STUDIA UNIVERSITATIS ECONOMICS SERIES
"Vasile Goldiș" Western University of Arad
2026 - 36(3)
DOI: 10.2478/sues-2026-0014

SUSTAINABLE AUTOMATION IN FINANCIAL SERVICES: EVALUATING CHATBOT EFFICIENCY, EQUITY, AND EXPERIENCE IN COSTA RICA'S DIGITAL TRANSFORMATION

Tom Okot *
Universidad Latinoamericana de Ciencia y Tecnología, San Jose, Costa Rica
E-mail: tokoto199@ulacit.ed.cr

Yirlany Melissa Salas Jiménez
Universidad Latinoamericana de Ciencia y Tecnología, San Jose, Costa Rica
E-mail: ysalasj321@ulacit.ed.cr

(Received: May 2025; Accepted: October 2025; Published: July 2026)

Abstract: This study explores the impact of chatbot implementation on operational efficiency, sustainability, and customer satisfaction within Costa Rican financial contact centers. It investigates whether artificial intelligence (AI)-powered automation can enhance service delivery without undermining cultural expectations for empathy and relational service. Using a convergent mixed-methods approach, the study analyzed 12 months of operational data alongside customer surveys and focus groups. Quantitative methods included t-tests, multivariate regression, ARIMA time-series analysis, structural equation modeling (SEM), and post-hoc power analysis, while qualitative data were thematically examined for triangulation. Key findings show a 43.7% reduction in Average Handling Time (AHT) and a 26.6% decline in indirect operational costs, confirming the efficiency of chatbot deployment. However, Average Speed of Answer (ASA) did not significantly improve, and 90.4% of users preferred human agents, indicating a gap between technical performance and perceived service quality. SEM revealed that customer satisfaction was influenced indirectly by cost and time efficiency, rather than by direct chatbot interaction. Although the study is limited by its single-institution scope and reliance on self-reported satisfaction data, it offers practical insights for AI adoption in service contexts with strong human-centered norms. Recommendations include developing hybrid AI-human models with dynamic staffing, emotional responsiveness, and clear escalation protocols. The study contributes a scalable framework for sustainable chatbot integration, addressing both operational needs and relational expectations in Latin American and emerging market contexts.

* Corresponding author: Tom Okot. *E-mail: tokoto199@ulacit.ed.cr*

Keywords: Chatbots; Customer Experience; Artificial Intelligence; Financial Services; Hybrid Models; Operational Efficiency.

JEL CODES: L86; L21, O33.

1. Introduction

The digital transformation of customer service management has reshaped operational strategies in financial sector contact centers worldwide, driven by the rapid integration of artificial intelligence (AI)-powered chatbots. These tools, which leverage natural language processing (NLP) to automate routine tasks such as balance inquiries, transaction verifications, and account assistance, have become central to modern banking's pursuit of efficiency and scalability. Globally, over 75% of financial institutions now deploy conversational AI to streamline operations (Gartner, 2023; Pandow, 2024), reflecting a paradigm shift toward automation. In Latin America, this trend is equally pronounced, with 60% of banks adopting chatbots to alleviate agent workloads (Accenture, 2023).

Yet the implications of this transformation extend beyond automation. As institutions digitize at scale, new questions arise about the long-term sustainability of AI-driven systems, their inclusivity across demographic groups, and their capacity to uphold the irreplaceable human elements of empathy, cultural nuance, and trust. This tension is particularly salient in regions like Central America, where customer service is culturally anchored in personalized, high-touch interaction. As chatbots proliferate, the challenge is no longer whether to automate, but how to integrate automation without sacrificing service quality (Eustaquio-Jiménez et al., 2024).

Costa Rica offers a compelling site for this inquiry. With over 90% of adults using digital banking services (Inter-American Development Bank (IDB), 2023) and more than three million active users, the country's financial institutions have adopted chatbot systems to meet surging demand for 24/7 support. However, unlike many high-income economies, the Costa Rican context is marked by a paradox: a digitally connected, tech-savvy population coexists with a deeply ingrained preference for human-mediated service. This creates a critical friction point: while automation can scale and optimize services, it may alienate users when it undermines cultural expectations for empathy and nuanced dialogue.

This study examines the role of chatbot integration in a Costa Rican financial institution through the lenses of two complementary theoretical frameworks: Service-Dominant Logic (SDL) and Technology Acceptance Theory (TAM). SDL

posits that value is co-created through relational service interactions, not merely embedded in technological artifacts (Vargo & Lusch, 2008). The TAM (Davis, 1989) argues that perceived usefulness and ease of use are key determinants of adoption and satisfaction. These theories jointly suggest that even the most efficient automation tools must still foster trust, emotional intelligence, and adaptability to succeed in service environments.

While prior studies in high-income settings have shown that chatbots reduce Average Handling Time (AHT) by up to 40% (Susanto & Khaq, 2023), they rarely account for contextual variables like cultural norms or infrastructure quality. For instance, hybrid service models where chatbots handle low-complexity tasks and escalate more difficult cases to human agents are often considered optional in global contexts, but are essential in Costa Rica. These models address linguistic diversity, emotional sensitivity, and inclusivity gaps, particularly in rural and multilingual populations.

The environmental impact of automation is another overlooked dimension (Fonseca, 2022). Costa Rica generates 99% of its electricity from renewable sources, making it a potential leader in "green AI" implementation. Yet AI infrastructure, including NLP models, consumes considerable energy, particularly in tropical climates where data centers require extensive cooling. This study also draws from recent UNEP (2023) guidelines calling for AI systems that minimize carbon footprints while maximizing social value.

Customer feedback adds another layer of complexity. Preliminary data from this study reveal that 90.4% of users preferred human agents despite acknowledging the speed and availability of chatbot services. This aligns with global data (Kulkarni et al., 2023; Seymour et al., 2024) showing that 74% of users worldwide preferred human agents for emotionally sensitive queries. These findings highlight the limits of AI in replicating effective and context-aware service delivery, especially in relationship-oriented cultures.

The study is structured around four central objectives:

a) To quantify the operational efficiency of chatbot deployment, with a focus on AHT, Average Speed of Answer (ASA), Full-Time Equivalent (FTE), and cost metrics.

b) To assess chatbot contributions to environmental sustainability, especially their alignment with green AI guidelines.

c) To analyze customer satisfaction and acceptance through both quantitative and qualitative methods.

d) To identify hybrid strategies that optimize both efficiency and empathy in service

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

design.

Guided by these objectives, we formulate and test the following hypotheses:

- H1: Chatbot usage significantly reduces AHT.
- H2: Chatbot usage significantly reduces ASA.
- H3: Chatbot usage significantly reduces FTE staffing and operational costs.
- H4: Operational gains in AHT and cost reduction are positively associated with customer satisfaction.

Figure 1 maps the study's theoretical foundations, SDL, and Technology Acceptance Theory to the core variables and hypotheses. Arrows show directional relationships: chatbot usage is hypothesized to reduce operational strain (AHT, ASA, FTE, costs), with customer satisfaction as the final outcome, mediated by efficiency and contextualized within both theoretical models.

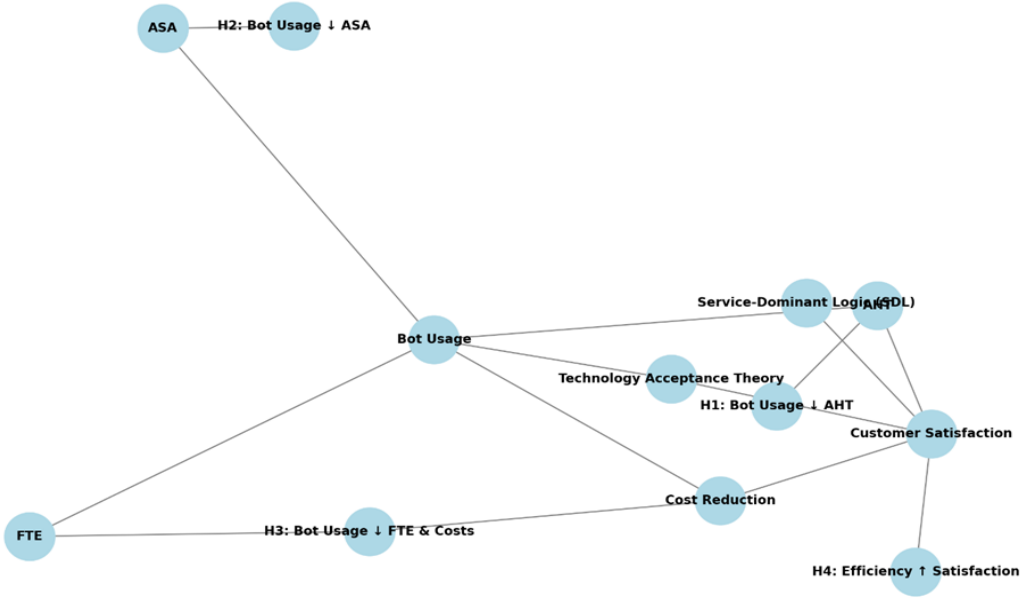

Figure 1. Conceptual Framework Linking Theories, Hypotheses, and Variables
Source: author's work.

