The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1756-378X.htm

Research on electronic word-of-mouth for product and service quality improvement: bibliometric analysis and future directions

Received 3 March 2022
Revised 24 March 2022
Accepted 25 March 2022

Yajun Wang, Xinyu Meng and Chang Xu
School of Business Administration, Northeastern University, Shenyang, China and Northeastern University at Qinhuangdao, Qinhuangdao, China, and
Meng Zhao
School of Business Administration, Northeastern University, Shenyang, China; Northeastern University at Qinhuangdao, Qinhuangdao, China and Sichuan University, Chengdu, China

Abstract

Purpose – This paper aims to analyze high-quality papers on the research of electronic word-of-mouth (eWOM) for product and service quality improvement from 2009 to 2022, in order to fully understand their historical progress, current situation and future development trend.

Design/Methodology/Approach – This paper adopts the bibliometrics method to analyze the relevant literature, including publishing trend and citation status, regional and discipline area distribution, and influential publications. Secondly, the VOSviewer is used for literature co-citation analysis and keyword co-occurrence analysis to obtain the basic literature and research hotspots in this research field.

Findings – Firstly, the study finds that the number of publications basically shows an increasing trend, and those publications are mainly published in tourism journals. In addition, among these papers, China has the largest number of publications, followed by the USA and South Korea. Through co-citation analysis of literature and keyword co-occurrence analysis, 22 foundational papers and six main research topics are obtained in this paper. Finally, this paper elaborates on the development trend of the research topic and future research directions in detail.

Originality/value – This is the first paper that uses bibliometrics to analyze and review relevant researches on eWOM for product and service quality improvement, which is helpful for researchers to quickly understand its development status and trend. This review also provides some future research directions and provides a reference for further research.

Keywords Electronic word-of-mouth, Product improvement, Service quality improvement, Bibliometric analysis, VOSviewer

Paper type Literature review

1. Introduction

With the popularization and development of the Internet, more and more customers express their opinions on the Internet. When customers use products or enjoy certain services, they are more willing to share their personal experiences on social media in the form of text reviews. With the increase of user-generated content (UGC) such as online reviews, the electronic form of word-of-mouth is flourishing. Electronic word-of-mouth (eWOM) refers to any statement that customers share via the Internet (e.g. websites, blogs, micro blogs and instant messages) about a product, service, brand, or company (Hennig-Thurau *et al.*, 2004).

The work presented in this paper is supported by the National Natural Science Foundation of China (No. 71701037), National Natural Science Foundation of Hebei province (No. G2021501004), Fundamental Research Funds for the Central Universities (No. N2123020), Youth Top-notch Talent Support Program of Hebei province (No. BJ2020211) and Postdoctoral Science Foundation of China (2019M663542).

International Journal of Intelligent Computing and Cybernetics
© Emerald Publishing Limited
1756-378X
DOI 10.1108/IJICC-03-2022-0065

eWOM spreads faster and wider than traditional word-of-mouth, showing customer perceptions in real time (Xu, 2018; Xu and Li, 2016). eWOM promotes the responsiveness of customers and managers and has a significant impact on enterprise performance and customer purchase intention (Zheng, 2021; Bilgihan *et al.*, 2018). For potential customers, they are more likely to read online reviews from customers who have purchased than sellers. Studies show that about 70% of customers rely on product reviews to make decisions, as they are the second most trusted information source after family and friends (Abbas and Malik, 2021). UGC (such as online reviews) is an important channel for managers to understand customer feedback, while negative reviews from customers will generate negative eWOM for the enterprise and affect its image. Therefore, managers must take measures to restore the corporate image. The more common measures are to improve product performance and service quality, which means managers must pay close attention to customer requirements. They provide products or services more in line with customer expectations according to customer preferences, so as to improve customer satisfaction and enable enterprises to better maintain competitiveness (Ayağ, 2021).

Traditional product and service quality improvement methods include quality function deployment (QFD) (Ping *et al.*, 2020), importance-performance analysis (IPA) model (Azzopardi and Nash, 2013) and KANO model (Kano *et al.*, 1984). The research data come from questionnaire surveys, interviews, marketing interviews, etc., which have defects such as high cost and difficult quality assurance. With the continuous segmentation and expansion of the consumer market, customers have more and more new requirements, and it is increasingly difficult to improve the quality of products and services. Kwok *et al.* (2017) analyzed 67 papers published between 2000 and 2015 that focused on online reviews. The results showed that online reviews can effectively understand the business performance. Schuckert *et al.* (2016) also proved that online reviews, especially negative reviews, are the best channel to assess information about service delivery, quality and customer requirements. Therefore, compared with traditional research methods, analyzing eWOM for product improvement and precision marketing can not only effectively save time and cost, but also ensure the authenticity and reliability of data, which is more helpful for enterprises to understand customer requirements (Xu and Li, 2016; Culotta and Cutler, 2016).

Because the improvement of product and service quality is of great significance to customers, enterprises and society, the research of eWOM oriented to the improvement of product and service quality is favored by many domestic and foreign researchers (Chen *et al.*, 2019; Jiang *et al.*, 2017; Jin *et al.*, 2016b; Qi *et al.*, 2016; Deng *et al.*, 2013). The research focuses on customer preference analysis, customer satisfaction analysis, requirement classification and so on (Bi *et al.*, 2020; Martí Bigorra *et al.*, 2019; Raditham *et al.*, 2022). For example, Oh and Yi (2021) provided a new method based on bigram natural language processing (NLP) analysis to evaluate the product feature level and its impact on customer satisfaction in view of the asymmetric impact of customer sentiment on the rating of wireless headset products. For the improvement of sports utility vehicles, Wang *et al.* (2020a) first considered the inconsistencies between the numerical product ratings and the textual product reviews to establish the inconsistent ordered choice model (IOCM) for measuring customer preferences concerning product features. In order to promote the research development of eWOM for product and service quality improvement, it is necessary to conduct more in-depth and comprehensive research in this field, and provide more enlightenment for the further research of researchers. However, it is found that no researchers have summarized relevant researches in this field to reveal current research hotspots and development trends. In view of this, this paper makes a bibliometric analysis of eWOM research for product and service quality improvement, in order to find the research hotspots and future development directions. Several research questions are proposed and solved during the research:

RQ1. What is the current research status in this area?

RQ2. What are the research directions in which scholars are interested and what are the relevant research bases?

RQ3. How to reveal the future research direction according to the research hotspots and development trend?

The rest of this paper is structured as follows: Section 2 provides the publication collection and processing process. Section 3 makes bibliometric analysis of publication and presents the research results. Section 4 reviews the research of eWOM in the direction of product and service quality improvement. The findings and future research directions are summarized in section 5. Section 6 presents the conclusions and limitations of the paper.

2. Method

2.1 Data collection

To access all publications on eWOM research for product and service quality improvement, we search papers from the Web of Science database, Elsevier database and SpringerLink database by title, keyword and abstract. They are recognized databases that include a variety of reference journals from reputable publishers and provide reliable and comprehensive journal data. The keywords retrieved are "Product Improvements/Online reviews/User-generated content/eWOM", "Service Quality Improvement/Online Reviews/User-generated content/eWOM", "Satisfaction/Online Reviews/User-generated content/eWOM". A total of 5,858 papers are retrieved in this way. In order to narrow the scope of papers, we only select publications between 2009 and 2022, and initially screened out 5,658 papers.

Due to the different level of academic journals, the quality of papers is also uneven. In order to obtain papers that meet the requirements for review, the following literature screening criteria are determined in this paper:

(1) Papers published in journals with ABS 2-stars or above.

(2) If (1) is not met, journals shall be screened according to the criteria of JCR partition in Q1 and Q2.

(3) The research content must be related to product/service quality improvement and eWOM.

(4) Eliminate papers published at conferences, book chapter/trade journals/book contributions.

(5) Papers involving CSA on languages other than English, for example, languages Hindi, Portuguese, Chinese, Bengali, Spanish, etc. are not considered.

According to the manual screening process, 110 papers are obtained for subsequent analysis. The literature screening process is shown in Figure 1.

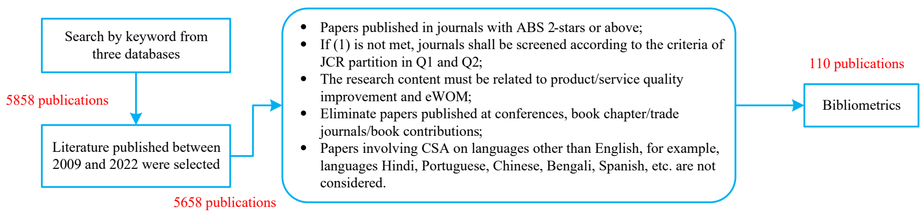

Figure 1. Literature screening process

2.2 Data analysis

Bibliometric analysis is a research method widely used by librarians and researchers. It uses quantitative analysis and statistical data to discover patterns in publications in a particular field. It has been widely used by researchers in literature analysis of the publications in journals and publications of special topics (Liao *et al.*, 2020; Wang *et al.*, 2020b; Yu *et al.*, 2020; Lin *et al.*, 2021a; Muritala *et al.*, 2020; Luo and Lin, 2021). Therefore, this paper uses manual review and VOSviewer bibliometric analysis tools to analyze the selected publications.

Firstly, this paper analyzes the publishing trend and citation of all publications. Secondly, in order to obtain the regional distribution of researchers and their main contribution fields, this paper makes statistics on the distribution of the regions and the discipline categories of the publications, and analyzes the 15 papers with the highest citations. Further, VOSviewer software is used to conduct the co-citation analysis, and the basic literature (the most commonly cited literature among all literature) in the research field is determined. Finally, the keywords and key terms co-occurrence analysis is carried out to determine the main research topic and the development trend. The research framework is shown in Figure 2. We summarize the specific analysis results in Section 3.

3. Results

3.1 Publishing trend and citation status

Firstly, the bibliometrics method is used to count the number of publications and citations of relevant studies over the years and show the development trend of this study. This paper evaluates the publication through indicators such as the total number of publications (TP), number of cited papers (CP), total number of citations (TC), citations per publication (C/P) and citations per cited paper (C/CP), and then quantifies 110 papers based on these indicators. To make the results more visual, TP and TC are drawn in Figure 3.

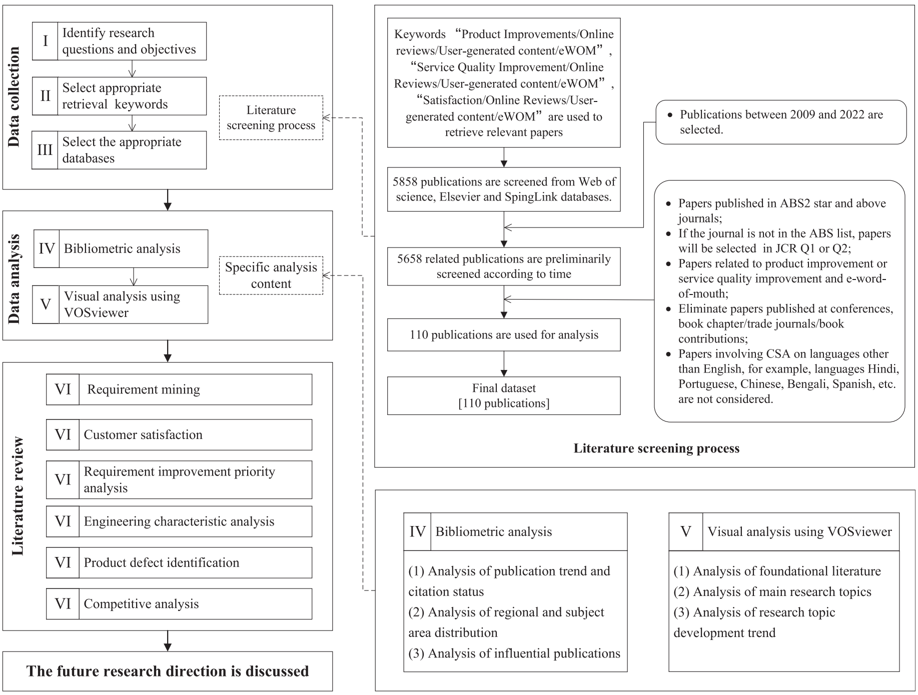

Figure 2.
The research framework

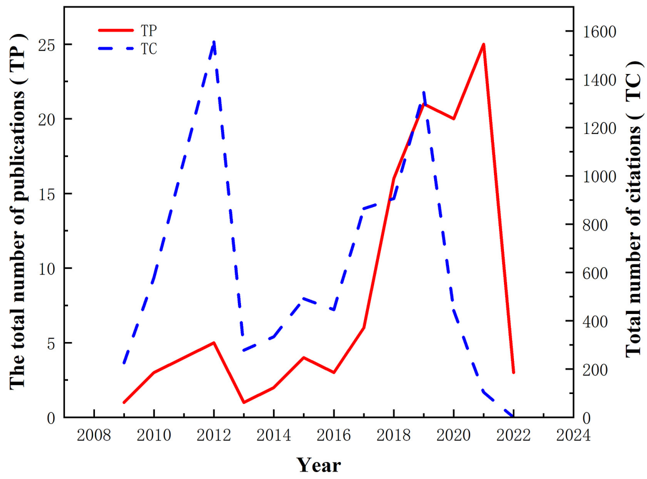

Figure 3.
TP and TC from 2009 to 2022

From Table 1 and Figure 3, some researchers began to pay attention to the relevant research direction of this field in 2009. As time goes by, the publishing trend basically maintained an increasing trend, and TC has reached its peak in 2012, which means more and more researchers begin to pay attention to this field. In addition, almost every paper has been cited until 2018. The C/CP in 2009, 2012 and 2013 are 226, 311 and 278, respectively, far exceeding the value after 2018. The reason for this phenomenon is that the number of publications each year did not reach 10 before 2018. Therefore, researchers expand their research based on limited research. Due to the accumulation of time or the quality of articles after 2019, the C/CP in the last three years is low.

