Journal of Research, Innovation and Technologies
https://www.acadlore.com/journals/JORIT

Exploring Research on ChatGPT in Online Consumer Reviews: A Systematic Literature Review and Thematic Analysis

Muhamad Izaidi Ishak[1*], Aini Khalida Muslim[2], Naveed Ahmad[3]

[1] Faculty of Business and Communication, Universiti Malaysia Perlis, 02600 Arau, Malaysia
[2] Fakulti Pengurusan Teknologi dan Teknousahawanan, Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Malaysia
[3] Department of Management Sciences, Faculty of Business Education, Social Sciences, & Humanities, Isra University, 71000 Hyderabad, Pakistan

[*] Correspondence: Muhamad Izaidi Ishak (izaidi@unimap.edu.my)

Received: 11-08-2025 **Revised:** 01-06-2026 **Accepted:** 01-08-2026

Citation: Ishak, M. I., Muslim, A. K., & Ahmad, N. (2026). Exploring research on ChatGPT in online consumer reviews: A systematic literature review and thematic analysis. *J. Res. Innov. Technol.*, *5*(1), 63–84. https://doi.org/10.56578/jorit050104.

Abstract: Studies on ChatGPT within the context of online consumer reviews (OCRs) have emerged as part of the broader exploration of generative AI across multiple disciplines. However, to date, no research has systematically examined the current research focus or other key aspects related to the application of ChatGPT in OCRs. To address this gap, this study conducts a systematic literature review to identify dominant research focus areas, highlight existing research gaps, and propose directions for future research. Guided by the PRISMA 2020 protocol and employing a thematic analysis approach, 22 relevant studies were analysed, revealing three overarching themes: (1) ChatGPT for review analytics, (2) ChatGPT for review modeling and evaluation, and (3) ChatGPT for review management. The findings indicate that current research primarily emphasizes ChatGPT's potential as an analytical tool for OCR datasets, enabling the extraction of valuable and actionable insights for both marketers and researchers. In addition, the review identifies growing concern regarding fake reviews and highlights the emerging use of ChatGPT-generated synthetic reviews as datasets for developing fake review detection models, offering a practical alternative for studies facing challenges in obtaining high-quality training data. Finally, findings related to the third theme demonstrate ChatGPT's utility in supporting managerial responses to customer reviews, providing insights into its role in enhancing customer relationship management. Overall, this review suggests that research on ChatGPT in OCRs remains at an early stage but offers significant insights and opportunities for future investigation in this emerging field.

Keywords: Artificial intelligent; ChatGPT, Consumer feedback; Online consumer reviews; Systematic literature review; Thematic analysis

JEL Classification: D12, O33, M31

1. Introduction

In the contemporary digital marketplace, consumers increasingly rely on online platforms to inform their purchasing decisions. Among these platforms, online consumer reviews (OCRs) have emerged as one of the most influential information sources, shaping consumer perceptions, guiding preferences, and influencing purchase intentions (Su et al., 2025). Prior research highlights that OCRs function as an essential decision-support mechanism, enabling consumers to evaluate service quality, credibility, and reliability before making choices (Jeong & Lee, 2024). This influence is particularly pronounced among younger generations, who demonstrate a strong dependence on online feedback when selecting products and services (Roy et al., 2017). From a managerial perspective, OCRs not only affect firm performance and sales outcomes (Babić Rosario et al., 2016; Harun et al., 2025; Ishak & Harun, 2023) but also offer rich insights into consumer needs, preferences, and satisfaction levels (Baier et al., 2025). Collectively, OCRs serve as a critical interface between consumers and firms, shaping both

purchasing behavior and strategic decision-making.

As the volume and complexity of OCRs continue to expand, both researchers and practitioners have increasingly turned to artificial intelligence (AI) to analyze and manage review content more efficiently. AI technologies are now deeply embedded across industries (Collins et al., 2021) and are commonly defined as systems capable of simulating human reasoning, learning, and problem-solving processes (Berente et al., 2021). Within this broader AI landscape, ChatGPT, which is powered by large language models (LLMs), has rapidly gained prominence. Trained on extensive textual datasets, LLMs are able to generate coherent, contextually relevant, and human-like text (Campbell IV et al., 2024). Reflecting its versatility, ChatGPT has been widely adopted across diverse domains, including healthcare (Garg et al., 2023; Miao et al., 2023), education (Xiao et al., 2025), tourism (Kumamoto & Joho, 2025), entrepreneurship (Duong et al., 2025), and content creation (Hwang & Lee, 2025).

Despite this growing adoption, research examining ChatGPT within the context of OCRs remains fragmented. Existing studies tend to focus on isolated applications, such as review generation (Amos & Zhang, 2024; Knoedler et al., 2024), sentiment analysis (Cheng et al., 2024), or review management (Tan et al., 2025; Wayne Litvin & Pei-Sze Tan, 2024), without offering a consolidated understanding of the broader research landscape. As a result, there is limited clarity regarding the dominant research themes, methodological approaches, industry contexts, and emerging trends that characterize this body of literature. This lack of systematic synthesis constrains theoretical development and hinders the identification of research gaps or the establishment of coherent future research agendas at the intersection of ChatGPT and OCRs. Taken together, this fragmentation highlights the need for a systematic literature review to organize existing knowledge and guide future research on the use of ChatGPT in the context of OCRs.

To address this gap, this study conducts a systematic literature review (SLR) integrating thematic analysis to explore the application of ChatGPT in OCRs. Specifically, this review aims to identify prevailing research focus areas, dominant themes, industry distributions, and methodological patterns in the existing literature, as well as to highlight underexplored issues and future research opportunities. Accordingly, this study is guided by the following research questions:

RQ1: What are the current publication fields, industry distributions, and research methodologies in studies examining ChatGPT and OCRs?

RQ2: What key themes and dominant research focus areas have emerged in studies on the application of ChatGPT in OCRs?

RQ3: What research gaps and future directions can be identified based on the current state of scholarship in this area?

This review offers several contributions. First, it provides a structured and comprehensive synthesis of prior research to map the intellectual landscape surrounding ChatGPT and OCRs. Second, it identifies thematic and methodological trends, clarifying where scholarly attention has been concentrated and where it remains limited. Third, by highlighting research gaps and emerging directions, this study offers guidance for future investigations and supports the responsible, evidence-based application of generative AI in online review contexts.

The remainder of this paper is organized as follows. Section 2 describes the SLR methodology, including the search strategy and selection criteria. Section 3 presents the descriptive and thematic findings. Section 4 discusses the theoretical and managerial implications, and Section 5 concludes by outlining research gaps, limitations, and directions for future studies.

2. Methodology

This study adopted an SLR approach combined with thematic analysis, drawing on established methodological guidance (Braun & Clarke, 2006; Nasrabadi et al., 2024; Page et al., 2021; Pham et al., 2023). The selection of an SLR over alternative review methods, such as bibliometric or meta-analytical analyses, was guided by several considerations. First, the SLR approach offers richer qualitative insights by emphasising content-level and contextual understanding rather than relying solely on quantitative indicators such as citation counts (Ishak et al., 2025; Nikseresht et al., 2024; Rojas-Sánchez et al., 2023). This qualitative orientation is particularly well suited to thematic analysis, as it enables the systematic identification, organisation, and interpretation of recurring patterns and meanings across studies, thereby facilitating deeper theoretical integration and synthesis. Such an approach allows researchers to capture nuanced insights that purely quantitative review methods may overlook (Marzi et al., 2024). As noted by Paul & Criado (2020), a well-structured SLR plays a critical role in consolidating existing knowledge within a research domain, thereby enhancing conceptual clarity and contextual depth.

Second, compared with other review methodologies, the SLR technique provides a more systematic and rigorous framework for addressing clearly defined research objectives (Ahmad et al., 2022; Angioi & Hiller, 2023). When combined with thematic analysis, the SLR process enables the transparent and replicable identification of key themes, sub-themes, and research patterns across the reviewed studies. This involves the formulation of explicit research aims, the selection of appropriate academic databases, and the application of a structured procedure for coding, theme development, and synthesis (Braun & Clarke, 2006; Nasrabadi et al., 2024). Together,

this methodological integration strengthens the credibility, interpretive depth, and comprehensiveness of the review findings. Accordingly, the combined SLR and thematic analysis approach was deemed the most appropriate method for achieving the objectives of this study.

2.1 Review Protocol

This study employed an SLR approach to explore the current research focus on the application of ChatGPT in the context of OCRs, guided by the PRISMA 2020 protocol proposed by Page et al. (2021). The PRISMA 2020 framework provides a set of 27 reporting items intended to enhance transparency, rigor, and reproducibility in systematic reviews. Accordingly, the review process was structured into three main phases—identification, screening, and inclusion—as illustrated in Figure 1.

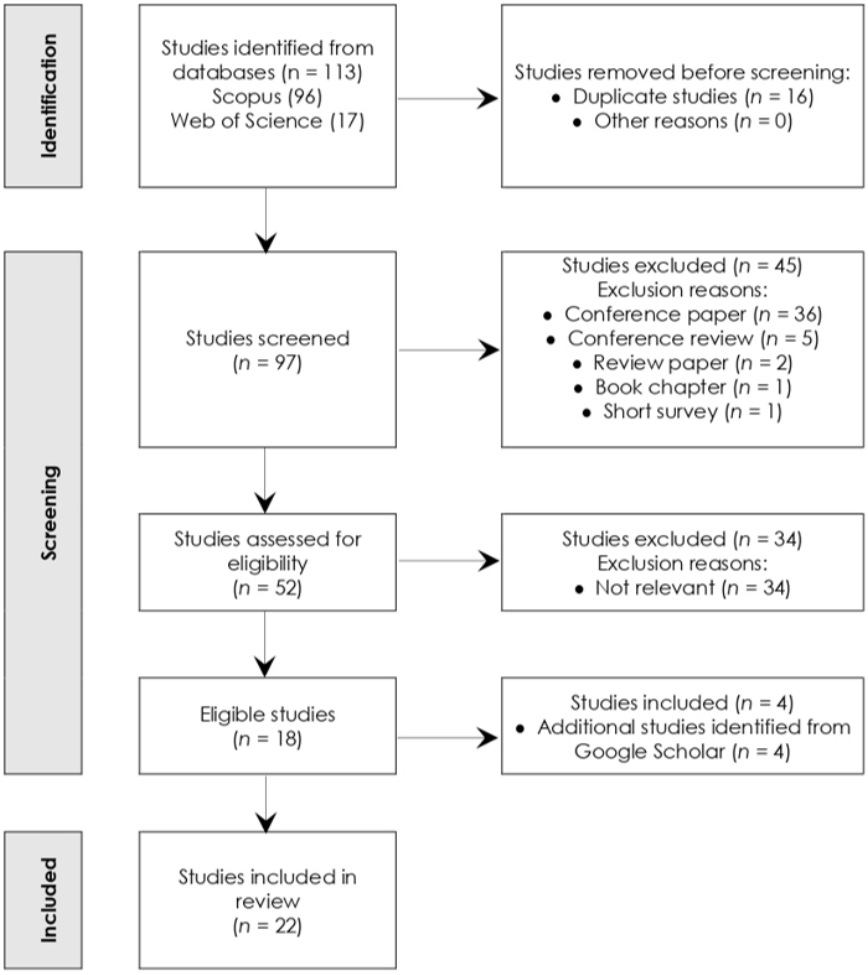

Figure 1. Systematic review search process following PRISMA 2020 protocol

2.1.1 Identification

The literature search was conducted using two prominent academic databases, Elsevier's Scopus and Clarivate's Web of Science (WoS), selected for their extensive coverage and comprehensive indexing compared with other academic databases (Carrera-Rivera et al., 2022; Nikseresht et al., 2024). The eligibility criteria restricted inclusion to peer-reviewed journal articles published between 2022 and 2025 to ensure the relevance and timeliness of the evidence base (Muslim & Harun, 2022; Snyder, 2019). The year 2022 was selected as the initial timeframe, as ChatGPT was introduced at the end of that year and no relevant studies were available prior to its emergence.

