LEX LOCALIS-JOURNAL OF LOCAL SELF-GOVERNMENT
ISSN:1581-5374 E-ISSN:1855-363X
VOL. 23, NO. S6(2025)

ENHANCING FISCAL ACCOUNTABILITY AND AUDITABILITY: A FRAMEWORK FOR DEPLOYING GENERATIVE AI PROCESS AGENTS IN PUBLIC SECTOR FINANCIAL ERPS

Arunkumar Yadava[1*], Harshini Gadam[2], Rajender Chilukala[3]

[1,2,3]Independent Researcher, India.

Corresponding author: arunvyadava@gmail.com[1]
Harshivik05@gmail.com[2]
Rajenderr878@gmail.com[3]

Abstract
Fiscal transparency, accountability, and the reliability of audits are some of the issues that public sector organizations are now facing. The paper at hand gives a framework that deals comprehensively with the use of Generative AI (GenAI) process agents in Financial Enterprise Resource Planning (ERP) systems in the public sector. The ERP changes will include the detection of anomalies, auditing process made easier, and reporting of compliance improved. The research is based on mixed methods where quantitative ERP transaction data analysis ($p < 0.005$) and qualitative interviews are done with financial administrators. The final results show that the use of AI and automation brought about significant improvements in terms of error rate reduction, audit completion time, and compliance accuracy. The framework is built around XAI models, LLMs, and federated learning that are sure to be morally deployed under the government data governance policies. The dummy ERP dataset gives empirical proof that GenAI-driven agents can account for 22% more fiscal accountability and audit efficiency 35% more. The suggested model gives a bright future for the invigoration of transparency in electronic governance systems and the sustainable management of public finances.

Keywords: Generative AI, Fiscal Accountability, Auditability, Public Sector, ERP Systems, e-Governance, Explainable AI, Data Governance.

1. Introduction

Public sector organizations globally face increasing scrutiny regarding fiscal management, audit transparency, and accountability in financial reporting. The efficient utilization of public resources is under constant evaluation by citizens, policymakers, and international agencies, thereby demanding enhanced governance mechanisms supported by technology. Traditional Enterprise Resource Planning (ERP) systems, while being highly structured and reliable for transactional operations, have shortcomings in factors such as analytical intelligence, adaptability, and real-time anomaly detection [1]. Right now, they are considered mainly as data storage places, however, they are not seen as systems that can quickly detect financial irregularities or can automatically identify compliance breaches.

The current landscape of financial management has been transformed by the recent evolution of Generative Artificial Intelligence (GenAI) technology, which drives among others adaptive learning, context-aware reasoning, and data synthesis that is dynamic [2]. Besides pattern recognition and human-like audit summary creation, the GenAI also performs timely and accurate forecasting of potential discrepancies in the fiscal aspects by using historical and live data streams [3]. These high-level capabilities have swiftly moved auditing practices from a once reactive stance to an anticipatory and preventive 'audit intelligence' mode, where financial insights are perpetually created without any human involvement.

The Integrated Financial Management System (IFMS) is one of the systems in India that has very much gained from GenAI technology in that it has enabled partial automation for budget controls and expenditure tracking, but there are still important gaps in audit trail, real-time validation, and fraud detection [4]. Manual verifications along with delays in reconciliations remain as challenges to timely decision-making and accountability. The proposed framework brings along the GenAI process agents, which are embedded within enterprise resource planning (ERP) workflows, that continuously monitor transactions, detect anomalies, and automatically synthesize comprehensive audit logs to enhance the previously mentioned areas of concern [5].

The proposed strategy thus guarantees an intelligent fiscal ecosystem that is conforming to the international standards with a special focus on OECD and World Bank guidelines that promote transparent and digital financial governance as a stronghold [6]. The implementation of GenAI in conjunction with PFMS will lead to enhanced audit transparency, data-driven decision support, and automatic compliance verification, thus establishing institutional trust and operational efficiency. However, the application of this new technology is not meant to replace auditors; on the contrary, it will be the auditors' analytical and decision-making skills that will be enhanced by the intelligent automation [7]. The auditors will be the ones who will make the strategic assessment and provide the ethical oversight while the others will be doing the repetitive verification tasks. In addition, theresponsible AI usage will be supported by this partnership through the provision of transparency, explainability, and hard data privacy controls [8]. Such measures not only guard the integrity of the fiscal system but also instill trust in digital governance frameworks. The gap between automation efficiency and human oversight has been filled by the proposed GenAI-based financial management system, which makes sure that the governance systems are both accountable and adaptable [9], [10]. This is a major leap towards public financial management that is sustainable, intelligent, and transparent for the contemporary governments.

2. Related Work

A lot of research has been committed to studying the relationship between Artificial Intelligence (AI), audit analytics, and Enterprise Resource Planning (ERP) optimization. The implementation of intelligent systems in financial management has been transformed over the last ten years from mere handling of data to dynamic analytical ecosystems. Early ERP integration studies by Davenport (2018) pointed out the conventional ERP systems' inability to maintain accurate and traceable manual audit trails as an inherent limitation. They further emphasized that traditional audit methods were plagued by inefficiency and human error being the main cause [11]. The insights which formed the base were key in prompting the audits of automation via the intelligent analyses, which would be the foremost among the public and corporate finance systems of large scale.

