Article

VoiceBack: Design of Artificial Intelligence-Driven Voice-Based Feedback System for Customer-Agency Communication in Online Travel Services

Anniki Skeidsvoll Edén *, Pernilla Sandlund , Montathar Faraon and Kari Rönkkö

Department of Design, Kristianstad University, 291 88 Kristianstad, Sweden
* Correspondence: jasmine.skeidsvoll0005@stud.hkr.se

Abstract: Online travel booking has become increasingly popular; however, most travel websites do not yet offer voice interaction. This study introduces VoiceBack, an artificial intelligence (AI)-driven voice-based feedback system conceptualized to support both customers and online travel agencies during the booking process. It proposes a theoretically and empirically underpinned design concept that involves a voice user interface (VUI) for customer feedback. This feedback, collected by an AI agent, is analyzed and converted into actionable statistics, which are then presented to online travel agencies through a visual interface. The interface is designed to highlight problem areas and usability issues during the booking process. This study contributes to the field of human-centered AI, by offering insight into the complex process of designing and integrating voice, emotion, and feedback within user interfaces. This integrated approach can enrich the user experience of customers when booking travel online, and pave the way for more intuitive and responsive interaction designs in the future.

Keywords: artificial intelligence; voice user interface; feedback; online travel booking

1. Introduction

Artificial intelligence (AI) has recently received considerable attention in society. Increased computing power, access to data, advances in machine learning, deep neural networks, and natural language processing have all contributed to the prominence of AI [1]. Furthermore, the openness to sharing AI technology through open-source platforms has facilitated the development of sophisticated AI systems [2]. AI-driven voice-user interfaces (VUIs), such as Apple's Siri, Amazon's Alexa, and Google Assistant, are gaining popularity in society because of their time-saving, hands-free, and attention-liberating interaction capabilities. They increasingly support users in complex tasks such as managing health-related activities, booking restaurant visits, and planning travel [3]. Voice interaction is perceived as intimate and warm compared to text interaction, which fosters user trust in systems [4].

The voice is the primary medium of human communication [5,6]. It plays a crucial role in the interaction between users and AI systems, a field known as voice-based human–agent interaction (vHAI) [7]. The nuances of the voice convey emotional information through variations in pitch, texture, volume, and duration [8,9]. Humans instinctively interpret these nuances, which allows them to detect underlying emotions, such as joy, fear, and sadness, in a voice [9].

Most AI systems generally only focus on the semantics and meaning of a statement and often overlook emotions in the voice [10]. This oversight has significant implications, because emotions in a voice can change semantic meaning. As emotions are a fundamental aspect of social communication, the ability of AI systems to interpret emotions is crucial for user interaction [5]. This requires that VUIs offer personalized and natural conversations

with users [11]. Previous research has stressed the importance of user feedback in VUI systems to measure customer experience (CX) [12]. Although some studies have investigated how users handle problems in VUI-specific situations [13,14], more studies are needed to explore the nuanced relationship between emotional expression in voice and its impact on VUI design and user experience.

Research on voice-based feedback systems that extract feedback from voice communication is crucial for developing advanced AI systems. Booking flights and train tickets is one of the most common Internet activities and generates significant traffic on the websites of online travel agencies [15,16]. The sale of travel-related products and services such as flights, cruises, package tours, and hotel stays online is a growing market [17]. This growth creates a high demand for online travel agencies to meet customer needs in a competitive market [16,17].

Currently, most travel websites do not feature AI-based VUI systems. Hence, this study aimed to explore, conceptualize, and evaluate the design of a voice-based AI agent for online travel agencies that supports customers during the booking process. Although it is important to consider user privacy, system effectiveness, data handling of travel information, and the efficacy of voice communication, such aspects are beyond the scope of this research. Correspondingly, this study was guided by the following research question: How can a design concept be theoretically anchored, conceptualized, and evaluated to interpret and visualize customer voice feedback to support online travel agencies' ability to understand and improve customer experiences during the booking process? This question was addressed by applying the methodology of concept-driven design research [18].

The remainder of this article is organized as follows. The second section presents the background and theoretical foundations, covering artificial intelligence, voice user interfaces, and the design of AI-driven voice-based feedback systems. The third section describes the methodological approach, detailing the concept-driven design research method, participants, materials and measures, procedure, and ethical considerations. Next, the fourth section presents the results and analysis of the design concept process. The fifth section discusses the findings, their implications for online travel services, practical challenges, and directions for future research. Finally, the sixth section offers concluding remarks on the VoiceBack design concept.

2. Background

2.1. Artificial Intelligence

The term artificial intelligence (AI) was coined in 1956 by John McCarthy, who defined it as "the science and engineering of making intelligent machines, especially intelligent computer programs" [19]. Subsequent research has expanded this definition, describing AI as systems that perform tasks through intelligent actions [1] and possess the ability to learn, understand, or manage new challenging situations [20].

Two key properties are fundamental to AI-driven services. The first is *adaptivity*, in which a system learns from experience and becomes more intelligent. The second is *autonomy*, which refers to a system's ability to perform tasks automatically, without human support [1]. These properties are essential for building intelligent systems that perform specific tasks and have paved the way for specialized branches of AI. One such branch is natural language processing (NLP). NLP's ability to process human language forms the basis for more complex functions in AI, particularly in understanding and interpreting spoken and written languages. It has laid the foundation for voice-based human–agent interaction (vHAI) [6] and practical applications, including voice-based assistants.

Natural language understanding (NLU), a subset of NLP, focuses on understanding the nuances of human language, including context, intent, and the meaning of words and phrases. This understanding is fundamental to AI-driven speech recognition [21] and is critical for voice assistants to handle voice or text input from users [22]. Intertwined with NLU is the concept of social AI, which aims to understand people's social contexts, to

promote natural interaction [22]. This integration improves the ability of voice assistants to grasp semantics and effectively analyze social communication patterns.

In recent years, significant advances have been made in AI conversational capabilities, notably through the development of sophisticated language models [1]. Text-to-speech (TTS) technology is a crucial element of this evolution, which significantly enhances automated interactions by making them more human-like, natural, and credible. The foundation of NLP has been pivotal in the advancement of TTS technologies, which can be found in voice assistants, navigation systems, customer support, and audio books. As AI continues to evolve, new challenges and opportunities arise, particularly in understanding and responding to human emotions.

One of the current challenges for AI-driven services is to identify and respond to emotions using emotional intelligence. The need for AI to achieve a conversational flow that feels safe, seamless, and emotionally intelligent has led to the rise of a new field called emotional AI [23]. Emotional AI focuses on understanding and reacting to human emotions, to foster authentic and natural interactions between users and systems. This capability is critical for moving beyond mere speech or text recognition toward understanding the emotional context of interactions. Emotional AI seeks to achieve this by analyzing people's tone of voice [23]. This capability not only enhances communication effectiveness but also plays a vital role in how an AI is perceived and trusted by users.

Humans can perceive the emotional tone of a conversation and engage in an emotionally intelligent dialogue. The lack of this dialogue and emotional awareness in machines makes it challenging to develop an intelligent conversational agent that can increase user engagement and promote trust in it [24]. Similarly, when voice assistants employ a human voice in dialogue with users, they are perceived as more relatable and trustworthy [25,26].

This focus on trust extends beyond the interaction level to encompass the overall design of AI systems, particularly in situations where users are expected to rely on, follow, and use the information provided by these systems [27]. The design of interfaces that prioritize user trust in AI is particularly relevant for the protection of identity and privacy [27]. One approach to fostering trust is through the development of explainable AI (XAI), which aims to make AI operations transparent by providing users with clear feedback and understanding of AI functionalities [28]. This transparency can foster trust and enhance user control, which is particularly relevant as AI technologies become more integrated into everyday devices and interactions, such as those that involve VUIs.

2.2. Voice User Interfaces

Voice user interfaces (VUIs) are becoming increasingly popular and integrated into more devices and contexts. These include applications for customer support, car assistants, smartphones, smart speakers, social robots, and voice assistants [4,8,29]. Interactions in these systems primarily occur through the voice [6,7] and are often powered by AI technology [6,30]. Voice assistants serve as the primary platform for VUIs, and their market share is expected to grow. The growth of VUIs reflects a deeper and more fundamental aspect of human interaction: the use of the voice as the primary medium of communication.

The human voice comprises several components, including speech, a mixture of vowels, consonants, sounds, and phonetics. Speech encompasses both non-linguistic elements and emotional expressions [6]. The new generation of AI-driven voice assistants enables users to communicate intuitively and naturally [26,31]. Incorporating anthropomorphic (i.e., human-like) properties into these systems further strengthens the user–system relationship [6]. Voice interaction offers unique advantages such as flexibility and time efficiency, and provides more control to users [32]. Additionally, it is practical to free users' hands and attention for other tasks, particularly benefiting people with visual impairments [4,7,29].

The combination of a VUI and a graphical user interface (GUI) has been shown to positively impact the user experience. In this combination, the voice component provides auditory signals, such as personality and emotions, while the graphical interface delivers

visual information through text, graphics, and images [6]. This design of voice-based multi-modal user interfaces offers significant support for users struggling with complex systems, helping them understand the interface, grasp the intent of the interaction, and choose appropriate communication commands [33,34].

