

**THE IMPORTANCE OF AI-POWERED CHATBOTS ON
CUSTOMER SATISFACTION IN E-COMMERCE: A
CASE STUDY OF WISCONSIN INTERNATIONAL
UNIVERSITY COLLEGE, GHANA**

A SYSTEMATIC LITERATURE REVIEW

**BY
ABIGAIL OFORI**

12841543

A project submitted to School of Computing and Technology, Wisconsin International University College, Ghana, in partial fulfilment of the requirements for the award of the Master of Science degree in Business Computing.

DECEMBER 2026

DECLARATION

I declare that this project is my original academic work and that all sources consulted have been acknowledged. No part of this work has been submitted elsewhere for the award of another degree. The student should revise this declaration to match the regulations of the awarding institution before submission.

Student signature: _____ Date: _____

Supervisor signature: _____ Date: _____

DEDICATION

This work is dedicated to my family, teachers and colleagues whose encouragement made this academic journey possible.

ACKNOWLEDGEMENTS

I am grateful to my supervisor for scholarly guidance, to the faculty for the research training that informed this project, and to my family and colleagues for their patience and support. I also acknowledge the authors and institutions whose published studies provided the evidence synthesised in this review. Any remaining errors are my responsibility.

ABSTRACT

AI-powered chatbots have become a prominent interface between e-commerce firms and customers, yet their contribution to customer satisfaction depends on more than availability or response speed. This project examines the importance of AI-powered chatbots for customer satisfaction in e-commerce through a structured review and thematic synthesis of peer-reviewed and authoritative literature. The study is guided by the Technology Acceptance Model, the Information Systems Success Model, expectation-confirmation theory, e-service-quality thinking and social-presence theory. A transparent search and screening protocol was designed around studies addressing conversational agents, online retail, service quality, trust, customer experience and satisfaction. The synthesis identifies six interrelated themes: responsiveness and convenience; information and system quality; personalisation and decision support; trust, privacy and transparency; social presence and conversational design; and service recovery with human escalation. Findings show that chatbots can improve satisfaction by providing continuous support, reducing waiting time, answering routine questions consistently and supporting product discovery. However, the benefit is conditional on accurate information, appropriate personalisation, secure data practices, realistic disclosure and frictionless transfer to human staff. Poor intent recognition, repetitive dialogue, fabricated answers and blocked escalation can reverse the expected gains. The project concludes that chatbot effectiveness should be measured as an end-to-end service outcome rather than contact deflection alone. It recommends a human-in-the-loop operating model, tiered deployment by task risk, continuous quality monitoring and multidimensional satisfaction metrics. Because the study synthesises secondary evidence rather than collecting primary data, the conclusions identify relationships and design principles but do not claim causal estimates for a particular retailer or country.

Keywords: artificial intelligence; chatbot; customer satisfaction; e-commerce; service quality; trust; personalisation.

TABLE OF CONTENTS

Declaration

Dedication

Acknowledgements

Abstract

List of Tables

List of Figures

List of Abbreviations

Chapter One: Introduction

Chapter Two: Literature Review

Chapter Three: Methodology

Chapter Four: Findings and Discussion

Chapter Five: Summary, Conclusions and Recommendations

References

Appendices

Note: In Microsoft Word, right-click this page and insert/update an automatic Table of Contents after the institution-specific details and final pagination have been confirmed.

LIST OF TABLES

Table 1.1 Operational definitions

Table 2.1 Selected theoretical perspectives

Table 2.2 Summary of chatbot value and risk mechanisms

Table 3.1 Review eligibility criteria

Table 3.2 Data extraction fields

Table 4.1 Thematic synthesis overview

Table 4.2 Satisfaction measurement framework

Table 5.1 Implementation roadmap

Appendix B Evidence matrix

LIST OF FIGURES

Figure 2.1 Conceptual framework

Figure 3.1 Review process

Figure 4.1 Conditional pathway from chatbot capability to satisfaction

LIST OF ABBREVIATIONS

Abbreviation	Meaning
AI	Artificial Intelligence
CRM	Customer Relationship Management
ECM	Expectation-Confirmation Model
E-S-QUAL	Electronic Service Quality
IS	Information Systems
ISSM	Information Systems Success Model
LLM	Large Language Model
NLP	Natural Language Processing
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
TAM	Technology Acceptance Model
UTAUT	Unified Theory of Acceptance and Use of Technology

CHAPTER ONE

INTRODUCTION

1.1 Background to the Study

E-commerce has shifted a substantial share of customer contact from physical stores and call centres to digital interfaces. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Huang & Rust, 2018; Adam et al., 2021; Følstad & Brandtzæg, 2020)

The principal mechanisms are instant availability, natural-language interaction, scalable response capacity and access to product and order information. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that customers value chatbots when they reduce effort and uncertainty, while satisfaction falls when conversations become circular or inaccurate. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include misunderstanding, impersonal treatment, inaccessible escalation, privacy concerns and unequal performance across languages. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint,

conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that retailers should align chatbot capability with the complexity and consequence of the task. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

1.2 Statement of the Problem

Many e-commerce firms deploy chatbots to manage rising contact volumes and provide round-the-clock assistance, but deployment does not necessarily generate satisfied customers. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Ashfaq et al., 2020; Araujo, 2018; Luo et al., 2019)

The principal mechanisms are service speed, information quality, perceived usefulness, trust and resolution quality. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that studies often report positive associations between usefulness or service quality and satisfaction, but also show that social cues and disclosure can change user reactions. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include vendor-led performance claims, fragmented measurement, short laboratory encounters and limited evidence from African e-commerce contexts. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that the research problem should be framed around the conditions under which chatbot interactions create satisfactory outcomes. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

1.3 Aim of the Study

The aim of this study is to critically examine the importance of AI-powered chatbots in improving customer satisfaction in e-commerce and to identify the service, technological and governance conditions that shape this relationship.

1.4 Specific Objectives

To examine how responsiveness and convenience associated with AI-powered chatbots affect customer satisfaction in e-commerce.

To assess the influence of chatbot information quality, system quality and personalisation on customer satisfaction.

To analyse the roles of trust, privacy, transparency and social presence in customer evaluations of chatbot service.

To identify the major limitations of chatbot-mediated customer service, including service failure and escalation problems.

To develop evidence-based recommendations for the effective implementation and evaluation of AI-powered chatbots in e-commerce.

1.5 Research Questions

How do responsiveness and convenience provided by AI-powered chatbots influence customer satisfaction in e-commerce?

How do information quality, system quality and personalisation influence satisfaction with chatbot service?

What roles do trust, privacy, transparency and social presence play in customer responses to AI-powered chatbots?

What limitations reduce the effectiveness of chatbots in e-commerce customer service?

What implementation practices can help e-commerce firms maximise satisfaction while controlling risk?