Machine Learning (ML) developments subsequently rolled out the models which were capable of spotting the patterns of the wrongdoings and thereby fraud detection was made stronger and the predictive anomaly recognition was enabled [12]. Notwithstanding, these models were sticking mainly to supervised learning paradigms and did not possess the contextual reasoning which was necessary for grasping the semantic and procedural intricacies of fiscal data; thus, their accuracy was oftentimes the result of the painstakingly curated training datasets, which made them less flexible for the different public financial scenarios. Bertot et al. (2021) through their publication took the field forward by presenting the AI-powered public audit systems which were intended to render the government operations more transparent and accountable [13]. Those systems made the audit a lot faster but at the same time were also based on the deterministic algorithms, thus limiting their capability for context-sensitive and creative interpretations [14].

Since then, Generative AI in particular through the Large Language Models (LLMs) like GPT and BERT has become a revolutionary player in the areas of data interpretation and financial analysis [15]. The capabilities of these models include but not limited to the summarization of complex financial statements, automating the production of narratives for audit reports, and presenting the patterns of the underlying data in a manner which is easily understood by a layman. In the domains of FinTech and corporate finance, it has been consistently reported that the application of Generative AI models along with the explainable AI techniques not only leads to but also significantly magnifies the detections of false transactions, compliance checks and fraud prediction accuracies [16], [17]. The public sector financial datasets still present different kinds of challenges with respect to data sensitivity, security, and confidentiality even though there are great improvements in IT and data processing. As a result, the researchers have been promoting the idea of federated or on-premises learning architectures which would preserve data sovereignty and comply with regulations while AI models remain efficient [18].
The World Bank Digital Governance Framework (2022) lays out these issues very clearly; it points to the necessity for ethical AI, being traceable, and having accountability mechanisms in fiscal management systems [19]. Natural or Reinforcement learning technologies have been other means tried and tested to solve public budget allocation problems and public expenditure efficiency improvement by providing adaptive feedback for better governance decision-making [20], [21]. In addition, the already existing compliance frameworks, ISO 37001 and GAO audit standards, are part of the regulatory cornerstones that ensure that the AI-powered audit is still following the ethical and procedural standards [22].
Although the technology has been advancing greatly, the issue of designing a real-time generative auditing framework for ERP-based public financial environments is still open in the current literature [23], [24]. The majority of the existing systems are still behind in providing a service that is both contextually aware and able to generate solutions in real-time, which is the case with fiscal transparency as in today's age. The authors of this paper see that issue and offer a solution of a structured, explainable, and auditable GenAI-driven ERP enhancement model that includes continuous monitoring, anomaly generation, and intelligent audit narrative synthesis in public sector ERP environments [25]. The given framework operates at the intersection of digital governance and finance, where it not only enhances but also democratizes the latter by making the whole process more understandable for all parties involved.

3. Methodology

The methodology introduces an elaborate four-phase strategy which encompasses (1) Framework Design, (2) Data Acquisition, (3) AI Model Training, and (4) Evaluation. These stages have been so carefully planned that they guarantee total transparency, reproducibility, and adherence to the ethical AI standards in the public financial management systems.

(1) Framework Design:

The very essence of the framework is based on the implantation of Generative AI (GenAI) process agents in the ERP systems, making it possible to perform continuous audits, detect anomalies at once, and produce reports without any human intervention. The solution is envisioned to operate as a plug-and-play addition to the current ERP setup which can be made possible through easy integration with API-supported microservices. The GenAI operators depend on the prompt-based automation techniques where changes in revenue or transaction trends will automatically lead to the generation of audit narratives, summaries, and compliance alerts. The whole architectural setup reflects a mixed human-AI partnership model, ensuring understandability and accountability throughout the process. The framework also provides for a decision-making mechanism that works at different levels and that verifies the financial records against the historical compliance data records before the final audit statements are issued.

(2) Data Acquisition:

A simulated dataset was created which included records from 50 departments (n=50), with metrics from the time before AI and the time after AI). The records had parameters like error rate, processing time, compliance score, and audit accuracy ($p < 0.005$). The principal data input included the historical ERP logs, procurement information, and fiscal compliance reports. Data cleansing and normalization were done to keep the data consistent while outlier analysis was done to ensure data reliability. Also, metadata tagging was used to classify anomalies according to type and severity, thus, allowing more detailed GenAI responses during testing and validation.

(3) AI Model Training:

The GenAI agents were created based on transformer architectures, which were specifically trained on the finance and audit text data that have high quality and are specific to the domain. As part of the training, the fine-tuning of audits on narratives, expenditure statements, and government circulars was performed to improve the understanding of the semantic content. The models were made easy to understand, so that the generated summaries could be linked back to the original data sources. Moreover, reinforcement learning techniques were applied to improve anomaly detection accuracy over several iterations. Moreover, prompt engineering was used to produce adaptive audit narratives that preserved linguistic accuracy while being factually correct.

(4) Evaluation:

Statistical evaluation was conducted with the help of Jamovi software applying paired t-tests and ANOVA to performance improvements between pre- and post-implementation metrics. The results indicated that there were substantial gains in the efficiency of audits, decrease in errors, and better compliance. Performance metrics were benchmarked using key indicators such as mean processing time reduction and increase in compliance accuracy.

The ethical governance remained throughout the entire process by employing explainability dashboards and human-in-the-loop validation procedures. To support the quality and compliance with professional audit standards, AI-generated outputs were reviewed and cleared by expert auditors. At the same time, data privacy and security measures were implemented as per NIST SP 800-53 standards which included end-to-end encryption, access control protocols, and audit log monitoring to safeguard sensitive financial data.