The search was conducted between July and December 2025. The development of the search strategy, including keyword selection, was adapted from Nasrabadi et al. (2024). Keywords were designed to capture studies examining ChatGPT in the context of OCRs. The final list of search terms included "ChatGPT," "online reviews," "online customer reviews," "online consumer reviews," "customer reviews," "user reviews," "consumer reviews," and "user-generated content." Notably, only "ChatGPT" was used, rather than broader terms such as large language models or generative AI, to maintain conceptual specificity. This restriction ensured the identification of studies explicitly addressing ChatGPT. Although some relevant studies may employ broader terminology in their

titles or abstracts, author-assigned keywords typically reflect the core content and research focus of an article (Corrin et al., 2022). Our review found that 80% (n = 20) of the selected articles included "ChatGPT" as an author-assigned keyword. Accordingly, prioritizing "ChatGPT" as the primary search term enhanced relevance and reduced the inclusion of studies examining generative AI more broadly without directly focusing on ChatGPT in the context of OCRs.

The search was executed using the advanced document search function by combining the selected keywords with the Boolean operators "OR" and "AND" across the title, abstract, and keyword fields (Khan et al., 2024). For the Scopus database, the following comprehensive search string was applied: (TITLE-ABS-KEY("ChatGPT") AND TITLE-ABS-KEY ("online reviews" OR "online customer reviews" OR "online consumer reviews" OR "customer reviews" OR "user reviews" OR "consumer reviews" OR "user-generated content")). For the Web of Science (WoS) database, the corresponding search string was: TS = ("ChatGPT") AND TS = ("online reviews" OR "online customer reviews" OR "online consumer reviews" OR "customer reviews" OR "user reviews" OR "consumer reviews" OR "user-generated content"). The search yielded 113 records from Scopus and 17 records from WoS. After removing 16 duplicate records, 97 unique records remained and were subsequently subjected to the screening phase.

2.1.2 Screening

This phase is a critical component of the SLR process, as it enhances reliability and mitigates potential bias. During this stage, 97 identified papers were screened for eligibility. Table 1 summarises the inclusion and exclusion criteria applied during the screening process. The screening process was facilitated by the detailed publication metadata provided by the selected databases, which categorized documents as research articles, conference papers, or other publication types. Publications categorized as conference papers (n = 36), conference reviews (n = 5), review articles (n = 2), book chapters (n = 1), and short surveys (n = 1) were excluded, resulting in a total of 45 exclusions. The decision to exclude conference papers, conference reviews, book chapters, and short surveys aligns with methodological studies suggesting that such documents may provide limited detail and hinder the extraction of meaningful and reliable data (Scherer & Saldanha, 2019; Taylor et al., 2014). Review articles were excluded to avoid duplication of evidence and potential bias, as they synthesize findings from primary studies rather than provide original empirical data (Ahmad et al., 2022). Consequently, these exclusions, supported by the literature, help to maintain a focused and rigorous evidence base.

Following this eligibility screening, the remaining records were subjected to a more detailed examination at the title and abstract level in accordance with the PRISMA 2020 protocol (Page et al., 2021). For this reason, two researchers independently reviewed the titles and abstracts of the 52 records and discussed any inconsistencies until consensus was reached (Zupic & Čater, 2015). In cases of disagreement regarding which articles should proceed to full-text screening, consensus was achieved through discussion. Next, the same two researchers independently screened the full-text articles for inclusion. Any disagreements at this stage were resolved through discussion to reach a final inclusion or exclusion decision. As a result, 34 articles were identified as not relevant, and 18 articles were included in the final review.

Table 1. The inclusion and exclusion criteria applied during the screening process

Criterion	Inclusion Criteria	Exclusion Criteria
Research relevance	Studies examining the application, evaluation, or implications of ChatGPT in the context of OCRs	Studies focusing on AI, LLMs, or generative AI without explicit relevance to ChatGPT in OCRs
Document type	Peer-reviewed journal articles	Conference papers, conference reviews, book chapters, short surveys, editorials, and review articles
Language	English-language publications	Non-English publications
Publication period	Studies published between 2022 and 2025	Studies published outside the defined period
Data availability	Full-text articles accessible	Abstract-only or inaccessible full texts
Indexing source	Studies indexed in Scopus or Web of Science	Studies indexed exclusively in other databases

2.1.3 Inclusion

To ensure completeness, an additional search was conducted using Google Scholar with the same set of keywords (Ahmad et al., 2022; Pham et al., 2023). This process identified four additional papers, which were subsequently checked in the Scopus and WoS databases and found to be indexed in Scopus. The full texts of these papers were then retrieved. Two researchers carefully examined the titles and abstracts of these papers and reached a consensus to include them for further review. Consequently, a total of 22 studies were finalized for inclusion in the review. The final step of this phase involved compiling a summary table containing essential details from each selected study, including authors, titles, publication years, journal sources, methodologies, key findings, and other

relevant information (see Appendix A).

2.1.4 Quality assessment and risk of bias

Consistent with prior exploratory and thematic systematic literature reviews, this study did not conduct a formal risk-of-bias or quality assessment of the included studies (Paul & Criado, 2020; Snyder, 2019). The primary objective of the review was to map research themes, methodological trends, and dominant research focus areas rather than to evaluate effect sizes or intervention outcomes (Marzi et al., 2024). Furthermore, all included studies were peer-reviewed journal articles indexed in Scopus and Web of Science, which provides an initial level of quality assurance (Carrera-Rivera et al., 2022). This approach aligns with previous SLRs that emphasize conceptual synthesis and theory development over methodological appraisal (Nikseresht et al., 2024; Rojas-Sánchez et al., 2023).

2.2 Thematic Analysis

The selection of relevant papers in this study was conducted in accordance with the PRISMA 2020 protocol (Page et al., 2021), resulting in the identification of 22 studies for analysis. Subsequently, thematic analysis was employed to identify key themes and to establish the dominant research focus areas within the selected literature, thereby directly addressing the second research question of the study. The analysis followed the well-established framework proposed by Braun & Clarke (2006), which is widely regarded as one of the most influential approaches in the social sciences. In addition, the guidance provided by Maguire & Delahunt (2017) was used to support the systematic execution of the thematic analysis.

Following Braun & Clarke (2006), thematic analysis was conducted through six sequential phases: (1) familiarisation with the data, (2) generation of initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the final write-up. During the familiarisation phase, the reviewers examined the titles, objectives, methodologies, and key findings of the selected articles, with additional full-text readings undertaken where necessary to develop a deeper understanding of the studies (Maguire & Delahunt, 2017; Nasrabadi et al., 2024). Building on this process, initial codes were generated manually using an inductive, open-coding approach, with no pre-established codes applied. These codes were progressively compared, sorted, and refined to identify broader themes (Braun & Clarke, 2006; Clark et al., 2019; Liñán & Fayolle, 2015). Manual coding was considered appropriate given the manageable sample size of 22 studies, as it enabled close engagement with the data and facilitated deeper analytical insight.

To ensure coding consistency and reliability, two independent reviewers conducted the coding process. Inter-coder reliability was assessed using Cohen's Kappa based on a contingency table constructed from their coding decisions. The analysis yielded a Kappa value of 0.775, indicating substantial agreement beyond chance. In cases of coding discrepancies, the coders discussed differences to clarify code definitions and reach consensus, thereby ensuring consistent application of the coding framework and enhancing reproducibility.

The third phase focused on identifying relevant themes. Consistent with Braun & Clarke (2006), themes were defined based on their relevance to the research question rather than their frequency of occurrence. At this stage, the analysis shifted from individual codes to a broader thematic level, whereby related codes derived from the reviewed studies were systematically organised, compared, and grouped into potential themes. This process involved examining how different codes converged to form overarching themes that represent the dominant research focus areas within the literature. To support this organisation and synthesis, a detailed table illustrating the development of initial codes into sub-themes and overarching themes is provided in Appendix B. In the fourth and fifth phases, the preliminary themes were reviewed and refined to minimise overlap and enhance conceptual clarity, resulting in a coherent and well-defined thematic structure that captured distinct and meaningful aspects of the research focus across the included studies (Ryan & Bernard, 2003).

In the final phase, the refined themes were consolidated and articulated in relation to the research objectives. This phase involved synthesising the themes into a coherent narrative and presenting them with supporting evidence from the reviewed studies to ensure transparency and analytical rigour. The final write-up aimed to clearly communicate the dominant research focus areas and their interrelationships within the existing literature (Braun & Clarke, 2006; Thorpe et al., 2005). The entire thematic analysis procedure was conducted manually to maintain close interpretive engagement with the data.

3. Findings

This study conducted an SLR on the application of ChatGPT in OCRs, with a primary emphasis on thematic analysis. A descriptive quantitative analysis was first undertaken, followed by an inductive thematic analysis to identify and synthesise key themes characterising current research on ChatGPT in OCRs. The findings from the thematic analysis constitute the core qualitative contribution of this study.

3.1 Descriptive Quantitative Analysis

This section presents a descriptive quantitative analysis of the research landscape, focusing on trends across publication fields, industries, and research methodologies. The analysis provides an overview of the structure of existing research and highlights the dominant areas of scholarly attention.

3.1.1 Publication field

The 22 reviewed papers were published across a diverse range of academic journals, underscoring the multidisciplinary nature of research on ChatGPT in OCR, as shown in Table 2. Publications span several major fields, including Hospitality and Tourism, featuring journals such as Cornell Hospitality Quarterly (Wayne Litvin & Pei-Sze Tan, 2024), Tourism Management (Tan et al., 2025), and the International Journal of Hospitality Management (Choi et al., 2024), and Retail and Consumer Services, represented by the Journal of Retailing and Consumer Services (Baier et al., 2025; Zhao et al., 2025). Additionally, a significant portion of the research appeared in Computer Science and Information Technology outlets, including Telematics and Informatics (Amos & Zhang, 2024), Information Processing and Management (Xylogiannopoulo et al., 2024), and IEEE Access (Mathebula et al., 2024), reflecting the technological foundations of ChatGPT-related studies. Other contributions emerged from Healthcare and Medicine journals such as PLoS ONE (Li et al., 2025) and the Journal of Plastic, Reconstructive and Aesthetic Surgery (Knoedler et al., 2024), as well as multidisciplinary platforms including Scientific Reports (Su et al., 2025), Sustainability (Switzerland) (Jeong & Lee, 2024), and Technology in Society (Koc et al., 2023). These journals show how the topic bridges multiple disciplines, from tourism and consumer behavior to computing and health sciences.

