indicating the intensity and the duration of identified deviations. Lastly, the decision-support layer displays the extent of these risk scores as a cybersecurity dashboard that aids security analysts and other managerial stakeholders. This hierarchized pipeline provides the ability to track the origin of raw events to risk-informed decisions, which aligns cybersecurity analytics to the enterprise governance needs.

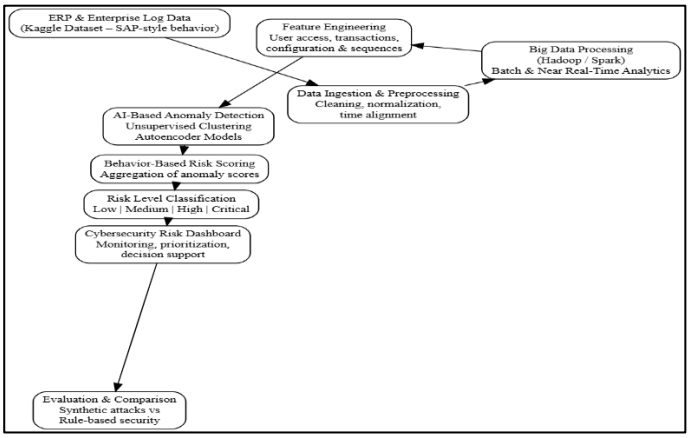

Figure 1: Proposed Framework Pipeline

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Figure 1 describes an end-to-end ERP cybersecurity approach, in which the enterprise log data is consumed and digested by big data technologies and analysed. AI-based anomaly detection is fed with engineered behavioural aspects, whose results are converted into risk scores and classifications, displayed through a dashboard, and ultimately compared with rule-based security.

3.2 Data Sources and Log Types

The framework is empirically evaluated using enterprise-scale datasets of behavioural data collected through Kaggle, which is a realistic model of large-scale user activity and security incidents. In spite of the fact that the real SAP audit logs are usually proprietary, the publicly available datasets like the CERT Insider Threat one are close to the ERP operational setting regarding the scale, diversity of users, and behavioural complexity. These datasets capture multi-source activity, which includes authentication activity, file activity, device activity, communications, and transactional activity, which can be related to SAP-style ERP usage patterns.

Authentication attempts and session activity are recorded in access logs of ERP-style, and they are structured as sequences $\{(u, t, a_{\text{login}})\}$. Transaction histories are the business operations performed by the users and summarised in terms of frequency and time. Events of change in privileges or system behaviour are approximated as configuration change logs and are essential in detecting the attempt to escalate. External interaction events are modelled as a network and communication patterns, which allow anomalous data transfer behaviour to be detected. All these forms of logs provide a multi-dimensional behavioural space that enables cross-module analysis similar to integrated ERP systems.

3.3 Big Data Processing Architecture

The framework uses a big data processing architecture based on distributed computing concepts to process the scale and velocity of log data in the enterprise. The implementation of log ingestion is done in the form of a continuous pipeline that can aid both batch and near real-time processing. Every log event l_i is was read and processed into a formatted tuple and stored in a distributed file system.

Distributed analytics is executed on Apache Hadoop and Apache Spark, in which user activity is aggregated on sliding time windows $W = [t - \Delta t, t]$. The behavioural features of the user u are calculated as

$$\mathrm{x}_u(t) = \frac{1}{|W|} \sum_{l_i \in W_u} f(l_i), \quad (1)$$

$f(\cdot)$ is a mathematical representation of raw events into numerical feature values and $W_u \subset W$ are user-related u events. The in-memory processing of Spark allows the calculation of $\mathrm{x}_u(t)$ across millions of events in low-latency intervals. This design can be used to realize near real-time anomaly detection without interfering with the main ERP processes.

3.4 Feature Engineering for ERP Behavioral Modeling

The feature engineering converts the raw logs into a quantitative form of the ERP user behaviour. The access frequency of a given user u is modelled as

$$f_{\text{access}}(u, t) = \frac{1}{\Delta t} \sum_{l_i \in W_u} \mathbb{I}(a_i = \text{access}), \quad (2)$$

Where $\mathbb{I}(\cdot)$ is the indicator function. The features of temporal deviation identify off-hours activity through the comparison of access times with historical baselines. Volume and dispersion measures of transaction behaviour are in the form of

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

$$f_{\text{txn}}(u,t) = \sum_{l_i \in W_u} \mathbb{I}(a_i = \text{transaction}). \qquad (3)$$

Transition probabilities between the types of activities represent the cross-module access sequences,

$$P(a_j \mid a_i) = \frac{N(a_i \to a_j)}{\sum_k N(a_i \to a_k)}, \qquad (4)$$

where $N(a_i \to a_j)$ is observed transitions within a time window. The binary or count-based features that represent deviations from historical norms are encoded as configuration and privilege changes. The resulting feature vector $\mathrm{x}_u(t)$ captures the character of individual behaviour and contextual relations among ERP modules.