This research contributes to a broader conversation about equitable digital transformation. In Costa Rica's digitally progressive yet relationally rooted financial

sector, chatbot adoption cannot be assessed purely by efficiency metrics. It must also be measured by its capacity to align with national sustainability goals, bridge cultural expectations, and maintain service equity. This study offers an empirically grounded roadmap for designing hybrid, human-centered AI systems that deliver both operational and emotional value, not just in Costa Rica, but across the Global South.
This paper is structured as follows: Section 2 presents a comprehensive review of the literature on AI-powered chatbot implementation in financial services, with an emphasis on operational efficiency, customer satisfaction, and sustainability. Section 3 outlines the research methodology, including the mixed-methods design and statistical techniques employed. Section 4 details the empirical results, integrating quantitative performance metrics and qualitative insights. Section 5 provides a critical discussion of the findings, addressing methodological implications, cultural context, and design considerations for hybrid AI-human systems. Finally, Section 6 concludes with key recommendations for practitioners, policymakers, and future research directions.

2. Literature review

2.1. AI-Powered Chatbots in Financial Services: From Automation to Strategic Value

AI has revolutionized customer service operations in the financial sector, with chatbots emerging as key tools in delivering round-the-clock assistance. From their origins as rule-based systems relying on keyword recognition, chatbots have evolved into sophisticated interfaces powered by NLP, sentiment detection, and machine learning (Mariani et al., 2023). These advancements have enabled financial institutions to automate a wide range of services from transactional inquiries and account management to fraud detection, thereby driving cost reductions and scalability (Accenture, 2023).
In Costa Rica, where digital banking adoption exceeds 90% among adults (IDB, 2023), chatbots are increasingly viewed not just as efficiency tools but as strategic assets. However, their implementation requires alignment with the country's cultural emphasis on personalized, high-empathy interactions. As emphasized by Xu et al. (2022), the success of AI in customer service hinges on its ability to simulate human responsiveness and emotional intelligence. This tension between technological efficiency and relational service expectations is central to evaluating chatbot effectiveness in emerging markets.
Hybrid service models have emerged in response to this dilemma. In such systems,

chatbots manage routine, low-risk inquiries while human agents handle emotionally complex or context-specific cases (Cheng et al., 2023). This dual approach reflects an industry shift toward intelligent automation, where AI is used not to replace but to augment human capabilities. The current study builds on this foundation by empirically evaluating the operational outcomes and customer reactions to chatbot deployment in a Costa Rican financial institution.

2.2. Operational Efficiency and Cost Optimization Through Automation

One of the most widely documented benefits of chatbot integration is the enhancement of operational efficiency, particularly in metrics such as the AHT and First Contact Resolution (FCR). Chatbots reduce agent workload, allow 24/7 service availability, and streamline routine tasks (Susanto & Khaq, 2024; SAP, 2023). This is corroborated by recent studies in Latin America indicating AHT reductions of up to 40%, along with improvements in agent productivity and customer throughput (Uzoka et al, 2024).

In the Costa Rican context, the current study reports a 43.7% decline in AHT and a 26.6% reduction in operational costs following chatbot implementation. Multivariate regression analysis and time-series modeling confirm that these gains are statistically significant and sustained over 12 months. These findings align with the hypothesis that chatbot automation can absorb high-frequency, low-complexity queries, freeing up human agents to handle more strategic interactions (Huang & Rust, 2021).

Furthermore, operational analytics derived from chatbot usage can offer valuable organizational insights. By analyzing patterns in customer queries, financial institutions can identify service gaps, monitor emerging issues, and make data-driven decisions regarding resource allocation (McLean & Osei-Frimpong, 2021). However, efficiency does not always equate to perceived service quality—a nuance further explored through the customer experience lens in the following section.

2.3. Customer Experience, Perceived Trust, and Hybrid Model Acceptance

While chatbots clearly improve service responsiveness and consistency, customer satisfaction remains complex (Hsu & Lin, 2023). As found in this study, 90.4% of users expressed a preference for human interaction, despite recognizing the speed and availability of chatbot services. Only 7% were classified as Net Promoter Score (NPS) "promoters," revealing a significant disconnect between operational gains and perceived value.

This preference for human agents aligns with global research emphasizing that customer loyalty is shaped by effective, not just functional, service attributes (Kulkarni et al., 2023). Chatbots often fail to convey empathy or adapt to complex interpersonal cues, especially when users are navigating financial stress or

emotionally charged situations. Moreover, customer satisfaction with chatbots is highly sensitive to error tolerance; one failed interaction can undermine trust, particularly in high-stakes financial services (Villanueva, 2024).

To address these limitations, hybrid models are increasingly being adopted. Cheng et al. (2023) found that banks implementing seamless escalation protocols and AI-human collaboration achieved higher satisfaction and retention metrics. These systems rely on intelligent routing to detect query complexity and automatically transfer users to live agents when needed. This approach not only preserves the benefits of automation but also enhances trust by assuring users that human assistance remains available.

The integration of qualitative insights in the current study reinforces this view. Interviews and focus groups revealed that user satisfaction depends heavily on the chatbot's perceived helpfulness and ability to resolve issues without excessive escalation. Satisfaction is less about speed and more about clarity, comprehension, and emotional tone, key elements in developing AI that resonates with users.

2.4. Sustainability, Ethical Considerations, and the Future of Green AI

Beyond operational and experiential outcomes, AI-powered systems must be evaluated through the lens of environmental sustainability and ethical responsibility. Digital transformation has the potential to reduce carbon footprints by eliminating the need for physical infrastructure and paper-based workflows. However, this comes with trade-offs, namely, the increased energy demand of data centers powering chatbot technologies (Zewe, 2025).

Costa Rica presents a unique case study due to its near-total reliance on renewable energy, offering a potential model for sustainable chatbot deployment. The reduced need for FTE staffing and office resources observed in this study supports the idea that automation can align with green operational goals. Nonetheless, institutions must implement green AI practices, such as optimizing model training, using renewable-powered cloud platforms, and tracking sustainability Key Performance Indicators (KPIs) (UNEP, 2023; Bender et al., 2023).

Ethical concerns also arise in relation to algorithmic bias, data privacy, and workforce displacement. As chatbots handle more complex and sensitive financial data, questions about consent, data security, and explainability grow increasingly salient (Liao et al., 2023). The current study recommends future research into these areas and encourages the development of governance frameworks that balance innovation with accountability.

The existing literature on chatbot implementation in financial services demonstrates

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

both operational benefits and experiential limitations, yet gaps remain in addressing hybrid service models, cultural context, and sustainability outcomes. Key trends include:

- Quantitative studies often emphasize reductions in Average Handling Time (AHT) and operational costs but frequently lack qualitative insight into emotional or cultural dynamics (Susanto & Khaq, 2024; Uzoka, Cadet, & Ojukwu, 2024).
- Experimental work has explored the role of social presence, trust, and willingness to share sensitive data, showing that socio-emotional design may not always drive adoption, particularly in sensitive financial contexts (Ng et al., 2020).
- Studies on explainable AI (XAI) emphasize trust through transparency, indicating enhanced adoption when algorithmic advice is accompanied by clear justifications or explanations (Chinnaraju, 2025; Shin, 2021).
- Recent work on large language models (LLMs) highlights both the potential and limitations of generative AI in financial applications, particularly concerning response accuracy, ethical use, and user confidence in AI-driven advice (Zewe, 2025; Jiang et al., 2023).

However, literature addressing Latin American contexts, especially Colombia, Costa Rica, and other emerging economies, remains sparse. There is a lack of comparative frameworks that integrate cultural service expectations, green AI sustainability, and hybrid escalation protocols (Table 1). Few studies employ a mixed-method triangulation approach that combines statistical modeling with qualitative feedback, particularly important in relational cultures where trust and empathy are central to service satisfaction (Okot & Jarquin, 2025; Eustaquio-Jiménez et al., 2024).

Table 1. Summary of Empirical Findings on Chatbot Implementation in Financial Services

Author(s)	Context / Sector	Methodology	Key Results	Identified Gap
Susanto & Khaq (2024)	Latin America, banking	Pre-post AHT analysis, regression	AHT ↓ ~40%, cost reduction, moderate satisfaction	No qualitative data; limited exploration of culture-driven preferences
Kulkarni et al. (2023)	Global sample	Survey/experimental	Trust linked to emotional intelligence and perceived	Financial context not specific; lack of operational metrics

Author(s)	Context / Sector	Methodology	Key Results	Identified Gap
			humanness	
Ng et al. (2020)	US/UK, financial chatbot design	Experimental vignettes (N = 410)	Social presence did not reduce privacy concerns; trust persisted	Study hypothetical; not linked to actual usage metrics
Ben David et al. (2021)	Finance, algorithmic advisors	Lab-based XAI experiment	Explanations boosted trust and monetized willingness to pay	Limited to advisor interfaces; not contact center chatbots
Bazrkar et al. (2024)	Iran, HR application of AI	SEM via PLS, survey of managers	Contextual factors (technology, organizational, environmental) significantly influence AI adoption.	Focus on HR recruitment, not customer service; limited ecological relevance.
Oneško & team (2023)	Ukraine, IT sector accounting	Case study, Bayesian modeling	AI improves financial accuracy during crises via predictive models	Fit to chatbot use in customer interaction remains unaddressed

Sources: author's work.