Table 2. Distribution of reviewed studies across academic journals

Journal Field	Journal Name	No. of Study	No. of Citation
Hospitality & Tourism	*Cornell Hospitality Quarterly*	1	6
	Journal of Hospitality and Tourism Technology	2	4
	Tourism Management	1	6
	International Journal of Hospitality Management	1	10
	International Journal of Contemporary Hospitality Management	1	3
	Journal of Qualitative Research in Tourism	1	0
Computer Science & Information Technology	*Telematics and Informatics*	1	12
	Knowledge-Based Systems	1	0
	Information Processing and Management	1	8
	Computers	1	0
	IEEE Access	2	15
Multidisciplinary & General Science	*Scientific Reports*	1	1
	Sustainability (Switzerland)	1	14
	Technology in Society	1	63
Medical & Surgical	*Journal of Plastic, Reconstructive and Aesthetic Surgery*	1	9
	Journal of Foot and Ankle Surgery	1	3
	PLoS ONE	1	1
Retailing & Consumer Services	*Journal of Retailing and Consumer Services*	2	2
Operations Research & Management Science	*Annals of Operations Research*	1	9

3.1.2 Industry distribution

The reviewed studies span a range of industry contexts, demonstrating ChatGPT's broad relevance and adaptability across service-oriented sectors, as shown in Figure 2. Research contexts are predominantly concentrated in the hospitality and tourism industry, followed by e-commerce and retail and healthcare. Hospitality and tourism emerge as the primary research focus, with 12 papers, of which 10 examine hotel reviews and 2 focus on tourism platforms such as TripAdvisor and Airbnb (Cheng et al., 2024; Morini-Marrero et al., 2025; Tan et al., 2025; Wayne Litvin & Pei-Sze Tan, 2024). The e-commerce and retail sector represents the second major area, encompassing four studies that analyze online reviews from platforms such as Amazon, Yelp, Google Play Store, and restaurant feedback datasets (Baier et al., 2025; Su et al., 2025; Zhao et al., 2025). In contrast, healthcare and medical contexts form a smaller yet emerging category, with three studies focusing on patient and surgeon reviews on platforms such as Haodf.com and Healthgrades (Knoedler et al., 2024; Li et al., 2025). Additional sectors appear less frequently: the food and beverage industry is represented by one study (McCloskey et al., 2024), cross-industry service contexts—including airlines and the clothing industry—are examined in one study (Rosete et al., 2025), and the finance sector appears in one study (Mathebula et al., 2024). Collectively, these patterns highlight the

predominance of hospitality- and retail-focused research while underscoring ChatGPT's expanding role in analyzing consumer experiences across diverse service industries.

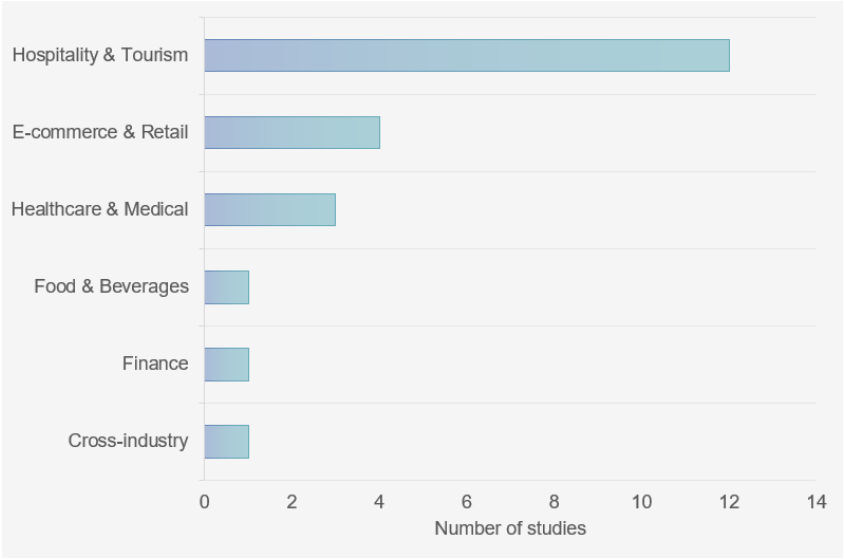

Figure 2. Industry contexts of the reviewed studies

3.1.3 Methodology analysis

The selected studies employed a range of methodological approaches, as summarised in Table 3. Computational text mining and NLP-based approaches were most prevalent, with many studies applying large-scale text analysis, machine learning, predictive modelling, and sentiment or topic modelling to analyse OCRs and examine ChatGPT's analytical capabilities (Baier et al., 2025; Botunac et al., 2024; Casciato & Mateen, 2024; Cheng et al., 2024; Choi et al., 2024; Koc et al., 2023; Li et al., 2025; Mathebula et al., 2024; Ramos-Henriquez & Morini-Marrero, 2025; Su et al., 2025; Xylogiannopoulos et al., 2024; Zhao et al., 2025). Experimental designs formed the second largest group, employing perception, comparative, detection, and quasi-experimental approaches to assess the credibility and behavioural effects of ChatGPT-generated reviews (Amos & Zhang, 2024; Knoedler et al., 2024; Morini-Marrero et al., 2025; Wayne Litvin & Pei-Sze Tan, 2024). Perception and behavioural studies were less common, focusing on user or managerial evaluations of AI-generated content through surveys or qualitative assessments (Hajra, 2023; Jeong & Lee, 2024; Rosete et al., 2025). Only one study adopted a mixed-methods approach, integrating computational and experimental insights to examine managerial response contexts (Tan et al., 2025). Overall, the dominance of computational and experimental approaches reflects the field's emphasis on empirical validation, while the limited use of perception-oriented and mixed-methods designs indicates opportunities for deeper contextual and theory-driven research. Figure 3 provides a visualization of the methodologies utilized in the reviewed studies.

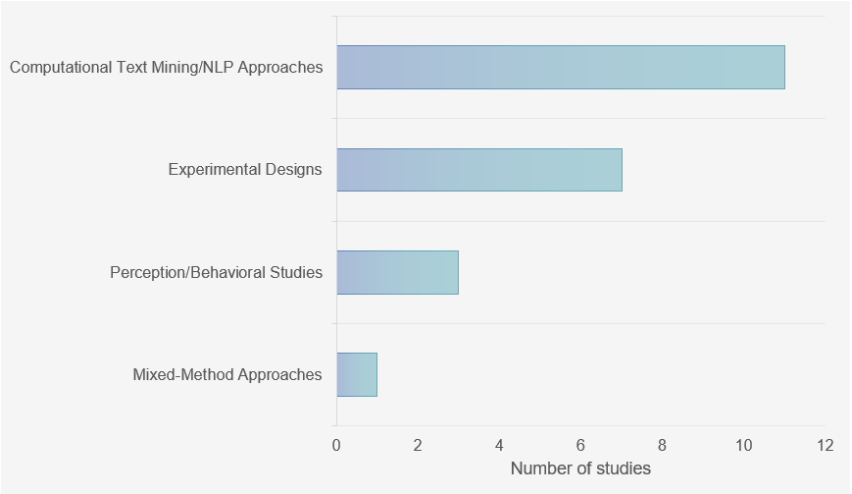

Figure 3. Methodological distribution of the reviewed studies

Table 3. Summary of methodological approaches in the selected studies

Methodological Approach	No. of Papers	Description
Computational Text Mining/NLP Approaches	11	This category comprises studies that primarily apply computational techniques to analyse or generate OCRs, including large-scale text analysis (Zhao et al., 2025), supervised machine learning (Su et al., 2025), predictive modelling (Baier et al., 2025), sentiment and topic modelling (Cheng et al., 2024; Li et al., 2025), and other NLP-based or AI-driven analytical approaches (Botunac et al., 2024; Casciato & Mateen, 2024; Choi et al., 2024; Koc et al., 2023; Mathebula et al., 2024; Ramos-Henriquez & Morini-Marrero, 2025; Xylogiannopoulos et al., 2024).
Experimental Designs	7	This category includes studies employing experimental or quasi-experimental designs to examine the performance, credibility, or behavioural effects of ChatGPT-generated reviews, such as perception experiments (Litvin & Tan, 2024), controlled comparative experiments (Amos & Zhang, 2024), human–AI comparative detection experiments (Knoedler et al., 2024), and quasi-experimental designs (Morini-Marrero et al., 2025).
Perception/Behavioral Studies	3	This category covers studies that focus on users' or managers' perceptions, evaluations, or behavioural responses to ChatGPT-generated reviews, typically using surveys or evaluative assessments (Hajra, 2023; Jeong & Lee, 2024; Rosete et al., 2025).
Mixed-Methods Approaches	1	This category includes studies that explicitly combine computational analysis with experimental or qualitative techniques, such as mixed-method experimental designs comparing managerial responses to ChatGPT-generated reviews (Tan et al., 2025).

3.2 Results of Thematic Analysis

The thematic analysis of the selected studies, following Braun & Clarke (2006), resulted in the emergence of three overarching themes reflecting the current research focus in this area: (1) ChatGPT for review analytics, (2) ChatGPT for review modeling and evaluation, and (3) ChatGPT for review management. The formation of these overarching themes from the initial codes is visualized in Figure 4. These findings provide insight into how ChatGPT has been applied within OCR research and highlight potential avenues for future investigation.

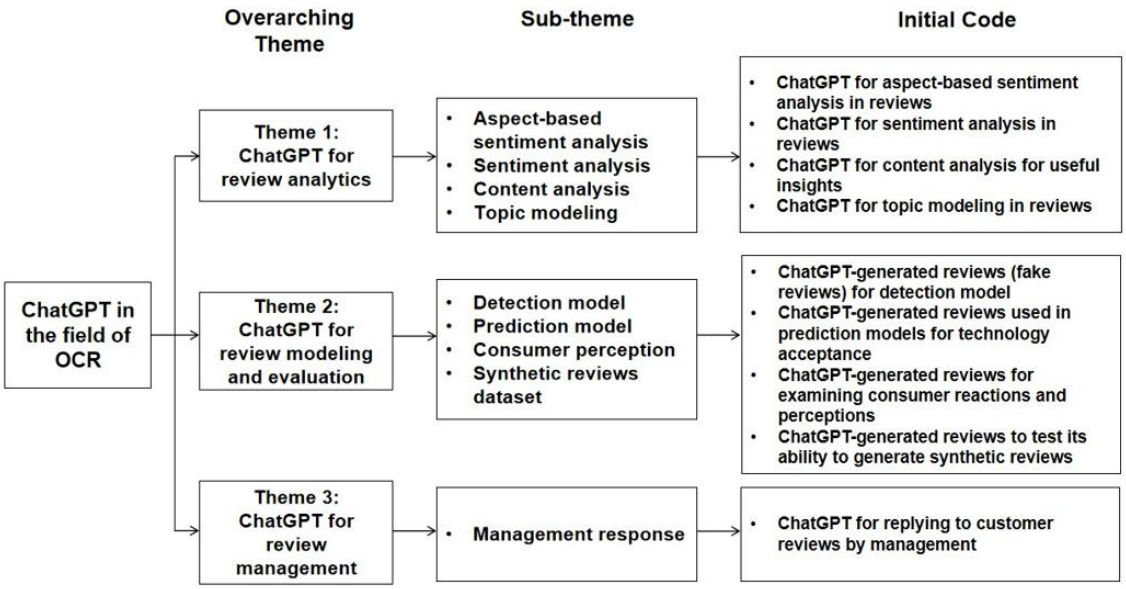

Figure 4. Thematic framework of ChatGPT research in OCRs

3.2.1 ChatGPT for review analytics

ChatGPT for review analytics emerged as the dominant overarching theme in the selected studies (n = 11; 50%), indicating that prior research primarily positions ChatGPT as an analytical tool for extracting meaning from OCRs. As summarised in Appendix B, this theme comprises several analytically oriented sub-themes, reflecting the diverse applications of ChatGPT in review analysis across domains. The largest sub-theme involves aspect-based sentiment analysis (n = 4), which enhances the precision and interpretability of OCR analysis by capturing sentiment at the attribute level (Botunac et al., 2024; Falatouri et al., 2024). A substantial subset of studies focuses

on sentiment analysis (n = 3), examining ChatGPT's ability to identify emotional polarity and evaluative tone in OCRs and demonstrating accuracy comparable to established NLP approaches, particularly in large-scale textual datasets (Casciato & Mateen, 2024; Cheng et al., 2024; Mathebula et al., 2024). Rather than replacing traditional models, ChatGPT is commonly positioned as a complementary tool that enhances sentiment interpretation through more context-aware language processing.

