EXPLORING THE AUTONOMY POTENTIAL OF ARTIFICIAL INTELLIGENCE AGENTS IN OMNICHANNEL CUSTOMER SERVICE

Lappeenranta–Lahti University of Technology LUT

Master's degree programme in Industrial Engineering and Management

2026

Oskari Alamäki

Examiners: Docent, Post-Doctoral Researcher Lasse Metso

Post-Doctoral Researcher Minnu Joseph

ABSTRACT

Lappeenranta–Lahti University of Technology LUT
School of Engineering Sciences
Industrial Engineering and Management

Oskari Alamäki

Exploring the autonomy potential of artificial intelligence agents in omnichannel customer service

Master's thesis
2026
95 pages, 22 figures, 7 tables and 4 appendices
Examiners: Docent, Post-Doctoral Researcher Lasse Metso and Post-Doctoral Researcher Minnu Joseph
Keywords: Artificial Intelligence, Omnichannel Customer Service, Customer Experience, Service Quality, Natural Language Processing, Agent-Based Systems, Customer Relationship Management

This thesis examines how autonomous AI agents can be utilised in omnichannel customer service in a controlled and meaningful manner using current technologies. The study addresses a gap in existing research, which has focused primarily on single-channel chatbot deployments rather than coordinated multi-agent architectures in enterprise customer relationship management environments. A design science research approach is adapted, combining a structured literature review, qualitative interviews with six industry professionals, and the design, implementation, and evaluation of a multi-agent customer service artefact within Microsoft Dynamics 365.

The artefact consists of four coordinated agents operating across four intake channels. Empirical evaluation was conducted through two simulation rounds comprising 15 test cases across four complexity tiers, with performance assessed using the SERVQUAL service quality framework. The refined simulation round produced a mean SERVQUAL score of 4.83 out of 5.00, with correct escalation decisions in all cases and a 71% reduction in hallucination events compared to the baseline round.

The findings demonstrate that autonomy in customer service is a gradual, context-dependent property shaped by task complexity, emotional sensitivity, and operational risk rather than a binary capability. Autonomous handling proved effective in routine and transactional scenarios, while emotionally complex and legally sensitive cases consistently required escalation. A key practical finding is that effective autonomous operation depends on well-defined service processes. Without formally documented policies and escalation criteria, agents cannot apply them consistently. The study contributes an empirically grounded account of how controlled, incremental autonomy can be operationalised within an existing enterprise CRM environment.

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT
LUTin insinööritieteiden tiedekunta
Tuotantotalous

Oskari Alamäki

Tekoälyagenttien itsenäisen toiminnan mahdollisuuksien tutkiminen monikanavaisessa asiakaspalvelussa

Tuotantotalouden diplomityö
2026
95 sivua, 22 kuvaa, 7 taulukkoa ja 4 liitettä
Tarkastajat: Dosentti, Tutkijatohtori Lasse Metso ja Tutkijatohtori Minnu Joseph
Avainsanat: Tekoäly, Monikanavainen asiakaspalvelu, Asiakaskokemus, Palvelun laatu, Luonnollisen kielen käsittely, Agenttipohjaiset järjestelmät, Asiakassuhteiden hallinta

Tässä diplomityössä tutkitaan, miten autonomisia tekoälyagentteja voidaan hyödyntää monikanavaisessa asiakaspalvelussa hallitusti ja tarkoituksenmukaisesti nykyteknologialla. Tutkimus vastaa olemassa olevaan tutkimusaukkoon, sillä aiempi kirjallisuus on keskittynyt pääasiassa yksikanavaisiin chatbot-ratkaisuihin eikä koordinoituihin moniagenttiarkkitehtuureihin asiakkuudenhallintaympäristöissä. Tutkimusmenetelmänä käytetään suunnittelutieteellistä tutkimustapaa, joka yhdistää systemaattisen kirjallisuuskatsauksen, haastattelut sekä moniagenttisen asiakaspalveluartefaktin suunnittelun, toteutuksen ja arvioinnin Microsoft Dynamics 365 -ympäristössä.

Artefakti koostuu neljästä koordinoidusta agentista, jotka toimivat neljän sisäänottokanavan kautta. Empiirinen arviointi toteutettiin kahdessa simulaatiokierroksessa 15 testitapauksen avulla neljällä monimutkaisuustasolla, ja suorituskykyä arvioitiin SERVQUAL-palvelunlaatuviitekehyksen avulla. Toinen simulaatiokierros tuotti SERVQUAL-keskiarvon 4,83/5,00. Samalla kaikissa tapauksissa tehtiin oikea päätös siirtää tapaus ihmisen käsiteltäväksi, ja hallusinaatiotapaukset vähenivät 71% lähtötasoon verrattuna.

Tulokset osoittavat, että autonomia asiakaspalvelussa ei ole binäärinen vaan asteittainen ja kontekstisidonnainen ilmiö, johon vaikuttavat tehtävän monimutkaisuus, emotionaalinen herkkyys ja operatiivinen riski. Autonominen käsittely osoittautui tehokkaaksi rutiininomaisissa ja selkeästi rajatuissa palvelutilanteissa, kun taas emotionaalisesti monimutkaiset ja oikeudellisesti herkät tapaukset edellyttivät johdonmukaisesti siirtämistä ihmisen käsiteltäväksi. Keskeinen käytännöllinen löydös on, että hallittu autonominen toiminta edellyttää hyvin määriteltyjä palveluprosesseja. Ilman selkeästi dokumentoituja käytäntöjä ja siirtämiskriteerejä agentit eivät pysty soveltamaan niitä johdonmukaisesti. Tutkimus tuottaa empiirisesti perustellun kuvauksen siitä, miten kontrolloitu ja asteittainen autonomia voidaan toteuttaa olemassa olevassa CRM-ympäristössä.

ACKNOWLEDGEMENTS

The completion of this work is a sign that my six-year journey at LUT is nearing its end. A lot has happened in those six years: many new friendships, unforgettable memories, student association activities, finding a life partner, and a little bit of studying as well. I wouldn't trade that time for anything.

Thank you, Matti and Johan, for your guidance on this thesis; your insights and expertise kept me on the right track whenever I needed it. Thank you to Bofor for making this thesis possible, and to all my colleagues for your thoughts and interest in my work. Thank you, Lasse and Minnu, for your guidance and feedback throughout the process.

A huge thank you to my family and my partner, Kaisa. Your belief in me and the encouragement you have shown throughout this journey means the world to me. You have always pushed me to pursue my dreams, and I am grateful for that every day.

Espoo 24.5.2026

Oskari Alamäki

ABBREVIATIONS

ACS	Azure Communication Services
AI	Artificial Intelligence
API	Application Programming Interface
ARC	Automatic Record Creation
BERT	Bidirectional Encoder Representations from Transformers
CA	Conversational Agents
CRM	Customer Relationship Management
CX	Customer Experience
DL	Deep Learning
DSR	Design Science Research
EU	European Union
GPT	Generative Pre-trained Transformer
HRL	Hierarchical Reinforcement Learning
IVR	Interactive Voice Response
LLM	Large-Language Model
MAS	Multi-agent Systems
ML	Machine Learning
NER	Named Entity Recognition
NLP	Natural Language Processing
VA	Voice Agents

Table of contents

Abstract

Acknowledgements

Abbreviations

1 Introduction 11
1.1 Background 11
1.2 Research objective and limitations 12
1.3 Methodology and data 14
1.4 Research structure 15
2 Customer service and its impact on customer experience 17
2.1 Customer experience 17
2.2 Understanding the importance of customer service 20
2.3 Service quality, customer satisfaction and customer loyalty 21
2.4 Customer service quality measurement 24
3 Natural language processing and agentic artificial intelligence for omnichannel customer service 27
3.1 Definition of natural language processing and artificial intelligence 27
3.2 Key components of natural language processing 29
3.3 Applications of natural language processing in customer service 32
3.4 Artificial intelligence agent architecture 41
4 Design science research: empirical part 48
4.1 Design science research method 48
4.2 Design science research process 50
4.3 Problem identification and objective definition 52
4.4 Solution artefact design 58
4.5 Solution development 61
4.6 Solution demonstration and evaluation 69
5 Discussion 78
6 Conclusions 82
References 86

Appendices

Appendix 1. Interview topics, questions and goals

Appendix 2. SERVQUAL evaluation rubric

Appendix 3. Voice channel intake agent interaction – T06

Appendix 4. Customer service agent interaction – T06

Figures

Figure 1: TCQ definition for customer experience

Figure 2: Impact of customer support service

Figure 3: Relationship between service quality, customer satisfaction and customer loyalty

Figure 4: Impact of service quality dimensions on customer satisfaction and customer loyalty

Figure 5: Relationship between AI, ML, DL, LLMs and NLP

Figure 6: Categorisation of NLP

Figure 7: NLG process stages

Figure 8: NLP implementation steps in Customer Service

Figure 9: NLP applications

Figure 10: Example architecture of a chatbot

Figure 11: Example of sentiment analysis

Figure 12: Example of NER

Figure 13: Example of document summarisation

Figure 14: Example of an individual AI agent architecture

Figure 15: Difference between AI agent and agentic AI

Figure 16: The DSR methodology process model

Figure 17: DSR process implemented

Figure 18: The problem and objective definition process

Figure 19: High-level architecture

Figure 20: Architecture of AI Voice Channel

Figure 21: Architecture of Chat Agent

Figure 22: ER diagram of the Dataverse

Tables

Table 1: Research structure

Table 2: NLP applications

Table 3: Semi-structured interview background

Table 4: Requirement for the DSR artefact

Table 5: Case evaluation

Table 6: SERVQUAL dimension scores across simulation rounds

Table 7: Synthesis of theoretical themes and empirical findings

DECLARATIONS

Turnitin

The originality of this thesis has been reviewed with the Turnitin similarity checking service.

AI usage

The author of the thesis, Oskari Alamäki, used the following AI tools during the preparation of the thesis:

1. Litmaps
    a. Purpose of use: Reference exploring
    b. Explanation of the use of the tool: I used Litmaps to explore relevant literature for my literature review.
2. Copilot
    a. Purpose of use: Transcription of interviews
    b. Explanation of the use: Transcription and recording of interviews
3. Grammarly
    a. Purpose of use: For checking grammar and spelling
    b. Explanation of the use of the tool: Checking spelling and consistency
4. Claude
    a. Purpose of use: Setting up the simulation environment
    b. Explanation of the use: Claude and the Dataverse MCP server were used to create flights, contact details, and existing reservations.

Responsibility

The author, Oskari Alamäki, takes full responsibility for the content of this thesis and has reviewed and edited the content generated by the possible use of AI tools.

1 Introduction

The purpose of this thesis is to investigate how autonomously artificial intelligence (AI) agents can operate in the current Microsoft Dynamics 365 customer relationship management (CRM) system. The study designs and develops an agent-based AI solution architecture that operates as independently as possible within customer service channels, including chat, email, support ticket, and voice. The goal of the solution is to evaluate how autonomous customer service can be developed in the Dynamics CRM system using current technologies.

1.1 Background

Customer satisfaction and trust generated through quality customer service are crucial factors in today's competitive market. Companies that offer excellent customer service often see improvements in customer experience (CX), leading to better financial results and longer-lasting customer relationships. (Sharif & Sidi Lemine, 2024) At the same time, customer demands for service levels have increased significantly, with customers expecting quicker, more personalised service, prompting companies to accelerate and enhance their customer service processes. (Schweiger et al., 2024) With the rapid development of AI, companies have begun to use it to streamline their customer service processes and reduce costs (Pathak & Bansal, 2025).

Although AI agents and their use in customer service have been studied, the current literature focuses primarily on single-channel solutions, limited automation, or chatbot functionality (Khanna et al., 2025; Yatawara et al., 2025). Research on the autonomy of AI agents, particularly in an omnichannel Microsoft Dynamics 365 CRM environment, remains limited. Therefore, this thesis addresses the extent to which agent autonomy can be achieved with current technology and the conditions under which such autonomy is justified from a service quality perspective.

1.2 Research objective and limitations

AI agents have become a prominent topic in the IT industry in recent years. The company conducting research has recognised the potential of AI in customer service and aims to investigate how autonomously AI can perform various customer service tasks within a controlled enterprise environment. In addition to autonomy, the study assesses how effectively AI agents can serve customers across different interaction types and levels of complexity. By understanding the current capabilities and boundaries of agent-based systems, the research company will be better equipped to serve its customers and provide grounded insights into the current state of AI agents and the near-term conditions for their responsible deployment.

The objective of this thesis is to examine how autonomous AI agents can be utilised in customer service in a controlled and meaningful manner using current technologies. Rather than aiming to maximise agent autonomy, the study focuses on identifying the conditions under which autonomous behaviour can be applied without compromising service quality, transparency, or organisational accountability. The research is conducted within an omnichannel customer service context and is grounded in a concrete enterprise environment based on Microsoft Dynamics 365 and associated cloud technologies.

Based on the research objective, three research questions have been formulated to guide the investigation:

"What are the key factors that determine customer service quality in AI supported omnichannel environments, and under what conditions is autonomous handling appropriate?"

"How do autonomous AI agents perform across different customer service scenarios, and what service quality outcomes emerge across interaction types of varying complexity?"

"What is the current potential of autonomous AI agents in customer service, and what architectural and operational conditions are necessary for their controlled and incremental deployment?"

The first research question establishes an analytical basis for the study by clarifying the role of customer service quality in AI supported environments and identifying the key factors

that determine when autonomous handling is appropriate and when escalation to a human agent is required. The question is addressed through a structured literature review and qualitative interviews with industry professionals, which together establish a theoretical and empirically grounded framework for evaluating the agent solution.

The second research question moves from concepts to empirical evaluation. It is addressed through a Design Science Research (DSR) process in which a multi-agent software architecture is specified and implemented, and subsequently evaluated against fifteen test cases spanning four complexity tiers. Agent performance is assessed using the SERVQUAL service quality framework across two simulation rounds, enabling systematic comparison of baseline and refined agent behaviour across routine, transactional, complaint-handling, and crisis scenarios.

The third research question integrates and interprets the findings from a forward-looking perspective. It is addressed by triangulating results from the literature review, qualitative interviews, and the empirical evaluation of the implemented artefact. The synthesis assesses the practical implications of controlled autonomy for the future of customer service, examines the architectural and operational mechanisms that enable safe, incremental deployment of agents within an enterprise CRM environment, and outlines near-term opportunities and constraints grounded in evidence accumulated across the previous stages.

This study focuses on customer service interactions handled through the Dynamics 365 Customer Service Workspace, including voice, email, chat, and ticketing channels. The empirical evaluation, however, is conducted through simulation using the ticketing channel as the primary interface for test case execution, as this provides the most controlled and reproducible evaluation environment. The designed and developed artefact is limited to Microsoft technology, including Azure, Teams, Power Platform, and Microsoft Dynamics business applications. The study does not include live production deployment, nor does it evaluate the full end-to-end performance of all intake channels in real customer use. Findings are therefore interpreted within this specific technical and organisational context and should not be generalised to other CRM platforms or service domain configurations without further validation.

1.3 Methodology and data

This study follows a qualitative research design that combines a structured literature review, qualitative interviews, and a DSR approach. The methodological logic of the study is sequential. First, a literature review was conducted to establish the theoretical foundation of the research. Second, qualitative interviews were conducted to deepen understanding of the practical problem context, current customer service practices, and organisational expectations regarding AI supported customer service. Third, the study applied DSR to design, develop, demonstrate, and evaluate a concrete artefact informed by insights from the literature and interviews.

The literature review provides the conceptual basis of the study. It brings together models, theories, and prior research related to customer experience, customer service, service quality, natural language processing, and agentic AI in customer service environments. Through this review, the study establishes the key concepts, identifies the research gap, and defines the theoretical criteria against which the artefact is later evaluated.

The interview phase was used to anchor the research in practice. Initial unstructured discussions helped narrow the topic and define the organisational problem space, while semi-structured interviews with industry professionals provided a more systematic understanding of customer service processes, AI maturity, service quality requirements, and expectations regarding autonomous customer service agents. These interviews informed the identification of business and technical requirements for the artefact and helped contextualise the theoretical findings within real organisational settings.

The empirical part of the study is conducted using the DSR approach. Within this framework, the study proceeds from problem identification and objective definition to artefact design, development, demonstration, and evaluation. The artefact is an agent-based customer service solution developed in a Microsoft Dynamics 365 environment. Its purpose is not only to demonstrate technical feasibility but also to examine the conditions under which controlled and meaningful autonomy can be achieved in customer service.

The data used in the study consists of two main sources. The first source is the literature reviewed in the theoretical section. The second source comprises qualitative interview material and customer service interaction data generated in the simulation-based evaluation

environment. The empirical evaluation is conducted in controlled scenarios using synthetic customer service cases in Finnish and English. These data sources support the study's aim of combining theoretical grounding, practical relevance, and artefact-based evaluation within a single research design.

1.4 Research structure

The research comprises two sections: a literature review and an empirical section. The first chapter introduces the thesis background, outlines the research objectives and limitations, and describes the methodology and data used in the study. The second chapter presents the theoretical foundation by examining customer service and its connection to customer experience. It introduces the core concepts of CX, discusses the role of customer service in shaping customer satisfaction and loyalty, and reviews established models for evaluating service quality.

The third chapter continues the literature review by focusing on Natural Language Processing (NLP). It defines NLP and its relation to AI, introduces key NLP techniques, and outlines their use cases in customer service. The chapter also discusses AI agents, their architecture, and the main challenges associated with their use in customer service environments.

The fourth chapter forms the empirical part of the research and follows the DSR methodology. It describes the problem identification, the goal definition, and the design and development of the solution artefact. The chapter also explains how the solution was demonstrated and evaluated in practice.

The last two chapters present the discussion and conclusions, where the results are interpreted, the research contributions and limitations are assessed, and potential directions for future research are proposed.

Table 1 Research structure

Input	Chapter	Output
Introduction to the research methods and research topic.	1. Introduction	Research background, objectives, research questions, scope, methodology, data, and overall structure of the thesis.
Literature and scientific research related to customer experience, customer service, and service quality.	2. Customer service and its impact on customer experience	Understanding of CX, its multidimensional nature, the role of customer service, and frameworks for evaluating service quality and satisfaction.
Literature and scientific research about NLP, AI, and related customer service applications.	3. Natural language processing and agentic artificial intelligence for omnichannel customer service	Knowledge of NLP within AI, its core techniques, and how NLP and AI agents can be used to support customer service tasks across channels.
Theoretical framework, insights from literature reviews, interview findings, and tools required for software development.	4. Design science research: empirical part	A designed and implemented agent-based solution architecture, including development, demonstration, and evaluation of the artefact in customer service scenarios.
Findings from literature and empirical research, including identified limitations.	5. Discussion	Deeper analysis of the results, comparison with prior research, reflection on limitations, and interpretation of how the findings relate to the research objectives.
Results of the research and research questions.	6. Conclusions	Summary of key findings, answers to research questions, concluding insights, and suggestions for future research.

2 Customer service and its impact on customer experience

This chapter provides an overview of how customer service quality is linked to CX, customer satisfaction, and customer loyalty. The chapter discusses the importance of customer service to companies and identifies the key factors that affect its quality. By identifying the factors that influence customer service quality, we can evaluate the quality of complex customer service encounters. This chapter is essential to the study, as it aims to identify which factors positively affect CX. With a solid understanding, we can comprehensively evaluate the autonomy of AI agents across different customer service channels and situations at the end of the study.

2.1 Customer experience

Every customer journey involves multiple touchpoints, their context, and the qualities they provide. CX can be regarded as a multidimensional construct encompassing the customer's emotional, cognitive, sensory, behavioural, and social responses to products, services, and interactions offered by the business throughout the customer journey. This perspective highlights that CX is not limited to a single dimension but comprises a set of interconnected experiential elements that collectively influence the customer's overall experience and value creation. (De Keyser et al., 2020) Figure 1 brings together the perspectives described above and illustrates the concept of CX, defining its key dimensions.

Figure 1 TCQ definition for customer experience (Adapted from De Keyser et al. 2020)

Pires et al. (2025) define customer experience as a multifaceted concept encompassing cognitive, emotional, sensory, and behavioural responses arising from interactions with the business, its employees, partners, and the broader environment. This multidimensional perspective extends beyond individual touchpoints, capturing the entire journey a customer undertakes across various service cycle phases (Siebert et al., 2020). Given its holistic nature, understanding customer experience necessitates an appreciation for how each interaction shapes a customer's perception of value, especially in omnichannel environments where seamless integration across diverse channels and touchpoints is paramount (Gerea et al., 2021). The cumulative effect of these interactions shapes a customer's overall perception and evaluation of a brand or product, encompassing both deliberate and spontaneous reactions to stimuli (Blümel et al., 2024). In summary, CX is the overall perception and emotional response to all interactions with a company, from initial contact through the post-purchase phase.

De Keyser et al. (2020) define the above in terms of three building blocks: touchpoints, context, and qualities (TCQ). Touchpoint is any individual moment of contact or interaction between a customer and a service provider during the customer journey. Context refers to the personal, social, market, and environmental factors that shape how a customer perceives and interacts with a service provider at each touchpoint, thereby making every experience unique. Qualities are the attributes that describe how customers respond to interactions with a service provider. (De Keyser et al., 2020; Hung et al., 2025)

Touchpoints are the most crucial aspect of CX, because without any actual or perceived interaction between the business and the customer, there can be no CX. These touchpoints vary depending on the control, nature, and stage of the customer journey. (De Keyser et al., 2020; Hung et al., 2025) Numerous studies have found that factors shaping customer experience are either within the business's control or beyond its influence. Business-controlled touchpoints are planned and managed by the company. They typically include physical stores, websites, advertising, and employees. Non-business-controlled touchpoints are typically managed by customers, influencers, and other companies and can also be used to manage the overall customer experience. (Hung et al., 2025; Kim & Fung So, 2024) The nature of the touchpoint reflects how the company is represented there. A touchpoint can be human (e.g., a customer service representative), digital (e.g., a mobile app), physical (e.g., a store environment), or a combination of these. (De Keyser et al., 2020; Kim & Fung So,

2024) The touchpoint level refers to the specific stage in the customer journey at which the touchpoint occurs. In the literature, these are divided into three phases: pre-purchase, purchase, and post-purchase. (Blümel et al., 2024; De Keyser et al., 2020; Hung et al., 2025) The pre-purchase phase encompasses all interactions between the customer and the touchpoint prior to the purchase decision. The purchase phase encompasses all interactions and actions during the purchase decision, whereas the post-purchase phase is defined by touchpoints during actual use or consumption. (De Keyser et al., 2020)

Even if touchpoints are identical across customers, a customer may experience them very differently at different times, depending on the context of their interaction with the company (De Keyser et al., 2020). According to Hung et al. (2025), context can be divided into two categories: internal and external. External context elements include social, market, and environmental factors, while internal context includes individual components such as cognitive, emotional, social, physical, and economic factors. The internal context reflects the customer's current personal state at the touchpoint and can hence be multidimensional. (Blümel et al., 2024; De Keyser et al., 2020; Hung et al., 2025) Unlike internal elements, external context is not person-dependent. It is shaped by the prevailing environment and norms. The social context encompasses social rules and standards that reflect the customer's current social circumstances. The market context refers to the circumstances arising from the market players with whom the customer interacts. These may include complementary products, competitors, substitute products, and new market entrants. (De Keyser et al., 2020; Hung et al., 2025) The environmental context is a broader concept, encompassing external factors such as natural, economic, and political ones, or a combination of these. Natural environmental factors include, for example, outdoor temperature or prevailing climate. Economic and environmental factors include, for example, commodity prices such as gasoline, gross domestic product growth, and the overall economic situation. Finally, political environmental factors refer to changes in consumption habits due to political decisions or events. (De Keyser et al., 2020)

Finally, we define qualities that describe customers' reactions to interactions at touchpoints, comprising five parts: participation level, time flow, valence, ordinariness, and dimensionality (Hung et al., 2025). De Keyser et al. (2020) define participation level as the extent to which people actively engage with the business. Some experiences require more customer input than others, such as the design of kitchen materials and surfaces.