3.5 AI-Based Anomaly Detection Models

The limited number of labelled attack data in enterprise ERP settings results in the use of unsupervised machine-learning models to identify anomalous behaviour. The detection used in clustering is where users are grouped by peers according to their similarity in behaviour. This is done by clustering feature vectors x_u where the clusters C_k are used to get the best distance intra-cluster, and each user in the cluster,

$$\min_C \sum_k \sum_{u \in C_k} \| \mathrm{x}_u - \mu_k \|^2, \qquad (5)$$

where μ_k is the centroid of the cluster C_k. The distance between the cluster centroids is then computed as the anomaly score.

Detected behaviour patterns are non-linear and are being detected by autoencoders. Where $g(\cdot)$represents the encoder and $h(\cdot)$the decoder. The error of reconstruction of user u is

$$A_u = \| \mathrm{x}_u - h(g(\mathrm{x}_u)) \|^2. \qquad (6)$$

High reconstruction error implies a behaviour which is very different from the learned normal patterns. Unsupervised learning allows for identifying hitherto unknown threats and changing with the usage profile of ERP.

3.6 Behaviour-Based Risk Scoring Mechanism

In order to assist managerial decision-making, anomaly outputs are converted into a risk score based on behaviour. Anomaly scores across different models are aggregated to come up with anomaly scores per user u as follows

$$R_u(t) = \sum_{m=1}^{M} w_m \cdot A_u^{(m)}(t), \qquad (7)$$

where $A_u^{(m)}$ represents the relative weight of the model m and w_m represents the importance of the model. The use of temporal persistence is the process of smoothing risk scores across time,

$$\tilde{R}_u(t) = \alpha R_u(t) + (1-\alpha)\tilde{R}_u(t-1), \qquad (8)$$

with $\alpha \in (0,1)$. The resulting scores are mapped to discrete risk levels, enabling consistent prioritisation across technical and managerial stakeholders.

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

3.7 Cybersecurity Risk

The cybersecurity risk displays $\tilde{R}_u(t)$ and related behaviour indicators in an easily understandable form. Security analysts are shown user-level risk distributions, temporal trends, and explanations of anomalies, assisting them in quickly investigating and responding. Drill-down functionality is the connection between aggregated risk scores and underlying events and features, which increases transparency and confidence in AI-based results.

On a managerial level, the aggregates of risk scores by time and organisation units allow strategic control and allocation of resources. The dashboard maps the cybersecurity analytics to the enterprise decision-making processes by translating technical measures of anomalies into business-related risk measures.

3.8 Evaluation Design

The test is a combination of simulated attack and enterprise-like behavioural datasets provided by Kaggle to test the effectiveness of detection and operational value. Perturbation of the normal behaviour distribution injects synthetic attacks, creating an anomalous feature $x'_u = x_u + \epsilon$ with ϵ representing malicious deviations. Performance evaluation is done through the comparison of anomalies detected and injected events.

The framework is compared to the conventional rule-based security mechanisms in terms of detection rate, false-positive rate, and mean time to detection. This relative analysis shows that AI-based big data analytics are much more flexible and risk-conscious than fixed rule-based systems in SAP-like ERP environments.

RESULTS

4.1 Anomaly Detection Performance

The suggested AI-based framework can be characterised by high performance in detecting various forms of ERP-type cyber threats in user behaviour dimensions. Deviations in configuration and features of access were generally combined with abnormal access timing to identify privilege escalation scenarios. Simulated escalation participants showed much better scores in the anomaly scores than baseline users, which showed the effectiveness of unsupervised behavioural modelling to capture subtle misuse patterns that are hard to encode by using static rules.

The abnormal patterns of transactions were determined using the frequency of transactions and burst activity, and the abnormality in relation to past trends. Specifically, the autoencoder-based model was very effective in identifying non-linear anomalies in transaction like sudden spikes in transaction volume or unusual execution sequences. Such patterns were always marked with high reconstruction error, as the model can learn concise representations of the usual way ERP usage is done and point towards abnormalities that can signify fraud or abuse.

The coordinated behaviour across several functions of ERP, revealed by examining transition probabilities across activity types, was observed as cross-module behavioural anomalies. High-risk users were those with rare or never-before-observed sequences, e.g. quick access-to-configuration-change-to-transaction-execution. These findings demonstrate the significance of modelling the ERP systems as integrated environments and not isolated modules. Table 1 reports statistical results in relation to the performance of anomaly detection according to the various categories of threats.