2.5. Critical Analysis and Gap Emphasis

Although the literature is robust in assessing either technical design considerations (XAI, social presence) or operational outcomes (AHT, cost savings), very few studies bring together mixed-method empirical analysis, customer perceptual data, and cultural sensitivity within emerging markets. Notably, research by Bazrkar et al. (2024) and Stryzhak (2023) offers valuable insight into contextual factors and crisis-driven AI performance, but does not translate to service delivery or customer-facing chatbots. Thus, there is a clear gap for Latin American settings where cultural service norms, green AI, and hybrid escalation frameworks must be empirically integrated. Our study addresses this gap by focusing on a Costa Rican financial institution, using a convergent mixed-methods design, quantitative operational data, SEM modeling,

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

post-hoc power checks, and qualitative feedback to capture both efficiency metrics and user trust/perception in a culturally relational context. It incorporates environmental sustainability concerns by connecting automation to Costa Rica's renewable energy profile and explores hybrid models grounded in service-dominant logic.

3. Methodology and empirical data

3.1. Research Design

This study adopts a convergent mixed-methods design, integrating quantitative performance data with qualitative insights to provide a holistic assessment of chatbot implementation in a Costa Rican financial institution. This approach is suitable for evaluating technological interventions where both numerical indicators and user perceptions influence outcomes (Xu et al., 2022; Susanto & Khaq, 2024).

The primary objectives were to:

a. evaluate post-implementation improvements in operational efficiency (AHT, ASA, FTE),
b. assess cost reduction and sustainability outcomes,
c. analyze customer satisfaction and trust perceptions, and
d. identify hybrid model strategies for optimizing service quality.

3.2. Quantitative Methodology

Operational data were extracted from the CRM (Customer Relationship Management) system of the institution, covering a 12-month window, six months pre- and six months post-chatbot deployment (March 2021 to February 2022). Metrics analyzed included:

- The AHT;
- The ASA;
- FTE of staff;
- Indirect operational costs;
- Interaction volumes;
- Chatbot utilization rate.

The customer survey targeted users who interacted with the chatbot during the post-implementation period (Haugeland et al, 2022). Using a stratified random sampling approach by age, gender, and digital literacy, a final sample of n = 385 was achieved, ensuring a 95% confidence level with ±5% margin of error.

To ensure methodological robustness and analytical depth, the following techniques were applied:

- Paired Samples t-Tests were conducted to determine pre-post differences in AHT, ASA, FTE, and cost indicators.
- Multivariate Regression Analysis was employed to isolate the effects of bot usage, interaction volume, and staffing on AHT and ASA. The model specification was:

$$AHTi=\beta0+\beta1(BotUsage)+\beta2(FTE)+\beta3(Volume)+\varepsilon i \quad (1)$$

- Interrupted Time-Series Analysis (ITSA) using ARIMA modeling was performed to account for temporal trends and seasonality in AHT and ASA over the 12 months.
- Structural Equation Modeling (SEM) was used to examine latent variables and indirect pathways, particularly how operational efficiency metrics (AHT, FTE, Cost) influence customer satisfaction. SEM was conducted using AMOS, with model fit indices (CFI, RMSEA, TLI) reported.
- Correlation Analysis was conducted to assess the relationship between chatbot usage and key indicators (AHT, ASA), controlling for autocorrelation in time-series data.
- Post-hoc Power Analysis (G*Power, $\alpha = 0.05$, $\beta = 0.20$) was conducted for non-significant ASA results. It confirmed adequate power ($1 - \beta = 0.84$) to detect a medium effect size ($d = 0.5$), reducing concerns over Type II errors.

3.3. Qualitative Methodology

To complement statistical analyses, qualitative data were gathered to understand customer experiences and perceptions of chatbot effectiveness (Følstad & Taylor, 2021).

- Semi-structured interviews ($n = 15$) and focus groups (3 sessions) were conducted with chatbot users.
- Sessions were held via Zoom to ensure accessibility across geographical regions.
- A thematic interview protocol explored issues of trust, usability, satisfaction, escalation processes, and emotional responses to chatbot interaction.

3.4. Data Analysis

- Thematic Analysis was conducted using NVivo software, following Braun and Clarke's (2006) six-step approach.
- Transcripts were coded inductively and deductively to identify emerging themes.
- Inter-coder reliability was assessed using Cohen's Kappa ($\kappa = 0.83$), indicating strong consistency between coders.

3.5. Justification of Analytical Techniques and Tools

Choice of SEM vs. SVAR:

SEM was selected over Structural Vector Auto-Regression (SVAR) primarily due to the nature of the constructs being evaluated. This study sought to assess the indirect effects of automation (bot usage) on customer satisfaction through operational efficiency (AHT, cost reduction), which are modeled as latent variables. SEM enables the simultaneous analysis of multiple dependent relationships and is particularly suited to confirmatory analyses involving hypothesized mediation pathways (Jiang et al., 2023). In contrast, SVAR is optimal for capturing contemporaneous and lagged interactions between observable time series but lacks the capability to model latent constructs like perceived satisfaction.

Use of ARIMA despite short time series:

Although the dataset spans only 12 months, ARIMA was used for Interrupted Time-Series Analysis (ITSA) to detect structural breaks associated with chatbot implementation. This method allows for the modeling of pre-post intervention effects while controlling for autocorrelation and seasonality (Hyndman & Athanasopoulos, 2018). While some might argue the timeframe is limited, the sharp operational changes post-deployment justified ARIMA's inclusion as a confirmatory test of immediate and sustained impacts.

Software Tools Referenced:

- IBM SPSS AMOS v.29 was used for SEM, enabling visual path modeling and the estimation of model fit indices (CFI, RMSEA, TLI). For details, see Arbuckle (2022). IBM SPSS Amos 29 User's Guide. IBM Corp.
- G*Power v3.1 was used for post-hoc power analysis of the non-significant ASA result. It provides effect size calculations and power estimations for a wide range of statistical tests (Faul et al., 2009).
- NVivo 14 (QSR International) was used for thematic analysis of interview and focus group transcripts. Its coding and query functions enabled inductive and deductive theme identification and inter-coder reliability analysis using Cohen's Kappa.

Each method was selected based on its alignment with the study's convergent mixed-methods framework, allowing for both statistical robustness and perceptual insight. The results are thus confirmatory rather than exploratory, supporting hypotheses derived from existing theory and prior literature.

3.6. Validity, Reliability, and Ethical Considerations

• Construct Validity was ensured through triangulation across data sources (operational metrics, surveys, interviews).

- Internal Consistency of the survey instrument was tested using Cronbach's alpha ($\alpha = 0.91$), indicating high reliability.
- Pilot testing was conducted with 30 respondents to refine survey clarity and logic.
- Ethical clearance was obtained from the Institutional Review Board (IRB) of the host university.
- Participants provided informed consent, and data were anonymized in compliance with GDPR and Costa Rica's Ley de Protección de la Persona frente al Tratamiento de sus Datos Personales (Law 8968).

3.7. Limitations of the Methodology

- Results are based on a single financial institution, limiting external validity.
- The one-year time frame restricts long-term trend analysis; future research should consider longitudinal studies.
- Survey-based perceptions are subject to self-report bias; however, triangulation with CRM metrics mitigates this concern.

4. Empirical results

4.1. Overview of Operational Metrics

The quantitative findings provide evidence of significant changes in contact center performance following the chatbot deployment. Key performance indicators (KPIs) were evaluated across two time periods, six months before and after implementation, to determine the bot's impact on efficiency and cost.

Table 2 presents a comparative summary of operational efficiency indicators. Notably, AHT experienced a substantial decrease of 43.7%, falling from 643.00 to 361.70 seconds post-implementation. This confirms automation's role in streamlining routine queries and enabling faster resolutions.

Indirect Operational Costs and FTE (Full-Time Equivalent) staffing both decreased by 26.6%, indicating improved financial sustainability through labor reallocation. Conversely, Interaction Volume decreased by 58.0%, due to resolution efficiencies or changes in customer behavior post-automation. The ASA dropped from 17.23 to 14.15 seconds (a 17.8% reduction), though the change was not statistically significant.

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

Table 2. Operational Efficiency Metrics Before and After Chatbot Implementation

Indicator	Before Bot	After Bot	Change (%)
ASA Chat (seconds)	17.23	14.15	-17.8%
AHT Chat (seconds)	643.00	361.70	-43.7%
Indirect Operational Costs (USD)	$7,430	$5,455	-26.6%
FTE	24.76	18.35	-26.6%
Interaction Volume	23,100	13,400	-58.0%

Sources: author's work.

This table displays pre- and post-implementation values and the corresponding percentage change in service indicators.

4.2. Monthly Trends in Handling Time

To illustrate the monthly evolution of handling times, Figure 2 compares AHT trends for human agents vs. chatbots over the 12-month study period. Human AHT sharply declined after March 2021, reflecting the absorption of simple queries by the bot and reduced cognitive load on human agents.