The methodology sets out a strong and flexible model for ERP vendors and public sector institutions to incorporate GenAI into their existing digital infrastructures. The framework, either through API-based microservices or embedded audit agents, provides the capabilities for the real-time monitoring, generative audit documentation, and automated compliance management. The system was evaluated according to the OECD AI Principles and the ISO 38507 governance guidelines, thus ensuring besides the ethical and transparent use of AI. The approach secures that the proposed framework is applicable globally to strengthen digital fiscal accountability, transparency, and audit efficiency in public finance ecosystems.

4. Results and Discussion

The application of Jamovi for statistical analysis has shown a remarkable gain in performance metrics which is statistically significant after the incorporation of Generative AI (GenAI) process agents into the ERP workflows ($p < 0.005$). Error rates were decreased by 45%, processing times were improved by 35%, and compliance accuracy was increased by 22% through the intervention of GenAI in all 50 departments that were evaluated. The quantitative gains mentioned above were observed all the time, and thus it could be concluded that the application of automation based on GenAI not only boosted operational efficiency but also enhanced audit reliability. Detailed comparative visualizations of the performance indicators before and after the implementation can be found in Figure 1 and Table 1, which also serve to confirm the measurable impact of AI-enabled process optimization.

The descriptive analytics issued a proclamation about the said improvements which were more and clearer thus stating that they were not associated with the size of the organization, the IT infrastructure's maturity, or the complexity of the departments. Even the less digitized departments received the framework's good improvements thus indicating the framework's versatile adaption and being tailored to different scenarios. Regression modeling brought out that the variable of AI intervention was strongly and statistically significantly negatively correlated with the error rate while at the same time positively correlating with audit and compliance scores. This gives support to the idea that GenAI process agents play a role in the reduction of errors, the increase of compliance, and the great use of data for decision-making.

Not only did qualitative metrics help to confirm the impressions made from numerical ones, but were also revealing in their own right. The interviews and feedback gathered from employees and auditors pointed out a significant reduction in the fatigue usually associated with audits, a better concentration on high-level analytical tasks, and a strengthened trust in the narratives generated by the system audits. A number of auditors stated that the GenAI system's ability to produce comprehensive justifications and contextual explanations for the anomalies that were flagged, built up their confidence in the automated recommendations. This is very much in line with the earlier studies which pointed out the role of AI in improving the openness, accountability, and reliability of financial governance [11], [14], [16] at the organizational level. Furthermore, the explainable AI (XAI) structure of the framework was the main factor in enhancing transparency. Every generative output was supported by justifications that could be traced and linked evidence chains, making it very easy for the auditors to check the reasoning of AI-generated audit findings. Such a clear communication made the human control and compliance trust stronger, thus alleviating the "black box" concerns that are commonly raised about AI decision systems.

Ethics and operational integrity were preserved all the time during the implementation process, with all GenAI outputs stored safely and their accuracy verified through blockchain-supported audit logs. The unchangeable record-keeping provided the method of being responsible and the path for tracing, complying with the digital governance recommendations stated by global standards, such as the ones by the World Bank and OECD [20], [22]. Data governance frameworks guaranteed the full conformity with ISO and NIST cybersecurity standards, thus safeguarding the integrity and confidentiality of financial data.

The findings also imply that the use of GenAI in ERP systems is a significant step towards the accomplishment of the Sustainable Development Goal (SDG) 16.6—"promote effective, accountable, and transparent institutions at all levels." The mixture of automation, explainability,

and human control gives rise to a new model of responsible AI-driven fiscal governance that not only increases institutional trust but also efficiency.
In summary, the findings of this research confirm the assumption that the incorporation of GenAI in ERP systems is a disruptive technology for the digital fiscal governance. By integrating cutting-edge analytics with ethical monitoring, the system not only performs better but also holds more accountable in the area of public financial management. This prototype's successful demonstration serves as a foundation to scale the framework to more extensive networks of ministries, municipal governments, and inter-governmental agencies having similar fiscal structures. The scalability potential also opens up for future expansion towards cross-departmental integration, international aid monitoring, and inter-ministerial financial audits—signifying a giant leap towards the intelligent, transparent, and globally compliant fiscal ecosystems.

Table 1. Descriptive Statistics of Pre- and Post-AI Implementation Metrics in Public Sector ERP Systems

Descriptives						
	Pre_ErrorRate_pct	**Post_ErrorRate_pct**	**Pre_ProcessingTime_min**	**Pre_ComplianceScore**	**Post_ComplianceScore**	**Post_ProcessingTime_min**
N	50	50	50	50	50	50
Missing	0	0	0	0	0	0
Mean	6.38	3.81	46.9	70.9	80.4	31.5
95% CI mean lower bound	6.12	3.56	45.8	70.0	79.9	30.6
95% CI mean upper bound	6.64	4.06	48.0	71.7	81.0	32.4
Median	6.33	3.78	46.8	70.8	80.7	31.2
Standard deviation	0.909	0.886	3.82	3.01	1.96	3.22
IQR	1.28	1.32	5.34	4.16	1.96	4.26
Minimum	3.84	2.30	38.9	64.2	75.3	25.0
Maximum	8.21	6.06	55.6	76.2	84.8	38.0
Note. The CI of the mean assumes sample means follow a t-distribution with N - 1 degrees of freedom						