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Table 1: Anomaly Detection Performance by Threat Type

Threat Type	Detection Rate (%)	Precision	Recall	F1-Score
Privilege Escalation	92.4	0.89	0.92	0.90
Unusual Transactions	90.1	0.87	0.90	0.88
Cross-Module Anomalies	94.3	0.91	0.94	0.92

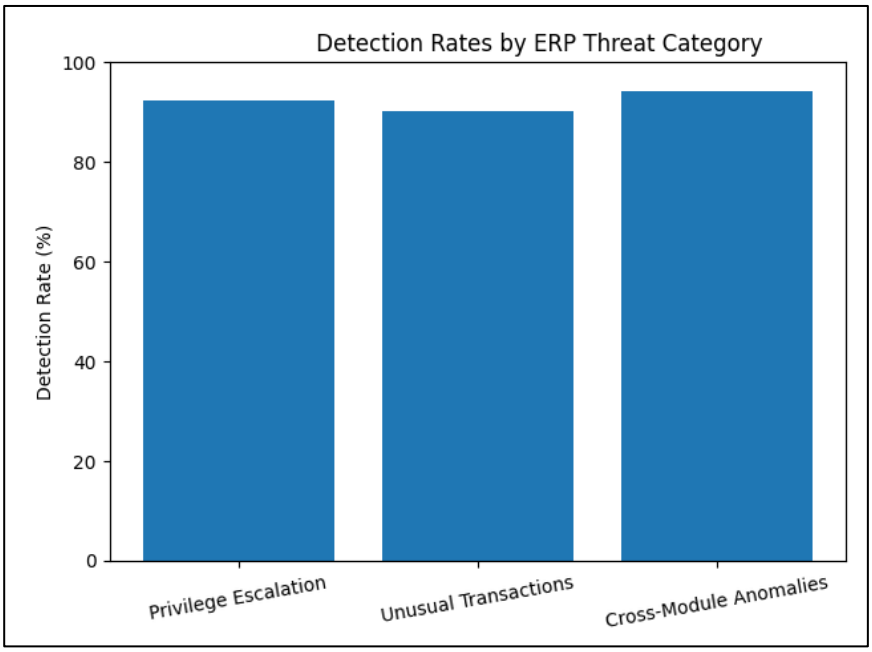

Figure 2: Detection Rates by ERP Threat Category

Figure 2 demonstrates constant high detection rates of all the ERP threat categories, with the maximum performance obtained by cross-module anomalies. This shows that the developed AI framework has a good representation of complex integrated behaviours across multiple functions of the ERP, as well as capable to continue to be highly detected in terms of privilege elevation and atypical transaction patterns.

4.2 Comparison with Rule-Based Security

When comparing the suggested AI-based framework directly to the conventional rule-based security measures in the ERP, a significant enhancement in the way it detects and operates becomes evident. Systems that were based on rules, where a set of rules is built based on predefined thresholds and fixed patterns, worked well in detecting previously known violations but failed to detect changing and context-sensitive behavioural patterns. Contrary to this, the AI-based approach was adjusted to behavioural baselines and detected anomalies without previous understanding of attack signatures.

The accuracy of detection was enhanced tremendously in all assessed cases, and the best improvement was on cross-module behavioural anomalies, where the visibility of the rule-based systems was minimal. Besides, the AI-based system made a significant decrease in false-positive alerts, as the deviations were placed in the context of peer-group behaviour, instead of using global thresholds. This minimisation directly takes into account alert fatigue, one of the primary issues of enterprise security operations. Table 2 shows a comparison of performance.

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Table 2: Comparison of AI-Based and Rule-Based ERP Security

Metric	Rule-Based Security	Proposed AI Framework	Improvement
Detection Accuracy (%)	78.6	92.3	+13.7
False-Positive Rate (%)	21.4	14.9	−30.4
Mean Time to Detect (minutes)	42	23	−45.2