In addition, we analyze journal publication trends, listing the most prolific 14 journals in Table 2. As the most prolific journals in this field, Tourism Management and International Journal of Hospitality Management have published 12 papers, with a total number of citations of 1425 and 577 respectively. Secondly, many papers were published in Decision Support

Year	TP	CP	TC	C/P	C/CP
2009	1	1	226	226	226
2010	3	3	580	193.3333	193.3333
2012	5	5	1,555	311	311
2013	1	1	278	278	278
2014	2	2	334	167	167
2015	4	4	491	122.75	122.75
2016	3	3	446	148.6667	148.6667
2017	6	6	864	144	144
2018	16	16	906	56.625	56.625
2019	21	20	1,345	64.04762	67.25
2020	20	20	442	22.1	22.1
2021	25	16	104	4.16	6.5
2022	3	0	0	0	0
Total	110	97	7,571	1737.683	1743.225

Note(s): The total number of publications (TP), number of cited papers (CP), total number of citations (TC), citations per publication (C/P) and citations per cited paper (C/CP)

Table 1.
Publication trend and citation from 2009 to 2022

Table 2.
List of most prolific journals

Rank	Journals	Total papers	Total citations	Quartile
1	*Tourism Management*	12	1,425	Q1
2	*International Journal of Hospitality Management*	12	577	Q1
3	*Decision Support Systems*	7	375	Q1
4	*International Journal of Contemporary Hospitality Management*	6	466	Q1/Q2
5	*International Journal of Production Research*	6	191	Q1
6	*Expert Systems with Applications*	5	511	Q1
7	*Advanced Engineering Informatics*	4	84	Q1
8	*Engineering Applications of Artificial Intelligence*	4	292	Q2
9	*Electronic Commerce Research and Applications*	4	106	Q1/Q2
10	*Information and Management*	3	503	Q1
11	*International Journal of Information Management*	3	303	Q1
12	*Journal of Ambient Intelligence and Humanized Computing*	3	9	Q1
13	*Technological Forecasting and Social Change*	3	12	Q1
14	*Journal of Hospitality and Tourism Research*	3	303	Q2

Systems and International Journal of Contemporary Hospitality Management, with 7 and 6 articles respectively. Therefore, tourism journals published the most related research, followed by engineering and computer science journals.

3.2 Regional and discipline area distribution

We list 10 countries/regions that have made great contributions to this field and their number of citations in Table 3, and the proportion of journals published in each country/region is plotted in Figure 4. As we can see, 52 papers were published in China in the past 14 years, accounting for 47.27%, far exceeding other countries. The USA and South Korea published 19 and 8 papers, respectively, accounting for 17.27% and 7.27%, and the number of other countries published was three or less. The result shows that this field has received higher attention in China. However, the total number of citations in China was 2,669, while the USA was 2,838. This indicates that Chinese researchers receive less attention in this research field than American, which can be increased by improving research quality and innovation. In addition, as developing countries, Malaysia, Brazil and Turkey have also made significant contributions to develop this field.

To find out which disciplines have contributed more to the field, we analyze the distribution of the discipline categories of the papers. From Figure 5, There are 38 papers in

Table 3.
Top regions having more than 2 publications

Rank	Country/region	Publications	Percentage	Citations
1	China	52	47.27%	2,669
2	USA	19	17.27%	2,838
3	South Korea	8	7.27%	277
4	Australia	3	2.73%	496
5	Malaysia	3	2.73%	122
6	Singapore	3	2.73%	289
7	UK	3	2.73%	100
8	Brazil	2	1.82%	50
9	Turkey	2	1.82%	11
10	Italy	2	1.82%	80

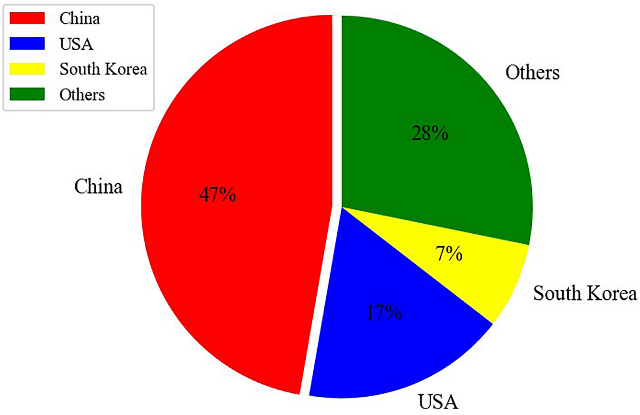

Figure 4.
Top regions having more than 2 publications

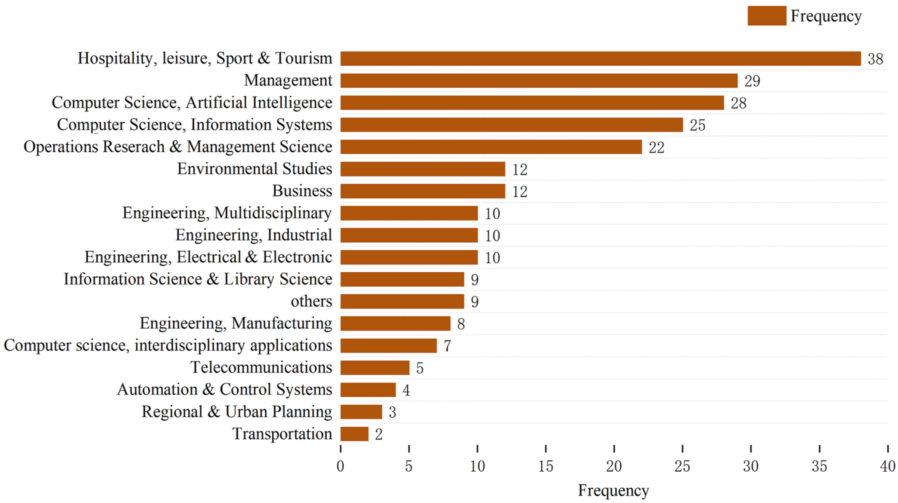

Figure 5.
Categories of reviewed papers

Hospitality, leisure, Sport and Tourism. The reason is that the field has a strong application and large demand for service improvements. The following categories are Management, Computer Science Artificial Intelligence, Computer Science Information Systems, Operations Research and Management Science.

3.3 Influential publications

In this section, we select the 15 most highly cited papers to analyze the influence of 110 papers in this field, and the results are shown in Table 4.

As shown in Table 4, the paper titled "Mine Your Own Business: Market-structure Surveillance Through Text Mining" is the most highly cited paper, which was written by Netzer *et al.* (2012). The number of citations reached 737 times, indicating that the research content of this paper is the research hotspot. Guo *et al.* (2017) followed with 544 citations. According to the analysis, 7 of the top 15 literature are related to customer satisfaction

Rank	Author	Journals	Citations
1	Netzer *et al.* (2012)	*Marketing Science*	737
2	Guo *et al.* (2017)	*Tourism Management*	544
3	Decker and Trusov (2010)	*International Journal of Research in Marketing*	391
4	Lu and Stepchenkova (2012)	*Tourism Management*	326
5	Browning *et al.* (2013)	*Journal of Travel and Tourism Marketing*	278
6	Pizam *et al.* (2016)	*International Journal of Contemporary Hospitality Management*	275
7	Zhao *et al.* (2019)	*International Journal of Hospitality Management*	248
8	López *et al.* (2012)	*Journal of General Internal*	242
9	Ye *et al.* (2014)	*Journal of Hospitality and Tourism Research*	235
10	He *et al.* (2015)	*Information and Management*	230
11	Zhan *et al.* (2009)	*Expert Systems with Applications*	226
12	Zhang *et al.* (2012)	*Expert Systems with Applications*	217
13	Timoshenko and Hauser (2019)	*Marketing Science*	200
14	Qi *et al.* (2016)	*Information and Management*	191
15	Li *et al.* (2015)	*Tourism Management*	170

Table 4.
Top 15 most highly cited papers of reviewed publications

analysis, indicating that customer satisfaction based on social media data is a research hotspot, followed by papers related to customer requirement mining. From the perspective of issuing journals, most of the journals are related to tourism management, such as Tourism Management, Journal of Travel and Tourism Marketing, International Journal of Contemporary Hospitality, Journal of Hospitality and Tourism Research, which means researchers are most interested in the hotel and tourism industry. Finally, by observing the published years, we can find that eWOM research for product and service improvement is emerging in recent years. Based on the background of human history into the era of big data, it can be understood that researchers are interested in this field recently.

3.4 Foundational literature

Through the research of foundational literature, we can find influential and pioneering papers in this field. In this section, foundational papers are obtained by co-citation analysis of selected high-quality papers. It is worth noting that this section is fundamentally different from the influential publications in Section 3.3. Specifically, Section 3.3 obtains the top 15 most highly cited papers among 110 papers, which reflects the research interests of researchers. This section analyzes the co-citation of 110 papers and reveals the main basis on which researchers expand and improve. Therefore, the papers analyzed in this section are the foundational papers in this field.

A total of 5,265 references are cited by drawing a co-citation network graph through VOSviewer. In order to obtain the suitable number of foundational papers, the minimum threshold of the number of cited references is set to 11, and finally 22 co-cited references are obtained as foundational literature, as shown in Figure 6.

It can be seen from Figure 6 that the nodes of the nine papers in the green cluster are relatively larger, indicating that these papers are cited most frequently. It can be divided into three sub clusters according to the year: Sub cluster 1 includes two early and influential conceptual papers (Oliver, 1980; Parasuraman *et al.*, 1988); Sub cluster 2 includes three classic papers published from 2009 to 2010 which, respectively, studied review usefulness and the impact of eWOM on customer behavior and product sales (Mudambi and Schuff, 2010; Ye *et al.*, 2009; Zhang *et al.*, 2010); Sub cluster 3 includes four papers related to customer requirement mining published in 2015–2016 (Berezina *et al.*, 2016; Schuckert *et al.*, 2015;

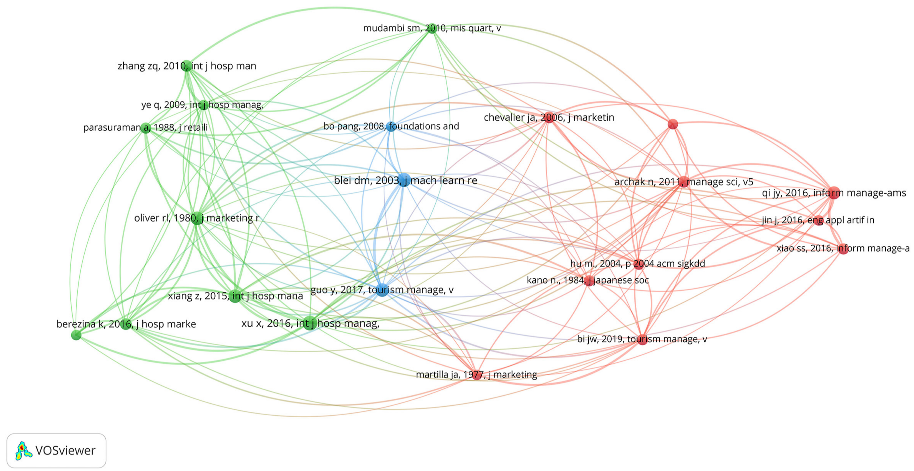

Figure 6.
Co-citation network map

Xiang *et al.*, 2015; Xu and Li, 2016). It is worth noting that the research fields of the latter two sub clusters are mainly in the tourism and hotel industry.