Dimensionality refers to the scope and variety of different dimensions, including cognitive, emotional, sensory, social, and behavioural aspects that are activated during the customer journey. The time flow describes the perceived duration and rhythm of an experience, ranging from brief, momentary experiences to long, dynamic ones. Valence refers to the perceived tone of an experience, which can range from negative to positive and also includes simultaneous, conflicting emotions. Ordinariness describes the degree of everydayness on a continuum ranging from low-intensity, everyday experiences to exceptional, rare, and intense ones.

2.2 Understanding the importance of customer service

Customer service encompasses all interactions between the customer and the company, aiming to ensure customer satisfaction and smooth business operations through trust, communication, and ease of doing business. It plays a pivotal role in establishing competitive advantage and facilitating sustained organisational success by improving the overall customer experience (Metz et al., 2020; Slack & Singh, 2020). According to Sheth et al. (2023), approximately 90% of customers are willing to return or provide positive word-of-mouth (WOM) recommendations after a positive interaction with support services. Sharif and Sidi Lemine (2024) emphasise the importance of customer service in strengthening the brand. Through positive customer service experiences, customer relationships become profitable for the organisation, helping to create and enhance competitive advantage. Patti et al. (2020) note that integrating customer service into broader business strategies is recognised as essential for value creation and sustained competitive advantage.

Initially, customer service mainly involved direct, personal interactions centred on resolving problems, often limited to the point of sale or immediate post-purchase support. (Metz et al., 2020). Over time, as markets expanded and communication technologies advanced, businesses began to recognise the broader implications beyond simple transactional closure. This realisation led to the creation of dedicated customer service departments, often operating via telephone, to handle enquiries and complaints more systematically. However, this centralised approach, while seeking efficiency, often led to customer dissatisfaction due to impersonal interactions, long wait times, and standardised responses that failed to adequately address specific issues. (Hodge et al., 2019)

These encounters have been referred to metaphorically as "moments of truth", emphasising their importance for companies, as 92% of customers form their opinion of a company based on customer service. Each such moment presents an opportunity to strengthen or weaken the customer relationship, making it critical to shape the customer experience and brand loyalty. (Metz et al., 2020; Ribeiro et al., 2024; Slack & Singh, 2020) Costa et al. (2024) define customer service as assisting customers before, during, and after the purchase of a product or service. This type of service allows customers to perform various actions, such as asking questions, accessing information about their purchases, receiving support from sales staff, submitting complaints, and requesting technical help, among other interactions. Figure 2 defines the impact of customer support services throughout the entire customer journey.

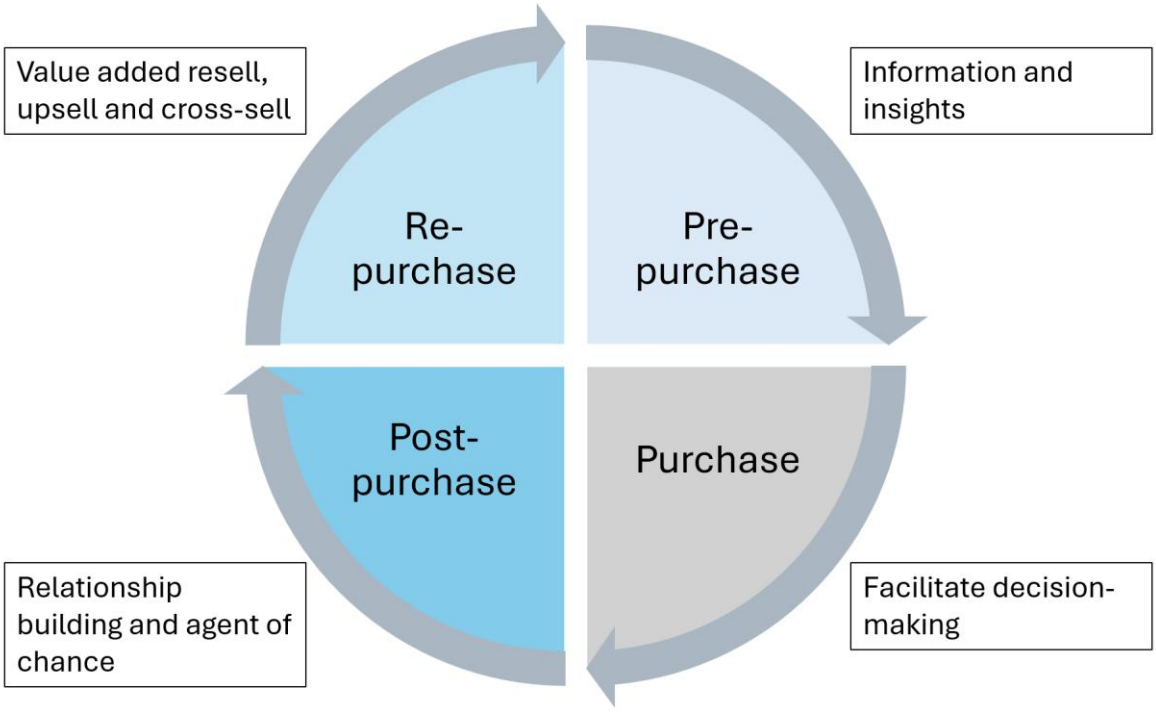

Figure 2 Impact of customer support service

2.3 Service quality, customer satisfaction and customer loyalty

Customer satisfaction is a highly subjective concept, as evidenced by numerous attempts to define it. Ahmad and Allan (2014) define customer satisfaction as the comparison that customers make between their expectations before purchase and their actual experience after purchase. Customers are satisfied when the product or service exceeds their expectations and dissatisfied when it does not, as customer satisfaction is a key indicator of the quality of the

interaction. (Ahmad & Allan, 2014) This definition dates back to 1988, when Tse and Wilton (1988) proposed the following definition of customer satisfaction: "The consumer's reaction when evaluating the perceived differences between their prior expectations and the actual performance of the product after use." To summarise the definition, customer satisfaction can be measured according to this theory: if a product or service exceeds expectations, the result is a "positive disconfirmation". If the expectations are not met, the result is a "negative disconfirmation" and if the expectations are met, the result is "confirmation" (Slack & Singh, 2020).

Among researchers, the prevailing view is that service quality plays a significant role in building customer satisfaction and loyalty (Slack & Singh, 2020). Customer satisfaction directly reflects the CX when expectations are exceeded, and it also increases the likelihood of experiencing positive feelings about a particular service (Mamakou et al., 2024). This positive evaluation, stemming from actual performance surpassing prior expectations, is a critical precursor to cultivating enduring customer loyalty (Hodge et al., 2019). When customer expectations are met or exceeded, satisfaction rises, leading to increased repurchases and, in turn, loyalty. This underscores customer satisfaction as a strong predictor of loyalty, making it a fundamental focus for businesses seeking sustainable growth and market share. Furthermore, a lack of essential service components can lead to profound dissatisfaction, even when other aspects of the experience are satisfactory. (Sheth et al., 2023) This heightened satisfaction, influenced by subjective emotions and perceptions of service performance, directly correlates with higher customer retention rates and augmented revenue at the account level (Hodge et al., 2019).

As satisfaction strengthens and becomes more consistent across interactions, it naturally evolves into customer loyalty, providing the foundation for long-term behavioural commitment. Building customer loyalty is essential for businesses, as it encourages repeat purchases, helps resist competing offers, and helps customers overlook customer service failures and WOM recommendations. (El-Manstrly, 2016; Özkan et al., 2020). Customer loyalty is based on positive attitudes and behaviours toward the company, thereby preventing customers from switching to another company. Customer loyalty can be increased by improving quality, satisfaction, perceived value, reputation, and other CX factors across the company's operations. (Özkan et al., 2020) As competition intensifies, customer loyalty has become increasingly important. The costs of acquiring new customers are significantly

higher than those of maintaining existing customer relationships. (El-Manstrly, 2016; Slack & Singh, 2020) Slack and Singh (2020) point out that customer loyalty is becoming increasingly important for the following reasons.

Figure 3 shows the relationship between service quality, customer satisfaction, and customer loyalty. The relationships between them have been studied extensively (El-Manstrly, 2016; Hsu & Lin, 2023; Özkan et al., 2020; Slack & Singh, 2020), and a strong correlation has been observed between them.

Figure 3 Relationship between service quality, customer satisfaction and customer loyalty

The relationship between service quality and customer satisfaction has been examined multiple times and is widely recognised as significant. Research shows that higher perceived service quality leads to greater customer satisfaction. The quality of customer service is also directly linked to customer loyalty, though this connection is somewhat indirect. When customer satisfaction is high, it serves as a key driver of customer loyalty. Therefore, service quality can be regarded as directly influencing customer satisfaction and indirectly influencing customer loyalty. (Özkan et al., 2020; Saneva & Chortoseva, 2020; Tedjokusumo & Murhadi, 2023; Win et al., 2024)

2.4 Customer service quality measurement

Service quality has traditionally been defined as the consumer's assessment of the overall excellence of the service. Today's dynamic and competitive operating environment imposes new demands on service quality. An omnichannel, multi-touchpoint environment requires high-quality customer service across multiple channels. (Sheth et al., 2023). In the face of rapid technological change, companies increasingly seek to gain a competitive advantage by improving service quality. (Hsu & Lin, 2023) Service quality has been studied for over three decades and is widely regarded as a difficult concept to measure. However, various frameworks for measuring service quality have been developed over the years, and it is essential to use an appropriate model for service quality management. (Park et al., 2021)

Developed by Parasuraman et al. (1985), SERVQUAL has become the most widely used model for measuring service quality over the decades (Oh & Elango, 2024; Park et al., 2021). Although it was initially developed for only five industries, it has also been widely applied to other industries with minor modifications. According to SERVQUAL, service quality perception is determined by the gap between customer expectations and actual performance. Service quality is summarised in SERVQUAL into five dimensions: reliability, responsiveness, tangibility, assurance, and empathy. (Chatterjee et al., 2023; Park et al., 2021) Figure 4 illustrates the five dimensions of service quality and the previously defined relationships between service quality, customer satisfaction, and customer loyalty.

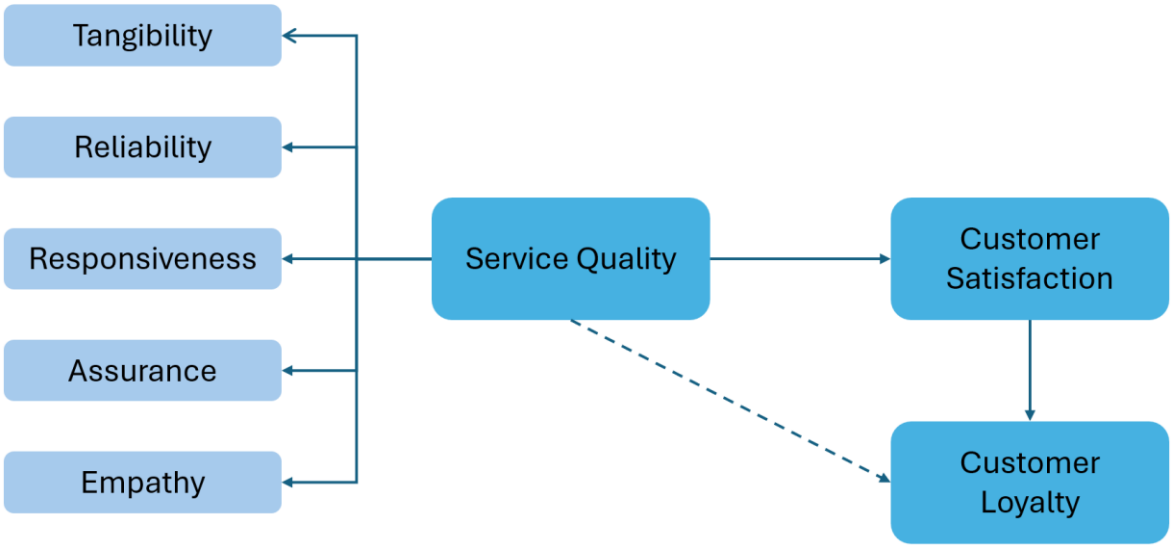

Figure 4 Impact of service quality dimensions on customer satisfaction and customer loyalty

In the SERVQUAL model, tangibility refers to the physical evidence of a service, including the appearance of facilities, equipment, and communication materials. It encompasses the visible and sensory aspects that shape customers' perceptions of service quality. (Chatterjee et al., 2023) In traditional contexts, this includes, for example, facility cleanliness and physical resources. In digital services, tangibility extends beyond traditional physical cues to include digital and experiential aspects. (Arli et al., 2024)

Reliability describes a service provider's ability to deliver promised services accurately and consistently (Tedjokusumo & Murhadi, 2023). It highlights stability, accountability, and faultless performance, which are vital for establishing trust and confidence among customers (Chatterjee et al., 2023). In the realm of online services, reliability is often seen as the key factor influencing customer trust and satisfaction. These qualities indicate operational accuracy and dedication to service standards, directly shaping customer perceptions of quality. High reliability promotes loyalty, repeat business, and positive WOM, whereas reliability failures can cause dissatisfaction and customer loss. (Arli et al., 2024)

Responsiveness refers to a service provider's willingness and ability to help customers swiftly and respond effectively to their needs. It emphasises speed, timeliness, and proactive communication. (Chatterjee et al., 2023; Tedjokusumo & Murhadi, 2023) Studies consistently show that responsiveness and customer loyalty are strongly positively correlated, and research indicates that responsiveness may even surpass reliability as the primary driver of satisfaction. Furthermore, responsiveness in digital environments extends to website functionality, where quick response times and clear communication foster trust and enhance overall service quality. Therefore, responsiveness is vital to building strong customer relationships and maintaining a competitive edge.(Arli et al., 2024)

In the SERVQUAL framework, assurance pertains to the service provider's capacity to instil confidence and certainty through their expertise, courtesy, and credibility. It encompasses factors such as security, privacy, and professionalism, which are vital for reducing customer uncertainty and perceived risk. (Chatterjee et al., 2023) Studies emphasise that assurance includes personalisation, responsiveness, and system security, all of which foster reliability and ease of use. Attractive, user-friendly websites enhance trust by conveying reliability and security, thereby encouraging repeat purchases and loyalty. For organisations, maximising

assurance through secure systems and intuitive design is a strategic necessity, as it directly influences customer satisfaction and long-term engagement in digital environments. (Arli et al., 2024)

Empathy refers to a service provider's capacity to understand and respond to each customer's needs with care and attention. It highlights appreciation and understanding, fostering trust and emotional bonds. (Chatterjee et al., 2023) Along with respect, empathy greatly enhances customer satisfaction. However, overdoing or insincerely demonstrating empathy can cause discomfort, underscoring the importance of balance. Overall, empathy remains a vital element in distinguishing service quality and creating enduring customer relationships in both traditional and digital settings. (Arli et al., 2024)

In summary, the findings in chapter 2 show that service quality is determined by several factors related to both technical performance (e.g., responsiveness, reliability) and soft elements of interaction (e.g., empathy, ease of use). These factors will later serve as evaluation criteria for AI agents in the study, enabling us to examine the extent to which agent autonomy supports or undermines the overall service experience.

3 Natural language processing and agentic artificial intelligence for omnichannel customer service

This chapter introduces NLP and defines it within the scope of AI. It explores recent advances in NLP and offers some background on its history. The significance of NLP is highlighted through various applications and by familiarising readers with their use cases in customer service. The chapter provides a thorough overview of the current state of NLP and the opportunities it offers in customer service.

3.1 Definition of natural language processing and artificial intelligence

AI is one of the fastest-growing markets. The global AI market was estimated to be approximately US$371.71 billion in 2025 and is projected to reach US$2,407.02 billion by 2032. (MarketsandMarkets, 2025) Although the AI market is growing rapidly, people do not fully understand the relationships among terms such as AI, machine learning (ML), deep learning (DL), large language models (LLMs), and NLP. Understanding these terms and their relationships is essential for a deeper study of NLP. Figure 5 illustrates the relationships between the terms.

AI refers to machine-based systems that make predictions, generate recommendations, or make decisions based on objectives defined by humans, thereby affecting real or virtual environments (Dalton, 2024). In simple terms, AI is a field of computer science that focuses on machines' ability to perform cognitive functions we associate with the human mind, such as perception, problem-solving, learning, decision-making, and creativity. (Dennis et al., 2023) ML is a subfield of AI that studies how to build programs that can independently improve their performance through experience. One of the main goals of ML is to teach algorithms to interpret data successfully. ML has become one of the most widely used AI techniques, and most ML-based tasks are precisely defined to optimise specific functions. (Dalton, 2024) DL is a branch of ML based on computational models called neural networks. The learning and information-processing mechanisms of the human brain have inspired the development of neural networks. (Bishop & Bishop, 2024) NLP is a branch of AI that combines computational linguistics, ML, and DL. NLP focuses on converting spoken or

written text into datasets that can be interpreted by computer systems. (Just, 2024) LLM is a combination of DL, ML, and NLP definitions (Amaratunga, 2023). LLM relies heavily on ML techniques, particularly deep neural networks, for processing human language and text. By leveraging NLP, LLMs achieve better results in complex tasks such as sentiment analysis, translation, text generation, text analysis, and question answering. (De Nicola et al., 2024; Raiaan et al., 2024)

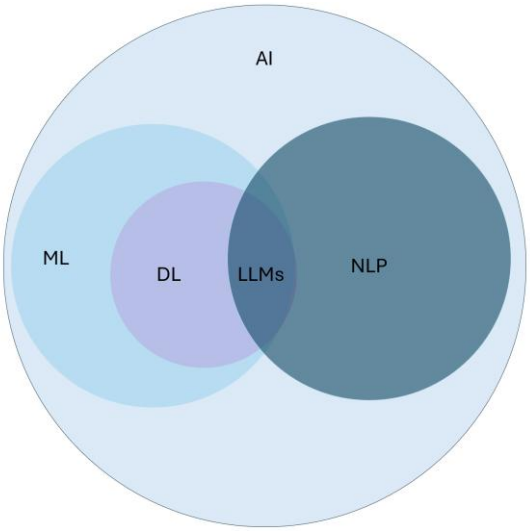

Figure 5 Relationship between AI, ML, DL, LLMs and NLP (adapted from Amaratunga, 2023)

NLP technology was initially developed to ease users' work by allowing them to communicate in natural language rather than machine language. With NLP, machines can understand, analyse, and generate human language, reducing the gap between how humans communicate and how machines process information. (De Nicola et al., 2024; Khurana et al., 2023) Feng (2023) proposes that the study and processing of natural language by computers should include four areas:

1. Formalisation: The problem to be investigated should be formulated in terms of linguistics, and a formal model created for the language. After this, the problem can be precisely formulated in mathematical terms.
2. Algorithmization: The mathematical form should be expressed as an algorithm.
3. Programming: Writing and implementing an algorithm-based computer program so that we can create a practical NLP system.

4. Practicalization: Evaluating the developed NLP system and continuously enhancing its quality and performance based on user requirements.

As shown in the four areas defined above, NLP is a multidisciplinary field of research encompassing linguistics, information technology, and mathematics. (Feng, 2023)

NLP research has made great strides in recent years, with the development and widespread adoption of bidirectional encoder representations from transformers (BERT) and generative pre-trained transformer (GPT) models. These models have achieved state-of-the-art results across many NLP tasks, including text classification, machine translation, and question answering. (Mariprasath et al., 2025) The primary reason for the rise of BERT and GPT models is the introduction of the Transformer architecture in 2017. These models differ from traditional models in that they employ an attention mechanism. Designed to avoid recurrence and convolution, they improved the models' ability to process tasks in parallel and reduced training costs. With the Attention mechanism, each word in the input text is assigned a specific weight, enabling the model to ignore meaningless sentence structures and focus on the parts that matter most for the outcome. (Seidel et al., 2025)

3.2 Key components of natural language processing

The literature suggests that NLP tasks can be broadly divided into two categories: natural language understanding (NLU) and natural language generation (NLG) (Joshi et al., 2025; Mishra et al., 2024; Shaik, 2025). NLU covers tasks related to language comprehension, such as dialect recognition, emotion classification, and morphosyntactic analysis. NLG, on the other hand, includes tasks where both the input and output are text sequences, such as summarisation. Figure 6 shows the classification of NLP into NLG and NLU.