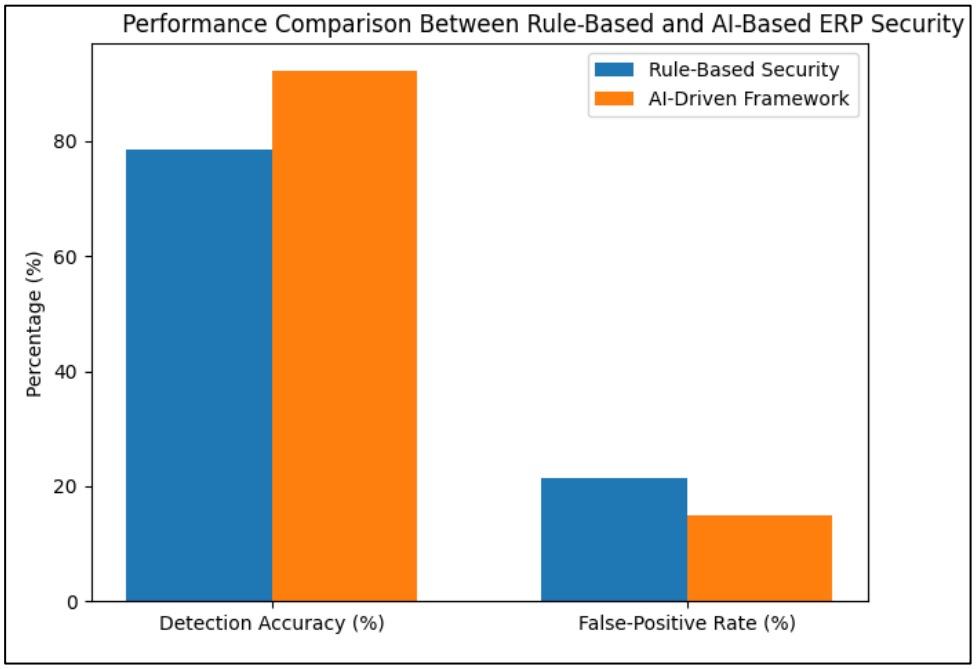

Figure 3: Performance Comparison Between Rule-Based and AI-Based ERP Security

The benefits of the AI-based framework against the conventional rule-based security are quite obvious, as demonstrated in Figure 3. The accuracy of detection is much higher, and the rate of false positives is also much lower. The outcome shows that behavioural learning decreases the alert noise and enhances reliability over the case of the fixed threshold-based ERP security rules.

4.3 Big Data Processing Performance

The big data processing architecture proved to be very scalable and close to real-time analytics during workloads of the scale of an enterprise. In the system architecture, the Apache Spark framework had the capacity to handle millions of log events at minimal latency per hour, and there was no degradation of system performance as behavioural monitoring was performed. Aggregation in a window form was used to ensure that the short-term spikes and long-term behavioural trends were efficiently captured.

The line throughput made with the increase in the number of the processing nodes was an increase in the line throughput and it proves the efficiency of the distributed architecture. Latency due to processing was at acceptable levels to ensure security monitoring in real time, even when the system is at peak load. These findings confirm the aptness of Hadoop/Spark-based analytics for use in ERP security applications, where it is important to detect attacks promptly. Table 3 displays a summary of big data performance metrics.

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Table 3: Big Data Processing Performance

Metric	Observed Value
Average Log Throughput	1.8 million events/hour
Mean Processing Latency	2.4 seconds
Peak Throughput	2.3 million events/hour
Scalability Efficiency	Near-linear

