Check for updates

Wiley
Human Behavior and Emerging Technologies
Volume 2025, Article ID 1205909, 21 pages
https://doi.org/10.1155/hbe2/1205909

WILEY

Review Article

Natural Language Processing–Based Technologies Along the Customer Journey—A Systematic Review and Co-Occurrence Analysis

Tom Ferber, Daryoush Vaziri, and Alexander Boden

Department of Economics, Bonn-Rhein-Sieg University of Applied Sciences, Sankt Augustin, Germany

Correspondence should be addressed to Tom Ferber; tom.ferber@smail.wis.h-brs.de

Received 21 October 2024; Accepted 29 May 2025

Academic Editor: Sanmugam Annamalah

Copyright © 2025 Tom Ferber et al. Human Behavior and Emerging Technologies published by John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Interactions between consumers and companies are increasingly relying on technologies such as chatbots and voice assistants that are based on natural language processing (NLP) techniques. With the advent of more sophisticated technologies such as transformers and generative artificial intelligence, this trend will likely continue and further solidify. To our knowledge, this study is the first to systematically review the current scientific discourse on NLP-based technologies in the context of the customer journey and attempts to outline existing knowledge and identify gaps before the onset of a new era in NLP sophistication. Employing the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method and co-occurrence analysis, we offer new and nuanced insights into the prevailing discourse. From a sample of 734 articles, 41 studies were selected and analyzed. Our findings shed light on the current research focus, exploring various technologies, concepts, and challenges. We also offer a starting point for how emerging NLP-based technologies could impact the customer journey, as well as future research directions.

Keywords: chatbots; customer journey; generative AI; marketing; natural language processing

1. Introduction

The natural language processing (NLP) market is steadily gaining traction and is expected to increase in value from $29.1 billion in 2023 to $92.7 billion by 2028, reflecting a compound annual growth rate of 26.1%, according to a market analysis report by Singh [1]. This expansion is driven by increasing consumer adoption of NLP-based technologies, such as chatbots and the widespread integration of these tools into business operations [2]. As organizations seek to enhance customer engagement and streamline interactions, NLP-powered applications are becoming a core component of the customer journey. By 2027, chatbots are expected to serve as the primary customer service channel for around a quarter of companies [3].

NLP, as a subset of artificial intelligence (AI), is a generic term for technologies concerned with the handling of human languages [4]. Unlike other AI domains, such as computer vision, which interprets visual data, or machine learning-based analytics, which identifies patterns in structured data, NLP is specifically concerned with linguistic data. It encompasses various subareas such as speech recognition, which enables machines to interact with humans in a more natural and context-aware way. Thus, we define NLP-based technologies as applications whose core functionalities stem from this field of research, particularly in processing and generating human language. These systems must incorporate at least one core NLP functionality—such as language understanding or language generation—and are, in the case of this study, primarily deployed at customer touchpoints.

NLP-based technologies include, for example, chatbots, which simulate human conversation by analyzing user inputs and producing text-based responses, or voice assistants, which use NLP in combination with speech recognition and synthesis to facilitate hands-free interactions.

Emerging and more powerful technologies, such as large language models (LLMs), represent a significant advancement and also fall under this category. However, unlike traditional NLP models, which rely on predefined rules or statistical methods for specific language tasks, LLMs are advanced models based on transformer technology, featuring unprecedented parameter sizes and learning capabilities that enable the analysis and generation of new texts [5]. While traditional diffusion theories (e.g., [6]) emphasize early adopters as key drivers, an LLM like ChatGPT reaching the mark of 100 million monthly active users in 2 months after launch [7] suggests that its appeal has swiftly extended beyond early adopters to a broader, mainstream audience. These new technologies therefore seem to accelerate market development even further and transform how customers interact with companies throughout their journey. Given this development, it is essential to explore and analyze the limitations of traditional NLP models and investigate how advanced models like LLMs can be beneficially implemented to enhance customer interactions.

Despite the clear relevance and timeliness of this topic, there is, to our knowledge, no comprehensive overview of customer interactions with various NLP-based technologies across the customer journey—only reviews on adjacent topics exist (e.g. [8–11]). The customer journey is a framework rooted in customer experience management [12] that describes the various phases a customer experiences and thus encompasses the entirety of the experiences that a customer has when interacting with a company or brand [11]. A structured analysis of these interactions would be highly valuable, offering practitioners insights into current approaches for implementing these technologies and their impact on customer behavior and thus address a critical knowledge gap in a rapidly growing market poised for significant expansion in the coming years. The current absence of a systematic overview on the specific topic of NLP-based technologies in conjunction with the customer journey framework may limit researchers' and practitioners' ability to fully capture the nuances of how these technologies impact customer behavior and could hinder the ability to project the possibilities and create advanced customer journeys.

To address this knowledge gap, we propose the following research questions:

RQ1: How is the current state of research on NLP-based technologies distributed across the different stages of the customer journey in terms of research quantity and technological focus, and what underlying technological trends or limitations characterize this distribution?

RQ2: How do key concepts identified in our curated database affect customer behavior across different stages of the customer journey, and how are these concepts distributed and interrelated within these stages?

Given these research questions, from both an academic and practical standpoint, understanding their implications is essential for ensuring the effective integration of new technologies and the creation of satisfying customer experiences, making this a critical area of study.

We have taken this development as an opportunity to address these challenges by performing a systematic literature review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology, along with a co-occurrence analysis. We aim to analyze the previous scientific discourse regarding NLP-based technologies along all phases of the customer journey, namely, the prepurchase, purchase, and postpurchase phases. Specifically, the objectives and contributions of this research are to (1) address the current lack of overview by identifying and categorizing specific gaps in the existing literature on NLP-based technologies and their interplay with customers to provide actionable insights guiding future research directions, (2) propose a structured framework that synthesizes current knowledge and highlights key concepts and their impact that must be considered when designing innovative customer interactions across different journey stages, and (3) offer a multidisciplinary perspective that bridges management sciences and technical research fields, offering actionable insights into how NLP-based technologies influence different facets of consumer decision-making and behavior—from initial rapport building to increasing purchase intention. The significance of this research lies in its potential to serve as a foundational reference for researchers and practitioners navigating this rapidly evolving landscape, enabling the development of innovative and effective customer interactions powered by NLP-based technologies.

2. Related Literature Reviews

AI is a major point of discussion in scientific discourse across disciplines. Therefore, it is not surprising that various authors have engaged in the preparation of literature reviews on this topic from different viewpoints. Haleem et al. [13], for example, explore the topic from a general marketing perspective, noting that AI can create a more personalized brand experience, improve user experience, and optimize campaigns and strategies by analyzing user data. The authors investigate how AI can be applied to different marketing domains like pricing, product management, promotion management, or strategy and planning. Besides mentioning the growing importance of NLP, the report remains relatively broad, without focusing on a specific subset of AI or a particular technology. Mustak et al. [14] add to this with their article highlighting key research topics related to AI in marketing, including understanding consumer sentiment, analyzing customer satisfaction, and improving market performance. These areas are significant for the application of NLP-based technologies, yet the review falls short in this area. Instead of providing a detailed exploration, it offers only a broad overview by mentioning that big data could be leveraged. The role of the customer journey is not discussed. Instead, the authors analyze the literature using software such as VOSviewer to examine the major themes of AI research in marketing since the 1960s and find

that these can be divided into two categories: consumer research or organization and strategy, giving general insights. Other reviews commit themselves to specific application areas, addressing, for instance, the way in which AI can improve customer experience management [10] or how AI helps to improve brand awareness and allows personalized product modifications [11]. While Nwachukwu and Affen [10] address customer experience, a closely related topic to the customer journey, they do not discuss it in sufficient detail. The same goes for exploring the role of NLP-based technologies, as the authors remain general and superficial. Rana et al. [11], on the other hand, go into more depth regarding the concept of the customer journey but do not explore NLP-based technologies in detail, remaining more on the surface and discussing areas such as machine learning and recommendation systems. In addition to these more general literature reviews that explore AI as a whole, there are also more focused articles discussing specific AI-based technologies in the context of marketing practice. For example, there are several reviews investigating the implementation of augmented reality into retail, showing a positive impact on attitude, decision-making, behavioral intentions, and customer engagement [15, 16]. In addition, chatbots and conversational agents, both of which are part of the NLP field, seem to be of great interest to researchers as they appear relatively frequently as the main topic in review articles [8, 9, 17, 18]. These reviews have shown that a connection between chatbots and customer loyalty is very likely [9], which could make them a viable option throughout the customer journey, especially for postpurchase interactions. Unfortunately, the authors do not link this finding to the customer journey in more detail. Further, Mariani et al. [17] found that the adoption of conversational agents has various user-related outcomes on variables like attitude towards an offering, advice adherence, or purchase intention, leading to promising results. However, the authors again do not link these results to the customer journey framework.

The body of existing literature reviews highlights that this research area holds relevance not only for academic scholars but also for industry practitioners. For instance, AI tools have been found to possess increasing potential in helping marketers identify optimal promotional strategies while addressing trust and privacy concerns [11]. Additionally, practitioners can enhance chatbot sales performance by leveraging anthropomorphic features [18]. Jenneboer et al. [9] have even created a table that lists various practical implications regarding the impact of chatbots on customer loyalty. While these findings are promising for practitioners, the authors do not integrate them into the customer journey framework. For instance, enhanced sales performance could be critical during the purchase stage, yet the authors fail to make that connection.

The previous discussion shows that the current review literature tends to (1) either study AI as a whole (e.g. [13]), (2) study multiple technologies (NLP, computer vision, etc.) at once without differentiating and focusing on specific areas of AI research (e.g. [11]), or (3) limit the scope of research to very specific AI-based technologies (e.g. [17]). These approaches, while valuable, leave a critical gap: they do not provide a detailed overview of a field like NLP as a complex area with its various technologies and how they fit into marketing practice and the customer journey. To address this shortcoming, the following literature review focuses specifically on the integration of NLP-based technologies into the customer journey. The results can serve as a basis for decision-makers, be they practitioners or researchers, to decide how NLP-based technologies can improve the customer journey and, thus, the overall customer experience.

3. Methodology

In this article, two methodological approaches are followed, namely, (1) the PRISMA method introduced by Moher et al. [19], which is a recognized reviewing system frequently used in other literature reviews on similar topics [16, 20–22] and (2) co-occurrence analysis using the software VOSviewer. By applying the standardized PRISMA method, we aim to ensure transparency, reproducibility, and comprehensiveness from the identification of articles to the final determination of study inclusion or exclusion. VOSviewer, on the other hand, is a tool for constructing and visualizing bibliometric networks, enabling the presentation of scientific landscapes in a clear and easily understandable format, making it ideal for review articles. VOSviewer allows us to identify underlying topical trends within the scientific discourse and gaps in terms of terminology. The software uses advanced layout, clustering, and NLP techniques to create these bibliometric networks [23].

At the beginning of each literature review, a search string needs to be defined that contains the most important keywords to cover the research field in the best way possible. In the case of this article, the search string is composed of two parts that were entered into the databases one after the other and are linked by a Boolean AND operator. The first part of the search string contains frequently used keywords of NLP-based technologies and their abbreviations if any. The terms in the first part of the search string are then linked with a Boolean AND operator to the second part, which contains keywords related to the customer journey and adjacent topics such as customer experience (see Table 1). As a result, an article appeared in the search results only if it included at least one keyword from each group.

The various keywords were selected based on other research articles and were discussed and evaluated with two other researchers who are experts in the field of NLP and marketing. To further ensure that the search string satisfactorily represented the subject matter and achieved the maximum number of matching search hits, it was subjected to an iterative testing procedure. For example, we examined whether using the plural or singular form of keywords affected the number of search hits and whether omitting certain keywords influenced the results. Ultimately, we selected the search string that produced the highest number of relevant hits. This fine-tuning was done in the databases/scientific search engines used in this article, namely, EBSCO, Web of Science, IEEE, and ACM. Within EBSCO, the databases Business Source Elite, Communication & Mass Media

TABLE 1: Search string.

