Designing a conversational AI agent: Framework combining customer experience management, personalization, and AI in service techniques

Jan H. Blümel
University of Cambridge
jhb65@cam.ac.uk

Gautam Jha
University of Cambridge
gj298@cam.ac.uk

Abstract

Conversational AI agents are fundamentally changing how firms are delivering service to their customers. The rapid advancement of technology and ready tools means that deploying a conversational AI agent has become far simpler than ever imagined. However, customers remain unsatisfied with their experience and firms are unable to demonstrate value of conversational AI agents. Drawing on the theoretical notions of customer experience (CX) management, personalization, and AI in service, we develop a framework to design conversational AI agents. We propose a six-stage iterative design for conversational AI agents that begins with sensing customer intent, adapting to journey context, assigning tone to the conversation, delegating to humans, orchestrating processes to service requests and training of AI agents to adaptively improve and loopback into prior design stages. Additionally, we recognize that firms need to holistically qualify and allocate service requests to such conversational AI agents based on the firm’s purpose and CX strategy.

Keywords: Conversational AI agents, Customer experience management, Service personalization, Relational personalization

1. Introduction

Creating a pleasant customer experience (CX) has widely been acknowledged as a critical factor for achieving and sustaining competitive advantage (Lemon and Verhoef, 2016; McColl-Kennedy et al., 2019). CX management has been conceptualized as a firm’s capabilities to adapt customer experiences delivered by a firm leading to outcomes of sustained customer loyalty (Homburg et al., 2017). Firms need to constantly adjust and expand these capabilities to cope with permanently changing customer behavior caused by external effects, such as the Covid-19 pandemic and technological advancements in AI.

Autonomous service systems, such as service robots and chatbots, play an ever-increasing role in services and have lastingly impacted the way customers and firms interact with each other (Hollebeek et al., 2021).

One thriving area is conversational AI, a technology users can talk to like chatbots or virtual agents. A conversational AI agent imitates human interactions, recognizes speech and text inputs, and translates their meaning across various languages (IBM, 2021). It is one of the AI domains with the highest number of patents and currently represents the top use of AI in enterprises (Comes et al., 2021). The consumer retail spend on conversational AI will increase on average over 400% rising from $2.8 billion in 2019 to $142 billion by 2024 (Yuen, 2022).

Conversations in customer service can have a decisive impact on the overall customer service experience (Packard and Berger, 2021). The language used in a conversation can convey affective components such as emotions (Hennig-Thurau et al., 2006). It is therefore particularly important in customer service situations, where the customers are emotionally vulnerable (Groth et al., 2019). If deployed conversational AI agents are poorly managed and lacking technological capabilities customers are left unsatisfied and with a bad customer service experience.

Therefore, it is paramount to have a conversational AI agent that goes beyond AI that can chat. Instead, conversational AI agents that can truly meet customer expectations and benefit firms deliver better service. From a managerial perspective, the challenge is to design, implement and assess such a conversational AI agent that enhances the customer service experience. Although some approaches in the literature help to select the technical capabilities of a chatbot (Adamopoulou and Moussiades, 2020), or specifically classify individual avatars (Miao et al., 2022), to the best of the authors’ knowledge there are currently no approaches in the literature that help managers design and assess conversational AI agents for better customer service experiences.
This paper aims to address this challenge by combining the literature streams of technological conversational AI and managerial customer experience. We take a perspective where personalization is one way to improve the CX (as shown in Hänninen et al., 2019; Rieger et al., 2021; Sujata et al., 2019; Tyrväinen et al., 2020) and develop a framework to design and assess personalized conversational agents based on customer intent and context. We identify contextual factors effecting the customer intent and transform these into requirements for conversational AI agents.

The contribution of this paper is three-fold. First, we combine the academic literature of customer experience management with the technological advancements of conversational AI by merging requirements of successfully managing customer service experience with capabilities of conversational AI. Second, we provide CX managers with a framework to assess, and design personalized conversational agents based on customer intent and context. Third, we apply the framework to assess existing conversational AI agents deployed in various industry sectors and provide the reader with a typology of existing approaches.