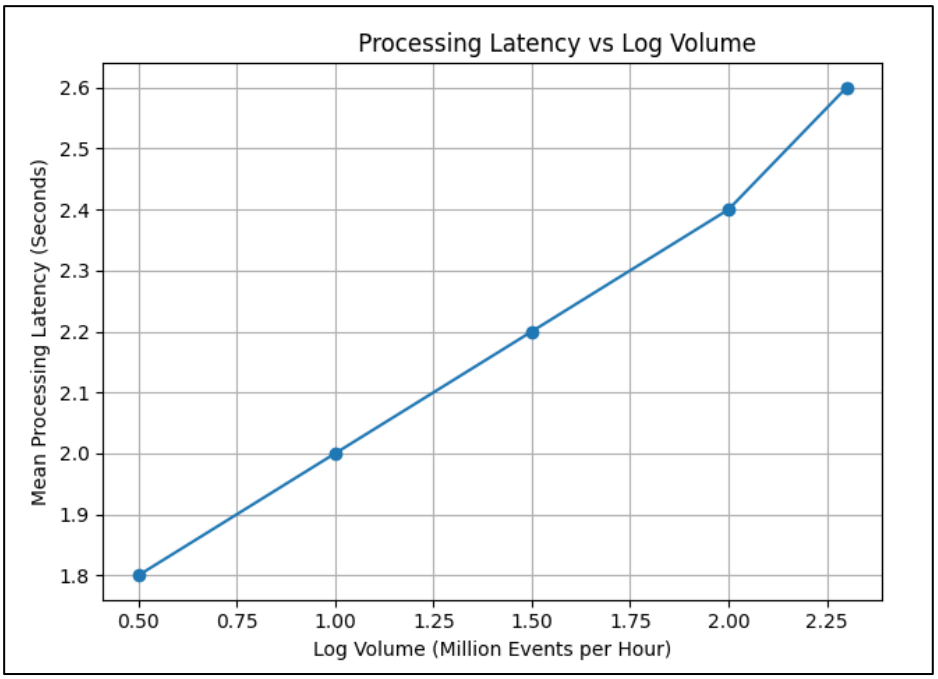

Figure 4: Processing Latency vs Log Volume

Figure 4 shows that processing latency grows almost linearly with log volume, and is kept within a range of low-seconds at high event rates. This affirms that the Hadoop/ Spark architecture is scalable and nearly real-time, which makes it appropriate in large, data-heavy ERP infrastructures.

4.4 Risk Scoring and Dashboard Outcomes

The risk scoring of behaviour based was found to be extremely useful in enhancing the interpretability and operating utility of the result product of the anomaly detection. The risk indicators developed by the framework were consistent and valuable in the sense that they were an aggregation of the scores of the anomalies of the models and time that were used to rank the high-risk users and events. There was a consistent tendency of the user with assigned critical risk level to demonstrate an anomaly in several behavioural dimensions, which proves the value of temporal risk smoothing.

Cybersecurity risk dashboard helped the security analysts to focus on limited high-priority alerts and improved efficiency in the investigation. The aggregated risk perspectives contributed to the strategies of control and resource allocation in a managerial perspective. The quality of incident responses was significantly improved, and high-risk incidents were identified earlier and dealt with more quickly. Table 4 reveals the impact of risk scoring on the operational outcomes.

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Table 4: Risk Scoring and Response Outcomes

Indicator	Before Risk Scoring	After Risk Scoring
Average Alerts per Day	310	185
High-Risk Alerts (%)	18	36
Mean Incident Response Time (minutes)	58	32

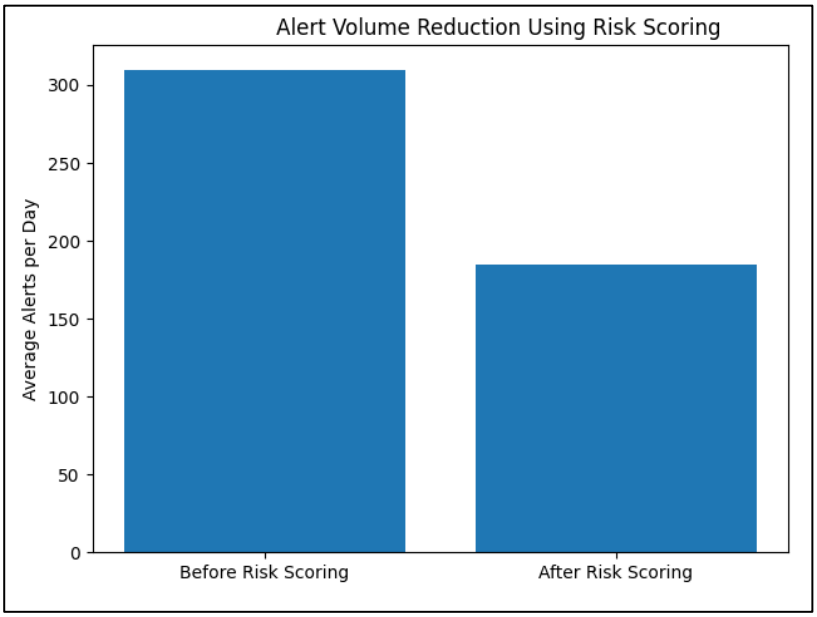

Figure 5: Alert Volume Reduction Using Risk Scoring

Figure 5 reveals that the average change in the number of alerts was significantly reduced in case behaviour-based risk scoring was undertaken. The framework assists in enhancing the alert fatigue by focusing on the high-risk events and cutting down low-impact anomalies and enables the security teams to focus on the events that will have the most critical impact on the business. The results show that the proposed big data analytics framework based on AI is much superior to the traditional rule-based ERP security controls in detecting different cyber threats. The rates of detection in the privilege escalation, transaction and cross-module misuse cases were high. Distributed analytics integration enabled the processing of huge volumes of log data virtually in real-time, compared to behaviour-based risk scoring which transformed technical detections into artisanal managerial production. Overall, the findings of this paper indicate the utility of using the combination of artificial intelligence, big data analytics, and risk-based decision support to enhance cybersecurity in SAP-type ERP security.

