Unlocking Business Value with Generative AI! Economic Value Assessment for Chatbots and Gen AI ROI Discovery

Jim Samuel
Rutgers, The State University of New Jersey

Manideep Pendyala
Rutgers, The State University of New Jersey

Kushal Patel
Rutgers, The State University of New Jersey

Unlocking Business Value with Generative AI! Economic Value Assessment for Chatbots and Gen AI ROI Discovery

Jim Samuel
Bloustein, Rutgers University
AIXosphere
New Brunswick, NJ 08901

Julia Esguerra
Bloustein, Rutgers University
33 Livingston Ave,
New Brunswick, NJ 08901

Udayan Bose
NetElixir
500 College Road East, Suite #410,
Princeton, NJ 08540

Manideep Pendyala*
Bloustein, Rutgers University
33 Livingston Ave,
New Brunswick, NJ 08901

Esha More
Bloustein, Rutgers University
33 Livingston Ave,
New Brunswick, NJ 08901

Kushal Patel
Bloustein, Rutgers University
33 Livingston Ave,
New Brunswick, NJ 08901

Rimshah Jawad
Bloustein, Rutgers University
33 Livingston Ave,
New Brunswick, NJ 08901

Accepted at NBEA 2025, Cite as below (APA) :

Samuel, J., Esguerra, J., Bose, U., Pendyala, M., More, E., Patel, K., & Jawad, R. (2025). Unlocking business value with generative AI! Economic value assessment for chatbots and Gen AI ROI discovery. The 52nd Annual Northeast Business & Economics Association (NBEA) Conference, Seaview Hotel, Galloway, NJ, USA.

ABSTRACT
The rapid rise of generative artificial intelligence (AI) and generative AI-based chatbots has prompted growing concerns over their true economic value. While industry sources are reporting projected or anticipated productivity gains and cost savings, a dependable analysis of these claims remains absent. This paper reviews extant research, economic studies, industry reports, and other pertinent information to evaluate the impacts of generative AI and chatbots, especially on return on investment (ROI) measures. We offer an analytical view of the financial impacts of generative AI, especially chatbots, with cases and pricing information from generative AI service providers. Our goal is to provide policymakers, business leaders, and researchers with an exploratory understanding of how business value can be unlocked with generative AI. This is an abbreviated version of the paper to meet the page limit of the conference.

Keywords
Generative Artificial Intelligence, Chatbots, Return on Investment

1 INTRODUCTION

Generative artificial intelligence (Gen AI) based chatbots are conversational agents that generate human-like text for tasks ranging from customer service to content creation. These systems mimic human reasoning and language dynamics to produce original responses (Samuel, 2021). As capabilities improve, Gen AI is reshaping business operations and the broader landscape of information creation and consumption. Gen AI has gained attention for its potential to make work faster and more cost-efficient (Samuel et al., 2024). At the center of these expectations lies a compelling economic proposition: if the marginal cost of automated conversations drops to fractions of a cent, large-scale value creation becomes feasible, so long as chatbot performance can effectively augment or replace human labor (Samuel, 2023). Despite this potential, concerns persist. Building and operating advanced chatbots remains costly, and whether they consistently deliver sufficient returns is uncertain. This paper examines the economic implications of Gen AI chatbots by synthesizing findings from academic research, industry reports, and government assessments. As these systems expand globally, ensuring accuracy, transparency, and adaptability across contexts, such as real-time updates, diverse languages, and shifting public sentiment, will be essential to realizing their full economic and social value (Chidipothu et al., 2025; Samuel et al., 2020a; Samuel et al., 2020b; Anderson et al., 2024).

The paper proceeds as follows: Section 2 reviews research on chatbot adoption, productivity, and policy issues. Section 3 outlines the cost framework and methods, followed by pricing and ROI findings in Section 4. Section 6 identifies research gaps, and Section 7 concludes with insights for policymakers and industry. Empirical data and forecast-based assumptions are clearly distinguished throughout.

2 LITERATURE REVIEW

Gen AI chatbots are powered by large language models (LLMs), which leverage vast amounts of unstructured textual data and transformer-based architectures to generate human-like responses to user queries and can differ in openness and deployment flexibility (Chidipothu et al., 2025). McKinsey estimates that Gen AI could add $2.6 to $4.4 trillion annually across sectors such as customer service, software development, and marketing, but such productivity gains may not be evenly distributed (McKinsey & Company, 2023). Beyond adoption and labor, media framing also matters as it directly affects the user attitude and organizational readiness. A recent study on AI-phobia applies topic modeling and fear-sentiment classification (BERT, LLaMA, Mistral) to approximately 70,000 news headlines drawn from various sources and shows that recurrent fear frames shape public perception on AI (Samuel et al., 2025). Even a minority of alarmist headlines can have outsized effects on attitudes and behavior, creating ROI risk for customer-facing deployments that rely on trust.