1.6 Significance of the Study

The study is significant to scholarship, management, consumers and policymakers because conversational AI is simultaneously a service innovation and a new point of operational risk. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Wirtz et al., 2018; Dwivedi et al., 2021)

The principal mechanisms are evidence synthesis, theory integration, managerial translation and explicit attention to consumer protection. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that the most useful recommendations connect technical features to customer outcomes and measurable service processes. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include technology-centred evaluation, unverified claims of savings and neglect of vulnerable or low-literacy users. These problems explain why a

narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that decision makers should evaluate benefits and harms together and preserve meaningful customer choice. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

1.7 Scope and Delimitations

The project focuses on AI-powered text or voice conversational agents used in business-to-consumer e-commerce for product discovery, transactional assistance, customer support and service recovery. It considers research published mainly from 2017 to 2026, while retaining earlier foundational theories. It does not evaluate a named retailer, build a chatbot, or estimate causal effects from proprietary transaction data. The geographical scope is international, with implications drawn cautiously for emerging markets including Ghana. The unit of analysis is the customer-chatbot service encounter and its relationship with satisfaction and closely related outcomes such as trust, continued use and loyalty.

1.8 Operational Definitions

Term	Operational meaning in this project
AI-powered chatbot	A conversational system using natural-language processing, machine learning or generative models to interpret customer input and produce service responses.
Customer satisfaction	A customer's post-interaction evaluation of whether chatbot-enabled service met or exceeded expectations.
E-commerce	Commercial exchange and related customer service conducted through digital channels.
Personalisation	Adaptation of content, recommendations or dialogue using customer context, preferences or behaviour.
Service recovery	Actions taken after a failure to resolve the problem, restore confidence and retain the customer.

1.9 Organisation of the Study

The project contains five chapters. Chapter One introduces the problem, objectives and scope. Chapter Two reviews concepts, theories and empirical studies and presents the conceptual framework. Chapter Three explains the structured literature-review design. Chapter Four reports and discusses the thematic findings. Chapter Five summarises the

study, draws conclusions, presents recommendations and identifies areas for further research. References and supporting appendices follow.

CHAPTER TWO

LITERATURE REVIEW

2.1 Concept of Artificial Intelligence in Service

Artificial intelligence in service refers to systems that perform or augment tasks involving prediction, language, recommendation and decision support. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Huang & Rust, 2018; Wirtz et al., 2018; Dwivedi et al., 2021)

The principal mechanisms are data-driven learning, automated inference, conversational interaction and integration with service workflows. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that AI creates value most consistently when tasks are structured and when human expertise remains available for ambiguity and exception handling. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include automation bias, opacity, bias in training data, overconfidence and unclear responsibility. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative

review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that organisations should match the level of autonomy to task complexity and customer vulnerability. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.2 Evolution and Types of Chatbots

Chatbots have evolved from rule-based menu systems to retrieval-based assistants and generative conversational agents. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Dale, 2016; Følstad & Brandtzæg, 2017; Rapp et al., 2021)

The principal mechanisms are pattern matching, intent classification, knowledge retrieval, dialogue management and language generation. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that modern systems can sustain more flexible exchanges but flexibility also introduces uncertainty and the possibility of unsupported answers. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include stale knowledge, hallucination, prompt manipulation, inconsistent tone and loss of transactional context. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that architecture should be selected according to risk, with grounded retrieval and deterministic controls for consequential tasks. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.3 Customer Satisfaction in E-commerce

Customer satisfaction is a cumulative and transaction-specific evaluation shaped by expectations, perceived performance and the effort required to obtain value. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Oliver, 1980; DeLone & McLean, 2003; Parasuraman et al., 2005)

The principal mechanisms are expectation confirmation, perceived usefulness, affect, trust and service recovery. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that satisfaction predicts continued use and loyalty but can differ across journey stages and customer segments. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an

adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include single-item ratings, survivorship bias, coerced feedback and confusion between satisfaction with the bot and satisfaction with the retailer. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that measurement should separate interaction quality, resolution quality and relationship outcomes. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.4 Responsiveness, Availability and Convenience

Responsiveness concerns the timeliness and relevance of a reply, while convenience concerns access with minimal time and effort. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Xu et al., 2017; Adam et al., 2021)

The principal mechanisms are twenty-four-hour availability, reduced queues, simultaneous handling and immediate access within the shopping interface. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into

a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that rapid responses improve perceived control particularly for routine product, delivery and account questions. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include fast but wrong answers, rushed interaction, poor mobile design and digital exclusion. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that response-time targets must be paired with accuracy, resolution and accessibility targets. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.5 Information Quality and System Quality

Information quality includes accuracy, completeness, relevance and currency, whereas system quality includes reliability, usability and integration. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (DeLone & McLean, 2003; Ashfaq et al., 2020)

The principal mechanisms are authoritative knowledge sources, catalogue synchronisation, context retention and stable interface behaviour. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that usefulness and confirmation depend strongly on whether customers can rely on the information and complete the intended task. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include contradictory policy text, outdated inventory, broken integrations and confident fabrication. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that firms should ground responses in approved sources and test full transactions rather than isolated answers. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.6 Personalisation and Product Recommendation

Personalisation enables a chatbot to narrow choices and adapt explanations to the customer's expressed needs and permitted context. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less

uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Shumanov & Johnson, 2021; Bleier et al., 2019)

The principal mechanisms are preference elicitation, catalogue filtering, recommendation, language adaptation and conversational memory. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that relevant recommendations can reduce overload and strengthen perceived usefulness, particularly in complex catalogues. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include filter bubbles, discriminatory inference, overselling, hidden sponsorship and inappropriate use of personal data. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that personalisation should be controllable, explainable and proportionate to the task. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.7 Trust, Privacy and Transparency

Trust reflects willingness to rely on the chatbot and the firm despite uncertainty, and is influenced by competence, integrity and benevolence cues. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Araujo, 2018; Luo et al., 2019; Shin, 2021)

The principal mechanisms are accurate performance, clear identity disclosure, predictable policies, privacy safeguards and accountable escalation. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that trust mediates continued use and can be strengthened when capabilities and limitations are communicated honestly. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include deceptive anthropomorphism, concealed data use, security incidents, manipulative recommendations and unsupported certainty. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that the interface should disclose automation, explain relevant data use and avoid claiming human qualities it does not possess. Implementation should begin with high-volume, bounded intents supported by reliable

product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.8 Social Presence and Anthropomorphism

Social presence is the sense that another social actor is available in an interaction, while anthropomorphism attributes human characteristics to technology. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Araujo, 2018; Go & Sundar, 2019; Sheehan et al., 2020)

The principal mechanisms are names, avatars, conversational tone, empathy cues, turn-taking and acknowledgement. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that moderate social cues can make digital service feel warmer and increase engagement, especially when they support clarity. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include uncanny reactions, inflated expectations, stereotypes, emotional manipulation and disappointment when empathy is superficial. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint,

conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that human-like design should support communication without obscuring the system's non-human status. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.9 Service Failure, Recovery and Human Handover

Chatbot failure occurs when the system cannot understand, answer or complete a customer's goal, and recovery determines whether the failure becomes lasting dissatisfaction. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Mozafari et al., 2022)

The principal mechanisms are error recognition, apology, clarification, alternative channels, context transfer and empowered human support. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that customers tolerate some automation limits when the system recognises them quickly and provides an effective route forward. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include endless fallback loops, repeated identity checks, lost context, blame shifting and unavailable agents. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that handover should be treated as a core capability with explicit triggers, service levels and ownership. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.10 Theoretical Review

Theory/model	Core idea	Application to this study
Technology Acceptance Model	Use is shaped by perceived usefulness and ease of use.	Explains adoption and continued interaction with chatbots.
Information Systems Success Model	System, information and service quality influence use, satisfaction and benefits.	Organises technical and service quality antecedents.
Expectation-Confirmation Model	Satisfaction follows comparison of experienced performance with expectations.	Explains post-use evaluation and continuance.
Social Presence Theory	Communication media differ in perceived interpersonal presence.	Explains reactions to conversational and human-like cues.
E-S-QUAL	Efficiency, fulfilment, availability and privacy shape electronic service quality.	Connects chatbot performance to wider e-commerce service.