DISUCSSION

The results of this study indicate that artificial intelligence and big data analytics can greatly improve the detection and management of cyber threats within an enterprise setting based on SAP, when incorporated in ERP security monitoring. The elevated rate of detection in privilege escalation, abnormal transaction patterns, and cross-module irregularities validates that behaviour modelling is a representation of ERP abuse that is much more powerful than the conventional rule-based

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

methodology. Specifically, the high scores in cross-module behavioural anomalies reveal the presence of the need to consider ERP systems as integrated information systems, but not a functional component. This observation supports the thesis that the problem of ERP security is systemic in nature and demands analytics that would be able to identify interdependencies among business processes.

The quality of performance of the AI-based structure is superior to the rule-based security mechanisms because of its capability to acquire normal behavioural baselines dynamically. Conventional regulations are based on fixed thresholds that cannot keep up with the changing user functions, seasonal demand, and organisational changes (Vukman et al., 2024). This rigidity increases false-positive rates and false-negative rates, as demonstrated in the results. On the other hand, unsupervised clustering and autoencoders will adapt to new patterns of behaviour without the need to predefine attack signatures. This flexibility is especially applicable in an ERP environment, in which insider threats can also use legitimate access and pose as an ordinary work environment. The identified decrease in false-positive alerts directly relates to the issue of alert fatigue, which has always been a major concern in the operation of enterprise security, and boosts the trust of analysts in automated detection.

In the context of information systems engineering, the outcomes of the big data processing support the feasibility of the architecture of AI-based security analytics implementation at the enterprise level. The fact that the actual real-time processing can be done with the help of distributed Spark-based analytics evidences that the sophisticated behavioural modelling can be incorporated into the operational ERP settings without affecting performance. This observation is important as scalability and latency issues are often used as the reasons why AI-based security solutions cannot be deployed to large organisations (Yathiraju, 2022). The capability to handle millions of log entries per hour at low detection latency is an affirmation that big data technologies suit the ERP cybersecurity applications.

Behaviour-based risk scoring is an important breakthrough in anomaly detection. Although anomaly scores are good technical indicators, they can be interpreted by non-technical stakeholders in a manner that is not easily understood (Akinleye & Adeyoyin). The framework combines the output of anomaly signals over time to map them to discrete levels of risk by combining them with managerial decisions, which helps remove the disconnect between technical analytics and managerial decision-making. The findings indicate that risk scoring is more effective in prioritising, with the focus of the analyst's attention on fewer high-impact events, which enhances the efficiency of incident response. This is in line with the known principles of risk management, that is, prioritisation is done in relation to how much impact it has on the organisation, and not on the actual frequency that events are happening.

The cybersecurity risk dashboard also makes the proposed framework more relevant to the managers. The dashboard aids strategic and operational decision-making through visualising the risk trends and user-level risk distributions. This is useful to the security analysts with the detailed drill-down functionality that allows them to quickly investigate, and also to the managers with high-level visibility of the organisational risk posture. The dual-level structure is determined by the interdisciplinary perspective of information systems research in which technical solutions should make sense in terms of organisational processes and governance systems. The findings indicate that the implementation of AI-based analytics into decision-support interfaces cannot be done without converting the performance of detection into real business value.

Although these strengths are present, one should consider the findings as they are characterised by some limitations. Kaggle datasets, as realistic and large-scale as they are, cannot be used to completely recreate the semantic richness of proprietary SAP audit logs. ERP-related objects, which include transaction code, authorisation objects and business workflow context, are approximated instead of being explicitly modelled. However, the good results found with the enterprise-like data on behaviour point out the fact that the suggested framework does reflect the essential patterns applicable to ERP security. The strategy would further be validated and developed by using real SAP audit data in future research. Generally, it is possible to note that the discussion highlights that the effective protection of ERP cybersecurity presupposes a comprehensive combination of AI-based behavioural analytics and

scalable data processing, as well as risk-driven decision support. The suggested framework can be used to meet these needs and support both the technical and management aspects of the security of information systems.

CONCLUSION

This paper presented and assessed an AI-based big data analytics platform to forecast a cyber-threat and risk management in SAP-like ERP systems. The deficiencies of the traditional rule-based security of an ERP, the combination of distributed data processing, unsupervised machine-learning-based anomaly detection, and risk scoring on the basis of behaviour-based risk scoring in a unit decision-support architecture, drive the framework. As the study depicts, the user behavior modeling on the ERP module can improves the detection of insider misuse, privilege elevation and cross-module threats of complexable module threats more effectively than the non-evolving security policies.