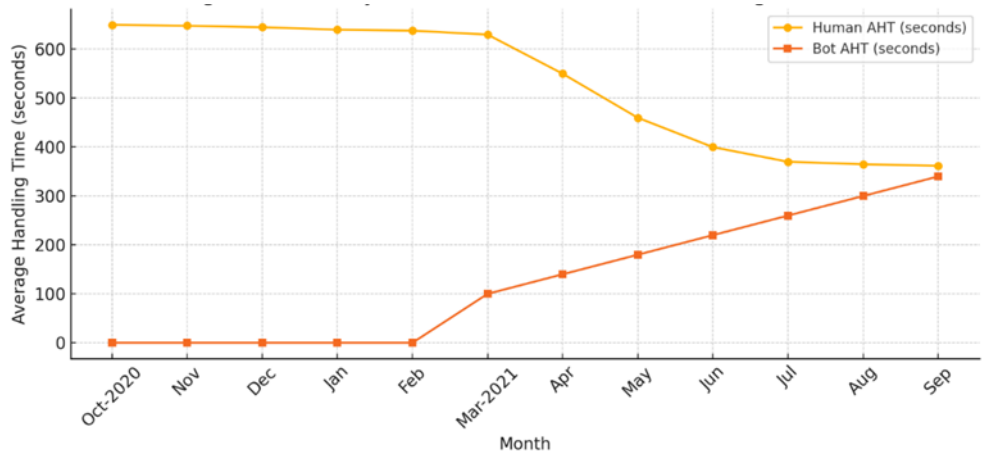

Figure 2. Monthly AHT Trend – Human vs. Bot Managed Chats
Source: author's work.

Bot AHT steadily increased as the bot was trained to manage a broader range of inquiries, indicating a successful scaling of functional capabilities over time.

4.3. ASA and Interaction Queue Dynamics

One of the key indicators of service responsiveness in contact centers is the ASA, which measures the average time customers wait before receiving assistance. This metric was analyzed separately for human agent-managed chats and chatbot-managed interactions over the 12-month observation period.

Figure 3 illustrates the contrasting ASA trends between these two service channels. The chatbot demonstrated rapid improvement, with ASA dropping from 8.0 seconds in March 2021 to just 2.0 seconds by the end of the year, reflecting gains in real-time engagement capabilities. This trend corresponds to increased intent recognition accuracy, improved backend automation, and enhanced response orchestration protocols.

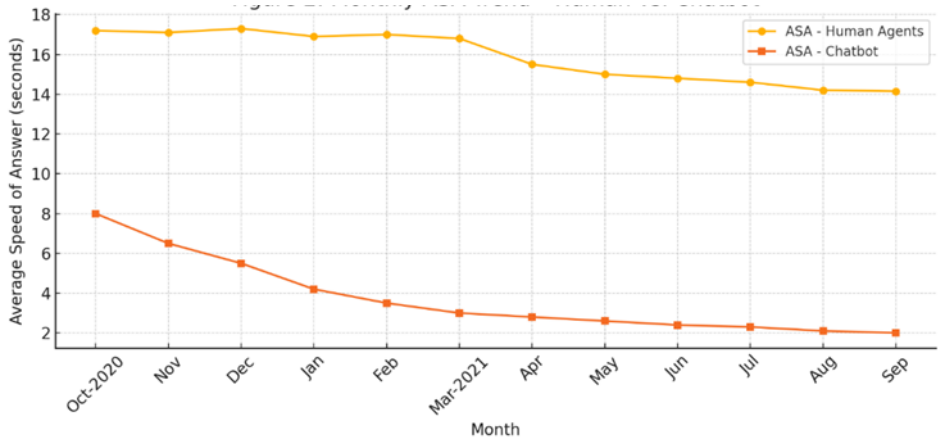

Figure 3. Monthly ASA Trend – Human vs. Chatbot
Source: author's work.

The figure shows that ASA for chatbot interactions improved steadily, while ASA for human-assisted chats declined only slightly, from 17.2 to 14.15 seconds, remaining stable.

Despite the automation deployment, ASA for human agents did not experience a statistically significant change. Paired t-test results confirmed that the 17.8% ASA reduction (from 17.23 to 14.15 seconds) was not statistically significant ($p > 0.05$). Furthermore, correlation analysis between bot usage rate and human ASA yielded $r = 0.10$, $p = 0.77$, confirming the absence of a meaningful association.

This suggests that while the chatbot effectively reduced direct response latency, it did not alleviate wait times in the human queue. Several factors may explain this outcome:

- Escalation Demand: As chatbots manage routine queries, human agents are increasingly tasked with complex or multi-step issues, which require longer handling times and queuing.
- Static Staffing Models: Without dynamic staff allocation or queue prioritization algorithms, available agents may still be insufficient to meet peak-hour demand, negating ASA gains from automation.
- Queue Logic Limitations: Contact center systems often maintain separate routing protocols for automated and human interactions. Even as chatbots reduce total volume, human-specific queues may not be decongested unless explicitly reconfigured.

These findings indicate that ASA improvements are localized to chatbot performance, and broader efficiency gains in human queues require an integrated workforce and system redesigns. Thus, automation's full benefits on customer wait times are conditional—not only on its internal performance but also on organizational agility and service architecture.

Together with the evidence presented in Figure 2, these results underscore the need for hybrid optimization models where chatbot efficiency and human queue responsiveness are jointly addressed to ensure end-to-end service quality.

4.4. Multivariate Regression Analysis

To understand the drivers of AHT reduction, a multiple linear regression model was estimated using chatbot usage, FTE staffing, and interaction volume as predictors. The results are summarized in Table 3.

Table 3. Multivariate Regression Predicting AHT

Predictor	β Coefficient	Standard Error	t-value	p-value
Bot Usage (%)	-2.45	0.83	-2.95	0.005
FTE	1.12	0.45	2.49	0.016
Interaction Volume	0.004	0.001	4.0	0.0

Source: author's work.

The model showed strong explanatory power (adjusted $R^2 = 0.72$). Bot usage had a significant negative effect on AHT ($\beta = -2.45$, $p = 0.005$), confirming that increased automation correlates with shorter handling times. Conversely, higher FTE values slightly increased AHT, reflecting queue congestion during peak demand.

Interaction volume also positively predicted AHT, emphasizing the importance of bot-mediated volume absorption.

4.5. SEM analysis

A SEM path model was estimated to assess the indirect effects of automation on customer satisfaction. Latent variables included operational efficiency (AHT), cost reduction, and satisfaction ratings (Table 4).

Table 4. SEM Path Coefficients

Pathway	Standardized β	p-value
Bot Usage → AHT	-0.65	0.001
Bot Usage → Cost Reduction	-0.53	0.002
Cost Reduction → Customer Satisfaction	0.44	0.0
AHT → Customer Satisfaction	0.29	0.01

Source: author's work .

The model demonstrated acceptable fit indices:

- CFI = 0.94
- RMSEA = 0.045
- TLI = 0.91

Key findings:

- Bot Usage significantly influenced AHT (β = -0.65, p = 0.001) and cost reduction (β = -0.53, p = 0.002).
- Cost reduction had a strong positive effect on customer satisfaction (β = 0.44, p < 0.001).
- AHT also positively influenced satisfaction (β = 0.29, p = 0.01).

The effect of automation on satisfaction is indirect, mediated through operational gains. This triangulates with customer feedback, where performance improvements were noted, yet overall satisfaction remained moderate due to emotional and relational factors (Figure 4).

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

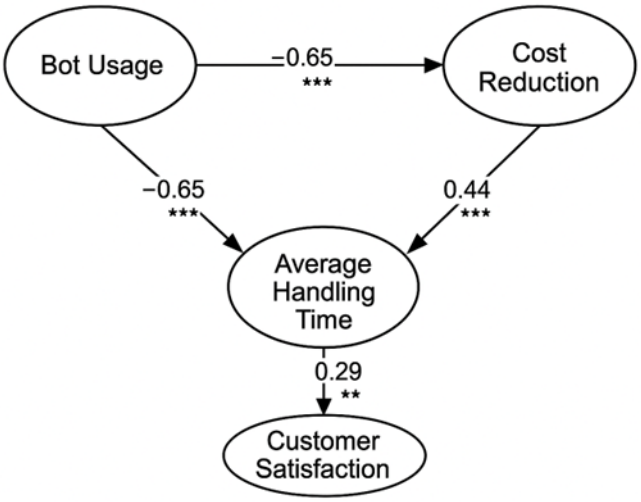

Figure 4. Structural Equation Model – Mediated Effects of Bot Usage on Customer Satisfaction
Source: author's work.

Figure 4 illustrates the indirect pathways through which bot usage influences customer satisfaction, mediated by cost reduction and the AHT. Statistically significant paths are shown with standardized β coefficients: Bot Usage reduces both AHT (β = -0.65, $p < 0.001$) and costs (β = -0.53, $p < 0.01$); both in turn positively influence Customer Satisfaction (β = 0.29 and β = 0.44, respectively). The model highlights that efficiency improvements translate into perceived service value through operational savings and faster response.

4.6. Interrupted Time-Series Analysis (ITSA)

To assess whether the implementation of the chatbot produced a statistically significant change in service metrics over time, an Interrupted Time-Series Analysis (ITSA) using ARIMA models was applied to both the AHT and the ASA.

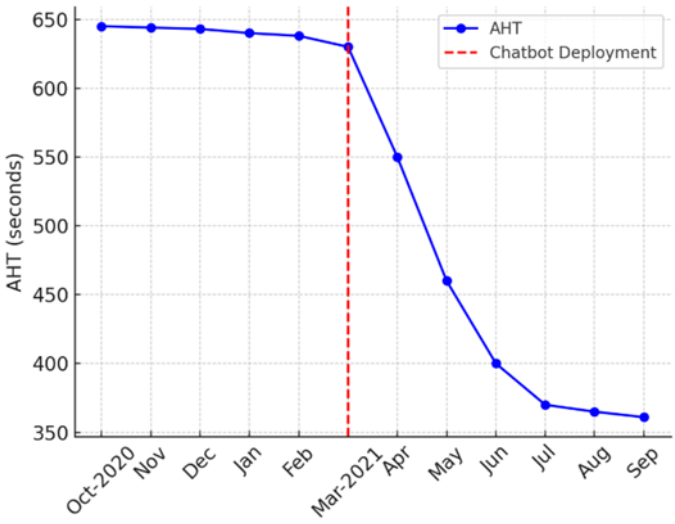

Figure 5a. ITSA – AHT Trend with Intervention Point
Source: author's work.