2. Theoretical background

2.1 Managing CX delivered by conversational AI agents

Research on conversational AI agents needs to shift from an anthropomorphic human vs AI perspective towards a human-AI collaborative perspective (Blut et al., 2021; de Keyser and Kunz, 2022). We concur because advancements in natural language processing (NLP) mean resemblance with human agents is within the realm of possibility and that businesses should now assess conversational AI agents using a combined human-AI view of enabling customer experience. The introduction of deep learning-based architectures and improved hardware capabilities have allowed for processing large amount of data and creating pre-trained language models such as GPT-3, OPT, or GATO. These models, build on transformer architectures, can generate short stories, songs (Heaven, 2020) or even caption images, chat, and stack blocks with a real robot arm at the same time (Reed et al., 2022). Critical tasks of conversational AI have been advanced with the help of transformer-based architectures (Ni et al., 2021).

Managing CX delivered through conversational AI agents within the service ecosystem can benefit by applying the journey context across pre-purchase, purchase, and post-purchase journeys (Grewal et al., 2009; Lemon and Verhoeef, 2016). Further, such journeys designed in conversational AI agents need to factor experiential dimensions like sensory, cognitive, emotional, behavioral, and social (Lemon and Verhoeef, 2016) alongside market and environmental contextual factors (de Keyser et al., 2020).

The firm’s capabilities to manage CX are constantly evolving across the building blocks of CX (de Keyser et al., 2020). Whether it’s ‘touchpoints’ like a mobile app or ‘contextual’ factors like altered preferences of a customer or ‘qualities’ like sensorial factors for example, smell in an augmented reality environment. Advances in technology have meant that firms are able to rapidly deploy conversational AI agents within their CX capabilities. For example, building and deploying a basic chatbot using services like Google Dialogflow or Azure bot service is sometimes just a matter of minutes (Kapoor, 2021). However, disillusionment with conversational AI agents is a norm rather than exception as illustrated by scrapping of several chatbot services in the financial services sector (Finextra, 2018).

Managing such a touchpoint requires a multi-faceted approach to derive value from a diverse set of actors that includes customers, employees, trainers, AI agents and AI technology vendors across the service ecosystem. First, a nuanced understanding of the channel context across digital, physical, and social realms (Bolton et al., 2018) is essential to manage control of the touchpoints where conversational AI agents can be of value for a firm. For example, enabling turning the thermostat up or down related conversations by an energy provider on smart speakers like Amazon Alexa services may be more relevant to customers than such a service on a chatbot on the energy provider’s website.

Second, applying service experience design concepts (Andreasen et al., 2016; Bellos and Kavadias, 2021; Patricio et al., 2008) to tailor conversational AI agent experiences are needed to improve the service interaction as well as the overall customer journey experience. For example, the sequential nature of conversations compared to the dynamic nature of interaction on a visual medium like a website or app need to be considered in conversational service designs.

Third and finally, the strategic goals need to be aligned with the benefits that a firm can derive from conversational AI agents. The goals alignment is well articulated in the management literature on AI strategy (Kiron and Schrage, 2019). Measurement of the goals and performance of conversational AI agents is key to the holistic success for CX managers designing and deploying conversational AI agents.
Further, such metrics and scripts can be helpful in continuous training to deliver data-driven experience (Holmlund et al., 2020; McColl-Kennedy et al., 2019) using conversational AI agents.

### 2.2. Conversational AI for personalizing interactions to improve customer experience

To find out how we can best improve customer service experiences using conversational AI, we use the perspective of personalization. Personalizing the customer interaction and information provided by delivering relevance to the customer has been shown as critical factor in improving CX (Zanker et al., 2019). Firms have significantly advanced personalization efforts in areas such as advertisement and product recommendations. Spotify, Google, or Amazon are building their business models based on learning from data about customers’ history, their background, and preferences.

“Personalization tailors certain offerings by providers to customers based on certain knowledge about them, on the context in which these offerings are provided and with certain goals in mind” (Tuzhilin, 2009, p. 8). This definition of personalization fits well into the framework of conversational AI. To interact with the customer, the conversational AI agent first needs to understand what the customer is saying (natural language understanding – NLU). It then needs to connect the intent of the customer with a database, where domain-specific knowledge is stored, to produce meaningful answers. Thirdly it needs to generate a response by using natural language generation (NLG) (Bhashkar, 2019). While offerings in early personalization literature were focused on physical products or services, the era of web-technologies and introduction of chatbots or voice assistances have shifted the focus towards online content, such as websites, information searches, or user interfaces and communication (Zanker et al., 2019).