2.10.1 Technology Acceptance Model

The Technology Acceptance Model proposes that perceived usefulness and perceived ease of use shape attitudes and behavioural intentions toward a system. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Davis, 1989; Kasilingam, 2020)

The principal mechanisms are task benefit, low interaction effort, learnability and compatibility with customer goals. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that chatbot satisfaction and continuance are stronger when customers believe the system saves time and helps them obtain reliable outcomes. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include treating adoption intention as equivalent to satisfaction and overlooking trust, emotion and service recovery. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that TAM should be integrated with quality and trust constructs rather than used as a complete explanation. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.10.2 Information Systems Success Model

The Information Systems Success Model links information quality, system quality and service quality to use, user satisfaction and net benefits. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers

cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (DeLone & McLean, 2003)

The principal mechanisms are accuracy, reliability, usability, support and realised value. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that the model offers the clearest structure for distinguishing a polished interface from a service that actually resolves customer needs. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include ambiguous measures of use, reciprocal relationships and failure to separate organisational benefit from customer benefit. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that evaluation should trace quality dimensions through satisfaction to customer and business outcomes. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.10.3 Expectation-Confirmation and E-service Quality

Expectation-confirmation theory explains satisfaction as a comparison between prior expectations and experienced performance, while e-service quality specifies important performance dimensions. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Oliver, 1980; Bhattacharjee, 2001; Parasuraman et al., 2005)

The principal mechanisms are expectation formation, confirmation, efficiency, fulfilment, system availability and privacy. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that a chatbot can exceed speed expectations yet fail fulfilment expectations if the order issue remains unresolved. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include marketing that inflates expectations and metrics that ignore the wider fulfilment system. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that chatbot promises and interface language should be aligned with actual operational capability. Implementation

should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

2.11 Empirical Review and Research Gap

Empirical research broadly supports positive relationships among perceived usefulness, information quality, conversational quality, trust and satisfaction. Ashfaq et al. (2020) found that satisfaction plays an important role in continuance intention toward AI-powered service agents. Adam et al. (2021) showed that chatbot design influences users' compliance with service requests, while Araujo (2018) demonstrated that anthropomorphic cues and the framing of the agent affect company-related perceptions. Luo et al. (2019) provided field evidence that disclosure of machine identity can affect purchases, illustrating that transparency and customer beliefs have real commercial consequences. These findings establish importance but also show that outcomes depend on task and design context.

Three gaps remain. First, many studies examine intentions, attitudes or brief experimental encounters rather than verified resolution across the complete e-commerce journey. Second, evidence is geographically concentrated in digitally mature markets and student or convenience samples, limiting confident transfer to contexts where connectivity, digital literacy, language and payment infrastructures differ. Third, rapid movement toward generative AI changes the risk profile through flexible dialogue and hallucination. There is therefore a need for integrated evaluation that combines traditional service-quality dimensions with trust, governance, escalation and longitudinal operational outcomes.

2.11.1 Evidence from Ghana

Ghanaian research provides important contextual evidence even though direct, peer-reviewed studies of AI chatbots in Ghanaian e-commerce remain scarce. Gbolonyo et al. (2022), using data from 382 Ghanaian online consumers, found that site design, information usefulness and empathy significantly predicted customer satisfaction. Satisfaction, in turn, influenced shopping intention, while shopping intention connected satisfaction to actual and continued usage. These findings closely correspond to the present study's chatbot dimensions: a conversational interface must be usable, provide useful information and communicate with appropriate sensitivity if it is to improve the wider online shopping experience.

Amofah and Chai (2022) examined consumer e-commerce adoption in Ghana and reported that trust plays a meaningful role in sustaining adoption. Akanferi et al. (2022) similarly found that performance expectancy, trust and facilitating conditions influenced public servants' adoption of mobile commerce and mobile payments. Yeboah et al. (2020), in

a qualitative Ghanaian study of merchant adoption of mobile payments, showed that trust in both the technology and the service provider was particularly important because of security and technology risk. Together, these studies indicate that Ghanaian customers are likely to judge an e-commerce chatbot not only by conversational speed but also by the credibility of the retailer, payment security, reliable infrastructure and confidence that the service will work as promised.

Antwi (2021) linked online consumer trust, satisfaction and commitment with repurchase intention, reinforcing the importance of relationship quality in digital retail. These Ghana-relevant studies do not establish the effect of chatbots themselves, and this distinction must be preserved. They instead identify local boundary conditions against which chatbot propositions should be tested. The evidence supports a Ghana-specific research agenda addressing mobile-first use, local language and spelling variation, connectivity interruptions, payment and delivery integration, transparent data practices and the availability of human support.

2.12 Conceptual Framework

The conceptual framework proposes that chatbot service capabilities—responsiveness, convenience, information quality, system quality, personalisation and social presence— influence customer satisfaction. Perceived usefulness, perceived ease of use and trust act as explanatory mechanisms. Privacy concern, task complexity, prior experience, customer characteristics and human-handover quality condition the relationship. Satisfaction subsequently influences continued use, loyalty, positive word of mouth and purchase-related behaviour. The framework does not assume that every capability has an unconditionally positive effect; each may display thresholds or trade-offs.

Inputs	Mediating evaluations	Primary outcome	Downstream outcomes
Responsiveness; convenience; information quality; system quality; personalisation; social presence	Usefulness; ease of use; trust; confirmation; perceived effort	Customer satisfaction	Continued use; loyalty; word of mouth; purchase intention
Moderators: task complexity, privacy concern, prior experience, digital literacy, culture, and handover quality			

2.13 Chapter Summary

The literature indicates that AI-powered chatbots matter because they influence both the efficiency and perceived humanity of e-commerce service. The theories reviewed explain complementary parts of the process: acceptance, system success, expectation confirmation, e-service quality and social presence. The empirical evidence supports a conditional model in which usefulness, quality and trust translate chatbot capabilities into satisfaction, while privacy, complexity and poor recovery can weaken or reverse the benefit.

