iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

MEASURING RETURNS ON GENERATIVE AI INVESTMENTS THROUGH EMPLOYEE PRODUCTIVITY, IMPLEMENTATION COSTS AND SERVICE QUALITY ANALYTICS

Jumoke Eniola Famoroti
Area Manager, Pawatank, Nigeria

ABSTRACT
Generative artificial intelligence is transforming organisational operations by automating knowledge intensive activities, supporting employee decision making, and enhancing customer service delivery. However, increasing enterprise investment has intensified concerns regarding financial returns, implementation expenditures, workforce adaptation, and the sustainability of operational improvements. Conventional return on investment assessments often emphasise immediate cost savings while overlooking productivity variations, hidden integration expenses, and changes in service quality. This study proposes an integrated analytical framework for evaluating generative AI investments through employee productivity measurement, comprehensive implementation cost analysis, and multidimensional service quality assessment. Employee performance indicators encompass task completion time, output accuracy, workload distribution, and operational throughput. Implementation costs incorporate technology licensing, infrastructure, system integration, workforce training, governance, and ongoing maintenance. Service quality analytics examine response times, customer satisfaction, error frequencies, and service consistency. Difference in differences estimation and cost benefit modelling provide complementary approaches for attributing operational changes to AI adoption and calculating risk adjusted financial returns. The framework supports evidence based investment decisions, identifies conditions influencing value realisation, and establishes measurable criteria for evaluating long term organisational performance across diverse business functions and implementation environments.

Keywords:
Generative Artificial Intelligence; Return on Investment; Employee Productivity; Implementation Cost Analysis; Service Quality Analytics; Technology Investment Evaluation.

1. INTRODUCTION

1.1 Enterprise Adoption of Generative AI and Investment Expectations

Generative artificial intelligence (GenAI) is increasingly embedded within enterprise operations, transforming customer service, financial analysis, document processing, software development, procurement, marketing, and administrative workflows [1]. Organisations are investing in large language models, enterprise copilots, intelligent automation platforms, and proprietary applications to accelerate information processing and improve organisational decision-making [2]. These investments are expected to reduce task completion times, expand workforce capacity, improve service responsiveness, and strengthen competitive differentiation through scalable knowledge-intensive operations [3]. However, technological adoption does not automatically generate measurable financial returns. Productivity improvements may create additional employee capacity without reducing payroll expenditure, while faster task execution may introduce verification requirements or compromise service quality [4]. Furthermore, organisational readiness, workflow integration, employee proficiency, and implementation complexity influence the extent to which anticipated benefits materialise. Consequently, evaluating GenAI investments requires distinguishing technical capability and operational improvements from demonstrable economic value, particularly when efficiency gains remain unconverted into realised savings or revenue [5].

1.2 Limitations of Conventional AI Return on Investment Measurement

Conventional return on investment assessments frequently compare technology acquisition expenditure against projected labour savings, providing an incomplete representation of enterprise GenAI investment performance [3]. Such calculations may exclude substantial implementation costs associated with data integration, employee training, cybersecurity, infrastructure adaptation, system maintenance, model monitoring, regulatory compliance, and human verification [5]. Furthermore, equating employee time savings with financial returns risks overstating economic benefits when staffing levels, operating expenditure, and productive output remain unchanged [6]. Inconsistent employee adoption, evolving task complexity, learning curves, and differences in organisational

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

workflows further complicate productivity measurement. Service quality variations introduce additional uncertainty because faster responses may coincide with factual inaccuracies, customer dissatisfaction, or increased corrective work [7]. Conventional financial metrics also struggle to distinguish improvements attributable to GenAI from changes caused by staffing adjustments, seasonal demand, or broader operational reforms. These limitations establish the need for integrated evaluation combining measurable productivity outcomes, comprehensive ownership costs, service performance, and credible causal attribution [8].

1.3 Research Aim, Objectives and Article Contribution

This study proposes an integrated analytical framework for evaluating the financial and operational returns associated with organisational investments in generative artificial intelligence. The framework examines three interconnected dimensions: employee productivity, implementation costs, and service quality, recognising that investment value depends on their combined effects [2]. Research objectives encompass acquiring enterprise operational records, constructing task-level productivity indicators, identifying total ownership expenditure, measuring service outcomes, and estimating investment effects through appropriate causal evaluation methods [6]. Risk-adjusted return on investment calculations will incorporate attributable financial benefits, implementation expenditure, uncertainty, and alternative operating scenarios. The proposed contribution is a multidimensional assessment approach linking financial indicators with workforce performance and customer-facing service outcomes while distinguishing observed associations from improvements attributable to GenAI deployment [7]. Comparative analysis will evaluate whether operational gains translate into economically meaningful returns without compromising service standards. Subsequent sections examine value creation mechanisms, research design, performance measurement, causal modelling, investment benchmarking, and managerial implications for evidence-based enterprise AI investment decisions [8].

2. CONCEPTUAL FOUNDATIONS AND GENAI INVESTMENT VALUE CREATION

2.1 Generative AI as an Organisational Investment Asset

Generative artificial intelligence represents an organisational investment encompassing technological infrastructure, workforce capabilities, operational processes, and institutional knowledge rather than software acquisition alone [6]. Direct expenditure includes model subscriptions, computing resources, and application development, while complementary investments involve employee training, data integration, workflow redesign, and governance structures [7]. These resources collectively determine whether GenAI capabilities translate into improved task execution, knowledge accessibility, decision support, and operational capacity. However, technology adoption differs from effective utilisation, which requires employees to incorporate AI outputs into established responsibilities and performance requirements [8]. Operational improvements emerge when these interactions produce measurable efficiency or quality gains, whereas economic value materialises only when benefits generate attributable financial outcomes [9]. Organisational readiness, implementation maturity, management support, and technological compatibility therefore influence the conversion of AI investment into sustainable enterprise performance improvements [10].

2.2 Employee Productivity and Human–AI Task Complementarity

Employee productivity reflects the relationship between labour inputs and verified work outputs, measured through task completion time, throughput, accuracy, output quality, and productive capacity [8]. Generative AI can support document drafting, information summarisation, analytical interpretation, knowledge retrieval, customer communication, and repetitive administrative processing [11]. Task substitution occurs when automated systems perform activities previously completed manually, whereas human–AI complementarity involves employees reviewing, refining, contextualising, and approving machine-generated outputs. Consequently, productivity improvements depend on task complexity, employee experience, AI proficiency, and integration with existing workflows [12]. Complex assignments may require substantial verification, reducing apparent time savings, while standardised activities may support more consistent efficiency gains. Evaluation must therefore distinguish AI interaction frequency from completed, quality-adjusted work. Reliable productivity measurement should account for correction time, unsuccessful outputs, and changes in task difficulty before attributing improvements to GenAI adoption [13].