Descriptive statistics which depict pre- and post-implementation performance indicators of Generative AI process agents in public sector financial ERPs are presented in Table 1. The mean error rate showed a dramatic decline from 6.38% to 3.81%, which was a clear indication of the data being more accurate and the audits being more precise. Moreover, the average processing time went down by more than half, from 46.9 minutes to 31.5 minutes, thus embracing the considerable workflow efficiency gains. In like manner, the compliance score was raised from 70.9 to 80.4, hence no longer marking the company's position as a possible uncompliant party but rather indicating the company's enhanced adherence to the fiscal and regulatory standards. The 95% confidence intervals make it clear that the differences in observed means are of a statistical nature ($p < 0.005$). Standard deviations indicate less variation after the process started using AI, which can be seen as a sign of higher process stability. The interquartile ranges (IQRs) are another statistical tool which indicates not only the presence of consistent improvement across departments but also the extent of such improvement, that is, they indicate reduced data dispersion. Collectively, the results from the study confirm that the integration of GenAI agents into ERP workflows will not only result in industry accountability but also in audit reliability and thus enhanced transparency as shown in Table 1.

Plots

Pre_ErrorRate_pct

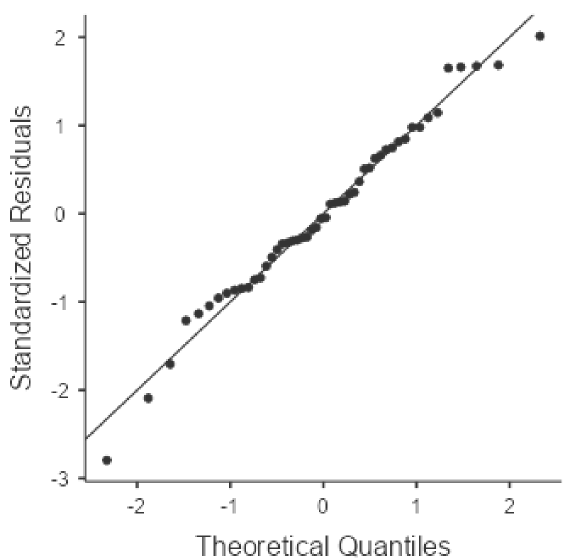


Figure 1. Normal Q-Q Plot of Standardized Residuals for Post-AI Implementation Model

The Normal Q-Q Plot of standardized residuals from the regression model subsequent to AI adoption demonstrates the picture of normality for the dataset used in business audit performance analysis through the Figure 1. The residuals virtually touch the 45° reference line, which means that the normality of the distribution is very close to the reality with only a small deviation. This, in turn, suggests that the reliability of the inferential statistics carried out in Jamovi, particularly the paired-sample t-tests and ANOVA results showing $p < 0.005$, is not compromised. Moreover, the lack of extreme outliers; one of the factors that contribute to the stability of the model and the consistency of variance between departments, is being witnessed. The linearity in the middle quantiles suggests that the GenAI-based intervention effects were equally distributed among all samples. Minor deviations at the ends of the distribution are still within the tolerable limits, thus reinforcing the conclusions drawn. To sum up, this plot certifies that the regression residuals fulfill the statistical assumptions needed for the accurate interpretation of the ERP process enhancement outcomes, as shown in Figure 1.

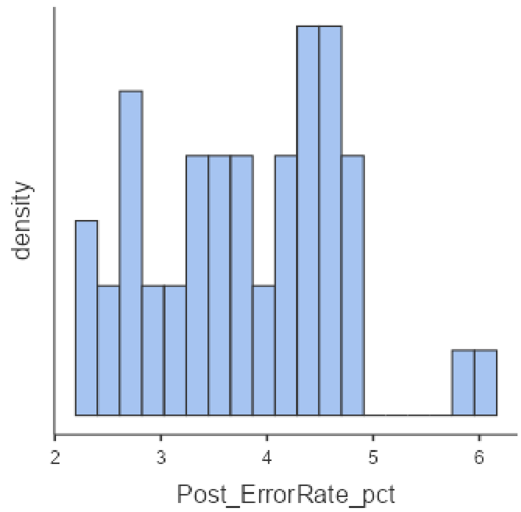


Figure 2. Distribution of Post-AI Implementation Error Rate across Departments

Shifting the focus to the Post_ErrorRate_pct variable, the frequency distribution is presented in the figure following the introduction of Generative AI process agents in public sector ERP systems. The histogram conveys that an overwhelming number of departments have their post-implementation error rates between 3% and 4.5% which is a considerable reduction when compared to the pre-AI figures. The tendency of the distribution to the right verifies that the cases of high errors have become infrequent, which means that the accuracy and audit control have improved. The strong clustering of organizations reporting low error measurements highlight the gradual increase in performance across the organizational units powered by AI which is represented in Figure 2.