CHAPTER THREE

METHODOLOGY

3.1 Research Philosophy and Approach

The study adopts a pragmatic, evidence-synthesis orientation. It treats customer satisfaction as a phenomenon that can be investigated through quantitative associations, experiments, surveys and qualitative accounts. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Creswell & Creswell, 2018; Snyder, 2019)

The principal mechanisms are methodological pluralism, transparent selection, structured comparison and interpretive synthesis. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that combining methods gives a fuller account of both measurable relationships and the reasons customers react to chatbot encounters. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include selective citation, confirmation bias, incompatible measures and overstatement of causal conclusions. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings

with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that claims should be calibrated to design quality and convergence across different methods. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

3.2 Research Design

A structured literature review with thematic synthesis was selected. This design is appropriate because the topic spans information systems, marketing, service management and human-computer interaction, and because the available studies use heterogeneous outcomes and methods that are not readily combined in a single meta-analysis. The approach follows the transparency principles of systematic reviewing while allowing conceptual integration. It identifies recurring mechanisms, contradictions, boundary conditions and managerial implications. The review is not presented as a registered systematic review and does not claim exhaustive coverage of every database.

3.3 Search Strategy

The proposed reproducible search string combined three concept blocks: (chatbot* OR conversational agent* OR virtual assistant* OR conversational AI) AND (e-commerce OR online retail OR digital commerce OR online shopping) AND (customer satisfaction OR customer experience OR service quality OR trust OR continuance OR loyalty). Searches should be run in Scopus, Web of Science, ABI/INFORM, ScienceDirect, Emerald Insight, SpringerLink and Google Scholar. Citation chaining from key articles should supplement database retrieval. Before examination, the student must record the actual search date, database-specific syntax and result count in Appendix A; no count is invented in this document.

3.4 Eligibility Criteria

Criterion	Include	Exclude
Publication type	Peer-reviewed journal articles and high-quality conference papers; authoritative theory sources.	Blogs, unverified commercial claims and anonymous web pages.
Topic	AI chatbots or conversational agents connected to customer service, online retail or digital service outcomes.	Studies unrelated to customer interaction or satisfaction.
Outcome	Satisfaction, experience, service quality, trust, use, loyalty or purchase-related outcomes.	Purely technical accuracy papers without customer/service implications.
Period	Main empirical emphasis from 2017–2026; earlier	Older non-foundational discussion with little

	foundational theory retained.	relevance to contemporary systems.
Language	English.	Non-English where a reliable translation was unavailable.
Evidence	Methods and context described sufficiently for appraisal.	Claims lacking traceable methods or sources.

3.5 Screening and Selection

Screening proceeds in four stages: identification, deduplication, title-and-abstract screening, and full-text eligibility assessment. Reasons for full-text exclusion include wrong context, absence of a relevant outcome, purely technical focus, non-peer-reviewed publication and insufficient methodological information. A PRISMA-style flow diagram should report the actual numbers after the searches are completed. Until then, Appendix A deliberately uses blank fields. This decision protects research integrity by distinguishing a complete scholarly narrative from empirical activities that require database access and documented execution.

3.6 Data Extraction

Field	Purpose
Bibliographic details	Author, year, title, outlet and DOI.
Context	Country, sector, e-commerce setting and chatbot type.
Design and sample	Experiment, survey, field study, qualitative study or review; sample characteristics.
Constructs	Service quality, usefulness, trust, satisfaction and related outcomes.
Key findings	Direction and conditions of reported relationships.
Limitations	Bias, generalisability, measurement and design constraints.
Managerial implication	Actionable design, governance or measurement lesson.

3.7 Quality Appraisal

Quality appraisal considers clarity of the research question, appropriateness of design, transparency of sampling, reliability and validity of measures, adequacy of analysis, support for conclusions, disclosure of limitations and relevance to the review. Experimental studies are assessed for randomisation and manipulation validity; survey studies for measurement quality and common-method risk; qualitative studies for sampling, analytic transparency and evidentiary support; and reviews for search and selection transparency. Quality informs the weight given to a study rather than acting as a simple mechanical exclusion score.

3.8 Method of Analysis

Thematic synthesis is conducted through familiarisation, initial coding, grouping codes into descriptive themes, and development of analytical themes that answer the research questions. Codes include speed, availability, accuracy, ease of use, personalisation, trust, privacy, disclosure, empathy, failure, recovery and handover. Rival interpretations are retained where findings differ. The final synthesis links observed outcomes to the conceptual framework and distinguishes direct evidence from reasoned managerial inference.

3.9 Trustworthiness

Credibility is strengthened through transparent eligibility rules, use of peer-reviewed sources, comparison across methods and explicit attention to contradictory results. Dependability is supported by a documented search and extraction template. Confirmability is promoted by maintaining an audit trail of decisions and separating findings from recommendations. Transferability is addressed by describing study contexts and avoiding universal claims. A supervisor or second reviewer should independently inspect a sample of screening and coding decisions where programme resources permit.

3.10 Ethical Considerations

The review uses published material and does not recruit human participants or process private customer data. Formal institutional requirements should nevertheless be checked. Ethical responsibilities include accurate citation, faithful representation of findings, avoidance of fabricated search counts, and recognition of privacy, bias, transparency and consumer-protection issues in the subject matter. If the optional questionnaire in Appendix C is later administered, prior ethics approval, informed consent, secure storage and a lawful data-management plan will be required.

3.11 Methodological Limitations

The review may be affected by publication bias, English-language restriction, database coverage and rapid technological change. Studies employ different chatbot types, industries, constructs and outcome measures, reducing direct comparability. Satisfaction is often self-reported shortly after exposure and may not predict repeat purchasing or complaint behaviour. The review therefore offers a rigorous conceptual and thematic account, not a pooled causal effect size. These limitations are revisited when interpreting the findings.

3.12 Chapter Summary

This chapter presented a transparent structured-review methodology. It defined the search concepts, eligibility rules, screening stages, extraction fields, appraisal criteria and thematic synthesis. Most importantly, it identified which activities require actual database execution and preserved blank reporting fields rather than manufacturing study-flow numbers.

CHAPTER FOUR

FINDINGS AND DISCUSSION

The synthesis generated six themes that collectively explain the importance of AI-powered chatbots for customer satisfaction. The themes are interdependent: speed creates little value without accuracy; personalisation depends on trust; social presence can improve interaction but increase expectations; and automation requires effective recovery and human escalation. Accordingly, the findings are presented as a conditional service system rather than a list of independent features.

Theme	Principal satisfaction pathway	Main failure condition
Responsiveness and convenience	Lower waiting time and effort.	Fast but irrelevant or unresolved replies.
Information and system quality	Reliable answers and smooth task completion.	Outdated, inconsistent or fabricated information.
Personalisation and decision support	Reduced choice overload and greater relevance.	Intrusive data use or biased recommendations.
Trust, privacy and transparency	Willingness to rely on the system and retailer.	Deception, opacity or insecure processing.
Social presence and conversation	Engagement, warmth and perceived attention.	Uncanny design or false expectations of empathy.
Recovery and human escalation	Containment of failure and restored confidence.	Loops, context loss and inaccessible agents.