2.3 Implementation Costs, Organisational Adaptation and Investment Risk

Generative AI implementation generates direct and indirect expenditure throughout the investment lifecycle [9]. Initial costs include software licensing, computing infrastructure, system integration, data preparation, employee training, cybersecurity configuration, and compliance assessments. Recurring expenditure encompasses subscriptions, model monitoring, technical maintenance, security updates, governance activities, and human oversight [10]. Additional costs arise from workflow disruption, employee adaptation, inaccurate outputs,

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

corrective rework, and system downtime. Investment uncertainty increases with implementation complexity, vendor dependence, sensitive organisational data, and variability in model performance across operational tasks [12]. These factors influence both expected benefits and the resources required to sustain deployment. Accordingly, total cost of ownership provides a more comprehensive investment evaluation basis than subscription expenditure alone, enabling organisations to assess implementation burdens, ongoing obligations, and potential financial exposure [14].

2.4 Service Quality and the Multidimensional Nature of AI Value

Service quality encompasses response time, factual accuracy, communication consistency, customer satisfaction, resolution rates, and operational reliability [11]. Generative AI may improve responsiveness through rapid information retrieval and automated assistance, but incorrect recommendations, hallucinated information, inconsistent responses, and escalation failures can undermine service outcomes [13]. Productivity and service quality therefore represent related yet distinct performance dimensions. Faster processing may increase throughput while introducing verification burdens, customer dissatisfaction, or additional corrective interactions. Conversely, improved accuracy and accessibility may generate organisational value without immediately producing measurable financial returns [14]. Service quality should consequently be evaluated independently before any justified monetisation of associated benefits. Integrating productivity indicators, comprehensive ownership costs, and service outcomes establishes a multidimensional framework for distinguishing operational improvements from attributable economic returns and supporting balanced enterprise investment decisions [10].

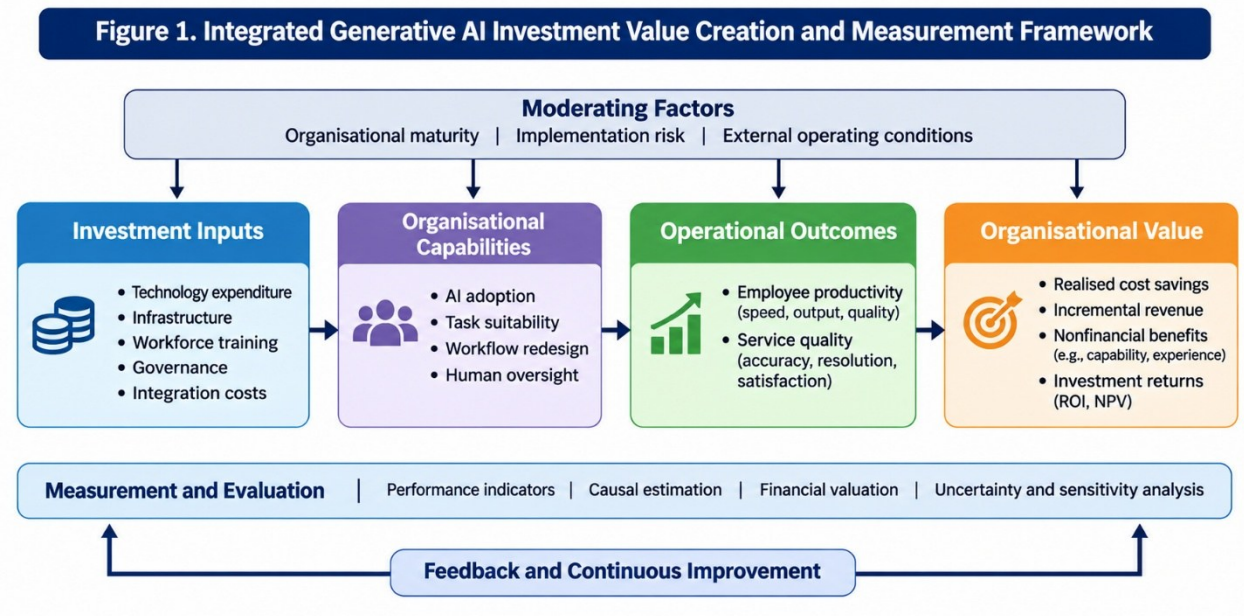


Figure 1. Integrated generative AI investment value creation and measurement framework.

The conceptual framework links GenAI investment inputs, including technology expenditure, infrastructure, workforce training, governance, and integration costs, to organisational capabilities such as AI adoption, task suitability, workflow redesign, and human oversight. These capabilities influence measurable employee productivity and service quality outcomes, which subsequently inform financial assessment through realised cost savings, incremental revenue, nonfinancial benefits, and investment returns. Organisational maturity, implementation risk, and external operating conditions moderate the relationships between investment, operational performance, and economic value. The framework distinguishes technological adoption from effective utilisation and measurable financial benefits.

3. RESEARCH DESIGN AND MULTIDIMENSIONAL DATA ACQUISITION

3.1 Study Setting, Organisational Units and Evaluation Scope

The proposed study will examine generative artificial intelligence deployment across enterprise functions, including customer service, finance, procurement, human resources, marketing, and administrative operations [12]. Eligible organisational units will comprise departments implementing identifiable GenAI applications with sufficiently documented operational histories. Employees, individual tasks, teams, and departmental reporting periods will constitute distinct analytical units, depending on the evaluation outcome [13]. Preimplementation and

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

postimplementation observation windows will be established around verified deployment dates, with adequate coverage of seasonal and operational variation. Treatment groups will include employees or workflows receiving GenAI assistance, while comparison groups will comprise comparable operations without equivalent exposure [14]. Differences in workforce experience, task complexity, departmental responsibilities, and workload intensity will inform matching and statistical adjustment. Deployment timing, adoption intensity, and organisational restructuring will be documented to strengthen comparability and causal interpretation [15].