Another stream of research applies content analysis (n = 2) and topic modeling (n = 2), in which ChatGPT is used to extract themes, patterns, and higher-level insights from unstructured OCR data. Prior studies indicate that ChatGPT can synthesise large volumes of textual reviews into coherent analytical outputs, supporting the identification of recurring issues, consumer expectations, and experiential dimensions embedded in OCRs (Hajra, 2023; Morini-Marrero et al., 2025). In this context, ChatGPT is employed either as a primary coding aid or as an interpretive layer that enhances traditional qualitative content analysis. Similarly, in topic modeling applications, ChatGPT is used to identify latent themes by summarising recurring patterns and discussion points expressed by consumers, thereby facilitating a more interpretable synthesis of review content (McCloskey et al., 2024; Ramos-Henriquez & Morini-Marrero, 2025). Taken together, these findings indicate that research on ChatGPT for review analytics is methodologically diverse, with applications spanning multiple analytical functions rather than a single homogeneous approach.

3.2.2 ChatGPT for review modeling and evaluation

ChatGPT for review modeling and evaluation emerged as the second dominant overarching theme in the reviewed literature (n = 8; 36%), with the majority of studies (n = 5) focusing on fake review detection. This emphasis reflects growing concern over the authenticity and credibility of AI-generated online reviews. As detailed in Appendix B, this theme comprises four closely related sub-themes: ChatGPT-generated reviews for detection models, prediction models, consumer perception evaluation, and ChatGPT's capability in generating synthetic review datasets. The most prominent sub-theme centres on the use of ChatGPT-generated reviews for detection models, with five studies employing these generated reviews to train or test models designed to distinguish AI-generated from human-authored content (Knoedler et al., 2024; Su et al., 2025; Xylogiannopoulos et al., 2024; Xylogiannopoulos et al., 2025; Zhao et al., 2025). These studies highlight the increasing difficulty users face in reliably identifying AI-generated reviews, underscoring the heightened risk of misleading or deceptive content.

Complementing these detection-focused investigations, other studies examine alternative applications of ChatGPT-generated reviews for modeling and evaluation purposes. One study incorporates these reviews into prediction models to assess technology acceptance outcomes (Baier et al., 2025), while another explores consumer perceptions of AI-generated reviews, demonstrating that suspected AI authorship negatively affects perceived usefulness and authenticity (Amos & Zhang, 2024). In addition, Rosete et al. (2025) evaluated ChatGPT's capacity to generate synthetic review datasets by comparing AI-generated and human-written content across multiple datasets, identifying limitations related to linguistic diversity and repetitiveness. Collectively, these studies indicate that while ChatGPT-generated reviews offer methodological utility for model development and evaluation, they also raise substantive concerns regarding credibility, detection, and consumer trust within OCR ecosystems.

3.2.3 ChatGPT for review management

ChatGPT for review management emerged as a distinct but less prevalent overarching theme in the reviewed studies (n = 3; 14%). In line with Appendix B, this theme is represented by a single, clearly defined sub-theme—management response—which focuses on the use of ChatGPT to assist managers in replying to online customer reviews. Within this sub-theme, prior studies consistently examine ChatGPT's ability to generate managerial responses that meet service recovery and communication standards. For example, Koc et al. (2023) compared human-authored management replies on TripAdvisor with responses generated by ChatGPT and found that AI-generated responses were more efficient and, in some cases, more effective in addressing customer concerns. Similarly, Wayne Litvin & Pei-Sze Tan (2024) examined consumer perceptions of human versus ChatGPT-generated management responses and reported that ChatGPT could convincingly replicate authentic managerial communication, suggesting its potential value for responding to reviews at scale.

Despite these advantages, the reviewed studies also identify important limitations. Tan et al. (2025) further explored ChatGPT's role in online service recovery, highlighting both its strengths and constraints. In particular, they found that although customers often struggled to distinguish between human- and AI-generated responses, explicit disclosure of AI authorship led to lower satisfaction and reduced purchase intentions. This finding suggests that while ChatGPT can support efficiency and consistency in review management, maintaining perceived authenticity remains critical. Overall, the evidence indicates that ChatGPT functions most effectively as a supportive tool for review management rather than a full substitute for human managerial engagement.

4. Discussion

Prior studies have examined ChatGPT's applications in OCRs; however, the literature remains fragmented, with

no systematic synthesis of the current research landscape. To address this gap, this study conducted an SLR of ChatGPT research in the context of OCRs. The analysis identified three overarching themes: (1) ChatGPT for review analytics, (2) ChatGPT for review modeling and evaluation, and (3) ChatGPT for review management. Among these, review analytics emerged as the most dominant theme, indicating the primary focus of existing OCR research.

4.1 Theoretical Implications

Our findings indicate that current research on the application of ChatGPT in the OCR field predominantly focuses on analytical purposes, with particular emphasis on content and sentiment analysis. This finding provides a significant contribution to the literature, as existing methods for analyzing OCR data typically rely on traditional text-mining software such as Leximancer and ATLAS.ti. For example, Arasli et al. (2020) utilized Leximancer to analyze OCRs from cruise travelers in order to identify key service quality perceptions influencing value-for-money ratings. Similarly, Olorunsola et al. (2024) analyzed OCRs from eco-friendly hotels using Leximancer to identify key themes related to customer satisfaction. Further details on other analytical tools used for OCR data analysis, including ATLAS.ti, KH Coder, R software, and related tools, are extensively reviewed by Ishak et al. (2025). Although these tools perform effectively in processing textual data, they present notable limitations, most prominently the need for highly skilled personnel to operate the software and interpret the analyses (Engstrom et al., 2022). In contrast, AI tools such as ChatGPT are more user-friendly (Skjuve et al., 2023), thereby facilitating the analysis of OCR datasets and enabling marketers to obtain timely insights into consumers' current perceptions of a brand or company. These findings suggest that ChatGPT has the potential to enhance OCR data analysis and support faster, more valuable insight generation.

However, existing studies report mixed findings regarding ChatGPT's analytical accuracy, particularly in sentiment analysis tasks. While some research shows that ChatGPT's classification performance can be competitive with, or even exceed, traditional NLP and machine learning approaches (Fatouros et al., 2023; Lossio-Ventura et al., 2024), other studies report greater variability and context-dependent weaknesses. For example, comparative evaluations indicate that although ChatGPT demonstrates competitive overall accuracy, it may lag traditional algorithms on certain datasets and task types (Arif & Aladdin, 2025). Similarly, sentiment classification studies suggest that while ChatGPT performs well overall, it struggles with specific categories, particularly neutral sentiment, reflecting challenges in interpreting ambiguous emotional expressions (Jonathan et al., 2025). Likewise, Rebora et al. (2023) found that although ChatGPT's sentiment analysis performance is comparable to established automated tools at the aggregate level, it aligns less well with individual human judgments. Overall, these findings indicate that ChatGPT's effectiveness in sentiment analysis is context-dependent and should be interpreted with caution in OCR research.

Our thematic analysis of the selected papers revealed growing scholarly attention to fake review detection within the second theme, reflecting increasing concern about the prevalence and impact of fake reviews on digital platforms. Prior research has provided substantial evidence of the severity of this issue; for example, He et al. (2022) demonstrate that fake reviews significantly distort product ratings on Amazon, showing that fake reviews are purchased across a broad range of products on the platform. This concern can be interpreted through a Signaling Theory lens, whereby OCRs function as market signals that convey cues about underlying product or service quality and reduce information asymmetry (Siering et al., 2018; Spence, 1978; Xu et al., 2024). Fake reviews, including AI-generated reviews, thus operate as deceptive signals that imitate credible experiential feedback without incurring genuine consumption-based costs, thereby undermining the reliability of online review ecosystems.

Building on this concern, a key finding within this theme is the growing number of studies that develop fake review detection models using ChatGPT-generated reviews, as evidenced by five reviewed studies (Knoedler et al., 2024; Su et al., 2025; Xylogiannopoulos et al., 2024; Xylogiannopoulos et al., 2025; Zhao et al., 2025). Earlier research highlighted the limited availability of high-quality datasets as a major challenge in developing reliable fake review detection models (Aslam et al., 2019; Wu et al., 2020). In contrast, our findings indicate that recent studies increasingly leverage ChatGPT-generated reviews to train and test detection models, signaling a clear methodological shift away from data scarcity toward AI-enabled data generation. Accordingly, this study contributes by systematically identifying this emerging shift in fake review research and clarifying ChatGPT's evolving role as a data-generation tool in advancing fake review detection within the OCR literature.

In addition to its role in modeling fake review detection, ChatGPT also demonstrates constructive applications within the online review ecosystem, particularly in supporting managerial responses to customer feedback. Prior studies highlight its practical benefits in enhancing customer relationship management. For example, Wayne Litvin & Pei-Sze Tan (2024) showed that ChatGPT is effective and useful in improving management responses to online reviews in the hotel industry. Similarly, Koc et al. (2023) found that ChatGPT is both effective and efficient in summarizing and generating responses to customer complaints within the hospitality context. Our findings regarding ChatGPT's usefulness for management responses are consistent with the conclusions of Simetgo et al.

(2025), who systematically reviewed 40 articles and established the positive role of ChatGPT in customer support, particularly in providing 24/7 assistance, rapid responses, and consistent service quality.

From a Trust Theory perspective, these benefits primarily strengthen functional dimensions of trust, such as perceived ability and reliability, by signalling managerial competence and responsiveness (Mayer & Davis, 1999; Seok & Chiew, 2013). However, our review also reveals a more nuanced trust-related challenge. While AI-assisted responses may enhance operational efficiency, they may simultaneously weaken relational trust by raising concerns about authenticity and benevolence, particularly when customers become aware that responses are AI-generated. This concern is consistent with prior research on Trust Theory showing that user trust in AI-enabled systems is undermined when transparency, perceived benevolence, and socio-ethical expectations are weakened (Bach et al., 2024). Notably, our SLR identified evidence that prospective customers report lower purchase intentions when management responses are explicitly disclosed as AI-generated (Tan et al., 2025). Collectively, these findings suggest that organizations need to balance the operational advantages of ChatGPT with careful consideration of how AI-mediated interactions shape customer trust perceptions.