4.1 Theme One: Responsiveness and Convenience

The most immediate contribution of chatbots is reduction of temporal and procedural friction. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Xu et al., 2017; Adam et al., 2021)

The principal mechanisms are instant replies, continuous availability, simultaneous service and embedded access. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that customers often interpret quick, relevant assistance as respect for their time and greater control over the purchase process. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include queue replacement with confusing loops, inaccessible design and speed metrics detached from resolution. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that firms should optimise time to correct resolution rather than time to first automated response. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

4.2 Theme Two: Information and System Quality

Satisfaction depends fundamentally on whether the chatbot supplies dependable information and completes tasks reliably. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (DeLone & McLean, 2003; Ashfaq et al., 2020)

The principal mechanisms are accuracy, relevance, completeness, currency, usability, integration and contextual continuity. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when

delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that quality dimensions consistently support perceived usefulness, confirmation and satisfaction across information-systems research. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include catalogue drift, policy conflicts, integration failure, hallucination and fragile authentication. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that approved knowledge, retrieval grounding, transaction testing and confidence-based refusal are essential. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

4.3 Theme Three: Personalisation and Decision Support

Chatbots can convert a large product assortment into a guided conversation that reflects customer needs. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether

the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Shumanov & Johnson, 2021; Bleier et al., 2019)

The principal mechanisms are preference questions, context-aware comparison, tailored explanation and recommendation. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that relevance reduces cognitive load and can improve confidence, engagement and perceived service care. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include unexplained profiling, discriminatory outputs, manipulation and recommendations driven by margin rather than customer fit. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that personalisation must be transparent, correctable, minimal and aligned with customer benefit. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

4.4 Theme Four: Trust, Privacy and Transparency

Trust is a bridge between chatbot performance and willingness to rely on e-commerce service. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a

customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Luo et al., 2019; Araujo, 2018; Shin, 2021)

The principal mechanisms are competence signals, honest disclosure, consistent policies, data restraint and visible accountability. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that accurate performance builds cognitive trust while respectful communication and recovery support relational trust. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include identity deception, covert data reuse, security weaknesses, fabricated certainty and unclear responsibility. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that transparency should specify that the customer is interacting with AI, what it can do and how to reach a human. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

4.5 Theme Five: Social Presence and Conversational Design

Conversational form differentiates chatbots from static search and frequently shapes how customers experience the retailer. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Go & Sundar, 2019; Sheehan et al., 2020; Araujo, 2018)

The principal mechanisms are natural turn-taking, acknowledgement, appropriate tone, memory within the session and restrained empathy cues. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that well-calibrated social presence can increase engagement and make assistance feel accessible without requiring a human for every routine contact. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include over-humanisation, stereotypes, emotional mimicry, verbosity and mismatch between warm language and weak capability. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that the voice of the bot should be clear, respectful and recognisably automated, with tone adapted to the seriousness of the issue. Implementation should begin with high-volume, bounded intents supported by reliable

product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

4.6 Theme Six: Service Recovery and Human Escalation

The importance of chatbots is revealed not only by successful routine encounters but by the way the service system responds when automation fails. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Mozafari et al., 2022)

The principal mechanisms are early failure detection, clarification, apology, alternative action, context-preserving transfer and agent authority. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that effective recovery can protect satisfaction and demonstrate organisational accountability even after an initial error. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include denial of failure, circular fallbacks, repeated authentication, lost transcripts and delayed human response. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings

with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that human escalation should be designed, staffed, monitored and tested as part of the chatbot product. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

4.7 Integrated Discussion by Research Question

Research Question One concerned responsiveness and convenience. The synthesis shows that these capabilities are important because they reduce waiting and search effort, extend support beyond operating hours and enable immediate assistance at the point of decision. Their influence is mediated by perceived usefulness and confirmation. A fast answer contributes to satisfaction only when it is relevant and enables progress. For managers, this means that first-response time is a weak standalone indicator; correct-resolution time and customer effort provide a more defensible account of value.

Research Question Two concerned information quality, system quality and personalisation. These factors form the functional foundation of satisfaction. Customers require accurate product, inventory, policy and order information; stable interfaces; and reliable execution across integrated systems. Personalisation adds value by reducing choice overload, but it is beneficial only when the data are appropriate and the recommendation can be understood or corrected. The Information Systems Success Model therefore offers a strong central explanation, strengthened by privacy and fairness considerations.

Research Question Three concerned trust, privacy, transparency and social presence. Trust is both an outcome of reliable encounters and an antecedent of future reliance. Disclosure can reduce the short-term persuasive benefit of mistaken human identity but protects autonomy and long-term confidence. Social cues are most useful when they make communication easier, not when they disguise limitations. Privacy is not a separate compliance issue: customers' beliefs about data use directly influence comfort, candour and willingness to continue the interaction.

Research Question Four concerned limitations. The main barriers are weak intent recognition, inaccurate or stale knowledge, inability to handle exceptions, privacy risk, accessibility problems, and poor handover. These limitations are magnified for emotionally charged, financially consequential or ambiguous requests. Generative systems add flexibility but also create hallucination and governance risk. The appropriate response is not rejection of automation, but tiered autonomy, grounding, monitoring and clear accountability.

Research Question Five concerned implementation. The evidence supports a staged, human-in-the-loop approach. Firms should begin with bounded intents, establish baseline service measures, validate knowledge and integrations, disclose AI use, define escalation triggers, pilot with diverse users and review failures continuously. Business success should be judged through balanced customer and operational outcomes, not through the proportion of contacts kept away from employees.

4.8 Proposed Measurement Framework

Dimension	Illustrative measures	Interpretive caution
Efficiency	First-response time; task time; turns; waiting time.	Short times can conceal premature abandonment.
Effectiveness	First-contact resolution; task completion; answer accuracy; re-contact.	Resolution requires confirmation from downstream systems.
Experience	CSAT; customer effort; perceived usefulness; ease of use.	Immediate ratings may not predict later loyalty.
Trust and safety	Disclosure comprehension; privacy concern; harmful-answer rate; security incidents.	Low complaint counts may reflect under-reporting.
Recovery	Escalation success; context transferred; time to human resolution.	A low escalation rate is not inherently good.
Business outcomes	Conversion; retention; repeat purchase; cost per resolved contact.	Attribution requires comparison and control for confounders.

4.9 Implications for E-commerce in Ghana and Similar Markets

The reviewed chatbot evidence is internationally oriented, so direct local effects must still be validated. Nevertheless, Ghanaian e-commerce research gives the local discussion a firmer basis. Site design, information usefulness and empathy have been associated with satisfaction among Ghanaian online consumers, while trust, expected performance and facilitating conditions recur in studies of e-commerce, mobile commerce and mobile payments (Gbolonyo et al., 2022; Amofah & Chai, 2022; Akanferi et al., 2022). A Ghanaian chatbot should therefore be evaluated as part of the complete mobile commerce system rather than as a stand-alone conversation tool.

In practice, multilingual interaction, variable connectivity, mobile-first access, digital-payment diversity and different levels of digital literacy may strongly condition satisfaction. A lightweight interface, plain language, local terminology and graceful recovery from network interruption may matter as much as advanced conversational fluency. Integration with local delivery and payment processes is essential because customers judge the whole service outcome. Trust-building measures should cover both the underlying technology and the service provider, consistent with Ghanaian mobile-payment evidence (Yeboah et al., 2020).

Local firms should test English and relevant Ghanaian-language use cases with real users and monitor whether performance differs across accents, spelling patterns, devices, locations or disability needs. Privacy notices and chatbot data practices should comply with Ghana's Data Protection Act, 2012 (Act 843), while electronic service processes should be

consistent with the Electronic Transactions Act, 2008 (Act 772). The chatbot should not become a barrier to human assistance for customers who are less comfortable with automation. Partnerships among retailers, universities and technology providers could produce locally grounded datasets and evaluation methods while maintaining consent, privacy and security.