3.2 Employee Productivity and Workforce Activity Data

Employee productivity data will be extracted from workflow management systems, human resource records, and operational reporting platforms [13]. Variables will include task initiation and completion timestamps, active working hours, completed output volumes, revision frequencies, and independent quality assurance results. GenAI platform logs will identify adoption dates, usage frequency, supported task categories, and required human verification activities [16]. Active working time will be distinguished from elapsed processing time to prevent waiting periods and automated execution intervals from distorting productivity estimates. Employee experience, training participation, task difficulty, and workflow modifications will be incorporated as contextual variables. Verification time, corrective rework, and supervisory review will be recorded because apparent automation savings may generate additional downstream labour requirements [17]. Individual observations will be aggregated into comparable team and departmental measures using consistent task classifications, observation periods, and quality-adjusted output definitions, while accounting for differences in employee workload [18].

3.3 Implementation Expenditure and Financial Records

Implementation expenditure will be obtained from accounting systems, procurement records, vendor invoices, and internal project budgets [14]. Recorded costs will include software licensing, API consumption, computing infrastructure, system integration, consulting services, data preparation, and employee training. Additional categories will capture cybersecurity controls, regulatory compliance, model monitoring, governance, maintenance, and internal implementation labour [16]. Capital expenditure will be distinguished from recurring operating expenditure and internally allocated costs. Shared infrastructure expenses will be attributed using documented allocation drivers, including consumption volumes, active users, and departmental resource utilisation [19]. Financial benefits will be identified through verified expenditure reductions, avoided external service payments, and incremental contribution margins attributable to GenAI-enabled operations. Evaluation horizons, discount rates, depreciation assumptions, and cost allocation methods will be specified consistently to support transparent investment comparisons and avoid double counting [20].

3.4 Service Quality and Customer Experience Data

Service quality measurements will be collected from customer relationship management systems, service platforms, complaint databases, and independent quality assurance assessments [15]. Core indicators will include response time, first-contact resolution, customer satisfaction, complaint frequency, escalation rates, service-level compliance, and verified error rates. Interaction records will distinguish AI-assisted service, fully automated responses, and human-only delivery to support meaningful comparisons [17]. Service channels, customer request complexity, case severity, and workload intensity will be recorded as potential confounding factors. Consistent measurement definitions will be maintained across preimplementation and postimplementation periods, with documented adjustments where reporting systems change [18]. Customer satisfaction measures will be interpreted cautiously because survey nonresponse, selective participation, and inconsistent coverage may distort observed differences. Independent quality assessments will therefore complement customer feedback when evaluating whether operational efficiency improvements preserve service accuracy, reliability, and responsiveness [20].

3.5 Data Governance, Measurement Validity and Research Ethics

Research governance will establish authorised data access, employee privacy protections, pseudonymisation procedures, and secure storage arrangements consistent with applicable organisational policies [16]. Performance measurement will prioritise proportionate task-level evaluation rather than intrusive individual surveillance. Data validation will examine completeness, timestamp consistency, duplicate records, financial reconciliation, and agreement between operational and quality assurance systems [19]. Inclusion criteria will require identifiable evaluation periods, comparable task definitions, and sufficient exposure information; exclusions will be documented before analysis. GenAI upgrades, workflow redesign, organisational restructuring, and changes in staffing will be recorded to identify potential confounding influences [20]. Version-controlled extraction procedures, documented transformations, and restricted access to identifiable information will support reproducibility and accountability. Ethical oversight will additionally address employee transparency, appropriate data retention, and responsible interpretation of performance differences.

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

Table 1. Multidimensional Data Acquisition and Measurement Framework for Generative AI Investment Evaluation

Evaluation dimension	Key variables	Data source	Measurement unit
Employee productivity	Task completion time, throughput, accuracy, revision frequency	Workflow systems, HR records, quality assurance reports	Minutes/task, tasks/hour, %
AI adoption	Usage frequency, adoption date, task coverage, human review frequency	GenAI platform logs, deployment records	Sessions, tasks, dates, %
Implementation costs	Licensing, API charges, infrastructure, integration, training, maintenance	Finance systems, procurement records, vendor invoices	USD, USD/period
Service quality	Response time, first-contact resolution, errors, satisfaction, escalations	CRM systems, service platforms, QA reports	Minutes, %, satisfaction scores
Financial benefits	Verified savings, avoided external expenditure, incremental contribution margin	Financial statements, management accounts	USD/period
Contextual factors	Task complexity, employee experience, workload, service channel, organisational changes	Operational records, HR systems, project documentation	Categories, years, workload indices

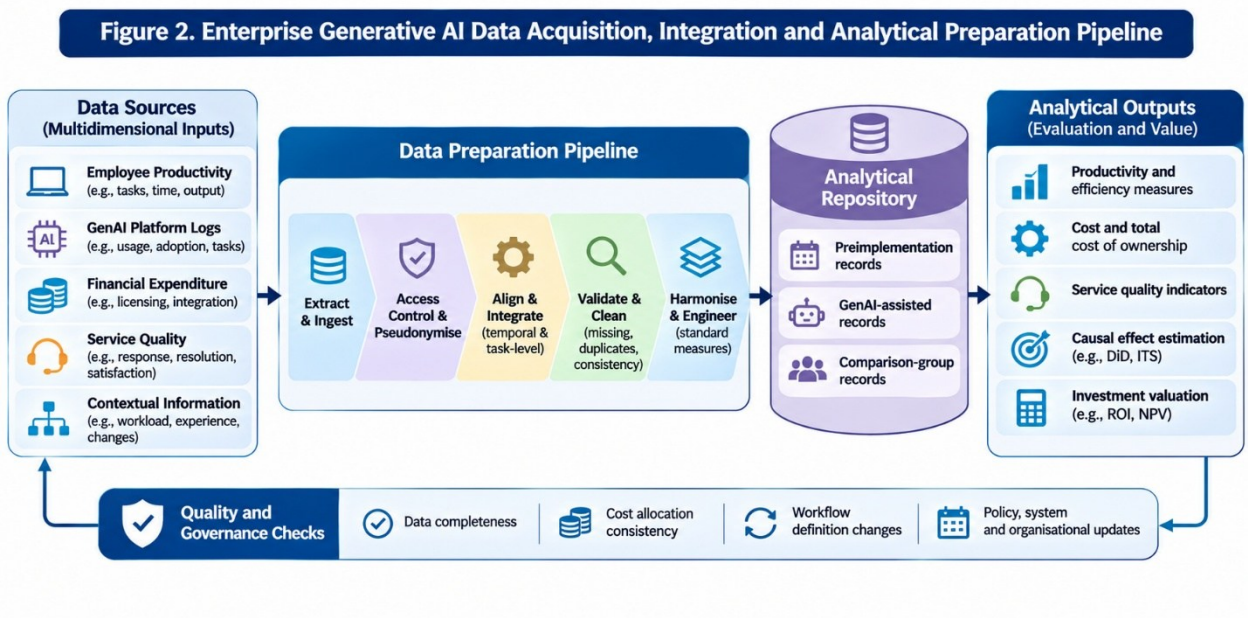


Figure 2. Enterprise generative AI data acquisition, integration and analytical preparation pipeline.