4.2 Practical Implications

This SLR provides targeted practical insights for marketers, platform managers, and policymakers involved in the management and use of OCRs. For marketers and managers, the findings highlight ChatGPT's strong potential as an analytical tool for extracting valuable insights from OCR datasets. Within the hospitality and tourism sector, ChatGPT can be deployed for near real-time review analysis on platforms such as TripAdvisor to monitor guest feedback related to service quality dimensions such as cleanliness, staff professionalism, accommodation comfort, food and beverage quality, and overall travel experience. By enabling rapid identification of emerging issues, dominant consumer sentiments, and key drivers of satisfaction and dissatisfaction, ChatGPT supports timelier and data-driven decision-making aimed at improving service quality and overall performance. Hotel managers, destination marketers, and tourism operators may leverage ChatGPT to support proactive service recovery and continuous experience enhancement throughout the customer journey, helping to sustain positive experiences while addressing areas requiring improvement.

For platform managers, however, the findings also raise important concerns regarding the increasing prevalence of fake reviews. The generation of false or AI-generated feedback poses a direct threat to the credibility of OCR platforms and may erode consumer trust if left unaddressed. In the e-commerce context, ChatGPT-generated reviews may be used constructively as training data to enhance fake review detection systems tailored to different product categories. As such, review platforms such as Amazon should invest in more robust detection mechanisms and consider integrating ChatGPT-powered review verification or screening tools as decision-support systems for human moderators, while actively collaborating with researchers and technology providers to identify and filter out deceptive content. Enhanced transparency mechanisms, such as review verification labels or disclosure indicators, may further help maintain platform integrity.

From a policy and regulatory perspective, the findings indicate a growing need for governance frameworks that address the use of generative AI in online reviews. Policymakers may consider developing and enforcing explicit AI-generated content disclosure standards that require review platforms to clearly label ChatGPT-generated or AI-assisted reviews. Such measures would establish accountability mechanisms for non-compliance, thereby protecting consumers from misleading information and enhancing transparency across various industries, particularly e-commerce platforms. At the platform governance level, managers are encouraged to integrate ChatGPT-powered review verification and screening tools into existing moderation systems to automatically flag suspicious or AI-generated content, prioritise reviews for human evaluation, and monitor emerging manipulation patterns. The implementation of standardised disclosure indicators and AI-assisted moderation guidelines may further help balance technological innovation with consumer protection, while consumers are advised to cross-check information across multiple sources to make more informed purchasing and travel decisions.

4.3 Limitations and Future Research

This study has several limitations that should be considered when interpreting the findings. First, although the review employed two leading academic databases, Scopus and Web of Science, to enhance coverage and rigour, relevant studies indexed in other databases or emerging outlets may not have been captured. Second, the review was restricted to English-language publications, which may introduce language bias and limit representation of research conducted in non-English-speaking contexts. Third, the relatively narrow date range reflects the emerging nature of ChatGPT research but may constrain the generalisability of the findings as the field continues to develop rapidly. Fourth, while the thematic analysis followed established procedures, a degree of subjectivity in coding and theme development is inherent in qualitative synthesis. Finally, the review did not incorporate backward and forward citation tracking, which may have resulted in the omission of influential studies not identified through the database search strategy alone.

Beyond these methodological considerations, a substantive limitation lies in the absence of a detailed, practice-oriented framework to guide marketers in applying ChatGPT for online review analysis. This limitation is closely related to the exploratory nature of the present study, which aims to map and synthesise existing research rather than to develop prescriptive models. Although this systematic review identifies key themes and research patterns, it does not propose a structured or step-by-step framework that practitioners can readily implement.

Building on these limitations, future research could be advanced through a more structured agenda encompassing theoretical, methodological, and contextual dimensions. From a theoretical perspective, future studies could integrate established theories from marketing, information systems, and communication such as trust, persuasion, and technology acceptance theories to explain how AI-assisted and AI-generated reviews influence consumer judgment, credibility assessment, and decision-making. Such theoretical integration would help move the literature beyond descriptive applications toward deeper explanatory and theory-building contributions.

From a methodological perspective, future research could extend beyond the predominantly experimental and modeling-based approaches by employing mixed-method designs, longitudinal analyses, and qualitative inquiries. Combining computational techniques with interviews or surveys involving managers and consumers would allow for a more nuanced understanding of how ChatGPT is used, evaluated, and managed in real-world review environments.

From a contextual perspective, future studies could broaden the scope of investigation by examining a wider range of industries and cross-cultural settings. Most existing research remains concentrated in service-oriented and digitally mature contexts, limiting the generalisability of findings. Expanding empirical coverage to diverse cultural, regulatory, and industry environments would enhance understanding of how contextual factors shape perceptions, ethical concerns, and managerial adoption of ChatGPT in OCRs.

In addition, the increasing use of ChatGPT-generated content calls for more rigorous empirical investigations into its reliability and ethical implications. While this review focuses on OCRs, future research could extend to high-stakes domains such as healthcare, finance, journalism, and customer service, where misinformation and automation bias may have more serious consequences. Such efforts would contribute to the development of stronger governance mechanisms, improved detection techniques, and the responsible integration of generative AI across digital platforms.

5. Conclusion

This study systematically reviewed recent literature on the application of ChatGPT within the context of OCRs to identify key themes, current research focus, and directions for future inquiry. Guided by the PRISMA 2020 protocol, three dominant themes emerged: ChatGPT for review analytics, ChatGPT for review modeling and evaluation, and ChatGPT for review management. Across the reviewed studies, research is largely situated within marketing and information systems disciplines, predominantly examines service-based industries, and relies mainly on experimental, modeling, and computational text analysis approaches. The findings indicate that existing research primarily investigates the use of ChatGPT as an analytical tool for extracting insights from consumer feedback, particularly in sentiment analysis and pattern identification. In parallel, a growing body of work examines ChatGPT-generated reviews for modeling, validation, and authenticity assessment, reflecting increasing concerns regarding the trustworthiness and ethical implications of AI-generated content. A smaller but notable stream of studies focuses on review management, highlighting ChatGPT's role in supporting managerial responses and customer communication. Overall, this review suggests that research on ChatGPT and OCRs remains at an early stage of development. The emerging emphasis on review analytics, synthetic review evaluation, fake review detection, and responsible AI use presents promising opportunities for future research. Future studies could expand this field by incorporating diverse industry contexts, adopting cross-cultural perspectives, and employing mixed-method approaches to deepen understanding of ChatGPT's role in digital marketing and online review ecosystems.

Author Contributions

Conceptualization, M.I.I.; methodology, M.I.I. and N.A; validation, M.I.I., A.K.M., and N.A.; formal analysis, M.I.I. and A.K.M.; investigation, M.I.I.; data curation, M.I.I.; writing—original draft preparation, M.I.I.; writing—review and editing, M.I.I. and N.A.; supervision, M.I.I.; project administration, M.I.I.; funding acquisition, M.I.I. and A.K.M. All authors have read and agreed to the published version of the manuscript.

Funding

This work is funded by Universiti Malaysia Perlis (UniMAP) Publication Fund.

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Acknowledgements

The authors would like to express their sincere appreciation to Universiti Malaysia Perlis (UniMAP), particularly the Research Management Centre (RMC), through the UniMAP Publication Fund, for their encouragement in completing this research.

Conflicts of Interest

The authors declare no conflict of interest.

Declaration on the Use of Generative AI and AI-assisted Technologies

Artificial intelligence (AI)-based tools were used solely to improve the clarity, grammar, and readability of this manuscript. The use of AI did not influence the study's conceptualization, data analysis, interpretation of results, or conclusions. All intellectual contributions, ideas, and content remain the sole responsibility of the authors.

References

Ahmad, N., Harun, A., Khizar, H. M. U., Khalid, J., & Khan, S. (2022). Drivers and barriers of travel behaviors during and post COVID-19 pandemic: A systematic literature review and future agenda. *J. Tour. Futures.*, 1–23. https://doi.org/10.1108/jtf-01-2022-0023.

Algaraady, J. & Mahyoob, M. (2025). Exploring ChatGPT's potential for augmenting post-editing in machine translation across multiple domains: Challenges and opportunities. *Front. Artif. Intell.*, *8*, 1526293. https://doi.org/10.3389/frai.2025.1526293.

Amos, C. & Zhang, L. (2024). Consumer reactions to perceived undisclosed ChatGPT usage in an online review context. *Telemat. Inform.*, *93*, 102163. https://doi.org/10.1016/j.tele.2024.102163.

Angioi, M. & Hiller, C. E. (2023). Systematic literature reviews. In *Research Methods in the Dance Sciences* (pp. 265–280). University Press of Florida.

Arasli, H., Saydam, M. B., & Kilic, H. (2020). Cruise travelers' service perceptions: A critical content analysis. *Sustainability*, *12*(17), 6702. https://doi.org/10.3390/su12176702.

Arif, B. K. & Aladdin, A. M. (2025). A comparative analysis of ChatGPT and traditional machine learning algorithms on real-world data. *Kurd. J. Appl. Res.*, *10*(2), 93–118. https://doi.org/10.24017/science.2025.2.8.

Aslam, U., Jayabalan, M., Ilyas, H., & Suhail, A. (2019). A survey on opinion spam detection methods. *Int. J. Sci. Technol. Res.*, *8*(9).

Atalan, A., Keskin, A., & Özer, S. (2024). Exploring ChatGPT's role in healthcare management: Opportunities, ethical considerations, and future directions. *Int. J. Healthc. Manag.*, 1–13. https://doi.org/10.1080/20479700.2024.2444102.

Babić Rosario, A., Sotgiu, F., De Valck, K., & Bijmolt, T. H. A. (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *J. Mark. Res.*, *53*(3), 297–318. https://doi.org/10.1509/jmr.14.0380.

Bach, T. A., Khan, A., Hallock, H., Beltrão, G., & Sousa, S. (2024). A systematic literature review of user trust in AI-enabled systems: An HCI perspective. *Int. J. Hum. Comput. Interact.*, *40*(5), 1251–1266. https://doi.org/10.1080/10447318.2022.2138826.

Baier, D., Karasenko, A., & Rese, A. (2025). Measuring technology acceptance over time using transfer models based on online customer reviews. *J. Retail. Consum. Serv.*, *85*, 104278. https://doi.org/10.1016/j.jretconser.2025.104278.

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS Q.*, *45*(3), 1433–1450. https://doi.org/10.25300/misq/2021/16274.

Bettayeb, A. M., Abu Talib, M., Sobhe Altayasinah, A. Z., & Dakalbab, F. (2024). Exploring the impact of ChatGPT: Conversational AI in education. *Front. Educ.*, *9*, 1379796. https://doi.org/10.3389/feduc.2024.1379796.

Botunac, I., Brkić Bakarić, M., & Matetić, M. (2024). Comparing fine-tuning and prompt engineering for multi-class classification in hospitality review analysis. *Appl. Sci.*, *14*(14), 6254. https://doi.org/10.3390/app14146254.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa.

Campbell IV, W. A., Chick, J. F., Shin, D., & Makary, M. S. (2024). Understanding ChatGPT for evidence-based utilization in interventional radiology. *Clin. Imaging.*, *108*, 110098. https://doi.org/10.1016/j.clinimag.2024.110098.