As VUIs become more sophisticated and emotionally aware, they are increasingly being adopted for more complex customer–business interactions. Customers have started to communicate with businesses through voice assistants and can perform actions using voice commands that were previously unavailable [32]. To effectively meet evolving user expectations, companies need to offer voice interaction experiences that optimize the benefits of the voice and provide services that exceed the capabilities of existing interfaces [35]. An area that has the potential for significant improvement through voice interaction is online travel services, where an AI-driven voice-based feedback system could improve user experience in planning and managing travel arrangements.

Although AI and VUI technologies show great promise, it is necessary to acknowledge their limitations. AI systems still struggle to understand context and nuance in human communications, often leading to misinterpretations [36]. The accuracy of voice recognition has improved over the years, but remains imperfect, especially for unusual accents or in noisy environments [37]. Moreover, the emotional intelligence of AI is still far from a human level, which can lead to inappropriate responses in sensitive situations. These limitations underscore the need for the careful design and implementation of AI-driven voice interfaces.

2.3. Design of AI-Driven Voice-Based Feedback System

Research indicates that customers are often dissatisfied with the functionality and design of the websites of online travel agencies. They express dissatisfaction and share their feelings and opinions through online reviews [38]. As a result, these agencies must adopt novel strategies to meet the expectations of their customers.

Online travel agencies can mitigate customer dissatisfaction by optimizing their efforts to understand customer needs and preferences [15,39]. A potential strategy to achieve this is to integrate an AI agent into their booking systems. An AI agent may be able to reduce company costs, increase turnover, and provide services 24/7 [40]. It could also respond faster and handle larger volumes of data than traditional customer services. Integrating an AI agent into booking systems would enable online travel agencies to improve customer's booking experience and make transactions more efficient. These improvements in service quality are likely to increase customer trust and purchase intentions [35,41].

The online purchase process, such as booking a trip, involves five distinct stages: (1) problem identification (i.e., customer needs are confirmed and related to services or products); (2) information collection (i.e., information is collected from multiple sources to ensure the quality of decision making); (3) scheme evaluation (i.e., evaluation of the features of the service in relation to other services); (4) purchase decision making (i.e., depending on whether a customer is satisfied or dissatisfied, a decision will be made on whether to proceed with the purchase); and (5) post-purchase behavior (i.e., how a customer behaves after the purchase) [40]. Incorporating these five stages into the design of an AI agent could improve online booking experiences for customers. An AI agent has the potential to better understand and predict customer needs and preferences throughout the booking process. It could provide valuable information for online travel agencies, help them tailor their offerings, and improve customer satisfaction [40]. Consequently, this integration would bridge the gap between customer behavior and optimal service delivery, which could lead to a satisfying user experience for customers and more actionable feedback for travel agencies [40].

Developing an identity for voice assistants in the design of VUIs is also critical. This identity can be achieved by creating a persona for the system, in which its personality and characteristics are more prominently featured. This persona can act as a bridge between a customer's mental model and a design concept [31].

Research indicates that both positive (i.e., satisfaction) and negative (i.e., frustration) customer emotions are critical in a dialog-based AI feedback system, where the focus is on analyzing and assessing the nuances of a customer's voice [42]. This understanding of emotional nuances in customer feedback is particularly relevant in the evolving landscape of VUIs. Currently, VUIs leverage crowd-sourced voice feedback (CVF) to actively solicit user responses, where users can respond to questions by voice with a yes, no, or a numerical rating of one to five [12]. In addition, feedback can also be obtained through explicit user feedback, which involves users sharing their experiences directly with a system. This form of feedback has shown considerable value, particularly in settings that use voice assistants [12,43]. Furthermore, the provision of continuous and timely feedback has the potential to positively influence the user experience of VUIs [44].

Building on this foundation of user feedback in VUIs, the field of AI is advancing towards more nuanced forms of interaction, such as automated emotion recognition. This technology can enhance the user experience by personalizing interactions based on emotions detected in user feedback. Therefore, the ability of AI systems to accurately identify and interpret emotions is becoming increasingly crucial, offering more engaging and interactive experiences [45]. Alternative methods of expressing and interpreting emotions are also emerging. The use of emojis for feedback provides a visual way for users to express their feelings, as they are familiar with their use and associated emotions [46]. Similarly, color serves as another means to depict emotions, further broadening the spectrum of tools available for emotion representation and interpretation in AI systems [46,47].

This evolving understanding of emotional recognition and expression in AI mirrors broader trends in recognizing emotions in human interactions. Fundamental emotions, such as joy, sadness, anger, disgust, and fear, form the basis for more complex emotions [48]. These emotions, influenced by various factors such as circumstances, mood, and relationships, play a pivotal role in effective communication [49]. In recent years, businesses have shown an increased interest in the emotional experiences of customers during interactions, which constitutes a significant aspect of interactive systems [50].

The literature review presented here provides the characteristics, values, principles, functions, and technologies that are instrumental in the theoretical underpinning and crafting of a design concept. The next section will elaborate on this concept, its foundational elements, and its potential application.

3. Methodological Approach
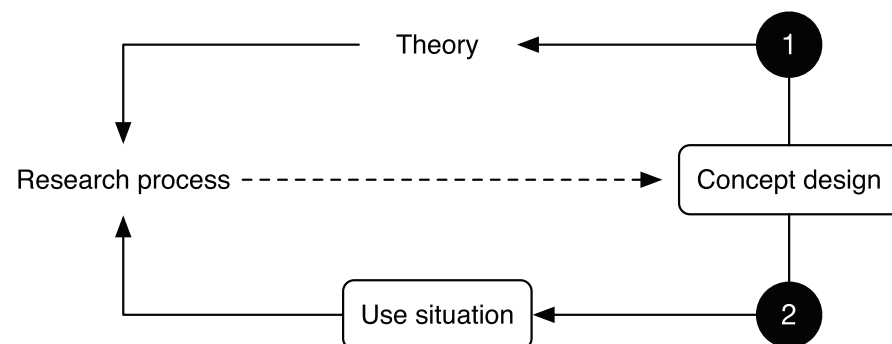
3.1. Concept-Driven Design Research

Concept-driven design research was used in this study, as it was considered suitable to address the research question [18]. The purpose of the concept-driven design approach is to manifest theoretical concepts in concrete designs [18]. In concept-driven design, new ideas and approaches are explored, rather than looking for a specific design solution for a certain problem in a certain context [18]. The concept-driven design approach is based on three principles ([18], p. 98):

1. The point of departure is conceptual/theoretical rather than empirical.
2. The research furthers conceptual and theoretical explorations through hands-on design and development of artifacts.
3. The end result–that is, the final design–is optimized in relation to a specific idea, concept, or theory, rather than to a specific problem, user, or a particular use context.

The relationship between interaction research in terms of theory and use situation is illustrated in Figure 1 [18]. The purpose of the concept-driven approach is knowledge production, and the question of whether this can "lead to improvements of a situation is of lesser interest, or maybe even no interest at all" ([18], p. 101). Several studies have adopted the concept-driven design research approach and, although they generally resulted in a prototype [51–53], others concluded in concept designs [54–56]. The research approach consists of seven methodological activities: *concept generation*, *concept exploration*,

internal concept critique, *design of artifacts*, *external design critique*, *concept revisited*, and *concept contextualization* [18].

Figure 1. The concept-driven design research approach and its relation to theory and use situation.

The first activity, *concept generation*, aims to generate ideas for novel concepts based on previous research findings. This activity focuses on discovering creative and unexpected solutions. This can be achieved by associating and comparing contradictory theoretical concepts, theories, and historical data from existing research and other design areas to create innovative concepts [18].

The second activity is *concept exploration*, which involves designers exploring the design space by venturing beyond the initial idea and what is already known. This can be achieved through practical work with different materials and by creating models and prototypes. This activity focuses on finding alternative and unknown possibilities in what is already known. The objective is to discover novel ideas that challenge existing theories. Concept exploration aims not to further develop or test well-functioning ideas, but to search for new areas in the design space [18].

The third activity entails *internal design critique* of the concept. In this activity, key concepts are identified, mapped, and represented through simple sketches, prototypes, and mock-ups. The feasibility and potential of these sketches and prototypes is examined before the concept is advanced to a defined form in the next activity. The internal critique must clearly show how design and concepts connect to theoretical foundations. For a concept to succeed, it should exhibit three qualities: first, the core concept should be novel; second, the concept should relate to existing theory; and third, the concept should be expressible in a concrete design [18].

The fourth and fifth activities comprise the *design of artifacts* and the *external design critique*, respectively. In the fourth activity, the artifact should express the design of the concept as a whole and be underpinned by the theoretical development. In this activity, concept-driven design research is integrated with theoretical development, leading to the creation of a theoretically underpinned concept [18]. The fifth activity includes an evaluation of the design concept. The design concept is presented to the public, who can critique the idea, the overall design concept, and the theoretical principles underlying the design concept. External design critique serves as a crucial step in validating a concept and its theoretical basis, as manifested in the current design [18].

The sixth activity, *concept revisited*, involves reexamining and revising the concept based on the critique received in the previous activity. Depending on the critique of the concept, different measures may be appropriate in different cases. Common criticisms include a flawed basic idea, problems with the composition, or incomplete parts. Anticipating user reactions to the concept beforehand poses a significant challenge [18].

The seventh activity, *concept contextualization*, entails aligning the revised design concept with existing theories and illustrating its connections to prior research. The research contribution poses a significant challenge, as it is critical to clearly distinguish the unique parts of the proposed concept from the already established concepts and theories, while also showing how they relate to each other [18].