First part of the search string	Operator	Second part of the search string
"Natural Language Processing" OR NLP OR "Large Language Models" OR LLM OR "Language Models" OR "Voice Assistants" OR "Smart Devices" OR Chatbots OR "Natural Language Understanding" OR NLU OR "Natural Language Generation" OR NLG OR "Conversational Agents" OR CA OR "AI Agents" OR "Service Robots" OR "Virtual Assistants" OR "Virtual Agent"	AND	"Customer Journey" OR "Customer Experience" OR "User Experience" OR UX OR "Consumer Experience" OR "Customer Satisfaction" OR "Consumer Satisfaction" OR "Consumer Journey" OR "Service Journey" OR "Customer Process" OR "Customer Touchpoints" OR "Digital Touchpoints" OR "Pre-purchase" OR Purchase OR "Post-purchase" OR "Pre-service" OR "Post-service"

Complete, Psychology and Behavioral Sciences Collection, and APA PsycInfo were chosen. Since the authors wanted this work to bridge the gap between the marketing literature with its related fields and the more technical publications, an important concern was to select databases that are commonly used in both fields. IEEE and ACM serve as primary hubs for academic research in computing and information technology, while EBSCO, as a specialized search engine, enables a targeted exploration of specific research areas through its customizable database selection, in our case, focusing on the field of management science. To ensure comprehensive coverage of relevant articles spanning multiple research fields, we supplemented this approach by incorporating the overarching database Web of Science into the search process. The search string was applied to the titles, abstracts, and keywords of potentially fitting articles.

In addition to using an appropriate search term, the database query was also narrowed down by defining suitability criteria based on the PRISMA method.

Search criteria:

- We followed highly cited reviews on similar topics [11, 16, 17, 24] and focused on publications from peer-reviewed journals to ensure better comparability and a high level of scientific quality.
- Publications must be full original primary research articles (no literature reviews, no books, no extended abstracts, no bachelor or master theses, and no dissertations) published within the last 5 years (July 2018 to December 2023) and contain the search keywords in the title, abstract, or keywords.
- Publications must be written in the English language

Screening criteria:

- Articles must address NLP and the customer journey. Papers only discussing one of the two topics are excluded
- We checked again if all of the search criteria were being met

Eligibility criteria:

- Articles must be open access or accessible through university licenses
- Articles must address NLP-based technologies in the context of the customer journey. The integration of the customer journey must be recognizable either by the direct naming of the term or its phases. If the terms are not directly mentioned, the description of the research context must indicate whether or not the article addresses the general concept of the customer journey. An example of this can be found in the study by Han [25]. Here, the author examines chatbots for food orders and the influence of their anthropomorphic features on consumers' purchasing decisions. While the article does not explicitly use the relevant terminology, it clearly addresses an NLP-based technology at the purchase stage. Therefore, despite not mentioning terms such as customer journey, NLP, or related concepts, the study was still included in our database.
- Only articles investigating NLP-based technologies as the research object are included. Articles that use NLP as a research method (e.g., semantic analysis) are dismissed. This was decided as the focus of this work is on the interaction of NLP-based technologies and consumers rather than on the capabilities of the technologies as research tools.

Infringement on one or more of the aforementioned criteria served as an exclusion criterion. These articles were not considered for further investigation in this review. The initial search query, according to the search criteria, returned 899 results. These results were then screened and filtered using the PRISMA method, which is shown in the following flowchart (Figure 1).

In the first step, these 899 search results were checked for duplicates, resulting in 734 initial publications. Articles were classified as duplicates if they had the same title and author. The abstracts of the 734 articles were then examined for the screening criteria. Because this screening was based only on the abstracts, it was sometimes difficult to assess whether the criteria were met. Therefore, articles were only excluded if the abstracts provided explicit evidence that one or more of the screening criteria had been violated. This was the case, for example, with abstracts that contained search string

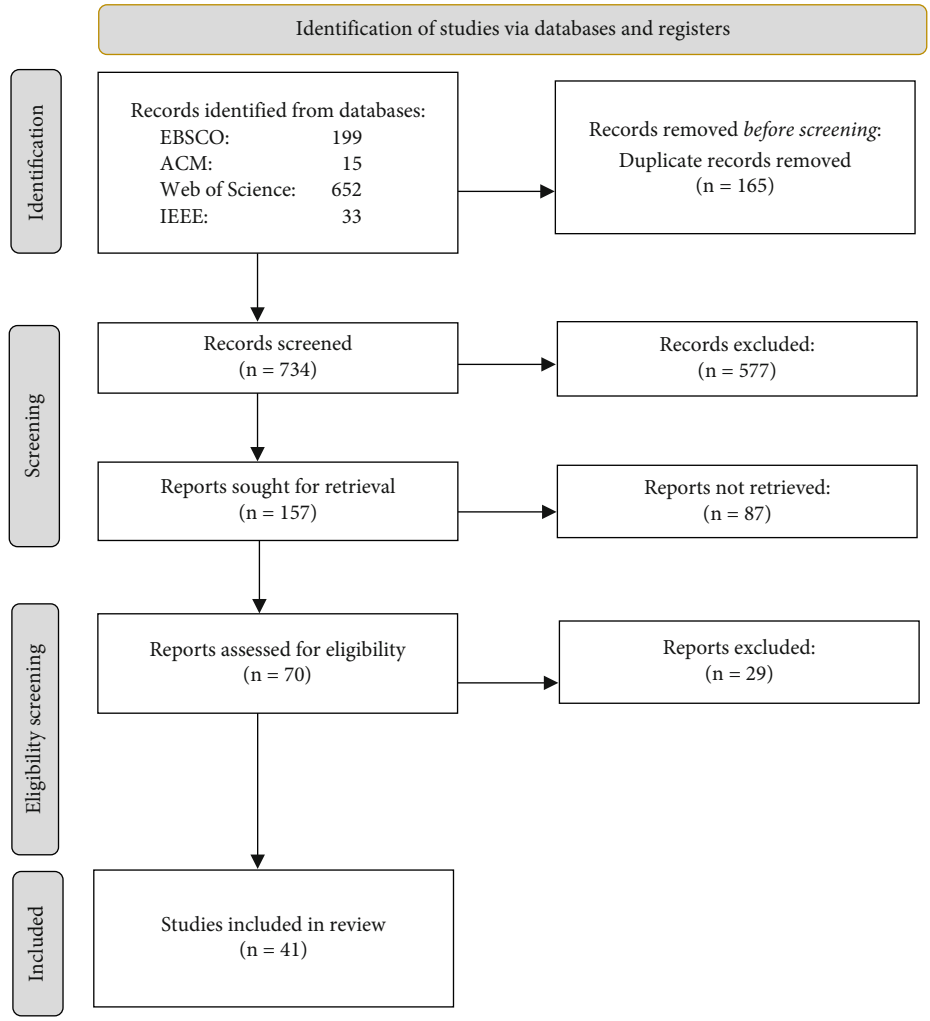

Figure 1: PRISMA flow chart [26].

terms, especially abbreviations, that have a different meaning in a different research area. For example, Malekian et al. [27] use the abbreviation CA for corporate ability instead of conversational agent. However, the main reason for exclusion in the screening phase was that some articles did not examine NLP-based technologies and the customer journey but focused on one of the two topics. The various reasons for the exclusion of the individual articles were recorded in the form of a spreadsheet. After the screening stage, 157 articles remained, 70 of which were accessible through the university licensing pool. Then, the full texts of these 70 publications were reviewed against the eligibility criteria, excluding an additional 29 articles. A total of 734 publications were reviewed, of which 41 articles qualified for further investigation and, thus, integration into this literature review. To ensure sufficient objectivity, the process from the creation of the search term to the screening and assessment of suitability was carried out by a team of researchers. The screening and eligibility process was recorded and structured in the form of a spreadsheet that included, among others, information about the authors, the title, the name of the journal, the decision on the outcome of the selection, and the reasons for the respective decision. Critical screening decisions were discussed in special meetings in order to reach joint decisions and ensure the greatest possible objectivity. The information was then categorized, processed, and synthesized in a separate spreadsheet.

4. Results

The following chapter, in which the results of this review are presented, begins with the presentation of the descriptive results. Afterwards, the focus will be on the co-occurrence analysis to gain a better understanding of the most prevalent themes of the papers. This will be supplemented by findings regarding NLP-based technologies in the context of the customer journey.

4.1. Descriptive Results

4.1.1. Chronology of Publications. This literature review includes two articles from 2018 and three from 2019. The number of publications present in this study then fell to zero in 2020 before rising sharply in 2021 to 9 and in 2022 to 10

publications. One year later, in 2023, the number of articles on NLP in the context of the customer journey peaked at 17 publications. The course of this timeline is also illustrated in Figure 2.

4.1.2. Distribution of Publications Among the Individual Journals. In addition to the chronology of the publications, it is also insightful to examine in which journals the respective articles were published.

With four articles, the journal *Psychology & Marketing* published the most studies on the topic, followed by the *Journal of Business Research*, the *Journal of Marketing*, the *International Journal of Retail & Distribution Management*, *Information System Frontiers*, *Frontiers in Psychology*, and *Electronic Commerce Research* with two articles each. It seems that journals from the fields of management and psychology are particularly interested in this topic. For the sake of clarity, these journals with two or more articles are also shown in Table 2 with the respective papers.

4.1.3. Method of Data Collection. The analysis has shown that both quantitative and qualitative methods are used, albeit to varying frequencies. Three times, a qualitative method was employed, while quantitative approaches were taken 43 times. It must be said that the number of different methodologies is higher than the number of articles since some articles used multiple methods for their studies. The three most frequently employed methods were the online experiment ($n = 20$), the online questionnaire ($n = 14$) and the laboratory experiment ($n = 5$). Other methods used in the articles were field experiments ($n = 3$), interviews ($n = 2$), EEG ($n = 1$), and content analysis ($n = 1$). The distribution is also illustrated in Figure 3.

4.1.4. Research Settings. The publications in our database cover a wide range of different research settings. To improve readability, we have decided to focus on research settings that occur twice or more rather than displaying all 24 different settings in the reviewed articles (see Figure 4).

As Figure 4 illustrates, there was a sharp drop in frequency between settings, while most settings only occurred once. Online shopping was by far the most common research setting ($n = 20$). This was followed by research settings relating to the hospitality and telecom sectors, with three articles each. The other four investigation contexts were grouped into banking ($n = 2$), online ticket sales ($n = 2$), tourism ($n = 2$), and no specific research setting ($n = 2$). Examples of more niche application settings that occurred only once include car rental services and music streaming.

4.1.5. Participant Groups. The selection of participants is a recurring process that researchers must deal with. To gain a better understanding of a population, it is important that the different groups within the population are adequately represented in terms of gender and age. For this reason, the authors decided to examine the selection of participant groups in the selected articles. It was found that none of the articles tailored their research to a specific gender. However, the situation is different with regard to age: although most articles do not focus on a specific age group ($n = 30$), a nonnegligible share of studies specifically investigates younger participants ($n = 9$) in the form of university students [25, 33, 35, 44–46], millennials [32, 43], and 23–30-year-olds [47]. Only two articles specifically focused on older participant groups, for example, 30 and upwards [48, 49].

4.2. Co-Occurrence Analysis. To gain a deeper insight into the predominant themes of the papers, it was decided to opt for a co-occurrence analysis. Co-occurrence analysis is a method for visualizing terms according to their frequency and association with each other in a given dataset [18]. The frequency of a term is illustrated by its node size, while the association strength is visualized by the distance between the nodes, meaning the smaller the distance between words, the higher the association [50].

To visualize the data set of this study in such a way, VOSviewer, a program for creating, exploring and analyzing maps of network data, was used [51]. Here, groups of terms associated with each other can also be grouped into different clusters. In addition, the degree to which a term is integrated into a network can be assessed based on the link strength. Before visualizing and analyzing this study's data set, which is based on the keywords of the different papers, the data was checked so that terms such as "chatbot" and "chatbots" are displayed as one node to avoid ambiguity. The threshold for the minimum number of occurrences of a keyword to be displayed on the map was set at three, resulting in a total network of 13 terms. The software then grouped these 13 terms into three different clusters, which are listed in Table 3.

4.2.1. Network Visualization of Keywords. Figure 5 illustrates the findings of the network visualization. It is evident from the node sizes that the keywords "artificial intelligence" and "chatbots" appear most frequently in the various papers. This visual observation is also backed by the total link strength, which is an attribute indicating how strongly an item is connected to other items within the network [51]. Here, "artificial intelligence" has an overall link strength of 39 and "chatbots" of 40. It is not surprising that the term "chatbots" occupies a central position in the network, as it was one of the components of the search string. However, the term "artificial intelligence" was not explicitly mentioned in the search string but is still one of the two most common keywords for the description of the articles in the paper pool, although much more specific terms were offered in the search string. This could be an indication that the current scientific discourse lacks specific terminology and tends to fall back on the very broad and unspecific term "artificial intelligence" instead of, for example, "NLP," as already suggested in the Related Literature Reviews section.