Taking personalization to the next level, Huang and Rust (2020) proposed the usage of feeling AI for learning and adapting from experience-based data. This should allow for “relationalization”, building personalized relationships. Building relationships in customer service with personalization predominantly involves adjusting the communication and conversation with the customer according to their personal preferences and needs. Conversational AI needs to build on customer knowledge and experiential data to meet the customer where they are and tailor the communication. For example, Hamilton et al. (2021) have highlighted the relevance of social others along the customer journey and shown that service agents or AI agents can act as surrogates. Social chit-chat conversations can allow to meet the customer on that level for building relationships.
Further, responding to customers’ affective states such as emotions can lead to higher customer satisfaction (Kernbach and Schutte, 2005). Theories such as emotional contagion and emotional display (Hennig-Thurau et al., 2006) show that responding appropriately to customers’ affective state can allow to evoke desired responses and therefore build relationships and improve customer service experience. Individual NLU models have been trained to detect customer’s emotion (Mundra et al., 2017) or personality traits (Herzig et al., 2016) from textual customer service conversations. These models allow for a contextualization of customer service conversations and in combination with the customer intent for a more accurate understanding of the customer’s request.

Combining personalization approaches with conversational AI to improve customer service experiences has not yet been done. Especially with the background of helping companies to design, implement and evaluate conversational AI agents.

As part of the literature review, we reviewed representative literature from the streams of CX(M), conversational AI in service, and personalization (see Table 1). As part of this, we analyzed the individual papers regarding the factors discussed for designing conversations with customers. We focused on the relevant elements discussed above: intent detection, customer journey application, context understanding and technological advancements in NLP. We could observe that the customer experience literature is predominantly speaking about the role of the customer journey for designing superior customer experiences. The technological side of the literature discusses either technological advances in conversational AI to improve understanding of intent (Ni et al., 2021) or sometimes of individual conceptual factors such as emotions (Herzig et al., 2016). However, in the mostly stand-alone studies, no reference is made to the influence of the customer journey. This is mainly also true with approaches from the personalization literature, where personalization is discussed in context (Zanker et al., 2019) and referred to understanding customer intent (Tuzhilin, 2009).

In summary, however, there is no approach that considers all relevant factors and applies this to the design of conversations. Therefore, the research question we address is how can firms design conversational AI agents to personalize and improve the customer service experience?

3. A framework for managing conversational AI agents

Drawing on the above portrayal of personalization in conversational AI agents, we developed a framework for designing conversational AI agents to manage customer experience. The stages apply contextual factors based on the journey and experiential context by combining the literature across CX(M), personalization and conversational AI agents. The framework to design and deploy conversational AI agents comprises of six dynamic stages, namely, sense, adapt, assign, delegate, orchestrate and train. Figure 1 depicts the framework along with the contextual factors that need to be continuously considered to match customer intent at each of the design stages of conversational AI agents.

To refine the framework, we assessed a set of nine conversational AI agents that have been deployed by established firms like Amazon Alexa and Vodafone Tobi across sectors (see Table 2). We gathered the data by participating as prospective customers using a diverse set of pre-purchase and purchase intents. Further, for firms where we held personal accounts, we assessed typical purchase (upsell) and post purchase intents by making hypothetical service requests. Although the sampling is limited, such an exercise allowed us to apply and test the framework to make a qualitative assessment of the various conversational AI agents that led to adjustments in the framework. For example, we derived an overarching stage of ‘qualify and allocate’ when we observed a pattern from all the firms to be funneling selective journeys to their conversational AI agent. For example, Monzo bank qualifies queries that a customer can’t self-serve using help text before allocating its chatbot into action.