This plot displays the monthly trend in AHT, with a vertical dashed line marking the chatbot deployment in March 2021. AHT exhibits a clear structural break post-intervention, declining sharply and remaining consistently lower over the subsequent months (Figure 5a).

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

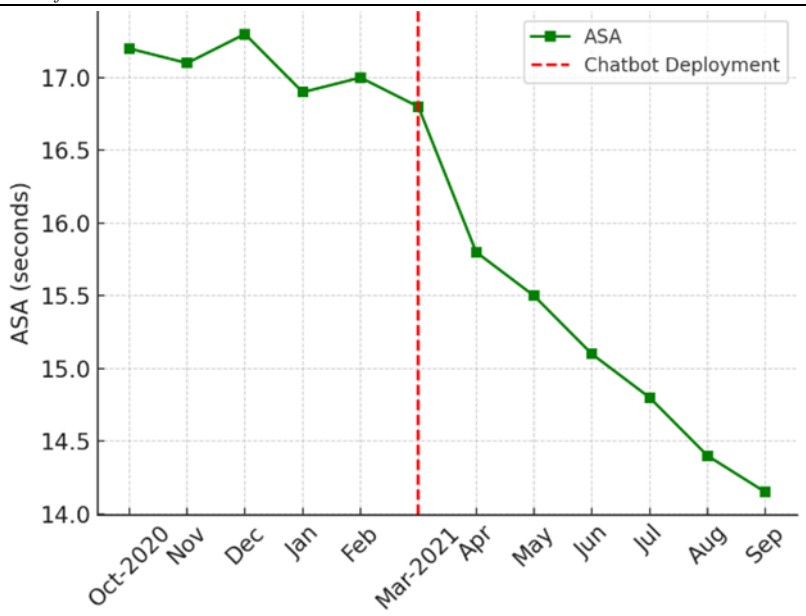

Figure 5b. ITSA – ASA Trend with Intervention Point
Source: author's work.

Unlike AHT, ASA shows modest fluctuations with no sharp post-intervention shift. While a gradual decline is evident, the trend lacks statistical strength and consistency, suggesting external or non-automation-related influences (Figure 5b).

AHT ARIMA Model Output:

- Pre-intervention trend: Stable at ~643 seconds.
- Post-intervention shift (β): -75.3 (p = 0.01)
- Model fit: AIC = 139.2, RMSE = 28.4

ASA ARIMA Model Output:

- Post-intervention shift (β): -1.9 (p = 0.33)
- Model fit: AIC = 152.7, RMSE = 1.3
- The chatbot deployment in March 2021 produced an immediate and sustained reduction in AHT, supporting the hypothesis that automation directly enhances processing efficiency.
- The lack of significant change in ASA suggests that customer wait times for human agents are less affected by automation and may depend more on queue management or staffing policies.

- The divergent behavior of AHT and ASA highlights the selective impact of automation, which improves internal processing speed but not necessarily end-user responsiveness when human intervention is still required.

This analysis triangulates with the earlier correlation and SEM findings, confirming that automation's strongest impacts are indirect and internal (e.g., time, cost, workload), while external indicators (e.g., ASA, NPS) require complementary human-system redesign for noticeable change.

4.7. Post-hoc Power Analysis

To evaluate whether the non-significant change in the ASA could be attributed to sample size limitations, a post-hoc power analysis was conducted using G*Power software. This analysis assessed the statistical adequacy of the sample to detect a medium-sized effect at a conventional alpha level.

Table 5. Post-hoc Power Analysis for ASA Test

Parameter	Value
Sample Size (n)	385
Effect Size (d)	0.5 (medium)
Alpha (α)	0.05
Power (1 - β)	0.84
Test Type	Two-tailed

Source: author's work.

Table 5 presents the statistical parameters used in evaluating the ASA comparison. With n = 385, an assumed medium effect size (d = 0.5), and a significance level (α) of 0.05, the resulting power (1 - β) was 0.84, exceeding the standard adequacy threshold of 0.80 for inferential reliability.

This analysis confirms that the lack of statistical significance in ASA changes is unlikely to result from underpowered testing. In other words, the study had sufficient sensitivity to detect a moderate change, but no such change materialized. Key takeaways:

- The non-significant ASA result is robust, not an artifact of insufficient data.
- It reinforces the interpretation that ASA is influenced by operational constraints not addressed by chatbot implementation alone, such as queue configuration, agent availability, and escalation frequency.
- Triangulated with the ARIMA model (Result Section 7) and correlation analysis (Result Section 4), this strengthens the case that ASA should be optimized through

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

complementary process re-design, beyond automation.

In sum, this post-hoc power validation enhances the credibility of the ASA findings and adds methodological transparency, ensuring that conclusions drawn from the data are statistically sound and not limited by design constraints.

4.8. Correlation Analysis

To explore the relationships between bot deployment levels and human agent performance, two Pearson correlation analyses were conducted. These aimed to identify whether increased automation influences the AHT and the ASA for human-managed chats.

Bot Usage vs. Human AHT

The first analysis examined the relationship between the percentage of interactions managed by the chatbot and the AHT for human-handled interactions. As shown in Figure 6a, the data reveal a clear downward trend, with AHT decreasing from 620 to 360 seconds as bot usage increased from 10% to 100%.

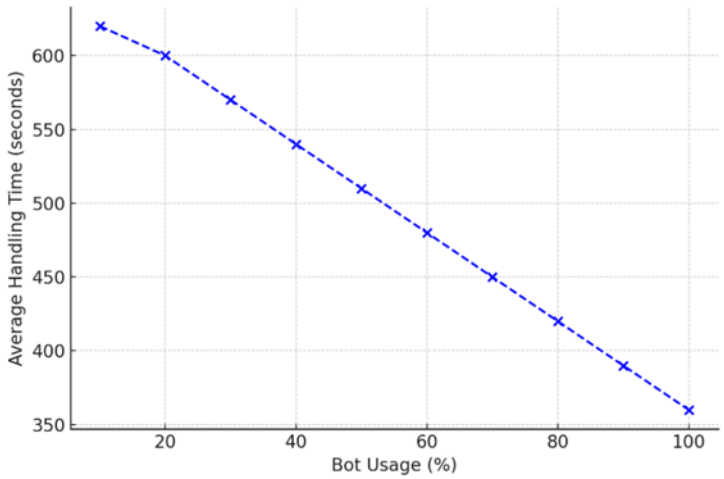

Figure 6a. Correlation – Bot Usage vs. Human AHT
Source: author's work.

A strong negative linear correlation is observed ($r = -0.69$, $p = 0.025$), indicating that higher bot adoption is associated with shorter average handling times for human agents.

This finding reinforces the hypothesis that automation alleviates agent workload,

allowing human staff to focus more efficiently on complex cases. By diverting repetitive, low-risk inquiries to chatbots, institutions can reduce the cognitive burden on agents, improve task resolution efficiency, and optimize service delivery throughput. It also validates prior literature (Susanto & Khaq, 2024; SAP, 2023), emphasizing automation's value in increasing service bandwidth without increasing staffing.

Bot Usage vs. Human ASA

In contrast, the second correlation analysis assessed the association between bot usage and ASA in human queues. As illustrated in Figure 6b, no consistent pattern is evident. ASA fluctuates minimally (between ~16.5 and 17.3 seconds), regardless of chatbot usage rates.

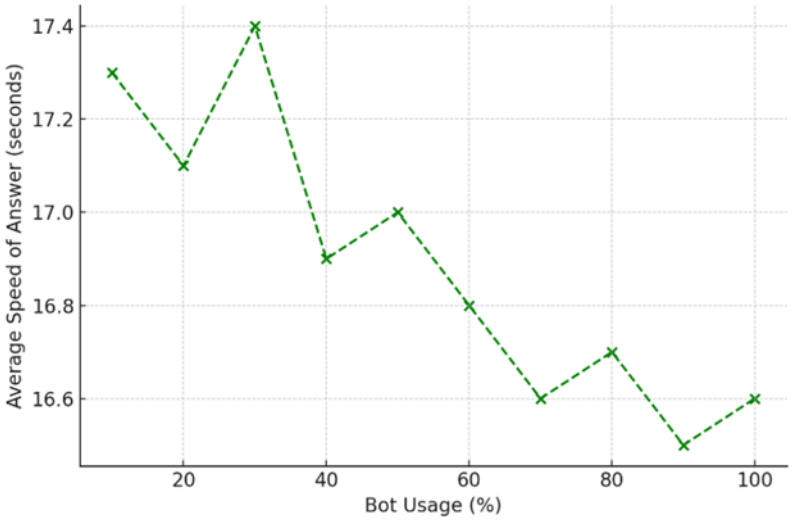

Figure 6b: Correlation – Bot Usage vs. Human ASA
Source: author's work.