Carrera-Rivera, A., Ochoa, W., Larrinaga, F., & Lasa, G. (2022). How-to conduct a systematic literature review: A quick guide for computer science research. *MethodsX.*, *9*, 101895. https://doi.org/10.1016/j.mex.2022.101895.

Casciato, D. J. & Mateen, S. (2024). AI-assisted sentiment analysis of ACFAS fellowship-trained foot and ankle surgeon online reviews. *J. Foot Ankle Surg.*, *63*(5), 577–579. https://doi.org/10.1053/j.jfas.2024.05.014.

Caumartin, G., Qin, Q., Chatragadda, S., Panjrolia, J., Li, H., & Costa, D. E. (2025). Exploring the potential of Llama models in automated code refinement: A replication study. In *2025 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)* (pp. 681–692). https://doi.org/10.1109/saner64311.2025.00070.

Cheng, X., Chen, Y., Wang, P., Zhou, Y., Wei, X., Luo, W., & Duan, Q. (2024). A novel ChatGPT-based multimodel framework for tourism review mining: A case study on China's five sacred mountains. *J. Hosp. Tour. Technol.*, *15*(4), 592–609. https://doi.org/10.1108/jhtt-06-2023-0170.

Choi, M., Choi, Y., Bangura, E., & Kim, D. (2024). ChatGPT or online review: Which is a better determinant of customers' trust in Airbnb listings and stay intention? *Int. J. Hosp. Manag.*, *122*, 103852. https://doi.org/10.1016/j.ijhm.2024.103852.

Clark, C., Harrison, C., & Gibb, S. (2019). Developing a conceptual framework of entrepreneurial leadership: A systematic literature review and thematic analysis. *Int. Rev. Entrep.*, *17*(3), 347–384.

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *Int. J. Inf. Manag.*, *60*, 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383.

Corrin, L., Thompson, K., Hwang, G., & Lodge, J. M. (2022). The importance of choosing the right keywords for educational technology publications. *Australas. J. Educ. Technol.*, *38*(2), 1–8. https://doi.org/10.14742/ajet.8087.

De Maeyer, P. (2012). Impact of online consumer reviews on sales and price strategies: A review and directions for future research. *J. Prod. Brand Manag.*, *21*(2), 132–139. https://doi.org/10.1108/10610421211215599.

Duong, C. D., Nguyen, T. H., Nguyen, M. H., Dang, N. S., Vu, A. T., & Do, N. D. (2025). Exploring the role of generative artificial intelligence (ChatGPT) adoption in digital social entrepreneurship: A serial mediation model. *Soc. Enterp. J.*, *21*(5), 910–936. https://doi.org/10.1108/sej-03-2024-0029.

Engstrom, T., Strong, J., Sullivan, C., & Pole, J. D. (2022). A comparison of Leximancer semi-automated content analysis to manual content analysis: A healthcare exemplar using emotive transcripts of COVID-19 hospital staff interactive webcasts. *Int. J. Qual. Methods.*, *21*. https://doi.org/10.1177/16094069221118993.

Falatouri, T., Hrušecká, D., & Fischer, T. (2024). Harnessing the power of LLMs for service quality assessment from user-generated content. *IEEE Access*, *12*, 99755–99767. https://doi.org/10.1109/ACCESS.2024.3429290.

Fatouros, G., Soldatos, J., Kouroumali, K., Makridis, G., & Kyriazis, D. (2023). Transforming sentiment analysis in the financial domain with ChatGPT. *Mach. Learn. Appl.*, *14*, 100508. https://doi.org/10.1016/j.mlwa.2023.100508.

Garg, R. K., Urs, V. L., Agrawal, A. A., Chaudhary, S. K., Paliwal, V., & Kar, S. K. (2023). Exploring the role of ChatGPT in patient care (diagnosis and treatment) and medical research: A systematic review. *Health Promot. Perspect.*, *13*(3), 183–191. https://doi.org/10.34172/hpp.2023.22.

Giorgino, R., Alessandri-Bonetti, M., Luca, A., Migliorini, F., Rossi, N., Peretti, G. M., & Mangiavini, L. (2023). ChatGPT in orthopedics: A narrative review exploring the potential of artificial intelligence in orthopedic practice. *Front. Surg.*, *10*(10), 1284015. https://doi.org/10.3389/fsurg.2023.1284015.

Hajra, V. (2023). Unveiling the aesthetics of sacred elephant interactions: A ChatGPT-driven analysis of Tripadvisor reviews. *J. Qual. Res. Tour.*, *4*(2), 175–183. https://doi.org/10.4337/jqrt.2023.02.05.

Harun, A., Ishak, M. I., & Muslim, M. A. (2025). Exploring the influence of online consumer reviews on sales of Umrah travel agencies: The moderating role of package pricing. *Int. Rev. Manag. Mark.*, *16*(1), 1–11. https://doi.org/10.32479/irmm.20473.

He, S., Hollenbeck, B., & Proserpio, D. (2022). The market for fake reviews. *Mark. Sci.*, *41*(5), 896–921. https://doi.org/10.1287/mksc.2022.1353.

Hwang, Y. & Lee, J. H. (2025). Exploring students' experiences and perceptions of human-AI collaboration in digital content making. *Int. J. Educ. Technol. High. Educ.*, *22*(1), 44. https://doi.org/10.1186/s41239-025-00542-0.

Ishak, M. I. & Harun, A. (2023). How review valence benefits umrah sales performance: The moderating role of price. *Technol. Univ. Dublin.*, *11*(2), 3. https://doi.org/10.21427/EXWX-FD39.

Ishak, M. I., Harun, A., Ahmad, N., & Muslim, A. K. (2025). Exploring key industry attributes from online customer reviews: A systematic literature review and network analysis. *PaperASIA*, *41*(4b), 152–180.

https://doi.org/10.59953/paperasia.v41i4b.574.
Jeong, N. & Lee, J. (2024). An aspect-based review analysis using ChatGPT for the exploration of hotel service failures. *Sustainability*, *16*(4), 1640. https://doi.org/10.3390/su16041640.
Jonathan, A., Halim, C. K., Wulandhari, L. A., & Nabiilah, G. Z. (2025). Exploring sentiment patterns in ChatGPT interactions: A machine and deep learning approach to sentiment classification. In *2025 International Conference on Smart Computing, IoT and Machine Learning (SIML)* (pp. 1–6). https://doi.org/10.1109/siml65326.2025.11081155.
Khan, N., Khan, Z., Koubaa, A., Khan, M. K., & Salleh, R. b. (2024). Global insights and the impact of generative AI-ChatGPT on multidisciplinary: A systematic review and bibliometric analysis. *Connect. Sci.*, *36*(1). https://doi.org/10.1080/09540091.2024.2353630.
Kim, J. B. (2014). Impact of online customer reviews and incentives on the product sales at the online retail store: An empirical study at Amazon. com. *AMCIS 2014 Proc.*
Knoedler, S., Sofo, G., Kern, B., Frank, K., Cotofana, S., von Isenburg, S., Könneker, S., Mazzarone, F., Dorafshar, A. H., et al. (2024). Modern Machiavelli? The illusion of ChatGPT-generated patient reviews in plastic and aesthetic surgery based on 9000 review classifications. *J. Plast. Reconstr. Aesthet. Surg.*, *88*, 99–108. https://doi.org/10.1016/j.bjps.2023.10.119.
Koc, E., Hatipoglu, S., Kivrak, O., Celik, C., & Koc, K. (2023). Houston, we have a problem!: The use of ChatGPT in responding to customer complaints. *Technol. Soc.*, *74*, 102333. https://doi.org/10.1016/j.techsoc.2023.102333.
Koçak, D. (2025). Examination of ChatGPT's performance as a data analysis tool. *Educ. Psychol. Meas.*, *85*(4), 641–671. https://doi.org/10.1177/00131644241302721.
Kumamoto, H. & Joho, H. (2025). Exploring the role of large language model in collaborative travel planning task. *Inf. Res. Int. Electron. J.*, *30*(iConf), 1123–1130. https://doi.org/10.47989/ir30iconf46966.
Kumar, Y., Marchena, J., Awlla, A. H., Li, J. J., & Abdalla, H. B. (2024). The AI-powered evolution of big data. *Appl. Sci.*, *14*(22), 10176. https://doi.org/10.3390/app142210176.
Lee, K. D., Han, K., & Myaeng, S. (2016). Capturing word choice patterns with LDA for fake review detection in sentiment analysis. In *Proceedings of the 6th International Conference on Web Intelligence, Mining and Semantics* (pp. 1–7). https://doi.org/10.1145/2912845.2912868.
Li, J., Yang, Y., Chen, R., Zheng, D., Pang, P. C., Lam, C. K., Wong, D., & Wang, Y. (2025). Identifying healthcare needs with patient experience reviews using ChatGPT. *PLoS ONE.*, *20*(3), e0313442. https://doi.org/10.1371/journal.pone.0313442.
Liñán, F. & Fayolle, A. (2015). A systematic literature review on entrepreneurial intentions: Citation, thematic analyses, and research agenda. *Int. Entrep. Manag. J.*, *11*(4), 907–933. https://doi.org/10.1007/s11365-015-0356-5.
Lossio-Ventura, J. A., Weger, R., Lee, A. Y., Guinee, E. P., Chung, J., Atlas, L., Linos, E., & Pereira, F. (2024). A comparison of ChatGPT and fine-tuned open pre-trained transformers (OPT) against widely used sentiment analysis tools: Sentiment analysis of COVID-19 survey data. *JMIR Ment. Health.*, *11*, e50150. https://doi.org/10.2196/50150.
Maguire, M. & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *All Irel. J. High. Educ.*, *3*(3). https://doi.org/10.62707/aishej.v9i3.335.
Mallett, R., Hagen-Zanker, J., Slater, R., & Duvendack, M. (2012). The benefits and challenges of using systematic reviews in international development research. *J. Dev. Eff.*, *4*(3), 445–455. https://doi.org/10.1080/19439342.2012.711342.
Marzi, G., Balzano, M., Caputo, A., & Pellegrini, M. M. (2024). Guidelines for bibliometric-systematic literature reviews: 10 steps to combine analysis, synthesis and theory development. *Int. J. Manag. Rev.*, *27*(1), 81–103. https://doi.org/10.1111/ijmr.12381.
Mathebula, M., Modupe, A., & Marivate, V. (2024). ChatGPT as a text annotation tool to evaluate sentiment analysis on South African financial institutions. *IEEE Access.*, *12*, 144017–144043. https://doi.org/10.1109/access.2024.3464374.
Mayer, R. C. & Davis, J. H. (1999). The effect of the performance appraisal system on trust for management: A field quasi-experiment. *J. Appl. Psychol.*, *84*(1), 123–136. https://doi.org/10.1037/0021-9010.84.1.123.
McAlister, A. R., Alhabash, S., & Yang, J. (2024). Artificial intelligence and ChatGPT: Exploring Current and potential future roles in marketing education. *J. Mark. Commun.*, *30*(2), 166–187. https://doi.org/10.1080/13527266.2023.2289034.
McCloskey, B. J., LaCasse, P. M., & Cox, B. A. (2024). Natural language processing analysis of online reviews for small business: Extracting insight from small corpora. *Ann. Oper. Res.*, *341*(1), 295–312. https://doi.org/10.1007/s10479-023-05816-2.
Miao, J., Thongprayoon, C., Suppadungsuk, S., Garcia Valencia, O. A., Qureshi, F., & Cheungpasitporn, W. (2023). Innovating personalized nephrology care: Exploring the potential utilization of ChatGPT. *J. Pers. Med.*, *13*(12), 1681. https://doi.org/10.3390/jpm13121681.