To design conversational AI agent experiences, we apply the journey and experiential context from an understanding of CX management literature (de Keyser et al., 2020; Lemon and Verhoef, 2016). Based on the journey and experiential factors, sensing customer intent is the most vital stage in the design of conversational AI agents. For instance, from the conversation scripts, it can be observed that the Virgin media chatbot Terri directly sensed intent when we made a straightforward request about switching providers, however, the chatbot was unable to recognize a check coverage request. Instead, the conversation progressed towards whatsapp messaging with a human. Further, the sense intent stage is a dynamic one that continually determines context as the conversation between AI agent and a customer progresses.
Next, as the chat progresses, the conversational AI agents design should be able to dynamically adapt to changing context and continually assess intent. For example, a Marks & Spencer chatbot was able to adapt to a specific gift card related query as we progressed from inspirations for gifts to provide options for purchase. Meanwhile, as the time to resolve was increasing, the conversational AI agent quickly adapted and made an offer to transfer us to a human agent.

Infusion of tone to conversations on text or voice or visual expression is a key design consideration for managing conversational AI agents. Studies on such human-AI encounters such as, (Filieri et al., 2022; Pantano and Scarpi, 2022) have illustrated differing emotions need to be considered in service robot design and hence this is applicable to conversational AI agents too. Further, understanding intent through changing communication etiquettes such as emojis and response through emojis in text (Riordan, 2017) are factors that firms need to constantly adapt to provide a suitable emotional response to customers through conversational AI agents. Several conversational AI agents we observed remained low on emotion and sentiments except for Amazon Alexa that did assign tone to conversations, for example, the conversational AI agent’s tone changed when talking about the weather or expressing disappointment when unable to handle a music related query.

Further, a conversational AI agent needs to be able to delegate seamlessly to a human agent for requests that it can’t handle autonomously or is not tasked to handle by design. Here, the balance between AI agents substituting humans vis a vis augmenting humans needs to be clear at the design stage (de Keyser et al., 2019; Larivière et al., 2017). Frontline service employees need to be enabled with the relevant context when dealing with requests arriving from conversational AI agents that can be applied from frontline service-AI theories applicable to service (de Keyser et al., 2019; Robinson et al., 2020). For example, Monzo have designed progressive service escalation from customer self-service to chatbot support to human agent support with context being passed on seamlessly. As early as the chatbot sensed our intent that ‘debit card was not working’ and as the chatbot was designed to handle typical problems like loss or damage or contactless failures, we were passed on to a human agent who seamlessly took over without repeating the earlier diagnostic enquiries by the chatbot and the human agent was trained to handle complex queries that have a superior level of regulated security procedures.

Two further dimensions relate to a firm’s capabilities that are not directly visible to the customers but are critical elements in fulfilment of the customer experience on a conversational AI agent. The first one is the ability to automate the underlying process autonomously to fulfil customer’s requests during the conversations with an AI agent. For example, we attempted can complete purchases while having chats with agents on retailers like Sephora and Marks & Spencer. However, this was not possible due to various reasons including security. We were directed to the websites or call centers, alternatively. Studies on conversational AI agents in the service context have focused on AI agents emoting.
resembling human behavior or personalizing experience or human emotion when dealing with robots but have not looked at autonomous process orchestration within customer conversations with AI agents. Thus, we contend that the further empowerment of conversational AI agents to automate processes to fulfill customer requests, the greater their level of adoption and benefits from conversational AI agents. However, further empirical research is needed in this area to explore autonomous conversational AI agents.

The second critical dimension that remains invisible to customers is training of conversational AI agents. While many chatbots and voicebots come pre-trained with NLP capabilities, continuously learning from real customer conversations can improve the designs of conversational AI agents. Using customer journey data and analytics along with conversational analytics need to be utilized to train AI agents. Theories on data driven customer experience are equally applicable to designing data driven conversational AI agents (Holmlund et al., 2020; McColl-Kennedy et al., 2019). We observed market leaders like Google and Amazon Lex powering the conversational AI agents in the firms we interacted with but were unable to assess the level of training. Hence, research from a firm’s perspective in understanding how conversational AI agents are trained outside of market standard NLP training will be impactful in improving conversational AI agent designs.