3.2. Participants

The sample size in qualitative research should be determined based on the principle of saturation [57]. Specifically, data reach a saturation point when no new information, insights, or themes emerge from further data collection [58,59]. The participants in this study provided sufficient data to achieve saturation. They were selected using a convenience sample of online travel agency employees. The extensive experience of the participants enabled valuable insights during the external design critique activity. The study included eight participants, five women and three men, with an age range of 38–59 years ($M = 51.1$; $SD = 7.8$). The participants were contacted by email, social media, and phone calls. To ensure anonymity, the participants will henceforth be referred to as participants 1–8 (P1–P8).

3.3. Materials and Measures

Twenty-one qualitative questions underpinning the interviews originated from previous related research (see the Appendix A) [54,56,60]. They were adapted according to the needs of VoiceBack and classified as follows: The first category was demographics and collected information about participants, for example, age and gender. The second category consisted of questions related to the concept of design, that is, how participants perceived the concept in terms of its strengths and weaknesses. The third category addressed questions related to artificial intelligence and ethical considerations for AI-powered services. The fourth category included questions about voice user interfaces, such as what experience participants had with VUIs. The final category was feedback, in which participants expressed their preferences on how they would like user feedback to be presented in an interface.

3.4. Procedure

In the external design critique, semi-structured interviews with eight participants were conducted, with each interview lasting approximately 30 min. Zoom (https://zoom.us), a video conferencing tool, was used for each interview. The interview sessions started with a presentation explaining the purpose of the study. Thereafter, the design concept and a video scenario were shown that visualized a conversation between a user and an AI agent. The participants were also shown a presentation of feedback within an online travel agency interface. Subsequently, the design concept (Figure 2) was displayed, accompanied by a verbal explanation and clarification of how the theoretically underpinned design concept was intended to function. The participants were then asked to answer twenty-one questions about the design concept. This formed the basis for their reflections and evaluations of the design concept. The average interview time was about 30 min. After the interviews had been completed, the recordings were transcribed with the AI-powered transcription tool Cockatoo (https://www.cockatoo.com). Following data collection, the material was analyzed using thematic analysis [61].

3.5. Ethical Considerations

Because this study involved participants, careful consideration was given to the four ethical principles of the Swedish Research Council [62]. This study applied the principles of secrecy, professional secrecy, anonymity, and confidentiality. Consent was obtained verbally from all participants at the beginning of the interview, when they could accept or decline participation in the study. The participants were informed that the information would be kept confidential [62]. They were informed that the collected data would only be used to revise the design concept and would not be shared with third parties. Furthermore, it was also explained that their real names would be replaced with pseudonyms to ensure anonymity.
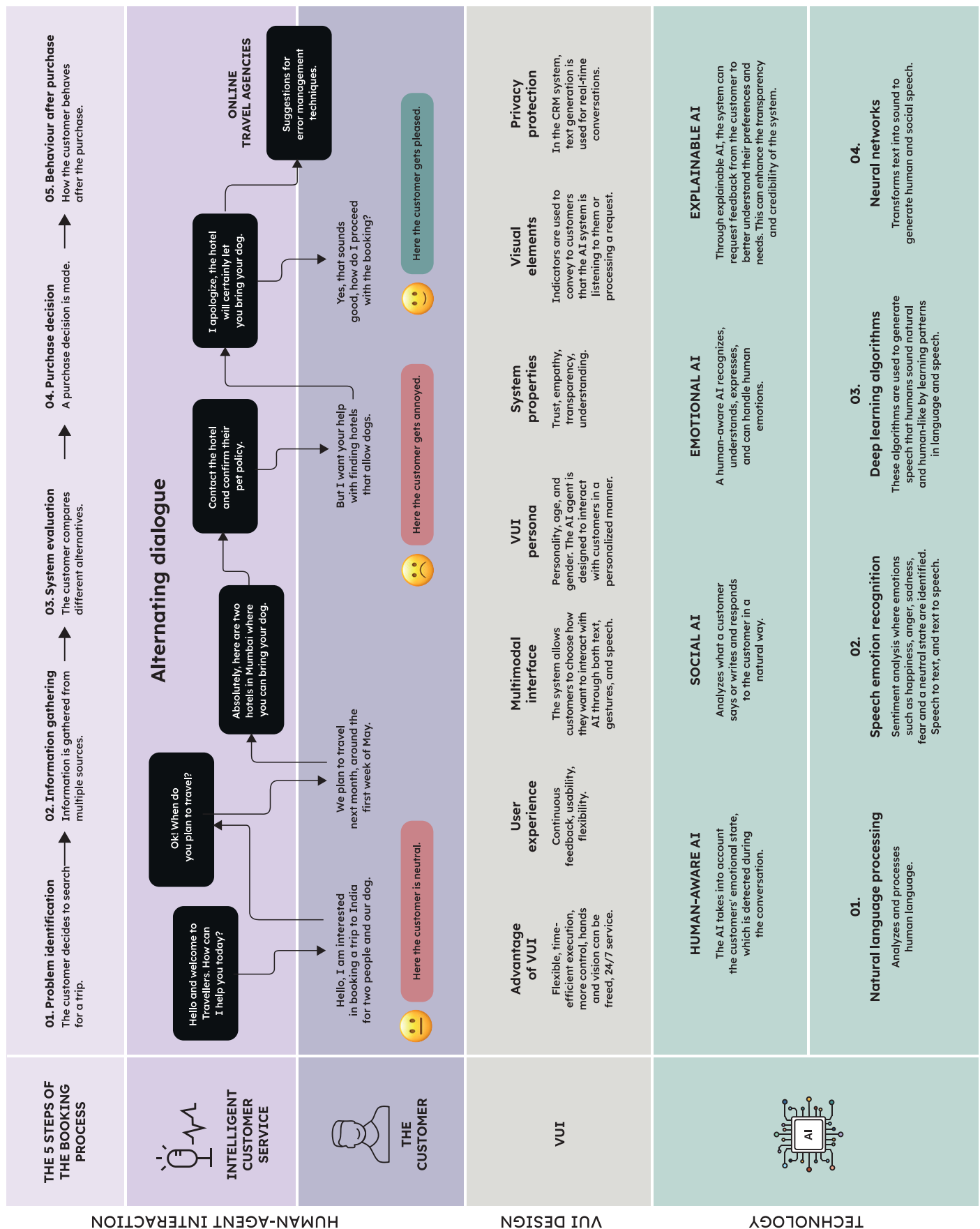
Figure 2. The theoretically underpinned initial design concept of VoiceBack, featuring an AI-driven voice-based feedback system.

4. Results and Analysis

4.1. Concept Generation, Concept Exploration, and Internal Concept Critique

In the *concept generation* activity, a theoretical foundation for the concept was established through a review of the literature of existing research in the field of HCI. This review focused on AI, its history, the current state of AI technology, and an outlook on future AI technologies. VUIs were also an integral part of the literature review, which focused on user experience by presenting an account of how VUIs operate today, the advantages they offer, and how these systems should be designed in a human-centric manner. The study also addressed voice-based human–agent interactions and the development of a voice-based feedback system for online travel agencies. Empirical research was conducted to examine the relationships between various concepts, theories, and historical contexts related to AI, VUIs, and the booking process. This was done to discover solutions that were not immediately apparent from what was already known. The design space was explored using empirical data during the development of the concept. This was carried out to understand how the theoretical concept could be visualized as a concrete design. This approach produced unique insights, culminating in a novel, innovative, and theoretical design concept that challenged existing theories in the field.

During *concept exploration*, initial theoretical concepts were generated using the brainwriting design method. This method was adopted in the collaborative tool Miro (https://miro.com), where keywords were carefully selected to see if the design concept and research theories underpinning this study were theoretically feasible. These keywords, chosen from previous research and based on the research question, were intended to define the type of information necessary to help online travel agencies capture customer feedback, especially by focusing on emotions in their voices. The first keywords included the five critical steps of a booking process: *problem identification*, *information gathering*, *system evaluation*, *purchase decision*, and *behavior after purchase* [40]. These are significant because they establish a framework for analyzing customer interactions at every stage of the booking process. In addition, they can ensure that the feedback system is technically efficient and emotionally intelligent in identifying and responding to the various needs of customers throughout their booking journey [40].

Furthermore, *voice interaction* was adopted to provide customers with a flexible way to interact with the system. It facilitates greater control for customers and is more time efficient than text interaction [32]. In addition, it offers customers *flexibility* to use their hands and vision for other tasks [4,7,29,63]. Feedback within a voice agent and its ability to handle errors in the interaction should be designed so that it is natural to customers [6,64]. Providing timely and *continuous feedback* is essential for enhancing the user experience of customers who interact with a VUI [44]. This can improve the overall usability of VUIs and foster a sense of fluency in the dialogue between customers and a system.

Other factors that can positively influence the user experience of VUIs, which were incorporated into the design concept, include *trust*, *empathy*, *transparency*, and *understanding*. Establishing trust is essential in the design of VUIs, and using a human-like voice in interactions contributes to building trust between customers and a system. The human voice element in dialogue enhances the relatability and credibility of a system [25,26]. Empathy in VUI, mirroring the human ability to perceive and respond to emotional tones in conversations, can foster trust and improve the user experience [24]. Trust can also be built through transparent AI systems and by providing customers with an understanding of AI's capabilities [28].