The three papers in blue cluster include two papers on opinion mining and sentiment analysis (Blei *et al.*, 2003; Pang and Lee, 2008) and a paper using text mining technology to analyze online reviews (Guo *et al.*, 2017). This cluster has a strong correlation with the papers of feature extraction and sentiment analysis based on online reviews in the other two clusters.

The 10 papers in the red cluster are similar to other clusters because their published years span the other two clusters. In terms of time, it can be divided into three sub clusters: Sub cluster 1 includes two classic pioneering papers related to requirement classification (Martilla and James, 1977; Kano *et al.*, 1984); Sub cluster 2 includes four papers published from 2004 to 2011, which mainly studied the impact of eWOM on product sales and requirement mining methods. The difference between them and papers in green cluster is that the research field is no longer limited to the tourism and hotel industry, but in the field of books and electronic products (Chevalier and Mayzlin, 2006; Decker and Trusov, 2010; Archak *et al.*, 2011; Hu and Liu, 2004); The four papers of sub cluster 3 focus on the research of competition analysis (Jin *et al.*, 2016a) and requirement classification, respectively (Qi *et al.*, 2016; Xiao *et al.*, 2016; Bi *et al.*, 2019b).

3.5 Main research topics

In this section, we conduct topic co-occurrence analysis on 110 papers to understand the research hotspots in this field. In order to make the co-word network map more convincing, this paper uses two different quantitative methods to obtain the main research topics. One is to use the keywords in the literature to draw the keywords co-occurrence network map, the other is to extract the terms from the title and abstract of the literature to draw the key terms co-occurrence network graph. The minimum thresholds for the occurrence times of keywords and terms are set to 7 and 12, respectively. A total of 31 keywords and terms are used to draw co-occurrence graphs. The results are shown in Figures 7 and 8.

In this paper, similar word lists in the two quantitative methods are constructed manually to form thesaurus file and uploaded to VOSviewer. The purpose is to replace low-frequency words with high-frequency words.

Figures 7 and 8 show the research topics in this field. As can be seen, the nodes of “customer satisfaction” and “online reviews” are relatively large and surrounded by dense

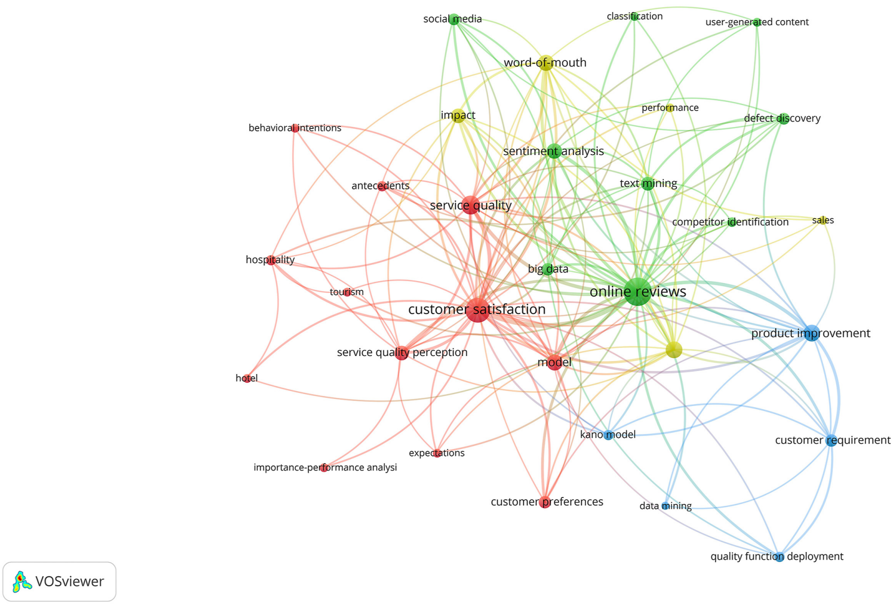

Figure 7.
Keywords co-occurrence network map

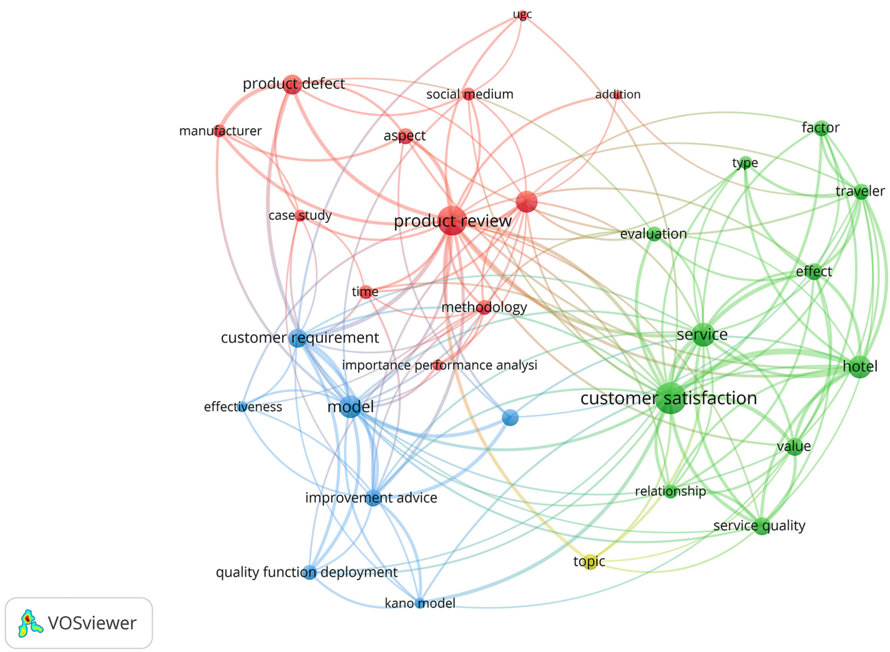

Figure 8.
Key terms co-occurrence network map

lines, indicating that these two topics are widely discussed in the literature. In addition, “product improvement”, “service quality”, “competitiveness”, “customer requirement”, "competitor identification", "customer preferences" and "defect discovery" are regarded as important topics and research hotspots in this field. According to Figures 6 and 7, "Kano Model", "customer preferences", "quality function deployment" and "importance performance analysis", "defect discovery" appear at the edges of the two figures in different forms. Further research directions in this field can be understood through these topics.

Through the analysis and classification of keywords co-occurrence network map and key terms co-occurrence network map, we summary six research topics in this field based on the similarity between keywords, namely requirement mining, customer satisfaction, requirement improvement priority analysis, engineering characteristic analysis, product defect identification and competitive analysis.

3.6 Research topic development trend

Considering that the research topics obtained from keywords co-occurrence network map has fewer noise words, we use keywords to generate overlay visualization (Figure 9) to analyze the development trend of the research topic. The node size reflects times the topic appears in the keyword list, and the dark of the node color represents the time series of the average publication year of the literature where the keyword is located. Observing the color bar in the lower right corner of Figure 9, the closer the node is to blue (yellow), the smaller (larger) the publication year of the topic is. We combine the changes of node size and node color in Figure 9 to analyze the development trend of research topics in this field.

Firstly, it can be seen that the average publication year of the literature on the 31 topics is concentrated between 2017 and 2020. Figure 9 shows that "customer satisfaction" and "online reviews" are relatively popular topics. According to the appearance time of "online reviews" (2018.68) and "word of mouth" (2017.75), "online reviews" has been mentioned more frequently in the literature in recent years. Secondly, we find the development trend of text analysis technology. Among them, "text mining" (2018.05) technology is a research hotspot in processing social media data. In recent years, "big data" (2018.94) and "sentient analysis" (2018.96) technologies associated with this topic have become more concerned research subjects in this field. Thirdly, "importance of product features" (2019.71), "classification"

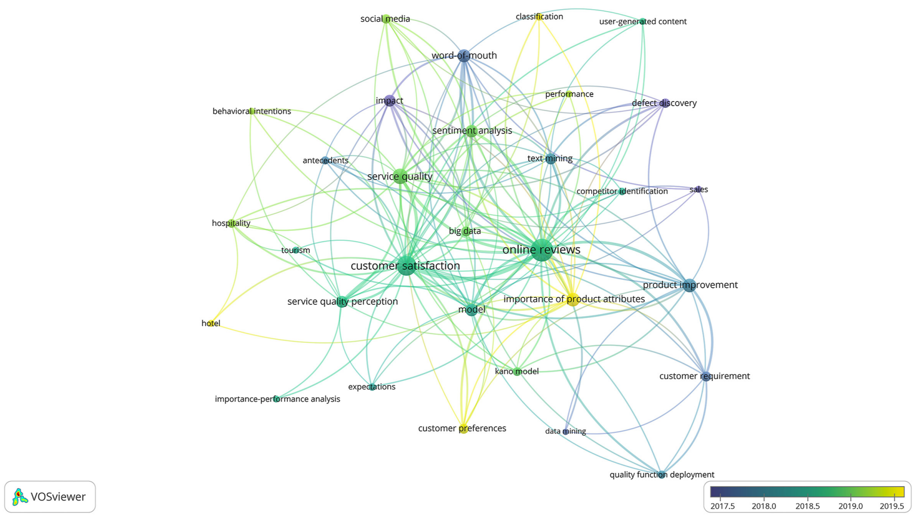

Figure 9.
Research topic development trend

(2019.63) and "customer preferences" (2019.4) are words that researchers have mentioned more in recent years. There is a strong correlation between "importance of product features" and the other two topics. Therefore, researchers prefer to consider the importance index of product features in classification and customer preference analysis. Fourthly, from the correlation between "behavior intention" and "service quality", some researchers have considered the behavior of customers in the research on service quality. Finally, from the perspective of the application of related research, the use of hotel reviews to improve the service quality of the hotel industry has been favored by many researchers in recent years.

4. Research on eWOM for product and service quality improvement

Based on the six research hotspots obtained in Section 3.5, this paper divides the literature and makes a systematic review to reveal the research status and future research direction. Related papers on six topics are shown in Table 5.

4.1 Requirement mining

Only by capturing customer requirements ahead of competitors and improving product or service quality, can an enterprise be in an advantageous position in market competition. Therefore, many researchers focus on customer requirement mining. We summarize the text analysis methods and tools mentioned in Table 6.

Some researchers believe that there is a certain relationship between customer requirements and product features. Features that are often mentioned by customers and have an important impact on customer perception can be regarded as customer requirements. In view of this, some researchers have adopted different text mining methods to identify customer requirements from the perspective of customer preferences (Jiao and Qu, 2019; Yu and Wang, 2010; Zhang *et al.*, 2019; Yadav and Roychoudhury, 2019; Rodrigues *et al.*, 2020; Ibrahim and Wang, 2019; Sun *et al.*, 2020; Jin *et al.*, 2016b). Li *et al.* (2015) applied the emerging pattern mining (EPM) concept to discover changes and trends in travelers' attention. They focus on some features of specific interest, which can help managers develop effective and targeted improvement plans. Wang *et al.* (2019) proposed a novel data-driven graph-based requirement elicitation framework in the Smart PSS, so as to assist engineers/designers to make better design improvement or new design concept generation in a closed-loop manner. Lee and Tse (2020) proposed a feature recognition method combining topic modeling using Latent Dirichlet Allocation (LDA), sentiment analysis and process analysis based on process chain network (PCN). They aim to identify key service features from online reviews and develop service improvement strategies based on unsatisfactory services. In addition, there are also researchers who predict the future importance of products to support for the formulation of product improvement strategies from a new perspective (Yakubu and Kwong, 2021).