### Table 2. Assessed industry chatbots using developed framework

<table>
<thead>
<tr>
<th>Firm</th>
<th>Virgin media</th>
<th>Vodafone</th>
<th>Monzo</th>
<th>Sephora</th>
<th>M&amp;S</th>
<th>Amazon</th>
<th>Aviva</th>
<th>Expedia</th>
<th>Tesla</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sector</td>
<td>Telecom</td>
<td>Telecom</td>
<td>Retail banking</td>
<td>Consumer retail</td>
<td>Consumer retail</td>
<td>Big Tech</td>
<td>Insurance</td>
<td>Travel</td>
<td>Automotive</td>
</tr>
<tr>
<td>Name</td>
<td>Terri / Tobi</td>
<td>Chatbot</td>
<td>None</td>
<td>Chatbot</td>
<td>Chatbot</td>
<td>Alexa</td>
<td>Vivy</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Type</td>
<td>Chatbot</td>
<td>Messenger</td>
<td>Web</td>
<td>Web</td>
<td>Web</td>
<td>Smart speaker Voice</td>
<td>Chatbot</td>
<td>Chatbot</td>
<td>Voice based In-car</td>
</tr>
<tr>
<td>Channel</td>
<td>Web</td>
<td>Web</td>
<td>App</td>
<td>Chatbot</td>
<td>Chatbot</td>
<td>Voice directed</td>
<td>Web</td>
<td>Web</td>
<td>In-car</td>
</tr>
<tr>
<td>Qualify &amp; Allocate</td>
<td>Found on contact us</td>
<td>Progressive escalation</td>
<td>Help pages replicated in a bot</td>
<td>Self-directed to Help page</td>
<td>Help pages replicated in a bot</td>
<td>Directly takes you to chatbot</td>
<td>Press of a button</td>
<td>Press of a button</td>
<td>Press of a button</td>
</tr>
</tbody>
</table>

4. Applied Framework

4.1 Assessing existing conversational AI agents

While the secondary data collection has served as a refinement of the developed framework, we can further use the data to develop a taxonomy of existing conversational AI agents. This can help us to better understand the capabilities and identify future research. It also helps forecast the success of individual chatbots. We distinguish chatbots along two dimensions, the level of personalization the chatbot can enable and the level of empowerment the chatbot possesses. The conversational AI agent stages as shown in figure 1 can be loosely assigned to the category of personalization and empowerment. Therefore, we will use the stages above to determine the level of personalization and empowerment.

At the beginning of each personalization effort, there is the need to collect information about the customer (Fan and Poole, 2006). To measure the capabilities of a conversational AI agent to personalize the interaction, we are measuring what customer information the conversational AI agent is capable of detecting and reacting to. This consist of three parts (stages). When a customer first approaches a conversational AI agent with a complaint or problem, the agent needs to detect what the customer wants and needs (intent). This capability can vary in strength depending on training and implementation. For example, a rule-based chatbot can only understand pre-determined intentions. When using pre-trained language models or dynamic training methods, the chatbot can learn and even understand intentions that are not company-specific or pre-installed. Now during a conversation, the customer’s intent can change, and additional requests might need to be addressed. Therefore, we assess whether the conversational AI agent can further adapt to the changing intent of the customer and respond dynamically. As already highlighted in the theoretical background, recognizing the affective state of the customer represents a next
level of personalization. If the chatbot is able to react to the customer's emotions and feelings and adapt the conversation accordingly, this enables the level of relational personalization (Huang and Rust, 2020). To measure the level of personalization capabilities, messages with clear intent were sent to the conversational AI agent and then changed after two utterances. Further in a second run keywords with clear display of emotion (e.g., sad, happy) were included into the conversation to see if the response changes.

In addition to the ability of the chatbot to understand the customer and enable personalization, we also differentiate by the level of empowerment. In simple terms, what the company allows the chatbot to do. One element of empowerment is the ability to start processes autonomously. These processes can range from looking up product information (e.g., price, delivery time) to ordering products (e.g., Amazon Alexa). One orchestration that is often implemented is escalating to customer service and connecting to a phone. This is mostly applied when the chatbot itself has a low level of empowerment. Another element that influences the assessment of the level of empowerment is the chatbot's access to the customer journey. As mentioned above, some chatbots only cover part of the customer journey, while others have the ability to operate across the entire journey. As with the personalization level, we examined the individual cases by asking whether the conversational AI agent can connect us with a human agent, can initiate a typical process and what areas (pre-, post-, purchase) it covers and can inform about.