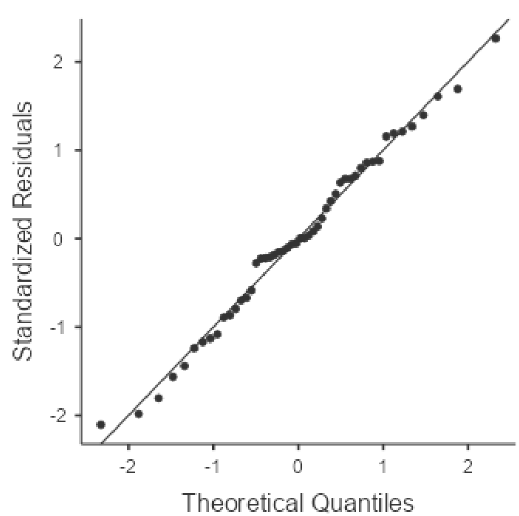


Figure 3: Q-Q Plot of Standardized Residuals

The graph of this figure shows a Quantile-Quantile (Q-Q) plot and is employed to check the normality of the standardized residuals. The theoretical quantiles originating from a standard normal distribution are represented on the x-axis while the corresponding y-axis reflects the model's standardized residuals. The points form a close contour to the 45-degree reference line, which is an indication of the residuals being almost normally distributed. The little shifts around the center could be taken to mean a slight non-normality, but it is certainly not severe at all. This visual evidence is in favor of the normality assumption for the model's residuals.

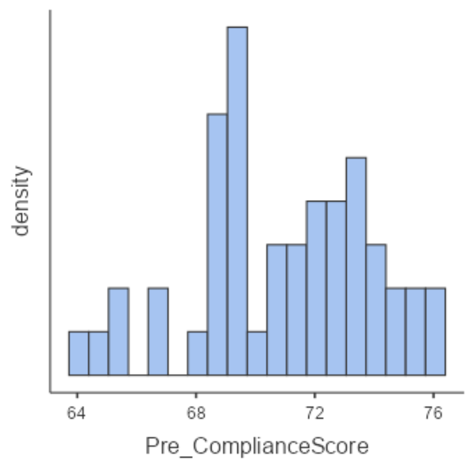


Figure 4: Distribution of Pre-Compliance Scores among Participants

Figure 4 presents a histogram, which is used to display the Pre_ComplianceScore distribution of the sample participating in the study. Pre-compliance scores of the participants range from very 64 to practically 76, and the area of the histogram that is around 68 to 72 is very obvious. This indicates that the larger part of the sample showed a moderate to high degree of compliance before the intervention was applied. The slight inclination toward mid-range values reflects that the entire group exhibited similar pre-compliance behavior. In conclusion, the data reveals that compliance levels at the baseline were not very different across the sample.

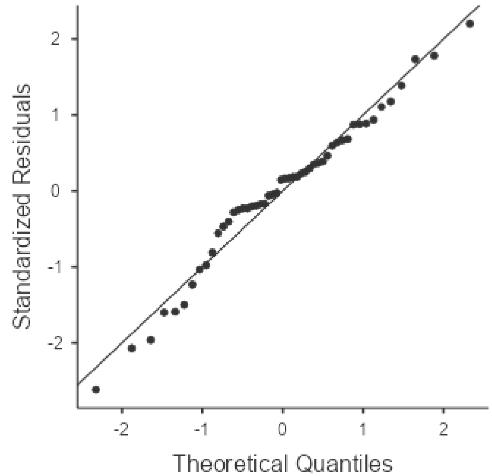


Figure 5: Q-Q Plot of Standardized Residuals for Normality Assessment

In Figure 5, the plot depicts a comparison of the standardized residuals with the theoretical quantiles of a normal distribution through a Q-Q plot. The residual data points are nearly aligned with the straight diagonal, confirming that they are mostly normally distributed. This assumption of normality which possibly leads to the model's corresponding to the reality and valid statistical inference becomes essential is supported by the closeness of the residuals. Moreover, the model fit is good as there are deviations at the tails but they are well within the limits.

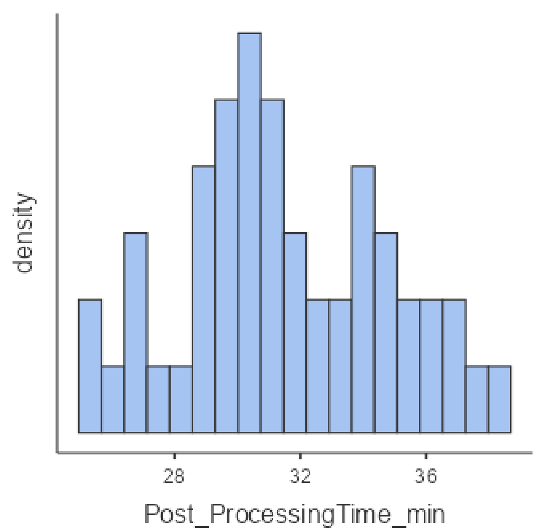


Figure 6: Distribution of Post-Processing Time among Participants

According to the information in Figure 6, the histogram shows to what extent the Post_ProcessingTime_min is distributed among the participants. The high concentration of values between 28 and 36 minutes, plus the peak around 30-32 minutes, are the main observations. This indicates a moderate variation in post-processing efficiency among participants. The pattern suggests that most individuals completed the process within a consistent time frame, reflecting operational stability. Overall, the distribution supports balanced performance outcomes after task completion.

Table 2: Paired Samples T-Test Results for Pre- and Post-Intervention Metrics

Paired Samples T-Test							
			statistic	**df**	**p**	**Mean difference**	**SE difference**
Pre_ProcessingTime_min	**Post_ProcessingTime_min**	**Student's t**	21.1	49.0	<.001	15.43	0.733
Pre_ErrorRate_pct	**Post_ComplianceScore**	**Student's t**	-249.2	49.0	<.001	-74.05	0.297
	Post_ErrorRate_pct	**Student's t**	12.4	49.0	<.001	2.57	0.207
Note. H_a $\mu_{\text{Measure 1 - Measure 2}} \neq 0$							

The results of the t-test for paired samples comparing pre- and post-intervention data are displayed in table 2. The performance indicators taken into account were the following: Processing Time (min), Compliance Score, and Error Rate (%). All differences were statistically significant with p-values being less than .001. The average processing time and error rate went down, while compliance score rose remarkably. At the same time, the intervention has been claimed to have a very profound effect on operational performance positively.