In addition, some researchers believe that there is a certain difference between customer preferences and customer requirements, and it is not comprehensive to express customer requirements only by using product or service characteristics. Therefore, some researchers directly analyze the text and obtain more comprehensive customer requirements information by forming text summaries (Zhan *et al.*, 2009). Wang *et al.* (2018a, b) proposed a Kansei text mining approach, which summarizes the reviews by counting the frequency of the features, Kansei words, and Kansei features, and their associations, then a list of summaries can be generated. Timoshenko and Hauser (2019) constructed a convolutional neural network model to identify the usefulness information for the requirement mining of oral care products, and further avoided the repetition of sampling content through clustering. This method can directly identify customer requirements from

Topics	Sub-topics	Citations
Requirement mining	Customer preference mining	Ibrahim and Wang (2019), Rodrigues *et al.* (2020), Sun *et al.* (2020), Wang *et al.* (2019), Yadav and Roychoudhury (2019), Yakubu and Kwong (2021), Jin *et al.* (2016b), Li *et al.* (2015), Lee and Tse (2020), Jiao and Qu (2019), Yu and Wang (2010), Zhang *et al.* (2019)
	Summary document generation	Chen *et al.* (2019), Timoshenko and Hauser (2019), Wang *et al.* (2018b), Zhan *et al.* (2009)
Customer satisfaction	Customer satisfaction	Barnes *et al.* (2020), Bi *et al.* (2020), Chen *et al.* (2021a, b), Francesco and Roberta (2019), Kim and Canina (2015), Kwon *et al.* (2022), Lim and Lee (2020), Lv *et al.* (2021), Nilashi *et al.* (2021), Padma and Ahn (2020), Park *et al.* (2021), Park *et al.* (2020), Pokryshevskaya and Antipov (2017), Rajaguru and Hassanli (2018), Stamolampros *et al.* (2020), Tontini *et al.* (2017), Xu (2019, 2020a), Zhang *et al.* (2021a–c), Browning *et al.* (2013), Guo *et al.* (2017), Nakayama and Wan (2019), Pizam *et al.* (2016), Radojevic *et al.* (2018), Xu (2018), Xu *et al.* (2017), Ye *et al.* (2014), Kwon *et al.* (2020), Jiang *et al.* (2010), Lai *et al.* (2018), López *et al.* (2012), Lu and Stepchenkova (2012), Mathayomchan and Taecharungroj (2020), You *et al.* (2012), Zhao *et al.* (2019), Bilgihan *et al.* (2018), Sezgen *et al.* (2019)
Requirement improvement priority analysis	Requirement improvement priority analysis based on requirement classification	Chen *et al.* (2021a), Choi *et al.* (2020), Hu *et al.* (2020, 2021), Jin *et al.* (2021), Joung and Kim (2021), Li *et al.* (2021), Luo *et al.* (2021), Martí Bigorra *et al.* (2019), Ou *et al.* (2018), Ranjbari *et al.* (2020), Shi and Peng (2021), Wang *et al.* (2020a), Zhang *et al.* (2021c), Bi *et al.* (2019a, b), Jeong *et al.* (2019), Park *et al.* (2021), Qi *et al.* (2016)
	Requirement improvement priority analysis based on customer preference analysis	Cai *et al.* (2021), Hu *et al.* (2019), Oh and Yi (2021), Yadegaridehkordi *et al.* (2021), Zhang *et al.* (2021b), Nilashi *et al.* (2021), Du *et al.* (2022), Zhou *et al.* (2022), Decker and Trusov (2010), Xu (2020b)
Engineering characteristic analysis	QFD	Chin *et al.* (2019), Ha and Geum (2022), Jin *et al.* (2015), Özdağoğlu *et al.* (2018), Trappey *et al.* (2018)
	QFD and Kano	Ji *et al.* (2014), He *et al.* (2017)
Product defect identification	Product defect identification	Liu *et al.* (2018), Mummalaneni *et al.* (2018), Zheng *et al.* (2020, 2021), Abbas and Malik (2021), Law *et al.* (2017), Zhang *et al.* (2012)
Competitive analysis	Analysis of competitive advantages and disadvantages	Albayrak *et al.* (2021), Gang and Chenglin (2021), Jin *et al.* (2019), Liang *et al.* (2020), Liu *et al.* (2021), Rodríguez-Díaz and Espino-Rodríguez (2018), Abbas and Malik (2021), Marcolin *et al.* (2021), Kim and Kang (2018), Wang *et al.* (2018a), Netzer *et al.* (2012)
	Competitor identification	Gao *et al.* (2018), Hu and Trivedi (2020), Liu *et al.* (2019), Park and Lee (2021)

Table 5. Literature information corresponding to the research topic

References	POS	LDA	Unsupervised text mining method	TF-IDF	Rule-based feature extraction	Sentiment analysis based on dictionary	Binary Naive Bayes classifier	Text analysis tool
Zhan *et al.* (2009)					✓			
Yu and Wang (2010)	✓				✓			
Li *et al.* (2015)	✓						✓	✓
Jin *et al.* (2016)	✓					✓	✓	
Wang *et al.* (2018b)	✓		✓		✓	✓		✓
Zhang *et al.* (2019)			✓			✓		
Jiao and Qu (2019)	✓				✓	✓		
Yadav and Roychoudhury (2019)	✓					✓		
Chen *et al.* (2019)					✓			
Timoshenko and Hauser (2019)			✓					
Ibrahim and Wang (2019)		✓						✓
Sun *et al.* (2020)				✓				
Yakubu and Kwong (2021)	✓			✓				
Wang *et al.* (2019)					✓			✓
Rodrigues *et al.* (2020)								✓
Lee and Tse (2020)		✓						

Table 6.
Text analysis methods and tools

the text summaries of clustering. Chen *et al.* (2019) developed a context-aware word segmentation method for requirement mining of mobile applications to extract opinion targets from online reviews and obtain useful opinion-level text information.

The above studies mainly focus on topic extraction and sentiment analysis, mainly using POS Tagging, LDA, convolutional neural network, word frequency statistics and other methods. Considering that customer requirements are constantly changing, it is an important research direction in the future to analyze customer preferences from a dynamic point of view for requirement analysis and thus to forecast customer requirements.

4.2 Customer satisfaction

We find that relevant researches on customer satisfaction mainly focus on the relationship between service perception and satisfaction in the hotel industry, and others focus on the satisfaction of electronic products, restaurants, airlines and tourism. In addition, some researchers have extended their satisfaction research methods to online health communities and shared accommodation.

For the study of hotel industry, to make the results more accurate, some researchers usually use some text mining methods to obtain key features related to customer satisfaction to replace the existing service features of websites before establishing satisfactions. For example, Latent Semantic Analysis (LSA) model (Xu *et al.*, 2017; Xu, 2019, 2020a), LDA (Guo *et al.*, 2017) and text link analysis (Francesco and Roberta, 2019). Then, some methods are used to determine the relationship between service characteristics and satisfactions, such as regression analysis (Ye *et al.*, 2014; Kim and Canina, 2015; Xu *et al.*, 2017; Guo *et al.*, 2017; Xu, 2018), penalty reward contrast analysis (Tontini *et al.*, 2017; Bi *et al.*, 2020; Zhang *et al.*, 2021a–c), Kano model (Park *et al.*, 2021), and provide suggestions for the improvement of hotel service quality. Although many researchers focus on hotel industry, there is heterogeneity in research issues (Pokryshevskaya and Antipov, 2017). It is mainly divided into the influence of service perception among different market segments on satisfaction (Bi *et al.*, 2020; Radojevic *et al.*, 2018), the influence of service perception among different star hotels on satisfaction, and whether transnational culture has an impact on satisfaction (Padma and Ahn, 2020; Rajaguru and Hassanli, 2018). For example, regarding the question whether the core features of different star chain hotels and independent hotels have different effects on customer satisfaction, Xu (2019) extended the research by incorporating statistical tests into LSA to reveal the relevance customers give to each core feature of hotels' products and services offered by chain/independent hotels and those with different star levels. Regarding the impact of cross-border differences on online reviews, Francesco and Roberta (2019) adopted text link analysis to extrapolate the frequencies of predefined hotel features on which the multivariate analyses and tests were performed. In addition, some researchers also studied the role of health features in customer satisfaction and dissatisfaction and the prediction of customer satisfaction (Francesco and Roberta, 2019; Zhao *et al.*, 2019).

Customer satisfaction studies in other areas are divided into those related to electronics (Jiang *et al.*, 2010; You *et al.*, 2012), restaurants (Bilgihan *et al.*, 2018; Mathayomchan and Taecharungroj, 2020; Kwon *et al.*, 2022), airlines (Sezgen *et al.*, 2019; Abbas and Malik, 2021), tourism (Lu and Stepchenkova, 2012; Lai *et al.*, 2018; Park *et al.*, 2020), online health communities (López *et al.*, 2012; Chen *et al.*, 2021b) and shared accommodation (Lv *et al.*, 2021). For example, Nilashi *et al.* (2021) used classification and regression trees (CART) method to find the relationship between quality factors in order to predict customer preferences on healthy food choice. In the field of airlines, Lim and Lee (2020) used LDA thematic model to identify key features of airline service quality and reveal customer sentiment or attitude toward service quality of full service airlines (FSCs) and low cost airlines (LCCs). Additionally, some researchers have studied whether culture has an impact on restaurant and airline industry satisfaction (Nakayama and Wan, 2019; Stamolampros *et al.*, 2020).

Most researchers study customer satisfaction based on expectancy disconfirmation theory. They prefer to use online reviews of TripAdvisor.com and Booking.com to analyze customer satisfaction in the hotel industry. Online reviews from TripAdvisor.com and Yelp.com are used for airline and restaurant customer satisfaction analysis. Most of these studies have considered the linear and asymmetric influence between service perception and satisfaction. However, there may be multiple linear relationships between them rather than simple linear relationships. Therefore, choosing deep learning algorithm to analyze customer satisfaction is worth studying.

4.3 Requirement improvement priority analysis

Many researchers not only study customer satisfaction, but also analyze the improvement priority of requirements (product and service characteristics) on this basis, so as to provide more effective improvement suggestions for enterprise management. The research on the improvement priority of requirements is mainly carried out from two perspectives: requirement classification and customer preference analysis. The following is a summary of these two aspects.

4.3.1 Requirement improvement priority analysis based on requirement classification. By sorting out relevant literature, we find that researchers mainly use the Kano model, IPA model and chance algorithm to classify requirements, so as to obtain the improvement priority of different requirements categories. The number of studies based on the Kano model reached 11, followed by the application of the IPA model (6) and chance algorithm (2). Most of the research results are applied to the hotel industry, and detailed classification is shown in Table 7.

Kano model bridges between customer requirements and product design from a perspective of distinguishing the nonlinear relation between the fulfillment of product features and the improvement of customer satisfaction, including must-be (M), one-dimensional (O), attractive (A), indifference (I) and reverse (R). In general, the priority of improved requirements is arranged in order of M, O, A. According to the research on requirement classification based on the Kano model, the general steps can be divided into (1) Extracting product and service features as customers' concerns; (2) Sentiment analysis; (3) Obtaining the relationship between product/service characteristics and customer satisfaction; (4) Feature classification using the

References	Journals	Kano model	IPA model	Opportunity algorithm
Qi *et al.* (2016)	*Information and Management*	✓		
Ou *et al.* (2018)	*Electronic Commerce Research and Applications*	✓		
Bi *et al.* (2019a)	*International Journal of Production Research*	✓		
Martí Bigorra *et al.* (2019)	*International Journal of Information Management*	✓		
Jin *et al.* (2021)	*International Journal of Production Research*	✓		
Joung and Kim (2021)	*International Journal of Production Research*	✓		
Chen *et al.* (2021a)	*Journal of Ambient Intelligence and Humanized Computing*	✓		
Zhang *et al.* (2021c)	*Knowledge-Based Systems*	✓		
Shi and Peng (2021)	*Advanced Engineering Informatics*	✓		
Li *et al.* (2021)	*Research in Engineering Design*	✓		
Park *et al.* (2021)	*International Journal of Contemporary Hospitality Management*	✓		
Bi *et al.* (2019b)	*Tourism Management*		✓	
Wang *et al.* (2020a)	*Information Systems and E-Business Management*		✓	
Ranjbari *et al.* (2020)	*International Journal of Contemporary Hospitality Management*		✓	
Hu *et al.* (2020)	*International Journal of Hospitality Management*		✓	
Luo *et al.* (2021)	*Tourism Management Perspectives*		✓	
Hu *et al.* (2021)	*Tourism Management*		✓	
Jeong *et al.* (2019)	*International Journal of Information Management*			✓
Choi *et al.* (2020)	*Technological Forecasting and Social Change*			✓

Table 7.
Distribution of publications based on different classification models

Kano model. Among them, it can be classified according to whether to build satisfaction function or not. Some researchers analyze the relationship between product/service characteristics and customer satisfaction by constructing satisfaction function (Qi *et al.*, 2016; Shi and Peng, 2021). Some researchers use machine learning methods to capture the relationship between product/service features and customer satisfaction (Bi *et al.*, 2019a; Joung and Kim, 2021; Li *et al.*, 2021). Ou *et al.* (2018) input the non-linear relationship between feature performance and customer satisfaction into the support vector machine (SVM) model to obtain the classification of attractive features in order to determine the attractiveness features of products or services. In order to evaluate the impact of product customer requirements on overall satisfaction, Chen *et al.* (2021a) introduced a multiple neural network (MNN) model that is advanced on the basis of back propagation neural network (BPNN). This approach considered the implicit characteristics of a product. Some other researchers classify features through the relationship between product feature importance and emotional score, providing reference for the improvement of product and service quality.