Mohawesh, R., Xu, S., Tran, S. N., Ollington, R., Springer, M., Jararweh, Y., & Maqsood, S. (2021). Fake reviews detection: A survey. *IEEE Access.*, *9*, 65771–65802. https://doi.org/10.1109/access.2021.3075573.

Mondal, H. & Mondal, S. (2025). The role of large language model chatbots in sexual education: An unmet need of research. *J. Psychosex. Health.*, *7*(2), 120–127. https://doi.org/10.1177/26318318251323714.

Morini-Marrero, S., Ramos-Henriquez, J. M., & Bilgihan, A. (2025). Analyzing the concordance and consistency of AI and human ratings in hospitality reviews. *J. Hosp. Tour. Technol.*, *16*(5), 937–965. https://doi.org/10.1108/jhtt-04-2024-0251.

Morris, R. (1994). Computerized content analysis in management research: A demonstration of advantages & limitations. *J. Manag.*, *20*(4), 903–931. https://doi.org/10.1177/014920639402000405.

Mudrik, A., Nadkarni, G. N., Efros, O., Glicksberg, B. S., Klang, E., & Soffer, S. (2025). Exploring the role of large language models in haematology: A focused review of applications, benefits and limitations. *Br. J. Haematol.*, *205*(5), 1685–1698.

Muslim, A. K. & Harun, A. (2022). Exploring the concept of Muslim-friendly tourism. *Technol. Univ. Dublin.*, *10*(3), 55–89. https://doi.org/10.21427/N4FM-GB33.

Nasrabadi, M. A., Beauregard, Y., & Ekhlassi, A. (2024). The implication of user-generated content in new product development process: A systematic literature review and future research agenda. *Technol. Forecast. Soc. Change.*, *206*, 123551. https://doi.org/10.1016/j.techfore.2024.123551.

Naznin, K., Al Mahmud, A., Nguyen, M. T., & Chua, C. (2025). ChatGPT integration in higher education for personalized learning, academic writing, and coding tasks: A systematic review. *Computers*, *14*(2), 53. https://doi.org/10.3390/computers14020053.

Nikseresht, A., Golmohammadi, D., & Zandieh, M. (2024). Sustainable green logistics and remanufacturing: A bibliometric analysis and future research directions. *Int. J. Logist. Manag.*, *35*(3), 755–803. https://doi.org/10.1108/ijlm-03-2023-0085.

Olorunsola, V. O., Saydam, M. B., Arasli, H., & Sulu, D. (2024). Guest service experience in eco-centric hotels: A content analysis. *Int. Hosp. Rev.*, *38*(1), 81–100. https://doi.org/10.1108/ihr-04-2022-0019.

Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia J. Math. Sci. Technol. Educ.*, *19*(8), em2307. https://doi.org/10.29333/ejmste/13428.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71.

Paul, J. & Criado, A. R. (2020). The art of writing literature review: What do we know and what do we need to know? *Int. Bus. Rev.*, *29*(4), 101717. https://doi.org/10.1016/j.ibusrev.2020.101717.

Pham, D. T. T., Steinmann, S., & Jensen, B. B. (2023). The design of retailers' online review systems—A systematic literature review and future research agenda. *Int. J. Retail Distrib. Manag.*, *51*(9/10), 1255–1287. https://doi.org/10.1108/ijrdm-11-2022-0423.

Preiksaitis, C., Ashenburg, N., Bunney, G., Chu, A., Kabeer, R., Riley, F., Ribeira, R., & Rose, C. (2024). The role of large language models in transforming emergency medicine: Scoping review. *JMIR Med. Inform.*, *12*, e53787. https://doi.org/10.2196/53787.

Ramos-Henriquez, J. M. & Morini-Marrero, S. (2025). Airbnb customer experience in long-term stays: A structural topic model and ChatGPT-driven analysis of the reviews of remote workers. *Int. J. Contemp. Hosp. Manag.*, *37*(1), 161–179. https://doi.org/10.1108/ijchm-01-2024-0034.

Rebora, S., Lehmann, M., Heumann, A., Ding, W., & Lauer, G. (2023). Comparing ChatGPT to human raters and sentiment analysis tools for german children's literature. *CEUR Workshop Proc.*, *1613*, 0073.

Rojas-Sánchez, M. A., Palos-Sánchez, P. R., & Folgado-Fernández, J. A. (2023). Systematic literature review and bibliometric analysis on virtual reality and education. *Educ. Inf. Technol.*, *28*(1), 155–192. https://doi.org/10.1007/s10639-022-11167-5.

Rosete, A., Sosa-Gómez, G., & Rojas, O. (2025). A black-box analysis of the capacity of ChatGPT to generate datasets of human-like comments. *Computers*, *14*(5), 162. https://doi.org/10.3390/computers14050162.

Roy, G., Datta, B., & Basu, R. (2017). Effect of eWOM valence on online retail sales. *Glob. Bus. Rev.*, *18*(1), 198–209. https://doi.org/10.1177/0972150916666966.

Ryan, G. W. & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods.*, *15*(1), 85–109. https://doi.org/10.1177/1525822x02239569.

Salminen, J., Kandpal, C., Kamel, A. M., Jung, S. G., & Jansen, B. J. (2022). Creating and detecting fake reviews of online products. *J. Retail. Consum. Serv.*, *64*, 102771. https://doi.org/10.1016/j.jretconser.2021.102771.

Scharkow, M. (2017). Content analysis, automatic. *Int. Encycl. Commun. Res. Methods.* https://doi.org/10.1002/9781118901731.iecrm0043.

Scherer, R. W. & Saldanha, I. J. (2019). How should systematic reviewers handle conference abstracts? A view from the trenches. *Syst. Rev.*, *8*(1), 264. https://doi.org/10.1186/s13643-019-1188-0.

Seok, C. B. & Chiew, T. C. (2013). Trust, trustworthiness and justice perception toward the head of department.

Glob. J. Arts Humanit. Soc. Sci., *1*(1), 20–29.
Siering, M., Muntermann, J., & Rajagopalan, B. (2018). Explaining and predicting online review helpfulness: The role of content and reviewer-related signals. *Decis. Support Syst.*, *108*, 1–12. https://doi.org/10.1016/j.dss.2018.01.004.
Simetgo, K., Giovanis, A. N., & Kallivokas, D. (2025). The role of ChatGPT and artificial intelligence in customer management strategy transformation: A systematic literature review. *Corp. Bus. Strateg. Rev.*, *6*(1), 254–275. https://doi.org/10.22495/cbsrv6i1siart3.
Skjuve, M., Følstad, A., & Brandtzaeg, P. B. (2023). The user experience of ChatGPT: Findings from a questionnaire study of early users. In *Proceedings of the 5th International Conference on Conversational User Interfaces* (pp. 1–10). https://doi.org/10.1145/3571884.3597144.
Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *J. Bus. Res.*, *104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039.
Spence, M. (1978). Job market signaling. In *Uncertainty in Economics* (pp. 281–306). Elsevier. https://doi.org/10.1016/b978-0-12-214850-7.50025-5.
Su, Z., Yang, M., Zhai, Q., Guo, K., Huang, Y., & Cong, Y. (2025). A multigrained preference analysis method for product iterative design incorporating AI-generated review detection. *Sci. Rep.*, *15*(1), 2528. https://doi.org/10.1038/s41598-025-86551-5.
Tan, K. P., Liu, Y. V., & Litvin, S. W. (2025). ChatGPT and online service recovery: How potential customers react to managerial responses of negative reviews. *Tour. Manag.*, *107*, 105057. https://doi.org/10.1016/j.tourman.2024.105057.
Taylor, S. J., Pinnock, H., Epiphaniou, E., Pearce, G., Parke, H. L., Schwappach, A., Purushotham, N., Jacob, S., Griffiths, C. J., et al. (2014). A rapid synthesis of the evidence on interventions supporting self-management for people with long-term conditions: PRISMS—Practical systematic revIew of self-management support for long-term conditions. *Health Serv. Deliv. Res.*, *2*(53), 1–580. https://doi.org/10.3310/hsdr02530.
Thorpe, R., Holt, R., Macpherson, A., & Pittaway, L. (2005). Using knowledge within small and medium-sized firms: A systematic review of the evidence. *Int. J. Manag. Rev.*, *7*(4), 257–281. https://doi.org/10.1111/j.1468-2370.2005.00116.x.
Trappl, R. (1986). Impacts of artificial intelligence: An overview. *Impacts Artif. Intell.*, 31–51.
Wayne Litvin, S. & Pei-Sze Tan, K. (2024). ChatGPT: It's here, whether we want it or not! *Cornell Hosp. Q.*, *65*(3), 393–402. https://doi.org/10.1177/19389655241229615.
Wu, Y., Ngai, E. W. T., Wu, P., & Wu, C. (2020). Fake online reviews: Literature review, synthesis, and directions for future research. *Decis. Support Syst.*, *132*, 113280. https://doi.org/10.1016/j.dss.2020.113280.
Xiao, F., Zhu, S., & Xin, W. (2025). Exploring the landscape of generative AI (ChatGPT)-powered writing instruction in English as a foreign language education: A scoping review. *ECNU Rev. Educ.*, 20965311241310881. https://doi.org/10.1177/20965311241310881.
Xu, X., Fan, R., Wang, D., Wang, Y., & Wang, Y. (2024). The role of consumer reviews in e-commerce platform credit supervision: A signaling game model based on complex network. *Electron. Commer. Res. Appl.*, *63*, 101347. https://doi.org/10.1016/j.elerap.2023.101347.
Xylogiannopoulos, K. F., Xanthopoulos, P., Karampelas, P., & Bakamitsos, G. A. (2024). ChatGPT paraphrased product reviews can confuse consumers and undermine their trust in genuine reviews. Can you tell the difference? *Inf. Process. Manag.*, *61*(6), 103842. https://doi.org/10.1016/j.ipm.2024.103842.
Xylogiannopoulos, K. F., Xanthopoulos, P., Karampelas, P., & Bakamitsos, G. A. (2025). Is AI-assisted paraphrase the new tool for fake review creation? Challenges and remedies. *Knowledge-Based Systems*, *327*, 114151. https://doi.org/10.1016/j.knosys.2025.114151.
Zamith, R. & Lewis, S. C. (2015). Content analysis and the algorithmic coder: What computational social science means for traditional modes of media analysis. *Ann. Am. Acad. Polit. Soc. Sci.*, *659*(1), 307–318. https://doi.org/10.1177/0002716215570576.
Zhao, Y., Tang, S., Zhang, H., & Lyu, L. (2025). AI vs. human: A large-scale analysis of AI-generated fake reviews, human-generated fake reviews and authentic reviews. *J. Retail. Consum. Serv.*, *87*, 104400. https://doi.org/10.1016/j.jretconser.2025.104400.
Zupic, I. & Čater, T. (2015). Bibliometric methods in management and organization. *Organ. Res. Methods.*, *18*(3), 429–472. https://doi.org/10.1177/1094428114562629.