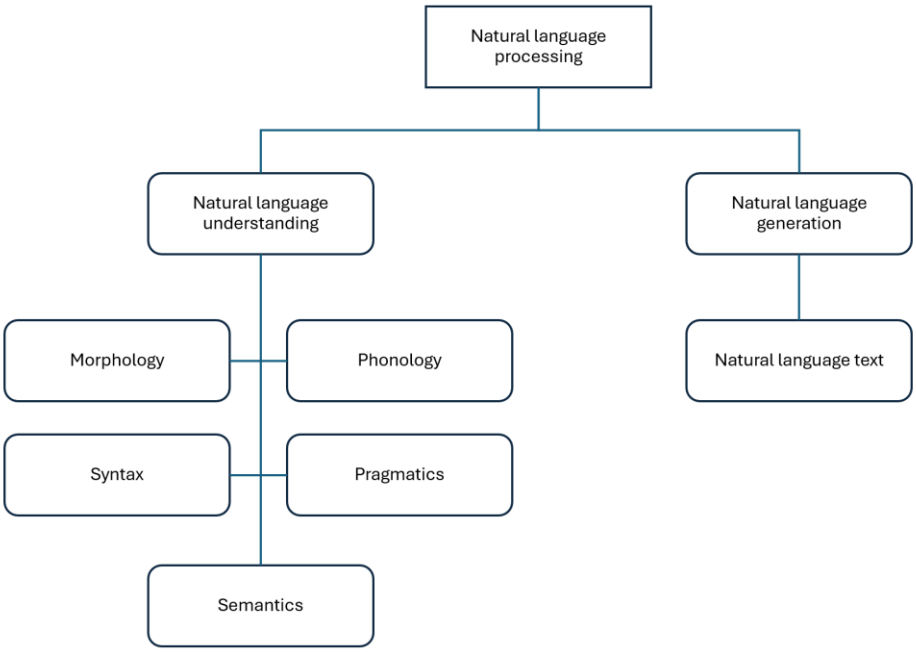

Figure 6 Categorisation of NLP (Mishra et al., 2024)

NLU is considered the foundation of NLP, enabling machines to interpret and analyse natural language by extracting concepts, emotions, and entities. It is often used, for example, in customer service applications to understand problems reported verbally or in writing by customers. (Khurana et al., 2023; Mishra et al., 2024; Shaik, 2025) To better understand NLU, we need to examine the five linguistic domains it covers: phonology, morphology, pragmatics, syntax, and semantics. Phonology is a branch of linguistics that studies the functioning, behaviour, and organisation of speech sounds as linguistic units. (Mishra et al., 2024) Morphology looks into the structure and meaning of words. Morphemes are the smallest meaningful segments, divided into lexical (e.g., table) and grammatical (e.g., -ed, -ing). Bound morphemes can be inflectional, which modify grammatical forms, or derivational, which change a word's class and meaning. (Khurana et al., 2023; Mishra et al., 2024) Syntax involves the rules of language structure that determine how words are combined to form sentences and clauses. It explores word order, grammatical relationships, and structural dependencies to ensure sentences are meaningful and grammatically correct. (Khurana et al., 2023) Semantics concerns the study of meaning in language. It focuses on what words, sentences, and texts communicate and how meaning is constructed within context. Semantics not only considers word order but also aims to understand the relationships between concepts and words. (Mishra et al., 2024) Pragmatics is a branch of linguistics that examines how language is used in context and how meaning arises from the

situation, background knowledge, and the speaker's intentions. Unlike semantics, which focuses on the literal meaning of words, pragmatics investigates the meanings words imply or infer. (Khurana et al., 2023; Mishra et al., 2024)

Although NLG is the opposite of NLU, it is related to the same areas of linguistics. In NLG, structured data or prompts are converted into clear and fluent text. Traditional NLG methods were often based on rules or models that could limit flexibility and adaptability; however, transformation-based architectures now enable data-driven, context-aware text generation. (Shaik, 2025) Khurana et al. (2023) suggest that NLG occurs in four stages: identifying goals, planning how to achieve them, assessing the situation and available communication tools, and implementing plans in text form. Language production is therefore a multi-stage process involving various tasks, each of which is important for the final outcome. These tasks include content selection, text organisation, language resource selection, and implementation. (Mishra et al., 2024) Figure 7 illustrates the NLG process flow.

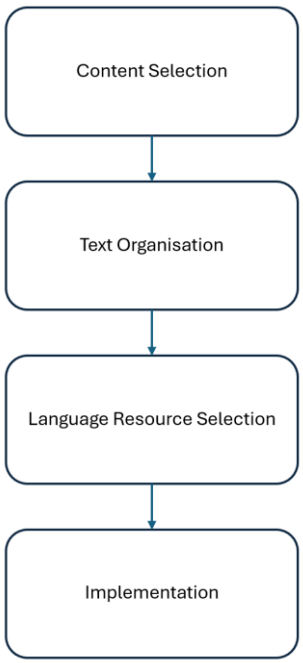

Figure 7 NLG process stages (Adapted from Khurana et al., 2023; Mishra et al., 2024)

In the content selection stage, relevant information is selected and organised according to linguistic relationships. Because NLG systems may contain redundant information, some

details may need to be removed, while others may need to be added. During the text organisation phase, the selected information from the previous stage is organised according to grammatical rules, taking into account linguistic connections, such as modifications. In selecting language resources, the necessary linguistic tools are chosen to support the realisation of the information. Ultimately, these resources are transformed into words, idioms, and syntactic structures. Finally, in the implementation phase, the chosen and organised resources are realised as actual text or speech. (Khurana et al., 2023; Mishra et al., 2024)

3.3 Applications of natural language processing in customer service

NLP has facilitated interactions between individuals and machines. Although NLP has long served as a conceptual framework, it has only recently achieved sufficient accuracy to contribute meaningfully to customer service. By utilising NLP, customer service representatives can focus on more complex tasks that require human expertise and engagement. (Olujimi & Ade-Ibijola, 2023) As digitalisation accelerates, companies are constantly looking for new NLP technologies that will help them maintain high-quality customer service while lowering the operational costs of customer service (Yu et al., 2024).

The nine-step implementation framework outlined by Juglan et al. (2023) describes the stages involved in successfully applying NLP in customer service. The initial stage consists of defining the problem that requires NLP solutions. The next stage is identifying various NLP use cases. The third stage focuses on collecting and preparing the necessary data for testing and practising NLP. Once this data has been gathered and prepared, the fourth stage begins, during which NLP is built and trained on the collected data. The fifth stage involves evaluating and testing the extent to which the NLP application meets its objectives. If it passes the criteria, it can be integrated into existing customer service systems and then transferred to monitoring mode. During this phase, it is actively supervised and optimised to ensure peak performance. Once performance is sufficiently refined, the eighth stage involves training employees to use the application and interpret its results accurately. As a final step, the model is continually refined based on user feedback or new data to improve performance. (Juglan et al., 2023) Figure 8 depicts the implementation framework.

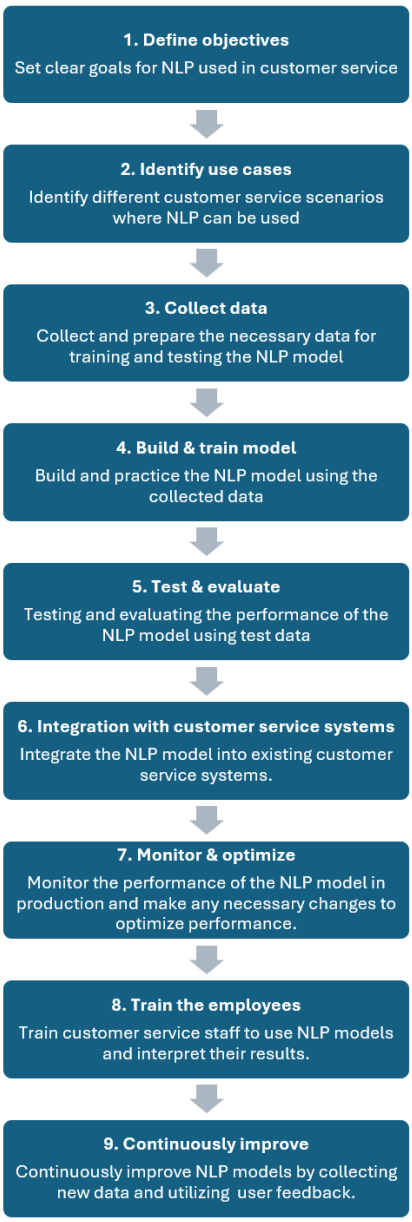

Figure 8 NLP implementation steps in Customer Service

Industrial NLP applications can generally be categorised based on their primary function and the type of human-computer interaction they facilitate. It is essential to recognise that some NLP functions act as standalone use cases, while others are incorporated as sub-tasks within larger system applications. (Pakray et al., 2025). Patwardhan et al. (2023) classified applications based on their modal degree into unimodal and multimodal applications. Unimodal models focus primarily on text processing and analysis. These categories illustrate versatile applications and features that can be achieved by focusing on text-based information. For example, sentiment analysis, machine translation, and text classification

belong to this group. Multimodal applications utilise multiple modalities, such as text, speech, images, and videos, for data processing and analysis. By integrating different data formats, these applications enable more versatile solutions. Chatbots and virtual assistants are examples of such applications. (Patwardhan et al., 2023) To gain a more detailed understanding of the various applications, we examine seven in the following subsections. Figure 9 shows all the applications that have been reviewed.

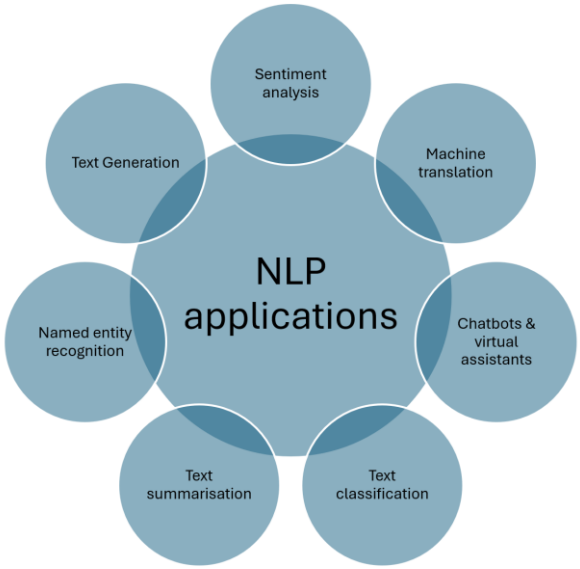

Figure 9 NLP applications (adapted from Kadre et al., 2025)

Chatbots and virtual assistants

NLP can be used in initial customer service interactions to answer questions about technology or services. However, it is essential to distinguish between rule-based and AI chatbots from the outset. (Mariprasath et al., 2025) Rule-based chatbots are based on decision trees, where the user's input determines the action or response of the if-then branch. This method is easy to implement, but its scalability and flexibility are limited. A rule-based chatbot is a viable solution when the set of possible user inputs is limited and well-defined. (Benaddi et al., 2024) AI chatbots, on the other hand, are based on DL and NLP technologies, allowing them to manage more complex and dynamic conversations. AI-driven dialogue management can process a wider variety of inputs, understand context, and adapt to different conversation styles. These features enable more personalised responses to users, which is

why AI chatbots are popular as virtual assistants or customer service bots. (Mariprasath et al., 2025)

Olujimi and Ade-Ibijola (2023) note that chatbots are commonly used in customer service, both for customer-facing purposes and as internal tools for customer service representatives. Chatbots and virtual assistants are often found on company websites, portals, social media, apps, and other digital channels. By utilising chatbots and virtual assistants across different channels, companies can serve customers much faster than with human customer service representatives. In a study conducted by Sood et al. (2024) for example, resolution time decreased by 30.4%, engagement increased by 25.7%, and customer satisfaction was high when they compared chatbot performance to human customer service.

Chatbot architecture can be classified into four categories based on query and response types: text-to-text (TTT), text-to-speech (TTS), speech-to-text (STT), and speech-to-speech (STS). TTT is a typical architecture in which the user inputs text and the chatbot responds with text. TTS receives text input and produces speech output. STT takes speech input from the user and produces text output. STS involves the user providing speech input, and the bot replying with speech. (Nirala et al., 2022)

A robust knowledge base and other data sources are essential for the chatbot to function correctly. The knowledge base may include, for example, frequently asked questions, solutions to various identified problem cases, and other internal company guidelines. Other data sources may include integrations with ERP or CRM systems, from which current or past order or customer information can be retrieved. With this information, the chatbot can provide accurate and relevant answers to customer queries. (Horvat et al., 2025; Sood et al., 2024) Figure 10 shows an example of chatbot architecture.

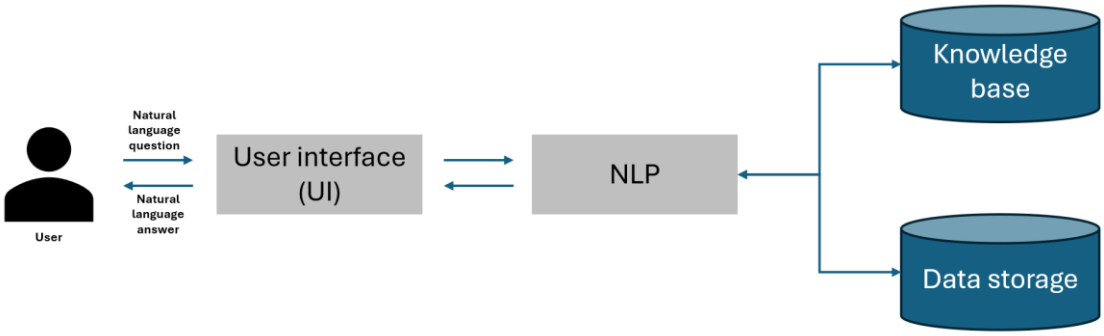

Figure 10 Example architecture of a chatbot (adapted from Sood et al., 2024)

The chatbot interacts with the user through the user interface (UI), where the user can type their question and the chatbot responds. The NLP phase is part of the chatbot's backend, which uses NLP techniques, NLU to understand and parse the question, and NLG to generate a reply. The backend can also connect to, for example, the company's knowledge base and other external sources that help create the answer. (Sood et al., 2024)

Machine translation

Machine translation is one of the earliest applications of NLP, having been identified as early as the 1950s (Lee, 2024). Machine translation refers to the process of transforming spoken or written language from one language to another. Although translating individual words is simple, translating complex texts with multiple sentences requires NLP. (Ahammad et al., 2024; Juglan et al., 2023) With the help of NLP, these difficult phrases containing idiomatic expressions, rare words, or words with multiple meanings can be translated while preserving the original text's tone and meaning (Mishra et al., 2024).

In customer service, machine translations create many opportunities for companies. Companies can serve their customers in multiple languages across different channels, even if the customer service representative and the customer do not share a common language. Machine translations can also be integrated with other NLP applications, such as chatbots, text generation, sentiment analysis, and text summarisation. (Ahammad et al., 2024; Olujimi & Ade-Ibijola, 2023) The need for intercultural communication will only grow as globalisation continues, which means that the need for accurate and efficient translations will also grow in customer service (Mohamed et al., 2024).

Sentiment analysis

Sentiment analysis (SA) is a widely used NLP application in customer service (Patwardhan et al., 2023). It helps detect emotions and assess customer moods in real time, enabling customer service agents to adapt their behaviour accordingly. SA generally considers three levels: neutral, positive, and negative. More levels can be included if greater detail is required. (Amaratunga, 2023; Gowri et al., 2025) SA can be used for many different purposes and data sources. Companies can use SA, for example, for different channels of an omnichannel contact centre, such as emails, phone calls, and support requests, as well as their own social media messages or comments. (Gowri et al., 2025; Juglan et al., 2023) Figure 11 shows an example of SA with three categories.

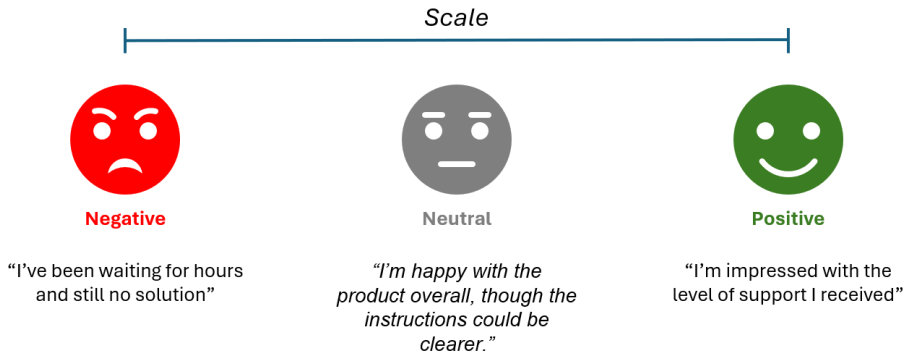

Figure 11 Example of sentiment analysis (adapted from Gowri et al., 2025)

SA can be divided into three sub-levels based on the depth of text analysis: document level, sentence level, and aspect level. Document-level SA examines the overall emotional tone and sentiment of the entire document. For instance, a customer review of a service would be analysed to understand the general opinion it conveys. Sentence-level SA involves segmenting a document or paragraph into individual sentences and assessing each sentence's polarity. This enables a more accurate understanding of variations in opinion within the text, as different sentences can express different opinions or emotions. Aspect-level SA is the most accurate approach, focusing on identifying opinions about specific aspects, characteristics, or entities mentioned in the text. Instead of just determining the general opinion of a review, for example, in a review of a phone, the opinion expressed about "battery life" is analysed. (Ali et al., 2025; Frigau et al., 2024; Liu et al., 2025)

In customer service, SA can be utilised in various ways, such as analysing customer feedback, managing branding, examining customer interactions, identifying issues in customer support, and detecting emotions during real-time customer interactions (Gowri et al., 2025; Juglan et al., 2023). SA can be used to identify potential problems in processes, products, or services. For example, companies can detect problems in their customer service through dissatisfaction and improve their customer service processes based on feedback (Ali et al., 2025; Taherdoost & Madanchian, 2023). Statistics demonstrating positive feedback on customer service can be utilised in marketing campaigns, showcased on the company's website or social media to enhance the brand. (Juglan et al., 2023).

SA can be used as a standalone application, but when combined with chatbots and virtual assistants, it can deliver more personalised and effective customer service. SA's role here is to recognise emotional states from customer inquiries and feedback. This enables chatbots and virtual assistants to consider the customer's emotional condition when formulating a

response. Taking the customer's emotional state into account positively influences the overall customer experience. (Taherdoost & Madanchian, 2023)

Named entity recognition

Named entity recognition (NER) is an NLP application that recognises and distinguishes named entities from data and groups them into predefined categories, such as places, organisations, order numbers, times, or customer names (Khurana et al., 2023; Mishra et al., 2024). By converting unstructured text into structured data, NER enables the extraction of meaningful information, which is vital for NLP applications in information retrieval and question answering (Mariprasath et al., 2025). The quality of NER results can be affected by factors such as the ambiguity of entity names, mentions of entities in different forms, and the context in which the entity is mentioned (Patwardhan et al., 2023).

In customer service, NER can be used to extract information from customer inquiries, such as order or customer IDs, products, and delivery dates, which can help provide more personalised service and positively influence the CX. For example, by extracting the product name from a customer's query, the customer can be better informed about the product's price, features, and store availability. In addition, order numbers can be obtained from the customer and later used to review order details, thereby enabling more efficient service. In some cases, NER can also help identify the customer's intent by recognising keywords such as "cancellation" or "refund" in customer queries. (Olujimi & Ade-Ibijola, 2023) Figure 12 shows an example of NER.

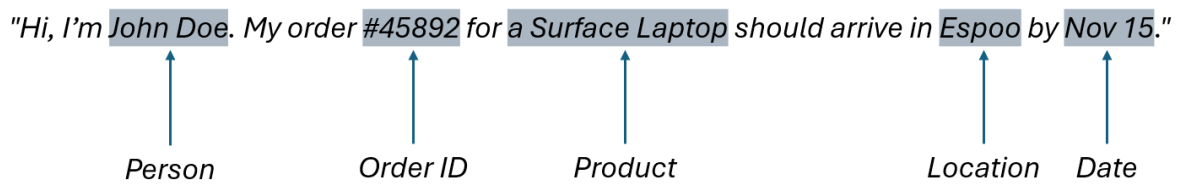

Figure 12 Example of NER

Text classification

Text classification is an NLP application in which text is categorised into one or more predefined groups. The goal of text classification is to automatically assign a label to a given

text based on its content, so that the text can be organised and categorised for easier analysis and management. (Murrar et al., 2025; Patwardhan et al., 2023) Text classification models are usually trained on labelled text data, so they learn to associate content with tags. The model can then be used to classify new text inputs. (Kadre et al., 2025) Typically, text classification is divided into two subcategories: document classification and cause-and-effect classification. In document classification, the identifier or category must be defined for the entire document, such as a service request or request for quotation. Cause-and-effect classification refers to identifying the cause-and-effect relationship between two events described in a sentence or paragraph. (Patwardhan et al., 2023)

In customer service, text classification can be used in many tasks, such as filtering spam emails, classifying service requests, and placing customers in customer service queues (Mishra et al., 2024). Text classification can therefore significantly improve the efficiency of customer service. For example, requests that were previously processed manually can now be handled by automatically classifying the request to the right customer service representative, which increases operational efficiency and customer satisfaction. (Murrar et al., 2025)

Text summarisation

Text summarisation produces a summary of the original text that is quick to read but contains all the important details. Text summarisation uses NLP to identify the most important points in a text and convert them into a clear, concise form that allows the reader to understand the entire text without reading the original. (Juglan et al., 2023; Patwardhan et al., 2023) Text summarisation typically takes place in three different stages: pre-processing, processing, and post-processing. In the pre-processing stage, the original text is cleaned up and prepared for summarisation. This involves several linguistic procedures, such as sentence segmentation and punctuation removal. The actual summarisation method is applied during the processing stage. At this stage, essential features such as keywords and meaningful phrases are extracted from the text. Each sentence is scored based on its importance, and the best sentences are selected for the summary. In the post-processing phase, the selected sentences are arranged in a logical and understandable order. The goal is to ensure that the final summary is clear, consistent, and easy to read. (Kadre et al., 2025; Mishra et al., 2024) Figure 13 shows an example of document summarisation.

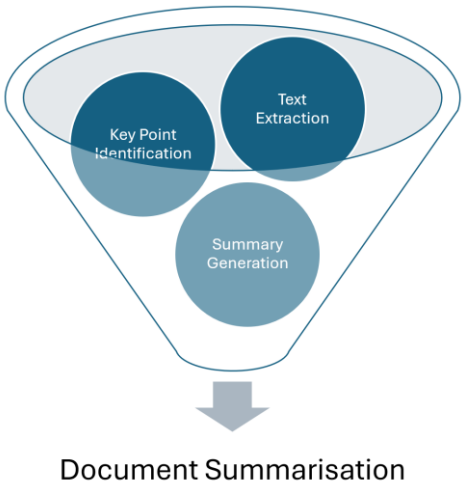

Figure 13 Example of document summarisation (adapted from Kadre et al., 2025)

In customer service, text summarisation can be used, for example, in call centres, where even long calls can be summarised concisely. Text summarisation can also reduce the time spent processing customer feedback, as customer service representatives do not need to read lengthy messages. By combining this with text categorisation and sentiment analysis, companies can focus on the most relevant feedback, thereby improving their customer service. By utilising text summarisation, companies can reduce manual work, streamline their operations, and improve CX. (Kadre et al., 2025; Mishra et al., 2024)

Voice recognition

Nowadays, there is an emphasis on using modern contact centers, where all customer communication, emails, phone calls, live chat, and online form submissions take place on the same platform (Shah et al., 2023). Voice recognition uses NLP to enable real-time understanding of spoken language and convert it to text, facilitating structuring and searching. By converting audio data, which is difficult to analyze, into text format, companies can utilize their call data much more effectively in their reporting. (Alqurni et al., 2025; Juglan et al., 2023)

Voice recognition can improve the efficiency and effectiveness of customer service in many ways. For example, it can be integrated into a chatbot or virtual assistant, enabling these systems to communicate with customers via voice technology. This means customer service phone lines can be open 24/7 without requiring human presence. Another example is the use of voice recognition to automatically transcribe calls for quality assurance or training, thereby improving the quality of customer service staff's work and, in turn, customer

satisfaction. (Juglan et al., 2023) The various applications presented in the literature have been combined and summarised in Table 2 below.