To collect and represent customer emotions effectively based on their feedback, the design concept incorporates *visual elements*, such as *emojis* and *color* schemes [46,47]. Emojis serve as universal and intuitive symbols that convey a wide range of emotions. They allow customers to express their feelings quickly and in a user-friendly way. Additionally, the use of color as a visual tool helps in the emotional categorization and representation of feedback for online travel agencies.

Moreover, the inclusion of a *persona* in this concept is critical to the shaping of a voice assistant identity in the design of *VUIs* [31]. A well-crafted persona for a voice assistant is

necessary, as it determines how customers perceive and interact with the system. This requires defining characteristics such as the tone, speech style, gender, and personality traits of a voice assistant.

The design concept also recognizes the need for versatility in user input methods. To address this, a multimodal interface was integrated to allow customers to choose input options and support situations where the voice cannot be used [65]. For example, in noisy environments or situations that require discretion, customers may find voice input inconvenient or inappropriate. In such cases, the multimodal interface allows alternative interaction methods such as text input or touch-based commands. The design concept further enhances the capabilities of the system by incorporating four AI technologies: *human-aware AI* to communicate with customers through *natural language processing* and *deep learning algorithms* [21,22]; *social AI* to promote natural interaction [22]; *emotional AI* for understanding, reacting to, and detecting human emotions in a voice [23]; and *explainable AI* to provide transparency of the AI system and understand what the system is doing [28].

Subsequently, the design concept underwent an *internal design critique*, during which key concepts were selected and classified within the collaborative tool Miro. This activity reinforced the central concepts and highlighted several unique characteristics that could be linked to existing theories. After reviewing the design concept, revisions were made to clarify its theoretical underpinnings. The goal was to facilitate the participants' understanding of the overall concept. The revisions also enabled certain connections between the accepted theoretical foundation and the concept to be clarified or added.

4.2. Design of Artifacts

In the fourth activity, *design of artifacts*, digital sketches, wireframes, and prototypes were created in the collaborative interface design tool Figma (https://figma.com). Based on these sketches and wireframes, an initial design concept of VoiceBack was crafted with clear connections to the theoretical foundation (see Figure 2). Graphic elements, such as arrows, lines, colors, emojis, icons, and illustrations, were used to represent the interconnections between all parts of the design concept. In addition, two prototypes (see Figure 3) were created. The first prototype was used to illustrate how VoiceBack would look from a customer's perspective on the websites of online travel agencies. The second prototype was designed to show how feedback can be visualized on the interfaces of online travel agencies. A video scenario was also produced to illustrate the potential interaction between customers and VoiceBack. The video scenario can be viewed at: https://youtu.be/7dPNJXw3Cxg (accessed on 28 July 2024). The video depicts an interaction between a hypothetical user and a customer service representative of a fictitious travel company. Throughout the video, the user's location is displayed on an accompanying map.

The following section elaborates on the fundamental concepts, functions, methods, and processes that shaped the design of the theoretically underpinned concept. The five steps of the booking process were central to this design process: problem identification, information collection, scheme evaluation, purchase decision making, and post-purchase behavior [40]. The VoiceBack design solution introduces a voice user interface for the websites of online travel agencies with the aim of enabling customers to communicate with an AI agent using their voices. This advancement facilitates a more direct and intuitive interaction on a website or within an application.

VoiceBack uses a voice user interface and capitalizes on its unique ability to achieve flexible and time-efficient execution [32]. Adding a voice as an interaction option is based on the fact that the voice is the primary means of human communication [5,6], making it an essential component of voice-based human–agent interaction, or vHAI [7]. If fully implemented, VoiceBack has the potential to meet the customer's need for efficient 24/7 services and handle larger data volumes than traditional customer services [40]. Voice interaction is not only a functional improvement, it is also perceived positively by customers as convenient, pleasant, sincere, and warm [4,35]. This perception is supported by practical benefits, such as freeing the hands and attention for other tasks during interaction [4,7,29,63]. For this purpose, a multimodal

interface was introduced to support customers in various situations. This interface offers a customizable solution to meet the needs of the customer, with text and voice as input options [65].

The emotional aspect of voice interaction plays an essential role. Emotions are inherently embedded in the human voice [5,6] and were a vital consideration in the designing of VoiceBack. Hence, the system uses voice recognition technology to identify the emotional states of customers [45]. The adoption of automated emotion recognition enables the generation of customer feedback by extracting emotions from voices in VUIs. This approach not only measures the customer experience of a booking process, but can also improve it [12].

A key conclusion of the literature review was that voice assistants should communicate in a human, natural, and intuitive way [6,26,31], which was then incorporated into the design concept. These aspects underscore the importance of creating voice assistants that can interact in a way that feels familiar and comfortable to customers. The design of AI-driven services should also include characteristics such as trust [27], empathy [24], transparency, and understanding [28]. Trust and empathy are vital to ensure that customers feel safe and understood when interacting with AI systems. Transparency and understanding are crucial for ensuring that AI's actions and decisions are clear to customers, which can foster a sense of reliability and predictability.

These characteristics are necessary to build intelligent systems capable of performing specific tasks [1]. To meet these requirements effectively, adaptability and autonomy were introduced into the design concept. Adaptability ensures that the system responds dynamically to diverse customer needs and scenarios. Autonomy allows a system to operate independently, make decisions, and adapt responses, without continuous human oversight. Consequently, this feature can improve the efficiency and effectiveness of online travel agencies.

The following AI technologies have been integrated into the VoiceBack design concept to achieve automatic, natural, and credible conversations with customers [66]: (1) emotional AI to understand and respond to customer emotions [23]; (2) human-aware AI to communicate with customers in natural language, allowing intelligent and human-like interactions [22]; (3) social AI to understand the social contexts of customers and promote natural interactions [22]; and (4) explainable AI to make the system transparent and meet customers' needs for feedback and understanding of its actions [28]. Other core techniques embedded within the AI technologies mentioned above include machine learning, deep learning algorithms, and neural networks, which form the foundation of AI systems [21,25].

Moreover, text-to-speech (TTS) was included in the design concept, because it allows an AI to use voice to communicate with customers [66]. Natural language processing (NLP) was added to analyze spoken language and perform AI-driven speech recognition [21]. Furthermore, natural language understanding (NLU) was used to learn to process the input of customers' voice and text [22]. Another technology theoretically incorporated into the concept was OpenAI's GPT-4 (https://openai.com). This AI service enables natural conversations between customers and the proposed AI-driven voice-based feedback system [67]. Its integration into VoiceBack can serve as a valuable tool for online travel agencies to respond to questions, provide personalized recommendations about destinations, and translate customer feedback into recommendations for improvement.

A successful user experience can be achieved by designing usable, flexible, and satisfying websites [15,17,38,41]. The application of usability principles can improve the user experience and has been recommended when creating online services [68,69]. In practice, this approach is exemplified in how VoiceBack displays voice-based customer feedback. During the booking process, customer feedback is visually represented on online travel agencies' interfaces using charts, statistics, emojis, and colors. Color is a practical means of describing emotions [46,47]. Similarly, emojis are effective because people understand their use and the associated emotions [46]. Additionally, a real-time emotion indicator displays the communication between customers and the AI system through a graphical blob that changes color based on customers' emotions.

Figure 3. A theoretically and empirically underpinned revised design concept of VoiceBack.

4.3. External Design Critique

In general, the participants were impressed with the design concept and recognized its value in the business development of online travel agencies. Through the process of thematic analysis, three themes were identified from the interview data as follows:

1. **User experience challenges:** The participants frequently mentioned difficulties in understanding and interacting with existing VUIs.
2. **Personalization needs:** A strong desire for human-like and personalized AI interactions emerged among the participants.
3. **Implementation concerns:** The participants expressed concerns about the technical challenges of integrating AI systems into existing platforms.

These themes formed the basis for the subsequent design revisions and concept development. With regard to the first theme, several issues emerged concerning participants' past interactions with voice user interfaces (VUIs). Each participant had experience communicating with a voice assistant. The interviews revealed that they faced challenges in understanding the existing VUIs. Three participants (P5, P6, P8) emphasized the importance of understandable communication to prevent customer frustration. Their frustration, stemming from the voice assistant's failure to understand them, led to decreased trust in VUIs and resulted in discontinued use.

Two participants (P5 and P7) were generally skeptical of AI and held this view before being introduced to the design concept. Existing voice assistants were mentioned as having challenges in understanding dialects, which can cause irritation and frustration for users. One participant described their experience: *"That damn service just doesn't get my skånska [dialect], so we end up in conflict because it doesn't understand what I want"* (P6). Another participant shared a similar experience: *"I speak skånska [dialect], and when I talk, it never quite gets through"* (P5). A third participant also had a negative experience with voice assistants: *"Yes, I have a phone that I can interact with, but it doesn't always work as I want"* (P8). Another participant (P3) expressed concern about the VoiceBack AI agent's ability to understand language: *"Of course, there might be misunderstandings. It depends on how you speak and the type of words you use. It might misinterpret someone who's not very proficient in a language"* (P3). These results indicate that users of existing VUI interfaces were often excluded due to their dialects and how they were understood. The following revision was made to the design concept based on the previous feedback:

> **Design revision 1:** A diverse range of data sources can be utilized to ensure that VoiceBack remains inclusive of all groups (i.e., the data are not sourced from the same places, sources, or groups). Ethical AI was added to the design concept to ensure that the system is fair, secure, reliable, and free of biased data. This was integrated to maintain equal value for all individuals.