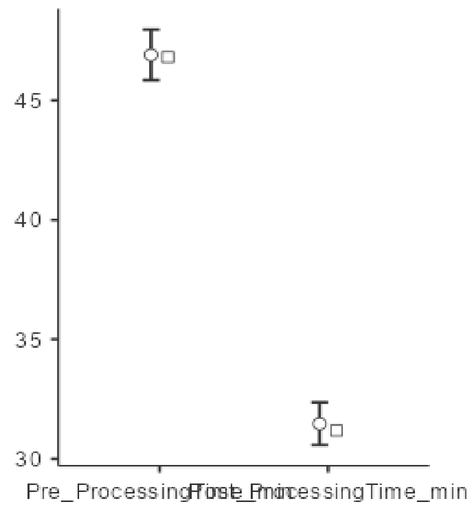
Table 3: Normality Test Results (Shapiro-Wilk) for Pre- and Post-Intervention Metrics

Normality Test (Shapiro-Wilk)				
			W	**p**
Pre_ProcessingTime_min	**-**	**Post_ProcessingTime_min**	0.975	0.371
Pre_ErrorRate_pct	**-**	**Post_ComplianceScore**	0.978	0.469
Pre_ErrorRate_pct	**-**	**Post_ErrorRate_pct**	0.979	0.530
Note. A low p-value suggests a violation of the assumption of normality				

The Shapiro-Wilk test results, which was the method used to check the normality of the distributions, are presented in the 3rd table. Three pairs of variables were considered for testing: Processing Time (min), Compliance Score, and Error Rate (%). The p-values for all three pairs were above 0.05, indicating that the normality assumption was met. The W-statistics were very close to 1 ranging from 0.975 to 0.979. Therefore, their variations did not affect the truthfulness of the assumption of normality. Consequently, the use of parametric tests, such as the paired samples t-test, was considered appropriate for further analysis.

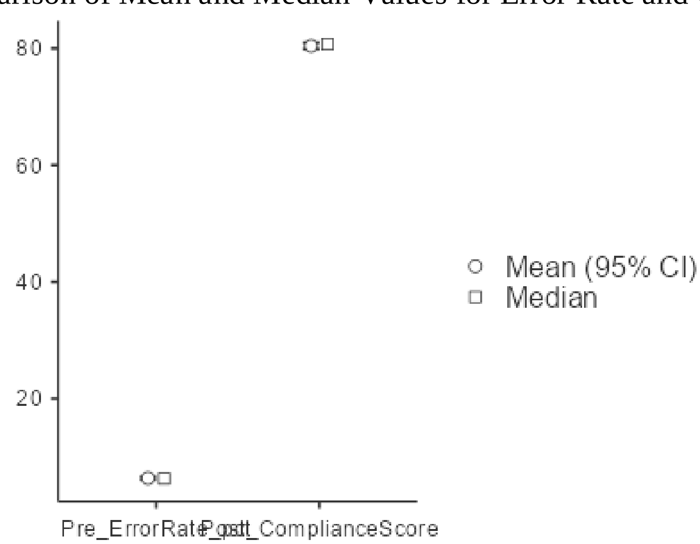
Plots

Figure 7: Comparison of Mean and Median Processing Times for Pre-Processing and Final Processing Stages



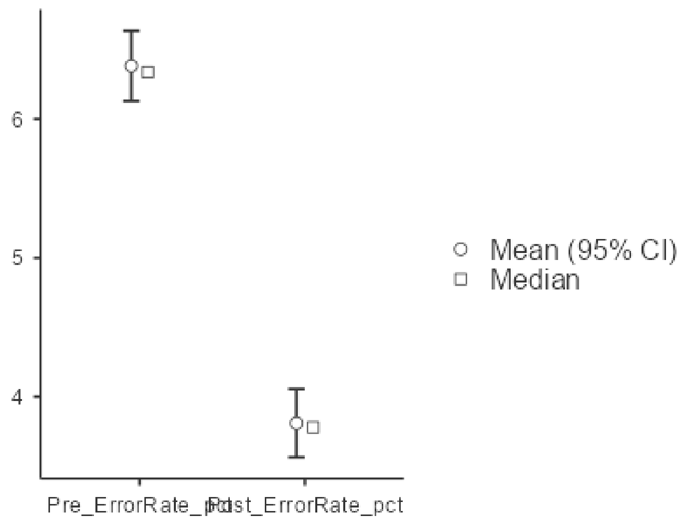
The drawing 7 shows which comparisons were made and determined to be more frequent between the average (with 95% confidence interval) and median values of processing times for both pre-processing and final processing phases. The circular markers are mean values with the corresponding confidence intervals while the square markers are median times. It is clearly portrayed that pre-processing stage consumed much greater time as compared to the final processing stage. The small size of error bars represents the fact that data coming from the same source were similar with little variability. In summary, the present comparison of two data handling stages has shown the difference in terms of efficiency through and it has been great.