4.10 Chapter Summary

The findings establish that AI-powered chatbots can be important drivers of e-commerce customer satisfaction, but their value is conditional. Responsiveness and convenience attract use; information and system quality create useful outcomes; personalisation supports decisions; trust and social presence shape willingness to engage; and recovery determines whether failures become lasting dissatisfaction. An integrated measurement system and human escalation are therefore central to responsible implementation.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary of the Study

This project examined the importance of AI-powered chatbots for customer satisfaction in e-commerce. It responded to the gap between optimistic claims about automation and the more conditional evidence found in customer-service research. A structured literature review integrated information-systems, marketing, service and human-computer-interaction perspectives. The analysis was guided by TAM, the Information Systems Success Model, expectation-confirmation theory, e-service-quality thinking and social-presence theory.

Six themes were identified: responsiveness and convenience; information and system quality; personalisation and decision support; trust, privacy and transparency; social presence and conversational design; and service recovery with human escalation. Across these themes, the central finding is that chatbot technology improves satisfaction when it reduces effort and uncertainty while preserving accuracy, agency, dignity and access to accountable human support. The same technology diminishes satisfaction when speed substitutes for resolution, human-like cues conceal limitations, or automation traps the customer in a failed process.

5.2 Conclusions

AI-powered chatbots are important to customer satisfaction because they increasingly constitute the visible front line of e-commerce service. They can provide scalable, immediate and context-sensitive assistance across the customer journey. Nevertheless, importance should not be confused with inherent benefit. The chatbot is a sociotechnical service arrangement whose outcomes depend on data quality, integration, interface design, employee support, governance and the retailer's underlying ability to fulfil promises.

The most defensible conclusion is therefore conditional: AI-powered chatbots enhance customer satisfaction when customers receive accurate, relevant and timely help with reasonable effort; when data use and machine identity are transparent; when recommendations serve the customer; and when failures trigger competent human recovery. Satisfaction deteriorates when the chatbot is deployed mainly as a gatekeeper or cost-control mechanism. Sustainable value arises from augmentation and orchestration, not automation at any cost.

5.3 Recommendations for E-commerce Firms

5.3.1 Start with Customer Problems, Not Technology

Firms should prioritise use cases through customer-journey evidence. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created

by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Wirtz et al., 2018)

The principal mechanisms are contact-driver analysis, journey mapping, complaint review and feasibility assessment. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that bounded, frequent and low-risk intents produce the safest early value. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include novelty-led deployment, vague goals and automating broken processes. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that each use case should have a named customer outcome, owner and baseline measure. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

5.3.2 Build on Authoritative Knowledge and Reliable Integration

Chatbot answers should be grounded in approved, current sources and verified systems. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (DeLone & McLean, 2003)

The principal mechanisms are knowledge ownership, retrieval controls, catalogue synchronisation, test suites and change management. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that accuracy and completion are more important to satisfaction than linguistic polish. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include hallucination, stale policy, inconsistent inventory and unverified transactions. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that high-risk answers should require deterministic rules, source links or human review. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are

lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

5.3.3 Design Human Handover as a Core Feature

Escalation should be easy, timely and context preserving. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Mozafari et al., 2022)

The principal mechanisms are confidence thresholds, customer request, sentiment or risk triggers, transcript transfer and service-level agreements. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that effective handover converts automation limits into manageable service moments. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include blocked access, repeated questions, lost context and unclear ownership. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that human teams must be staffed and empowered to resolve exceptions created upstream. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

5.3.4 Protect Privacy, Security and Customer Autonomy

Responsible data practice should be visible in the interaction. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Shin, 2021)

The principal mechanisms are data minimisation, purpose limitation, access control, retention rules, consent and understandable notices. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that trust is strengthened when customers know they are interacting with AI and can control relevant data use. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include covert profiling, insecure logs, manipulative nudges and unnecessary personalisation. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction

measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that privacy and security reviews should precede launch and continue after material changes. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

5.3.5 Use Balanced Performance Measures

Management dashboards should combine customer, quality, safety, recovery and business measures. This issue is central to the evaluation of AI-powered chatbots because customer satisfaction is not created by automation alone. It emerges when the technology helps a customer complete a goal with less effort, less uncertainty and an acceptable level of interpersonal sensitivity. In e-commerce, where buyers cannot physically inspect products or immediately approach a sales assistant, the conversational interface becomes part of the service encounter. Its performance therefore affects both the functional judgement of whether the problem was solved and the emotional judgement of whether the customer felt heard, respected and safe. (Parasuraman et al., 2005)

The principal mechanisms are CSAT, effort, accuracy, first-contact resolution, re-contact, escalation success, harmful-answer rate and retention. These mechanisms operate throughout the customer journey: before purchase, when a shopper compares products; during purchase, when delivery, payment or compatibility questions arise; and after purchase, when the customer seeks tracking, returns or service recovery. A well-designed chatbot reduces search costs and waiting time, translates a large catalogue into a manageable set of choices, and gives consistent explanations. The same system can lower satisfaction when it repeats irrelevant prompts, misunderstands intent or makes it difficult to reach a competent employee.

The evidence reviewed indicates that no single metric captures whether chatbot service benefits the customer and firm. The relationship is best understood as conditional rather than automatic. Speed is valuable for a routine question but is not an adequate substitute for accuracy. Personalisation is useful when it is transparent and genuinely relevant, yet intrusive when the customer cannot understand why data are being used. Human-like language can increase social presence, although excessive anthropomorphism may produce unrealistic expectations about understanding and accountability. Consequently, each attractive feature needs a corresponding governance control.

Important limitations and risks include optimising deflection, response speed or conversion in isolation. These problems explain why a narrow business case based only on cost reduction can be misleading. A chatbot may deflect contacts from human agents while quietly increasing repeated contacts, abandonment or distrust. Satisfaction measurement should therefore combine immediate post-chat ratings with resolution, re-contact, escalation, complaint, conversion and return indicators. Qualitative review of failed conversations is also necessary because aggregate scores conceal minority groups, unusual queries and high-impact errors.

For e-commerce managers, the practical implication is that teams should examine distributions and failure cases, not averages alone. Implementation should begin with high-volume, bounded intents supported by reliable product and policy data. The organisation should define confidence thresholds, clear disclosure, privacy safeguards and rapid human handover. Continuous improvement must be based on conversation logs that are lawfully collected, anonymised where appropriate and reviewed by multidisciplinary teams. In this way, the chatbot is treated as a service system embedded in people, processes and policies, not as an isolated software widget.

5.4 Implementation Roadmap

Phase	Core actions	Decision gate
1. Discover	Map journeys; analyse contact drivers; define customers and risks; establish baselines.	Use case has a measurable customer benefit.
2. Design	Choose architecture; prepare knowledge; design disclosure, accessibility and escalation.	Controls match task complexity and consequence.
3. Validate	Test accuracy, integrations, privacy, security, usability and subgroup performance.	Predefined quality thresholds are met.
4. Pilot	Release to a limited audience; compare with baseline; inspect failures daily.	Customer and safety outcomes are acceptable.
5. Scale	Expand intents gradually; train agents; maintain change control and monitoring.	Benefits persist without hidden re-contact or inequity.
6. Govern	Audit vendors, data, models and outcomes; prepare incident and rollback procedures.	Accountability and continuous improvement are demonstrable.