Standard importance-performance analysis (SIPA) was first proposed by Martilla and James (1977), which helps managers to allocate limited resources or formulate improvement strategies for products/services by investigating the feature performance and importance of the product/service (Bi *et al.*, 2019b). In SIPA, the features of a product/service are divided into four quadrants using their performance and importance. Quadrant 1 (Q1), Quadrant 2 (Q2), Quadrant 3 (Q3) and Quadrant 4 (Q4) are referred to as "keep up the good work", "concentrate here", "low priority" and "possible overkill", respectively. Features located in Q1, Q2, Q3 and Q4 are regarded as the main advantage and potential competitive advantage, the main weakness, the secondary weakness and the waste of limited resources of the product or service, respectively. Therefore, customer requirements in Q1 should be improved first, followed by Q2, Q3 and Q4. Similar to that based on the Kano model, the extended research based on IPA model is also mainly focused on the hotel field (Ranjbari *et al.*, 2020; Hu *et al.*, 2020). Bi *et al.* (2019b) proposed a feature importance assessment method based on the Ensemble Neural Network Based Model (ENNM) for the service improvement of two five-star hotels in Singapore, and used IPA model to classify features. Hu *et al.* (2021) using Kh-coder to identify hotel features. Moreover, the comparative Importance–Performance Analysis (CIPA) is constructed to classify features. This paper is one of the first studies to investigate and compare traveler preferences during different phases of a public health crisis. In addition, other researchers have focused their research on the improvement of products and services in automobiles (Wang *et al.*, 2020a), geoparks and geotourism (Luo *et al.*, 2021).

The opportunity algorithm is used to identify the extent to which each product topic is a potential opportunity for improvement from a customer-centered view. For the improvement of smart speakers, Jeong *et al.* (2019) used sentiment analysis and an opportunity algorithm to evaluate time-evolving events and provides quantified clues regarding product-development directions based on events with high opportunity. Finally, the improved event is obtained based on opportunity landscape map. In view of the improvement of Samsung Galaxy Note 5 (SGN5), Choi *et al.* (2020) proposed an opportunity mining method based on topic modeling and sentiment analysis of social media data to identify product opportunities and determine the opportunity value and improvement direction of each product theme.

4.3.2 Requirement improvement priority analysis based on customer preference analysis. Some researchers focus on customer preference analysis and support managers to obtain customer preferences from online reviews, thus improving customer satisfaction and enterprise performance.

Their research on customer preferences focuses on hotels (Xu, 2020b; Nilashi *et al.*, 2021; Yadegaridehkordi *et al.*, 2021) and electronics (Decker and Trusov, 2010; Oh and Yi, 2021; Zhou *et al.*, 2022). In the hotel industry, Hu *et al.* (2019) collected 500,000 online customer reviews from TripAdvisor to determine the customer requirements that affect customer satisfaction and actual repeat customer visits through empirical analysis. It is found that the influence of service-related features on overall customer satisfaction increases with the increase of visits. To better deal with the implicit aspect-level terms extraction, Zhang *et al.* (2021b) proposed an unsupervised approach for aspect-level sentiment analysis with the implicit hotel features into consideration by integrating word embedding, co-occurrence and dependency parsing. The study found that different types of customers pay different attention to hotel features, which is helpful for hotels to identify service characteristics that need to be improved. For the study of electronic products, aiming at the improvement of new energy vehicles, in order to identify the requirement differences between different customer groups, Cai *et al.* (2021) proposed a product-and-user oriented approach (PURA) for requirement analysis to identify the difference in requirements between customers. Multi-attribute decision making method is a common ranking method in decision making field (Lin *et al.*, 2020), so some researchers used extended multi-attribute decision making method to analyze customer preferences (Yadegaridehkordi *et al.*, 2021; Du *et al.*, 2022).

Through the summary of the above literature, we find that most studies need to use positive and negative reviews to analyze the emotion of requirement characteristics when analyzing the priority of requirement improvement. In this process, they regard positive and negative reviews as homogenous. Some studies show that there are significant differences between positive and negative evaluations in the responsiveness of customers. In addition, when analyzing customer preferences, most studies do not take customers' psychological and other behavioral characteristics into account, which will affect decision making. Therefore, it is a future research direction to make use of the heterogeneity of positive and negative evaluation to conduct sentiment analysis and customer preference analysis considering the psychological behavior of customers.

4.4 Engineering characteristic analysis

In customer-driven product design and product improvement, after successfully identifying customer requirements, designers begin to consider how to interpret customer requirements to improve products. They want to use their limited resources to maximize customer satisfaction, profit and social prestige. The Quality Function Development (QFD) model has been widely used to bridge the gap between customer requirements and design characteristics. Some researchers have carried out research on how to transform customer requirements (CRs) into engineering characteristics (ECs) in QFD.

To help designers translate online reviews into ECs, Ha and Geum (2022) proposed a probabilistic linguistic analysis method. Based on the unigram model and the bigram model, an integrated impact learning algorithm is advised to estimate the impacts of keywords and nearby words, respectively. The estimated impacts are implied which ECs are implied in a given context. This is the first attempt to integrate a large number of online reviews directly into QFD. In order to reduce the time complexity of voice of the customer (VoC) analysis, Özdağoğlu *et al.* (2018) proposed the QFD method and topic modeling aggregation decision framework to extract the real customer requirements. To improve key functions of products with short life cycles, Trappey *et al.* (2018) proposed a computer supported approach based on technical function deployment (TFD) and an extended eQFD approach to analyze R&D priorities based on customer requirements. Chin *et al.* (2019) used the interval-valued hierarchy process (I-AHP) to process the collected data and improve the cabin design from the passenger's perspective. This method is not only convenient for respondents to express evaluation, but also reflects the uncertainty and fuzziness of subjective evaluation.

In addition, some researchers combine the QFD model with Kano model to analyze CRs and ECs from qualitative and quantitative perspectives. In order to quantitatively analyze the target value of ECs, Ji *et al.* (2014) combined QFD and Kano model to establish a hybrid nonlinear integer programming model under cost and technical constraints, so as to define and set the target value of ECs. He *et al.* (2017) proposed an improved Kano model named as importance- Frequency Kano (IF-Kano) Model. The qualitative and quantitative results of the IF-Kano model are integrated into QFD using nonlinear programming model, and the target value of engineering features is obtained under the condition of optimal balance between enterprise satisfaction and customer satisfaction.

Some researchers combined the Kano model with QFD to qualitatively analyze the improvement priorities of ECs. However, few studies have obtained the definite target value of ECs. Therefore, how to use a large number of online reviews and combine the Kano and QFD models to obtain the improvement degree of ECs is the future research direction. In addition, the traditional house of quality (HOQ) can also be improved. For example, the uncertainty of decision evaluation can be considered when the correlation between CRs and ECs is obtained. Considering the interdependence of factors, the autocorrelation of CRs or ECs was studied during the calculation of correlation matrix.

4.5 Product defect identification

In addition to developing product improvement strategies from the perspective of customer requirements, product defect identification is also an important research direction. However, there are still few researches on product defect identification based on online reviews.

Product performance defects and safety defects are increasingly discussed on social media. Defective products will generate negative word-of-mouth, and large performance defects will lead to unimaginable economic and safety losses for customers. Therefore, product improvement based on product defect identification will help to improve customer satisfaction and enterprise competitiveness. Traditionally, product defect information collection sources have been mainly quality tests and feedback from after-sales service centers. However, this method has some disadvantages such as high cost, incomprehensiveness and untimely. Online reviews become a valuable resource for detecting defects in designing and manufacturing processes.

The existing researches are mainly distributed in the defect identification of automobiles, electronic products, electrical appliances and other products, focusing on feature selection (Zheng *et al.*, 2020), sentiment analysis (Zhang *et al.*, 2012; Abbas and Malik, 2021) and classification methods (Liu *et al.*, 2018). To understand the severity of the defect, Zheng *et al.* (2021) proposed a novel approach that integrates the probabilistic graphic model named product defect identification and analysis model (PDIAM) with failure mode and effect analysis (FMEA) to derive product defect information. They utilized FMEA based on the results of PDIAM and alleviated the inherent subjectivity brought by expert evaluation. For defect identification of household appliances, Law *et al.* (2017) obtained online product reviews from Amazon.com and Bestbuy.com. The result shows that domain-specific sparkle and smoke words for the dishwasher product category are critical to efficient automated defect discovery in dishwashers. In order to identify which words and phrases in crib reviews are indicators of defects, Mummalaneni *et al.* (2018) first constructed smoke term lists, and then created training sets and retention sets for defect classification. They found that sentiment analysis serves as a useful tool for automated defect discovery in the baby crib industry. In addition, they created a supplementary set of "smoke terms" that are strong indicators of safety defects in online reviews of baby cribs.

The basic process of most product defect identification studies is as follows: (1) Selecting appropriate features (linguistic features, social features, sentiment features, etc.) or constructing the smoke word list by using text analysis method; (2) Using machine learning method to identify product defect text. However, they rarely further analyze the detailed defect information and the severity of the defect, which is especially important for the manufacturer to improve the product. Therefore, how to analyze the detailed defect information and the severity of the defect is worth studying in the future.

4.6 Competitive analysis

In order to gain competitive advantage and effectively assess the business environment, it is often necessary for enterprises to pay attention to the views of customers on competitors' products and services. When the enterprise has its own competitive advantage, it can make reasonable marketing strategies according to the results of competition analysis to continue to expand the gap; when an enterprise has a disadvantage in some aspects, it can formulate accurate and reasonable product and service improvement strategies according to the results of competition analysis, so as to narrow the gap with competitors.

The research on competition analysis mainly involves the following problems: key customer requirement mining (Albayrak *et al.*, 2021; Gang and Chenglin, 2021), market structure analysis (Netzer *et al.*, 2012), competitors and competitive groups identification (Hu and Trivedi, 2020; Liu *et al.*, 2019; Marcolin *et al.*, 2021), main competitive advantages and disadvantages analysis (He *et al.*, 2015; Jin *et al.*, 2016a; Kim and Kang, 2018; Wang *et al.*, 2018a), how to improve the performance level of products or services in the competitive environment (Rodríguez-Díaz and Espino-Rodríguez, 2018). In order to identify competitors, Gao *et al.* (2018) proposed a new model to extract comparison relationships from online reviews and constructed three types of comparison relationship networks. They have realized three kinds of competitiveness analysis: analyzing market structure, identifying top competitors, and identifying strengths and weaknesses. In order to identify benchmark hotels as competitors, Park and Lee (2021) proposed an output-oriented data envelopment analysis to calculate the degree of guest satisfaction and service positioning for each hotel, achieved the degree of feature improvement compared with competitors. Competitiveness analysis should not only focus on identifying competitors, but also improve the competitiveness of enterprises by analyzing the competitive advantages and disadvantages of all competitors (Liu *et al.*, 2021). Therefore, in view of the market advantages and disadvantages of online celebrity shops, Liang *et al.* (2020) used fuzzy cognitive graph and association rules to construct a relationship graph between the shop features of web celebrity shop. Through competitive analysis, suggestions for improvement were put forward for each store. To address the common advantages and disadvantages of different batteries, Jin *et al.* (2019) proposed a framework of representative sentence sampling from online reviews of series products. Representative sentences sampling from online customer concerns is formulated as an optimization problem and a greedy algorithm is designed to obtain sampling results effectively.

In the existing studies, the competitiveness indicators are usually the importance of requirements, customer sentiment score, etc., which are relatively simple. Other indicators, such as sales volume, number of customer reviews and likes, can be considered. What's more, as customer reviews become more syntactically and semantically complex, comparison statements between products are harder to extract, so more effective methods need to be developed to further explore the impact of comparative opinions on competitive analysis.

5. Main findings and future research

By reviewing the research on eWOM for product and service quality improvement, this section summarizes the main findings, and discusses the future development direction of this field. Based on the results of bibliometric analysis, we can get the following findings:

(1) In terms of the number of publications each year, there is an increasing trend, mainly published in Tourism Management, International Journal of Hospitality Management

and other tourism periodicals. In addition, comparing with previous publications, there are fewer citations in recent years, while the highest number of citations being 737, which also indicates that researchers need to further improve the quality of research.

(2) From the perspective of distribution of regional contribution, researchers all over the world are paying attention to the development of this field. China has the highest number of high-quality papers, with 52 published in the last 14 years. In the future, it is hoped that countries can strengthen international cooperation and provide fresh blood for research in this field according to their national conditions and research expertise.