Regarding the second theme, five participants (P4, P5, P6, P7, and P8) expressed the desire for greater personality and human-like qualities in the AI agent during the external design critique. One participant described the concept's AI agent as follows: *"I would like to add personality, but I'm not sure. Maybe change the agent if you are annoyed by a particular voice. It lacks humor and real human responsiveness, and people don't choose as rationally as the example suggests."* (P6). Another participant noted similar experiences: *"When selling trips, it's about creating a relationship with the person to understand their needs and desires and selling them the right trip."* (P5). A third participant expressed the following about the VoiceBack AI agent: *"It seems impersonal, doesn't it?"* (P8). Participant P7 commented on the humanity in the AI agent's voice: *"It always gets better the more personal it becomes, or the more personal it sounds."* (P8). Participant P4 appreciated natural interruptions in conversations with the VoiceBack AI agent: *"Good to have natural breaks like a human does."* (P4). One participant was concerned that discussions with an AI agent might be less effective: *"A customer has very detailed questions, and the AI might be limited, not enabling the same type of discussion as with a human."* (P3). The following revision was made to the design concept to address the feedback from the participants:

Design revision 2: Sounds and voice variations, such as sighs, hums, giggles, laughter, and humor, were added to create a human-like impression. For further personalization, customers can now select the voice of the AI agent.

The participants had varying experiences with regard to what customers dislike when booking trips on their websites. Four participants highlighted situations in which their customers encountered problems during the booking process. In particular, P1 and P4 identified issues at the point of purchase decision or that the booking process involved too many steps. Participant P3 stated that customers often needed help in deciding which trip might be suitable. One participant emphasized the importance of simplifying the booking process, which was one of the main problems their customers experienced: "*It's usually about payment, so it's often more technical errors. Also, I know that our users might find our booking procedure too lengthy.*" (P1). Another participant had a similar experience: "*We've switched to a new platform, so now we have a problematic booking process, which isn't good.*" (P4). A third participant described where in the booking process their customers encountered problems: "*Yes, they find it a bit difficult to choose tours. There we have to help them quite a lot to see what might suit them.*" (P3). The third revision was formulated on the basis of this feedback:

Design revision 3: A one-click payment solution was included as an additional payment option for customers. This was intended to accelerate the booking process for those who desired this. Consequently, this could lead to improved post-purchase behavior among customers, which could benefit online travel agencies through the potential positive feedback they provide.

It is crucial for companies to continuously provide and receive customer feedback when something does not work as expected during the booking process. The absence of built-in feedback in a system can cause customers to stop using it due to frustration and irritation. This could lead to a negative spiral in which the company may lose revenue. One participant emphasized the importance of giving feedback immediately when the feeling arises, rather than waiting for another time: "*It's very professional to be able to do it right away, to express yourself if you have an unpleasant feeling. To be able to sort it out immediately instead of carrying that feeling and sending feedback later.*" (P7). Another participant also pointed out that voice might not always be a suitable input option: "*If you're using your phone, you might not be in a place where you can use voice, like when you're on the subway or on a bus.*" (P1). Accordingly, these comments formed the basis for the fourth revision, which highlighted the need to provide feedback even when not interacting using voice communication:

Design revision 4: Direct feedback options in the form of thumbs-up and down buttons were added to the design concept to capture customer feedback when voice is not used as an interaction option. More dissatisfied customers can be detected and addressed, allowing their complaints to be resolved.

Several participants emphasized the importance of understanding the complexity of their travel arrangements. This complexity ranged from various rules regarding food and seat booking on flights to considerations for people with special needs: "*There are many things to keep track of, different rules for different airplanes, which is important for the customer to know when they book. Complex trips with specific rules. One must consider everything, such as food on the plane, seating, etc. There could be people with special needs.*" (P2). One participant pointed out: "*We have to delve deeply into certain customers. How much experience does the customer have to handle certain horseback riding tours? Not everyone can ride all tours, and there are many questions to find the right trip.*" (P3). Another participant highlighted the importance of knowing specific information about certain destinations where safety and vaccinations are crucial: "*We have very complex trips with a lot of information and special cases to consider that must be accurate. I don't think an AI can handle that. If it's just about going to India and staying in a hotel, that's fine. But there are many other things to consider. You need to be vaccinated and it's not recommended to travel alone as a woman. It can be quite dangerous to travel to India as a solo woman.*" (P8). This feedback from the participants formed the basis for the fifth revision:

Design revision 5: An onboarding feature was introduced where the AI directly asks about users' specific preferences, such as whether they have mobility impairments or medical restrictions. User accounts are automatically generated on the basis of the customer input. Onboarding can be skipped if the user only seeks travel inspiration and does not intend to book a trip or already has an existing customer account.

With regard to the third theme and the complexity of AI systems, four participants noted that handling large datasets poses significant challenges for online travel agencies. It takes a long time for an AI system to learn how to manage all the data, the millions of trips that online travel agencies offer, and the specific rules involved. A participant pointed out the long learning period for AI systems: "*Long training period. It's also about how quickly they learn things.*" (P1). Participant P4 stated that there are "*Big challenges with large datasets.*" (P4). Another participant was concerned about all the data that need to be managed for the millions of trips offered by online travel agencies: "*There's going to be a lot of background data that needs to be correlated.*" (P2). During the interview process, participant P6 identified a variety of obstacles to implementing the AI system in their organization and a desire to skip a few steps: "*High threshold for implementation in current systems. Overhauling our entire information management system. In our company, of course, one could have connected and skipped such a step and gone straight into some kind of AI-driven customer service thing.*" (P6). Hence, the design concept must address travel companies' data management needs through the following revision:

Design revision 6: Predictive analytics, with a machine learning algorithm, is used to identify customer trends and patterns from historical data, such as past bookings, travel plans, and preferences. This allows an AI system to effectively predict customer behavior, optimize pricing, and tailor recommendations to customer needs.

Another participant also mentioned the lack of options for customers to choose their type of interaction, such as connecting to a customer service representative when needed: "*We have older customers who book. They want that personal contact. They might be a bit restrictive. Some might want direct contact where they felt the information is not sufficient. To then be connected to a real person.*" (P3). A third participant had similar experiences with older people preferring more personal contact than talking to a system: "*The younger generation fits this because they don't want to talk to anyone on the phone. They just want to book their trip and be done. While the older generation prefers more personal contact.*" (P5). Taking into account the feedback of the participants, the following revision was made:

Design revision 7: The AI agent needs to be supported by multimodal options to ensure that its behavior is accurate and that inputs are understood. Such options enable customers to make choices that fit their context and personalities. There needs to be a way to communicate directly in a booking process through various options, such as voice and text. It should also be possible to contact a customer service representative via a button or voice command. The voice assistant should speak clearly and calmly, not too quickly, as some customers could need help keeping up with fast speech.

The seven design revisions previously described were integrated into a revised design concept, see Figure 3.

4.4. Concept Contextualization

The revised design concept is contextualized in the following section by relating it to similar concepts, notions, and theories within the existing field. Currently, booking a trip using voice interaction on most online travel agencies' websites is not available as a technology. Consequently, the design concept of VoiceBack contributes to the field of informatics by introducing this novel interaction. VoiceBack's AI agent can be customized by a customer to become personal and change the voice, name, tone, and dialect at any

point during the interaction. Due to VoiceBack's ability to immediately detect feedback from emotions in the voice, it can implement real-time measures to prevent frustration and ensure that a customer does not abandon the booking process.

The VoiceBack concept draws on existing theories about AI, VUI, and user feedback. It could be implemented in existing systems, where customer experience can be measured through voice-based feedback in VUIs [12]. The goal is to gain customer loyalty by providing high-quality services and satisfying their needs [38]. The design of a voice assistant must meet the expectations and needs of customers [68]. It is challenging to achieve a natural flow between customers and the current VUIs [22]. VoiceBack's novel concept optimizes the benefits of voice interaction. It incorporates emotional AI that recognizes and analyzes emotions from the voice of customers to pinpoint problems in the booking process. This provides invaluable insights into the reasons behind customer actions, such as cancellations or incomplete bookings, which can be used to improve the overall booking experience.

Customers often find that voice assistants do not understand voice input correctly, which can lead to frustration and abandonment. To address this issue, VoiceBack theoretically implemented language models that interpret and communicate in various dialects, to ensure that customers feel included in conversations with an AI agent. This enhancement allows customers to configure VoiceBack based on their preferences. For example, it allows them to dictate an agent's behavior, select their voice, and even determine whether they should possess a sense of humor. Such personalization can foster greater trust in a system and improve the user experience [70].

Currently, customer feedback is collected through crowdsourced voice feedback (CVF) [46] and explicit feedback [12,43]. For example, online travel agencies collect customer feedback through online forms, pop-up surveys, social media, and customer reviews. VoiceBack has potential as a future voice-based technology, because it can detect feedback in customers' voices using emotional recognition techniques such as sentiment analysis and adaptive learning. With adaptive learning, VoiceBack can continuously and autonomously improve its ability to handle and recognize feedback. Over time, it can learn from customer interactions and provide better responses than existing voice assistants.