Appendix

Appendix A. Summary of core information for the 22 studies included in this review

No.	Authors	Aim	Methodology	Key Findings	Data
1	Wayne Litvin & Pei-Sze Tan (2024)	This study aims to assess hotels' use of ChatGPT for responding to TripAdvisor reviews.	A quantitative research design involving a perception experiment using a modified Turing test.	ChatGPT performs effectively mimicking authentic responses written by hotel managers.	Hotel TripAdvisor reviews
2	Cheng et al. (2024)	This study aims to propose a user-friendly framework for mining tourism reviews with high sentiment analysis accuracy.	A qualitative research design involving computational sentiment and topic analysis.	ChatGPT outperforms traditional models and matches BERT in sentiment analysis.	Tourism reviews of China's Five Sacred Mountains
3	Amos & Zhang (2024)	This study aims to study consumer reactions to reviews perceived as ChatGPT-generated versus human-written.	A quantitative research design involving controlled comparative experiments on source perception.	Consumers rate reviews as less useful, trustworthy, and authentic when they perceive them to be generated by ChatGPT.	Hotel reviews on TripAdvisor and restaurant reviews on Yelp
4	Knoedler et al. (2024)	This study aims to compare real and ChatGPT-generated reviews of top US plastic surgeries and evaluate human and AI detection.	A quantitative research design involving a human–AI comparative detection experiment with computational text analysis.	ChatGPT can mimic real patient reviews, with humans identifying them correctly only 59.6%, highlighting the risk of fake reviews.	Plastic surgery patient reviews
5	Baier et al. (2025)	This study aims to develop transfer models predicting technology acceptance and its key factors using online customer reviews.	A quantitative research design involving predictive modeling using secondary OCR data	OCRs, combined with transfer models, AI, and machine learning, can predict technology acceptance over time and may replace traditional surveys.	Google play store reviews
6	Morini-Marrero et al. (2025)	This study aims to explore using ChatGPT to analyze hotel reviews and compare its rating accuracy with humans and classic machine learning methods.	A quantitative research design involving a structured comparative (quasi-)experiment	ChatGPT gives more moderate ratings than humans and can help analyze feedback in hospitality.	TripAdvisor reviews of five-star hotels
7	Li et al. (2025)	This study aims to propose using ChatGPT to analyze patient reviews for improved healthcare insights and resource allocation.	A qualitative research design involving aspect-based sentiment analysis and prompt-engineered ChatGPT	Using ChatGPT with aspect-based sentiment analysis templates achieved high precision, helping providers understand patient needs.	Patient reviews on Haodf.com
8	Su et al. (2025)	This study aims to develop a multi-level method combining AI review detection with product attribute analysis.	A quantitative research design involving supervised machine learning model	The method detects AI-generated fake reviews, analyzes user preferences, and outperforms other approaches.	Amazon reviews of sweeping robots

9	Tan et al. (2025)	This study aims to assess ChatGPT's effectiveness in online service recovery by comparing its managerial responses to human-written ones for hotel reviews.	A mixed-method experimental design	Customers couldn't reliably tell ChatGPT apart from human managerial responses, but knowing a reply was from ChatGPT lowered ratings due to perceived inauthenticity and uncanniness.	Responses to customer reviews in tourism and hospitality
10	Xylogiannopoulos et al. (2025)	To propose and validate a pattern-based method for detecting AI-generated (paraphrased) reviews.	Pattern-based detection evaluated on TripAdvisor reviews and ChatGPT-4.0 paraphrases.	The results show that the proposed method outperforms existing AI text detection approaches in identifying AI-assisted fake reviews.	TripAdvisor reviews
11	Zhao et al. (2025)	This study aims to study linguistic differences between AI, human, and real reviews to improve detection methods.	Quantitative apporach involving large-scale text analysis and statistical testing	AI-generated fake reviews are easier to read but less specific, exaggerated, and mechanical than human or real reviews, highlighting the need for AI-specific detection methods.	Yelp dataset reviews
12	Mathebula et al. (2024)	This study aims to develop a more accurate sentiment analysis model for financial reviews using advanced NLP techniques.	A quantitative research design involving computational machine learning experiment	ChatGPT with BERT and BiLSTM outperformed lexicon methods, improving sentiment analysis for financial decisions.	HelloPeter reviews
13	McCloskey et al. (2024)	This study aims to explore using NLP and LLMs to help small businesses analyze strengths and weaknesses from limited customer reviews.	Quantitative content analysis combined with topic modeling and large language model analysis	Combining NLP methods like topic modeling and ChatGPT yields clear, actionable insights for small businesses.	Altomonte's Italian Market reviews from Google, TripAdvisor, and Yelp
14	Hajra (2023)	This study aims to examine the emotional and spiritual impact of sacred elephant interactions in Asia and identify factors affecting visitor satisfaction.	A qualitative research design involving thematic analysis	Elephant interactions boost visitor satisfaction, and ChatGPT can analyze reviews to enrich tourism aesthetics research.	TripAdvisor reviews of Sacred Elephant
15	Casciato & Mateen (2024)	This study aims to analyze sentiment in foot and ankle surgeon reviews by sex and demographics.	A quantitative research design involving secondary data analysis	Foot and ankle surgeons get mostly positive reviews; males rate higher, and ratings dip with experience, highlighting the need to monitor reputation.	Healthgrades website reviews

16	Xylogiannopoulos et al. (2024)	This study aims to study AI-paraphrased reviews from ChatGPT 4.0 and show how text similarity aids in detecting fake reviews.	Quantitive approach involving comparative computational text similarity experiment	AI-paraphrased reviews differ from human ones but remain similar to each other, making text similarity a useful tool for detecting AI-generated fake reviews.	Reviews of 20 hotels worldwide
17	Botunac et al. (2024)	This study aims to compare fine-tuned Transformers with LLMs like ChatGPT and GPT-4 for classifying and analyzing hotel reviews.	A quantitative research design involving comparative NLP modeling experiment	RoBERTa is faster, while ChatGPT and GPT-4 capture sentiment better but need more resources.	Restaurant reviews from the SemEval-2014
18	Falatouri et al. (2024)	This study aims to assess the efficiency of LLMs, including ChatGPT, for sentiment analysis and service quality (SQ) dimension extraction from online reviews.	The study compares two LLMs (ChatGPT-3.5 and Claude-3) with three traditional NLP techniques using English and Persian customer review datasets.	The findings show that LLMs outperform traditional NLP methods, achieving higher accuracy and stronger agreement with human raters, particularly ChatGPT.	Mobile app reviews from Google Play Store and Cafebazaar.ir
19	Jeong & Lee (2024)	This study aims to use ChatGPT for aspect-based analysis of hotel reviews to identify detailed service failures.	A qualitative research design involving aspect-based content analysis	ChatGPT effectively summarizes hotel complaints, extracts key terms, and outperforms traditional methods.	TripAdvisor hotel reviews dataset
20	Koc et al. (2023)	This study aims to explore using ChatGPT-4 to generate TripAdvisor responses and evaluate their effectiveness in service recovery.	A quantitative research design involving controlled expert-rating experiment	ChatGPT-4's responses are high-quality, fast, and meet service recovery standards, while also assessing service failure severity and offering insights for recovery research.	TripAdvisor customer complaints and management responses
21	Ramos-Henriquez & Morini-Marrero (2025)	This study aims to examine remote workers' Airbnb experiences and compare cognitive outcomes between long- and short-term stays.	A quantitative research design involving structural topic modeling	Remote workers' Airbnb stays are mainly emotional, influenced by stay length and city, with high satisfaction and return intent; hosts should tailor amenities accordingly.	InsideAirbnb reviews for Lisbon and Austin
22	Rosete et al. (2025)	This study aims to test ChatGPT's ability to generate synthetic comments that mimic human vocabulary.	A quantitative research design involving comparative computational vocabulary analysis	ChatGPT's comments are less diverse and partly repetitive but can still support tasks like word clouds.	Reviews from four Kaggle datasets

Appendix B. Thematic coding hierarchy from initial codes to overarching themes

Theme	Theme Description	Sub-theme	Sample Quote	Initial Code	Representative Studies
ChatGPT for Review Analytics	This theme captures studies using ChatGPT as an analytical tool to extract meaning from reviews.	Aspect-based sentiment analysis	*"Using ChatGPT, we automatically identified aspects and sentiments within the reviews, specifically focusing on negative sentiments, with the aid of a pre-trained BERT model." (Jeong & Lee, 2024).*	ChatGPT for aspect-based sentiment analysis in reviews.	Botunac et al. (2024); Falatouri et al. (2024); Jeong & Lee (2024); Li et al. (2025)
		Sentiment analysis	*"ChatGPT was used to perform a sentiment analysis to describe the positivity, negativity, and neutrality of online physician reviews." (Casciato & Mateen, 2024).*	ChatGPT for sentiment analysis in reviews.	Cheng et al. (2024); Mathebula et al. (2024); Casciato & Mateen (2024)
		Content analysis	*"This study aims to explore the application of ChatGPT to analyze hotel guest satisfaction from online reviews." (Morini-Marrero et al., 2025).*	ChatGPT for content analysis for useful insights.	Morini-Marrero et al. (2025); Hajra (2023)
		Topic modeling	*"Do large language models (LLM) such as Chat Generative Pre-Trained Transformer (ChatGPT) provide equal or superior results to topic modeling without requiring the technical skills needed to implement topic modeling?" (McCloskey et al., 2024).*	ChatGPT for topic modeling in reviews.	Ramos-Henriquez & Morini-Marrero (2025); McCloskey et al. (2024)
ChatGPT for Review Modeling and Evaluation	This theme encompasses studies that use ChatGPT-generated reviews for review modeling and evaluation in the OCR context.	Detection model	*"ChatGPT3.5 (also known as ChatGPT in its first version) and ERNIE Bot, developed by Baidu, Inc., are used to generate fake reviews." (Su et al., 2025).*	ChatGPT-generated reviews (fake reviews) for detection model development.	Su et al. (2025); Zhao et al. (2025); Xylogiannopoulos et al. (2024); Xylogiannopoulos et al. (2025); Knoedler et al. (2024)
		Prediction model	*"Based on these findings, we used the chatbot OpenAI ChatGPT 4o (OpenAI, 2024) to generate non-human expert ratings of the six con- structs based on the comments in dataset Ikea1." (Baier et al., 2025).*	ChatGPT-generated reviews used in prediction models for technology acceptance.	Baier et al. (2025)

		Consumer perception	*"This research provides a much-needed investigation into consumer reactions to reviews perceived to be generated by ChatGPT vs. a human." (Amos & Zhang, 2024).*	ChatGPT-generated reviews for examining consumer reactions and perceptions.	Amos & Zhang (2024)
		Synthetic review dataset	*"This paper examines the ability of ChatGPT to generate synthetic comment datasets that mimic those produced by humans." (Rosete et al., 2025).*	ChatGPT-generated reviews to test its ability to generate synthetic reviews dataset.	Rosete et al. (2025)
ChatGPT for reviews management	This theme reflects ChatGPT's role in managerial interaction and response to online reviews.	Management response	*"This managerial perspective considers the use of ChatGPT by hotels as a tool for replying to their property's online consumer-generated media postings." (Wayne Litvin & Pei-Sze Tan, 2024).*	ChatGPT for replying to customer reviews by management.	Wayne Litvin & Pei-Sze Tan (2024); Tan et al. (2025); Koc et al. (2023)