(3) From the distribution of discipline contribution, the fields of "Hospitality, Leisure, Sport and Tourism" published the most relevant studies. Following by are Management, Computer Science Artificial Intelligence and Computer Science Information Systems. At present, the research trend of the academic tends to be the integration of disciplines, so researchers in different fields can strengthen cooperation and share the latest research results in the field, which is conducive to obtaining more cutting-edge scientific research results.

(4) From the influential literature, most scholars have expanded the research of 22 basic papers. The research directions include customer requirement mining, competition analysis, product defect identification, requirement classification, text analysis, etc. These 22 articles laid a foundation for further research.

(5) From the perspective and the development trend of research topics, this paper summarizes six research topics in this field. In addition, we find that the research topics gradually changed to sentiment analysis, behavior intention, multi-source data analysis and other directions, suggesting that researchers pay close attention to these aspects.

By summarizing the researches on eWOM for product and service quality improvement, we draw the following research directions in the future:

(1) How to use multi-source datasets to effectively identify implicit features in the process of requirement mining is the future research direction. In addition, customer requirements are still changing over time. In the future, we can analyze customer preferences from a dynamic perspective, so as to predict customer requirement accurately.

(2) As there is a complex asymmetric, nonlinear and multi-linear relationship between customer perception of service quality and customer satisfaction, it is worth studying to expand existing deep learning methods to analyze customer satisfaction. In addition, the expectancy disconfirmation theory can be combined with other theories to conduct more accurate customer satisfaction research.

(3) When analyzing customer satisfaction, it is worthwhile to consider the difference between positive and negative evaluation statements in customer perception. Customers' reactions to positive and negative evaluations are different, which will affect their purchase decisions and thus affect the formulation of improvement strategies by enterprise managers. In addition, it is necessary to consider the psychology and behavioral intention of customers.

(4) Customer preference analysis under uncertain environment is a research direction in the future. When customers make reviews, they often use different words to express

different sentiments and requirements. It is worth studying how to express customer evaluation information more reasonably and construct customer preference model (Li *et al.*, 2021).

(5) Considering the uncertainty of decision evaluation, quantifying the improvement degree of ECs is the direction of future research. Besides, in the calculation of the correlation matrix of the HOQ, the interdependence between factors can be considered to study the influence of the autocorrelation of CRs or ECs on the research results.

(6) It will be an interesting research topic to consider sales volume, number of customer reviews and likes as new indicators in sentiment analysis and competitiveness analysis. Finally, as customer reviews become more syntactically and semantically complex, comparison statements between products are harder to extract, so more effective methods need to be developed to further explore the impact of comparative opinions on competitive analysis.

In addition to the above future research directions, researchers can expand their research fields to other highly competitive product or service industries, so as to increase the competitiveness of enterprises and promote the benign development of society.

6. Discussion and conclusions

6.1 Conclusions

This paper selects 110 valuable papers published from 2009 to 2022 from Web of Science, Elsevier and SpingerLink databases. Using bibliometric methods, we summarize the research of eWOM for product and service quality improvement. Firstly, this paper conducts bibliometric analysis to obtain publishing trend and citation status, regional and discipline area distribution, influential publications. VOSviewer software is used to draw the co-citation network map, and the foundational papers in this field are obtained. In addition, the main research topics and the development trend of research hotspots are determined through keyword co-occurrence analysis. Secondly, according to the analysis chart of research topics, all papers are divided into six major topics and reviewed, respectively. Finally, some research findings and future development directions of eWOM research for product and service quality improvement are summarized. This paper provides a scientific way for researchers to understand the development trend and hotspots of research, and provides a solid and reliable foundation for future research and development. What's more, this paper can provide some guidance for eWOM, product improvement, product design, service improvement and word-of-mouth marketing related application fields.

6.2 Limitations and future research

One limitation of this work is that we do not consider any keywords in the content of the paper in the topic co-occurrence analysis, which may affect the accuracy of the keyword network map and the topic evolution time map. Secondly, when conducting literature retrieval, only three databases are used. This may lead to the defect of an incomplete literature search. What's more, VOSviewer only allows using the bibliographic data from one database, instead of the data combined from any two databases. Although the 110 papers can all be retrieved in the Web of Science database, it is not feasible to use the bibliographic data from both databases with VOSviewer. If there is a part of the document that only appears in one database and another document that only appears in the other database. Therefore, in the future, we can consider expanding the search scope, such as adding the Scopus database. Future studies may adopt other methods to review and analyze.

References

Abbas, Y. and Malik, M.S.I. (2021), "Defective products identification framework using online reviews", *Electronic Commerce Research*, pp. 1-22.

Albayrak, T., Cengizci, A.D., Caber, M. and Nang Fong, L.H. (2021), "Big data use in determining competitive position: the case of theme parks in Hong Kong", *Journal of Destination Marketing and Management*, Vol. 22, p. 100668.

Archak, N., Ghose, A. and Ipeirotis, P.G. (2011), "Deriving the pricing power of product features by mining consumer reviews", *Management Science*, Vol. 57 No. 8, pp. 1485-1509.

Ayağ, Z. (2021), "A comparison study of fuzzy-based multiple-criteria decision-making methods to evaluating green concept alternatives in a new product development environment", *International Journal of Intelligent Computing and Cybernetics*, Vol. 14 No. 3, pp. 412-438.

Azzopardi, E. and Nash, R. (2013), "A critical evaluation of importance–performance analysis", *Tourism Management*, Vol. 35, pp. 222-233.

Barnes, S.J., Mattsson, J., Sørensen, F. and Jensen, J.F. (2020), "Measuring employee-tourist encounter experience value: a big data analytics approach", *Expert Systems with Applications*, Vol. 154, p. 113450.

Berezina, K., Bilgihan, A., Cobanoglu, C. and Okumus, F. (2016), "Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews", *Journal of Hospitality Marketing and Management*, Vol. 25 No. 1, pp. 1-24.

Bi, J.W., Liu, Y., Fan, Z.P. and Cambria, E. (2019a), "Modelling customer satisfaction from online reviews using ensemble neural network and effect-based Kano model", *International Journal of Production Research*, Vol. 57 No. 22, pp. 7068-7088.

Bi, J.W., Liu, Y., Fan, Z.P. and Zhang, J. (2019b), "Wisdom of crowds: conducting importance-performance analysis (IPA) through online reviews", *Tourism Management*, Vol. 70 No. 1, pp. 460-478.

Bi, J.W., Liu, Y., Fan, Z.P. and Zhang, J. (2020), "Exploring asymmetric effects of attribute performance on customer satisfaction in the hotel industry", *Tourism Management*, Vol. 77, p. 104006.

Bilgihan, A., Seo, S. and Choi, J. (2018), "Identifying restaurant satisfiers and dissatisfiers: suggestions from online reviews", *Journal of Hospitality Marketing and Management*, Vol. 27 No. 5, pp. 601-625.

Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003), "Latent dirichlet allocation", *Journal of Machine Learning Research*, Vol. 3, pp. 993-1022.

Browning, V., So, K.K.F. and Sparks, B. (2013), "The influence of online reviews on consumers' attributions of service quality and control for service standards in hotels", *Journal of Travel and Tourism Marketing*, Vol. 30 Nos 1-2, pp. 23-40.

Cai, M., Tan, Y., Ge, B., Dou, Y., Huang, G. and Du, Y. (2021), "PURA: a product-and-user oriented approach for requirement analysis from online reviews", *IEEE Systems Journal*, Vol. 16 No. 1, pp. 566-577.

Chen, R., Wang, Q. and Xu, W. (2019), "Mining user requirements to facilitate mobile app quality upgrades with big data", *Electronic Commerce Research and Applications*, Vol. 38, p. 100889.

Chen, K., Jin, J. and Luo, J. (2021a), "Big consumer opinion data understanding for Kano categorization in new product development", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 13 No. 4, pp. 2269-2288.

Chen, Q., Jin, J. and Yan, X. (2021b), "Understanding online review behaviors of patients in online health communities: an expectation-disconfirmation perspective", *Information Technology and People*, ahead-of-print (ahead-of-print), doi: 10.1108/ITP-04-2021-0290.

Chevalier, J.A. and Mayzlin, D. (2006), "The effect of word of mouth on sales: online book reviews", *Journal of Marketing Research*, Vol. 43 No. 3, pp. 345-354.

Chin, K.S., Yang, Q., Chan, C.Y., Tsui, K.L. and Li, Y.L. (2019), "Identifying passengers' needs in cabin interiors of high-speed rails in China using quality function deployment for improving passenger satisfaction", *Transportation Research Part A: Policy and Practice*, Vol. 119, pp. 326-342.

Choi, J., Oh, S., Yoon, J., Lee, J.M. and Coh, B.Y. (2020), "Identification of time-evolving product opportunities via social media mining", *Technological Forecasting and Social Change*, Vol. 156, p. 120045.

Culotta, A. and Cutler, J. (2016), "Mining brand perceptions from twitter social networks", *Marketing Science*, Vol. 35 No. 3, pp. 343-362.

Decker, R. and Trusov, M. (2010), "Estimating aggregate consumer preferences from online product reviews", *International Journal of Research in Marketing*, Vol. 27 No. 4, pp. 293-307.

Deng, W.J., Yeh, M.L. and Sung, M.L. (2013), "A customer satisfaction index model for international tourist hotels: integrating consumption emotions into the American Customer Satisfaction Index", *International Journal of Hospitality Management*, Vol. 35, pp. 133-140.

Du, Y., Liu, D. and Duan, H. (2022), "A textual data-driven method to identify and prioritise user preferences based on regret/rejoicing perception for smart and connected products", *International Journal of Production Research*. doi: 10.1080/00207543.2021.2023776.

Francesco, G. and Roberta, G. (2019), "Cross-country analysis of perception and emphasis of hotel attributes", *Tourism Management*, Vol. 74, pp. 24-42.

Gang, Z. and Chenglin, L. (2021), "Dynamic measurement and evaluation of hotel customer satisfaction through sentiment analysis on online reviews", *Journal of Organizational and End User Computing*, Vol. 33 No. 6, pp. 1-27.

Gao, S., Tang, O., Wang, H. and Yin, P. (2018), "Identifying competitors through comparative relation mining of online reviews in the restaurant industry", *International Journal of Hospitality Management*, Vol. 71, pp. 19-32.

Guo, Y., Barnes, S.J. and Jia, Q. (2017), "Mining meaning from online ratings and reviews: tourist satisfaction analysis using latent dirichlet allocation", *Tourism Management*, Vol. 59 No. 2, pp. 467-483.

Ha, S. and Geum, Y. (2022), "Identifying new innovative services using M&A data: an integrated approach of data-driven morphological analysis", *Technological Forecasting and Social Change*, Vol. 174, p. 121197.

He, W., Wu, H., Yan, G., Akula, V. and Shen, J. (2015), "A novel social media competitive analytics framework with sentiment benchmarks", *Information and Management*, Vol. 52 No. 7, pp. 801-812.

He, L., Song, W., Wu, Z., Xu, Z., Zheng, M. and Ming, X. (2017), "Quantification and integration of an improved Kano model into QFD based on multi-population adaptive genetic algorithm", *Computers and Industrial Engineering*, Vol. 114, pp. 183-194.

Hennig-Thurau, T., Gwinner, K.P., Walsh, G. and Gremler, D.D. (2004), "Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the Internet?", *Journal of Interactive Marketing*, Vol. 18 No. 1, pp. 38-52.

Hu, M. and Liu, B. (2004), "Mining and summarizing customer reviews", in *Proceedings of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '04*, ACM Press, New York, USA, pp. 168-177, doi: 10.1145/1014052.1014073.

Hu, F. and Trivedi, R.H. (2020), "Mapping hotel brand positioning and competitive landscapes by text-mining user-generated content", *International Journal of Hospitality Management*, Vol. 84, p. 102317.

Hu, F., Teichert, T., Liu, Y., Li, H. and Gundyreva, E. (2019), "Evolving customer expectations of hospitality services: differences in attribute effects on satisfaction and Re-Patronage", *Tourism Management*, Vol. 74, pp. 345-357.

Hu, F., Li, H., Liu, Y. and Teichert, T. (2020), "Optimizing service offerings using asymmetric impact-sentiment-performance analysis", *International Journal of Hospitality Management*, Vol. 89, p. 102557.

Hu, F., Teichert, T., Deng, S., Liu, Y. and Zhou, G. (2021), "Dealing with pandemics: an investigation of the effects of COVID-19 on customers' evaluations of hospitality services", *Tourism Management*, Vol. 85, p. 104320.

Ibrahim, N.F. and Wang, X. (2019), "A text analytics approach for online retailing service improvement: evidence from Twitter", *Decision Support Systems*, Vol. 121, pp. 37-50.