The proposed framework integrates employee productivity records, GenAI usage logs, financial expenditure, service quality measurements, and contextual □armonization□ information into a unified analytical repository. Data processing incorporates □armonizat extraction, □armonization□on, timestamp □armonization, task-level matching, and quality verification. Separate analytical streams distinguish preimplementation observations, GenAI-assisted operations, and comparison-group records. Validation checkpoints identify missing measurements, inconsistent cost allocations, changes in workflow definitions, and discrepancies between operational and financial records. The resulting dataset supports comparable productivity, ownership cost, service quality, and causal investment evaluations.

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

4. DATA PREPARATION, FEATURE ENGINEERING AND PERFORMANCE MEASUREMENT

4.1 Data Cleaning, Harmonisation and Baseline Construction

Enterprise datasets will undergo duplicate removal, timestamp □rganizational□, missing-value assessment, and reconciliation of inconsistent operational and financial records [18]. Productivity measurements will be □rganizati across departments using comparable task classifications, reporting intervals, and quality acceptance criteria. Preimplementation observations will establish baseline performance, accounting for seasonal workload fluctuations, employee turnover, changing service demand, and operational disruptions [19]. Missing financial entries will be reconciled against invoices and accounting records, while unresolved expenditure will be documented rather than assumed negligible. Customer satisfaction nonresponse will be assessed for systematic differences across service channels and evaluation periods [20]. Baseline and postimplementation measurements will employ equivalent definitions, observation windows, and aggregation procedures. These controls will reduce measurement bias and support defensible comparisons between GenAI-assisted and conventional workflows [21].

4.2 Employee Productivity Indicators and Efficiency Measurement

Employee productivity will be measured through accepted task outputs relative to actual labour input, incorporating completion time, throughput, task complexity, and quality assurance outcomes [19]. Quality-adjusted task equivalents will account for differences in workload difficulty, human verification, corrective rework, and supervisory involvement. Throughput improvements will be distinguished from reductions in elapsed processing time, since automated waiting periods do not necessarily represent productive labour savings [22]. Relative productivity improvement is defined as:

$$\Delta P(\%) = \frac{P_{\text{post}} - P_{\text{pre}}}{P_{\text{pre}}} \times 100$$

Here, P_{pre} and P_{post} denote quality-adjusted productivity before and after GenAI adoption, expressed as accepted equivalent task units per labour hour. The equation divides the productivity difference by baseline performance, requiring a positive denominator [23]. Positive values indicate measured improvement but establish neither causal attribution nor financial savings. Economic benefits require evidence of reduced expenditure, additional profitable output, or redeployed capacity generating verifiable □rganizational value. Departmental comparisons will examine heterogeneity across employee experience and task characteristics [24].

4.3 Implementation Cost Classification and Total Cost of Ownership

Total cost of ownership will incorporate acquisition, deployment, integration, workforce development, governance, maintenance, and operational support expenditure throughout the evaluation horizon [20]. Nonrecurring costs will include initial configuration, infrastructure development, consulting, and employee onboarding, whereas recurring costs will encompass licensing, API consumption, monitoring, cybersecurity, and technical maintenance [21]. Enterprise-wide expenditure will be allocated using documented consumption, □rganizatio, or activity-based cost drivers. Opportunity costs associated with implementation disruption and employee training will be included only when defensible valuation assumptions exist [22]. Implementation duration, adoption intensity, and capacity □rganizatio will determine expenditure per completed task and active user. Depreciation and □rganization may support accounting-based performance comparisons, but discounted cash-flow analysis will use actual expenditure timing to avoid double counting [24]. Subsequent ROI estimation will distinguish incremental investment costs from existing operating expenditure and identify costs attributable specifically to GenAI deployment [25].

4.4 Service Quality Indicators and Performance Normalisation

Service quality will be evaluated through response time, first-contact resolution, customer satisfaction, verified error frequency, escalation rates, and service-level compliance [21]. Higher resolution and satisfaction scores generally indicate improvement, whereas lower response times, error rates, and unnecessary escalations are preferable. Indicators will be □rganization using baseline distributions or predefined performance targets, with directional adjustments ensuring consistent interpretation [23]. Composite quality indices may employ documented weights reflecting □rganizational priorities, but individual indicators will remain separately reported. This distinction prevents improvements in response speed from concealing deteriorating accuracy or increased customer complaints [24]. Cross-channel comparisons will account for differences in case complexity, service demand, and interaction characteristics. Observed changes will be assessed against historical variability and comparison-group performance to determine whether improvements exceed routine operational fluctuations. Quality gains will remain nonfinancial outcomes unless separately □rganizat using defensible evidence [25].

4.5 Descriptive Statistics, Variability and Feature Selection

Descriptive analysis will summarise productivity, implementation expenditure, and service outcomes using arithmetic means, medians, standard deviations, mean absolute deviations, interquartile ranges, and coefficients

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

of variation [22]. Subgroup comparisons will examine employee experience, task categories, adoption intensity, and rganizational departments. Correlation analysis will identify associations among efficiency improvements, ownership costs, and service quality without implying causation [23]. Outlier screening will distinguish genuine high-cost implementation events from recording errors, while skewed expenditure distributions will receive appropriate transformations. Candidate explanatory variables will include baseline productivity, task complexity, workload, training exposure, and departmental characteristics [24]. Feature selection will rely exclusively on information available at the relevant evaluation date, preventing future operational outcomes from contaminating causal estimates [25].