In addition, the network visualization shows that the variable "anthropomorphism," which was also not part of the search string, is a common theme in the present sample (link strength: 29), which is primarily examined in the context of customer service (link strength: 4). Based on this finding, current research is investigating human likeness of NLP-based technologies and their impact on customer

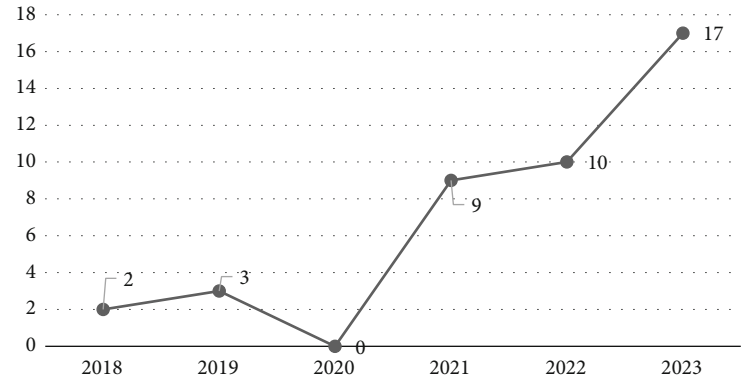

Figure 2: Number of publications per year.

TABLE 2: Most frequent journals and corresponding papers in our database.

Journal	Article
Psychology & Marketing	Chen et al. [28]; Dwivedi et al. [29]; Elmashhara et al. [30]; Pizzi et al. [31]
Journal of Business Research	Fernandes & Oliveira [32]; Roy and Naidoo [33]
Journal of Marketing	Crolic et al. [34]; Garvey et al. [35]
International Journal of Retail & Distribution Management	Chen et al. [36]; Silva et al. [37]
Frontiers in Psychology	Xu et al. [38]; Yun & Park [39]
Electronic Commerce Research	Klein & Martinez [40]; Wang et al. [41]
Information System Frontiers	Li, Gan, et al. [42]; Sharma et al. [43]

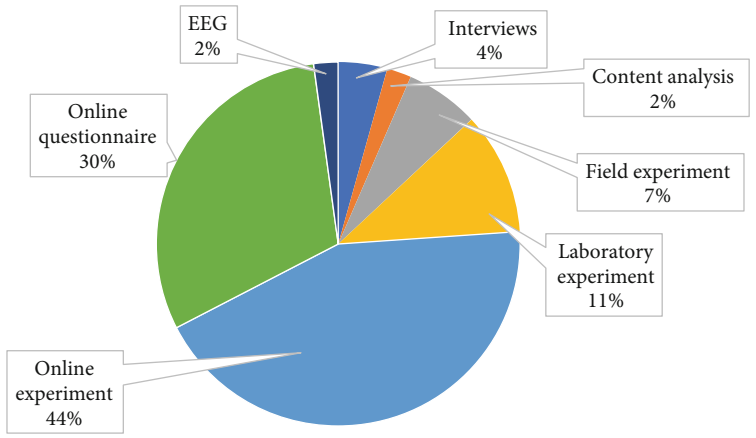

Figure 3: Frequency of methodologies.

interactions. The most common keywords that can be described as factors regarding the customers' interaction experiences according to the network visualization are trust (link strength: 10), personality (link strength: 8), and, of course, customer satisfaction (link strength: 13), customer experience (link strength: 17), and satisfaction (link strength: 9).

It is noticeable that customer characteristics, for example, in relation to demographics, are not present in the sample. In terms of analysis methods, structural equation modeling (link strength: 11) is the most common term that is explicitly mentioned as a keyword by the authors.

4.2.2. Overlay Visualization of Keywords. In addition to the visualization of the keywords in the form of a network, the sample was also edited as a so-called overlay visualization, which puts the terms in a temporal context (Figure 6). The values are formed by VOSviewer from the publication year of the articles to which the keywords belong. The overlay visualization shows that authors used the term "Structural Equation Modelling" the most around 2020, with an average publication year of 2020.50, making it the "oldest" keyword in the dataset. The central term of artificial intelligence had its height towards the end of 2021 (average publication year 2021.79). Current

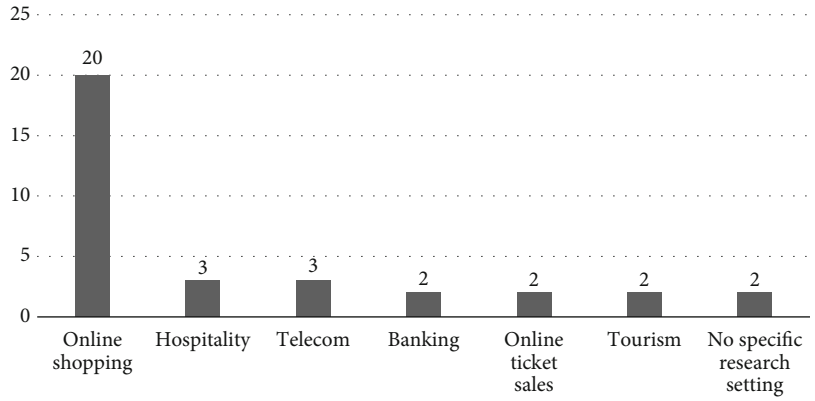

Figure 4: Most frequent research settings.

TABLE 3: Clustering of keywords.

Cluster 1	Cluster 2	Cluster 3
Anthropomorphism	Artificial intelligence	Chatbots
Conversational agents	Customer experience	Customer satisfaction
Customer services	Personality	
Robots	Structural equation modelling	Conversational commerce
Satisfaction	Trust	

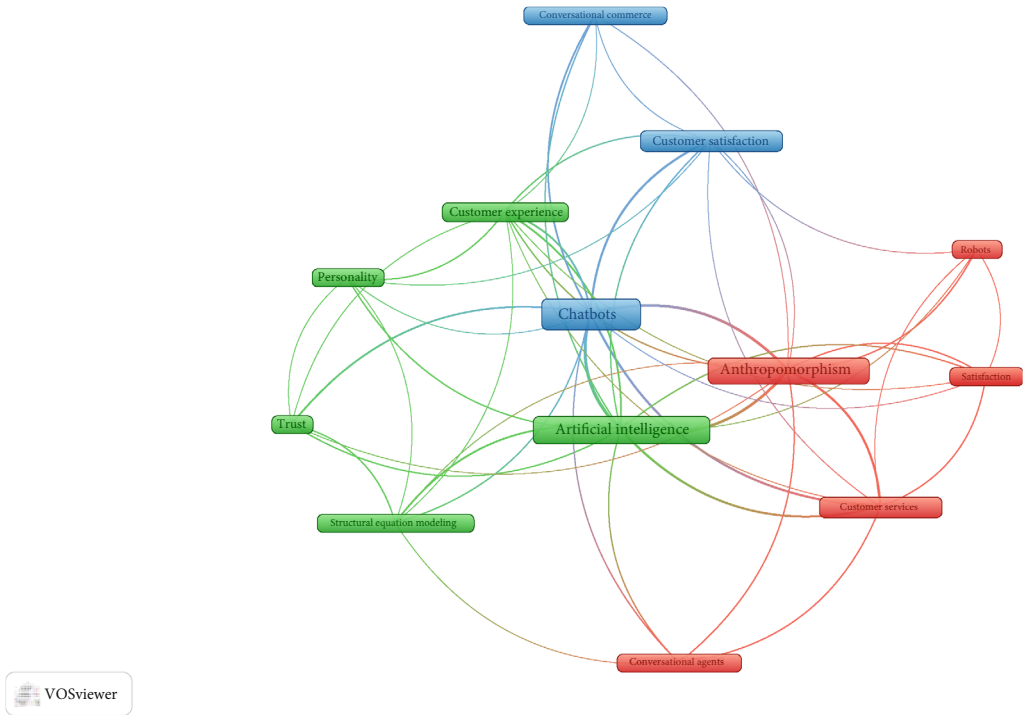

Figure 5: Network visualization of the keywords in the dataset.

research seems to be most interested in the aforementioned user experience factors like trust, anthropomorphism, and satisfaction, all having an average publication year of 2022 and upwards. Of course, the data should be interpreted with caution, as some keywords are used less frequently than others and could, therefore, be more prone to outliers. Nevertheless, the analysis reveals tendencies within the scientific discourse and its prioritization.

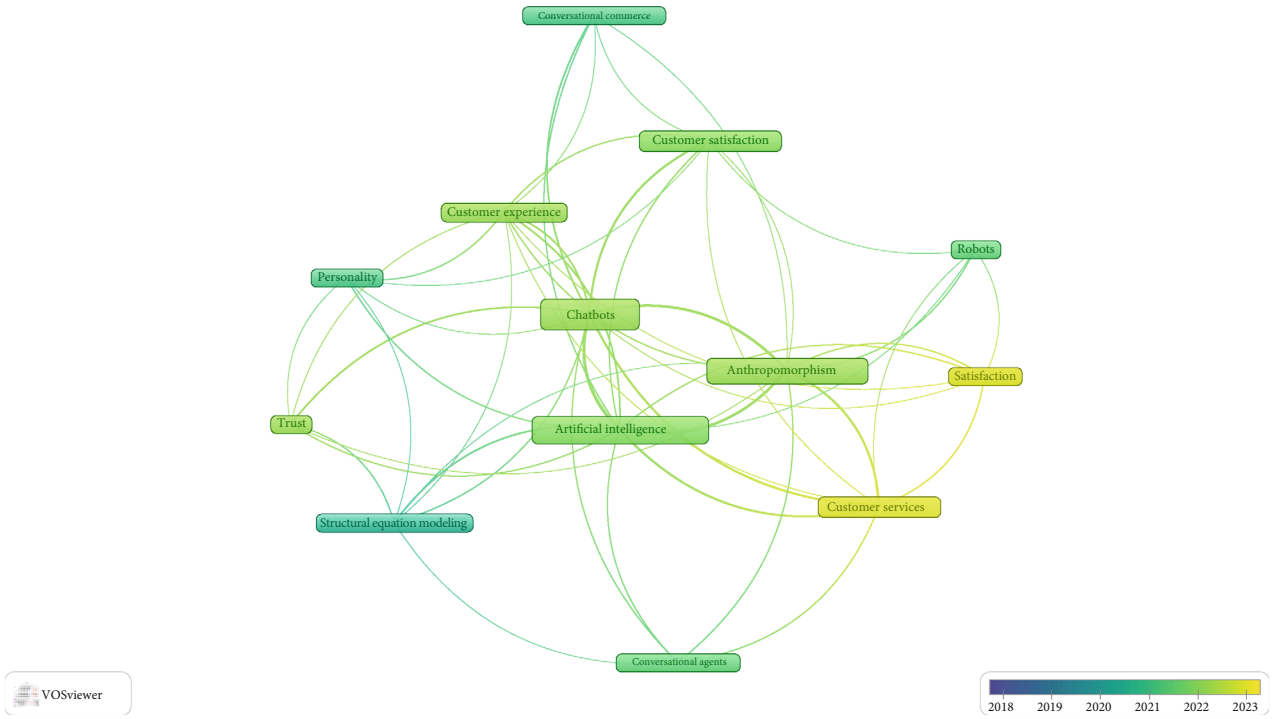

Figure 6: Overlay visualization of the keywords in the dataset.

4.3. Results Regarding NLP-Based Technologies Throughout the Customer Journey

4.3.1. Customer Journey Phases. Since the customer journey is one of the major concepts of interest in this literature review, the articles were analyzed regarding how the different technologies are distributed along its stages (Figure 7). The articles were grouped into the four groups "no specific stage," "prepurchase stage," "purchase stage," and "postpurchase stage." The label "no specific phase" was given for articles that are more exploratory or general and do not focus on a specific stage [52, 53]. If the customer journey phase was not already evident from the keywords in the article or because it was clearly mentioned in the text, the categorization was made based on the definition by Lemon and Verhoef [12] and by looking at the respective context in which the studies were conducted. Articles could be categorized into multiple stages since the stages are dynamic [12], and publications could investigate multiple stages within one study. The analysis showed that most of the articles investigated NLP in the purchase stage ($n = 27$), tightly followed by the prepurchase stage ($n = 23$). NLP-based technologies were significantly less studied in the postpurchase stage ($n = 9$). Five articles were not categorized in one of the specific stages and were labelled with "no specific phase." This underrepresentation of the postpurchase phase could be explained by the complexity of after-sale dialogues, which is difficult to handle with classic conversational platforms like rasa. As the power of customers has increased at this stage due to the widespread availability of rating options in e-commerce, the postpurchase stage could benefit strongly from more powerful NLP applications. With generative AI on the rise, this might change in the near future, as this kind of AI brings new capabilities in performing complex dialogues.

4.3.2. NLP-Based Technologies. The investigated publications used different NLP-based technologies for their research purposes. An overview of the variety of technologies can be seen in Table 4.