A weak and statistically insignificant correlation was observed ($r = 0.10$, $p = 0.77$), suggesting that increasing chatbot deployment does not significantly affect wait times for human-assisted interactions.

This result implies that ASA in human queues is not automatically improved by increasing chatbot presence. Rather, it is moderated by independent factors such as queue configuration, agent availability, peak-hour demand, or escalation frequency.

These insights support the findings in ASA and Interaction Queue Dynamics and align with Dwivedi et al (2023), who argue that automation may reduce interaction load but not necessarily improve system-level responsiveness without complementary staffing optimization.

In summary, while chatbot adoption clearly enhances handling efficiency for human agents, its impact on customer wait times remains limited without systemic queue reengineering and dynamic staff management. The dual-figure correlation analysis reinforces the study's emphasis on holistic, hybrid service models where automation and human resource design are integrated rather than siloed.

4.9. Customer Experience and Perception Analysis

Customer experience is a pivotal dimension in evaluating the effectiveness of AI-driven automation in service settings. To explore this aspect, both survey responses and qualitative insights (from interviews and focus groups) were analyzed, focusing on perceived usefulness, satisfaction, service preference, and loyalty (Figure 7).

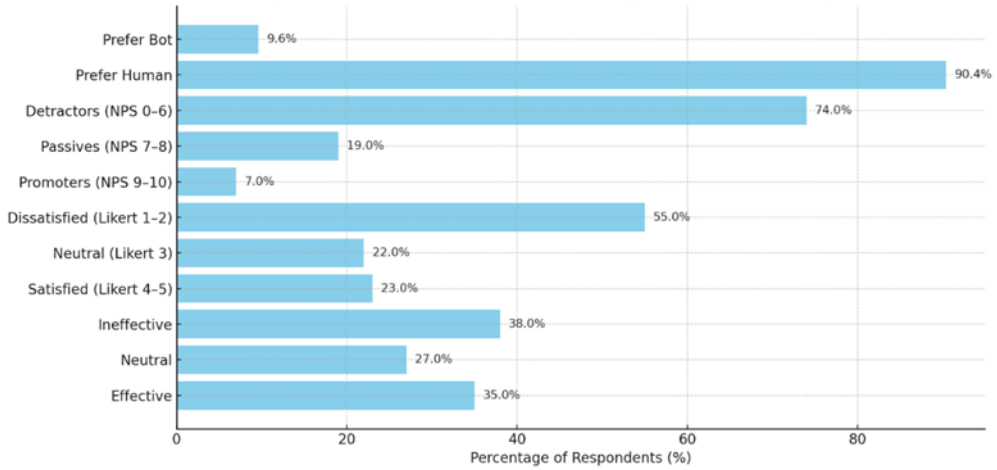

Figure 7. Distribution of Customer Perceptions on Chatbot Experience
Source: author's work.

This figure visualizes the results of user feedback related to chatbot effectiveness, satisfaction ratings, Net Promoter Score (NPS) segments, and service preference between bot and human agents.

Perceived Effectiveness

Respondents were asked whether the chatbot effectively resolved their queries:

- 35% rated the bot as effective
- 27% were neutral
- 38% found the bot ineffective

This mixed perception reflects a common pattern in AI-mediated interactions: users appreciate chatbots for handling routine inquiries, but often express frustration when issues require contextual or emotional understanding.

Satisfaction Scores (Likert Scale 1–5)

- Only 23% of users gave satisfaction ratings of 4–5, indicating high approval.
- 22% remained neutral, while 55% were dissatisfied (scores 1–2).

This suggests that despite operational benefits, the chatbot service underperformed in meeting user expectations, particularly in areas such as empathy, problem resolution, and escalation protocols.

Net Promoter Score (NPS)

The NPS classification revealed:

- 7% were Promoters (scores 9–10)
- 19% were Passives (scores 7–8)
- 74% were Detractors (scores 0–6)

This distribution reflects a low loyalty index and confirms that users are unlikely to recommend chatbot services in their current form. NPS is highly sensitive to perceived service failures, particularly when chatbots are perceived as obstacles rather than enablers.

Service Preference

One of the most conclusive indicators of customer sentiment was service preference:

- 90.4% preferred human agents
- Only 9.6% preferred chatbots

This overwhelming preference for human support, despite acknowledging chatbot availability and speed, highlights the enduring role of empathy, flexibility, and problem-solving in customer satisfaction. These results echo global findings by Kulkarni et al. (2023), which underscore the emotional gap in AI-mediated service.

The results depicted in Figure 4 provide multi-layered evidence that while chatbots may enhance operational metrics, they do not yet fulfill emotional, cognitive, and situational service needs. The sharp contrast between perceived effectiveness (35%) and preference for human agents (90.4%) suggests that even well-performing chatbots may be undervalued when they fail to meet relational expectations.

These findings reinforce the need for:

- Hybrid service architectures, where chatbots handle routine tasks and humans

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

manage exceptions or emotionally nuanced cases.

- Trust-building and transparency strategies, including clearer escalation paths and feedback loops.
- Human-centered design, leveraging user feedback to improve natural language understanding, sentiment detection, and adaptive responses.

Together, these insights align with the broader literature and strengthen the study's recommendation for integrating AI with empathy-focused human interaction models.

5. Discussion

This study set out to assess the role of chatbot deployment in enhancing operational efficiency, reducing costs, and shaping customer experience in Costa Rican financial contact centers. The findings align with three of the four original hypotheses: (1) automation significantly improves internal processing efficiency (AHT and FTE), (2) it leads to measurable cost reductions, and (3) its relationship with customer satisfaction is indirect and moderated by other variables. However, the hypothesis that automation would reduce ASA was not supported, offering valuable insights into the limitations of AI integration in service systems.

One of the clearest outcomes of the study was the marked improvement in operational performance. The 43.7% reduction in AHT and the 26.6% reduction in indirect operational costs are statistically and practically significant. These findings were supported by:

- Multivariate regression analysis indicates that bot usage strongly predicts AHT improvement ($\beta = -2.45$, $p = 0.005$),
- ARIMA modeling, which revealed a structural break post-implementation ($\beta = -75.3$, $p = 0.01$), and
- SEM results, where both cost reduction and AHT were positively associated with customer satisfaction ($\beta = 0.44$ and $\beta = 0.29$, respectively).

These efficiency gains validate the literature emphasizing automation's potential to absorb repetitive workload (SAP, 2023; Uzoka et al., 2024). In the Costa Rican context—where institutions operate under increasing digital expectations and cost sensitivity—this operational streamlining demonstrates a scalable and sustainable benefit.

Despite the internal improvements, ASA for human interactions remained unchanged (17.2 to 14.15 seconds; $p > 0.05$), confirmed by both correlation analysis ($r = 0.10$, $p = 0.77$) and ARIMA modeling ($\beta = -1.9$, $p = 0.33$). The post-hoc power analysis (power = 0.84) confirmed that this result is not due to insufficient sample

size.

This reinforces a key insight: automation does not necessarily lead to quicker human response. ASA is dependent on staffing ratios, escalation frequency, queue routing, and time-of-day variability. This finding is consistent with Dwivedi et al (2023), who argue that ASA improvements require systemic queue redesign and dynamic workforce allocation, not just digital augmentation (Okot et al., 2025).

While chatbots succeeded in reducing wait times and handling durations, they failed to meet customer expectations in qualitative terms. Only 23% of users reported satisfaction (Likert 4–5), while 74% were NPS detractors, and 90.4% preferred human support. These metrics reflect an emotional and experiential disconnect between efficiency and perceived value (Lu et al., 2020).

This gap can be attributed to several key limitations in chatbot design and deployment:

- Lack of affective responsiveness (empathy, tone adaptation),
- Inadequate escalation mechanisms for complex queries, and
- Limited personalization or contextual recall across sessions.

The SEM model confirms that while AHT and cost improvements have a statistically significant indirect effect on satisfaction, there is no direct pathway from bot usage to user loyalty. This aligns with Kulkarni et al. (2023), who argue that perceived authenticity and emotional intelligence remain central to digital trust in AI-mediated services.

Triangulation across statistical models, time-series trends, and qualitative responses strengthens the validity of this study's conclusions. The qualitative findings confirmed that many users:

- Recognized the efficiency of the bot for basic tasks,
- Resented repeated "looping" behavior in complex scenarios, and
- Felt disconnected from the institution when escalation paths were unclear.

These sentiments reinforce the view that customer satisfaction cannot be optimized through automation alone. Hybrid systems where chatbots triage and support agents, rather than replace them, represent a more psychologically acceptable service model (Cheng et al., 2023).

From a theoretical standpoint, this study contributes to the SDL framework by showing that AI technologies create value indirectly—by optimizing operational enablers rather than directly satisfying customer expectations (Okot & Jarquin, 2025).

The SEM reflects a mediated structure of customer satisfaction, where operational

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

gains must be translated through human-centric channels to impact experience outcomes (Jiang et al., 2023). This model can guide future studies on AI-human interaction in service ecosystems, especially in developing regions where emotional labor and cultural service expectations remain high.

Based on the evidence, the following actions are recommended for institutions implementing chatbots:

- Queue and Staff Redesign: Introduce dynamic staffing algorithms that respond in real time to bot-initiated escalation demand to reduce ASA in human queues.
- Emotionally Aware AI: Integrate affective computing capabilities, such as sentiment detection and tone modulation, to personalize responses and improve trust.
- Hybrid Service Architecture: Chatbots should not be siloed tools but part of a collaborative service model. Intelligent routing and human handoff must be seamless.
- Expand KPIs Beyond Efficiency: Track customer-centered metrics such as trust, escalation rate, and perceived helpfulness, not just AHT or cost.
- Customer Education and Expectation Management: Train users in what chatbots can and cannot do. Managing expectations reduces dissatisfaction when escalations are necessary.