Table 2 NLP applications

NLP method	Category	Technique	Use case scenarios	References
Chatbot and virtual assistants	NLU & NLG	NLU & NLG	Provide self-service and answer general questions across various channels.	Benaddi et al., 2024, Horvat et al., 2025, Mariprasath et al., 2025, Olujimi and Ade-Ibijola, 2023, Sood et al., 2024
Machine translation	Machine translation	NLU & NLG	Translates customer messages in a multilingual service.	Ahammad et al., 2024, Juglan et al., 2023, Lee, 2024, Mohamed et al., 2024, Olujimi and Ade-Ibijola, 2023
Sentiment analysis	Text analytics	NLU	Recognises the customer's emotional state and helps prioritise or direct cases.	Ali et al., 2025, Frigau et al., 2024, Liu et al., 2025, Taherdoost and Madanchian, 2023
Named entity recognition	Text analytics	NLU	Extracts relevant information (such as order or product numbers) from customer messages.	Khurana et al., 2023, Mariprasath et al., 2025, Mishra et al., 2024
Text classification	Text analytics	NLU	Classifies customer requests and directs them to the appropriate service channel.	Kadre et al., 2025, Murrar et al., 2025, Patwardhan et al., 2023
Text summarisation	Text analytics	NLU & NLG	Summarises long customer messages or transcripts into a short summary.	Khurana et al., 2023, Mariprasath et al., 2025, Mishra et al., 2024
Voice recognition	NLU & NLG	NLU & NLG	Automatic transcriptions from calls or combined with other NLP methods.	Alqurni et al., 2025, Juglan et al., 2023, Shah et al., 2023

3.4 Artificial intelligence agent architecture

AI-based systems have found applications across various environments, including call centers and factories. However, recent advances in AI development have led to a change in the conceptualisation of AI, transforming it from a simple automation tool into an intelligent artefact that can facilitate collaboration. These technologies can take many forms, such as recommendation systems, chatbots, and digital assistants, and have been adopted across a wide range of business environments. (Dennis et al., 2023)

These technologies are referred to as AI agents. An AI agent is an AI system that can pursue goals with limited human involvement by planning and executing complex tasks, such as decision-making, action-taking, and adapting to changing circumstances, rather than functioning solely as a static predictor or a passive assistant. (McLaughlin, 2025)

In practice, this agent capability is typically achieved through a set of architectural elements, in particular planning mechanisms, memory, and tool usage, which enable the system to act purposefully and influence its environment through its actions. Accordingly, an AI agent can be described as an intelligent system capable of independent understanding and perception, goal-oriented planning, memory maintenance and use, and the use of tools to perform tasks. In other words, it learns and performs tasks through repeated observation and decision-making processes. (Li et al., 2025) An example of an AI agent is shown in Figure 14.

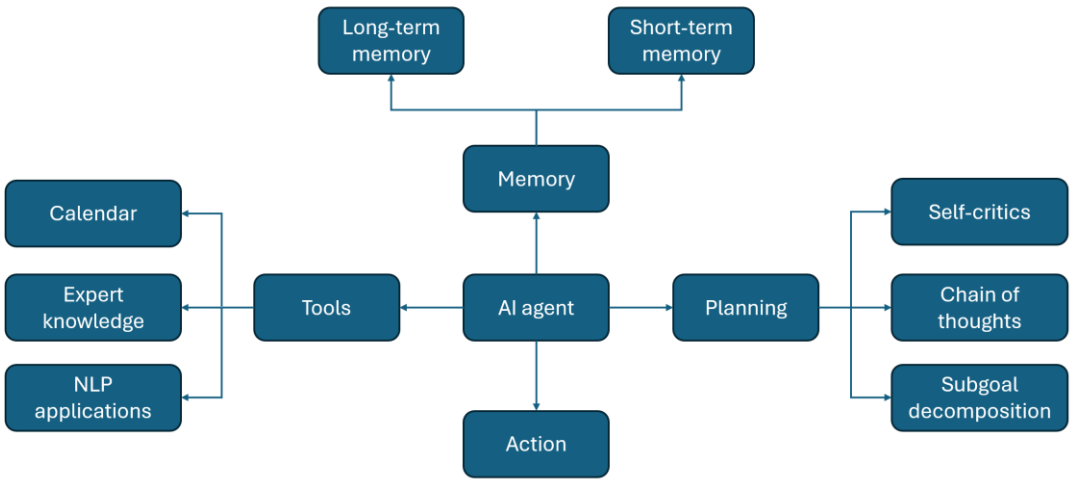

Figure 14 Example of an individual AI agent architecture (adapted from Li et al., 2025)

The memory module supports continuity by combining short-term working context, which is capacity-limited and fades as new information accumulates, with long-term memory that persists across interactions and enables retrieval of stable knowledge when needed. The planning module represents the deliberative layer typically centred on a large language model, where the agent translates the current goal and context into an ordered course of action and may employ structured prompting to improve multi-step reasoning and reduce errors. The action and tools module operationalises the plan by producing user-facing language outputs and by invoking external tools that implement domain rules and access

authoritative resources; the results of these tool calls feed back into memory, allowing the agent to update its internal state and refine subsequent plans. (Biswas & Talukdar, 2025; Li et al., 2025; Ranjan et al., 2025) Together, the components shown in the figure explain how agentic behaviour arises from an iterative cycle of stateful reasoning and environment interaction rather than from single-pass text generation (Abou Ali et al., 2025)

Typically, AI agents are modular, standalone entities that utilise AI models to perform precisely defined tasks autonomously. With NLU, tool integration, prompt chaining, and limited memory, these agents operate in clearly defined contexts, such as email filtering or content summarisation. (Dwivedi et al., 2025) Although these architectures enable autonomy and task-specific performance, traditional agents remain limited by their restricted memory, poor collaborative abilities, and linear sense-plan-act implementation model. As a result, they excel at well-defined, deterministic tasks but struggle in complex, high-complexity environments or in spontaneous behaviour. (Sapkota et al., 2026)

Individual agents are therefore unable to handle more complex tasks, such as handling customer service calls or evaluating customer service quality, on their own (Dwivedi et al., 2025; Huang, 2025). For this reason, recent developments in AI have shifted toward agentic AI, an approach that extends beyond the capabilities of a single agent by enabling multiple agents to work together toward a shared goal. Instead of relying on a single autonomous unit, agentic AI structures intelligence as a coordinated system in which multiple agents contribute their expertise, share information, and collectively manage tasks that would be too demanding for any single agent. This shift lays the foundation for systems that can address broader, multi-stage workflows and operate with greater adaptability and context awareness than traditional AI agents. (Acharya et al., 2025; Ranjan et al., 2025; Sapkota et al., 2026) Figure 15 demonstrates the difference between agentic AI and AI agents.

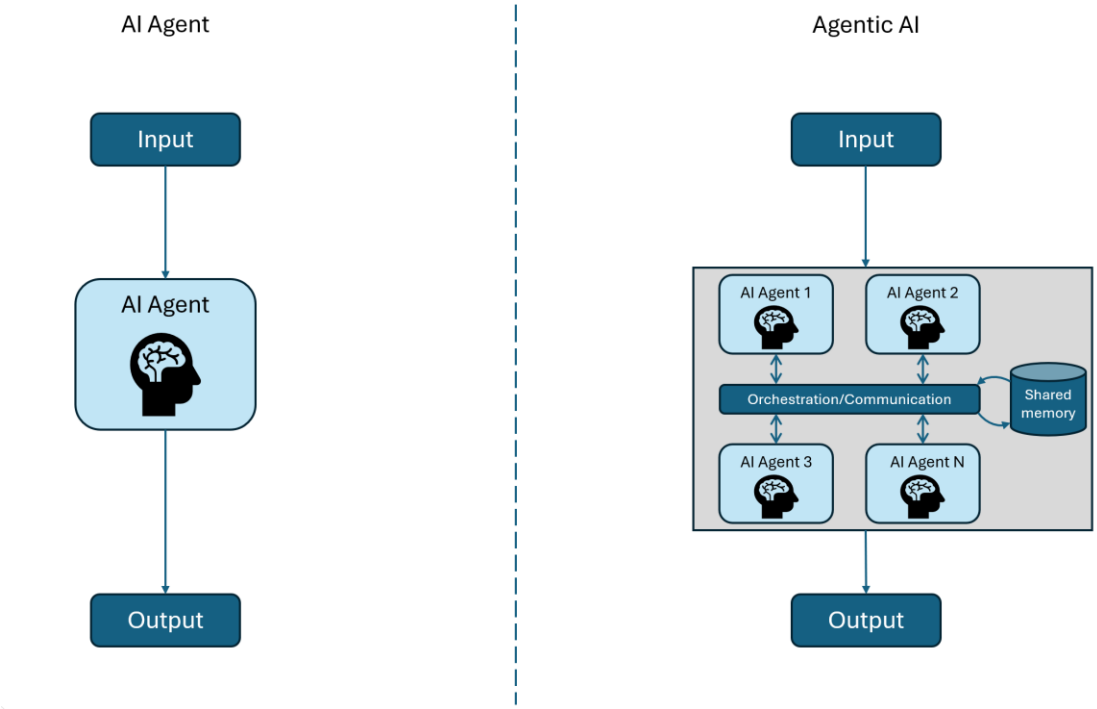

Figure 15 Difference between AI agent and agentic AI

Architectural approaches to agentic AI are typically based on modular and hierarchical structures that enable the system to manage complex objectives and adapt to dynamic environments. Common architecture includes multi-agent systems (MAS), hierarchical reinforcement learning (HRL), and goal-oriented modular architectures. (Acharya et al., 2025)

A MAS is a distributed system in which multiple autonomous agents interact in a shared environment to achieve individual or collective goals. The core of MAS systems is based on the autonomy, reactivity, proactivity, and social interaction capabilities of agents, which enable information exchange, coordination, and cooperation. The capability of these systems does not arise from the complexity of individual agents, but from the interaction between agents and emergent collective behaviour, which makes MAS suitable for complex, dynamic, and wide-ranging problems that cannot be solved by a single agent. (Acharya et al., 2025; Biswas & Talukdar, 2025; Sapkota et al., 2026)

HRL is a hierarchical decision-making model in which tasks are divided into several levels of abstraction. A high-level agent or controller defines sub-goals, while lower-level agents implement them in practice. This structure enables complex tasks by breaking a large whole into simpler, more manageable parts. The hierarchical task model improves the system's

scalability and clarifies task implementation, as sub-goals can be learned, executed, and reused in different situations. For this reason, HRL-based solutions are particularly well-suited to environments where decision-making requires layered organisation, long-term goal management, and efficient task distribution among agents. (Acharya et al., 2025; Huang, 2025; Pateria et al., 2022)

A goal-oriented modular architecture divides an agent's operations into separate modules, enabling it to flexibly combine the necessary functionality for different tasks. Modularity improves scalability, maintainability, and adaptability to complex, changing goals. (Abou Ali et al., 2025; Acharya et al., 2025; Sapkota et al., 2026)

Artificial intelligence agents in customer service

Customer service is entering a phase in which AI agents complement and partially replace traditional service models. This development has accelerated, particularly with the advent of NLP and agent-based AI, which enable natural interaction and real-time processing of large volumes of data. (Brohi et al., 2025) Practical implementations indicate that automated, voice-based service systems can reduce customer complaints and improve the service experience, particularly for routine service requests. At the same time, it has been observed that the option to switch to a human service representative does not increase customers' long-term need for human service, indicating that AI solutions function as a natural part of service environments. (Wang et al., 2023)

AI agents are used both in external customer interfaces and in internal service tasks, and they can operate independently, for example, by handling routine questions, retrieving information from different systems, and producing solutions without constant supervision (Ranjan et al., 2025). However, the effectiveness of these systems depends heavily on the type of service tasks (Wang et al., 2023). Simple, repetitive tasks are well-suited to AI, while tasks that require interpretation or are multi-layered are still handled by human experts (Acharya et al., 2025).

AI agents are particularly attractive in high-volume environments where service demand is constant, and the work is demanding (Brohi et al., 2025). Automated systems can respond to contacts immediately and consistently, reducing staff workload and freeing up resources for more complex customer cases (Huang et al., 2024). This speed and immediate availability have been found to influence customer behaviour, as many users find automated

services attractive for offering quick help, links to further information, and clear, structured solutions (Li & Zhang, 2023).

Overall, this development has led to a new service model in which AI agents are central to the operational foundation of customer service. Agents already operate independently on routine and information-intensive tasks, and their role is expanding toward greater autonomy as technology advances. At the same time, human employees are focusing on situations that require judgment, broader expertise, and case-by-case interpretation. The current situation can be seen as a transitional phase toward hybrid models, in which collaboration between AI and humans is becoming the new norm in customer service. (Brohi et al., 2025; Jang et al., 2025; Pathak & Bansal, 2025)

Artificial intelligence agent challenges in customer service

Although AI agents and agentic AI offer significant opportunities to improve customer service efficiency and scalability, their implementation involves several technical, ethical, and operational challenges that limit large-scale adoption of autonomy (Ranjan et al., 2025). One of the key challenges relates to system reliability and data security. Customer service often operates in an information-sensitive environment where agents handle personal data, contracts, and customer history. In such tasks, an agent must operate accurately and transparently while complying with strict data protection and control requirements. An agent's ability to use external tools, call interfaces, and take actions independently increases the risk of incorrect decisions, misuse, and data leaks. Managing these risks requires clear authorisation models, operational auditing, context monitoring, and restricted access rights, which are technically challenging to implement when decision-making is not fully predictable. (Acharya et al., 2025)

Another significant challenge relates to the transparency and explainability of decision-making. Much of the reasoning in agentic systems relies on assessments generated by large language models, whose internal mechanisms cannot be accurately interpreted. This creates problems, especially when the agent performs multi-step tasks or acts independently on the customer's behalf. (Ranjan et al., 2025) When the reasoning behind decisions cannot be verified retrospectively, the division of responsibility between humans and systems also becomes blurred. The lack of explainability also weakens the organisation's ability to evaluate the agent's performance, identify biases, and ensure that service processes meet

regulatory and quality requirements. This challenge is particularly acute in environments where customer-facing solutions require consistent and justifiable decision-making paths. (Acharya et al., 2025)

The ability of agents to maintain and utilise context is the third key challenge. Customer service is often long-term and multi-channel, meaning relevant information can be scattered across multiple systems and stages of interaction. Agents must be able to recognise the customer's situation history, the goal of the interaction, and the associated constraints, but current models are not always capable of managing large and changing contexts. (Ranjan et al., 2025) This can lead to inconsistent responses, increased repetition, or false assumptions that undermine CX and customer trust.

There are also challenges related to agents' social and linguistic capabilities. Although modern language models can produce fluent, natural language, they cannot engage in genuine emotional interaction. This limitation is particularly evident when customers expect empathetic service that recognises the seriousness of the situation. The "superficial empathy" the agent displays may seem insincere to the customer, undermining interaction and potentially increasing the need for human escalation. At the same time, customers already have negative experiences with previous AI agents, which increases scepticism towards new-generation agent systems. (Li & Zhang, 2023)

In addition, the use of autonomous agents requires significant changes in processes, division of labour, and control structures within organisations. Agents are not separate add-on components. Rather, they affect how customer service work is planned, how responsibilities are divided, and what control and escalation models are needed to support the service process. Without organisational maturity and clear principles regarding the authority, training, and quality control of agents, autonomy can lead to inconsistent operating models and random errors. (Ranjan et al., 2025)

In summary, the biggest challenges for AI agents in customer service include ensuring reliability, transparency, security, managing context, and addressing the lack of emotional interaction. Solving these challenges requires both technical innovations and organisations' ability to define clear operating models that ensure the responsible and transparent use of agent systems. (Acharya et al., 2025; Ranjan et al., 2025)

4 Design science research: empirical part

The aim of this thesis is to examine, in practice, how autonomously AI agents can operate in customer service situations using current technology. As noted in the theoretical section, AI has advanced rapidly in recent years, thanks to NLP and agent-based AI, which have opened new opportunities to improve customer service processes. In the empirical part, an agent-based customer service module is built into the CRM to identify the limits of current AI capabilities. The empirical section consists of research methodology and process, problem identification and goal setting, solution architecture design and development, and solution presentation and evaluation.

4.1 Design science research method

This study is based on the DSR approach, which is particularly well-suited to situations in which the aim of the research is to develop and evaluate concrete solutions to practical problems. DSR has been characterised as a research paradigm in which knowledge is produced through the design, implementation, and evaluation of artefacts. (Hevner & Chatterjee, 2010) Artefacts can be technical, organisational, or social in nature, and their value is primarily based on their ability to respond to challenges observed in a real operating environment. Although several process descriptions of the DSR method have been presented, most are based on the DSR model proposed by Peffers et al. (2007). (vom Brocke et al., 2020) Later models and improvements have simplified or grouped the model, while retaining its core elements: problem definition, solution development, and systematic evaluation. (Henriques & O'Neill, 2023)

Peffers et al. (2007) define the DSR process as consisting of six interrelated stages: problem identification and motivation, solution goal definition, artefact design and development, demonstration, evaluation, and communication of results. In this study, these stages are not viewed as a linear process, but as an iterative whole in which empirical data and artefact development guide one another. The DSR process is presented in Figure 16.

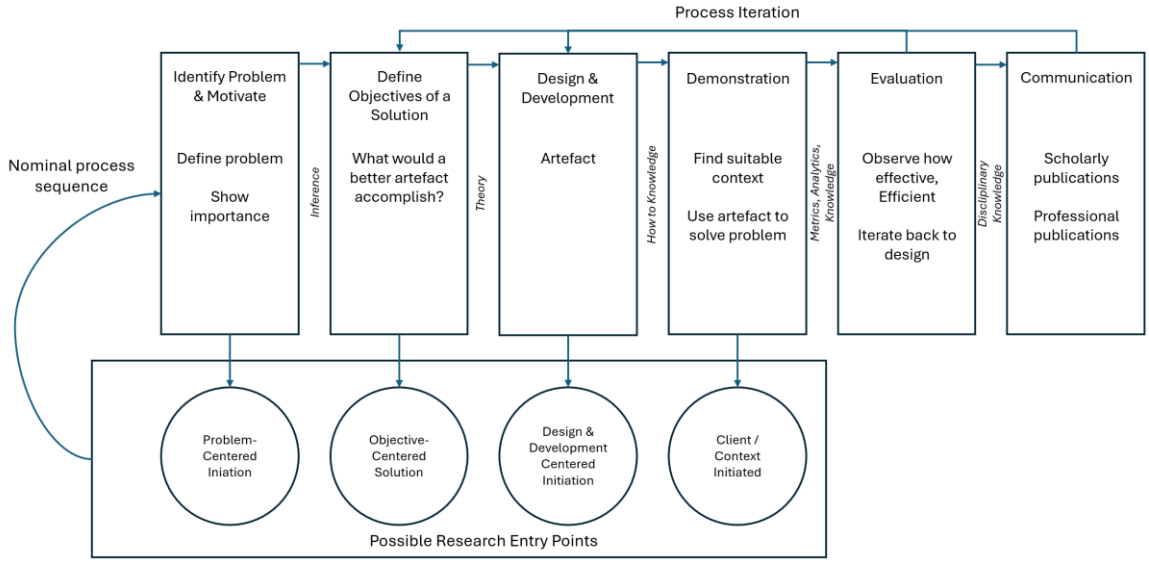

Figure 16 The DSR methodology process model (adapted from Peffers et al., 2007)

In the early stages of the process, the focus is on understanding the problem's context. Identifying the problem requires understanding the current situation and its limitations, and justification for why the identified challenge is significant for research and practical application. At this stage, the research's relevance is established with respect to both the organisation and the existing research literature.

Once the problem and its underlying objectives have been defined, the research proceeds to the design and development of a solution. This stage represents the core logic of DSR: building a solution grounded in empirical observations and existing knowledge (Hevner & Chatterjee, 2010).

The developed artefact is demonstrated in a real or simulated use situation, after which its functionality and usefulness are evaluated. The final stage of the process focuses on communicating the research results to stakeholders, including the scientific community and practitioners.

Empirical data were collected using qualitative methods that support the DSR's goal of building context-specific understanding. The study utilised both unstructured and semi-structured interviews, which complement each other in different stages of the research.

Unstructured interviews are particularly suitable for situations where the research topic or problem definition is not yet fully clear. This method allows for open discussion and gives the interviewee the opportunity to raise themes that the researcher might not have anticipated

(Qu & Dumay, 2011). In this study, unstructured interviews were used as a starting point for identifying the research problem from the organisation's perspective.

Semi-structured interviews were conducted later, after the key themes had emerged. The advantage of this method is the balance between a structured framework and flexibility. The researcher has a clear understanding of the topics to be discussed, while the interview situation allows for in-depth questions and clarification of perspectives. (Qu & Dumay, 2011) These interviews were used to gather detailed information about the current situation, challenges, and requirements for a solution.

DSR is part of a broader tradition of applied research, in which the research process is seen as iterative and step-by-step. However, a distinctive feature of DSR is its clear, solution-oriented approach and pragmatic validity. The value of the research is measured by how well the developed artefact responds to the identified problem in a real operating environment. (Hevner & Chatterjee, 2010; Peffers et al., 2007; vom Brocke et al., 2020) By combining a systematic process model, qualitative empirical data, and artefact development, this study aims to deliver both a practical solution and a scientifically sound research contribution. This is particularly relevant in the present thesis, because the research problem is not limited to describing AI agent autonomy as a phenomenon, but concerns how autonomy can be operationalised in customer service in a controlled and meaningful manner. In this context, the study must address not only technical feasibility, but also service quality, escalation logic, and organisational accountability, which makes DSR an appropriate methodological frame for the iterative design and evaluation of the artefact.

4.2 Design science research process

The DSR process described in the thesis aligns with the methodology discussed in the previous chapter. The process consists of ten stages, including stages related to DSR and data collection. During the process, a few rounds of iteration were conducted across design, development, demonstration, and evaluation. Each stage and the thinking behind the process are discussed in more detail in the following sections. The data sources and research stages are presented in Figure 17 below.

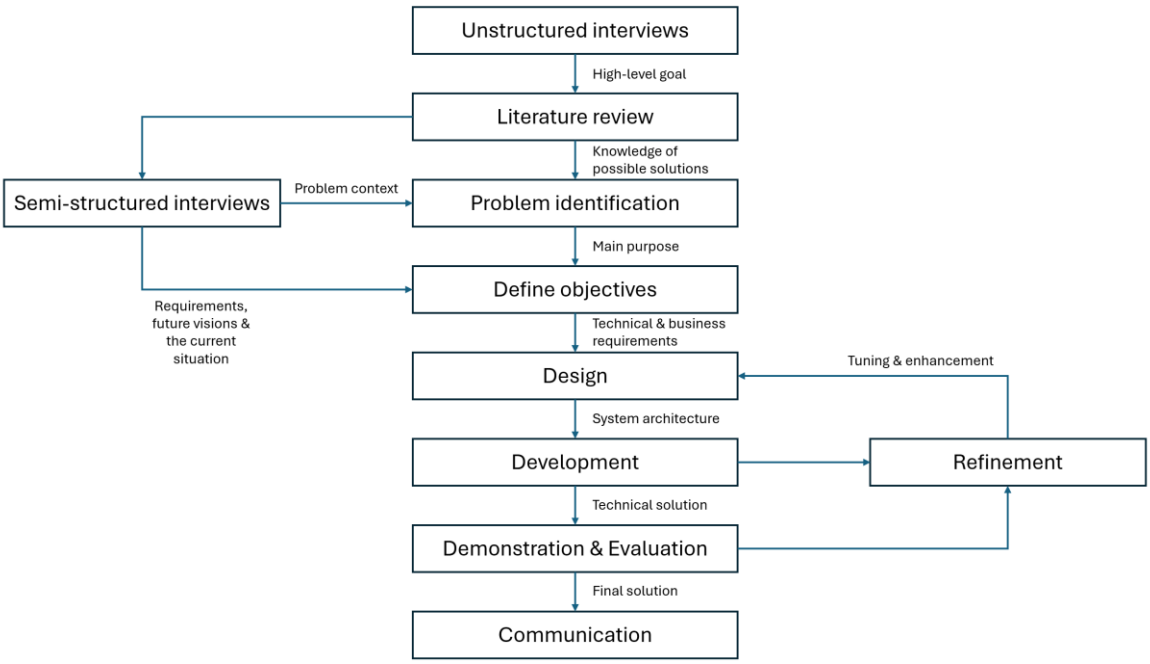

Figure 17 DSR process implemented

The DSR process for this study was carried out as part of an internal development project. The thesis was not a client-specific assignment, but was integrated into an ongoing internal project led by a dedicated project team. The study's topic and objectives were discussed and agreed upon as part of the project scope, ensuring alignment between the thesis's academic objectives and the organisation's development activities.