Chatbots were the most studied technology within the publication pool ($n = 26$). Other technologies examined were service robots and voice assistants, with five studies each. The least studied NLP-based technologies within our database were conversational agents ($n = 2$), AI agents ($n = 1$), digital assistants ($n = 1$), and virtual assistants ($n = 1$). During the analysis, it became clear that the terminology in the field of NLP-based technologies is not unambiguous. Conversational agents, virtual agents, and digital assistants, for example, all have chatbot characteristics, are closely related to each other and are frequently described as very similar if not as the same technology, across the reviewed papers [43, 44, 69]. Thus, it would be advantageous to agree on a certain terminology. It was decided not to cluster the mentioned terms into one of the other categories to better illustrate this ambiguity. In addition to this insight into terminology, the technologies were also clustered along the phases of the customer journey in Figure 8.

Figure 8 shows that although chatbots are distributed relatively proportionally to the total number in the individual phases of the customer journey, service robots have a strong overhang in the purchasing stage.

Another conspicuousness relates to the application of voice assistants. Here, the focus lies clearly on the first two stages of the customer journey, while the postpurchase stage

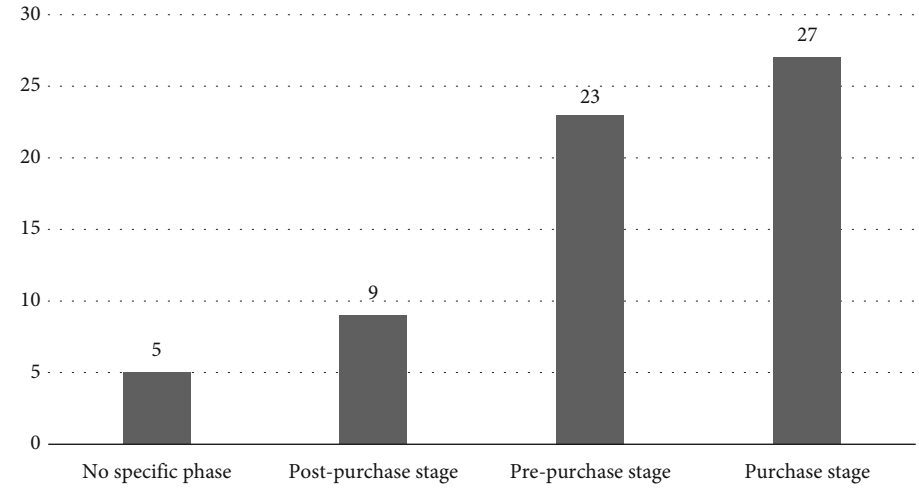

Figure 7: Distribution of the different customer journey stages within the dataset.

TABLE 4: Studied NLP-based technologies.

NLP technology	Publication
Chatbots	Chang et al. [49]; Cheng et al. [36]; Chen et al. [28]; Crolic et al. [34]; Dwivedi et al. [29]; Elmashhara et al. [30]; Han [25]; Haupt et al. [54]; Jansom et al. [55]; Kaushal & Yadav [53]; Klein & Martinez [40]; Gan et al. [42]; Luo et al. [56]; Pizzi et al. [31]; Roy & Naidoo [33]; Ruan & Mezei [57]; Sands et al. [58]; Shin et al. [46]; Shumanov & Johnson [59]; Silva et al. [37]; Tan & Liew [60]; Upadhyay & Kamble [61]; Wang et al. [41]; Xu et al. [38]; Yen & Chiang [62]; Yun & Park [39]
Service robots	Caleb-Solly et al. [48]; Liang et al. [63]; Mende et al. [45]; Okafuji et al. [64]; Wenguo & Yongwei [65]
Voice assistants	Bawack et al. [66]; Fernandes & Oliveira [32]; Kautish et al. [67]; Li et al. [68]; Wolbers & Walter [47]
Conversational agents	Danckwerts et al. [69]; Schmid et al. [70]
AI agents	Garvey et al. [35]
Digital assistants	Sharma et al. [43]
Virtual agents	Liew & Tan [44]

is mostly ignored. This uneven distribution of NLP-based technologies along the customer journey could have several explanations, for example, a more natural adaptation to certain application scenarios. An example of such a more natural adaptation to certain scenarios could be the aforementioned tendency to investigate chatbots in the first two phases of the customer journey, as communication in the postpurchase phase could be characterized by a higher degree of complexity. Nevertheless, researchers should be open to examining all phases of the customer journey so that no dark spots can arise, and rather, atypical application scenarios can still be identified and studied.

4.3.3. Investigated Concepts. One of the main objectives of this review is to get an overview of the critical variables and factors currently studied by researchers in the realm of NLP and the customer journey. Each article was checked for the most important variables examined, resulting in the extraction of around 100 terms. As this list itself was too extensive to draw conclusions from, it was then analyzed for the most common variables within the database, which were then clustered into different thematic groups. The analysis revealed that the publications in terms of frequency mainly focused on the four variables purchase intention ($n = 17$), anthropomorphism ($n = 15$), customer satisfaction ($n = 10$), and trust ($n = 10$), which were also examined together in various combinations. These concepts are also illustrated in Figure 9 and put into context according to their relationship to other variables. We categorized the relationship between overall positive influence, overall negative influence, and conflicting/ambiguous results by comparing the number of studies that reached each conclusion.

4.3.3.1. Purchase Intention. Purchase intention was one of the most frequently investigated variables in the database and describes the process in which customers indicate whether they are willing to continue to purchase or not [71]. Understanding this factor is important, as it can be influenced by various circumstances and elements within the shopping environment.

One factor influencing purchase intention is the perceived competence of NLP-based technologies. A study

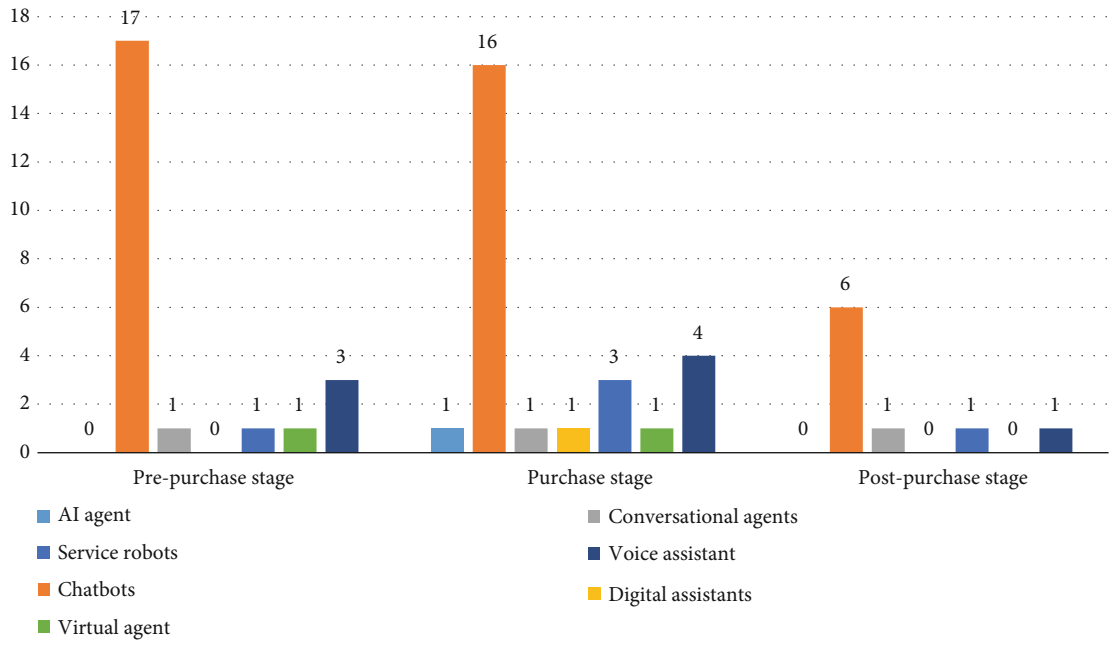

Figure 8: NLP-based technologies along the customer journey.

examining this variable in relation to purchase intention, signaling expertise through the use of specialized chatbots for specific product categories, came to conflicting results, with some users finding multiple chatbots more confusing than helpful [44, 60]. This indicates that while specialization can enhance competence perception, the implementation of multiple chatbots needs to be carefully planned to avoid diminishing user experience. One could speculate that the use of such specialized chatbots could reduce trust if the service does not meet customer expectations.

Building on this, the studies by Yen and Chiang [62] and Li et al. [68] investigated purchase intention in an online shopping environment and how, among other things, chatbot competence influences customers' purchase intention through mediating variables. They found that competence positively influences purchase intention through the variable of trust [42, 62]. Chang et al. [49] investigate the same phenomenon in a different setting but come to similar conclusions: The authors' results show that the purchase intention of solo travelers is positively influenced by the competence of the NLP-based technology. For solo travelers, who may place a higher value on personalized and reliable interactions, a competent chatbot can serve as a crucial factor in the decision-making process, providing a sense of security and enhancing the overall customer experience. The study highlights that competence seems to be an important and overarching variable impacting purchase intention in different contexts and customer segments.

In addition to competence and trust, anthropomorphism is another variable frequently studied in the context of purchase intention in our database. Roy and Naidoo [33] found that an anthropomorphic conversation style of NLP-based technologies can have a positive influence on purchase intention moderated by customers' time orientation. Time orientation refers to how individuals prioritize and focus on either immediate pleasures and rewards (present orientation) or long-term goals and consequences (future orientation) in their decision-making and behavior [33]. For instance, a chatbot that engages users with a human-like conversational style might increase purchase intention in scenarios requiring quick decision-making where underlying motives are hedonistic in nature, such as last-minute travel bookings. The implication for practitioners is that effectively matching the conversation style to consumer motives—aligned with the product type—determines whether a chatbot should emphasize warmth or competence in its communication to increase purchase intention. Crolic et al. [34] investigated a similar scenario and found that the effect of anthropomorphism on purchase intention is moderated by customer anger, revealing that while anthropomorphic features can improve purchase intention, negative emotions may dampen these effects. Building on this, beyond matching time orientation to conversation style, chatbot design should also account for the customer's emotional state to enhance purchase intention. Researchers are developing machine learning techniques to detect emotions through text analysis, which could serve as a foundation for even better alignment between chatbots and their human counterparts [72]. Furthermore, anthropomorphism can also positively influence purchase intention through mediating variables like trust [62] and perceived enjoyment and social presence [25]. Sharma et al. [43] also found a positive influence of anthropomorphism in NLP-based technologies on the variable specifically for the group of millennials. By integrating these findings, it becomes evident that the effectiveness of anthropomorphism is contingent on its ability to work with other psychological factors. The combined influence of these factors may enhance the overall appeal of the chatbot and make the purchasing process feel more natural and less transactional. This dynamic interaction illustrates the importance of a holistic approach in designing interactions.

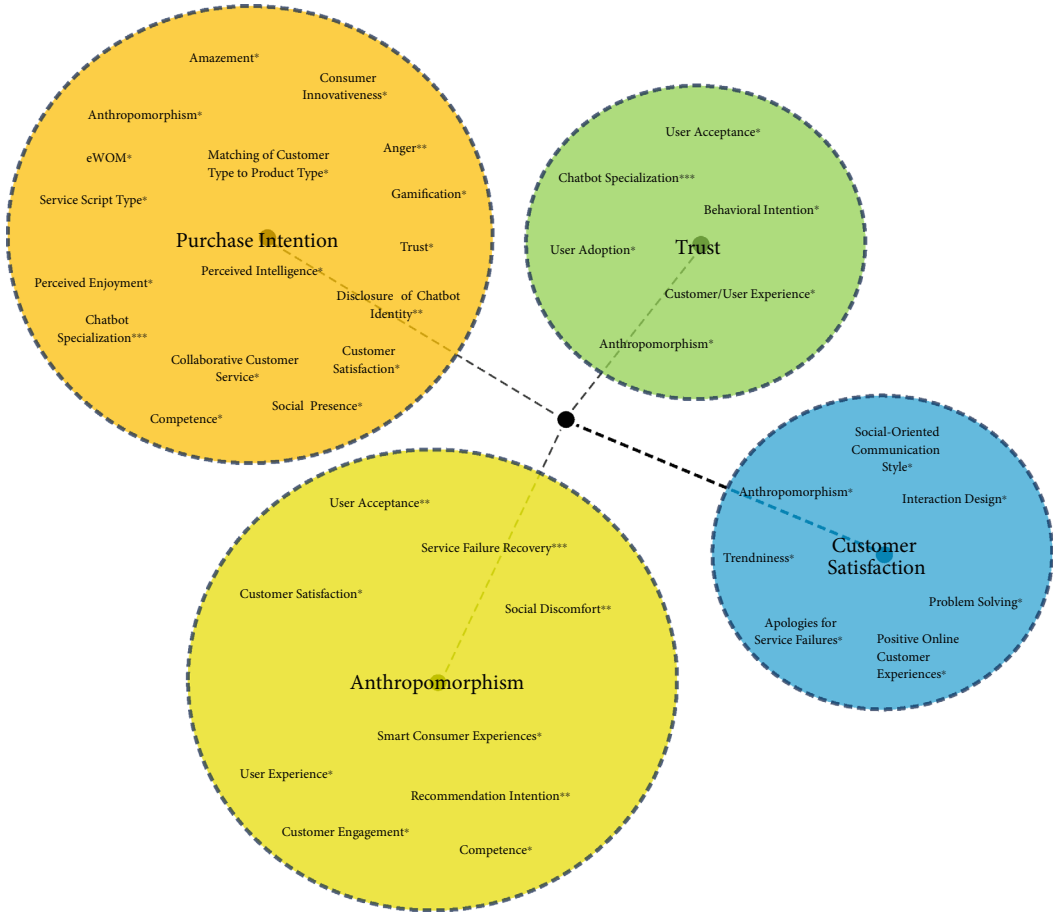

Figure 9: Illustration of the core concepts and their relationship to other variables. Note: *overall positive influence; **overall negative influence; ***conflicting/ambiguous results.