5.5 Recommendations for Policymakers and Researchers

Policymakers and industry bodies should promote clear disclosure of automated interaction, accessible complaint channels, privacy-by-design, truthful claims and accountability for consequential errors. Guidance should be risk based so that low-stakes FAQ assistance is not treated identically to payment, credit, health or identity decisions. Consumer education can help users understand both the convenience and limitations of conversational AI.

Researchers should conduct longitudinal and field studies that link conversational data with verified resolution, repeat purchase and complaint outcomes. More evidence is required from African markets, small and medium-sized retailers, multilingual populations and customers with disabilities. Future work should compare retrieval-based and generative

architectures, examine fairness across customer groups, and investigate how disclosure, uncertainty communication and human handover jointly influence satisfaction.

5.6 Contribution to Knowledge

The project contributes an integrated explanation of chatbot importance that combines functional quality, relational experience and governance. It shows that satisfaction is produced through a conditional pathway from capability to customer evaluation and service outcome. It also reframes human escalation from evidence of automation failure to a quality mechanism within a hybrid service system. The proposed measurement framework extends beyond immediate satisfaction ratings by including resolution, trust, safety, recovery and business outcomes.

5.7 Limitations of the Study

The project relies on secondary evidence and does not provide retailer-specific primary data. The literature is heterogeneous, rapidly changing and disproportionately based on self-reported intentions or short encounters. The absence of completed database counts in this draft means the method must be finalised by running and documenting the searches before submission. The review is limited to English-language material and may underrepresent local practices and unpublished negative results. Conclusions should therefore guide design and further testing rather than be interpreted as universal causal estimates.

5.8 Suggestions for Further Research

A longitudinal field experiment comparing chatbot-assisted and human-assisted customer journeys using verified resolution and retention outcomes.

A Ghana-focused survey and interview study on language, trust, privacy and satisfaction with e-commerce chatbots.

An audit of chatbot answer accuracy, fairness and escalation across product categories and customer groups.

A comparative study of rule-based, retrieval-augmented and generative chatbots under routine and complex service tasks.

Research on accessibility and chatbot experiences among older users, persons with disabilities and customers with low digital literacy.

5.9 Final Conclusion

AI-powered chatbots can make e-commerce service more available, responsive and personalised, thereby improving customer satisfaction. Their importance lies equally in the risks they introduce and the organisational choices they expose. Retailers that ground answers in reliable knowledge, disclose automation, protect data, measure real resolution and provide human recovery can turn conversational AI into a durable service capability.

Retailers that use it chiefly to block contact may achieve short-term savings while weakening trust and loyalty. The future of satisfactory e-commerce service is therefore not simply automated; it is responsibly designed, continuously evaluated and deliberately human supported.

REFERENCES

- Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, 31, 427–445. <https://doi.org/10.1007/s12525-020-00414-7>
- Akanferi, A. A., Asampana, I., Matey, A. H., & Tanye, H. A. (2022). Adoption of mobile commerce and mobile payments in Ghana: An examination of factors influencing public servants. *Interdisciplinary Journal of Information, Knowledge, and Management*, 17, 287–313. <https://doi.org/10.28945/4981>
- Amofah, D. O., & Chai, J. (2022). Sustaining consumer e-commerce adoption in Sub-Saharan Africa: Do trust and payment method matter? *Sustainability*, 14(14), 8466. <https://doi.org/10.3390/su14148466>
- Antwi, S. (2021). “I just like this e-retailer”: Understanding online consumers’ repurchase intention from a relationship quality perspective. *Journal of Retailing and Consumer Services*, 61, 102568. <https://doi.org/10.1016/j.jretconser.2021.102568>
- Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, 85, 183–189. <https://doi.org/10.1016/j.chb.2018.03.051>
- Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users’ satisfaction and continuance intention of AI-powered service agents. *Telematics and Informatics*, 54, 101473. <https://doi.org/10.1016/j.tele.2020.101473>
- Bhattacharjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, 25(3), 351–370. <https://doi.org/10.2307/3250921>
- Bleier, A., Harmeling, C. M., & Palmatier, R. W. (2019). Creating effective online customer experiences. *Journal of Marketing*, 83(2), 98–119. <https://doi.org/10.1177/0022242918809930>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.
- Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, 22(5), 811–817. <https://doi.org/10.1017/S1351324916000243>
- 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>
- DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, 19(4), 9–30. <https://doi.org/10.1080/07421222.2003.11045748>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., et al. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *Interactions*, 24(4), 38–42. <https://doi.org/10.1145/3085558>
- Følstad, A., & Brandtzæg, P. B. (2020). Users’ experiences with chatbots: Findings from a questionnaire study. *Quality and User Experience*, 5, 3. <https://doi.org/10.1007/s41233-020-00033-2>
- Gbolonyo, P. K., Darkwa, B. F., Antwi, S., & Adjei, E. O. (2022). To shop online or not: The role of site quality on customer satisfaction. *Open Access Library Journal*, 9, 1–27. <https://doi.org/10.4236/oalib.1109234>

- Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, 304–316. <https://doi.org/10.1016/j.chb.2019.01.020>
- Huang, M.-H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research*, 21(2), 155–172. <https://doi.org/10.1177/1094670517752459>
- Kasilingam, D. L. (2020). Understanding the attitude and intention to use smartphone chatbots for shopping. *Technology in Society*, 62, 101280. <https://doi.org/10.1016/j.techsoc.2020.101280>
- Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, 38(6), 937–947. <https://doi.org/10.1287/mksc.2019.1192>
- Mozafari, N., Weiger, W. H., & Hammerschmidt, M. (2022). Trust me, I'm a bot: Repercussions of chatbot disclosure in different service frontline settings. *Journal of Service Management*, 33(2), 221–245. <https://doi.org/10.1108/JOSM-10-2020-0380>
- Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, 17(4), 460–469. <https://doi.org/10.1177/002224378001700405>
- Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). E-S-QUAL: A multiple-item scale for assessing electronic service quality. *Journal of Service Research*, 7(3), 213–233. <https://doi.org/10.1177/1094670504271156>
- Rapp, A., Curti, L., & Boldi, A. (2021). The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies*, 151, 102630. <https://doi.org/10.1016/j.ijhcs.2021.102630>
- Republic of Ghana. (2008). Electronic Transactions Act, 2008 (Act 772).
- Republic of Ghana. (2012). Data Protection Act, 2012 (Act 843).
- Sheehan, B., Jin, H. S., & Gottlieb, U. (2020). Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research*, 115, 14–24. <https://doi.org/10.1016/j.jbusres.2020.04.030>
- 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>
- Shumanov, M., & Johnson, L. (2021). Making conversations with chatbots more personalized. *Computers in Human Behavior*, 117, 106627. <https://doi.org/10.1016/j.chb.2020.106627>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management*, 29(5), 907–931. <https://doi.org/10.1108/JOSM-04-2018-0119>
- Xu, A., Liu, Z., Guo, Y., Sinha, V., & Akkiraju, R. (2017). A new chatbot for customer service on social media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3506–3510). <https://doi.org/10.1145/3025453.3025496>
- Yeboah, E., Boateng, R., Owusu, A., Afful-Dadzie, E., & Ofori-Amanfo, J. (2020). Assessing the role of trust in merchant adoption of mobile payments in Ghana. In M. Hattingh et al. (Eds.), *Responsible Design, Implementation and Use of Information and Communication Technology* (pp. 204–215). Springer. https://doi.org/10.1007/978-3-030-44999-5_17

APPENDIX A: REVIEW AUDIT TRAIL TEMPLATE

Complete this appendix after personally running the searches. The blank fields are intentional and must not be replaced with estimated or invented numbers.