3 ECONOMIC FRAMEWORKS, METHODS, AND PRICING

Current Chatbot Pricing Models: Common approaches include (i) Subscription SaaS (e.g., $20-$30 per user/month for individual pro tiers), with enterprise variants (OpenAI, n.d.-a; Perplexity, n.d.); (ii) pay-as-you-go APIs (token/character based) for cost control based on use (Anthropic, n.d.-a); (iii) self-hosted open-weights models (eg., LLaMA 2, Mistral) with GPU and engineering costs (Schmid, 2023); and (iv) seat based enterprise add-ons (eg., Microsoft 365 Copilot at $30/user/month, Salesforce Einstein GPT at $50-70/user/month) (Microsoft, n.d.; TechTarget, 2024; Salesforce, 2023; Salesforce, n.d.).
Self-Hosting trade-offs: Open models reduce vendor fees but require infrastructure and MLOps. Cloud services (e.g., Amazon Bedrock) price open models at substantially lower per-token rates than frontier models (AWS, n.d).
Deployment Overheads: Integrating, hosting, monitoring, moderation, and safety add 10-30% over raw model costs at scale (Ever Efficient AI, 2024). Caching and retrieval strategies reduce cost/latency (Sinha, 2025). Moderation layers mitigate risk (Greyling, 2023).

4 ECONOMIC IMPACT

Klarna Customer Service: Klarna, a fintech company, has integrated OpenAI’s GPT models to improve customer service. A large-scale AI assistant (OpenAI, n.d.-b) handled 2.3M conversations in its first month, reaching ~67% of service chats and providing human-like performance, leading to resolution time falling from approximately 11 minutes to <2 minutes.
Illustrative Banking Scenario: For a bank of 1M customers, if 20% of consumers utilise the chatbot every month and each person asks an average of five questions, we may estimate the following costs, assuming 1 million customer queries per month and a chatbot deflection rate of 70% that is 700,000 interactions handled by the bot instead of live agents, and given an average cost of $7 per agent-handled interaction, the gross monthly savings from deflection are about $4.9 million; under a conservative realisation assumption of 10% of projected savings, this equates to approximately $490,000 per month. Industry reports indicate that banks deploying AI assistants commonly achieve deflection rates of 50% or more (Deloitte, 2023).

5 DISCUSSION

Cost Per Interaction: With smaller models, per-query costs can fall below a cent; with frontier models, a few cents is common, far below human-handled interactions (Desk365.io, 2025). **Value of 24/7 Availability**: Always-on chatbots reduce wait times, enabling humans to focus on more complex tasks. Process redesign with AI can achieve up to 25% cost savings while improving quality (Bain & Company, 2024). **Rapid Scaling and Flexibility:** Chatbots scale elastically to spikes in user demand, but guardrails, escalation paths, and SLAs are essential to preserve trust and brand safety. **Limitations, Risks, and Mitigations:** Hallucinations, domain errors, and compliance add monitoring costs; nonetheless, they are typically lower than full human handling when governance is in place. One of our detailed case studies illustrates this powerful potential for **Gen AI-driven ROI and organizational transformation** effectively: NetElixir's Gen AI program, launched in 2023, embedded an RAG-secured OpenAI API into its channel-marketing workflows and automated nearly half of the paid search activities, cutting cycle time by 30% in six months. The resulting labor savings funded a new experimentation unit, and across 50 e-commerce clients, the system achieved a 9% lift in campaign revenue while holding media costs constant, yielding a 7:1 ROI (Netelixir, 2025). This case demonstrates how Gen AI can simultaneously drive operational efficiency and revenue growth, validating its broader strategic impact.

6 FUTURE RESEARCH

Future research should focus on developing a standard ROI framework that captures hidden integration and compliance costs, monitors labor shifts, and closes pricing or performance gaps for low-resource languages and small firms. Publish open cost-performance data from real-world deployments to move from anecdotes to evidence. Moreover, research must integrate culturally adaptive and human-centered AI education principles, such as those articulated in the CATE-AI framework, to ensure AI systems and their evaluations are contextually relevant, culturally sensitive, and personally adaptive to diverse user needs (Samuel et al., 2023; Kashyap et al., 2024).. Finally, exploring AI's role in augmenting human cognitive fit could illuminate how adaptive AI representations improve decision-making performance, offering valuable insights for refining cost-benefit models and regulatory approaches in dynamic socio-technical environments (Samuel et al., 2022). Future studies should explore how these adaptations intersect with sentiment variation, digital infrastructure, and regulatory mandates, ultimately shaping effective, inclusive, and context-aware AI systems (Ali et al., 2021; Rahman et al., 2021; Pendyala et al., 2025).