Jeong, B., Yoon, J. and Lee, J.M. (2019), "Social media mining for product planning: a product opportunity mining approach based on topic modeling and sentiment analysis", *International Journal of Information Management*, Vol. 48, pp. 280-290.

Ji, P., Jin, J., Wang, T. and Chen, Y. (2014), "Quantification and integration of Kano's model into QFD for optimising product design", *International Journal of Production Research*, Vol. 52 No. 21, pp. 6335-6348.

Jiang, Y., Shang, J. and Liu, Y. (2010), "Maximizing customer satisfaction through an online recommendation system: a novel associative classification model", *Decision Support Systems*, Vol. 48 No. 3, pp. 470-479.

Jiang, C., Liu, Y., Ding, Y., Liang, K. and Duan, R. (2017), "Capturing helpful reviews from social media for product quality improvement: a multi-class classification approach", *International Journal of Production Research*, Vol. 55 No. 12, pp. 3528-3541.

Jiao, Y. and Qu, Q.X. (2019), "A proposal for Kansei knowledge extraction method based on natural language processing technology and online product reviews", *Computers in Industry*, Vol. 108, pp. 1-11.

Jin, J., Ji, P., Liu, Y. and Johnson Lim, S.C. (2015), "Translating online customer opinions into engineering characteristics in QFD: a probabilistic language analysis approach", *Engineering Applications of Artificial Intelligence*, Vol. 41, pp. 115-127.

Jin, J., Ji, P. and Gu, R. (2016a), "Identifying comparative customer requirements from product online reviews for competitor analysis", *Engineering Applications of Artificial Intelligence*, Vol. 49, pp. 61-73.

Jin, J., Ji, P. and Kwong, C.K. (2016b), "What makes consumers unsatisfied with your products: review analysis at a fine-grained level", *Engineering Applications of Artificial Intelligence*, Vol. 47, pp. 38-48.

Jin, J., Ji, P. and Yan, S. (2019), "Comparison of series products from customer online concerns for competitive intelligence", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 10 No. 3, pp. 937-952.

Jin, J., Jia, D. and Chen, K. (2021), "Mining online reviews with a Kansei-integrated Kano model for innovative product design", *International Journal of Production Research*. doi: 10.1080/00207543.2021.1949641.

Joung, J. and Kim, H.M. (2021), "Explainable neural network-based approach to Kano categorisation of product features from online reviews", *International Journal of Production Research*. doi: 10.1080/00207543.2021.2000656.

Kano, N., Seraku, N., Takahashi, F. and Tsuji, S. (1984), "Attractive quality and must-be quality", *Journal of the Japanese Society for Quality Control*, Vol. 14 No. 2, pp. 39-48.

Kim, J.Y. and Canina, L. (2015), "An analysis of smart tourism system satisfaction scores: the role of priced versus average quality", *Computers in Human Behavior*, Vol. 50, pp. 610-617.

Kim, S.G. and Kang, J. (2018), "Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews", *Information Processing and Management*, Vol. 54 No. 6, pp. 938-957.

Kwok, L., Xie, K.L. and Richards, T. (2017), "Thematic framework of online review research", *International Journal of Contemporary Hospitality Management*, Vol. 29 No. 1, pp. 307-354.

Kwon, W., Lee, M. and Back, K.J. (2020), "Exploring the underlying factors of customer value in restaurants: a machine learning approach", *International Journal of Hospitality Management*, Vol. 91, p. 102643.

Kwon, W., Lee, M. and Bowen, J.T. (2022), "Exploring customers' luxury consumption in restaurants: a combined method of topic modeling and three-factor theory", *Cornell Hospitality Quarterly*, Vol. 63 No. 1, pp. 66-77.

Lai, I.K.W., Hitchcock, M., Yang, T. and Lu, T.W. (2018), "Literature review on service quality in hospitality and tourism (1984–2014)", *International Journal of Contemporary Hospitality Management*, Vol. 30 No. 1, pp. 114-159.

Law, D., Gruss, R. and Abrahams, A.S. (2017), "Automated defect discovery for dishwasher appliances from online consumer reviews", *Expert Systems with Applications*, Vol. 67, pp. 84-94.

Lee, C.K.H. and Tse, Y.K. (2020), "Improving peer-to-peer accommodation service based on text analytics", *Industrial Management and Data Systems*, Vol. 121 No. 2, pp. 209-227.

Li, G., Law, R., Vu, H.Q., Rong, J., Zhao, X. and Roy. (2015), "Identifying emerging hotel preferences using emerging pattern mining technique", *Tourism Management*, Vol. 46, pp. 311-321.

Li, S., Liu, F., Lu, H., Zhang, Y., Li, Y. and Yu, Z. (2021), "Product family lean improvement based on matching deep mining of customer group preference", *Research in Engineering Design*, Vol. 32 No. 4, pp. 469-488.

Liang, D., Dai, Z., Wang, M. and Li, J. (2020), "Web celebrity shop assessment and improvement based on online review with probabilistic linguistic term sets by using sentiment analysis and fuzzy cognitive map", *Fuzzy Optimization and Decision Making*, Vol. 19 No. 4, pp. 561-586.

Liao, H., Mi, X. and Xu, Z. (2020), "A survey of decision-making methods with probabilistic linguistic information: bibliometrics, preliminaries, methodologies, applications and future directions", *Fuzzy Optimization and Decision Making*, Vol. 19 No. 1, pp. 81-134.

Lim, J. and Lee, H.C. (2020), "Comparisons of service quality perceptions between full service carriers and low cost carriers in airline travel", *Current Issues in Tourism*, Vol. 23 No. 10, pp. 1261-1276.

Lin, M., Huang, C., Xu, Z. and Chen, R. (2020), "Evaluating IoT platforms using integrated probabilistic linguistic MCDM method", *IEEE Internet of Things Journal*, Vol. 7 No. 11, pp. 11195-11208.

Lin, M., Chen, Y. and Chen, R. (2021a), "Bibliometric analysis on Pythagorean fuzzy sets during 2013-2020", *International Journal of Intelligent Computing and Cybernetics*, Vol. 14 No. 2, pp. 104-121.

Lin, M., Chen, Z., Xu, Z., Gou, X. and Herrera, F. (2021b), "Score function based on concentration degree for probabilistic linguistic term sets: an application to TOPSIS and VIKOR", *Information Sciences*, Vol. 551, pp. 270-290.

Liu, Y., Jiang, C. and Zhao, H. (2018), "Using contextual features and multi-view ensemble learning in product defect identification from online discussion forums", *Decision Support Systems*, Vol. 105, pp. 1-12.

Liu, Y., Jiang, C. and Zhao, H. (2019), "Assessing product competitive advantages from the perspective of customers by mining user-generated content on social media", *Decision Support Systems*, Vol. 123 No. 8, p. 113079.

Liu, Z., Qin, C.X. and Zhang, Y.J. (2021), "Mining product competitiveness by fusing multisource online information", *Decision Support Systems*, Vol. 143, p. 113477.

López, A., Detz, A., Ratanawongsa, N. and Sarkar, U. (2012), "What patients say about their doctors online: a qualitative content analysis", *Journal of General Internal Medicine*, Vol. 27 No. 6, pp. 685-692.

Lu, W. and Stepchenkova, S. (2012), "Ecotourism experiences reported online: classification of satisfaction attributes", *Tourism Management*, Vol. 33 No. 3, pp. 702-712.

Luo, Y. and Lin, M. (2021), "Flash translation layer: a review and bibliometric analysis", *International Journal of Intelligent Computing and Cybernetics*, Vol. 14 No. 3, pp. 480-508.

Luo, Y., He, J., Mou, Y., Wang, J. and Liu, T. (2021), "Exploring China's 5A global geoparks through online tourism reviews: a mining model based on machine learning approach", *Tourism Management Perspectives*, Vol. 37, p. 100769.

Lv, X., Liu, Y., Xu, S. and Li, Q. (2021), "Welcoming host, cozy house? The impact of service attitude on sensory experience", *International Journal of Hospitality Management*, Vol. 95, p. 102949.

Maditham, V., Reddy, N.S. and Kasa, M. (2022), "A multi-preference integrated algorithm (MPIA) for the deep learning-based recommender framework (DLRF)", *International Journal of Intelligent Computing and Cybernetics*. doi: 10.1108/IJICC-11-2021-0257.

Marcolin, C.B., Becker, J.L., Wild, F., Behr, A. and Schiavi, G. (2021), "Listening to the voice of the guest: a framework to improve decision-making processes with text data", *International Journal of Hospitality Management*, Vol. 94, p. 102853.

Martí Bigorra, A., Isaksson, O. and Karlberg, M. (2019), "Aspect-based Kano categorization", *International Journal of Information Management*, Vol. 46 No. 3, pp. 163-172.

Martilla, J.A. and James, J.C. (1977), "Importance-performance analysis", *Journal of Marketing*, Vol. 41 No. 1, p. 77. No.

Mathayomchan, B. and Taecharungroj, V. (2020), "How was your meal?" Examining customer experience using Google maps reviews", *International Journal of Hospitality Management*, Vol. 90, p. 102641.

Mudambi, S.M. and Schuff, D. (2010), "Research note: what makes a helpful online review? A study of customer reviews on Amazon.com", *MIS Quarterly*, Vol. 34 No. 1, pp. 185-200.

Mummalaneni, V., Gruss, R., Goldberg, D.M., Ehsani, J.P. and Abrahams, A.S. (2018), "Social media analytics for quality surveillance and safety hazard detection in baby cribs", *Safety Science*, Vol. 104, pp. 260-268.

Muritala, B.A., Sánchez-Rebull, M.V. and Hernández-Lara, A.B. (2020), "A bibliometric analysis of online reviews research in tourism and hospitality", *Sustainability*, Vol. 12 No. 23, p. 9977.

Nakayama, M. and Wan, Y. (2019), "The cultural impact on social commerce: a sentiment analysis on Yelp ethnic restaurant reviews", *Information and Management*, Vol. 56 No. 2, pp. 271-279.

Netzer, O., Feldman, R., Goldenberg, J. and Fresko, M. (2012), "Mine your own business: market-structure surveillance through text mining", *Marketing Science*, Vol. 31 No. 3, pp. 521-543.

Nilashi, M., Ahmadi, H., Arji, G., Alsalem, K.O., Samad, S., Ghabban, F., Alzahrani, A.O., Ahani, A. and Alarood, A.A. (2021), "Big social data and customer decision making in vegetarian restaurants: a combined machine learning method", *Journal of Retailing and Consumer Services*, Vol. 62, p. 102630.

Oh, Y.K. and Yi, J. (2021), "Asymmetric effect of feature level sentiment on product rating: an application of bigram natural language processing (NLP) analysis", *Internet Research*, ahead-of-print (ahead-of-print), doi: 10.1108/INTR-11-2020-0649.

Oliver, R.L. (1980), "A cognitive model of the antecedents and consequences of satisfaction decisions", *Journal of Marketing Research*, Vol. 17 No. 4, pp. 460-469.

Ou, W., Huynh, V.N. and Sriboonchitta, S. (2018), "Training attractive attribute classifiers based on opinion features extracted from review data", *Electronic Commerce Research and Applications*, Vol. 32, pp. 13-22.

Özdağoğlu, G., Kapucugil-İkiz, A. and Çelik, A.F. (2018), "Topic modelling-based decision framework for analysing digital voice of the customer", *Total Quality Management and Business Excellence*, Vol. 29 Nos 13-14, pp. 1545-1562.

Padma, P. and Ahn, J. (2020), "Guest satisfaction & dissatisfaction in luxury hotels: an application of big data", *International Journal of Hospitality Management*, Vol. 84, p. 102318.

Pang, B. and Lee, L. (2008), "Opinion mining and sentiment analysis", *Foundations and Trends in Information Retrieval*, Vol. 2 Nos 1-2, pp. 1-135.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), "SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality", *Journal of Retailing*, Vol. 64 No. 1, pp. 12-40.

Park, J. and Lee, B.K. (2021), "An opinion-driven decision-support framework for benchmarking hotel service", *Omega*, Vol. 103, p. 102415.

Park, S., Lee, J.S. and Nicolau, J.L. (2020), "Understanding the dynamics of the quality of airline service attributes: satisfiers and dissatisfiers", *Tourism Management*, Vol. 81, p. 104163.

Park, H., Lee, M. and Back, K.J. (2021), "Exploring the roles of hotel wellness attributes in customer satisfaction and dissatisfaction: application of Kano model through mixed methods", *International Journal of Contemporary Hospitality Management*, Vol. 33 No. 1, pp. 263-285.