Customer dissatisfaction with the websites of online travel agencies could be attributed to the complexity of booking a trip [38]. The number of trips offered by some online travel agencies, with varying travel operators, dates, and routes, adds to this complexity. In addition, the booking process becomes even more complicated when booking a combined trip, such as flights and hotels. The complexity increases further when travelers' particular needs and preferences are considered. Predictive analysis was added to the revised design concept to address these issues. VoiceBack offers advanced technology to facilitate more accurate analyses throughout the booking process to support customers and online travel agencies in complex situations.

With VoiceBack, online travel agencies can detect and address the challenges customers face in terms of website functionality. These include the usability of the interface, the frustration of finding information, the difficulty of understanding content, overload, navigation deficiencies, visual layout, and online help in real time [15]. VoiceBack's provision of real-time feedback sets it apart from other feedback methods. It can allow travel agencies to assist dissatisfied customers, improve customer satisfaction, and consequently increase revenue. Providing feedback after an interaction can change the initial impression of customers. An advantage of implementing VoiceBack compared to existing travel booking methods is that the AI agent rarely misses crucial information that a human might overlook, as an AI system is adept at handling large amounts of data compared to traditional customer service [40].

5. Discussion

Voice communication is a fundamental human interaction method, and its importance in customer–system interactions has been well documented [5–7]. Enabling customers to use voice interaction as a viable alternative to text input in the booking process offers several

advantages. These include saving time, freeing hands and sight for other tasks, and improving customer trust in AI systems [4]. Furthermore, research has attributed customer dissatisfaction with online travel bookings to poor service quality [38]. This dissatisfaction is often exacerbated by poorly designed or non-user-friendly interfaces, which hinder effective interactions and compromise the overall user experience of online travel services.

Based on previous reasoning, this article explored, conceptualized, and evaluated the design of VoiceBack, a voice-based AI agent for online travel agencies, to support customers during the booking process. The VoiceBack concept contributes to the fields of artificial intelligence, voice interaction, and HCI theory in several ways. It introduces a theoretically and empirically grounded design concept for a voice-based, AI-powered feedback system, while also presenting empirical evidence from online travel agencies to illustrate the practical challenges of implementation. By exploring the intersection of voice user interfaces, emotion recognition, and AI-driven customer service, VoiceBack demonstrates an approach to theory development that bridges theoretical concepts with practical design considerations. This aligns with the growing trend of developing sophisticated and context-sensitive AI systems [20,22] that aim to understand and respond to user needs.

The integration of VUIs into various contexts, including customer support, car assistants, and smart devices [4,8,29], underscores the growing importance of voice-based interactions. This trend aligns with the fundamental role of the voice in human communication [5,6], making it a natural choice to improve customer–agency interactions in the travel booking process. VoiceBack builds on the current AI-driven customer service systems in the travel industry by integrating real-time emotion recognition and adaptive responses. Although existing systems, such as chatbots, offer text-based support, VoiceBack's voice interface provides a more natural interaction medium. However, it is essential to note that current systems have advantages in terms of established infrastructure and user familiarity. VoiceBack's success will depend on overcoming these incumbent advantages through an enhanced user experience and demonstrable efficiency gains.

The implementation of VoiceBack has several practical hurdles that warrant careful consideration. Integration with existing travel systems requires significant technical efforts and cooperation from online travel agencies. Data privacy concerns must be carefully considered, particularly with regard to the collection and analysis of voice data and emotional states. The current limitations in AI and voice recognition technologies, such as the difficulty in understanding diverse accents and dialects, as highlighted by our participants, could impact VoiceBack's performance and user acceptance. These challenges echo the concerns raised in previous research on the complexity of NLU [21,22] and the need for AI systems to grasp semantics and social communication patterns.

The implications of VoiceBack can potentially foster more engaging interactions with customers and increase customer satisfaction and revenue for online travel agencies [40]. Additionally, VoiceBack can enable online travel agencies to pinpoint specific problem areas within a booking process and implement targeted solutions.

Although this article aims to theoretically contribute to academic discourse, it is essential to acknowledge a key limitation: the lack of practical application and real-life interaction with VoiceBack. Future research should focus on developing a functional VoiceBack prototype for testing in simulated travel booking scenarios. This should be followed by pilot studies with travel agencies, to assess real-world performance using empirical metrics such as task completion rate, customer satisfaction score, emotion recognition accuracy, and revenue impact. Such an evaluation could also provide valuable information on user acceptance and system effectiveness. Moreover, exploring the application of similar voice-based and emotion-aware AI systems in other service industries could produce interesting comparative data and extend the work on emotional AI and its applications [23,24].

Another direction for future work could be to focus on understanding customer interactions using customizable interfaces. How do customers perceive and engage with adaptive and changing interfaces? What are the nuances of such interactions and how can they enhance the customer experience? These questions align with ongoing research on

the design of natural and intuitive voice user interfaces [31,35]. A third direction involves implementing and evaluating VoiceBack in other settings, such as retail e-commerce environments, where it could transform how customers shop for clothes and offer them a level of personalization that was previously unattainable.

As AI technologies continue to advance and become more integrated into our daily lives, there is a need to collaboratively develop guidelines for their responsible implementation. These guidelines should address critical issues such as data privacy, informed consent for emotion analysis, and strategies to mitigate potential AI bias in response generation.

6. Conclusions

The revised theoretically and empirically underpinned design concept of VoiceBack contributes cumulatively to the promotion and expansion of the integration of AI technologies and voice-based interfaces in online travel services. VoiceBack presents a modest step towards enhancing customer–agency communication and supporting the future design of booking systems that can adapt to individual customer needs and emotions in real time. Its successful implementation requires addressing several technical, practical, and ethical challenges. The evaluation of VoiceBack provides insights into the specific challenges that customers can face during the booking process on the websites of online travel agencies. These insights could be valuable to VUI designers and developers to guide them in creating practical and user-friendly VUI solutions that address real-world customer needs. As AI continues to evolve, concepts such as VoiceBack may become integral in shaping more empathetic and efficient digital service experiences across various industries.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
CV	Customer experience
CVF	Crowd-sourced voice feedback
GUI	Graphical user interface
HCI	Human–computer interaction
NLP	Natural language processing
NLU	Natural language understanding
OTA	Online travel agency
VUI	Voice user interface
vHAI	Voice-based human–agent interaction

Appendix A. Interview Questions

Appendix A.1. Demographics

- Age
- Gender (female, male, non-binary, other, do not wish to disclose)

Appendix A.2. Design Concept

- What advantages and disadvantages do you identify in the design concept?
- What challenges or practical issues do you see in the design concept?
- What features would you like to add, remove, or change in the design concept?
- How could the design concept improve the user experience when booking trips through online travel agencies?

Appendix A.3. Artificial Intelligence

- What advantages and disadvantages do you see in an AI-based service that extracts emotions from voice?
- What ethical issues do you see with using an AI-based service that extracts emotions from voice?
- What features would you like an AI-driven intelligent customer service agent to have that they currently do not?
- To what extent would you like the concept to exhibit human characteristics in an AI agent?
- Would your company consider using an AI agent?
- What is your stance on the reliability of AI-generated feedback?

Appendix A.4. Voice User Interface

- What experience do you have with systems that use a voice user interface (VUI)?
- How inclined do you think customers are to use voice as an interaction option during a booking process?
- In what way would you personally like to use voice as an interaction option during a booking process?
- Are you currently using a voice assistant for customer service in your company? If so, which one?
- Have you ever discussed the implementation of a voice assistant at your company?

Appendix A.5. Feedback

- How do you currently collect digital feedback from your customers?
- What are the most common problems your customers encounter during a booking process that you receive feedback on?
- If you were to receive customer feedback, how would you like it presented to you in an interface?
- Is there anything in the visual presentation of feedback that you would like to add, remove, or change?

References

1. Slonim, N.; Bilu, Y.; Alzate, C.; Bar-Haim, R.; Bogin, B.; Bonin, F.; Choshen, L.; Cohen, E.; Dankin, L.; Edelstein, L.; et al. An autonomous debating system. *Nature* **2021**, *591*, 379–384. [CrossRef] [PubMed]
2. Thirumuruganathan, S.; Kunjir, M.; Ouzzani, M.; Chawla, S. Automated annotations for AI data and model transparency. *J. Data Inf. Qual.* **2021**, *14*, 1–9. [CrossRef]
3. Wei, J.; Tag, B.; Trippas, J.R.; Dingler, T.; Kostakos, V. What could possibly go wrong when interacting with proactive smart speakers? A case study using an ESM application. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 29 April–5 May 2022; Barbosa, S., Lampe, C., Appert, C., Shamma, D.A., Drucker, S., Williamson, J., Yatani, K., Eds.; Association for Computing Machinery: New York, NY, USA, 2022. [CrossRef]
4. Rzepka, C. Examining the use of voice assistants: A value-focused thinking approach. In Proceedings of the 25th Americas Conference on Information Systems (AMCIS 2019), Cancun, Mexico, 15–17 August 2019; Rodriguez-Abitia, G., Ferran, C., Eds.; Association for Information Systems: Atlanta, GA, USA, 2019; p. 20.