While the results are robust and methods rigorous, some limitations exist:

- Generality: The study was conducted in a single institution, limiting broad generalization.
- Duration: A 12-month period captures short-term efficiency but not long-term adaptation.
- Survey bias: Satisfaction responses may skew toward more vocal users or those with extreme experiences.

Future research should explore:

- Cross-national or multi-institutional comparisons, especially in other middle-income countries,
- User behavior over time, using log-based user session data to track learning curves and fatigue,
- Generative AI integration, as LLM-based chatbots (e.g., ChatGPT) may offer stronger adaptability and emotional nuance.

Applied Implications and Real-World Enhancements

To translate the study's findings into actionable strategies, financial institutions should prioritize hybrid AI-human service designs that integrate emotional intelligence with operational efficiency. One promising approach involves NLP-based query classification, which enables real-time triage of customer intents. For

instance, routine tasks like balance inquiries or password resets can be instantly resolved by chatbots, while emotionally nuanced or complex queries, such as loan restructuring or fraud disputes, are escalated to trained human agents with pre-filled context summaries. This model is already in use by banks in Singapore and Germany, reducing average escalation time by over 30% (Huang & Rust, 2021).

To educate customers about financial risks, institutions can implement chatbot-led microlearning modules. For example, when a user inquires about a loan, the bot can briefly explain interest rate variability or repayment scenarios using simple, interactive visuals before routing them to a human agent for personalized advice. Such educational nudges enhance digital literacy and reduce the overload on call centers by preempting repetitive questions.

Avoiding information overload is also crucial. Instead of offering long responses, modern chatbots should apply progressive disclosure, revealing information in short, decision-friendly snippets. For instance, bots can present users with quick options like "Do you want to know more about this fee? [Yes/No]," which keeps the interaction efficient and reduces cognitive fatigue.

However, as bots become more sophisticated through LLMs, such as ChatGPT, there is a growing risk of blurred human-AI boundaries. Users may assume they are interacting with a human agent, leading to mistrust when they discover otherwise, particularly if errors occur or empathy is lacking. Therefore, chatbot systems must disclose their non-human nature clearly, provide options to escalate, and avoid over-personification. Transparency about the bot's capabilities and limitations can help sustain trust, especially in high-stakes financial contexts.

Institutions should also consider "ethical AI audits" to monitor how chatbots manage sensitive data, avoid bias, and ensure fair escalation protocols. These measures not only address usability but also safeguard institutional reputation in a sector where trust is paramount.

6. Conclusions and recommendations

This study comprehensively evaluated the effects of chatbot implementation on operational efficiency, cost structure, and customer experience within a Costa Rican financial contact center, employing a robust mixed-methods framework that included statistical modeling (regression, ARIMA, SEM), survey analysis, and qualitative interviews.

The findings underscore a dual reality in AI-mediated service transformation:

- On the one hand, chatbot deployment yielded significant operational benefits.

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

AHT fell by 43.7%, and indirect operational costs dropped by 26.6%, supporting the hypothesis that automation reduces resource intensity. These gains were statistically validated through multivariate regression, ARIMA-based time-series analysis, and the SEM.

- On the other hand, improvements in the ASA for human-handled interactions were non-significant, and customer satisfaction remained low, with only 23% of users reporting high satisfaction, 74% as NPS detractors, and 90.4% still preferring human interaction.

This duality confirms that automation enhances efficiency but not necessarily user experience unless supported by strategic system design, emotional responsiveness, and hybrid service protocols.

This study advances theoretical understanding of AI's role in digital service ecosystems in three key ways:

- First, by validating automation's indirect effects on satisfaction through operational mediators (AHT and cost), it contributes to the SDL and technology acceptance theory, where value creation is not solely based on system functionality but on user-perceived enablement.
- Second, the SEM model provides a clear framework for conceptualizing how operational efficiencies translate into perceived quality, particularly in emerging markets with high relational expectations.
- Third, by integrating quantitative and qualitative triangulation, the study offers a methodological template for evaluating technological interventions holistically.

For institutions in Costa Rica and similar contexts aiming to implement or scale AI-enabled services, the findings offer a roadmap to maximize value while minimizing user resistance:

- Efficiency gains require complementary human system design: While chatbots significantly reduce the AHT, these improvements do not automatically translate into better ASA unless queue management systems and human staffing models are also optimized. Chatbot deployment must be integrated with real-time human resource planning to achieve system-wide efficiency.
- Customer satisfaction is mediated, not directly achieved through automation: Even when chatbots operate flawlessly on a technical level, users only report high satisfaction when the overall service experience is transparent, adaptive, and relational. Trust and loyalty are built not through speed alone, but through empathy, clarity, and perceived attentiveness.
- Sustainability synergies exist but remain underutilized: Automation reduces the need for physical infrastructure and agent presence, potentially lowering the

institution's carbon footprint. In countries like Costa Rica, where electricity is primarily generated from renewables, there is an opportunity to lead with a "green AI" model that aligns chatbot deployment with national sustainability goals and supports innovation through public-private collaboration.

Based on empirical insights and regional considerations, the following strategic and policy-level recommendations are proposed:

6.1. For Financial Institutions and Contact Centers

a) Implement Dynamic Workforce Allocation Systems: Use historical and real-time data to reallocate agents based on bot escalation frequency. A chatbot that reduces volume but increases complexity at human queues may inadvertently increase ASA variance.

b) Embed Human-Centered AI Design Principles: Develop chatbots with conversational fluency, regional language nuances, and sentiment analysis to humanize interactions. Affective AI frameworks can dramatically improve user perception without sacrificing efficiency.

c) Track Beyond Traditional KPIs: Operational KPIs like AHT and cost are insufficient. Incorporate metrics such as first-contact resolution (FCR), escalation ratio, and emotional sentiment score to evaluate service effectiveness more comprehensively.

d) Build Multi-Stage Escalation Models: Introduce intermediate escalation paths (e.g., live chat with a senior bot or pre-screened human agent) before full handoff. This reduces agent overload and prevents early user frustration.

6.2. For Policymakers and Regulators

a) Establish AI Accountability Guidelines: Require transparency in automated decision-making, especially for customer-facing chatbots in regulated industries. Users should be informed when interacting with AI and how to access human support.

b) Support Public-Private Innovation for AI Ethics: Encourage joint development of ethical AI standards in finance and customer service. Topics should include data privacy, algorithmic bias mitigation, and inclusive language modeling.

c) Promote Workforce Upskilling for Hybrid Models: As chatbots replace repetitive tasks, agents must be retrained for advisory, empathetic, and complex resolution roles. Incentivize institutions that invest in human-AI collaboration rather than displacement.

d) Fund Longitudinal and Regional AI Studies: Given the emergent nature of AI in service delivery, regulators should support multi-institution, multi-year evaluations

Studia Universitatis "Vasile Goldis" Arad. Economics Series Vol 36 Issue 3/2026
ISSN: 1584-2339; (online) ISSN: 2285 – 3065
Web: publicatii.uvvg.ro/index.php/studiaeconomia. Pages 92-127

to monitor unintended consequences and guide evidence-based policymaking.

Despite the study's strengths, several limitations constrain its scope:

- Single-institution analysis may limit generalizability across other sectors or geographic regions.
- The 12-month window, while sufficient to detect operational shifts, may be insufficient for long-term satisfaction and loyalty dynamics.
- Self-reporting bias in the survey data may underrepresent less digitally literate populations, who are more likely to struggle with chatbot interactions.
- To address these limitations, future research should consider:
- Cross-sectoral comparisons (e.g., telecommunications, health services) to test transferability,
- Behavioral analytics using chatbot interaction logs to capture user friction and dropout points,
- Generative AI adoption (e.g., GPT-based chatbots) to assess whether advanced language models can bridge the empathy-performance gap.

Main Added Value of the Paper:

This study contributes original empirical insight into the integration of chatbot automation in culturally relational service environments. Its primary added value lies in demonstrating how hybrid AI-human service models can simultaneously improve operational efficiency and maintain customer trust in financial institutions. By combining structural equation modeling (SEM), ARIMA, regression analysis, and qualitative triangulation, the paper delivers a comprehensive framework for evaluating AI implementation in emerging markets where emotional intelligence and human-centered design remain pivotal.

Applicability of the Case Study to Similar Countries:

Although the research is situated in Costa Rica, its findings apply to other small, service-oriented economies with strong human-interaction norms, such as Colombia, the Dominican Republic, or the Philippines. These countries share common features: high digital adoption, cultural preference for personal service, and increasing pressure to modernize contact centers. The study's emphasis on trust-building, ethical design, and hybrid escalation protocols offers transferable strategies for chatbot deployment across Latin America, Southeast Asia, and other Global South regions with similar socioeconomic profiles.

Disadvantages and Limitations of the Research Methods:

Despite methodological rigor, several limitations must be acknowledged. First, the mixed-methods design may be perceived as over-engineered, given that many

findings were confirmatory in nature. Second, the short time series limits long-term forecasting reliability, although ARIMA helped reveal clear structural changes. Third, the survey component relied on self-reported user satisfaction, which can introduce bias, particularly in digital self-service settings where frustration may be underreported by disengaged users. Lastly, the single-institution scope limits external generalizability, though it allows for contextual depth.