Other studies within the database suggested that NLP-based technologies may affect purchase intentions through even more factors like motivated consumer innovativeness, amazement with the technology, and eWOM [67]. The study by Kautisch et al. [67] suggests that consumers who are inherently more open to novel experiences may be more inclined to make purchases when interacting with advanced NLP-based technologies, particularly when these technologies evoke a sense of awe or amazement. This emotional engagement could draw customers into the purchasing funnel and create a memorable interaction that can again lead to positive electronic word-of-mouth (eWOM).

The deployment of different service scripts [58] is another factor illustrating the nuanced role of conversational design in shaping customer behavior. The choice between personalized, informal scripts versus more formal, structured interactions can significantly affect how customers perceive the chatbot and, consequently, their likelihood to make a purchase. This finding underscores the importance of tailoring the communication style of chatbots to match the context of the service and the preferences of the target audience. For instance, a casual script may work well in a fashion retail setting but might be less appropriate in a financial services context.

The importance of adjustments to individual needs and contexts is also visible for factors like collaborative customer service [64] and the matching of customer service type to product type [28] and highlights the importance of strategic use of NLP-based technologies. Okafuji et al. [64] demonstrate that combining chatbots with human agents can enhance purchase intentions by leveraging both human empathy and technological efficiency. Chen et al. [28] emphasize that the match between customer service type and product type positively affects consumers' purchase intentions. These insights suggest a complementary, context-sensitive approach to maximize purchasing behavior.

Additionally, visual information and perceived intelligence, as explored by Li et al. [68], underscore the role of sensory and cognitive elements. The study shows that visual cues provided by NLP-based technologies can enhance the perception of the chatbot's intelligence, which in turn positively influences purchase intentions. The interplay between visual appeal through these elements and perceived intelligence seems to be important, as it could bridge the gap between artificial and human-like interaction, making the technology more relatable and effective in driving conversions. Adding to the importance of sensory and cognitive elements, Elmashhara et al. [30] show that integrating the

element of gamification can boost engagement and motivation, making the shopping experience more enjoyable and likely to lead to purchases.

Another important factor ensuring a pleasant experience seems to be the disclosure or nondisclosure of chatbot identity as it affects purchasing behavior. Luo et al. [56] found that when customers know that they are communicating with a machine, they purchase less. This effect should be kept in mind to optimize purchasing outcomes in e-commerce.

While much research focuses on initial purchase intention, Yun and Park [39] uniquely examined repurchase intention. Their study reveals that repurchase intention is positively influenced by customer satisfaction. This finding shifts the focus from initial purchase to long-term customer loyalty, highlighting the importance of sustaining high levels of satisfaction through consistent and positive interactions with NLP-based technologies. Their research suggests that the key to driving repeat purchases lies in the chatbot's ability to continuously meet or exceed customer expectations, thereby fostering a lasting relationship between the customer and the brand.

To conclude, the interplay of the aforementioned factors shapes customer purchase intention, highlighting the need for a nuanced and holistic approach in the design and implementation of NLP-based technologies to optimize both initial and repeat purchases.

4.3.3.2. Anthropomorphism. Anthropomorphism, defined as the ascription of human attributes, physical characteristics, emotions, and traits to nonhuman entities [40], plays a significant role in how users interact with NLP technologies. This concept is important in understanding user behavior and is frequently studied within our database.

Research by Danckwerts et al. [69] and Fernandes and Oliveira [32] investigates the variable in the form of perceived humanness. It was found that perceived humanness has a positive impact on user experience [69] while not having a significant influence on user acceptance [32]. This suggests that while human-like interactions can enhance the enjoyment of using NLP technologies, they may not always translate to increased acceptance if the technology fails to meet user expectations. To leverage anthropomorphism effectively, developers of NLP technologies could ensure that the human-like qualities they design are supported by robust functionality and reliability.

Anthropomorphism was also studied in the context of customer satisfaction. The authors found that NLP-based technologies with human attributes can be positively associated with customer satisfaction [40], though the context has to be taken into consideration to ensure satisfying interactions [34, 35, 46]. However, not all outcomes are positive. For instance, one study on service robots found that anthropomorphism can negatively influence satisfaction by causing social discomfort [63]. This suggests that while a certain level of anthropomorphism can be beneficial, pushing the boundaries too far may backfire, leading to negative customer experiences. One takeaway could be that the degree of anthropomorphism needs to be carefully calibrated to match customer expectations and avoid triggering negative emotional responses.

Beyond customer satisfaction, human-like features in NLP-based technologies also affect competence [31], engagement and smart consumer experiences [61], service failure recovery [54], and chatbot experience and recommendation intention [29]. Pizzi et al. [31] have shown that higher levels of anthropomorphism increase the perception of competence, which could, in turn, impact user trust and willingness to disclose personal information, fostering deeper engagement and more effective interactions.

Additionally, Upadhyay and Kamble [61] demonstrate that the integration of humanlike features in service robots positively affects customer engagement and creates smarter consumer experiences. Engaging interactions can be a differentiator from competitors as they encourage customers to be more deeply involved with a product, service, or company, creating a stronger bond.

However, the role of anthropomorphism in service failure recovery is more nuanced. Haupt et al. [54] have shown that empathic and, therefore, more human-like recovery messages outperform solution-oriented messages primarily when the chatbot repeatedly fails to deliver an adequate response—a phenomenon known as double deviation. In contrast, solution-oriented messages may be more effective after a single failure, particularly when users attribute responsibility to themselves rather than the chatbot. This indicates that situational factors, such as failure frequency and failure attribution, play a crucial role in determining the effectiveness of recovery strategies. These findings were consistent across multiple studies involving different demographic groups, including a younger cohort from European universities and a more diverse sample from a US-based platform with an average age of 45 years. Consequently, tailoring the type of service recovery message to both the communication context and user demographics can significantly enhance interaction satisfaction and trust.

Finally, Dwivedi et al. [29] found that anthropomorphic features like warmth are particularly influential in the postpurchase stage, whereas in the purchase stage, cognitive cues play a more important role in a satisfying experience. This could suggest that optimizing customer interactions requires designing chatbots with human-like characteristics that align with the specific stage of the customer journey. For instance, incorporating more warmth can enhance customer satisfaction during postpurchase interactions. Furthermore, this study illustrates an important phenomenon in the current scientific discourse, which will be discussed in more detail later on.

In summary, while anthropomorphism can significantly enhance interactions on different variable levels, its effectiveness is highly context-dependent. Developers should carefully calibrate the level of human-like features in NLP-based technologies to ensure that they align with user expectations, the specific context of the interaction, and the stage of the customer journey, thereby avoiding potential drawbacks such as social discomfort or dissatisfaction.

4.3.3.3. Customer Satisfaction. Customer satisfaction is a multifaceted concept that encompasses a customer's expectations before a purchase, their interactions during service

encounters, and the fulfilment of their expectations post-purchase across all points of contact [55]. Within the context of NLP-based technologies, this concept seems to be central within our database to understand how these technologies influence the customer journey.

Our review reveals that several factors can enhance customer satisfaction when interacting with NLP-based systems. For instance, apologies for service failures [65] and the use of a social-oriented communication style [38] can significantly improve customer perceptions. Wang et al. [41] and Yun and Park [39] come to similar results and demonstrate that social-oriented and emotional communication styles are particularly effective during service recovery, where customers may tend to appreciate a more empathetic approach. This suggests that while technical proficiency is important, the ability of NLP technologies to convey empathy and understanding is equally important in fostering customer satisfaction. These elements can contribute to creating a more empathetic interaction, which could be important for mitigating the sometimes impersonal nature of AI-driven systems.

Positive online customer experiences [52] further bolster satisfaction by ensuring seamless and engaging interactions, while anthropomorphic cues [40], such as the ability to mimic human characteristics, can enhance satisfaction as well. However, as already mentioned in the prior chapter anthropomorphism can also have a negative effect on customer satisfaction in certain situations [34, 63], as it may create unrealistic expectations that the technology cannot meet.

Additionally, interaction design, trendiness, and problem-solving [55] were found to boost customer satisfaction, suggesting again that a well-designed NLP technology could potentially drive loyalty, positive brand perception, competitive advantage, and thus long-term business growth.

While the prior publications focused on NLP-based technologies alone, Ruan and Mezei [57] compared a chatbot to human frontline employees. Here, the authors showed that the technology only performed better than the employee in terms of customer satisfaction for functional products. This finding underlines again the importance of context when evaluating the effectiveness of NLP-based technologies in customer interactions.

In summary, while NLP-based technologies have the potential to enhance customer satisfaction significantly, their effectiveness is influenced by various factors, including the nature of the interaction, the customer's expectations, and the specific communication strategies employed.

4.3.3.4. Trust. Trust is another key variable in the field of NLP-based technologies according to our database. It appears to serve a moderating role in shaping user experience [69], enhancing user acceptance [32], and influencing behavioral intentions [37] and customer experience performance [66]. It seems that trust plays an important role for customers when deciding whether or not they are inclined to use an NLP-based technology and if they perceive the interaction as satisfying. This makes it all the more important to clarify the contradictory results regarding the influence of specialized chatbots as opposed to generalist chatbots on trust [44, 60]. On the other hand, it seems to be well-studied that trust can be positively impacted by the usage of anthropomorphic attributes [40, 42, 62]. In addition to these quantitative studies, Wolbers and Walter [47] found in in-depth interviews that trust is an important factor for user adoption throughout the customer journey, underlining the previously made assumption.

4.3.3.5. Other Concepts. Even though most of the publications have already been covered by the concepts mentioned above, some articles have not yet been discussed. Calleb-Solly et al. [48] investigated in a longitudinal study the acceptance of human-robot interaction, proposing a more pragmatic approach to the design of the technology. Their research highlights that practical design considerations can significantly impact user acceptance, suggesting that technological integration must address real-world user concerns and usability to enhance effectiveness. This is also evident in another study from our database that looked at service robots and found that the use of this technology can lead to increased compensatory consumption, driven by consumer discomfort [45]. This finding illustrates the complex interplay between technology adoption and consumer psychology, emphasizing the need for a deeper exploration of how these technologies influence consumer behavior beyond initial impressions.

Furthermore, Kaushal and Yadav [53] found through interviews with B2B professionals, among other things, that perceived risk, customization, integration with other platforms, scalability, omnichannel engagement, and personalization are important for a successful implementation of chatbots. The importance of tailored interactions was again underlined by the variable of personalization, which also has a positive impact on customer engagement with chatbots [59].

5. Discussion and Future Research Directions

Our systematic review led to several different results that need to be discussed and put into context. This discussion is structured according to our research questions and contains theoretical as well as managerial implications. This is followed by a chapter dedicated to discussing the other findings that go beyond the two research questions and an illustration of possible research questions for future studies.