5. Chamishka, S.; Madhavi, I.; Nawaratne, R.; Alahakoon, D.; De Silva, D.; Chilamkurti, N.; Nanayakkara, V. A voice-based real-time emotion detection technique using recurrent neural network empowered feature modelling. *Multimed. Tools Appl.* **2022**, *81*, 35173–35194. [CrossRef]

6. Seaborn, K.; Miyake, N.P.; Pennefather, P.; Otake-Matsuura, M. Voice in human-agent interaction: A survey. *ACM Comput. Surv.* **2021**, *54*, 1–43. [CrossRef]

7. Guglielmi, E.; Rosa, G.; Scalabrino, S.; Bavota, G.; Oliveto, R. Sorry, I don't understand: Improving voice user interface testing. In Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering, Rochester, MI, USA, 10–14 October 2022; Kessentini, M., Ed.; Association for Computing Machinery: New York, NY, USA, 2023. [CrossRef]

8. Kim, J.; Kim, W.; Nam, J.; Song, H. "I can feel your empathic voice": Effects of nonverbal vocal cues in voice user interface. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; Bernhaupt, R., Mueller, F., Verweij, D., Andres, J., Eds.; Association for Computing Machinery: New York, NY, USA, 2020. [CrossRef]

9. Zaheer, N.; Ahmad, O.U.; Ahmed, A.; Khan, M.S.; Shabbir, M. SEMOUR: A scripted emotional speech repository for Urdu. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; Kitamura, Y., Quigley, A., Eds.; Association for Computing Machinery: New York, NY, USA, 2021. [CrossRef]

10. Geng, J.; Zhu, H.; Li, X.Y. SeeSpeech: See emotions in the speech. In Proceedings of the 2nd Asia Service Sciences and Software Engineering Conference, Macau, China, 24–26 February 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 116–122. [CrossRef]

11. Sarikaya, R. Intelligent conversational agents for ambient computing. In Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, 11–15 July 2022; Amigo, E., Castells, P., Gonzalo, J., Eds.; Association for Computing Machinery: New York, NY, USA, 2022; p. 5. [CrossRef]

12. Deng, Y.; Murari, S. When a voice assistant asks for feedback: An empirical study on customer experience with A/B testing and causal inference methods. In Proceedings of the 2021 International Conference on Multimodal Interaction, Montreal, QC, Canada, 18–22 October 2021; Hammal, Z., Busso, C., Pelachaud, C., Oviatt, S., Salah, A.A., Zhao, G., Eds.; Association for Computing Machinery: New York, NY, USA, 2021; pp. 183–191. [CrossRef]

13. Myers, C.; Furqan, A.; Nebolsky, J.; Caro, K.; Zhu, J. Patterns for how users overcome obstacles in voice user interfaces. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; Mandryk, R., Hancock, M., Eds.; Association for Computing Machinery: New York, NY, USA, 2018. [CrossRef]

14. Murad, C.; Munteanu, C. "I don't know what you're talking about, HALexa": The case for voice user interface guidelines. In Proceedings of the CUI '19: Proceedings of the 1st International Conference on Conversational User Interfaces, Dublin, Ireland, 22–23 August 2019; Cowan, B.R., Clark, L., Eds.; Association for Computing Machinery: New York, NY, USA, 2019. [CrossRef]

15. Huang, Z.; Mou, J. Gender differences in user perception of usability and performance of online travel agency websites. *Technol. Soc.* **2021**, *66*, 101671. [CrossRef]

16. Nasution, F.I.H.; Khairunnisa, A.K.; Menawan, A.A.; Gobel, L.; Ardi, F.; Ulkhaq, M.M. Assessing website quality of online travel agency: An application of the importance-performance and gap analysis model. In Proceedings of the 3rd International Conference on Digital Technology in Education, Yamanashi, Japan, 25–27 October 2019; Evans, H., Dua, Y., Eds.; Association for Computing Machinery: New York, NY, USA, 2020; pp. 176–180. [CrossRef]

17. Talwar, S.; Dhir, A.; Kaur, P.; Mäntymäki, M. Why do people purchase from online travel agencies (OTAs)? A consumption values perspective. *Int. J. Hosp. Manag.* **2020**, *88*, 102534. [CrossRef]

18. Stolterman, E.; Wiberg, M. Concept-driven interaction design research. *Hum.-Comput. Interact.* **2010**, *25*, 95–118. [CrossRef]

19. McCarthy, J. What Is Artificial Intelligence? 2004. Available online: https://www-formal.stanford.edu/jmc/whatisai.pdf (accessed on 28 July 2024).

20. Gams, M.; Gu, I.Y.H.; Härmä, A.; Muñoz, A.; Tam, V. Artificial intelligence and ambient intelligence. *J. Ambient Intell. Smart Environ.* **2019**, *11*, 71–86. [CrossRef]

21. Collins, C.; Dennehy, D.; Conboy, K.; Mikalef, P. Artificial intelligence in information systems research: A systematic literature review and research agenda. *Int. J. Inf. Manag.* **2021**, *60*, 102383. [CrossRef]

22. Varitimiadis, S.; Kotis, K.; Pittou, D.; Konstantakis, G. Graph-based conversational AI: Towards a distributed and collaborative multi-chatbot approach for museums. *Appl. Sci.* **2021**, *11*, 9160. [CrossRef]

23. Peter, M.; Manh-Tung, H. Why we need to be weary of emotional AI. *AI Soc.* **2022**, *39*, 1447–1449. [CrossRef]

24. Pamungkas, E.W. Emotionally-aware chatbots: A survey. *arXiv* **2019**, arXiv:1906.09774. https://doi.org/10.48550/arXiv.1906.09774.

25. Hu, P.; Lu, Y.; Gong, Y. Dual humanness and trust in conversational AI: A person-centered approach. *Comput. Hum. Behav.* **2021**, *119*, 106727. [CrossRef]

26. Aronsson, J.; Lu, P.; Strüber, D.; Berger, T. A maturity assessment framework for conversational AI development platforms. In Proceedings of the 36th Annual ACM Symposium on Applied Computing, Virtual Event, 22–26 March 2021; Hung, C.C., Hong, J., Eds.; Association for Computing Machinery: New York, NY, USA, 2021; pp. 1736–1745. [CrossRef]

27. Edwards, J.; Sanoubari, E. A need for trust in conversational interface research. In Proceedings of the 1st International Conference on Conversational User Interfaces, Dublin, Ireland, 22–23 August 2019; Cowan, B.R., Clark, L., Eds.; Association for Computing Machinery: New York, NY, USA, 2019. [CrossRef]

28. Liao, V.Q.; Gruen, D.; Miller, S. Questioning the AI: Informing design practices for explainable ai user experiences. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; Bernhaupt, R., Mueller, F., Verweij, D., Andres, J., Eds.; Association for Computing Machinery: New York, NY, USA, 2020. [CrossRef]

29. Myers, C.; Furqan, A.; Zhu, J. The impact of user characteristics and preferences on performance with an unfamiliar voice user interface. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; Brewster, S., Fitzpatrick, G., Eds.; Association for Computing Machinery: New York, NY, USA, 2019. [CrossRef]

30. Ma, Y.; Drewes, H.; Butz, A. Fake moods: Can users trick an emotion-aware voicebot? In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Online, 8–13 May 2021; Kitamura, Y., Quigley, A., Isbister, K., Igarashi, T., Eds.; Association for Computing Machinery: New York, NY, USA, 2021. [CrossRef]

31. Kim, Y.; Reza, M.; McGrenere, J.; Yoon, D. Designers characterize naturalness in voice user interfaces: Their goals, practices, and challenges. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Online, 8–13 May 2021; Kitamura, Y., Quigley, A., Eds.; Association for Computing Machinery: New York, NY, USA, 2021. [CrossRef]

32. Poran, S.; Amsalem, G.; Beka, A.; Goldenberg, D. With one voice: Composing a travel voice assistant from repurposed models. In Proceedings of the Web Conference 2022, Lyon, France, 25–29 April 2022; Laforest, F., Troncy, R., Médini, L., Herman, I., Eds.; Association for Computing Machinery: New York, NY, USA, 2022; pp. 383–387. [CrossRef]

33. Cha, Y.; Jang, J.; Hong, Y.; Yi, M.Y. "Jack-of-All-Trades": A thematic analysis of conversational agents in multi-device collaboration contexts. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Scotland, UK, 4–9 May 2019; Brewster, S., Fitzpatrick, G., Eds.; Association for Computing Machinery: New York, NY, USA, 2019. [CrossRef]

34. Myers, C.; Pardo, L.F.L.; Acosta-Ruiz, A.; Canossa, A.; Zhu, J. "Try, try, try again": Sequence analysis of user interaction data with a voice user interface. In Proceedings of the 3rd Conference on Conversational User Interfaces, Bilbao, Spain, 27–29 July 2021; Schlögl, S., Ed.; Association for Computing Machinery: New York, NY, USA, 2021. [CrossRef]

35. Rzepka, C.; Berger, B.; Hess, T. Why another customer channel? Consumers' perceived benefits and costs of voice commerce. In Proceedings of the 53rd Hawaii International Conference on System Sciences, Honolulu, HI, USA, 7–10 January 2020; Bui, T., Ed.; University of Hawai'i at Mānoa Library: Maui, HI, USA, 2020; pp. 4079–4088. [CrossRef]

36. Bender, E.M.; Koller, A. Climbing towards NLU: On meaning, form, and understanding in the age of data. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, 5–10 July 2020; Jurafsky, D., Chai, J., Schluter, N., Tetreault, J., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2020; pp. 5185–5198. [CrossRef]