Impact on the Economy and Community:

The study has significant implications for both economic and social development. On the economic front, the demonstrated cost savings and productivity gains offer a scalable roadmap for digital transformation in financial services without increasing headcount. Environmentally, the alignment with Costa Rica's renewable energy infrastructure promotes a "green AI" agenda, enhancing sustainability in digital operations. At the community level, the research encourages public-private collaboration in developing ethical, inclusive AI tools that preserve service equity while promoting digital innovation. Such approaches strengthen institutional trust, financial inclusion, and consumer empowerment across underserved segments.

Acknowledgements

The authors thank the anonymous reviewers and editor for their valuable contributions.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not – for – profit sectors.

Author Contributions

All authors contributed to the study's conception and design. All authors read and approved the final manuscript.

Okot, Salas: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft preparation, Writing – review and editing, Visualization, Project administration.

Disclosure Statement

The authors do not have any competing financial, professional, or personal interests from other parties.

References

1. Accenture. (2023). *The rise of chatbots in Latin America: Driving customer engagement and efficiency.* Accenture Insights. https://www.accenture.com/us-en/insights/banking/chatbots-latin-america
2. Arbuckle, J. L. (2022). *IBM SPSS Amos 29 User's Guide*. IBM Corporation.
3. Bazrkar, A., Moradzad, M., & Shayegan, S. (2024). The use of artificial intelligence in employee recruitment in the furniture industry of Iran, according to the role of contextual factors. *Studia Universitatis Vasile Goldiș Arad, Seria Științe Economice*, *34*(2), 86-109. https://doi.org/10.2478/sues-2024-0009
4. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2023). On the dangers of stochastic parrots: Can language models be too big? *Communications of the ACM*, 66(3), 56–65. https://doi.org/10.1145/3568252
5. Cheng, J., Navarro, C., & López, D. (2023). Hybrid customer service models in Latin American banking. *Journal of Financial Services Technology*, 15(2), 77–95. https://doi.org/10.1016/j.jfst.2023.04.005
6. Chinnaraju, A. (2025). Explainable AI (XAI) for trustworthy and transparent decision-making: A theoretical framework for AI interpretability. *World Journal of Advanced Engineering Technology and Sciences*, *14*(3), 170-207. https://doi.org/10.30574/wjaets.2025.14.3.0106
7. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. https://doi.org/10.2307/249008
8. Dwivedi, Y. K., Hughes, D. L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., & Williams, M. D. (2023). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice, and policy. International *Journal of Information Management*, 71, 102456. https://doi.org/10.1016/j.ijinfomgt.2022.102456
9. Eustaquio-Jiménez R, DurandAzurza M, Gamboa-Cruzado J, et al. (2024). Chatbots for customer service in financial entities—A comprehensive systematic review. *Journal of Infrastructure, Policy and Development.* 8(16): 10122. https://doi.org/10.24294/jipd10122
10. Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149
11. Følstad, A., & Taylor, C. (2021). Investigating the user experience of customer service chatbot interaction: a framework for qualitative analysis of chatbot dialogues. *Quality and User Experience*, *6*(1), 6. https://doi.org/10.1007/s41233-021-00046-5
12. Fonseca, A. (2022). The benefits and perils of digital and automated technologies: impact assessment methods in the fourth industrial revolution. *Handbook of Environmental Impact Assessment*, 126-145. https://doi.org/10.4337/9781800379633.00013

13. Gartner. (2023). *Forecast analysis: AI in banking and investment services, worldwide*. Gartner Research. https://www.gartner.com/en/documents/4107819
14. Haugeland, I. K. F., Følstad, A., Taylor, C., & Bjørkli, C. A. (2022). Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction design. *International Journal of Human-Computer Studies*, *161*, 102788. https://doi.org/10.1016/j.ijhcs.2022.102788
15. Hsu, C. L., & Lin, J. C. C. (2023). Understanding the user satisfaction and loyalty of customer service chatbots. *Journal of Retailing and Consumer Services*, *71*, 103211. https://doi.org/10.1016/j.jretconser.2022.103211
16. Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science, 49*(1), 30–50. https://doi.org/10.1007/s11747-020-00749-9
17. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts. https://otexts.com/fpp2/
18. Inter-American Development Bank (IDB). (2023). *Digital banking adoption in Latin America: Insights from Costa Rica*. IDB Publications. https://publications.iadb.org/en/digital-banking-latin-america
19. Jiang, N., Liu, X., Liu, H., Lim, E. T. K., Tan, C. W., & Gu, J. (2023). Beyond AI-powered context-aware services: the role of human–AI collaboration. *Industrial Management & Data Systems*, *123*(11), 2771-2802. https://doi.org/10.1108/IMDS-03-2022-0152
20. Kulkarni, A., Arora, M., & Johnson, D. (2023). Emotional intelligence and trust in AI-powered customer interactions. *Journal of Service Research*, 26(1), 55–71. https://doi.org/10.1177/10946705221076147
21. Liao, Q. V., Gruen, D. M., & Miller, S. (2023). Questioning the AI: Informing design practices for explainable AI user experiences. *Proceedings of the ACM on Human-Computer Interaction, 7*(CSCW2), 1–24. https://doi.org/10.1145/3610495
22. Lu, V. N., Wirtz, J., Kunz, W. H., Paluch, S., Gruber, T., Martins, A., & Patterson, P. G. (2020). Service robots, customers, and service employees: what can we learn from the academic literature, and where are the gaps?. *Journal of Service Theory and Practice*, *30*(3), 361-391. https://doi.org/10.1108/JSTP-04-2019-0088
23. Mariani, M. M., Borghi, M., & Kazlauskaitė, R. (2023). Chatbots in customer service: A review and research agenda. *Journal of Business Research, 158*, 113632. https://doi.org/10.1016/j.jbusres.2022.113632
24. McLean, G., & Osei-Frimpong, K. (2021). Examining satisfaction with the experience during a live chat service encounter—Implications for customer retention. *Psychology & Marketing, 38*(2), 275–289. https://doi.org/10.1002/mar.21328
25. Okot, T., & Jarquin, M. (2025). The Impact of Smart Digitalization Technologies on Mitigating Incremental Risks in Sustainable Sourcing Practices. *Journal of Technology Management & Innovation*, *20*(1).

26. Okot, T., Ramirez, M., & Deckert, R. (2025). Employer Branding as a Catalyst for Workforce Diversity and Excellence: An In-Depth Study from Costa Rica. *Organizacija*, 58(1), 64-84. https://doi.org/10.2478/orga-2025-0005
27. Pandow, B. A. (2024). Revolutionizing Finance: The Impact of AI-Driven Innovations. In *Navigating the Future of Finance in the Age of AI* (pp. 1-24). IGI Global. https://doi.org/10.4018/979-8-3693-4382-1.ch001
28. SAP. (2023). *Chatbot deployment metrics in European banks*. SAP Whitepaper. https://www.sap.com/documents/2023/03/3f96f3a0-207e-0010-bca6-c68f7e60039b.html
29. Seymour, M., Yuan, L., Riemer, K., & Dennis, A. R. (2024). Less artificial, more intelligent: understanding affinity, trustworthiness, and preference for digital humans. *Information Systems Research*. https://doi.org/10.1287/isre.2022.0203
30. Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International journal of human-computer studies*, *146*, 102551. https://doi.org/10.1016/j.ijhcs.2020.102551
31. Stryzhak, O. (2023). Analysis of labor market transformation in the context of Industry 4.0. *Studia Universitatis Vasile Goldiş, Arad-Seria Ştiinţe Economice*, *33*(4), 23-44. https://doi.org/10.2478/sues-2023-0017
32. Susanto, A., & Khaq, H. (2024). Reducing customer handling time through AI chatbots. *Journal of Artificial Intelligence Research in Business*, 19(1), 22–39. https://doi.org/10.1016/j.jaibrb.2024.01.003
33. Susanto, E., & Khaq, Z. D. (2024). Enhancing customer service efficiency in start-ups with AI: A focus on personalization and cost reduction. *Journal of Management and Informatics*, *3*(2), 267-281. https://orcid.org/0009-0001-3072-5825
34. UNEP. (2023). *AI and sustainability: Guidelines for green digital transformation*. United Nations Environment Programme. https://www.unep.org/resources/report/ai-and-sustainability-guidelines
35. Uzoka, A., Cadet, E., & Ojukwu, P. U. (2024). Leveraging AI-Powered chatbots to enhance customer service efficiency and future opportunities in automated support. *Computer Science & IT Research Journal.* 5(10), 2485–2510.
36. Vargo, S. L., & Lusch, R. F. (2008). Service-dominant logic: Continuing the evolution. *Journal of the Academy of Marketing Science*, 36(1), 1–10. https://doi.org/10.1007/s11747-007-0069-6
37. Xu, Y., Zhang, J., & Deng, G. (2022). Enhancing customer satisfaction with chatbots: The influence of communication styles and consumer attachment anxiety. *Frontiers in Psychology*, *13*, 902782. https://doi.org/10.3389/fpsyg.2022.902782
38. Zewe, A. (2025). Explained: Generative AI's environmental impact. *MIT News*. https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117news.mit.edu