The empirical evidence based on enterprise-like data states that the proposed structure can achieve high detection rates and significant reduction of false-positive notifications and time to act on an incident. The issue of scalability of the current ERP systems is solved by Hadoop and Spark enabling the analysis of large amounts of log data in near real-time. It is worth noting that the implementation of anomaly detectors into a risk-scoring system and a cybersecurity dashboard can convert the technical information into intelligence that can be easily acted upon by the security analyst and managerial stakeholders.

In principle, this research contributes to the literature on information systems since it demonstrates that the ERP security and risk management can be logically aligned with AI and big data analytics. The study continues with the existing studies made on the behavioural analytics and insider threat detection, however the study explicitly examines the ERP environments and cross-module associations. In practice, the framework offers a adaptable and scalable approach to organisations, which require to enhance their ERP cybersecurity, in an endeavor to support the processes of governance, compliance, and strategic decision-making.

The proposed structure should be applied to the real SAP audit logs and more advanced sequence-based and graph-based learning models should be explored to capture more complex working processes of the ERP to be used in the future. In addition, the automatic response mechanisms, and explainable approaches with AI, may be applied to the AI to render it even more trusted and efficient in its work. Overall, the findings show the strengths of intelligent risk-oriented security analytics as one of the guiding principles of next-generation ERP cybersecurity strategies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Akinleye, O. K., & Adeyoyin, O. A Data Analytics–Driven Model for Supplier Onboarding and ERP-Based Compliance Management. https://doi.org/https://doi.org/10.54660/.IJMRGE.2021.2.6.538-555

[2] AlMarri, M., Al-Ali, M., Alzarooni, M., AlTeneiji, A., Al-Ali, K., & Bahroun, Z. (2025). Enterprise resource planning systems for health, safety, and environment management: Analyzing critical success factors. *Sustainability*, *17*(7), 2947. https://doi.org/https://doi.org/10.3390/su17072947

[3] Amini, M., & Abukari, A. (2020). ERP SYSTEMS ARCHITECTURE FOR THE MODERN AGE: A REVIEW OF THE STATE OF THE ART TECHNOLOGIES. *1*, 70-90.

https://doi.org/10.22034/jaisis.2020.232506.1009

[4] Atakari, C. (2024). A Multi-Layered Cybersecurity Model for ERP Systems Supporting National Critical Infrastructure: Threats, Challenges, and Solutions. *International Journal of Emerging Trends in Computer Science and Information Technology*, *5*(1), 94-101. https://doi.org/https://doi.org/10.63282/3050-9246.IJETCSIT-V5I1P110

[5] Rahul Reddy Bandhela, Abhishake Reddy Onteddu, RamMohan Reddy Kundavaram. (2022). Enhancing Precision Healthcare Machine Learning For Advanced Diagnostics And Personalized Treatment. South Eastern European Journal of Public Health. https://doi.org/10.70135/seejph.vi.6690

[6] Dachepalli, V. (2025). AI-Driven Decision Support Systems in ERP. *International Journal of Computer Science and Data Engineering*, *2*(2). https://doi.org/http://dx.doi.org/10.55124/csdb.v2i2.248

[7] Efe, A. (2024). Risk modelling of cyber threats against MIS and ERP applications. *Pamukkale Üniversitesi İşletme Araştırmaları Dergisi*, *11*(2), 502-530. https://doi.org/https://doi.org/10.47097/piar.1550812

[8] Faruk, O. M., & Khan, M. K. (2022). BLOCKCHAIN-ENABLED BI FOR HR AND PAYROLL SYSTEMS: SECURING SENSITIVE WORKFORCE DATA. *American Journal of Scholarly Research and Innovation*, *1*(02), 30-58. https://doi.org/https://doi.org/10.63125/et4bhy15

[9] Hannula, J. (2025). Development of access rights management in ERP systems. https://urn.fi/URN:NBN:fi-fe2025052148785

[10] Islam, M. M., & Hasan, M. M. (2023). Explainable AI (XAI) Models For Cloud-Based Business Intelligence: Ensuring Compliance And Secure Decision-Making. *American Journal of Interdisciplinary Studies*, *4*(03), 208-249. https://doi.org/https://doi.org/10.63125/5etfhh77

[11] Jha, K. M. (2023). AI-Powered Big Data and ERP Systems for Autonomous Detection of Cybersecurity Vulnerabilities. *Available at SSRN 5113376*. https://doi.org/https://dx.doi.org/10.2139/ssrn.5113376

[12] Khatri, D. K., Goel, P., & Renuka, A. (2024). Optimizing SAP FICO Integration with Cross-Module Interfaces. SHODH SAGAR: International Journal for Research Publication and Seminar, 15 (1), 188. Link,