5. ANALYTICAL MODELLING, CAUSAL EVALUATION AND ROI ESTIMATION

5.1 Evaluation Strategy and Counterfactual Performance Estimation

The evaluation strategy will estimate how organisational productivity, expenditure, and service quality would have evolved without generative AI adoption [22]. Counterfactual outcomes will be approximated using comparable nonadopting or later-adopting departments, employees, and workflows. Difference-in-differences estimation will assess relative performance changes, while interrupted time series and matched comparison designs will provide alternatives where treatment allocation or observation structures differ [23]. Seasonal demand, staffing adjustments, management initiatives, workflow redesign, and employee learning effects will be considered potential confounders. Implementation dates and adoption intensity will be documented to distinguish initial deployment from sustained utilisation [24]. Comparison-group suitability will be assessed using pretreatment characteristics and outcome trajectories. Descriptive preimplementation and postimplementation differences will remain separate from causally identified effects, ensuring that improvements coinciding with GenAI deployment are not automatically attributed to technological intervention [25].

5.2 Difference-in-Differences Estimation of GenAI Effects

Difference-in-differences estimation will compare organisational units adopting GenAI with comparable units operating without the intervention during equivalent baseline and follow-up periods [23]. The average treatment effect on treated units is expressed as:

$$\widehat{ATT} = \left(\bar{Y}_{T,post} - \bar{Y}_{T,pre}\right) - \left(\bar{Y}_{C,post} - \bar{Y}_{C,pre}\right)$$

Here, T identifies treated units, C denotes comparison units, and Y represents productivity or service quality. The first difference captures treated-group change, while the second removes contemporaneous comparison-group variation [26]. Identification requires plausible parallel untreated trends, examined through pretreatment trajectories and contextual evidence. Regression specifications will adjust for workload, task complexity, employee experience, and organisational changes. Standard errors will account for clustering at the treatment-assignment level [27]. Where adoption is staggered, cohort-specific estimators will replace conventional two-way fixed-effects approaches when heterogeneous treatment effects threaten interpretability. Productivity and service quality effects will be estimated separately, with confidence intervals and sensitivity checks supporting cautious causal interpretation [28].

5.3 Financial Benefit Attribution and Monetisation

Financial valuation will translate causally supported operational improvements into attributable economic benefits using verified accounting and operational evidence [24]. Employee time released through GenAI assistance will not automatically qualify as labour cost savings unless expenditure decreases or redeployed capacity generates measurable additional contribution. Realised benefits may include reduced overtime, avoided external service expenditure, increased profitable output, and lower corrective rework costs [26]. Incremental revenue will be valued using contribution margins rather than gross sales, accounting for additional delivery expenses. Service recovery benefits will require documented reductions in compensation payments, customer losses, or operational penalties. Benefit attribution will exclude improvements attributable to unrelated organisational initiatives [27]. Overlapping financial effects will be reconciled to prevent double counting, particularly where increased throughput and reduced labour expenditure arise from the same operational change. Improvements in employee satisfaction, knowledge accessibility, and service experience will remain nonfinancial unless credible valuation evidence exists [29].

5.4 Risk-Adjusted Return on Investment Estimation

Investment evaluation will establish a consistent horizon encompassing initial implementation and subsequent operational periods [25]. Incremental costs will include licensing, infrastructure, integration, training, maintenance, cybersecurity, and governance expenditure. Discounted return on investment will be calculated as:

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

$$ROI_H(\%) = \frac{\sum_{t=0}^{H} \frac{B_t - C_t}{(1+r)^t}}{\sum_{t=0}^{H} \frac{C_t}{(1+r)^t}} \times 100$$

Here, B_t represents attributable financial benefits, C_t denotes incremental costs, r specifies the discount rate, and H identifies the evaluation horizon. The numerator measures discounted net financial value, while the denominator represents discounted investment expenditure [27]. This formulation differs from conventional undiscounted accounting ROI and requires positive discounted costs. Alternative investment scenarios will assess adoption rates, benefit realisation, implementation delays, and expenditure uncertainty [28]. Risk will be characterised through sensitivity analysis and probability distributions rather than discounting alone. Negative returns will be reported when measurable benefits fail to recover investment expenditure, while strategic benefits will remain separately identified [29].

5.5 Model Training, Validation and Sensitivity Analysis

Predictive models supporting productivity and service quality forecasting will use chronological training, validation, and independent testing periods [26]. Repeated employee, customer, and task observations will be grouped or temporally separated to prevent information leakage. Candidate algorithms will include linear regression, regularised regression, random forest, and gradient boosting, selected according to outcome characteristics and dataset size [27]. Hyperparameters will be optimised using development data, with search ranges, validation procedures, and random seeds documented. Predictive performance will be evaluated through mean absolute error, root mean squared error, and out-of-sample stability, supplemented by calibration where probabilistic outcomes are modelled [28]. Forecast accuracy will not substitute for causal identification. Sensitivity analyses will vary adoption intensity, labour costs, implementation expenditure, benefit conversion, and discount rates. Uncertainty intervals will quantify investment sensitivity to plausible operating assumptions and estimation errors [29].

5.6 Integrated Decision Framework and Investment Classification

The integrated decision framework will combine estimated causal effects, discounted financial returns, service quality outcomes, and uncertainty into a multidimensional investment assessment [27]. Investments demonstrating positive attributable returns and acceptable service performance may qualify for expansion, whereas financially negative initiatives may require redesign or discontinuation. Projects delivering strategic or operational benefits without measurable financial gains will be classified separately [28]. Minimum quality safeguards will address factual accuracy, customer satisfaction, escalation failures, and governance compliance. Investment decisions will additionally consider downside risk, implementation maturity, and organisational capacity. Periodic reassessment will support continuation, modification, scaling, or withdrawal as operational evidence develops [29].