The process began with preliminary discussions within the project team, conducted as unstructured interviews. The purpose of these was to narrow the thesis topic and define the research's general objectives. After the unstructured interviews, a literature review was conducted to gain an understanding of the research field and its theoretical foundations. The insights from the literature were used to inform further discussions and interviews to clarify the problem context and identify the requirements for developing a solution.

In addition to internal stakeholders, customer organisations were included in the interview process to gain a broader understanding of the current state of customer service practices and the use of AI in real organisational contexts. The interviews provided information on customers' current challenges, the maturity of AI in customer service, and expectations for AI's future role. The views of customer organisations were used to supplement the internal view of the problem and to define business and technical requirements.

Based on insights from internal project discussions, customer interviews, and literature reviews, the business and technical requirements were formulated during the goal definition phase of the DSR process. These requirements were translated into a system architecture that guided the design and development of the artefact. The artefact was implemented, presented, and evaluated in accordance with the DSR method.

During the process, several rounds of iteration were performed to address technical challenges encountered during development and to incorporate feedback from project stakeholders. Although several versions of the artefact were developed, these iterations are not described in detail in this thesis, as most of the improvements were made at the source code and script level, which are not part of the research artefacts. Ultimately, a solution was achieved that adequately addressed the identified problem, and the results were communicated to stakeholders through this thesis and internal business presentations.

4.3 Problem identification and objective definition

The topic of this thesis has its roots in an internal need to create new concept templates using current technology. Within the company, there was a desire to investigate how well AI currently meets customer service needs. In internal discussions, it was agreed that customer service is increasingly moving toward agent-based environments in which AI supported service agents handle inquiries, feedback, and customer contacts. Accordingly, a new customer service module was defined as the practical focus of the study in order to examine how current AI capabilities can support autonomous customer service in a single environment. This led to the study's original problem statement: How autonomous can customer service be made with current technology in the Dynamics 365 CRM environment?

First, the study aimed to define the scope of the internal project and its key components. Although the system could support multiple channels and implementations, the scope of the project was limited to agent autonomy in email, chat, tickets, and phone calls.

Based on the literature review, the problem area was refined around the conditions under which autonomous AI agents can be deployed in customer service without compromising service quality, transparency, or organisational accountability. These insights informed the refinement of the semi-structured interview questions and guided the subsequent interview

discussions. Semi-structured interviews supported a more in-depth formulation of the problem and the definition of the solution objectives. A visual representation of the problem and objective definition process for this study is shown in Figure 18.

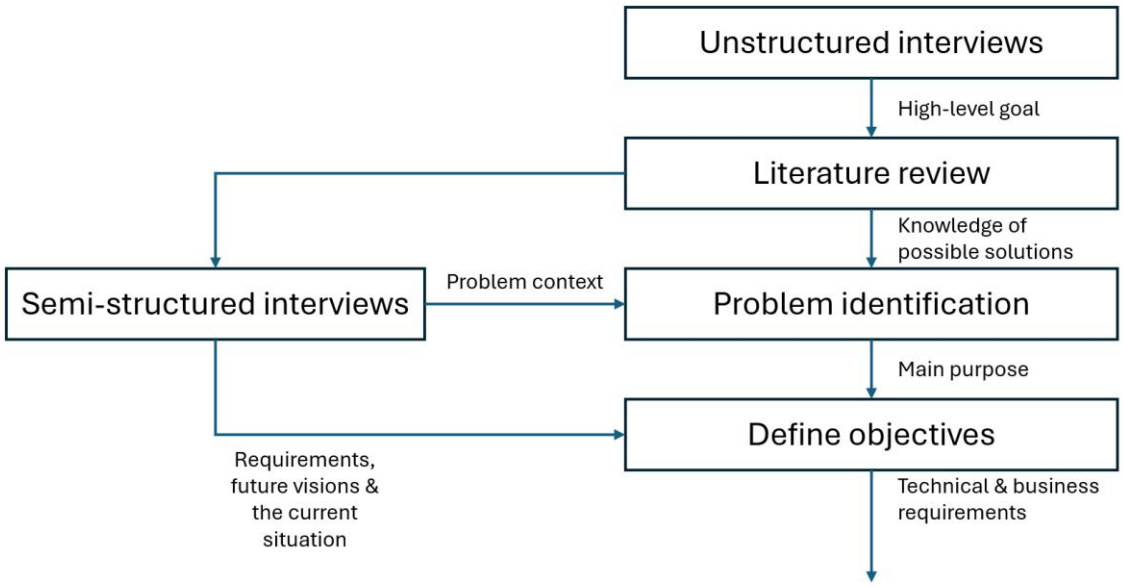

Figure 18 The problem and objective definition process

During the literature review, an overview of existing research on customer service, customer experience, and the application of AI in service contexts was developed. Particular emphasis was placed on recent developments in AI-enabled customer service, agent-based systems, and the autonomy of intelligent systems in customer interactions. The literature review supported the understanding of how customer service quality and customer experience are currently addressed, as well as how emerging AI technologies are expected to influence future customer service models.

The insights from the literature review directly contributed to refining the research problem and defining the study's objectives. In addition, the review informed the formulation of the semi-structured interview questions by highlighting key themes related to customer service processes, AI maturity, autonomy, and the relationship between customer service and customer experience. This ensured that the interviews focused on aspects relevant to the design and evaluation of an autonomous, AI-driven customer service solution within the Dynamics 365 CRM environment.

Following the literature review, semi-structured interviews were conducted to gain a deeper understanding of the problem context and to support the formulation of requirements for the solution artefact. The primary objective of the interviews was to examine the current state of customer service practices, customer experience management, and the use of AI across different business sectors, and to explore expectations regarding the future roles and autonomy of AI-driven customer service agents.

The interviews were conducted with experienced professionals representing multiple business sectors, including services, insurance, banking, and information technology. This cross-sectoral approach enabled the study to capture a broad range of perspectives on customer service operations, customer experience practices, and the adoption of AI technologies across different organisational contexts. An overview of the interviewed sectors, participants' experience levels, and interview dates is presented in Table 3.

The diversity of business sectors provided both strategic and operational insights into customer service and customer experience. Topics discussed during the interviews included the current maturity of customer service processes, existing customer service tools and measurement practices, the relationship between customer service and customer experience, the current and potential use of AI in customer interactions, and future visions for more autonomous, agent-based customer service environments.

To ensure the anonymity of the participating organisations and individuals, raw interview data, such as transcripts and detailed interview notes, are not included in this thesis. However, a high-level interview guide summarising the main themes, questions, and objectives of the semi-structured interviews is provided in Appendix 1. The same interview guide was used for all interviewees to support a coherent and comparable understanding of the topics discussed.

The insights gained from the interviews played a central role in refining the research problem and defining both the business and technical requirements for the artefact developed in this study. In particular, the interviews helped assess the current level of feasible autonomy in customer service and identify key constraints and opportunities for implementing AI-driven customer service agents within a Dynamics 365 CRM environment.

Table 3 Semi-structured interview background

Role of the interviewee	Sector	Experience	Duration	Date
Director of ICT & Compliance	Service	25 years	60 minutes	28.11.2025
Insurance Manager	Insurance	15 years		9.12.2025
Business Development Manager	Service	30 years		12.12.2025
Head of Operational Excellence	Service	20 years		15.12.2025
VP – Customer Support	IT	20 years		19.12.2025
Customer Service Manager	Banking	30 years		30.12.2025

Interview findings

The semi-structured interviews provided insight into the current state of customer service practices, the role of AI in existing processes, and expectations regarding future autonomy. Across all interviewed sectors, customer service was described as a critical customer-facing function with a direct influence on customer satisfaction and perceived service quality.

Interviewees highlighted response time, issue resolution capability, service availability across channels, and the expertise of customer service personnel as the most important drivers of service quality. At the same time, omnichannel service delivery was described as operationally complex. Fragmentation between communication channels and supporting systems was repeatedly identified as a significant challenge, as it limits end-to-end visibility into customer interactions and creates channel-specific data silos.

Customer service performance was primarily monitored using operational metrics such as average handling time, resolution time, and first-contact resolution. While these metrics support day-to-day management, interviewees emphasised that they provide only a partial view of service quality. In particular, unstructured interaction data generated during customer contacts, such as conversation transcripts, messages, and call content, was described as largely underutilised beyond basic reporting.

Regarding AI adoption, interviewees identified consistently near-term value primarily in high-volume and repetitive tasks, including initial contact handling, information retrieval, case classification, and interaction summarisation. AI was generally viewed as an efficiency-enhancing capability rather than a replacement for human customer service

representatives. Limitations related to contextual understanding, reliability, and the handling of complex or emotionally sensitive situations were emphasised.

Future visions described AI autonomy as incremental rather than immediate. Increased autonomy was associated with potential risks, including concerns about transparency, loss of control over customer-facing actions, and data quality. Consequently, interviewees consistently stressed the importance of monitoring mechanisms, validation practices, and human oversight as prerequisites for trustworthy AI supported customer service.

Interview analysis

Taken together, the interview findings reveal a structural mismatch between the growing availability of omnichannel customer service interaction data and organisations' limited ability to utilise this data through coordinated and autonomous AI systems. Although customer service interactions generate rich, unstructured information across text- and voice-based channels, existing AI applications remain fragmented, channel-specific, or limited to isolated tasks. As a result, organisations struggle to increase the level of autonomy in customer service without compromising service quality, transparency, or customer trust.

When combined with the literature review, these findings highlight a research gap regarding the architectural and operational conditions required for controlled, incremental autonomy in customer service. Prior research has largely focused on single-channel solutions, chatbot performance, or individual AI capabilities, while paying limited attention to how multiple AI components and agents can be composed into a coherent, interoperable system embedded within an existing CRM environment.

From a DSR perspective, this gap motivates a design-oriented contribution. Rather than proposing a fully autonomous customer service agent, this study focuses on how agent-based AI capabilities can be structured, combined, and reused to support increasing levels of autonomy while maintaining human involvement and oversight. Autonomy is therefore treated as a gradual property that depends on both technological maturity and organisational readiness.

Accordingly, the objective of this study is to design and describe an AI-centric, agent-based customer service solution that enables controlled and incremental autonomy within an omnichannel Microsoft Dynamics 365 CRM environment. The solution aims to design and

demonstrate how modular AI components and agents can interpret and enrich unstructured interaction data across channels in controlled customer service scenarios.

Based on the interview findings and the analysis presented above, a set of business and technical requirements was derived to guide the design of the DSR artefact. The requirements reflect the key constraints and opportunities identified in the empirical material, particularly those related to omnichannel complexity, the underutilisation of unstructured interaction data, and the need for controlled and incremental autonomy. Table 4 summarises these requirements and establishes an explicit link between the identified problem and the solution artefact designed in the following sections.

Table 4 Requirement for the DSR artefact

Requirement type	Requirement description
Business requirement	There is a need for AI-based capabilities that support customer service work by interpreting and enriching customer interactions.
	There is a need to enable a gradual transition to more autonomous customer service agents while maintaining human involvement and oversight.
	There is a need for AI components that can be reused and combined to support different customer service tasks and levels of autonomy.
	There is a need for AI solutions that can operate within existing customer service environments rather than as isolated systems.
Technical requirement	The AI solution should consist of modular AI components that support interoperability and coordinated operation.
	The AI components should be capable of analysing customer service interactions using intent recognition, sentiment analysis, interaction summarisation, and categorisation.
	The AI solution should support both text-based and voice-based customer service interaction data.
	The AI components should operate on standardised representations of interaction data to enable channel-independent analysis.
	The AI solution should support multilingual customer service interactions, with Finnish and English identified as primary languages.

The problem analysis, objectives, and derived requirements defined in this section provide the foundation for designing the solution artefact. In accordance with the DSR approach, the

identified requirements are translated into a concrete system architecture that demonstrates how agent-based AI capabilities can be embedded into an existing omnichannel customer service environment. The following section presents the design of this solution artefact and explains how the identified business and technical requirements are addressed through the proposed architecture.

4.4 Solution artefact design

Based on the identified business and technical requirements, a solution design process was initiated. The proposed solution architecture is built on three core Microsoft cloud components: Microsoft Dynamics 365, Microsoft Power Platform, and Microsoft Azure. Microsoft Dynamics 365 is a suite of business applications that support organisational processes across areas such as sales, marketing, service, finance, and supply chain. Dynamics is designed to help organisations manage their daily operations through integrated, modular business capabilities. In this architecture, the focus is on customer service functionality that enables handling interactions across multiple communication channels. (Microsoft, 2026a)

The second component of the architecture is the Microsoft Power Platform, which provides tools for low-code and no-code application development and process automation. Within this platform, Microsoft Dataverse serves as the unified data storage and management layer for operational information across applications. Power Automate automates workflows and triggers system processes in response to events in the CRM environment. (Microsoft, 2026b)

The third component is Microsoft Azure, a cloud computing platform that provides a wide range of services for application hosting, data processing, artificial intelligence, and integration. Azure provides the computing and service infrastructure required to run scalable and extensible cloud-based applications. (Microsoft, 2026c) The high-level solution architecture is illustrated in Figure 19 below.

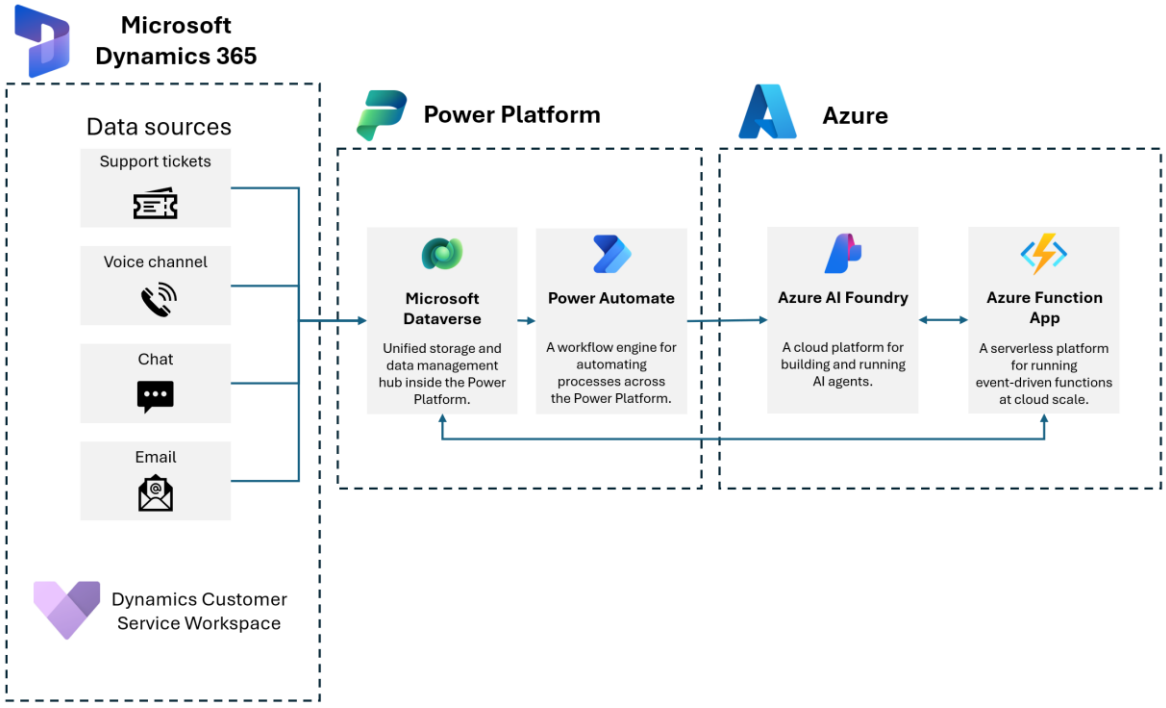

Figure 19 High-level architecture

In the architecture diagram, the flow of information originates from four primary data sources within the omnichannel customer service module: chat, email, support tickets and the voice channel. These sources represent the core channels through which customer interactions are created and managed in Dynamics 365 Customer Service Workspace. All structured interaction content generated through these channels is stored in Microsoft Dataverse, which functions as the centralised data hub of the Power Platform. Dataverse organises interaction data using a relational schema, where each customer event is represented as a row in the appropriate data table. This ensures consistent representation of interaction records and provides a unified foundation for subsequent analytical and agentic processing.

Voice channel interactions differ from other data sources because they produce audio files that cannot be stored directly in a relational database in Dataverse. Instead, Dataverse stores the associated metadata alongside a reference to Azure Blob Storage, where the MP4 audio files are stored. This separation allows the system to handle large audio files efficiently while preserving a clear logical linkage between the relational representation of an interaction and its corresponding audio content. Although this linkage is integral to the overall data model, it is not explicitly depicted in the architecture diagram.

The Dynamics voice channel includes built-in AI capabilities that generate text transcripts and perform sentiment analysis for each voice interaction. These results are stored directly in Dataverse, ensuring that both the textual representation and the automatically generated sentiment information are readily available for downstream processes. This unified and enriched data foundation provides the autonomous customer service agents described later in the thesis with access to a consistent set of input features regardless of the originating channel.

Within the Power Platform, Power Automate is used in this work as the most efficient and low-overhead mechanism for detecting changes in Dataverse and emitting reliable event-based notifications. Its purpose is not to orchestrate the agent logic, but to provide a lightweight, managed, and easily maintainable trigger layer that integrates natively with the existing data model. Power Automate is selected because it offers built-in connectors, governance, and lifecycle management for Dataverse events, removing the need to implement custom polling, webhook infrastructure, or separate monitoring services. Once these events are raised, the actual reasoning, planning, and orchestration of the autonomous agents takes place entirely within the Azure environment.

On the Azure side, the architecture introduces the components responsible for interpreting and processing the interaction data. Azure AI Foundry provides the environment in which autonomous agents are hosted and in which interaction inputs from Dataverse are transformed into higher-level representations. Although the purpose of this chapter is limited to describing how information is stored and made available rather than explaining how the agents operate internally, it is important to highlight that the consistent and structured data produced by the Dynamics and Dataverse layers form the essential basis on which the agentic reasoning processes rely.

Azure Function Apps complete the architectural flow by providing a serverless computational environment for executing system-level operations initiated by the orchestrator agent. These functions do not determine which actions should be performed; instead, they serve as callable operational capabilities that can be invoked as needed. Their event-driven nature ensures that operational steps can scale in accordance with demand while maintaining a clear separation between the reasoning processes executed within Azure AI Foundry and the concrete system actions executed in the Dynamics ecosystem.

Taken together, the architecture provides a coherent, structured information environment in which multimodal interaction data, metadata, transcription outputs, and sentiment information are captured, enriched, and made accessible. This environment provides the foundation for the autonomous customer service agents described in subsequent chapters to interpret customer interactions and support decision-making across all relevant service channels.

Overall, the architecture enables an end-to-end ecosystem for autonomous customer service agents across the supported service channels. It does so by separating customer interaction storage, event triggering, agent reasoning, and system-level execution into distinct but connected layers. Microsoft Dataverse provides a consistent, channel-independent case context for agents, while Power Automate enables reliable event-based triggering without orchestrating the agent logic. The actual reasoning, planning, and orchestration of the autonomous agents takes place within the Azure environment, and Azure Function Apps provide a serverless execution layer for system-level operations initiated by the orchestrator agent. This makes the Microsoft ecosystem particularly suitable for the present study, as it allows customer interaction data, workflow triggers, and AI services to operate within a shared enterprise environment. Together, these design choices make the architecture suitable for examining controlled and incremental autonomy in customer service.

4.5 Solution development

This section describes both the technical implementation of the artefact and the key design decisions that shaped its final structure. The development of the agent-based solution progressed through an iterative design process shaped by the operational and technical constraints of integrating autonomous reasoning into customer service channels. Rather than following a fully predetermined architecture from the outset, the solution structure evolved as requirements related to real-time audio processing, data integration, and agentic reasoning became clearer during the project.

A central architectural decision concerned the implementation of the voice channel. The initial plan was to integrate the autonomous voice agent via Microsoft Calling Plan, Microsoft’s cloud-based public switched telephone network (PSTN) service for Teams Phone. In this model, Microsoft functions as the telecommunications operator, and all

signalling and media traffic is routed entirely through Microsoft's managed infrastructure. While this approach is operationally convenient for standard enterprise calling scenarios, it imposes several technical restrictions that proved decisive in this work.

Most importantly, the Calling Plan permits the use only of conversational capabilities available in Microsoft Copilot Studio. Although Copilot Studio offers a built-in framework for voice agents, its language support is limited; for example, full-fidelity Finnish speech and NLP capabilities are not yet available. Because the project aimed to deploy a custom language model that supports multiple languages with high accuracy, the Calling Plan was incompatible with the requirements. Furthermore, Calling Plan does not provide the administrative control needed to redirect the audio stream to a self-managed Azure Communication Services (ACS) resource, nor does it permit the attachment of custom models or media processors to the voice pipeline.

Due to these limitations, an alternative architecture based on an external PSTN operator was implemented. This approach restored full administrative control over call routing, enabling incoming calls to be directed into ACS rather than remaining within Microsoft's locked telephony layer. Once calls reach ACS, they can be streamed in real time to the agent via a webhook and WebSocket pipeline, thereby supporting autonomous reasoning and multilingual interaction at the required quality level. Figure 20 shows how AI is integrated into the voice channel.