[13] Kocaoglu, B. (2024). Enterprise Applications in Logistics (Data Processing). In *Logistics Information Systems: Digital Transformation and Supply Chain Applications in the 4.0 Era* (pp. 121-180). Springer. https://doi.org/https://doi.org/10.1007/978-3-031-60290-0_5

[14] Mirhosseini, S. S., Ramezani, M., Khazaei, M., & Azar, A. (2021). Exploring and analysing the risks and challenges of implementing ERP systems: Critical system thinking. *International Journal of Information Systems and Change Management*, *12*(3), 234-258. https://doi.org/https://doi.org/10.1504/IJISCM.2021.120325

[15] Mishra, A. P., Dublish, M., & Kumar, D. (2022). Cyber security application in ERP implementation. *J. Pharm. Negat. Results*, *13*, 2507-2522. https://doi.org/10.47750/pnr.2022.13.S06.325

[16] Moore, C. (2023). AI-powered big data and ERP systems for autonomous detection of cybersecurity vulnerabilities. *Nanotechnology Perceptions*, *19*, 46-64. https://doi.org/https://dx.doi.org/10.2139/ssrn.5114902

[17] Mukkawar, A. (2025). Adaptive Security Framework for ERP Systems: Leveraging AI/ML with RBAC and ABAC to Combat Emerging Threats.

[18] Nittala, E. P. (2024). Secure data warehousing in ERP environments: An AI-based multimodal

threat detection framework. *International Journal of Emerging Trends in Computer Science and Information Technology*, *5*(3), 111-121. https://doi.org/https://doi.org/10.63282/3050-9246.IJETCSIT-V5I3P111

[19] Nittala, E. P. (2025). AI-Powered ERP Process Mining and Optimization Techniques for Agile Enterprise Transformation. *American International Journal of Computer Science and Technology*, *7*(6), 15-24. https://doi.org/https://doi.org/10.63282/3117-5481/AIJCST-V7I6P102

[20] Samson, O. (2025). Evaluating the Impact of Machine Learning on ERP Data Analytics and Reporting Capabilities.

[21] Sarja, J. (2023). DISCOVERING KNOWLEDGE WORK TASKS FROM SEQUENTIAL EVENT DATA.

[22] Shaik, M. (2023). SAP - ERP Software's Pivotal Role in Shaping Industry 4.0: Transforming the Future of Enterprise Operations. *Computer Science and Engineering*, *13*, 1-7. https://doi.org/10.5923/j.computer.20231301.02

[23] Sharma, G., Thakur, A., & Tiwari, C. (2024). Developing a comprehensive framework for user and entity behavior analytics (UEBA): Integrating advanced machine learning and contextual insights. *Journal of*, *14*(2). https://doi.org/10.37591/JoCES

[24] Shish, Z. H. (2025). Securing ERP Systems: The Role Of Information Security Analysts In US Textile And Manufacturing Enterprises. *International Journal of Business and Economics Insights*, *5*(3), 459-493. https://doi.org/https://doi.org/10.63125/y8evt228

[25] Singh, S., Alam, M. N., Kaur, B., Kaur, K., Kaur, S., & Hossain, S. (2025). Comparative analysis of Apache Hadoop and Apache Spark for business intelligence. AIP Conference Proceedings,

[26] Syed, S. (2025). Machine Learning Algorithms for Optimizing Big Data-Enhanced Cybersecurity in ERP Ecosystems. *Journal of Artificial Intelligence and Big Data Disciplines*, *2*(1), 36-44. https://doi.org/https://doi.org/10.70179/mnqh8179

[27] Tariq, S., Baruwal Chhetri, M., Nepal, S., & Paris, C. (2025). Alert fatigue in security operations centres: Research challenges and opportunities. *ACM Computing Surveys*, *57*(9), 1-38. https://doi.org/https://doi.org/10.1145/3723158

[28] Vukman, K., Klarić, K., Greger, K., & Perić, I. (2024). Driving efficiency and competitiveness: Trends and innovations in ERP systems for the wood industry. *Forests*, *15*(2), 230. https://doi.org/https://doi.org/10.3390/f15020230

[29] Yathiraju, N. (2022). Investigating the use of an artificial intelligence model in an ERP cloud-based system. *International Journal of Electrical, Electronics and Computers*, *7*(2), 1-26. https://doi.org/https://dx.doi.org/10.22161/eec.72.1

[30] Yu, J., Oh, H., Kim, M., & Jung, S. (2021). Unusual insider behavior detection framework on enterprise resource planning systems using adversarial recurrent autoencoder. *IEEE Transactions on Industrial Informatics*, *18*(3), 1541-1551. https://doi.org/https://doi.org/10.1109/TII.2021.3090362