7 CONCLUSION

Taken together, the results of this research imply cautious optimism. Organizations that (i) select use cases with clear cost displacement, (ii) budget realistically for integration, (iii) invest in upskilling, and (iv) manage the risks and limits of Gen AI, can expect to generate positive ROI. Conversely, projects that underestimate hidden costs or overestimate Gen AI risk value erosion. For the future, one of the most important dimensions for Gen AI, and for all AI across ecosystems, is the 'Human Enhancive AI' (HEAI) strategy, which goes beyond the much needed human centered AI (HAI or HCAI) approach, as HEAI emphasizes placing humans above AI on multiple dimensions. NetElixir, as discussed earlier in this research, demonstrates one of the dimensions of HEAI by following a strategy of repurposing humans for higher level tasks, based on Gen AI driven efficiency gains on lower level tasks. It is hoped that as human institutions and society march ahead, unlocking increasingly higher levels of business value and socioeconomic progress with Gen AI, human enhancive AI approaches are prioritized to ensure that generative AI benefits all humans equitably.

8 ACKNOWLEDGEMENTS

This project was funded by the Master of Public Informatics (MPI) program, Bloustein School, Rutgers University (https://informatics.rutgers.edu/), and the Rethinking AI for Shared Empowerment (RAISE) project (https://sites.rutgers.edu/raise/).

9 REFERENCES

Samuel, J. (2021). A call for proactive policies for informatics and artificial intelligence technologies. Scholars Strategy Network. Retrieved June 15, 2025, from https://scholars.org/contribution/call-proactive-policies-informatics-and

Samuel, J., Tripathi, A. and Mema, E., (2024). A new era of artificial intelligence begins… where will it lead us? Editorial - Journal of Big Data and Artificial Intelligence, Volume 2 Nu. 1.

Samuel, J. (2023). The Critical Need for Transparency and Regulation amidst the Rise of Powerful Artificial Intelligence Models, Scholars Strategy Network (SSN) Key Findings. Retrieved June 15, 2025, from: https://scholars.org/contribution/critical-need-transparency-and-regulation

Chidipothu, N., Anderson, R., Samuel, J., Pelaez, A., Esguerra, J., & Hoque, M. N. (2025). Improving Large Language Model (LLM) Performance with Retrieval Augmented Generation (RAG): Development of a Transparent Generative AI University Support System for Educational Purposes. Journal of Big Data and Artificial Intelligence,