5.1. How Is the Current State of Research on NLP-Based Technologies Distributed Across the Different Stages of the Customer Journey in Terms of Research Quantity and Technological Focus, and What Underlying Technological Trends or Limitations Characterize This Distribution? Our review has shown that there is currently an imbalance between the different stages of the journey in terms of occurrence. The analysis shows that the postpurchase stage must be described as a dark spot in the discourse at the moment. Researchers in the field of NLP currently tend to ignore the postpurchase stage and focus on the first two stages of the customer journey (see Figure 8). One possible explanation for this phenomenon could be the tendency towards more

complex interactions in this phase of the customer journey [36]. As shown by Dwivedi et al. [29], chatbots perform worse in the postpurchase phase than in the purchase phase for all measured variables. With the emergence of new technologies like generative AI, this might change. As generative AI becomes increasingly capable of handling complex communication situations with its ability to generate human-like content [73], postpurchase interactions between customers and technological counterparts appear to be more feasible than previously. Soni [74] supports this claim, stating that LLMs mark a significant advancement in customer relationship management. However, LLMs will probably not be the solution to all problems in this regard, since these models, as Haupt et al. [54] assess correctly, possibly hallucinate instead of giving a rather simple error message stating that the system does not know the answer. This could lead to incorrect system statements that may have significant consequences depending on the interaction context (e.g., purchasing process and troubleshooting instructions). Therefore, failure recovery measures could be even more important for this kind of NLP-based technology. The problem of hallucinations and their detection is currently part of the scientific discussion (e.g., [75]), so this problem may be solved by refining model architectures rather than through recovery messages. Nevertheless, future studies could investigate whether generative AI opens up the postpurchase stage or whether a human counterpart is still needed to handle interactions in this stage of the customer journey, as it seems to be currently the case. Practitioners could also benefit from deploying this new technology, as previously, all answers to possible questions that a customer might ask in postpurchase but also other interactions had to be determined manually. In contrast, now it could be sufficient to train an LLM based on FAQs. This could, of course, also be relevant for future research, as it is currently not clear whether such a model would perform better than a classic chatbot. However, according to our database, LLMs and generative AI are not currently being investigated in the context of the customer journey. Besides the inherent complexity of postpurchase interactions, another explanation for its underrepresentation could also reflect methodological factors. Data on acquisition and conversion metrics may be generally more readily available and easier to measure than latent variables like customer loyalty. Another reason for the current focus on the first two phases of the customer journey could lie in the concepts examined, in particular the one of purchase intention. As already shown, purchase intention is the most frequently investigated concept in our database. Repurchase intention, on the other hand, which relates to interactions after the first two phases of the customer journey, was only examined once [39]. Consequently, concepts that relate to customer retention, customer loyalty, and product or service advocacy are not widely discussed in the scientific discourse, at least not in the context of NLP-based technologies and the customer journey. This should change as the Internet allows customers to share their experiences, acting as ambassadors and influencing potential new customers. Another imbalance in the current discourse can be found in the technologies examined, as traditional chatbots seem to be, by far, the most common (see Table 4). Other NLP-based technologies like service robots or voice assistants are also present in our database, but far less frequent. At least in the case of service robots, this could, of course, be due to the limited access to such research objects. Nevertheless, such human-like technologies could be relevant for future research efforts, especially considering that anthropomorphism seems to be an important influencing factor in our database.

5.2. How Do Key Concepts Identified in the Scientific Discourse Affect Customer Behavior Across Different Stages of the Customer Journey, and How Are These Concepts Distributed and Interrelated Within These Stages? Concerning investigated concepts, our review found purchase intention, anthropomorphism, customer satisfaction and trust to be the most frequently studied. It was found that businesses should focus on improving the competence and anthropomorphic features of their NLP-based technologies to increase purchase intent [33, 43, 49, 62]. This could include the use of specialized chatbots for specific product categories [44], with the limitation of the conflicting findings by Tan and Liew [60]. Future studies should attempt to clarify these contradicting results, as it may be a feature that can be relatively easily implemented by practitioners and have a significant impact. In addition, integrating features that increase trust and perceived enjoyment can further increase purchase intent [25, 62]. The impact of trust is not only limited to purchase intention but seems to be a rather decisive variable already present in the adoption process of a technology [76]. Thus, this should definitely be taken into account when designing emerging technologies. Furthermore, regarding the design of the technologies, special attention should be on balancing human-like attributes to enhance positive user experiences [40] while mitigating potential negative effects like social discomfort [63]. This balance can be achieved through the context-appropriate use of anthropomorphic features, ensuring they are tailored to the specific interaction setting and customer expectations [34, 35, 46]. Here, future research could attempt to develop a framework that allows practitioners to derive the appropriate degree of anthropomorphic features based on the different interaction situations at hand. To improve customer satisfaction with NLP-based technologies, communication styles and interaction designs must be chosen that are socially oriented and emotionally intelligent [38]. Organizations should also ensure that NLP-based technologies can effectively address service failures and meet customer expectations at all touchpoints [39, 41, 65]. Current NLP-based technologies have their limitations, so an effective service recovery is important to ensure good performance. Our review also revealed a lack of studies that examine whether a particular concept is perceived and evaluated differently throughout the stages of the customer journey. The largest part of the studies focused on the evaluation of the concept and not on the possible interplay between the respective concept and the customer journey stage. In fact, only the article by Dwivedi et al. [29] compared the evaluation of different concepts between the purchase stage and postpurchase stage with important results. This could open up vast possibilities for

future research projects. The results of these future studies are urgently needed, as they are necessary to design customer interactions in which the concepts used are tailored to the respective point of the customer journey at which the customer is currently located.

The overview of currently discussed concepts in the area of NLP-based technologies provides researchers with a solid foundation for investigating the emerging technologies previously mentioned. To facilitate even easier access to this overview, we have decided to map the various concepts along the customer journey, as illustrated in Figure 10. This mapping not only clarifies the relationship between these technologies and customer interactions but also serves as a visual guide for further exploration. This visualization provides valuable new insights into the current scientific discourse as it offers an easy-to-understand illustration of which concepts are examined in which phases of the customer journey. This allows for easy identification of research gaps in relation to the different concepts and whether or not they have already been explored in the respective customer journey stages.

5.3. Other Findings. The descriptive findings revealed that the topic of NLP-based technologies in the customer journey has been gaining traction, especially since 2021 (see Figures 2 and 7), which means that the research interest in NLP-based technologies in the context of the customer journey is relatively new. The fluctuations in the number of publications per year, particularly the notable changes in 2020, could be explained by taking contextual factors such as the impact of the COVID-19 pandemic into account. Thus, this drop in publications should be considered an anomaly, and it can be assumed that the upward trend in research interest will continue, driven by emerging technologies that offer new possibilities for creating innovative customer journeys. Nevertheless, researchers should be prepared for the fact that the topic of generative AI could currently be tainted with unrealistic expectations and, according to the hype cycle model could enter the "Trough of Disillusionment" as proposed by King and Prasetyo [77]. Researchers should therefore continue to investigate technologies such as classic chatbots, as it is currently not possible to say whether the new technologies will be superior in all interaction situations.

The analysis also revealed anomalies referring to the terminology in the field of NLP-based technologies. Our results show that the terminology of some NLP-based technologies such as conversational agents, virtual agents and digital assistants is not clear as they are sometimes used interchangeably [43, 44, 69]. Thus, it is not always possible to identify why the authors use one term or the other, especially if the readers' expertise does not lie in the area. This makes the search for articles unnecessarily complicated and prevents researchers and practitioners from other disciplines from gaining a good overview of the literature. Another finding related to terminology: the analysis of co-occurrence revealed that authors often use broad, general terms such as "artificial intelligence" as keywords in the scientific discourse, although more precise terminology is available. Of course, the keywords should illustrate a research topic on a more abstract level. Nevertheless, this generalization could make it difficult to find specific research when conducting a keyword search. One possible explanation for this could be that authors from business and management fields, which are more strongly represented in our database, tend to use these general terms due to a lack of familiarity with the nuanced AI terminology. To improve scientific discourse and keyword search accuracy, researchers could adopt specific terms such as "natural language understanding" or "generative AI" instead of the broader "artificial intelligence."

Our analysis revealed that while most studies do not target specific age groups, those that do tend to focus primarily on younger participants. This may be due to the easier accessibility of groups like university students, who are frequently selected by researchers. Studies focusing specifically on senior age groups remain underrepresented. Future studies could specifically explore these age groups and their behavior when interacting with NLP-based technologies at various touchpoints, particularly as aging populations become more prevalent in most Western societies.

Most of the publications come from the field of business and management. This was partly to be expected, as the customer journey originates from this field. Nevertheless, the topic should also be of interest to more technical journals, as the concept could be used to further develop NLP-based technologies in a customer-centric way with a clear focus on different application scenarios. Many of these application scenarios will most likely have a commercial background. Therefore, early integration of the customer perspective with the various touchpoints could be useful for the design of the technologies. The current imbalance risks biasing research toward perspectives from psychology and management, thereby overlooking the technical, engineering, and data science dimensions of NLP.

In terms of application scenarios, we have identified an overhang in e-commerce (see Figure 5). This may be because online shopping is a familiar process for many respondents and allows for easy mapping and coverage of all phases of the customer journey. Furthermore, the sector's openness to and reliance on technological innovation, combined with its role as a global revenue driver, makes it highly relevant to researchers. The focus on e-commerce settings could also indicate that NLP-based technologies have not yet made the transition into offline scenarios, at least not in a noteworthy way. Future research should investigate which factors are currently hindering this transition, for example, to the point of sale.

The online focus is also evident in the data collection methods (see Figure 4), as most authors relied on online channels for the acquisition of data. The use of in-person data collection methods, supplemented by technologies such as EEG, is notably lacking. This may be due to the ease of use and efficiency of online data collection methods, which save both time and resources. However, the heavy reliance on online methods limits the diversity of insights. This issue becomes even more pronounced given the low number of field experiments combined with a high prevalence of

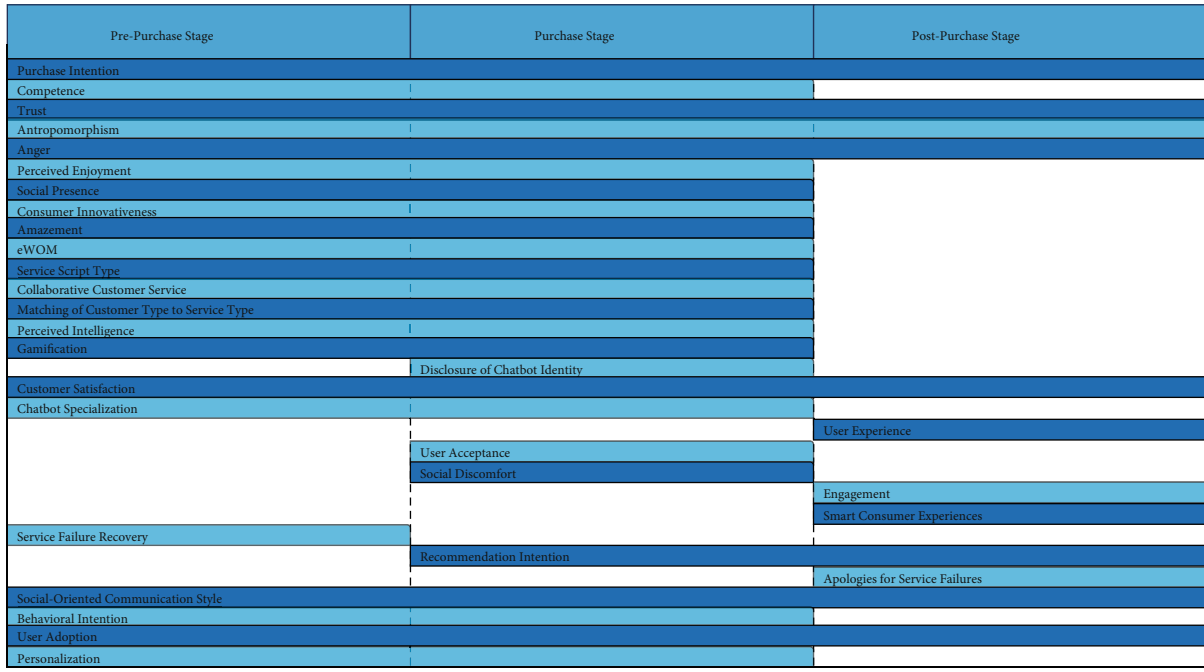

Figure 10: Mapping of concepts along the customer journey stages. Note: concepts out of studies that could not be categorized into a specific stage are not illustrated.

controlled experimental setups, ultimately reducing generalizability. Future research should attempt to compensate for this deficiency by incorporating methodologies like EEG, which could yield novel insights currently overlooked, and by focusing more on real-life settings to enhance the transferability of results. It also became clear that the authors have a strong tendency to use quantitative methods, as 94% of the articles in our database use such methods. Future research could seek to shift the focus to theory-building studies by using more qualitative research designs, particularly concerning emerging technologies. These qualitative insights may offer a more comprehensive perspective on customer sentiment, preferences, and emotional reactions to these technologies. To give the reader an overview of future research directions, we have decided to present a collection of possible research questions for future studies in Table 5.