37. Chavan, K.; Gawande, U. Speech recognition in noisy environment, issues and challenges: A review. In Proceedings of the 2015 International Conference on Soft-Computing and Networks Security (ICSNS), Coimbatore, India, 25–27 February 2015. [CrossRef]

38. Putri, N.N.S.; Alamsyah, A.; Widiyanesti, S. Service quality analysis of online travel agencies (OTA) using multiclass classification. *Almana J. Manaj. Dan Bisnis* **2020**, *4*, 72–78. [CrossRef]

39. Chen, J.S.; Kamalanon, P.; Janupiboon, T.P. Company websites and mobile apps versus social media: Which service experience creates more customer value for online travel agencies? *Serv. Bus.* **2022**, *16*, 1081–1110. [CrossRef]

40. Yu, S.; Qu, X. Research on the influence of artificial intelligence customer service on consumers' purchase process under the background of Internet. In Proceedings of the 13th International Conference on e-Business, Management and Economics, Beijing, China, 16–18 June 2022; Yin, L.R., Ed.; Association for Computing Machinery: New York, NY, USA, 2022; pp. 137–141. [CrossRef]

41. Zhu, Z.; Kim, S.; Shin, H.H. Does the website credibility of online travel agencies (OTAs) affect customers' evaluation and future intentions? *Asia Pac. J. Tour. Res.* **2022**, *27*, 842–855. [CrossRef]

42. Liu, X.; Xu, A.; Sinha, V.; Akkiraju, R. Voice of customer: A tone-based analysis system for online user engagement. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; Mandryk, R., Hancock, M., Eds.; Association for Computing Machinery: New York, NY, USA, 2018. [CrossRef]

43. Xiao, Z.; Mennicken, S.; Huber, B.; Shonkoff, A.; Thom, J. Let me ask you this: How can a voice assistant elicit explicit user feedback? *Proc. ACM Hum.-Comput. Interact.* **2021**, *5*, 1–24. [CrossRef]

44. Wang, J.; Li, Y.; Yang, S.; Dong, S.; Li, J. Waiting experience: Optimization of feedback mechanism of voice user interfaces based on time perception. *IEEE Access* **2023**, *11*, 21241–21251. [CrossRef]

45. Alsabhan, W. Human-computer interaction with a real-time speech emotion recognition with ensembling techniques 1D convolution neural network and attention. *Sensors* **2023**, *23*, 1386. [CrossRef]

46. Ertay, E.; Huang, H.; Sarsenbayeva, Z.; Dingler, T. Challenges of emotion detection using facial expressions and emotion visualisation in remote communication. In Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2021 ACM International Symposium on Wearable Computers, Virtual, 21–26 September 2021; Doryab, A., Lv, Q., Beigl, M., Eds.; Association for Computing Machinery: New York, NY, USA, 2021; pp. 230–236. [CrossRef]

47. Gupta, S.; Gupta, S. Investigating emotion—Color association in deep neural networks. *arXiv* **2020**, arXiv:2011.11058. https://doi.org/10.48550/arXiv.2011.11058.

48. Park, E.H.; Storey, V.C. Emotion ontology studies: A framework for expressing feelings digitally and its application to sentiment analysis. *ACM Comput. Surv.* **2023**, *55*, 1–38. [CrossRef]

49. Bashir, M.F.; Javed, A.R.; Arshad, M.U.; Gadekallu, T.R.; Shahzad, W.; Beg, M.O. Context-aware emotion detection from low-resource Urdu language using deep neural network. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* **2023**, *22*, 1–30. [CrossRef]

50. Klein, A.M. Toward a user experience tool selector for voice user interfaces. In Proceedings of the 18th International Web for All Conference, Online, 19–20 April 2021; Vazquez, S.R., Drake, T., Eds.; Association for Computing Machinery: New York, NY, USA, 2021. [CrossRef]

51. Johansson, D.; Wiberg, M. Conceptually advancing "application mobility" towards design: Applying a concept-driven approach to the design of mobile IT for home care service groups. *Int. J. Ambient Comput. Intell.* **2012**, *4*, 20–32. [CrossRef]

52. Johansson, D.; Lassinantti, J.; Wiberg, M. Mobile e-services and open data in e-government processes—Concept and design. In Proceedings of the Mobile Web and Intelligent Information Systems, Rome, Italy, 24–26 August 2015; Younas, M., Awan, I., Mecella, M., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Lecture Notes in Computer Science; pp. 149–160. [CrossRef]

53. Li, Z.; Wang, Y.; Andres, J.; Semertzidis, N.; Greuter, S.; Mueller, F. A design framework for ingestible play. *ACM Trans. Comput.-Hum. Interact.* **2023**, *30*, 1–39. [CrossRef]

54. Lövqvist, C.; Pinter, M.; Faraon, M.; Villavicencio, V. Crafting tasteful experiences: Designing artificial intelligence and voice user interfaces for home delivery contexts. In Proceedings of the 5th Artificial Intelligence and Cloud Computing Conference, Osaka, Japan, 17–19 December 2022; Ozawa, S., Ed.; Association for Computing Machinery: New York, NY, USA, 2023; pp. 175–189. [CrossRef]

55. Svenningsson, N.; Faraon, M.; Villavicencio, V. Assessment of theses in design education: Conceptualizing a co-creative process for grading by automated criteria evaluation. *Int. J. Innov. Teach. Learn. High. Educ. (IJITLHE)* **2021**, *2*, 1–17. [CrossRef]

56. Faraon, M.; Jaff, A.; Nepomuceno, L.P.; Villavicencio, V. Fake news and aggregated credibility: Conceptualizing a co-creative medium for evaluation of sources online. *Int. J. Ambient Comput. Intell. (IJACI)* **2020**, *11*, 93–117. [CrossRef]

57. Morse, J.M. The significance of saturation. *Qual. Health Res.* **1995**, *5*, 147–149. [CrossRef]

58. Guest, G.; Bunce, A.; Johnson, L. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* **2006**, *18*, 59–82. [CrossRef]

59. Hennink, M.; Kaiser, B.N. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Soc. Sci. Med.* **2022**, *292*, 114523. [CrossRef]

60. Svenningsson, N.; Faraon, M. Artificial intelligence in conversational agents: A study of factors related to perceived humanness in chatbots. In Proceedings of the 2nd Artificial Intelligence and Cloud Computing Conference, Kobe, Japan, 21–23 December 2019; Ozawa, S., Ed.; Association for Computing Machinery: New York, NY, USA, 2019; pp. 151–161. [CrossRef]

61. Clarke, V.; Braun, V. Thematic analysis. *J. Posit. Psychol.* **2017**, *12*, 297–298. [CrossRef]

62. Stafström, S. *Good Research Practice*; Swedish Research Council: Stockholm, Sweden, 2017.

63. Abdolrahmani, A.; Gupta, M.H.; Vader, M.L.; Kuber, R.; Branham, S. Towards more transactional voice assistants: Investigating the potential for a multimodal voice-activated indoor navigation assistant for blind and sighted travelers. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Online, 8–13 May 2021; Bjørn, P., Drucker, S., Eds.; Association for Computing Machinery: New York, NY, USA, 2021. [CrossRef]

64. Liu, Y.; Wang, L.; Kearns, W.R.; Wagner, L.; Raiti, J.; Wang, Y.; Yuwen, W. Integrating a voice user interface into a virtual therapy platform. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Online, 8–13 May 2021; Association for Computing Machinery: New York, NY, USA, 2021. [CrossRef]

65. Fahn, V.; Riener, A. Time to get conversational: Assessment of the potential of conversational user interfaces for mobile banking. In Proceedings of the Mensch und Computer 2021, Ingolstadt, Germany, 5–8 September 2021; Schneegass, S., Pfleging, B., Kern, D., Eds.; Association for Computing Machinery: New York, NY, USA, 2021; pp. 34–43. [CrossRef]

66. Xue, L.; Soong, F.K.; Zhang, S.; Xie, L. ParaTTS: Learning linguistic and prosodic cross-sentence information in paragraph-based TTS. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2022**, *30*, 2854–2864. [CrossRef]

67. Meyer, S.; Elsweiler, D.; Ludwig, B.; Fernandez-Pichel, M.; Losada, D.E. Do we still need human assessors? Prompt-based GPT-3 user simulation in conversational AI. In Proceedings of the 4th Conference on Conversational User Interfaces, Glasgow, UK, 26–28 July 2022; Halvey, M., Foster, M.E., Dalton, J., Munteanu, C., Trippas, J., Eds.; Association for Computing Machinery: New York, NY, USA, 2022. [CrossRef]

68. Branch, J.; Parker, C.J.; Evans, M. Do user experience (UX) design courses meet industry's needs? Analysing UX degrees and job adverts. *Des. J. Int. J. All Asp. Des.* **2021**, *24*, 631–652. [CrossRef]

69. Nielsen, J. 10 Usability Heuristics for User Interface Design. 1994. Available online: https://www.nngroup.com/articles/ten-usability-heuristics/ (accessed on 28 July 2024).

70. Murtarelli, G.; Gregory, A.; Romenti, S. A conversation-based perspective for shaping ethical human–machine interactions: The particular challenge of chatbots. *J. Bus. Res.* **2021**, *129*, 927–935. [CrossRef]