Ping, Y.J., Liu, R., Lin, W. and Liu, H.C. (2020), "A new integrated approach for engineering characteristic prioritization in quality function deployment", *Advanced Engineering Informatics*, Vol. 45, p. 101099.

Pizam, A., Shapoval, V. and Ellis, T. (2016), "Customer satisfaction and its measurement in hospitality enterprises: a revisit and update", *International Journal of Contemporary Hospitality Management*, Vol. 28 No. 1, pp. 2-35.

Pokryshevskaya, E.B. and Antipov, E.A. (2017), "Profiling satisfied and dissatisfied hotel visitors using publicly available data from a booking platform", *International Journal of Hospitality Management*, Vol. 67, pp. 1-10.

Qi, J., Zhang, Z., Jeon, S. and Zhou, Y. (2016), "Mining customer requirements from online reviews: a product improvement perspective", *Information and Management*, Vol. 53 No. 8, pp. 951-963. No.

Radojevic, T., Stanisic, N., Stanic, N. and Davidson, R. (2018), "The effects of traveling for business on customer satisfaction with hotel services", *Tourism Management*, Vol. 67, pp. 326-341.

Rajaguru, R. and Hassanli, N. (2018), "The role of trip purpose and hotel star rating on guests' satisfaction and WOM", *International Journal of Contemporary Hospitality Management*, Vol. 30 No. 5, pp. 2268-2286. No.

Ranjbari, M., Shams Esfandabadi, Z. and Scagnelli, S.D. (2020), "A big data approach to map the service quality of short-stay accommodation sharing", *International Journal of Contemporary Hospitality Management*, Vol. 32 No. 8, pp. 2575-2592.

Rodríguez-Díaz, M. and Espino-Rodríguez, T.F. (2018), "A methodology for a comparative analysis of the lodging offer of tourism destinations based on online customer reviews", *Journal of Destination Marketing and Management*, Vol. 8, pp. 147-160.

Rodrigues, H., Brochado, A. and Troilo, M. (2020), "Listening to the murmur of water: essential satisfaction and dissatisfaction attributes of thermal and mineral spas", *Journal of Travel and Tourism Marketing*, Vol. 37 No. 5, pp. 649-661.

Schuckert, M., Liu, X. and Law, R. (2015), "A segmentation of online reviews by language groups: how English and non-English speakers rate hotels differently", *International Journal of Hospitality Management*, Vol. 48, pp. 143-149.

Schuckert, M., Liu, X. and Law, R. (2016), "Stars, votes, and badges: how online badges affect hotel reviewers", *Journal of Travel and Tourism Marketing*, Vol. 33 No. 4, pp. 440-452.

Sezgen, E., Mason, K.J. and Mayer, R. (2019), "Voice of airline passenger: a text mining approach to understand customer satisfaction", *Journal of Air Transport Management*, Vol. 77, pp. 65-74.

Shi, Y. and Peng, Q. (2021), "Enhanced customer requirement classification for product design using big data and improved Kano model", *Advanced Engineering Informatics*, Vol. 49, p. 101340.

Stamolampros, P., Dousios, D., Korfiatis, N. and Symitsi, E. (2020), "The joint effect of consumer and service providers' culture on online service evaluations: a response surface analysis", *Tourism Management*, Vol. 78, p. 104057.

Sun, H., Guo, W., Shao, H. and Rong, B. (2020), "Dynamical mining of ever-changing user requirements: a product design and improvement perspective", *Advanced Engineering Informatics*, Vol. 46, p. 101174.

Timoshenko, A. and Hauser, J.R. (2019), "Identifying customer needs from user-generated content", *Marketing Science*, Vol. 38 No. 1, pp. 1-20.

Tontini, G., Bento, G.S., Milbratz, T.C., Volles, B.K. and Ferrari, D. (2017), "Exploring the nonlinear impact of critical incidents on customers' general evaluation of hospitality services", *International Journal of Hospitality Management*, Vol. 66, pp. 106-116.

Trappey, A.J., Trappey, C.V., Fan, C.Y. and Lee, I.J. (2018), "Consumer driven product technology function deployment using social media and patent mining", *Advanced Engineering Informatics*, Vol. 36, pp. 120-129.

Wang, W., Feng, Y. and Dai, W. (2018a), "Topic analysis of online reviews for two competitive products using latent Dirichlet allocation", *Electronic Commerce Research and Applications*, Vol. 29, pp. 142-156.

Wang, W.M., Li, Z., Tian, Z.G., Wang, J.W. and Cheng, M.N. (2018b), "Extracting and summarizing affective features and responses from online product descriptions and reviews: a Kansei text mining approach", *Engineering Applications of Artificial Intelligence*, Vol. 73, pp. 149-162.

Wang, Z., Chen, C.H., Zheng, P., Li, X. and Khoo, L.P. (2019), "A novel data-driven graph-based requirement elicitation framework in the smart product-service system context", *Advanced Engineering Informatics*, Vol. 42, p. 100983.

Wang, A., Zhang, Q., Zhao, S., Lu, X. and Peng, Z. (2020a), "A review-driven customer preference measurement model for product improvement: sentiment-based importance–performance analysis", *Information Systems and E-Business Management*, Vol. 18 No. 1, pp. 61-88.

Wang, X., Xu, Z. and Škare, M. (2020b), "A bibliometric analysis of Economic Research-Ekonomska Istra Zivanja (2007–2019)", *Economic Research-Ekonomska Istraživanja*, Vol. 33 No. 1, pp. 865-886.

Xiang, Z., Schwartz, Z., Gerdes, J.H. and Uysal, M. (2015), "What can big data and text analytics tell us about hotel guest experience and satisfaction?", *International Journal of Hospitality Management*, Vol. 44, pp. 120-130.

Xiao, S., Wei, C.P. and Dong, M. (2016), "Crowd intelligence: analyzing online product reviews for preference measurement", *Information and Management*, Vol. 53 No. 2, pp. 169-182.

Xu, X. (2018), "Does traveler satisfaction differ in various travel group compositions?", *International Journal of Contemporary Hospitality Management*, Vol. 30 No. 3, pp. 1663-1685.

Xu, X. (2019), "Examining the relevance of online customer textual reviews on hotels' product and service attributes", *Journal of Hospitality and Tourism Research*, Vol. 43 No. 1, pp. 141-163.

Xu, X. (2020a), "Examining an asymmetric effect between online customer reviews emphasis and overall satisfaction determinants", *Journal of Business Research*, Vol. 106, pp. 196-210.

Xu, X. (2020b), "How do consumers in the sharing economy value sharing? Evidence from online reviews", *Decision Support Systems*, Vol. 128, p. 113162.

Xu, X. and Li, Y. (2016), "The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: a text mining approach", *International Journal of Hospitality Management*, Vol. 55, pp. 57-69.

Xu, X., Wang, X., Li, Y. and Haghighi, M. (2017), "Business intelligence in online customer textual reviews: understanding consumer perceptions and influential factors", *International Journal of Information Management*, Vol. 37 No. 6, pp. 673-683.

Yadav, M.L. and Roychoudhury, B. (2019), "Effect of trip mode on opinion about hotel aspects: a social media analysis approach", *International Journal of Hospitality Management*, Vol. 80, pp. 155-165.

Yadegaridehkordi, E., Nilashi, M., Nizam Bin Md Nasir, M.H., Momtazi, S., Samad, S., Supriyanto, E. and Ghabban, F. (2021), "Customers segmentation in eco-friendly hotels using multi-criteria and machine learning techniques", *Technology in Society*, Vol. 65, p. 101528.

Yakubu, H. and Kwong, C.K. (2021), "Forecasting the importance of product attributes using online customer reviews and Google Trends", *Technological Forecasting and Social Change*, Vol. 171, p. 120983.

Ye, Q., Law, R. and Gu, B. (2009), "The impact of online user reviews on hotel room sales", *International Journal of Hospitality Management*, Vol. 28 No. 1, pp. 180-182.

Ye, Q., Li, H., Wang, Z. and Law, R. (2014), "The influence of hotel price on perceived service quality and value in E-Tourism", *Journal of Hospitality and Tourism Research*, Vol. 38 No. 1, pp. 23-39.

You, W., Xia, M., Liu, L. and Liu, D. (2012), "Customer knowledge discovery from online reviews", *Electronic Markets*, Vol. 22 No. 3, pp. 131-142.

Yu, L. and Wang, L. (2010), "Product portfolio identification with data mining based on multi-objective GA", *Journal of Intelligent Manufacturing*, Vol. 21 No. 6, pp. 797-810.

Yu, D., Xu, Z. and Wang, X. (2020), "Bibliometric analysis of support vector machines research trend: a case study in China", *International Journal of Machine Learning and Cybernetics*, Vol. 11 No. 3, pp. 715-728.

Zhan, J., Loh, H.T. and Liu, Y. (2009), "Gather customer concerns from online product reviews – a text summarization approach", *Expert Systems with Applications*, Vol. 36 No. 2, pp. 2107-2115.

Zhang, Z., Ye, Q., Law, R. and Li, Y. (2010), "The impact of e-word-of-mouth on the online popularity of restaurants: a comparison of consumer reviews and editor reviews", *International Journal of Hospitality Management*, Vol. 29 No. 4, pp. 694-700.

Zhang, W., Xu, H. and Wan, W. (2012), "Weakness finder: find product weakness from Chinese reviews by using aspects based sentiment analysis", *Expert Systems with Applications*, Vol. 39 No. 11, pp. 10283-10291.

Zhang, L., Chu, X. and Xue, D. (2019), "Identification of the to-be-improved product features based on online reviews for product redesign", *International Journal of Production Research*, Vol. 57 No. 8, pp. 2464-2479.

Zhang, C., Xu, Z., Gou, X. and Chen, S. (2021a), "An online reviews-driven method for the prioritization of improvements in hotel services", *Tourism Management*, Vol. 87, p. 104382.

Zhang, J., Lu, X. and Liu, D. (2021b), "Deriving customer preferences for hotels based on aspect-level sentiment analysis of online reviews", *Electronic Commerce Research and Applications*, Vol. 49, p. 101094.

Zhang, J., Zhang, A., Liu, D. and Bian, Y. (2021c), "Customer preferences extraction for air purifiers based on fine-grained sentiment analysis of online reviews", *Knowledge-Based Systems*, Vol. 228, p. 107259.

Zhao, Y., Xu, X. and Wang, M. (2019), "Predicting overall customer satisfaction: big data evidence from hotel online textual reviews", *International Journal of Hospitality Management*, Vol. 76, pp. 111-121.

Zheng, Lili. (2021), "The classification of online consumer reviews: a systematic literature review and integrative framework", *Journal of Business Research*, Vol. 135, pp. 226-251.

Zheng, L., He, Z. and He, S. (2020), "A novel probabilistic graphic model to detect product defects from social media data", *Decision Support Systems*, Vol. 137, p. 113369.

Zheng, L., He, Z. and He, S. (2021), "An integrated probabilistic graphic model and FMEA approach to identify product defects from social media data", *Expert Systems with Applications*, Vol. 178, p. 115030.

Zhou, L., Tang, L. and Zhang, Z. (2022), "Extracting and ranking product features in consumer reviews based on evidence theory", *Journal of Ambient Intelligence and Humanized Computing*. doi: 10.1007/s12652-021-03664-1.

About the authors

Yajun Wang is working toward the master degree within School of Business Administration, Northeastern University, China. Her research interests include data analysis, decision analysis and product improvement. She is a Member of the Decision Analysis Laboratory, Research Center for Science Education, Northeastern University at Qinhuangdao.

Xinyu Meng is working toward the master degree within School of Business Administration, Northeastern University, China. Her research interests include healthcare management and decision analysis. She is a Member of the Decision Analysis Laboratory, Research Center for Science Education, Northeastern University at Qinhuangdao.

Chang Xu received the bachelor of management from the School of Management, Hebei University, Baoding, China in 2021. She is currently pursuing the MS degree with the Business Administration School, Northeastern University, Shenyang, China. Her research interests are data analysis and product improvement.

Meng Zhao is an associate professor of School of Management, Northeastern University. She received her PhD degree in Management Science and Engineering from Beijing University of Aeronautics and Astronautics in 2011. She has published 2 monographs, and contributed over 40 journal articles to professional journals, such as Journal of the Operational Research Society, Fuzzy Optimization and Decision Making, Information Science, Computers & Industrial Engineering, Applied Soft Computing, and Group Decision and Negotiation Technological and Economic Development of Economy etc. Her current research interests include Decision analysis based on data mining and e-word-of-mouth analysis, Operations research, Multiple criteria decision analysis under uncertainty etc. Meng Zhao is the corresponding author and can be contacted at: ningmeng5072008@163.com