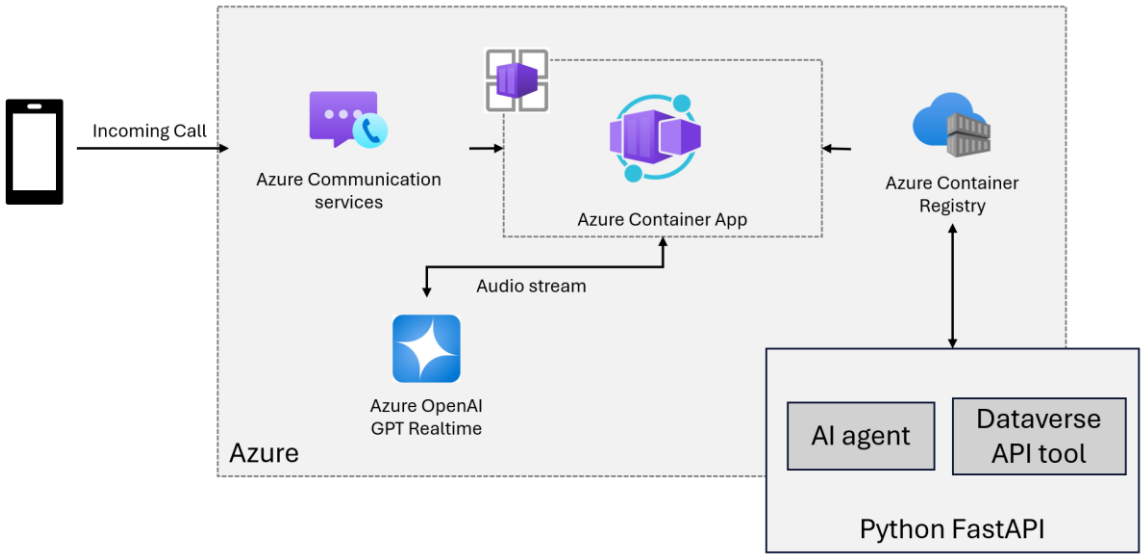

Figure 20 Architecture of AI Voice Channel

After call routing to ACS was established, the next step was to implement the ACS agent runtime integration. A containerised application served as the intermediary, coordinating call events and audio streams. ACS emitted events to a FastAPI-based webhook endpoint, which then established a persistent WebSocket connection. Through this connection, raw audio frames were forwarded continuously between ACS and the multimodal autonomous agent. This approach provided the low-latency loop required for the agent to process speech input, generate responses and maintain contextual continuity during the call.

The agent's functional capabilities were expanded through an integrated tool interface that exposed the Dataverse application programming interfaces (API). Through this interface, the agent could retrieve and update customer records, initiate service requests, or perform other operational tasks in real time. This enabled the agent to act not only as a conversational participant but also as an active component of the operational system, contributing to case creation and management during the voice interaction.

While the voice channel was being developed, the chat channel was implemented through a streamlined process that relied more heavily on Microsoft Copilot Studio's native capabilities. The customer-facing chat experience was embedded in a portal environment, where incoming messages were routed to an autonomous conversational agent built on a GPT-4.1 model hosted within Copilot Studio. Unlike the voice channel, the chat workflow did not require PSTN connectivity, ACS routing, or custom audio pipelines. Instead, it leveraged Copilot Studio’s built-in orchestration mechanisms, including Interactive Voice Response (IVR)-like dialogue flows, topic-based routing, and configurable conversational logic. Figure 21 below shows the architecture of the chat agent.

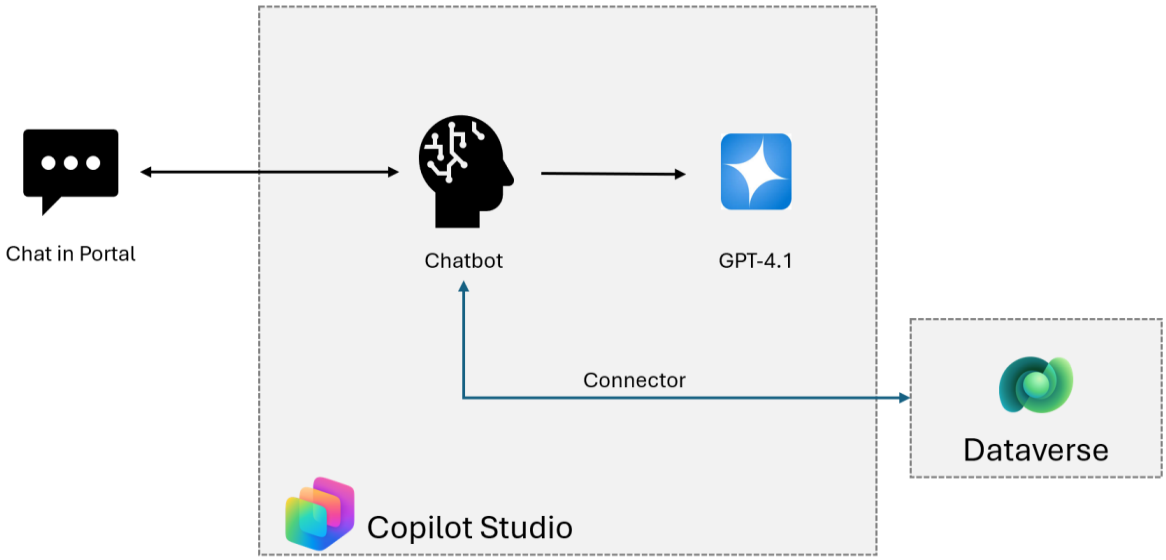

Figure 21 Architecture of Chat Agent

These IVR-style flows guided the customer through initial classification or triage steps, such as identifying the purpose of the inquiry or confirming the relevant service category. While the voice agent relied on a custom real-time model deployed outside Microsoft's agent framework, the chat agent operated exclusively through Copilot Studio. As a result, it did not share the same Dataverse tool interface as the voice agent, nor did it interact with Dataverse via custom API calls. Instead, Copilot Studio handled data operations via its own connectors and configurations, enabling the chat agent to create service requests or update selected records through its internal integration layer.

To ensure operational robustness, the chat channel was additionally equipped with a structured escalation mechanism. When the agent detected uncertainty, lacked sufficient confidence in its response or encountered a request requiring human judgment, it could escalate the conversation to a human support agent through the portal interface or the organisation's customer service queue. This escalation path provided continuity and ensured that complex or sensitive cases were handled appropriately.

Overall, the chat channel offered a lighter-weight, platform-aligned implementation compared with the voice agent. By combining GPT-4.1–based reasoning with Copilot Studio's IVR-driven conversational scaffolding and integrated data connectors, the chat agent delivered a consistent customer experience while operating within a more constrained yet operationally efficient architecture.

Other communication channels, such as email and support tickets, were incorporated using Automatic Record Creation (ARC) rules. These rules generated Dataverse records immediately upon receiving customer messages. Although these channels did not rely on real-time reasoning, their integration ensured that all customer communications were represented uniformly in the shared data environment. This also enabled the voice and chat agents to access the full history of customer interactions as part of their reasoning processes.

Through this sequence of design decisions and iterative refinements, the solution evolved into a unified architecture that supports autonomous agent functionality across multiple communication channels. The combination of external operator routing, ACS-based media control, multimodal reasoning capabilities and Dataverse-based operational tools established a flexible and extensible foundation for autonomous customer service.

After the interaction layer agents capture the customer's request and create the corresponding service record in Dataverse, the solution transitions from request intake to analytical processing. Before introducing the analytical layer, it is essential to outline the Dataverse data model that the agents operate on. The interaction layer agents do not attempt to resolve the customer's issue; rather, they structure and store the interaction as a set of linked Dataverse entities. These entities form the informational substrate on which the later analytical agents in Azure AI Foundry base their reasoning.

The core entity in this structure is the incident, created by the voice, chat, email or ticketing channels. Each case is associated with a contact and, where applicable, an account, establishing the customer context for subsequent analysis. Interaction-specific content is captured in conversation records, which reference the originating channel and are linked to transcripts or notes containing extracted utterances, summaries or classifications produced during intake. Email-based interactions are represented by the email entity, while asynchronously generated tasks and follow-up actions use the activity entity. These multiple sources of structured and semi-structured information are interconnected through Dataverse's relational schema, ensuring that all case-relevant data is available to downstream reasoning components.

Figure 22 illustrates the relevant subset of the Dataverse schema, showing how incidents, contacts, accounts, conversations, emails, transcripts, notes, channels and activities relate to one another. This data model provides the foundation for the analytical agents described in

the next section, enabling them to retrieve case context, interpret customer interactions, and synthesise diagnostic insights consistently.

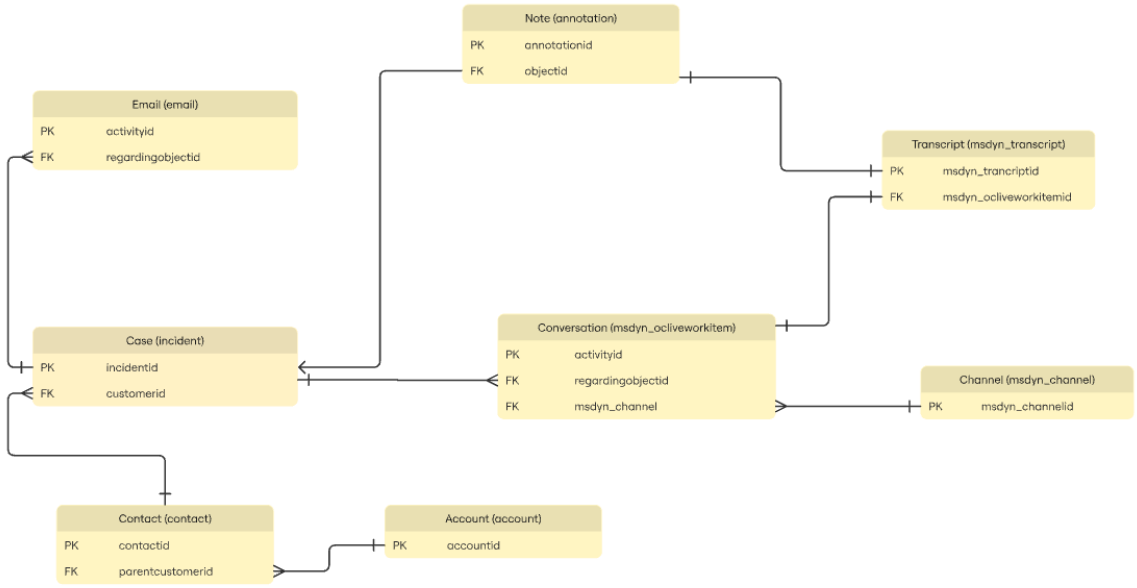

Figure 22 ER diagram of the Dataverse

The analytical layer is implemented as a Python 3.11 application deployed on Azure Functions. It relies on the Azure AI Projects SDK, which manages the tool-calling loop, function tool registration, and agent run lifecycle. Access to Dataverse is authenticated via an Entra ID service principal using the client credentials flow, while the connection to Azure AI Foundry is established through a managed identity resolved at runtime via DefaultAzureCredential. This separation of authentication mechanisms reflects the different trust boundaries between the operational CRM environment and the AI platform layer. These implementation choices were made to ensure that the analytical layer supports controlled autonomy, preserves clear trust boundaries between operational and AI components, and enables reproducible evaluation of agent behaviour within the DSR context.

When a new service request is created in Dataverse, a trigger pipeline invokes the orchestration agent via a dedicated HTTP endpoint, passing the service request identifier, customer identifier, service request description, current status code, and a set of control parameters that govern write-back behaviour and escalation handling. The orchestration agent assumes end-to-end responsibility for the reasoning trajectory. It begins by assessing the service request metadata, including interaction history, customer profile, and any

extracted summaries or classification results produced earlier in the process, then develops a reasoning plan comprising one or more subtasks.

To support continuity across multi-turn interactions, the system maintains a persistent memory layer backed by Azure Cosmos DB. This store maps each incident identifier to a corresponding agent thread identifier, enabling the orchestration agent to resume an existing reasoning thread when additional messages or updates arrive for the same service request. This design ensures that contextual continuity is preserved across asynchronous interactions, such as follow-up emails or delayed customer responses, without requiring the agent to reconstruct its reasoning from scratch.

The subtasks derived from the orchestration agent's reasoning plan are delegated to three specialised helper agents, each designed to support a distinct function within the analytical workflow. Rather than operating as fully autonomous units, these agents serve as bounded components that contribute targeted outputs to the orchestration agent's supervisory reasoning process. This division of responsibility is a deliberate architectural choice; by constraining each agent to a well-defined functional scope, the system maintains explainability and ensures that autonomous action remains incremental and supervised throughout the pipeline.

The diagnostics agent compiles the contextual evidence required for downstream analysis. It interfaces directly with Dataverse via its OData REST API to retrieve the active service request record, the customer's full service request history, customer profile information, and associated flight and booking records. Retrieving the service request history is particularly significant, as it enables detection of duplicate or previously unresolved incidents, thereby supporting the consolidation of handling into a single active record. All retrieved records are pre-processed before being passed to the language model: numeric option-set codes are translated into human-readable labels, and irrelevant metadata is removed. This normalisation step reduces the risk of hallucinations and improves the accuracy of the orchestration agent's subsequent reasoning. The orchestration agent itself, not a separate component, interprets the compiled evidence and infers the underlying cause of the customer's issue, drawing on the structured outputs provided by the diagnostics agent.

The resolution agent synthesises contextual inputs and the orchestration agent's diagnostic reasoning to determine and execute a set of actionable resolution steps. Unlike a purely

advisory component, the resolution agent has direct write access to Dataverse via its OData REST API: it can create, modify and cancel booking records, add observations or responses to the service request record, create follow-up tasks linked to the service request, and update the service request status to reflect the current handling stage. If a duplicate service request is detected, the resolution agent cancels the redundant record and consolidates handling into the original incident. This direct operational capability is a deliberate design choice that enables end-to-end case handling within the automated pipeline, with autonomous execution determined by the orchestration agent as appropriate, while escalation to human representatives remains available for service requests that exceed the system's confidence threshold.

The communication agent is responsible for producing all customer-facing and internal written outputs. It operates on consolidated inputs from the diagnostics agent and the orchestration agent's reasoning, producing natural-language content appropriate to the communication context. This includes drafted responses to customer enquiries, requests for additional documentation, clarification messages where required information is missing, and internal case notes written back to the service request record in Dataverse. The agent does not initiate contact independently; it generates output only when invoked by the orchestration agent, either to support an automated response pipeline or to assist a human customer service representative in composing an appropriate reply.

Each helper agent operates within a strictly bounded reasoning loop and contributes only within its own functional domain. Upon completing its subtask, the agent returns a structured output to the orchestration agent, which evaluates the combined results, determines whether additional subtasks are required, and manages the overall reasoning trajectory. This supervised design prevents any individual component from acting unilaterally while enabling the system to incorporate multiple functional perspectives in a controlled and transparent manner.

Once the orchestration agent determines that a coherent and actionable level of analysis has been reached, it consolidates the helper agents' outputs into a structured resolution package. When write-back is enabled, the package is persisted as an annotation on the incident record in Dataverse, where it can be reviewed and acted upon by customer service personnel or by downstream automation components. If the orchestration agent determines that the case exceeds the system's analytical scope or requires human judgment, it triggers an escalation

path: an annotation is written to the service request record indicating that the service request has been referred to a human representative, and the service request status is updated accordingly. This escalation mechanism ensures that complex or sensitive cases are handled appropriately without compromising the integrity of the automated pipeline.

Through this orchestrated multi-agent architecture, the system enhances analytical depth, modularity and explainability. The three helper agents collectively augment the customer service workflow by contributing targeted functional capabilities, while the orchestration agent retains overall reasoning authority and final decision-making responsibility. The resulting design operationalises the principle of controlled autonomy: agent behaviour is transparent, bounded and reproducible, supporting both the functional objectives of the customer service context and the evaluative requirements of the DSR methodology.

4.6 Solution demonstration and evaluation

Demonstration Setting and Simulation Environment

The demonstration phase operationalised the artefact in a simulated customer service environment built on Microsoft Dynamics 365 CRM. The evaluation was conducted in a controlled simulation representing a fictional travel service provider, with all customer profiles, bookings, flight records, and service requests synthetically generated for research purposes.

In the evaluated architecture, the intake and resolution phases are intentionally separated. The four channels are used exclusively to simulate the creation of service requests, each producing a structured incident record in Dataverse. The autonomous customer service agent evaluated in this study operates downstream from intake and is triggered only after the service request has been created. As a result, channel-specific interaction mechanics are not evaluated, the simulation focuses on autonomous case resolution based on a channel-independent incident representation.

In the full system architecture, customer contacts arrive through four intake channels: voice calls, emails, chat, and the ticketing interface, each of which generates a corresponding incident record in Dataverse upon receipt. Service requests from all four channels were represented in the simulation, reflecting the artefact's omnichannel scope. Rather than

deploying the system in a live production setting, a controlled simulation was constructed to enable systematic and reproducible evaluation of agent behaviour across a range of service scenarios. This approach was chosen because it allowed the artefact to be evaluated safely in a representative customer service context without exposing real customers or operational processes to potential failures during development.

The simulation environment was populated with realistic test data, including customer profiles, flight records with varying operational statuses, multi-passenger bookings with different insurance levels, and historical service request records. Each test case was created as a Dynamics 365 CRM incident record linked to the relevant booking, flight, and customer data, with the customer's enquiry recorded in the case description field. Upon creation, a Power Automate trigger invoked the orchestration agent, which processed the case and wrote its response back to Dataverse as an annotation on the service request record. This end-to-end execution preserved the structural form of a real intake-generated case while enabling controlled variation of case content.

Test Case Design

Fifteen test cases were constructed across four complexity tiers, each probing distinct dimensions of the agent's reasoning and operational capabilities. The tier structure was designed to reflect the study’s objective of evaluating autonomous customer service performance across cases of increasing complexity and to identify when autonomy remains appropriate and when human escalation becomes necessary. Tier assignment followed a consistent classification logic: Tier 1 cases involve unambiguous information requests that require no policy interpretation. Tier 2 cases involve straightforward transactional requests where the applicable policy is clear, but execution requires data retrieval and record manipulation. Tier 3 cases introduce policy ambiguity, emotional complexity, or historical context that require interpretive judgement beyond direct rule application, and Tier 4 cases are characterised by operational urgency, safety-relevant circumstances, or multi-dimensional resolution requirements that place autonomous handling at or beyond the system's confidence boundary, making escalation the expected outcome.

Tier 1 cases covered the most common service interactions encountered in day-to-day customer service operations. T01 tested the agent's ability to retrieve and present booking details, including flight itinerary, passenger information, and applicable insurance coverage,

from a linked CRM record. T02 assessed flight status enquiry handling, where the agent was expected to confirm the flight's operational status and provide check-in guidance. T03 involved an enquiry about a cancellation policy for a booking with a departure date within seven days, requiring the agent to identify and communicate the applicable 100% cancellation fee and any force majeure exceptions.

Tier 2 cases introduced transactional complexity requiring both policy application and record-level operations. T04 tested flight rebooking with a price difference calculation and was conducted in Swedish to assess multilingual handling. T05 assessed the registration of three simultaneous special service arrangements for a single booking, including wheelchair assistance, a gluten-free meal preference for two passengers, and a seating request near the exit, and was conducted in English. Both cases were designed to verify whether the agent could handle multi-attribute requests and produce accurate, channel-appropriate responses in the customer's language.

Tier 3 cases introduced conditions requiring interpretive judgment and contextual sensitivity. T06 involved a cancellation request supported by a medical certificate documenting a force majeure condition, testing the agent's ability to identify the applicable insurance coverage and waive the standard cancellation fee. T07 presented a price dispute in which the customer claimed a loyalty discount that the booking record did not confirm, requiring the agent to neither concede nor dismiss the claim, but instead to request supporting documentation and to record the dispute for internal review. To illustrate the interaction flow for a single Tier 3 case, example transcripts for T06 are provided in Appendix 3 and Appendix 4. Appendix 3 shows the voice intake interaction that captures the customer’s request and creates the incident record, while Appendix 4 presents the corresponding customer-facing interaction produced during case handling for the same incident.

Tier 4 cases represented the most demanding category and were designed to probe the boundaries of autonomous operation. T08 involved a family of four, including two young children, stranded at Rome airport following a same-day cancellation of their return flight, requiring the agent to identify the next available departure, confirm accommodation and meal entitlements under European Union (EU) Regulation 261/2004, and communicate insurance coverage within a single interaction. T09 received a complaint from a customer whose outbound flight arrived more than 3 hours late due to airport congestion, resulting in the loss of a pre-booked hotel night, as well as an aggressive communication style and a

reference to a prior unanswered service request. T10 involved a repeat complaint about a business-class seat assignment error that had remained unresolved beyond the stated 14-working-day response commitment, testing the agent's ability to identify the prior open incident and initiate appropriate remediation. T11 tested partial cancellation handling for a group booking in which one of three passengers wished to withdraw within seven days of departure, requiring calculation of the applicable per-person cancellation fee and confirmation that the remaining passengers' reservations would be unaffected. T12 assessed an EU261 compensation claim following a flight delay of two hours and ten minutes, which falls below the three-hour threshold for statutory compensation, requiring the agent to apply the eligibility threshold correctly and withhold confirmation of entitlements not supported by the available data. T13 involved an emotionally charged interaction from a customer whose flight to Rome had been cancelled at short notice, resulting in the loss of concert tickets and a prepaid hotel night, with an explicit threat to escalate to the consumer authority. T14 combined a price dispute regarding an unconfirmed discount with the same cancelled flight scenario as T13, requiring the agent to manage both issues within a single interaction while directing the pricing matter to the appropriate internal team. T15 presented a double-booking crisis in which a customer held two active reservations on the same departure, a business class booking and an economy class family booking, with thirty minutes remaining before gate closure, requiring immediate identification of the conflict and physical escalation to the airline check-in desk.

The evaluation was conducted across two consecutive simulation rounds. The first round (S1) used the baseline prompt architecture developed during the design and development phase. Following a systematic review of failure patterns identified in S1, a series of iterative refinements were applied to both the agent prompts and the underlying orchestration logic for the second round (S2). These modifications addressed six categories of deficiency: non-deterministic handling of multi-issue scenarios, hallucination of factual claims not supported by retrieved data, insufficient empathetic tone calibration in emotionally charged interactions, absent proactive behaviour in time-critical flight disruption cases, inconsistent application of pricing and cancellation policies, and unreliable language handling in multilingual interactions. The changes were implemented incrementally, with each revision targeting a specific failure pattern while preserving the overall agent architecture.

Evaluation Criteria and Measurement Approach

The evaluation assessed agent performance across four criteria: autonomy outcome per case, factual accuracy of responses, correctness of escalation decisions, and alignment with SERVQUAL service quality dimensions. The autonomy outcome was classified as autonomous, correctly escalated, or requiring human correction. Cases were pre-classified by the researcher as either autonomy-appropriate or escalation-required before evaluation, providing a reference against which actual agent behaviour could be compared. Factual accuracy was evaluated by comparing the agent's substantive claims against applicable service documentation and the simulation data. SERVQUAL alignment was assessed using a structured rubric operationalising the five dimensions on a five-point scale. As the classical operationalisation of tangibles refers to physical facilities and equipment, this dimension was reinterpreted in the agent context to refer to the clarity, structure, and professional quality of the written response.

The full rubric is provided in Appendix 2. This rubric-based approach constitutes a content analysis of agent outputs conducted from the perspective of a knowledgeable evaluator, consistent with established practice in DSR evaluation where expert judgement is a recognised method for assessing artefact quality against theoretically grounded criteria.

The empirical results of the evaluation are summarised in Tables 5 and 6. Table 5 presents the case-level outcomes across all fifteen test scenarios, including autonomy decisions, factual accuracy, and SERVQUAL scores, while Table 6 aggregates the results by SERVQUAL dimension to provide an overall performance overview.