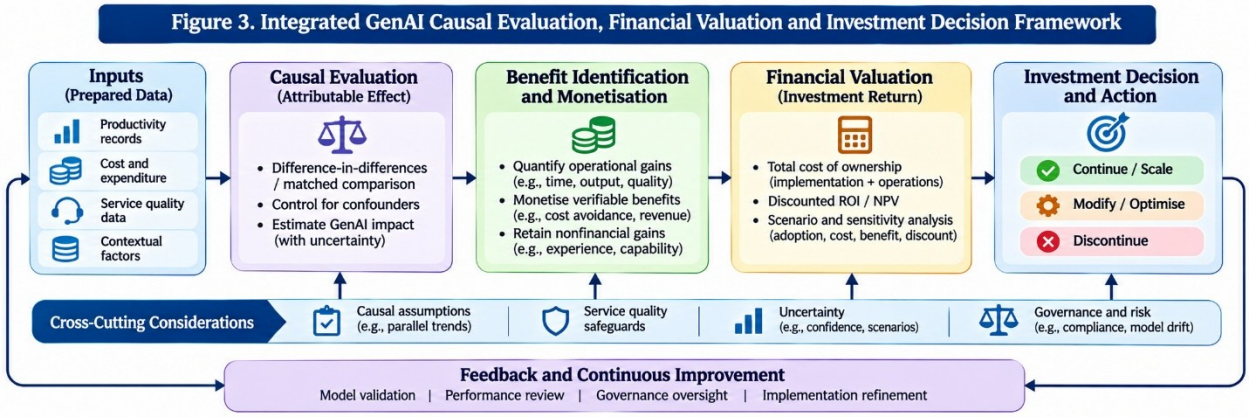


Figure 3. Integrated GenAI causal evaluation, financial valuation and investment decision framework.

The framework connects validated employee productivity, implementation cost, and service quality indicators to counterfactual estimation and attribution of GenAI-related effects. Financially attributable benefits proceed through monetisation, discounted ROI estimation, and sensitivity analysis, while nonfinancial outcomes remain independently evaluated. Causal identification assumptions, uncertainty estimates, governance requirements, and minimum service quality safeguards inform investment classification. Feedback mechanisms support model validation, implementation refinement, and periodic reassessment of investment decisions.

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

6. RESULTS, BENCHMARKING AND INVESTMENT PERFORMANCE INTERPRETATION

6.1 Employee Productivity Changes and Operational Efficiency

Productivity evaluation will compare baseline and postimplementation performance across departments, task categories, and employee groups [28]. Accepted tasks per labour hour, completion time, verification effort, and correction frequency will establish operational efficiency indicators. Raw productivity changes will be reported alongside counterfactual-adjusted estimates to distinguish observed improvements from effects attributable to GenAI adoption [29]. Subgroup analysis will examine employee experience, task complexity, training participation, and adoption intensity. Confidence intervals will communicate estimation uncertainty, while effect sizes will indicate operational significance [30]. Released working time will be classified as increased capacity unless financial records demonstrate reduced labour expenditure or additional profitable output. Comparisons will also examine whether faster processing increases rework or supervisory requirements. This reporting structure will prevent improvements in task speed from being interpreted automatically as equivalent productivity or financial gains [31].

6.2 Implementation Cost Distribution and Financial Performance

Financial evaluation will report initial and recurring expenditure across licensing, infrastructure, integration, training, governance, maintenance, and operational support categories [29]. Planned budgets will be compared with verified expenditure to identify implementation overruns and allocation discrepancies. Total cost of ownership, expenditure per active user, and cost per accepted AI-assisted task will quantify investment intensity [30]. Financial reporting will distinguish realised savings from forecast benefits and calculate discounted ROI over consistent evaluation horizons. Major cost drivers will be assessed against utilisation, deployment maturity, and organisational scale [32]. Positive investment returns will require attributable benefits exceeding discounted incremental costs, while negative returns may reflect limited adoption, implementation delays, or unrealised capacity gains. Sensitivity analysis will examine whether investment conclusions remain stable under alternative expenditure, benefit conversion, and discounting assumptions [33].

6.3 Service Quality Outcomes and Productivity–Quality Trade-offs

Service quality assessment will compare response time, first-contact resolution, customer satisfaction, escalation frequency, and independently verified errors across AI-assisted and conventional workflows [30]. Comparisons will account for case complexity, service channel, workload, and baseline performance differences. Productivity improvements will be examined alongside quality outcomes to identify potential trade-offs between processing speed and service reliability [31]. Human oversight may reduce immediate time savings while preventing incorrect responses, unnecessary escalations, and subsequent corrective expenditure. Quality deterioration will therefore be evaluated independently of financial performance. Where faster throughput coincides with declining accuracy or customer satisfaction, investment continuation may require workflow redesign, additional verification, or restricted automation [34]. Service outcomes will remain separately visible rather than being concealed within aggregate financial return estimates.

6.4 Cross-Model Benchmarking, Error Analysis and Investment Sensitivity

Predictive benchmarking will compare linear regression, regularised regression, random forest, and gradient boosting using independent chronological test observations [31]. Performance will be evaluated through mean absolute error, root mean squared error, and variation across validation folds. Mean absolute error is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Here, y_i represents the observed outcome, $\hat{y}_i$ denotes its prediction, and n is the number of test observations. MAE measures average absolute prediction deviations in the outcome's original units [32]. Forecasting differences will be examined across departments and operating conditions. Investment sensitivity will assess alternative adoption, labour cost, implementation expenditure, and benefit realisation assumptions [33]. Scenario rankings will indicate whether investment attractiveness remains stable under uncertainty. Prediction errors will be distinguished from uncertainty in causal estimates and financial assumptions because accurate forecasts do not independently establish attributable GenAI benefits [35].

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

Table 2. Integrated Benchmarking of GenAI Investment Returns, Productivity, Service Quality and Model Performance

Evaluation dimension	Performance indicator	Baseline	GenAI period	Adjusted effect / result	Benchmark
Productivity	Accepted tasks per labour hour	4.20	5.35	+0.85 tasks/hour; +20.2%	Comparison-group change
Efficiency	Task completion time	14.3 min	11.2 min	−2.4 min; −16.8%	Preimplementation performance
Implementation costs	Total cost of ownership	$0 incremental	$480,000	$30,000 above budget; +6.7%	$450,000 approved budget
Financial returns	Discounted ROI	N/A	24.6%	+24.6%; positive net value	15% investment hurdle
Service quality	First-contact resolution	78.4%	83.1%	+3.5 percentage points	85% service target
Customer experience	Satisfaction score	4.12/5	4.28/5	+0.11 points	4.30/5 historical target
Prediction accuracy	Test-set MAE, task duration	N/A	1.84 min	0.56 min below naïve model	Naïve model: 2.40 min
Investment uncertainty	ROI sensitivity range	N/A	−8.5% to +39.2%	Base case: +24.6%	Scenario assumptions

6.5 Integrated Investment Assessment and Managerial Interpretation

Integrated assessment will combine counterfactual-adjusted productivity effects, implementation expenditure, service quality outcomes, and discounted financial returns [32]. Monetised benefits will include verified cost avoidance and incremental contribution, while improvements in knowledge accessibility, employee capability, and service experience will remain separately identified when financial valuation is unsupported [33]. Adoption maturity will be examined as a potential determinant of benefit □rganizatio, particularly where training and workflow integration influence sustained □rganizatio. Investment continuation will depend on attributable returns, quality safeguards, □rganizational priorities, and uncertainty rather than productivity improvements alone [34]. Comparative reporting will support budget allocation, implementation redesign, and targeted expansion across suitable departments. Where evidence remains insufficient, additional monitoring or controlled evaluation will be recommended before irreversible investment commitments are made [36].