3(1).
Samuel, J., Ali, G. G., Rahman, M., Esawi, E., & Samuel, Y. (2020a). Covid-19 public sentiment insights and machine learning for tweets classification. Information, 11(6), 314.
Samuel, J., Rahman, M., Ali, Nawaz G. G. Md., Samuel, Y., Pelaez, A., Chong, P. H. J. and Yakubov, M (2020b) "Feeling Positive About Reopening? New Normal Scenarios From COVID-19 US Reopen Sentiment Analytics," in IEEE Access, vol. 8, pp. 142173-142190, 2020, doi: 10.1109/ACCESS.2020.3013933. https://ieeexplore.ieee.org/document/9154672
Anderson, R., Scala, C., Samuel, J., Kumar, V., & Jain, P. (2024). Are emotions conveyed across machine translations? establishing an analytical process for the effectiveness of multilingual sentiment analysis with italian text. Journal of Big Data and Artificial Intelligence, 2(1).
McKinsey & Company. (2023, June 14). The economic potential of generative AI: The next productivity frontier. Retrieved June 7, 2025, from https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier
Samuel, J., Khanna, T., Esguerra, J., Sundar, S., Pelaez, A., & Bhuyan, S. S. (2025). The rise of artificial intelligence phobia! Unveiling news-driven spread of AI fear sentiment using ML, NLP, and LLMs. IEEE Access, 13, 125944–125969. doi:10.1109/ACCESS.2025.3588179
OpenAI. (n.d.-a). ChatGPT pricing. Retrieved June 5, 2025, from https://openai.com/chatgpt/pricing/
Perplexity AI. (n.d.). Enterprise pricing. Retrieved June 5, 2025, from https://www.perplexity.ai/enterprise/pricing
Anthropic. (n.d.-a). Pricing. Retrieved June 5, 2025, from https://www.anthropic.com/pricing
Schmid, P. (2023, November 20). Benchmarking Llama 2 on Amazon SageMaker (Blog post). Retrieved June 5, 2025, from https://www.philschmid.de/sagemaker-llama-benchmark
Microsoft. (n.d.). Microsoft 365 Copilot pricing. Retrieved June 5, 2025, from https://www.microsoft.com/en-us/microsoft-365/copilot/pricing/
TechTarget. (2024, February 15). How does Microsoft 365 Copilot pricing and licensing work? Retrieved June 5, 2025, from https://www.techtarget.com/searchenterprisedesktop/answer/How-does-Microsoft-365-Copilot-pricing-and-licensing-work
Salesforce. (2023, June 28). Sales GPT and Service GPT now generally available (Press release). Retrieved June 5, 2025, from https://www.salesforce.com/news/stories/sales-gpt-service-gpt-ga/
Salesforce. (n.d.). Einstein for Service pricing. Retrieved June 5, 2025, from https://www.salesforce.com/service/ai/pricing/
AWS (n.d.). Amazon Bedrock pricing. Retrieved June 5, 2025, from https://aws.amazon.com/bedrock/pricing/
Ever Efficient AI (2024). Is AI Expensive for Small Businesses? A Cost Analysis. Retrieved June 8, 2025, from https://www.everefficient.ai/posts/is-ai-expensive-for-small-businesses-a-cost-analysis/
Sinha, R. (2025). Supercharging LLM Applications with Semantic Caching: Boost Speed, Cut Costs, and Maintain Accuracy. Retrieved June 8, 2025, from https://arkapravasinha.medium.com/supercharging-llm-applications-with-semantic-caching-boost-speed-cut-costs-and-maintain-accuracy-11f302464dff
Greyling C., 2023. Adding A Moderation Layer To Your OpenAI Implementations. Retrieved June 15, 2025, from https://cobusgreyling.medium.com/adding-a-moderation-layer-to-your-openai-implementations-3d4ea4f5d6ae
OpenAI (n.d.-b). MavenAGI launches automated customer support agents powered by OpenAI. Retrieved June 8, 2025, from https://openai.com/index/mavenagi/
Deloitte, 2023. AI Trends Outlook. Retrieved June 11, 2025, from https://www2.deloitte.com/us/en/pages/consulting/articles/ai-trends.html
Desk365.io, 2025. 61 AI Customer Service Statistics in 2025, Retrieved June 11, 2025 from https://www.desk365.io/blog/ai-customer-service-statistics/#:~:text=The%20future%20of%20AI%20in,(AI%20Business)
Bain & Company, 2024. Zero-Based Redesign: The Key to Realizing Gen AI's Cost Savings Potential. Retrieved June 11, 2025, from https://www.bain.com/insights/zero-based-redesign-the-key-to-realizing-Gen AI-cost-savings-potential/
NetElixir. (2025). NetElixir. Retrieved June 12, 2025, from NetElixir website: https://www.netelixir.com/

Samuel, Y., Brennan-Tonetta, M., Samuel, J., Kashyap, R., Kumar, V., Madabhushi, S. K.K., Chidipothu, N., Anand, I., and Jain, P., (2023). Cultivation of Human Centered Artificial Intelligence: Culturally Adaptive Thinking in Education for AI (CATE-AI). Frontiers in Artificial Intelligence, Volume 6. ISSN=2624-8212, URL=https://www.frontiersin.org/articles/10.3389/frai.2023.1198180 ; DOI=10.3389/frai.2023.1198180

Kashyap, R., Samuel, Y., Friedman, L. W., & Samuel, J. (2024). Artificial intelligence education & governance-human enhancive, culturally sensitive and personally adaptive HAI. Frontiers in Artificial Intelligence, 7, 1443386.

Samuel, J., Kashyap, Yana Samuel, and Alexander Pelaez. Adaptive cognitive fit: Artificial intelligence augmented management of information facets and representations. International Journal of Information Management 65 (2022) 102505, https://doi.org/10.1016/j.ijinfomgt.2022.102505

Ali, G. M. N., Rahman, M. M., Hossain, M. A., Rahman, M. S., Paul, K. C., Thill, J. C., & Samuel, J. (2021, August). Public perceptions of COVID-19 vaccines: Policy implications from US spatiotemporal sentiment analytics. In Healthcare (Vol. 9, No. 9, p. 1110). MDPI.

Rahman, M. M., Ali, G. M. N., Li, X. J., Samuel, J., Paul, K. C., Chong, P. H., & Yakubov, M. (2021). Socioeconomic factors analysis for COVID-19 US reopening sentiment with Twitter and census data. Heliyon (ScienceDirect by Elsevier), e06200.

Pendyala, M., Goel, U., Samuel, J., Patel, P., Kanakia, J., Pelaez, A., Savalia, N., & Khanna, T. (2025, May). Sentiment lost in preprocessing? Analysis of emoji-inclusive vs. emoji-exclusive methods with traditional lexicons-dictionaries and artificially intelligent ML-LLM strategies. ICTIS 2025, New York, USA.