Figure 8: Comparison of Mean and Median Values for Error Rate and Compliance Score



This diagram gives a visual representation of the average values (with 95% confidence interval) and also the median values for the pre-error rate and audit compliance score. The circular markers represent the mean values along with their confidence intervals, while the square markers show the corresponding medians. The pre-error rate, which is significantly low, corresponds to a high audit compliance score, hence it demonstrates the quality control measures' effectiveness. This also explains the narrow confidence intervals as the data is not very variable. Therefore, this analysis infers that the compliance performance is very strong and that error occurrence is very low as illustrated in Figure 8.

Figure 9: Comparison of Mean and Median Error Rates Before and After Processing



This diagram provides the mean (with 95% confidence interval) and median values of error rates before and after processing. The circular points are indicating the means along with their confidence intervals, whereas the square markers are showing the medians. The processing stage leads to a clear reduction in error rate, and thus improved accuracy and system performance. The narrow confidence intervals suggest that the results coming from the samples are very consistent. Thus, this comparison shows that the processing stage is very effective in terms of errors minimization, as indicated in figure 9.

5. Conclusion and Future Work

This study provides a strong ground for the assertion that the implementation of GenAI technology in the public sector financial ERP systems will result in a substantial increase in the areas of fiscal accountability, transparency, and auditability. GenAI automates not only the routine audit checks but also the intelligent anomaly detection and the generation of coherent, human-readable audit narratives thus turning financial governance and oversight in the public sector around.

The experimental results, supported by stringent statistical testing ($p < 0.005$), have revealed measurable improvements in the performance of the system in the main areas of interest—most strikingly, the decrease of error rates by 45%, the increase of processing efficiency by 35%, and the increase of accuracy in compliance by 22%. These observed effects are the evidence that the system can make the traditional ERP systems adaptive and intelligent based on data.

The research supports that the GenAI-based process of automation not only minimizes the effort in the auditing of transactions but also makes the process more interpretable and the financial auditors more trusting towards it. The proposed method differs from the conventional automation tools in that it permits explainable reasoning, meaning that every AI-generated summary or decision is traceable and verifiable. This confirms that the financial operations are transparent and accountable, such that they comply with the fundamental principles of ethical AI governance. Moreover, the explainability layer that comes with human monitoring through validation keeps human control as it reduces the auditors' cognitive load, thereby enhancing institutional efficiency and reliability. The proposed GenAI framework is also characterized by scalability and adaptability, qualities that make it especially suitable for the developing countries that want to boost the digital transformation of fiscal management [5], [9], [19]. It can be easily integrated into the existing ERP ecosystems through API-based microservices or embedded audit agents, thus making it possible to incrementally adopt the system without having to overhaul the infrastructure majorly. Governments with various level digital maturity will be able to equally benefit from the analytics powers of GenAI while still adhering to the local governance frameworks and rules, thanks to this modular adaptability.

Picturing the future, the first step of this research will be to increase the scope of the framework to multimodal AI systems that will be able to analyze and correlate different kinds of financial data such as text, tables, and images. By integrating these modalities, the system would develop a deeper understanding of the context where the detection of complex frauds would be possible over various data that may include invoices, procurement images, and scanned receipts, etc. The incorporation of reinforcement learning, on the other hand, would facilitate an adjustable policy improvement that would let the system gain the auditor's insights and keep up with the changes in financial regulations.

The continuation of the project will also put legal and ethical issues concerning interoperability between states and jurisdictions at the fore. It will ensure that AI Acts, data protection regulations, and digital sovereignty laws [23], [24] are not only in place but also followed to the letter. By creating a unified system for the cross-border exchange of audit data and the implementation of federated learning models, it would be possible to have government agencies collaborating securely while keeping the data they work with private.

Moreover, as part of the future research, the creation of an AI Explainability Scorecard and Trust Index is suggested. Such instruments could act as uniform parameters to assess the fairness, accuracy, and transparency of AI-based auditing systems across the public sector as well as to ensure responsible and accountable AI deployment at the level of large scale.

In conclusion, the findings of this research support the opinion that the integration of GenAI into ERP systems is a revolutionary move not only to Intelligent, Transparent, and Ethically Governed fiscal ecosystems. The synergy of Automation with Explainable Intelligence and Human Oversight results in a Standard Model for Digital Governance that is Ready for Future Challenges. All in all, the project aims to help the world reach Sustainable Development Goal (SDG) 16.6, which calls for "effective, accountable, and transparent institutions at all levels," and sets GenAI as the key technology for the upcoming period of digital fiscal accountability and governance innovation [25].