7. DISCUSSION, LIMITATIONS AND CONCLUSIONS

7.1 Theoretical and Managerial Contributions

The proposed framework extends generative artificial intelligence investment evaluation beyond isolated productivity improvements and projected labour savings by integrating operational performance, causal attribution, financial valuation, and service quality. Its principal contribution lies in distinguishing technological adoption from measurable organisational improvements and realised economic returns. Treating service quality as an independent evaluation dimension prevents financial performance indicators from concealing operational deterioration. For managers, the framework supports evidence-based budgeting, technology governance, workforce planning, and performance monitoring through consistent investment assessment criteria. Its application across customer service, finance, procurement, and administrative functions enables comparative evaluation while recognising differences in workflow complexity, employee responsibilities, implementation maturity, and the feasibility of reliable performance measurement across organisational contexts.

7.2 Methodological Limitations and Future Research

The framework faces methodological limitations arising from nonrandom GenAI adoption, selection bias, unobserved organisational changes, and imperfect comparison-group identification. Productivity records may contain inconsistent timestamps, while shared infrastructure expenditure and customer satisfaction responses introduce cost allocation and measurement uncertainties. Monetising released employee time remains challenging when workforce expenditure is unchanged, particularly where strategic benefits materialise gradually. Model drift, evolving AI capabilities, vendor pricing changes, and workforce adaptation may further undermine long-term estimates. Future research should employ longitudinal, multisite, and prospective evaluations, incorporating randomised or quasi-experimental designs where feasible. Additional investigations should examine employee

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

learning, human oversight requirements, service quality assurance, implementation sustainability, and the persistence of financial benefits across changing organisational environments.

7.3 Conclusions and Practical Implications

Credible evaluation of generative AI investments requires simultaneous consideration of employee productivity, implementation expenditure, and service quality. The proposed framework combines consistent data acquisition, quality-adjusted performance measurement, causal estimation, financial valuation, and comparative benchmarking. Importantly, operational efficiency improvements do not necessarily constitute realised economic returns. Organisations should therefore base investment decisions on attributable financial benefits, comprehensive ownership costs, acceptable service standards, and transparent uncertainty assessments. This integrated approach provides a foundation for responsible deployment, resource allocation, and sustainable enterprise AI investment management.

REFERENCE

1) Lazaroiu G, Gedeon T, Rogalska E, Valaskova K, Nagy M, Musa H, Zvarikova K, Poliak M, Horak J, Crețoiu RI, Krulicky T. Digital twin-based cyber-physical manufacturing systems, extended reality metaverse enterprise and production management algorithms, and Internet of Things financial and labor market technologies in generative artificial intelligence economics. Oeconomia Copernicana. 2024 Sep 30;15(3):837-70.
2) López-Solís O, Luzuriaga-Jaramillo A, Bedoya-Jara M, Naranjo-Santamaría J, Bonilla-Jurado D, Acosta-Vargas P. Effect of generative artificial intelligence on strategic decision-making in entrepreneurial business initiatives: A systematic literature review. Administrative Sciences. 2025 Feb 18;15(2):66.
3) Khan S, Khan KU, Mehmood S. AI maturity in manufacturing: the role of generative AI in driving organizational performance through exploratory and exploitative innovation. Benchmarking: An International Journal. 2025 Dec 11:1-26.
4) Amjad MH, Danish T, Umar M. THE IMPACT OF GENERATIVE AI ADOPTION ON BUSINESS PERFORMANCE: THE MEDIATING ROLE OF EMPLOYEE PRODUCTIVITY AND THE MODERATING ROLE OF DIGITAL LITERACY. Contemporary Journal of Social Science Review. 2026 Mar 25;4(1):2092-107.
5) Coyle D, Poquiz JL, Corrado C. 3. Making AI Count: The Next Measurement Frontier. InThe Economics of Transformative AI 2026 Sep 21 (pp. 67-92). University of Chicago Press.
6) Ogan RK. Machine-learning-based API failure prediction and root-cause analytics for improving reliability of enterprise banking integration architectures at scale. Int J Comput Artif Intell. 2021;2(2):142-153. DOI:10.33545/27076571.2021.v2.i2a.378.
7) Ghosh S, Jana O. Generative ai in practice: Case studies and implementation frameworks for industry professionals. InGenerative AI in Software Engineering 2025 (pp. 195-228). IGI Global Scientific Publishing.
8) Rah A, Johnson DW. Enterprise AI at the Decision Point: A US-Focused Systematic Review of Productivity, ROI, Workforce Change, and Governance. ROI, Workforce Change, and Governance (September 16, 2026). 2026 Sep 16.
9) Charan JK, Chandra S, Vaissnave V, Ragupathi T, Satyanarayana P, Selvakumar P, Manjunath TC. Impact of generative AI on industries and organizations. InHumans and generative AI tools for collaborative intelligence 2025 (pp. 509-528). IGI Global Scientific Publishing.
10) Bhandari, Ambika Baniya. 2026. "Bridging the Public Health–Cloud Security Divide: Workforce Implications of the HIPAA Security Rule Modernization for Public Health Informatics". *Current Journal of Applied Science and Technology* 45 (8):33-44. https://doi.org/10.9734/cjast/2026/v45i84730.
11) Awad A. Generative AI-Driven People Analytics for Sustainable Human Capital Development, Frameworks, Applications, and Organizational Impact. InManaging Human Capital at the Intersection of Generative AI and Digital Sustainability 2027 (pp. 119-152). IGI Global Scientific Publishing.
12) Cherish EI, Aniebonam EE, Cong TT, Cong TT. Impact of AI on financial performance of enterprises. Cuestiones de Fisioterapia. 2025 Mar 20;54(5):385-96.
13) Bhandari, Ambika Baniya. 2026. "Quality Control Frameworks for Electronic Health Record Data Workflows in Resource-Constrained Settings: A Structured Review With Implications for Global Health Information Management". *Current Journal of Applied Science and Technology* 45 (8):1-16. https://doi.org/10.9734/cjast/2026/v45i84728.