6. Limitations

This systematic overview provides a comprehensive picture of the current scientific discourse on NLP-based technologies along the customer journey. However, every study has its limitations, and this article is no different. Although we made great efforts in terms of screening, not all eligible articles could be retrieved due to licensing issues. In addition, it was decided to concentrate the search on journal articles, a specific timeframe and certain databases that should cover the most essential publications, but it cannot be ruled out that we missed interesting articles, which were not indexed in these databases or published outside the 5-year period examined. So, while our sample is significant, it may not fully encompass the entire scope of research on NLP in the customer journey, particularly in light of the rapid advances in AI and NLP-based technologies. However, the chronology of publications shows that the field of research is only currently gaining momentum. Another limitation of this study refers to the various decisions throughout the screening process. Although we have taken measures to avoid potential errors in the audit of studies like discussing critical decisions in special meetings or documenting each screening decision in detail, it cannot be ruled out that errors have occurred. In addition to our consensus-based screening approach, future studies could also include formal measures such as Cohen's kappa. Furthermore, we have tried to include all important terms from this research area in our search string, but of course, we cannot exclude the possibility that we have overlooked articles that use different terminology to describe their research. For example, we focused on the term "service robot" rather than "robot" in general, as we only wanted to identify studies that examined this technology in the context of human interaction, which may have meant that some studies that used the more general term were not included in our database. A further limitation concerns the heterogeneity of the studies in terms of methods, research settings, variables, or sample characteristics, which makes comparability difficult. The factors mentioned so far are inherent to literature reviews. A factor that is specific to this review is that it was not always possible to clearly assign the various publications to one of the customer journey phases. As the different stages are dynamic, some articles were sorted into multiple stages.

TABLE 5: Research questions for future studies.

Addressing postpurchase interaction gaps
Can advanced NLP-based technologies like Large Language Models help bridge the current research gap in the postpurchase stage?
How can generative AI be utilized to effectively handle postpurchase interactions in the customer journey? Are there specific challenges or advantages compared to traditional chatbots?
How can NLP-based technologies help drive customer loyalty, retention and advocacy through satisfying postpurchase interactions?
Conceptual focus in research
How could a framework look like to determine the appropriate level for each of the different concepts in NLP-based technologies and how can they be tailored to different stages of the customer journey and interaction contexts?
How do certain demographic groups, for example, seniors having significant buying power or young adults having strong technology affinity, respond to advanced NLP-based technologies, and which specific concepts (e.g. trust, ease of use) are most important to each group?
What role does failure recovery play in the effectiveness of LLMs across the customer journey, and how can these systems be optimized to reduce the occurrence of "hallucinations"?
Research settings and methods
How do different NLP-based technologies perform (e.g., user experience, customer satisfaction, and user acceptance) at the different customer journey stages in offline environments such as retail stores?
Which problems do customers typically report in in-depth interviews at the different customer journey stages and can certain themes be identified?

7. Conclusion

To our knowledge, this article is the first systematic review of NLP-based technologies in the context of the customer journey. Our review has shown that research on NLP-based technologies is currently focused on the prepurchase and purchase stage, with chatbots being the most deployed technology. Among a variety of other factors studied by researchers, purchase intention and anthropomorphism were the most extensively studied in our database. With technological advances, for instance, in generative AI and the development of LLMs, arguably more complex interaction settings from postpurchase interactions could eventually become more accessible. This could be a starting point for future research. As NLP-based technologies become more widespread and impact many people, our review can offer important new insights for researchers and practitioners alike.

Data Availability Statement

The table containing the extracted data supporting the findings of this systematic review is available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

No funding was received for this research.

References

[1] H. Singh, "The Future of Natural Language Processing Industry: A New Era" 2023, https://blog.bccresearch.com/natural-language-processing-industry.

[2] M. T. Myin and K. Watchravesringkan, "Investigating Consumers' Adoption of AI Chatbots for Apparel Shopping," *Journal of Consumer Marketing* 41, no. 3 (2024): 314–327, https://doi.org/10.1108/JCM-03-2022-5234.

[3] Gartner, "Gartner Predicts Chatbots Will Become a Primary Customer Service Channel Within Five Years" 2022, Gartner. https://www.gartner.com/en/newsroom/press-releases/2022-07-27-gartner-predicts-chatbots-will-become-a-primary-customer-service-channel-within-five-years.

[4] A. Chopra, A. Prashar, and C. Sain, "Natural Language Processing," *International Journal of Technology Enhancements and Emerging Engineering Research* 1, no. 4 (2013).

[5] Y. Chang, X. Wang, J. Wang, et al., "A Survey on Evaluation of Large Language Models," *ACM Transactions on Intelligent Systems and Technology* 15, no. 3 (2024): 1–45, https://doi.org/10.1145/3641289.

[6] E. M. Rogers, *Diffusion of Innovations* (Simon and Schuster, 5th edition, 2003).

[7] K. Hu, "ChatGPT Sets Record for Fastest-Growing User Base—Analyst Note" 2023, Reuters. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/.

[8] G. Bilquise, S. Ibrahim, and K. Shaalan, "Emotionally Intelligent Chatbots: A Systematic Literature Review," *Human Behavior and Emerging Technologies* 2022, no. 1 (2022): 9601630, https://doi.org/10.1155/2022/9601630.

[9] L. Jenneboer, C. Herrando, and E. Constantinides, "The Impact of Chatbots on Customer Loyalty: A Systematic Literature Review," *Electronic Commerce Research* 17, no. 1 (2022): 212–229, https://doi.org/10.3390/jtaer17010011.

[10] D. Nwachukwu and M. Affen, "Artificial Intelligence Marketing Practices: The Way Forward to Better Customer Experience Management in Africa (Systematic Literature Review)," *International Academy Journal of Management, Marketing and Entrepreneurial Studies* 9, no. 2 (2023): 44–62.

[11] J. Rana, L. Gaur, G. Singh, U. Awan, and M. I. Rasheed, "Reinforcing Customer Journey Through Artificial Intelligence: A Review and Research Agenda," *International Journal of Emerging Markets* 17, no. 7 (2022): 1738–1758, https://doi.org/10.1108/IJOEM-08-2021-1214.

[12] K. N. Lemon and P. C. Verhoef, "Understanding Customer Experience Throughout the Customer Journey," *Journal of Marketing* 80, no. 6 (2016): 69–96, https://doi.org/10.1509/jm.15.0420.

[13] A. Haleem, M. Javaid, M. Asim Qadri, R. Pratap Singh, and R. Suman, "Artificial Intelligence (AI) Applications for Marketing: A Literature-Based Study," *International Journal of Intelligent Networks* 3 (2022): 119–132, https://doi.org/10.1016/j.ijin.2022.08.005.

[14] M. Mustak, J. Salminen, L. Plé, and J. Wirtz, "Artificial Intelligence in Marketing: Topic Modeling, Scientometric Analysis, and Research Agenda," *Journal of Business Research* 124 (2021): 389–404, https://doi.org/10.1016/j.jbusres.2020.10.044.

[15] R. Chen, P. Perry, R. Boardman, and H. McCormick, "Augmented Reality in Retail: A Systematic Review of Research Foci and Future Research Agenda," *International Journal of Retail & Distribution Management* 50, no. 4 (2021): 498–518, https://doi.org/10.1108/IJRDM-11-2020-0472.

[16] H. Kumar, "Augmented Reality in Online Retailing: A Systematic Review and Research Agenda," *International Journal of Retail & Distribution Management* 50, no. 4 (2022): 537–559, https://doi.org/10.1108/IJRDM-06-2021-0287.

[17] M. M. Mariani, N. Hashemi, and J. Wirtz, "Artificial Intelligence Empowered Conversational Agents: A Systematic Literature Review and Research Agenda," *Journal of Business Research* 161 (2023): 113838, https://doi.org/10.1016/j.jbusres.2023.113838.

[18] A. Ramesh and V. Chawla, "Chatbots in Marketing: A Literature Review Using Morphological and Co-Occurrence Analyses," *Journal of Interactive Marketing* 57, no. 3 (2022): 472–496, https://doi.org/10.1177/10949968221095549.

[19] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and for the PRISMA Group, "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement," *BMJ* 339, no. jul21 1 (2009): b2535, https://doi.org/10.1136/bmj.b2535.

[20] N. Al-Ansari, D. Al-Thani, and R. S. Al-Mansoori, "User-Centered Evaluation of Explainable Artificial Intelligence (XAI): A Systematic Literature Review," *Human Behavior and Emerging Technologies* 2024, no. 1 (2024): 4628855, https://doi.org/10.1155/2024/4628855.

[21] J. Q. Pérez, T. Daradoumis, and J. M. M. Puig, "Rediscovering the Use of Chatbots in Education: A Systematic Literature Review," *Computer Applications in Engineering Education* 28, no. 6 (2020): 1549–1565, https://doi.org/10.1002/cae.22326.

[22] J. Yang, Y.-L. Chen, L. Y. Por, and C. S. Ku, "A Systematic Literature Review of Information Security in Chatbots," *Applied Sciences* 13, no. 11 (2023): 6355, https://doi.org/10.3390/app13116355.

[23] VOSviewer—Features—Highlights, "VOSviewer" Retrieved 24 January 2025, from https://www.vosviewer.com/features/highlights/.

[24] J. K. Hentzen, A. Hoffmann, R. Dolan, and E. Pala, "Artificial Intelligence in Customer-Facing Financial Services: A Systematic Literature Review and Agenda for Future Research," *International Journal of Bank Marketing* 40, no. 6 (2022): 1299–1336, https://doi.org/10.1108/IJBM-09-2021-0417.

[25] M. C. Han, "The impact of anthropomorphism on consumers' purchase decision in chatbot commerce," *Journal of Internet Commerce* 20, no. 1 (2021): 46–65, https://doi.org/10.1080/15332861.2020.1863022.

[26] PRISMA, 2020, http://prisma-statement.org/prismastatement/flowdiagram.aspx?AspxAutoDetectCookieSupport=1.

[27] A. R. Malekian, Y. Jabarzadeh, and A. Fazlzadeh, "The Impact of CSR and Corporate Ability on Purchase Intention in e-Commerce Websites: An Examination Across Genders," *The International Review of Retail, Distribution and Consumer Research* 33, no. 4 (2023): 415–440, https://doi.org/10.1080/09593969.2023.2235101.

[28] S. Chen, X. Li, K. Liu, and X. Wang, "Chatbot or Human? The Impact of Online Customer Service on Consumers' Purchase Intentions," *Psychology & Marketing* 40, no. 11 (2023): 2186–2200, https://doi.org/10.1002/mar.21862.

[29] Y. K. Dwivedi, J. Balakrishnan, A. M. Baabdullah, and R. Das, "Do Chatbots Establish "Humanness" in the Customer Purchase Journey? An Investigation Through Explanatory Sequential Design," *Psychology & Marketing* 40, no. 11 (2023): 2244–2271, https://doi.org/10.1002/mar.21888.

[30] M. G. Elmashhara, R. De Cicco, S. C. Silva, M. Hammerschmidt, and M. L. Silva, "How Gamifying AI Shapes Customer Motivation, Engagement, and Purchase Behavior," *Psychology & Marketing* 41, no. 1 (2024): 134–150, https://doi.org/10.1002/mar.21912.

[31] G. Pizzi, V. Vannucci, V. Mazzoli, and R. Donvito, "I, Chatbot! The Impact of Anthropomorphism and Gaze Direction on Willingness to Disclose Personal Information and Behavioral Intentions," in *Psychology & Marketing* 40, no. 7 (2023): 1372–1387, https://doi.org/10.1002/mar.21813.

[32] T. Fernandes and E. Oliveira, "Understanding Consumers' Acceptance of Automated Technologies in Service Encounters: Drivers of Digital Voice Assistants Adoption," *Journal of Business Research* 122 (2021): 180–191, https://doi.org/10.1016/j.jbusres.2020.08.058.

[33] R. Roy and V. Naidoo, "Enhancing Chatbot Effectiveness: The Role of Anthropomorphic Conversational Styles and Time Orientation," *Journal of Business Research* 126 (2021): 23–34, https://doi.org/10.1016/j.jbusres.2020.12.051.

[34] C. Crolic, F. Thomaz, R. Hadi, and A. T. Stephen, "Blame the Bot: Anthropomorphism and Anger in Customer-Chatbot Interactions," *Journal of Marketing* 86, no. 1 (2022): 132–148, https://doi.org/10.1177/00222429211045687.

[35] A. M. Garvey, T. Kim, and A. Duhachek, "Bad News? Send an AI. Good News? Send a Human," *Journal of Marketing* 87, no. 1 (2023): 10–25, https://doi.org/10.1177/00222429211066972.