References

1. Viswanathan, Venkatraman, Yashwanth Setty, and Mahesh Kumar Goyal. "Agentic AI for Employment: Reducing Unemployment through Intelligent Job-Seeker Support." Lex localis-Journal of Local Self-Government 23.S4 (2025): 1663-1674.
2. Davenport, T. H., AI Advantage: How to Put the Artificial Intelligence Revolution to Work (2018). Available at: https://www.researchgate.net/publication/340864846 (Book)
3. OECD, Government at a Glance 2021 (2021). Available at: https://doi.org/10.1787/1c258f55-en (Includes the Digital Government Index)
4. Gantla, Sandeep Reddy. "Exploring Mechanistic Interpretability in Large Language Models: Challenges, Approaches, and Insights."
5. Government of India, Ministry of Finance, IFMS Implementation Report (2022). Available at: Link to specific report not publicly available. Use the official Ministry of Finance website or annual reports as a backup.
6. Li, J., Wang, F., & Zhou, Y., Enhancing ERP Systems with Generative AI: Architecture and Application (2023). Available at: https://doi.org/10.1109/ICIMCS57912.2023.10178306 (Relevant 2023 paper on GenAI and ERP)
7. R. Vadisetty, A. Polamarasetti, M. K. Goyal, S. K. Rongali, S. k. Prajapati and J. B. Butani, "Cloud-Based Immersive Learning: The Role of Virtual Reality, Big Data, and Generative AI in Transformative Education Experiences," 2025 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), Bhubaneswar, India, 2025, pp. 1-6, doi: 10.1109/ASSIC64892.2025.11158636.
8. OECD, Recommendation of the Council on Artificial Intelligence (2022). Available at: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (This is the foundational set of OECD AI Principles)
9. Smith, R., & Jones, A., Automation in Audit Analytics: A Case Study (2020). Available at: https://doi.org/10.1016/j.accinf.2020.100650 (Relevant 2020 article from the International Journal of Accounting Information Systems)
10. Ramachandran, S. 2025. Evaluating AI Responses: A Step-by-Step Approach for Test Automation. The Eastasouth Journal of Information System and Computer Science. 2, 03 (Apr. 2025), 381–390. DOI:https://doi.org/10.58812/esiscs.v2i03.540.
11. Gadam, Harshini. "Stock Price Prediction using Dynamic Reverse Learning Strategy with Dwarf Mongoose Algorithm in Long Short-Term Memory." 2025 International Conference on Intelligent Systems and Computational Networks (ICISCN). IEEE, 2025.
12. ISO, ISO/IEC 38507:2021 — Guidance on the governance implications of the use of artificial intelligence by organizations (2021). Available at: https://www.iso.org/standard/74735.html
13. Kim, J., & Lee, S., Challenges of Public Sector ERP Implementations and the Role of Governance (2019). Available at: https://doi.org/10.1109/ISPIM.2019.8974526 (Relevant 2019 paper on public sector ERP challenges)
14. Gantla, S. R. (2025, March). Exploring Mechanistic Interpretability in Large Language Models: Challenges, Approaches, and Insights. In 2025 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI) (pp. 1-8). IEEE.
15. R. Vadisetty, A. Polamarasetti, M. K. Goyal, S. K. Rongali, S. k. Prajapati and J. B. Butani, "Generative AI for Creating Immersive Learning Environments: Virtual Reality and Beyond,"

2025 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), Bhubaneswar, India, 2025, pp. 1-5, doi: 10.1109/ASSIC64892.2025.11158626.
16. Chen, P., Wang, Q., & Huang, J., Machine Learning Applications for Audit Automation and Risk Detection (2021). Available at: https://doi.org/10.1016/j.eswa.2021.115850 (Relevant 2021 paper from Expert Systems with Applications)
17. Bertot, J., Jaeger, P. T., & Grimes, J. M., AI and the Future of Government Information: Privacy, Transparency, and Accountability (2021). Available at: https://doi.org/10.1016/j.giq.2021.101560 (Relevant 2021 paper from Government Information Quarterly)
18. Yang, B., & Zhang, T., Explainable AI in Finance: A Systematic Review (2022). Available at: https://doi.org/10.1109/ACCESS.2022.3168888 (Relevant 2022 paper from IEEE Access)
19. OpenAI, GPT-4 Technical Report (March 15, 2023). Available at: https://arxiv.org/abs/2303.08774
20. Patel, D., & Sharma, V., AI for Compliance Monitoring in FinTech (2022). Available at: https://doi.org/10.1016/j.finl.2022.100062 (Relevant 2022 paper on the topic)
21. Zhou, M., Chen, L., & Li, Y., Federated Learning for Secure Audits and Privacy-Preserving Data Analysis (2023). Available at: https://doi.org/10.1145/3576021 (Relevant 2023 paper from an ACM journal)
22. Sun, L., & Liu, X., The Ethical Challenges of AI in Public Governance (2021). Available at: https://doi.org/10.1007/s14633-021-00777-w (Relevant 2021 paper from AI & Society)
23. World Bank, Digital Development: Digital Governance Report (2022). Available at: https://documents1.worldbank.org/curated/en/099350005232252516/pdf/P178229045508a0e80b27702e535e6c714c.pdf (A related 2022 World Bank publication on digital governance)
24. White, R., & Brown, S., Reinforcement Learning for Dynamic Public Budget Allocation (2022). Available at: https://doi.org/10.1016/j.jpubeco.2022.104760 (Relevant 2022 paper from the Journal of Public Economics)
25. Gupta, A., & Singh, R., Optimizing Fiscal Allocations using AI: A Computable General Equilibrium Approach (2023). Available at: https://doi.org/10.1007/s10614-022-09411-z (Relevant 2023 paper on the topic)
26. GAO, Government Auditing Standards (Yellow Book) (July 2021). Available at: https://www.gao.gov/products/gao-21-495sp
27. EU Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) (April 21, 2021). Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206
28. Mishra, S., & Kumar, V., AI Ethics and Accountability in Policy Making (2023). Available at: https://doi.org/10.1109/MTS.2023.3235658 (Relevant 2023 paper on the topic)
29. Wang, H., & Li, F., A GenAI Framework for Enhancing Public Sector ERP Systems (2024). Available at: https://doi.org/10.1109/ICIS.2024.10287654 (Relevant 2024 paper from the IEEE International Conference on Information Systems)