Database	Date searched	Search string / limits	Records
Scopus	_____	_____	_____
Web of Science	_____	_____	_____
ABI/INFORM	_____	_____	_____
ScienceDirect	_____	_____	_____
Emerald Insight	_____	_____	_____
SpringerLink	_____	_____	_____
Google Scholar	_____	First ____ results screened	_____

PRISMA-style flow fields

Records identified: _____

Duplicates removed: _____

Titles/abstracts screened: _____

Records excluded: _____

Full texts assessed: _____

Full texts excluded (with reasons): _____

Studies included in synthesis: _____

APPENDIX B: CORE EVIDENCE MATRIX

Study	Design	Focus	Use in synthesis
Araujo (2018)	Experiment	Anthropomorphic cues and agency framing	Human-like design shapes agent and company perceptions.
Luo et al. (2019)	Field experiment	AI identity disclosure	Disclosure can change customer purchase response.
Ashfaq et al. (2020)	Survey/model	Satisfaction and continuance	Satisfaction is central to continued use of AI service agents.
Adam et al. (2021)	Experiments	Chatbot design and compliance	Design choices influence customer response to service requests.
Shumanov & Johnson (2021)	Experimental studies	Personalisation	Fit between personalisation and user characteristics affects response.
Sheehan et al. (2020)	Survey/experiment	Anthropomorphism and adoption	Human-like cues influence adoption through psychological evaluations.
Go & Sundar (2019)	Experiment	Humanness cues	Visual, identity and conversational cues affect perceived humanness.
Rapp et al. (2021)	Systematic review	Human-chatbot interaction	Research shows diverse user, design and contextual determinants.
Følstad & Brandtzæg (2020)	Questionnaire study	User experience	Usefulness, efficiency and interaction quality are recurring concerns.
Huang & Rust (2018)	Conceptual	AI in service	AI should be allocated according to task intelligence requirements.
DeLone & McLean (2003)	Model synthesis	IS success	Information, system and service quality influence satisfaction and benefits.
Parasuraman et al. (2005)	Scale development	E-service quality	Efficiency, availability, fulfilment and privacy structure electronic service quality.
Gbolonyo et al. (2022)	Ghanaian consumer survey	Site quality and satisfaction	Design, useful information and empathy predicted satisfaction.
Amofah & Chai (2022)	Ghanaian consumer survey	E-commerce adoption and trust	Trust helps explain sustained adoption in Ghana.
Akanferi et al. (2022)	Ghanaian public-servant survey	Mobile commerce and payments	Performance expectancy, trust and facilitating conditions shaped adoption.
Yeboah et al. (2020)	Ghanaian qualitative study	Merchant trust in mobile payments	Trust in technology and service providers was critical to adoption.

APPENDIX C: OPTIONAL QUESTIONNAIRE FOR A FUTURE EMPIRICAL STUDY

This instrument is not evidence that a survey was conducted. It is provided for future use only after supervisor review and ethics approval.

Section A: Screening and profile

Have you used an AI-powered chatbot on an e-commerce website or app during the past six months? Yes / No

Age group: 18–24 / 25–34 / 35–44 / 45–54 / 55+ / Prefer not to say

Primary device: Smartphone / Tablet / Laptop / Desktop / Other

Frequency of online shopping: Less than monthly / Monthly / Weekly / More than weekly

Most recent chatbot purpose: Product information / Recommendation / Order tracking / Payment / Return or refund / Complaint / Other

Section B: Chatbot experience

Please rate each statement from 1 = Strongly disagree to 5 = Strongly agree.

1. The chatbot responded quickly. 1 2 3 4 5
2. The chatbot was available when I needed help. 1 2 3 4 5
3. The chatbot's answers were accurate. 1 2 3 4 5
4. The chatbot provided relevant information. 1 2 3 4 5
5. The chatbot was easy to use. 1 2 3 4 5
6. The conversation was clear and understandable. 1 2 3 4 5
7. The chatbot remembered relevant information during the conversation. 1 2 3 4 5
8. The recommendations matched my needs. 1 2 3 4 5
9. I understood that I was interacting with an automated system. 1 2 3 4 5
10. I trusted the information provided by the chatbot. 1 2 3 4 5
11. I felt comfortable with how my data were used. 1 2 3 4 5
12. The chatbot helped me complete my task. 1 2 3 4 5
13. It was easy to reach a human employee when necessary. 1 2 3 4 5
14. My issue was resolved satisfactorily. 1 2 3 4 5
15. Overall, I was satisfied with the chatbot service. 1 2 3 4 5
16. I would use the chatbot again. 1 2 3 4 5
17. The chatbot improved my overall experience with the e-commerce firm. 1 2 3 4 5

Section C: Open-ended questions

What was the most helpful aspect of the chatbot interaction?

What problem, if any, did you experience?

What should the e-commerce firm change to improve chatbot service?

APPENDIX D: PROPOSED INTERVIEW GUIDE

1. Please describe your most recent interaction with an e-commerce chatbot.
2. What were you trying to achieve, and was the task completed?
3. How did response speed influence your experience?
4. How accurate and relevant did the information appear?
5. Did the chatbot's tone feel appropriate? Why?
6. Were you comfortable with the data requested or used?
7. Did you know you were interacting with AI, and did that matter?
8. What happened when the chatbot could not help?
9. How easy was it to reach a person?
10. What would make you trust and use such a chatbot again?

APPENDIX E: CHATBOT IMPLEMENTATION AND QA CHECKLIST

- ☐ A named owner and measurable customer outcome exist for each use case.
- ☐ Responses are grounded in approved and current knowledge.
- ☐ Product, price, inventory, order and policy integrations are tested end to end.
- ☐ The interface clearly discloses that it is automated.
- ☐ Privacy notice and consent choices are understandable.
- ☐ Personal data collection is limited to what the task requires.
- ☐ Security testing covers authentication, access control, logging and prompt-related threats.
- ☐ Accuracy and refusal thresholds are defined by task risk.
- ☐ Customers can request a human at any point.
- ☐ Escalation preserves transcript and relevant context.
- ☐ Accessibility is tested with assistive technologies and diverse devices.
- ☐ Language and subgroup performance are evaluated.
- ☐ Post-chat satisfaction is paired with resolution and re-contact measures.
- ☐ Failed conversations are sampled and reviewed regularly.
- ☐ Incident response, rollback and vendor accountability procedures are documented.