[36] X. Cheng, Y. Bao, A. Zarifis, W. Gong, and J. Mou, "Exploring Consumers' Response to Text-Based Chatbots in e-Commerce: The Moderating Role of Task Complexity and Chatbot Disclosure," *Internet Research* 32, no. 2 (2021): 496–517, https://doi.org/10.1108/INTR-08-2020-0460.

[37] S. C. C. Silva, R. De Cicco, B. Vlacic, and M. G. Elmashhara, "Using Chatbots in E-Retailing—How to Mitigate Perceived Risk and Enhance the Flow Experience," *International Journal of Retail & Distribution Management* 51, no. 3 (2023): 285–305, https://doi.org/10.1108/IJRDM-05-2022-0163.

[38] Y. Xu, J. Zhang, and G. Deng, "Enhancing Customer Satisfaction With Chatbots: The Influence of Communication Styles and Consumer Attachment Anxiety," *Frontiers in Psychology* 13 (2022): 902782, https://doi.org/10.3389/fpsyg.2022.902782.

[39] J. Yun and J. Park, "The Effects of Chatbot Service Recovery With Emotion Words on Customer Satisfaction, Repurchase Intention, and Positive Word-Of-Mouth," *Frontiers in Psychology* 13 (2022): https://doi.org/10.3389/fpsyg.2022.922503.

[40] K. Klein and L. F. Martinez, "The Impact of Anthropomorphism on Customer Satisfaction in Chatbot Commerce: An Experimental Study in the Food Sector," *Electronic Commerce Research* 23, no. 4 (2022): 2789–2825, https://doi.org/10.1007/s10660-022-09562-8.

[41] S. Wang, Q. Yan, and L. Wang, "Task-Oriented vs. Social-Oriented: Chatbot Communication Styles in Electronic Commerce Service Recovery," in *Electronic Commerce Research* (Business Source Elite, 2023), 1–33, https://doi.org/10.1007/s10660-023-09741-1.

[42] Y. Li, Z. Gan, and B. Zheng, "How Do Artificial Intelligence Chatbots Affect Customer Purchase? Uncovering the Dual Pathways of Anthropomorphism on Service Evaluation," *Information Systems Frontiers* 27, no. 1 (2023): 283–300, https://doi.org/10.1007/s10796-023-10438-x.

[43] M. Sharma, S. Joshi, S. Luthra, and A. Kumar, "Impact of Digital Assistant Attributes On Millennials' Purchasing Intentions: A Multi-Group Analysis Using PLS-SEM, Artificial Neural Network and fsQCA," *Information Systems Frontiers* 26, no. 3 (2022): 943–966, https://doi.org/10.1007/s10796-022-10339-5.

[44] T. W. Liew and S.-M. Tan, "Exploring the Effects of Specialist Versus Generalist Embodied Virtual Agents in a Multi-Product Category Online Store," *Telematics & Informatics* 35, no. 1 (2018): 122–135, https://doi.org/10.1016/j.tele.2017.10.005.

[45] M. Mende, M. L. Scott, J. van Doorn, D. Grewal, and I. Shanks, "Service Robots Rising: How Humanoid Robots Influence Service Experiences and Elicit Compensatory Consumer Responses," in *Journal of Marketing Research* 56, no. 4 (2019): 535–556, https://doi.org/10.1177/0022243718822827.

[46] H. Shin, I. Bunosso, and L. R. Levine, "The Influence of Chatbot Humour on Consumer Evaluations of Services," *International Journal of Consumer Studies* 47, no. 2 (2023): 545–562, https://doi.org/10.1111/ijcs.12849.

[47] K. O. Wolbers and N. Walter, "Silence Is Silver, but Speech Is Golden: Intelligent Voice Assistants (IVAs) and their Impact on a Brand's Customer Decision Journey With a Special Focus on Trust and Convenience—A Qualitative Consumer Analysis in the Netherlands," *IUP Journal of Brand Management* 18, no. 1 (2021): 7–31.

[48] P. Caleb-Solly, S. Dogramadzi, C. A. G. J. Huijnen, and H. Heuvel, "Exploiting Ability for Human Adaptation to Facilitate Improved Human-Robot Interaction and Acceptance," *Information Society* 34, no. 3 (2018): 153–165, https://doi.org/10.1080/01972243.2018.1444255.

[49] J. Y.-S. Chang, J.-H. Cheah, X.-J. Lim, and A. M. Morrison, "One Pie, Many Recipes: The Role of Artificial Intelligence Chatbots in Influencing Malaysian Solo Traveler Purchase Intentions," *Tourism Management Perspectives* 49 (2023): 101191, https://doi.org/10.1016/j.tmp.2023.101191.

[50] A. Kovács, B. Van Looy, and B. Cassiman, "Exploring the Scope of Open Innovation: A Bibliometric Review of a Decade of Research," *Scientometrics* 104, no. 3 (2015): 951–983, https://doi.org/10.1007/s11192-015-1628-0.

[51] N. J. van Eck and L. Waltman, "VOSviewer Manual," *Leiden: Univeristeit Leiden* 1, no. 1 (2013): 1–53.

[52] J.-S. Chen, T.-T.-Y. Le, and D. Florence, "Usability and Responsiveness of Artificial Intelligence Chatbot on Online Customer Experience in E-Retailing," in *International Journal of Retail & Distribution Management* 49, no. 11Emerald Group Publishing Ltd., 2021), 1512–1531, https://doi.org/10.1108/IJRDM-08-2020-0312.

[53] V. Kaushal and R. Yadav, "Learning Successful Implementation of Chatbots in Businesses From B2B Customer Experience Perspective," *Concurrency and Computation: Practice and Experience* 35, no. 1 (2023): e7450, https://doi.org/10.1002/cpe.7450.

[54] M. Haupt, A. Rozumowski, J. Freidank, and A. Haas, "Seeking Empathy or Suggesting a Solution? Effects of Chatbot Messages on Service Failure Recovery," *Electronic Markets* 33, no. 1 (2023): 1–22, https://doi.org/10.1007/s12525-023-00673-0.

[55] A. Jansom, T. Srisangkhajorn, and W. Limarunothai, "How Chatbot E-Services Motivate Communication Credibility and Lead to Customer Satisfaction: The Perspective of Thai Consumers in the Apparel Retailing Context," *Innovative Marketing* 18, no. 3 (2022): 13–25, https://doi.org/10.21511/im.18(3).2022.02.

[56] X. Luo, S. Tong, Z. Fang, and Z. Qu, "Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases," *Marketing Science* 38, no. 6 (2019): 937–947, https://doi.org/10.1287/mksc.2019.1192.

[57] Y. Ruan and J. Mezei, "When Do AI Chatbots Lead to Higher Customer Satisfaction Than Human Frontline Employees in Online Shopping Assistance? Considering Product Attribute Type," *Journal of Retailing and Consumer Services* 68 (2022): 103059, https://doi.org/10.1016/j.jretconser.2022.103059.

[58] S. Sands, C. Ferraro, C. Campbell, and H.-Y. Tsao, "Managing the Human-Chatbot Divide: How Service Scripts Influence Service Experience," *Journal of Service Management* 32, no. 2 (2021): 246–264, https://doi.org/10.1108/JOSM-06-2019-0203.

[59] M. Shumanov and L. Johnson, "Making Conversations With Chatbots More Personalized," *Computers in Human Behavior* 117 (2021): 106627, https://doi.org/10.1016/j.chb.2020.106627.

[60] S.-M. Tan and T. W. Liew, "Multi-Chatbot or Single-Chatbot? The Effects of M-Commerce Chatbot Interface on Source Credibility, Social Presence, Trust, and Purchase Intention," *Human Behavior and Emerging Technologies* 2022, no. 1 (2022): 2501538, https://doi.org/10.1155/2022/2501538.

[61] N. Upadhyay and A. Kamble, "Why Can't We Help but Love Mobile Banking Chatbots? Perspective of Stimulus-Organism-Response," *Journal of Financial Services Marketing* 29, no. 3 (2023): 855–872, https://doi.org/10.1057/s41264-023-00237-5.

[62] C. Yen and M.-C. Chiang, "Trust Me, if You Can: A Study on the Factors That Influence Consumers' Purchase Intention Triggered by Chatbots Based on Brain Image Evidence and Self-Reported Assessments," *Behaviour & Information Technology* 40, no. 11 (2021): 1177–1194, https://doi.org/10.1080/0144929X.2020.1743362.

[63] Y. Li, Z. Liang, Y. Wang, and Y. Chang, "The Negative Effect of Service Robots' Affective Human-Likeness on Consumer Satisfaction in Frontline Service Encounters," *International Journal of Hospitality Management* 115 (2023): 103603, https://doi.org/10.1016/j.ijhm.2023.103603.

[64] Y. Okafuji, S. Song, J. Baba, Y. Yoshikawa, and H. Ishiguro, "Influence of Collaborative Customer Service by Service Robots and Clerks in Bakery Stores," *Frontiers in Robotics and AI* 10 (2023): https://doi.org/10.3389/frobt.2023.1125308.

[65] S. Wenguo and W. Yongwei, "Facilitation of Customer Empathy: The Effect of Robot Apology on Customer Reaction Following a Service Failure," *Journal of Marketing Development & Competitiveness* 16, no. 2 (2022): 44–57.

[66] R. E. Bawack, S. F. Wamba, and K. D. A. Carillo, "Exploring the Role of Personality, Trust, and Privacy in Customer Experience Performance During Voice Shopping: Evidence From SEM and Fuzzy Set Qualitative Comparative Analysis," in *International Journal of Information Management* 58 (2021ELSEVIER SCI LTD., 2021), 102309, https://doi.org/10.1016/j.ijinfomgt.2021.102309.

[67] P. Kautish, S. Purohit, R. Filieri, and Y. K. Dwivedi, "Examining the Role of Consumer Motivations to Use Voice Assistants for Fashion Shopping: The Mediating Role of Awe Experience and eWOM," *Technological Forecasting and Social Change* 190 (2023): 122407, https://doi.org/10.1016/j.techfore.2023.122407.

[68] P. Li, C. Wu, and C. Spence, "Comparing the Influence of Visual Information and the Perceived Intelligence of Voice Assistants When Shopping for Sustainable Clothing Online," *Humanities and Social Sciences Communications* 10, no. 1 (2023): 1–12, https://doi.org/10.1057/s41599-023-02244-2.

[69] S. Danckwerts, L. Meißner, and C. Krampe, "Examining User Experience of Conversational Agents in Hedonic Digital Services—Antecedents and the Role of Psychological Ownership," *Journal of Service Management Research (SMR)* 3, no. 3 (2019): 111–125, https://doi.org/10.15358/2511-8676-2019-3-111.

[70] D. Schmid, D. Staehelin, A. Bucher, M. Dolata, and G. Schwabe, "Does Social Presence Increase Perceived Competence?," *Proceedings of the ACM on Human-Computer Interaction* 6, no. GROUP (2022): 1–22, https://doi.org/10.1145/3492845.

[71] H. J. Kang, J. Shin, and K. Ponto, "How 3D Virtual Reality Stores Can Shape Consumer Purchase Decisions: The Roles of Informativeness and Playfulness," *Journal of Interactive Marketing* 49, no. 1 (2020): 70–85, https://doi.org/10.1016/j.intmar.2019.07.002.

[72] K. Machová, M. Szabóova, J. Paralič, and J. Mičko, "Detection of Emotion by Text Analysis Using Machine Learning," *Frontiers in Psychology* 14 (2023): https://doi.org/10.3389/fpsyg.2023.1190326.

[73] R. Michel-Villarreal, E. Vilalta-Perdomo, D. E. Salinas-Navarro, R. Thierry-Aguilera, and F. S. Gerardou, "Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT," *Education Sciences* 13, no. 9 (2023): 856, https://doi.org/10.3390/educsci13090856.

[74] V. Soni, "Large Language Models for Enhancing Customer Lifecycle Management," *Journal of Empirical Social Science Studies* 7, no. 1 (2023): 67–89.

[75] S. Liu, K. Halder, Z. Qi, et al., "Towards Long Context Hallucination Detection" 2025, Amazon Science. https://www.amazon.science/publications/towards-long-context-hallucination-detection.

[76] T. Bahmanziari, J. M. Pearson, and L. Crosby, "Is Trust Important in Technology Adoption? A Policy Capturing Approach," *Journal of Computer Information Systems* 43, no. 4 (2003): 46–54, https://doi.org/10.1080/08874417.2003.11647533.

[77] S. King and J. Prasetyo, "Assessing Generative A.I. Through the Lens of the 2023 Gartner Hype Cycle for Emerging Technologies: A Collaborative Autoethnography," *Frontiers in Education* 8 (2023): 1300391, https://doi.org/10.3389/feduc.2023.1300391.