Table 5 Case evaluation

Case	Tier	Issue Type	Autonomy Outcome	Factual Accuracy	Escalation Correct	SERVQUAL avg.
T01	1	Booking enquiry	Autonomous	Correct	N/A	5.0
T02	1	Flight status	Autonomous	Correct	N/A	5.0
T03	1	Cancellation terms	Autonomous	Correct	N/A	5.0
T04	2	Flight rebooking	Autonomous	Correct	N/A	5.0
T05	2	Special arrangements	Autonomous	Correct	N/A	5.0
T06	3	Force majeure	Autonomous	Correct	N/A	5.0
T07	3	Price dispute	Correctly Escalated	Correct	Correct	4.6
T08	4	Cancelled flight	Correctly Escalated	Correct	Correct	5.0
T09	4	Delay and compensation	Correctly Escalated	Partial	Correct	4.4
T10	4	Business class complaint	Correctly Escalated	Correct	Correct	4.4
T11	4	Partial cancellation	Autonomous	Correct	N/A	4.8
T12	4	EU261 delay claim	Autonomous	Correct	N/A	4.4
T13	4	Angry customer and cancellation	Correctly Escalated	Correct	Correct	4.8
T14	4	Price dispute and cancellation	Correctly Escalated	Correct	Correct	4.8
T15	4	Double-booking crisis	Correctly Escalated	Correct	Correct	4.8

Across 15 test cases, the system achieved a mean SERVQUAL score of 4.83 out of 5.00 in S2. Seven cases (47%) were resolved autonomously, and eight (53%) were correctly escalated. No case required post-hoc human correction of an erroneous autonomous resolution in S2, in contrast to one such instance in S1. Nine cases achieved a perfect score of 5.0. The lowest scores were recorded in T10 (4.4, business class complaint) and T12 (4.4, EU261 delay claim). In T10, the primary deficiency was the omission of a specific complaint resolution timeline referenced in the service guidelines. In T12, the agent correctly withheld an unverifiable EU261 compensation amount, which reduced perceived completeness without constituting a factual error. In T09, the agent provided two explanations for the delay, citing both airport congestion, which was the documented cause, and a technical fault, which was not supported by the available data. As the correct cause was included in the

response, this instance is classified as a minor factual inaccuracy (unsupported supplementary information) rather than a major factual error, and is discussed further in the threats to validity section.

Table 6 SERVQUAL dimension scores across simulation rounds

Dimension	S1 Mean	S2 Mean	Difference	Key Improvement Factor
Reliability	4.80	4.87	+0.07	Elimination of hallucinated insurance benefits and process timelines
Responsiveness	4.53	4.80	+0.27	Proactive flight disruption handling; reduced confirmation latency
Assurance	4.67	4.87	+0.20	Structured escalation logic; mandatory policy verification
Empathy	4.00	4.67	+0.67	Tone calibration; contextual closings; first-name addressing
Tangibles	4.93	4.93	0.00	Consistent structured output across both rounds
Overall	4.59	4.83	+0.24	Compound effect of all prompts and code refinements

Several patterns emerged across the evaluation that are not fully captured by the quantitative criteria alone. The Empathy dimension showed the most pronounced improvement (+ 0.67), driven by tone calibration interventions that required the agent to explicitly acknowledge the customer's situation before proceeding to resolution steps. In S1, emotionally charged cases such as T13 and T15 received markedly lower Empathy scores because the agent addressed the factual substance of the claim before acknowledging the customer's distress; the S2 refinements inverted this sequence without reducing factual accuracy. Responsiveness improved by 0.27, with the largest single gain in T15, where a double-booking crisis with a 30-minute window before departure was handled passively in S1 (score: 3.2) and resolved with an immediate, structured escalation in S2 (score: 5.0). Reliability improved modestly, reflecting the already high baseline, while Tangibles remained stable at 4.93, indicating that structural response quality was near-optimal in both rounds. Total hallucination events decreased from seven in S1 to two in S2, a reduction of 71%, with fabricated insurance benefit details and invented process timelines entirely eliminated. The two remaining

instances involved EU261 compensation amounts stated before the underlying delay duration was fully confirmed by the flight record, indicating a residual gap in the orchestration agent's data-verification sequence.

A recurring pattern across SERVQUAL dimensions was that reliability and tangibles were more consistently achieved than empathy and assurance, which require sensitivity to emotional context and appropriate tone calibration. This asymmetry reflects a known limitation of large language model-based agents in affective communication and suggests that human oversight remains most valuable in emotionally charged interactions such as T09 and T13.

Taken together, the evaluation results contribute directly to the three research questions of this thesis. First, they show that service quality in AI-supported customer service is shaped not only by factual accuracy, but also by the agent's ability to align with key service quality dimensions and to recognise cases that require escalation. Second, the evaluation across 15 test cases and 4 complexity tiers demonstrates how agent performance varies across customer service scenarios of increasing complexity. Third, the findings indicate that the current potential of autonomous AI agents lies in controlled and incremental deployment supported by shared CRM data, coordinated multi-agent architecture, and explicit mechanisms for escalation and oversight.

Threats to Validity

Several threats to validity are acknowledged. With respect to internal validity, the simulation environment was constructed by the researcher and may not fully capture the variability and unpredictability of real customer interactions; the controlled nature of the test cases reduces ecological realism while increasing reproducibility. External validity is limited by the single organisational context and fixed data set used in the evaluation; the findings may not generalise directly to other service domains, data configurations, or organisational settings. Construct validity poses a specific challenge in applying SERVQUAL to the evaluation of autonomous agent outputs; the reinterpretation of dimensions such as tangibles and the absence of customer-side expectation measurement constitute adaptations whose validity rests on the theoretical argument that the underlying constructs remain applicable in the agent-mediated context. The reliability of the SERVQUAL scoring is subject to single-evaluator bias, as all rubric assessments were conducted by the researcher without

independent inter-rater verification. T09 required a manual adjustment to the flight arrival time in the CRM data after an inconsistency was detected between the test scenario specification and the underlying record; this intervention is documented as a threat to reproducibility. Future evaluations should consider involving multiple raters, larger and more varied case sets, and live deployment conditions to strengthen both internal and external validity.

5 Discussion

The findings of this thesis suggest that the meaningful use of autonomous AI agents in customer service does not lie in maximising autonomy as an end in itself, but in identifying the conditions under which autonomy can be applied without compromising service quality, transparency, or accountability. The empirical results demonstrate that autonomy in customer service is neither binary nor uniform across interaction types, but emerges as a gradual and context-dependent property shaped by task complexity, emotional sensitivity, operational risk, and time pressure (Dwivedi et al., 2025; Ranjan et al., 2025). This interpretation is consistent with the incremental autonomy trajectory identified in the interview phase, in which interviewees positioned full autonomy as a long-term prospect contingent on both technological maturity and organisational readiness, rather than an immediate operational reality.

Autonomous handling proved effective in routine and moderately complex scenarios involving information retrieval, transactional updates, and clearly defined policy application, achieving high factual accuracy and service quality without human intervention. Conversely, cases characterised by emotional distress, ambiguous responsibility, legal interpretation, or urgent operational constraints consistently required escalation, highlighting clear boundaries for autonomous operation. (Li & Zhang, 2023) These boundaries were not incidental but reflected deliberate design choices embedded in the agent's confidence thresholds, escalation rules, and verification mechanisms, which collectively ensured that human accountability was preserved where autonomous judgment was most likely to be insufficient. Autonomy should therefore be understood as controlled execution operating within predefined constraints rather than as independent decision-making authority. Table 7 summarises the main theoretical themes discussed in the literature and their corresponding empirical findings in this study.

Table 7 Synthesis of theoretical themes and empirical findings

Theoretical theme/expectation	**Empirical finding**	**Interpretation**
Autonomy is context-dependent	Strongest in routine and transactional cases	Supports bounded, non-binary autonomy
Emotional interaction is difficult for AI	Empathy remained the weakest SERVQUAL dimension	Human escalation remains necessary in affective cases
Control and architecture matter	Shared CRM data + orchestration + escalation improved outcomes	Controlled autonomy is architectural as well as agent-level
Process clarity affects performance	Poorly formalised policies reduced consistency	Readiness must be assessed at the process level

As shown in Table 7, the empirical findings broadly support the theoretical expectation that autonomy in customer service is context-dependent and also indicate that its practical feasibility depends on process formalisation, escalation logic, and architectural control mechanisms.

The application of SERVQUAL as an expert evaluation framework revealed that reliability, responsiveness, and assurance can be systematically achieved and improved through iterative refinement of agent logic and orchestration. Empathy, however, remained the most challenging dimension, particularly in emotionally charged interactions (Blümel et al., 2024). Although targeted refinements to the agent's tone calibration and acknowledgement sequencing reduced this gap measurably across the two simulation rounds, the results reinforce the continued necessity of human involvement in situations where emotional reassurance and contextual judgement are central to the service experience. This asymmetry between transactional and affective service quality dimensions has important practical implications for organisations evaluating the deployment boundaries of autonomous agents.

From an architectural perspective, the findings indicate that the current potential of autonomous AI agents in customer service lies in controlled, incremental deployment rather than uniform autonomy across all channels or interaction types. Measurable value was achieved through a shared CRM environment with consistent data representation, combined with selective deployment of autonomous reasoning governed by confidence thresholds and

escalation rules. In this way, human accountability could be preserved at the points where autonomous judgment was most likely to be insufficient. These findings extend prior research on agentic AI by showing that effective autonomy in customer service should not be understood as a binary end state, but as a context-dependent capability whose feasibility depends on architecture, control mechanisms, and the operational conditions of the service process. The variation in autonomy levels across case types should therefore be understood not as a limitation, but as a deliberate design condition.

A further contribution of this study is the finding that autonomous agent behaviour is structurally dependent on the existence of well-defined service processes. When policies and escalation criteria are not explicitly formalised, agents cannot apply them consistently. This suggests that readiness for incremental autonomy must be assessed not only at the technology level, but also at the process level. In theoretical terms, this refines prior work on AI-enabled service automation by indicating that the practical boundary of autonomy is shaped not only by the agent's technical capabilities but also by the extent to which service logic can be formalised into machine-executable structures. The iterative refinement process documented across the two simulation rounds further illustrates how prompt engineering and orchestration logic can serve as practical mechanisms to improve agent performance without changes to the underlying model infrastructure.

From a practical perspective, the results suggest that organisations can achieve substantial gains in efficiency and consistency without pursuing full autonomy, provided that autonomous behaviour remains carefully bounded and transparently governed. In the refined simulation round, hallucination events were reduced by 71% and escalation decisions were correct in all evaluated cases, indicating that targeted architectural controls can materially reduce the risk profile of autonomous deployment. When combined with shared CRM data and structured service logic, such controls enabled autonomous case handling while preserving accountability in situations where autonomous judgment was most likely insufficient. At the same time, the findings reinforce that autonomous agent behaviour cannot be separated from the organisational context in which it operates. In this sense, autonomy does not emerge from the agent alone, but from the clarity and completeness of the processes it is intended to execute. This implies that investments in agent technology should be preceded by, or at least conducted in parallel with, a systematic review and formalisation of the underlying service processes.

These findings are subject to limitations arising from the simulated evaluation environment, the domain-specific nature of the travel service context, and the single-evaluator assessment approach, all of which constrain the generalisability of the observed patterns. Nevertheless, the results provide a useful basis for interpreting how autonomous AI agents can be deployed in a controlled and meaningful manner within a representative enterprise CRM setting. Rather than supporting a general claim in favour of full autonomy, the findings point toward a more conditional model in which value emerges through bounded deployment, explicit escalation logic, and process environments that are sufficiently formalised to support consistent agent behaviour.

6 Conclusions

This thesis examined how autonomous AI agents can be utilised in omnichannel customer service in a controlled and meaningful manner using current technologies. The study addressed both a practical organisational need and a recognised research gap, as prior research has focused predominantly on single-channel chatbot deployments or isolated automation features rather than on coordinated multi-agent architectures embedded in enterprise CRM environments. Using a DSR approach, the study combined a structured literature review, qualitative interviews with industry professionals, and the design, implementation, and evaluation of a multi-agent customer service artefact within Microsoft Dynamics 365.

The first research question asked:

"What are the key factors that determine customer service quality in AI supported omnichannel environments, and under what conditions is autonomous handling appropriate?"

The literature review established that service quality in AI-supported contexts is shaped by the same fundamental dimensions identified in the SERVQUAL framework, namely reliability, responsiveness, assurance, empathy, and tangibles, while adding the distinctive challenge that autonomous agents must achieve these dimensions without the contextual flexibility and emotional adaptability that human agents possess. The qualitative interviews reinforced this theoretical foundation by identifying response time, issue-resolution capability, and channel consistency as primary quality drivers, and indicated that customer acceptance of AI-handled interactions is conditional on task simplicity and perceived reliability. The empirical evaluation confirmed these patterns: autonomous handling consistently achieved high scores on reliability, responsiveness, and assurance across routine and transactional scenarios, while the empathy dimension remained systematically lower in emotionally charged interactions. Together, the three sources of evidence converge on the conclusion that autonomous handling is appropriate when task complexity is low or clearly structured, when applicable policies are unambiguous, when emotional sensitivity is limited, and when operational risk is contained. When any of these conditions is absent, escalation to a human agent is the quality-preserving choice.

The second research question asked:

"How do autonomous AI agents perform across different customer service scenarios, and what service quality outcomes emerge across interaction types of varying complexity?"

The evaluation across 15 test cases spanning four complexity tiers yielded a mean SERVQUAL score of 4.83 out of 5.00 in the refined simulation round, compared to 4.59 in the baseline round. All fifteen cases were handled correctly in terms of escalation decision, with no erroneous autonomous resolution recorded in the second round. These results indicate that multi-agent architectures can deliver consistent, high-quality service outcomes across a broad range of routine and moderately complex interactions. Performance was strongest in Tier 1 and Tier 2 cases involving information retrieval, policy application, and transactional updates, and weakest in Tier 4 cases combining emotional distress with operational urgency, where empathy scores constrained the overall evaluation. This pattern is consistent with findings in prior chatbot research, which similarly reports that performance degrades as emotional complexity and contextual ambiguity increase (Li & Zhang, 2023). The improvement observed between the two simulation rounds further demonstrates that iterative prompt refinement and orchestration adjustments are a practically viable means of enhancing agent performance without altering the underlying model infrastructure.

The third research question asked:

"What is the current potential of autonomous AI agents in customer service, and what architectural and operational conditions are necessary for their controlled and incremental deployment?"

The interview findings indicated that the potential for autonomous handling is already meaningful in high-volume and clearly structured service interactions, while full autonomy across all interaction types remains contingent on further technological maturity and organisational readiness. This assessment is directly supported by the empirical results of this study, which demonstrate that controlled, incremental autonomy, bounded by confidence thresholds, escalation rules, and verification mechanisms, can already deliver substantive value in current deployments without requiring full autonomy. A finding of particular practical significance is that autonomous agent behaviour is structurally dependent on the existence of well-defined service processes. When service policies and escalation criteria are not explicitly documented and formalised, autonomous agents cannot

consistently apply them. Autonomy, therefore, does not emerge from the agent alone but from the clarity and completeness of the organisational processes that underpin its operation. From an architectural perspective, the study demonstrates that a shared CRM data layer, combined with a coordinated multi-agent design in which specialised agents handle communication, orchestration, and resolution, enables consistent service quality across heterogeneous case types without requiring uniform autonomy across all interaction channels.

Taken together, the findings of this thesis make contributions at three levels. Theoretically, the study extends prior research on agentic AI in customer service by showing that autonomy should be understood as a bounded, context-dependent capability rather than a uniform end state, and by demonstrating that the SERVQUAL framework can serve as a structured evaluation instrument for multi-agent systems. In practice, the results indicate that organisations do not need to pursue full autonomy to achieve meaningful gains in efficiency and consistency. Instead, carefully governed and incremental deployment can already reduce human workload in routine and transactional cases while preserving accountability in complex or sensitive situations. More broadly, the findings suggest that readiness for AI-enabled service delivery must be assessed not only at the technology level, but also at the process level.

The study is subject to several limitations that constrain the generalisability of its findings. The evaluation was conducted in a simulated environment using controlled test cases rather than live customer interactions, which limits how well the results reflect the variability and unpredictability of real-world service conditions. The assessment relied on a single evaluator applying the SERVQUAL framework, thereby introducing subjectivity and precluding the calculation of inter-rater reliability. The artefact is implemented exclusively within the Microsoft technology stack, and findings may not transfer directly to other CRM platforms or architectural configurations. The empirical evaluation was conducted in the travel service domain, which has specific regulatory characteristics, particularly under EU Regulation 261/2004, that may not generalise to other service sectors.

Future academic work could extend the present study, particularly regarding the empathy dimension, as emotionally charged interactions remain the most challenging area for autonomous agents. In addition, the role of service process formalisation warrants further investigation, as the findings indicate that effective autonomy depends not only on

technological capability but also on the existence of structured service processes. For the commissioning company, however, the present results already provide a sufficiently grounded basis for future solution development and customer advisory work. Meaningful progress does not require full autonomy across all interaction types, but can already be achieved through controlled, incremental deployment in scenarios where service processes, escalation criteria, and data structures are sufficiently well-defined. In this sense, the study contributes to both future academic refinement and the practical, responsible deployment of agent-based customer service capabilities within an existing enterprise CRM environment.

References

Abou Ali, M., Dornaika, F. & Charafeddine, J. 2025. Agentic AI: a comprehensive survey of architectures, applications, and future directions. The Artificial intelligence review. 59 (1), 11. Available at 10.1007/s10462-025-11422-4

Acharya, D. B., Kuppan, K. & Divya, B. 2025. Agentic AI: Autonomous Intelligence for Complex Goals-A Comprehensive Survey. IEEE access. 13, 18912–18936. Available at 10.1109/ACCESS.2025.3532853

Ahammad, S. H., Kalangi, R. R., Nagendram, S., Inthiyaz, S., Priya, P. P., Faragallah, O. S., Mohammad, A., Eid, M.M.A. & Rashed, A.N.Z. 2024. Improved neural machine translation using Natural Language Processing (NLP). Multimedia Tools and Applications. 83 (13), 39335–39348. Available at 10.1007/s11042-023-17207-7

Ahmad, H. & Allan, M. 2014. Customer Satisfaction. Sintok: Universiti Utara Malaysia.

Ali, H. M. U., Farooq, Q., Imran, A. & El Hindi, K. 2025. A systematic literature review on sentiment analysis techniques, challenges, and future trends: A systematic literature review on sentiment analysis techniques. Knowledge and information systems. 67 (5), 3967–4034. Available at 10.1007/s10115-025-02365-x

Alqurni, J. S., Alsmadi, M. K., Alfagham, H., Alzoubi, S., Ihab, S., Sameh, A., AbdElminaam, D.S & Khalaf, O.I. 2025. Streamlining Video Summarization with NLP: Techniques, Implementation, and Future Direction. SN computer science. 6 (2), 110. Available at 10.1007/s42979-024-03591-w

Amaratunga, T. 2023. Understanding Large Language Models: Learning Their Underlying Concepts and Technologies. 1. Berkeley, CA: Apress. Available at 10.1007/979-8-8688-0017-7

Arli, D., van Esch, P. & Weaven, S. 2024. The Impact of SERVQUAL on Consumers' Satisfaction, Loyalty, and Intention to Use Online Food Delivery Services. Journal of promotion management. 30 (7), 1159–1188. Available at 10.1080/10496491.2024.2372858

Benaddi, L., Ouaddi, C., Jakimi, A. & Ouchao, B. 2024. A Systematic Review of Chatbots: Classification, Development, and Their Impact on Tourism. IEEE access. 12, 78799–78810. Available at 10.1109/ACCESS.2024.3408108

Bishop, C. M. & Bishop, H. 2024. Deep learning: foundations and concepts. 1. Cham: Springer. Available at 10.1007/978-3-031-45468-4

Biswas, A. & Talukdar, W. 2025. Building Agentic AI Systems. Packt Publishing.

Blümel, J. H., Zaki, M. & Bohné, T. 2024. Personal touch in digital customer service: a conceptual framework of relational personalization for conversational AI. Journal of service theory and practice. 34 (1), 33–65. Available at 10.1108/JSTP-03-2023-0098

Brohi, S., Mastoi, Q., Jhanjhi, N. Z. & Pillai, T. R. 2025. A Research Landscape of Agentic AI and Large Language Models: Applications, Challenges and Future Directions. Algorithms. 18 (8), 499. Available at 10.3390/a18080499

Chatterjee, S., Ghatak, A., Nikte, R., Gupta, S. & Kumar, A. 2023. Measuring SERVQUAL dimensions and their importance for customer-satisfaction using online reviews: a text mining approach. Journal of enterprise information management. 36 (1), 22–44. Available at 10.1108/JEIM-06-2021-0252

Costa, A. C. F., Capelo Neto, F., Espuny, M., Rocha, A. B. T. d. & Oliveira, O. J. d. 2024. Digitalization of customer service in small and medium-sized enterprises: drivers for the development and improvement. International journal of entrepreneurial behaviour & research. 30 (2), 305–341. Available at 10.1108/IJEBR-10-2022-0953

Dalton, G. 2024. Artificial Intelligence: Background, Risks and Policies. New York, NY: Nova Science Publishers, Inc.

De Keyser, A., Verleye, K., Lemon, K. N., Keiningham, T. L. & Klaus, P. 2020. Moving the Customer Experience Field Forward: Introducing the Touchpoints, Context, Qualities (TCQ) Nomenclature. Journal of service research: JSR. 23 (4), 433–455. Available at 10.1177/1094670520928390

De Nicola, A., Formica, A., Mele, I., Missikoff, M. & Taglino, F. 2024. A comparative study of LLMs and NLP approaches for supporting business process analysis. Enterprise information systems. 18 (10) Available at 10.1080/17517575.2024.2415578

Dennis, A. R., Lakhiwal, A. & Sachdeva, A. 2023. AI Agents as Team Members: Effects on Satisfaction, Conflict, Trustworthiness, and Willingness to Work With. Journal of Management Information Systems. 40 (2), 307–337. Available at 10.1080/07421222.2023.2196773

Dwivedi, Y. K., Helal, M. Y. I., Elgendy, I. A., Alahmad, R., Walton, P., Suh, A., Singh, V. & Jeon, I. 2025. Agentic AI Systems: What It Is and Isn't. Global business and organizational excellence. Available at 10.1002/joe.70018

El-Manstrly, D. 2016. Enhancing customer loyalty: critical switching cost factors. International journal of service industry management. 27 (2), 144–169. Available at 10.1108/JOSM-09-2015-0291

Feng, Z. 2023. Formal analysis for natural language processing: a handbook. 1. Singapore: Springer Nature Singapore. Available at 10.1007/978-981-16-5172-4

Frigau, L., Contu, G., Ortu, M. & Carta, A. 2024. Gauging Airbnb review sentiments and critical key-topics by small area estimation. Statistical methods & applications. 33 (4), 1145–1170. Available at 10.1007/s10260-024-00764-y

Gerea, C., Gonzalez-Lopez, F. & Herskovic, V. 2021. Omnichannel Customer Experience and Management: An Integrative Review and Research Agenda. Sustainability. 13 (5) Available at 10.3390/su13052824

Gowri, R., Saivenkat, R., V, S., R, T. M., Karthikeyan, P. & M, Y. P. 2025. NLP in Customer Care Centers to Enhance Customer Service.