The assessment of the individual conversational AI agents was based on subjective assessments by the two authors. We evaluated each case according to the described six categories. By adding up each of these scores: High (3 points), Medium (2 points), Low (1 point), we got the empowerment and personalization score. This qualitative assessment shows that chatbots such as Amazon Alexa stand out and are ahead of Vodafone's and Virgin Media's chatbots in both areas.

4.2 Taxonomy of conversational AI agents

The differentiation between conversational AI agent characteristics allows us to introduce a 2x2 taxonomy (see Figure 2). The taxonomy allows to position currently existing conversational agents and to guide next design steps for improving the customer experience.

We refer to conversational agents with a low personalization and empowerment level as “fact-finders”, being characterized as only being able to answer questions of the customer, stored in the knowledge- or database. These are commonly deployed as a first conversational agent and aim to ease the access to the FAQ pages. Examples from above include Vodafone, Sephora, Tesla, Virgin Media, and M&S. If we increase the level of personalization capabilities, the conversational agent is able to pick up on non-task-related aspects within a conversation and perform chit-chat, as well as eventually relate to the customer’s affective state. Conversational agents can hold a conversation and relate to the customer and their issues. Therefore, we call this kind of chatbot “empathetic chit-chatter”.

On the other side, if companies want to increase the level of empowerment first, to move away from “fact-finders”, they allow the conversational agent for example to orchestrate processes such as booking or cancelling tickets. These “facilitator” conversational agents, low personalization and high empowerment, are task-oriented and aim to resolve the customer issue. However, these chatbots are not able to go beyond the recognized intent and struggle to deal with out-of-scope questions. Conversational agents who possess a high level of personalization and empowerment combine the benefits of “empathetic chit-chatters” with “facilitators” and are characterized by the ability to go beyond task-oriented conversations, perform chit-chat and answer out-of-scope questions. It further has the capabilities to start processes automatically and adjust their tone according to the customer’s mood. Some examples include Amazon’s Alexa, and Monzo. Because of their capabilities, we refer to these conversational agents as “assistant”.

To improve the CX, firms need to try advancing their conversational agents towards the “assistant” conversational agent. The outlined framework can help to allow building the capabilities to move towards a “assistant”-type conversational AI agent and improve the customer service experience.
5. Conclusion

As advancements in NLP and conversational AI agents become market standard, firms will need to make significant decisions on how to remain relevant and gain competitive advantage by using such technologies. Hence, managers will need to make a clear assessment on how personalized and how empowered their conversational AI agents need to be. In this paper, we have highlighted that conversational AI agents need to be designed by using techniques from multiple disciplines such as service personalization, AI in service, and CX management. The fundamental insight for CX managers is they need to prepare conversational AI agents by training the AI agents to sense customer intent based on available customer data and analytics both historical and in the real time.

However, this study also has some limitations. The small sample size used for validating the framework only gives an indication of existing approaches and their maturity for enabling improved customer experience. Further, there are sector specific design elements which might need to be considered (e.g., privacy and regulations for banking interactions) in addition to the framework described above. Although the modality of the conversational AI agent (e.g., voice, text, anthropomorph appearance) can play an important role in designing conversational AI agents (Miao et al., 2022), it is beyond the scope of this study.

Although, our framework offers an initial insight into designing conversational AI agents from a firm’s perspective, further empirical studies are needed in support of this emerging area that can be led by service research. Future research avenues are necessitated like, how to qualify and allocate service requests to a conversational AI agent? Should empowered agents be able to qualify requests and autonomously decide the next courses of action? Sector specific studies on conversational AI agents will add insights to both theory and practice. For example, are conversational AI agents more suitable for post purchase journeys in the retail banking sector and what are the implications for design of such agents within this regulated sector? In our paper, we included both voice-based, and text-based conversational AI agents but empirical studies comparing design of differing categories of conversational AI agents will also be useful in progressing our understanding of this area. Methodologically, studies can take a machine learning based approach to text mine conversational AI agent scripts (e.g., Ordenes et al., 2014) to gain understanding beyond human analysis of scripts. Further, conversational AI agents are being deployed to empower frontline service employees and studies on that combine employee experience will enrich this field.

6. References