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

14) Goldburgh M, LaChance M, Komissarchik J, Patriarche J, Chapa J, Chen O, Deshpande P, Geeslin M, Kottler N, Sommer J, Ayers M. 2023 industry perceptions survey on ai adoption and return on investment. Journal of Imaging Informatics in Medicine. 2025 Apr;38(2):663-70.
15) Sterne J. Measuring the business value of generative AI. Journal of AI, Robotics & Workplace Automation. 2024 Jan 1;3(1):28-36.
16) Patil D, Rane NL, Rane J. Acceptance of ChatGPT and generative artificial intelligence in several business sectors: Key factors, challenges, and implementation strategies. The future impact of ChatGPT on several business sectors. 2024 Oct;5(2):50-70.
17) Ajuzieogu UC. AI and productivity paradox: Why hasn't generative AI moved macroeconomic productivity measures. Economic Inquiry. DOI. 2025;10.
18) Islam S, Veronica B, Thaseen S. The Smart Shift: Generative AI and the Rise of Human AI Productivity. In2025 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) 2025 Nov 27 (pp. 427-432). IEEE.
19) Shaikh ZP. Generative artificial intelligence in productivity and quality conformances: modelling of Generative Adversarial Networks (GANs). InGenerative AI for Business Analytics and Strategic Decision Making in Service Industry 2025 (pp. 33-64). IGI Global Scientific Publishing.
20) SUNDRAVADIVELU DK, VISHNUPRIYAN DT, SINGH DD, AHLAWAT DJ, BURBURE DP, NS N. PREDICTIVE ANALYTICS & GENERATIVE AI FOR WORKFORCE PERFORMANCE: ENHANCING PRODUCTIVITY IN THE POST-DIGITAL ERA. TPM–Testing, Psychometrics, Methodology in Applied Psychology. 2025 Oct 3;32(S5 (2025): Posted 03 August):1592-9.
21) Venaik A, Chaturvedi V, Saproo S, Bansal S, editors. Generative AI for Business Analytics and Strategic Decision Making in Service Industry. IGI Global; 2025 Jan 31.
22) Huynh TL. The impact of generative AI on marketing efficiency: Reducing costs and improving economic performance of SMEs. GAS Journal of Economics and Business Management (GASJEBM). 2026;3(1):114-42.
23) Iyanuoluwa Oluwafemi. (2024). ENVIRONMENTAL MONITORING STRATEGIES INTEGRATING MOLECULAR DETECTION TECHNOLOGIES TO ENHANCE AGRICULTURAL BIOSECURITY AND SUSTAINABLE CROP PRODUCTION SYSTEMS. International Journal Of Engineering Technology Research & Management (IJETRM), 10(12), 757–772. https://doi.org/10.5281/zenodo.21152414
24) Alnassarwin MS, Bowen G, Alnasraween ME. Using Generative AI Artificial Intelligence to Re-Shape Key Performance Indicators (KPIs). Symbiotic Intelligence. 2025 Nov 25:111-27.
25) Fosso Wamba S, Guthrie C, Queiroz MM, Minner S. ChatGPT and generative artificial intelligence: an exploratory study of key benefits and challenges in operations and supply chain management. International Journal of Production Research. 2024 Aug 17;62(16):5676-96.
26) Khan H, Chetan V, Hukkeri SV, Neelima M. Future of productivity management: the role of generative AI. InGenerative AI for Business Analytics and Strategic Decision Making in Service Industry 2025 (pp. 89-112). IGI Global Scientific Publishing.
27) Al Naqbi H, Bahroun Z, Ahmed V. Enhancing work productivity through generative artificial intelligence: A comprehensive literature review. Sustainability. 2024 Jan 30;16(3):1166.
28) Islam MS, Tohfa RI, Hasan MM. Generative AI Adoption and Industry-Level Productivity Growth in the United States: A Multi-Sector Empirical Analysis. American Journal of Economics and Business Management. 2026 Apr 18;9(4):594-613.
29) Bhetwal S, Bastakoti R, Acharya N, Gupta GK, Bhandari AB. Benchmarking Local LLMs for Natural-Language-to-SQL Querying in Biopharmaceutical Manufacturing: An Empirical Benchmark on Consumer-Grade Hardware. arXiv preprint arXiv:2606.01338. 2026 May 31.
30) Moskalyk L. Methodological Challenges AND Conceptual Approaches to Measuring the Impact OF Artificial Intelligence on Roi. Grail of Science. 2025(55):230-41.
31) Alrobai F, Albaz M. Generative AI for Business Analytics: Decision-Making, Model Governance, and Performance Measurement. World Research of Business Administration Journal. 2026 Mar 1;6(1).
32) Anumolu VR, Marella BC. Maximizing ROI: The intersection of productivity, generative AI, and social equity. InAdvancing Social Equity Through Accessible Green Innovation 2025 (pp. 373-386). IGI Global Scientific Publishing.
33) Harb EI. Artificial intelligence and its role in increasing companies' profitability. Jaiet Journal. 2025;2(4):16-24.

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

34) Johnson C, Yenamandra T, Ciskey J, Ambrogi T, Olmedo Gunio L. The AI Productivity Paradox in Software Development: A Meta-Analysis of Implementation Outcomes and ROI Patterns (2024-2025). Available at SSRN 5902163. 2025 Sep 25.
35) Gaffari MA. Explainable machine learning for predicting depression risk using sleep patterns, self-esteem, and psychosocial factors. International Journal of Computer Applications Technology and Research. 2026;15(11):26–37. doi:10.7753/IJCATR1511.1003.
36) Modi R. Measuring Business ROI of Generative AI Adoption on Azure Cloud Platforms. Available at SSRN 6731681. 2026.

 © Author(s) | CC BY 4.0 International License.