Henriques, T. A. & O'Neill, H. 2023. Design science research with focus groups – a pragmatic meta-model. International journal of managing projects in business. 16 (1), 119–140. Available at 10.1108/IJMPB-01-2020-0015

Hevner, A. & Chatterjee, S. 2010. Design Research in Information Systems Theory and Practice. Boston, MA: Springer US. Available at 10.1007/978-1-4419-5653-8

Hodge, S. K., Honeycutt, E. D. & Shipley, D. 2019. Toward an Improved Understanding of Online Customer Service Delivery to Millennials. Services marketing quarterly. 40 (1), 33–47. Available at 10.1080/15332969.2019.1587865

Horvat, S., Horvat, T., Havaš, L. & Crčič, D. 2025. AI-Powered Chatbots for Enhancing Accessibility in Higher Education - Design and Implementation. Elektrotehniski Vestnik. 92 (1), 61–67.

Hsu, C. & Lin, J. C. 2023. Understanding the user satisfaction and loyalty of customer service chatbots. Journal of retailing and consumer services. 71, 103211. Available at 10.1016/j.jretconser.2022.103211

Huang, D., Markovitch, D. G. & Stough, R. A. 2024. Can chatbot customer service match human service agents on customer satisfaction? An investigation in the role of trust. Journal of retailing and consumer services. 76, 103600. Available at 10.1016/j.jretconser.2023.103600

Huang, K. 2025. Agentic AI: Theories and Practices. 1. Cham: Springer Nature Switzerland. Available at 10.1007/978-3-031-90026-6

Hung, H., Lee, N. & Hu, Y. 2025. Unlocking Service Provider Excellence: Expanding the Touchpoints, Context, Qualities Framework. Journal of service research: JSR. 28 (2), 336–354. Available at 10.1177/10946705231218406

Jang, I., Kim, J. & Park, I. 2025. When AI takes your call: Investigating the impact of conversational styles on customer service agents' affective delivery. Telematics and Informatics. 99, 102277. Available at 10.1016/j.tele.2025.102277

Joshi, A., Dabre, R., Kanojia, D., Li, Z., Zhan, H., Haffari, G. & Dippold, D. 2025. Natural Language Processing for Dialects of a Language: A Survey. ACM computing surveys. 57 (6), 1–37. Available at 10.1145/3712060

Juglan, K. C., Sharma, B., Gehlot, A., Singh, S. P., Hussein, A. & Alazzam, M. B. 2023. Exploring the Effectiveness of Natural Language Processing in Customer Service.

Just, J. 2024. Natural language processing for innovation search – Reviewing an emerging non-human innovation intermediary. Technovation. 129, 102883. Available at 10.1016/j.technovation.2023.102883

Kadre, S., Kadre, S. & Dey, S. 2025. Mastering Text Analytics: A Hands-on Guide to NLP Using Python. 1. Berkeley, CA: Apress. Available at 10.1007/979-8-8688-1582-9

Khanna, A., Sarasan, S. & Sunil, M. P. 2025. Navigating the AI revolution in customer service: a bibliometric analysis. Spanish journal of marketing-ESIC. Available at 10.1108/SJME-08-2024-0228

Khurana, D., Koli, A., Khatter, K. & Singh, S. 2023. Natural language processing: state of the art, current trends and challenges. Multimedia Tools and Applications. 82 (3), 3713–3744. Available at 10.1007/s11042-022-13428-4

Kim, H. & Fung So, K. K. 2024. Customer touchpoints: Conceptualization, index development, and validation. Tourism management (1982). 103, 104881. Available at 10.1016/j.tourman.2023.104881

Lee, R. S. T. 2024. Natural Language Processing: A Textbook with Python Implementation. 1. Singapore: Springer Nature Singapore. Available at 10.1007/978-981-99-1999-4

Li, C. & Zhang, J. 2023. Chatbots or me? Consumers' switching between human agents and conversational agents. Journal of retailing and consumer services. 72, 103264. Available at 10.1016/j.jretconser.2023.103264

Li, X., Li, W., Lu, L. & Fan, X. 2025. An Artificial Intelligence (AI) Agent-Based Question-and-Answer (Q&A) System for Tibetan Jiu Chess. ICGA journal. 47 (2), 95–111. Available at 10.1177/13896911251343450

Liu, H., Tsang, S., Wood, A. & Tong, X. 2025. Longitudinal Sentiment Analysis with Conversation Textual Data: Longitudinal Sentiment Analysis. Fudan journal of the humanities and social sciences. 18 (1), 193–214. Available at 10.1007/s40647-024-00417-0

Mamakou, X. J., Zaharias, P. & Milesi, M. 2024. Measuring customer satisfaction in electronic commerce: the impact of e-service quality and user experience. The International journal of quality & reliability management. 41 (3), 915–943. Available at 10.1108/IJQRM-07-2021-0215

Mariprasath, T., Marco, R. & Cheepati, K. R. 2025. Practical Guide to Machine Learning, NLP, and Generative AI: Libraries, Algorithms, and Applications. River Publishers. Available at 10.1201/9781003563945

MarketsandMarkets. 2025. Artificial intelligence (AI) market. [Online] [Cited at: 31.01.2026] Available at: https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-market-74851580.html

McLaughlin, L. 2025. Agentic AI: Nine Essential Questions. MIT Sloan management review., 2–4.

Metz, D., Ilieş, L. & Nistor, R. L. 2020. The impact of organizational culture on customer service effectiveness from a sustainability perspective. Sustainability (Switzerland). 12 (15) Available at 10.3390/SU12156240

Microsoft. 2026a. What is Dynamics 365? [Online] [Cited at: 31.01.2026] Available at: https://www.microsoft.com/en-us/dynamics-365/what-is-dynamics365

Microsoft. 2026b. Power Platform. [Online] [Cited at: 31.01.2026] Available at: https://www.microsoft.com/en-us/power-platform

Microsoft. 2026c. Azure. [Online] [Cited at: 31.01.2026] Available at: https://azure.microsoft.com/en-us

Mishra, A. D., Singh, P., Yadav, D. K. & Singh, Y. B. 2024. A Handbook of Computational Linguistics: Artificial Intelligence in Natural Language Processing: Volume 2. Singapore: Bentham Science Publishers. Available at 10.2174/9789815238488124020l

Mohamed, Y. A., Khanan, A., Bashir, M., Mohamed, A. H. H. M., Adiel, M. A. E. & Elsadig, M. A. 2024. The Impact of Artificial Intelligence on Language Translation: A Review. IEEE access. 12, 25553–25579. Available at 10.1109/ACCESS.2024.3366802

Murrar, S., Alhaj, F. M., Almasalha, F. & Qutqut, M. H. 2025. Transformers for Domain-Specific Text Classification: A Case Study in the Banking Sector. IEEE access. 13, 115063–115078. Available at 10.1109/ACCESS.2025.3585164

Nirala, K. K., Singh, N. K. & Purani, V. S. 2022. A survey on providing customer and public administration based services using AI: chatbot. Multimedia Tools and Applications. 81 (16), 22215–22246. Available at 10.1007/s11042-021-11458-y

Oh, D. & Elango, B. 2024. Quantitative analysis of the publication trends in four decades of SERVQUAL research: a bibliometric approach. Total quality management & business excellence. 35 (1-2), 297–319. Available at 10.1080/14783363.2023.2293877

Olujimi, P. A. & Ade-Ibijola, A. 2023. NLP techniques for automating responses to customer queries: a systematic review. Discover Artificial Intelligence. 3 (1), 20–19. Available at 10.1007/s44163-023-00065-5

Özkan, P., Süer, S., Keser, İ. K. & Kocakoç, İ. D. 2020. The effect of service quality and customer satisfaction on customer loyalty: The mediation of perceived value of services, corporate image, and corporate reputation. International journal of bank marketing. 38 (2), 384–405. Available at 10.1108/IJBM-03-2019-0096

Pakray, P., Gelbukh, A. & Bandyopadhyay, S. 2025. Natural language processing applications for low-resource languages. Natural Language Processing. 31 (2), 183–197. Available at 10.1017/nlp.2024.33

Parasuraman, A., Zeithaml, V. A. & Berry, L. L. 1985. A Conceptual Model of Service Quality and Its Implications for Future Research. Journal of Marketing. 49 (4), 41–50. Available at 10.1177/002224298504900403

Park, S., Yi, Y. & Lee, Y. 2021. Heterogeneous dimensions of SERVQUAL. Total quality management & business excellence. 32 (1-2), 92–118. Available at 10.1080/14783363.2018.1531700

Pateria, S., Subagdja, B., Tan, A. & Quek, C. 2022. Hierarchical Reinforcement Learning: A Comprehensive Survey. ACM computing surveys. 54 (5), 1–35. Available at 10.1145/3453160

Pathak, A. & Bansal, V. 2025. Factors affecting customer adoption of AI digital agents in service operations: an assessment of relative importance. Behaviour & information technology. , 1–23. Available at 10.1080/0144929X.2025.2490672

Patti, C. H., van Dessel, M. M. & Hartley, S. W. 2020. Reimagining customer service through journey mapping and measurement. European journal of marketing. 54 (10), 2387–2417. Available at 10.1108/EJM-07-2019-0556

Patwardhan, N., Marrone, S. & Sansone, C. 2023. Transformers in the Real World: A Survey on NLP Applications. Information (Basel). 14 (4), 242. Available at 10.3390/info14040242

Peffers, K., Tuunanen, T., Rothenberger, M. A. & Chatterjee, S. 2007. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems. 24 (3), 45–77. Available at 10.2753/MIS0742-1222240302

Pires, P. B., Perestrelo, B. M. & Santos, J. D. 2025. Unpacking Customer Experience in Online Shopping: Effects on Satisfaction and Loyalty. Journal of theoretical and applied electronic commerce research. 20 (3), 245. Available at 10.3390/jtaer20030245

Qu, S. Q. & Dumay, J. 2011. The qualitative research interview. Qualitative research in accounting and management. 8 (3), 238–264. Available at 10.1108/11766091111162070

Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mim, M. M. J., Ahmad, J., Ali, M.E. & Azam, S. 2024. A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges. IEEE access. 12, 1. Available at 10.1109/ACCESS.2024.3365742

Ranjan, S., Chembachere, D. & Lobo, L. 2025. Agentic AI in Enterprise. 1. Berkeley, CA: Apress. Available at 10.1007/979-8-8688-1542-3

Ribeiro, H., Barbosa, B., Moreira, A. C. & Rodrigues, R. 2024. Customer Experience, Loyalty, and Churn in Bundled Telecommunications Services. SAGE open. 14 (2) Available at 10.1177/21582440241245191

Saneva, D. & Chortoseva, S. 2020. Service Quality, Customer Satisfaction and Customer Loyalty: Testing a Structural Equation Model. Calitatea. 21 (179), 124–128.

Sapkota, R., Roumeliotis, K. I. & Karkee, M. 2026. AI Agents vs. Agentic AI: A Conceptual taxonomy, applications and challenges. Information fusion. 126, 103599. Available at 10.1016/j.inffus.2025.103599

Schweiger, S., Oed, S., Kemmer, J., de Geus, M., Buecheler, T., West, S. et al. 2024. Strengthening Customer Loyalty and Enhancing Resilience via Self-Service. In Buecheler, T., Meierhofer, J. & West, S. (eds.). Smart Services Summit. Switzerland: Springer. 273–278. Available at 10.1007/978-3-031-60313-6_21

Seidel, P., Preßler, W., Zielinski, S., Erfurth, C., Fahrnberger, G. & Eichler, G. 2025. Automated Text Classification in Maturity Models Using Transformer Architectures: An Encoder-Based Approach. In Zielinski, S., Erfurth, C., Eichler, G. & Fahrnberger, G. (eds.).

Innovations for Community Services. Cham: Springer Nature Switzerland. 51–60. Available at 10.1007/978-3-031-94263-1_4

Shah, S., Ghomeshi, H., Vakaj, E., Cooper, E. & Fouad, S. 2023. A review of natural language processing in contact centre automation. Pattern analysis and applications: PAA. 26 (3), 823–846. Available at 10.1007/s10044-023-01182-8

Shaik, R. 2025. Enhancing text generation in joint Nlg/Nlu learning through curriculum learning, semi-supervised training, and advanced optimization techniques. Multimedia Tools and Applications. 84 (30), 37377–37400. Available at 10.1007/s11042-025-20665-w

Sharif, K. & Sidi Lemine, M. 2024. Customer service quality, emotional brand attachment and customer citizenship behaviors: findings from an emerging higher education market. Journal of marketing for higher education. 34 (1), 18–43. Available at 10.1080/08841241.2021.1949659

Sheth, J. N., Jain, V. & Ambika, A. 2023. The growing importance of customer-centric support services for improving customer experience. Journal of business research. 164, 113943. Available at 10.1016/j.jbusres.2023.113943

Siebert, A., Gopaldas, A., Lindridge, A. & Simo˜es, C. 2020. Customer Experience Journeys: Loyalty Loops Versus Involvement Spirals. Journal of Marketing. 84 (4), 45–66. Available at 10.1177/0022242920920262

Slack, N. J. & Singh, G. 2020. The effect of service quality on customer satisfaction and loyalty and the mediating role of customer satisfaction: Supermarkets in Fiji. TQM journal. 32 (3), 543–558. Available at 10.1108/TQM-07-2019-0187

Sood, P., Tanwar, H., Singh, J., Ruhela, A. K., Gupta, N. & Kumar, R. 2024. Revolutionizing Customer Service: An AI-powered Chatbot Approach using Advanced NLP Techniques.

Taherdoost, H. & Madanchian, M. 2023. Artificial Intelligence and Sentiment Analysis: A Review in Competitive Research. Computers (Basel). 12 (2), 37. Available at 10.3390/computers12020037

Tedjokusumo, C. & Murhadi, W.R. 2023. Customer satisfaction as a mediator between service quality and customer loyalty: a case study of Bank Central Asia. Jurnal Siasat Bisnis. 27 (2), 156–170. Available at 10.20885/jsb.vol27.iss2.art3

Tse, D. K. & Wilton, P. C. 1988. Models of Consumer Satisfaction Formation: An Extension. Journal of Marketing Research. 25 (2), 204. Available at 10.2307/3172652

vom Brocke, J., Hevner, A. & Maedche, A. 2020. Design Science Research. Cases. 1. Cham: Springer International Publishing AG. Available at 10.1007/978-3-030-46781-4

Wang, L., Huang, N., Hong, Y., Liu, L., Guo, X. & Chen, G. 2023. Voice-based AI in call center customer service: A natural field experiment. Production and operations management. 32 (4), 1002–1018. Available at 10.1111/poms.13953

Win, C. C. M. T., Dodanwala, T. C. & Santoso, D. S. 2024. Synthesizing customer satisfaction and loyalty through contractors' service quality and brand image. Engineering, construction, and architectural management. Available at 10.1108/ECAM-06-2023-0633

Yatawara, K., Sampath, T., Kalupahana, P. L., Rathnayake, S., Jayasuriya, N. & Rathnayake, N. 2025. A Systematic Review on Consumer Adoption of AI-driven Chatbots. Vision (New Delhi, India). Available at 10.1177/09722629251332349

Yu, M., Xu, J. & Tang, J. 2024. Managing customer contact centers with delay announcements and automated service. IISE transactions. ahead-of-print (-), 1–13. Available at 10.1080/24725854.2023.2183532

Appendix 1. Interview topics, questions and goals

Topic	Question	Goal of the question
Current state of the customer service	How would you characterise the current state of your organisation's customer service and its areas for improvement?	To elicit the respondent's perspective on the present state and identify key areas requiring development.
	Which factors most significantly influence customer service quality and customer satisfaction?	To determine the primary drivers of service quality and customer satisfaction.
	What challenges are encountered across your existing customer service channels?	To gain insight into the complexities and limitations of multi-channel customer service delivery.
AI and Customer Service	In which tasks do you anticipate AI could provide added value within current customer service processes?	To explore potential AI applications and their perceived benefits in customer service workflows.
	Are you currently utilising AI or automation in customer service operations? What experiences have emerged from their use?	To assess the extent of AI adoption and gather experiential insights regarding its implementation.
	How have customers responded to AI-driven solutions compared to traditional human-mediated service?	To evaluate customer perceptions and acceptance of AI-based interventions relative to conventional service models.
Future of Customer Service	How do you envision the role of AI and autonomous agents in the future of customer service?	To capture forward-looking perspectives on the strategic integration of AI and autonomous systems in service delivery.
	What potential risks or challenges do you foresee in scaling AI adoption in customer service, and how might these be mitigated?	To identify anticipated risks and formulate strategies for mitigating challenges associated with broader AI deployment.

Appendix 2. SERVQUAL evaluation rubric

Dimension	1 — Poor	2 — Below average	3 —Adequate	4 — Good	5 — Excellent
Reliability	Response contains factual errors, incorrect policy application or misidentifies the customer's issue	Mostly correct but contains one significant factual error or policy misinterpretation	Correct on core facts but incomplete — omits relevant policy details or conditions	Factually accurate and policy-compliant with minor omissions	Fully accurate, correctly applies all relevant policies, no errors or omissions
Responsiveness	Ignores one or more of the customer's explicit requests	Addresses some requests but leaves key questions unanswered	Addresses all explicit requests but provides no proactive guidance	Addresses all requests and offers at least one proactive next step	Addresses all requests, provides clear next steps, specifies timelines and ownership
Assurance	Response is vague, contradictory or likely to reduce customer confidence	Limited assurance — lacks specifics or hedges excessively	Reasonable assurance but does not fully resolve uncertainty	Clear and confident response that addresses the customer's concerns	Highly confident, specific and credible response that fully resolves uncertainty and builds trust
Empathy	No acknowledgement of the customer's situation or emotional state	Superficial apology without genuine recognition of impact	Acknowledges the situation but does not adapt tone or content to context	Recognises emotional context and adjusts tone appropriately	Demonstrates genuine understanding of the customer's situation, adapts language and tone, prioritises the customer's experience
Tangibility	Disorganised, difficult to follow, missing key information	Some structure but inconsistent formatting or unclear sections	Adequately structured with clear sections but lacks polish	Well-structured, professional and easy to navigate	Exceptionally clear, well-formatted, professional and complete — all relevant information presented logically

Appendix 3. Voice channel intake agent interaction – T06

AGENT: Hei, Travel Agency Oy:n asiakaspalvelun keinoälyasiakaspalvelija. Miten voin auttaa?

CUSTOMER: Hi, I need to cancel my trip.

AGENT: I understand. Could you tell me more?

CUSTOMER: Yes, I have a booking to Barcelona, booking number BK-2026-0210. The flight is next week, on March 12th. But I got seriously ill last week, and my doctor has forbidden me from travelling.

AGENT: I'm sorry to hear that. Let me check your information in the system.

AGENT: I found your details. Oskari Alamäki, is that correct?

CUSTOMER: Kyllä, se olen minä. Voimmeko jatkaa suomeksi?

AGENT: Toki. Eli kyseessä on Barcelona-matka, varaus BK-2026-0210, lento 12. maaliskuuta. Milloin sairastuit?

CUSTOMER: Viime viikolla, olisko ollut tiistaina tai keskiviikkona. Kävin lääkärissä perjantaina ja sain todistuksen että en saa matkustaa. Mulla on se todistus tallessa.

AGENT: Selvä. Mitä muuta haluaisit vielä kertoa?

CUSTOMER: No siis, mulla on se Plus-vakuutus. Saanko mä ne peruutusmaksut takaisin kun mulla on lääkärintodistus? Ja mitä mun pitää tehdä, pitääkö se todistus lähettää johonkin vai? Tää on vähän kallista touhua niin ois kiva tietää et mitä tapahtuu.

AGENT: Ymmärrän huolesi. Kirjaan nämä tiedot palvelupyyntöön. Hetki.

CUSTOMER: Joo.

AGENT: Olen luonut palvelupyynnön asiastanne. Asiakaspalvelutiimimme käsittelee tämän ja ottaa teihin yhteyttä.

CUSTOMER: Okei, mut siis voitko sä sanoa ees jotain et miten tää menee? Pitääkö mun itse peruuttaa jostain vai hoitaako se teidän tiimi?

AGENT: Tiimimme ohjaa sinut oikean prosessin läpi ja kertoo mitä dokumentteja tarvitaan. He ottavat teihin yhteyttä.

CUSTOMER: Joo, okei. Kauanko se kestää? Lento lähtee viikon päästä niin...

AGENT: Tiimimme käsittelee asian mahdollisimman pian. Olemme teihin yhteydessä asian selvittämiseksi.

CUSTOMER: No hyvä. Toivottavasti hoituu. Pitääkö mun vielä tehdä jotain?

AGENT: Tiimimme on teihin yhteydessä. Onko vielä jotain lisättävää?

CUSTOMER: Ei kai. Odotellaan sitä yhteydenottoa sitten.

AGENT: Kiitos yhteydenotosta, hyvää päivänjatkoa.

CUSTOMER: Joo, kiitos. Moi.

Appendix 4. Customer service agent interaction – T06

AGENT: Hyvä Oskari, Pahoittelemme kuulla terveydentilastanne ja toivomme teille pikaista paranemista. Ymmärrämme, että tilanne on huolestuttava, ja haluamme auttaa teitä parhaalla mahdollisella tavalla. Varaamanne lennot (AY3421, lähtö 12. maaliskuuta 2026, ja AY3422, paluu 27. maaliskuuta 2026) ovat tällä hetkellä aktiivisia, mutta Plus-vakuutuksenne ehtojen mukaisesti voitte peruuttaa matkanne lääkärintodistuksella vahvistetun sairauden vuoksi. Pyydämme teitä toimittamaan lääkärintodistuksen esimerkiksi PDF- tai kuvamuodossa, jotta voimme käsitellä peruutusmaksujen palautuksen vakuutusehtojen mukaisesti. Kun todistus on vastaanotettu ja tarkistettu, peruutusmaksu voidaan palauttaa kokonaisuudessaan. Voitte lähettää lääkärintodistuksen suoraan tähän viestiin vastauksena. Jos tarvitsette lisäapua tai kysymyksiä herää, olemme täällä auttamassa. Toivotamme teille voimia ja pikaista toipumista! Ystävällisin terveisin, Travel Agency Oy.

Customer: Kiitos nopeasta vastauksesta. Lääkärintodistus liitteenä.
Toivottavasti tämä riittää peruutuksen käsittelyyn. (Viestin liitteenä 1kpl pdf-tiedosto)

AGENT: Hyvä Oskari, Kiitos toimittamastanne lääkärintodistuksesta. Olemme vahvistaneet, että varauksenne BK-2026-0210 voidaan peruuttaa Plus-vakuutuksen peruutusturvan ehtojen mukaisesti. Lääkärintodistuksenne, joka osoittaa matkustamiskieltonne ajalle 2.3.–15.3.2026 vakavan keuhkokuumeen vuoksi, täyttää kaikki vaaditut ehdot. Olemme aloittaneet peruutusmaksujen palautusprosessin, eikä teille jää maksuvelvoitteita tästä varauksesta. Palautus suoritetaan vakuutusehtojen mukaisesti mahdollisimman pian. Toivomme teille pikaista paranemista ja kaikkea hyvää! Ystävällisin terveisin, Travel Agency Oy