International Journal of Internet, Broadcasting and Communication Vol.16 No.3 281-289 (2024)
http://dx.doi.org/10.7236/IJIBC.2024.16.3.281

IJIBC 24-3-30

Detecting Fake Reviews: Exploring the Linguistic Characteristics by Computerized Text Analysis

Moon-Yong Kim

Professor, College of Business, Hankuk University of Foreign Studies, Seoul, Korea
E-mail: moonyong@hufs.ac.kr

Abstract

Online consumer reviews have become the most important basis for online shopping and product sales. Fake reviews are generated to boost sales because online consumer reviews play a vital role in consumers' decision making. The prevalence of fake reviews violates the regulations of the online business environment and misleads consumers in decision making. Thus, the present research investigates the effects of reviews' linguistic characteristics (i.e., analytical thinking, authenticity) on review fakeness. Specifically, this research examines whether (1) the level of analytical thinking is lower for fake (vs. genuine) reviews (hypothesis 1) and (2) the level of authenticity is lower for fake (vs. genuine) reviews (hypothesis 2). This research analyzed user-generated hotel reviews (genuine reviews, fake reviews) collected from MTurk. Linguistic Inquiry and Word Count (LIWC) 2022 was adopted to code review contents, and the hypotheses were tested using logistic regression. Consistent with the hypotheses 1 and 2, the results indicate that (1) analyticial thinking is negatively associated with review fakeness; and (2) authenticity is negatively associated with review fakeness. The findings provide important implications to identify fake reviews based on linguistic characteristics.

Keywords: *Fake Review, Genuine Review, Linguistic Text Analysis, Analytical Thinking, Authenticity*

1. Introduction

In recent years, as the e-commerce transaction model has been increasingly accepted by consumers, the success or failure of the business of online products and services is inseparable from the word-of-mouth from consumers. Previous studies have shown that online consumer reviews have become the most important basis for online shopping decisions and product sales [1, 2]. Based on the experience of prior consumers, online reviews significantly influence potential consumers in decision making [3]. For example, when selecting restaurants, consumers can learn about information, such as recommended dishes and dining environment from online reviews and then make purchase decisions.

Fake reviews are online reviews that are fabricated to boost sales and mislead consumers because online reviews play a vital role in consumers' decision making [4-7]. That is, fake reviews are generated to either promote one's business or damage the reputation of competitors [8]. The prevalence of fake reviews violates

Manuscript Received: June. 28, 2024 / Revised: July. 4, 2024 / Accepted: July. 11, 2024
Corresponding Author: moonyong@hufs.ac.kr (Moon-Yong Kim)
Tel: +82-2-2173-3157
Professor, College of Business, Hankuk University of Foreign Studies, Seoul, Korea

the regulations of the online business environment and misleads consumers in decision making [9]. Low credibility of online platforms and businesses will weaken consumers' continuance intentions and hinder the development of online businesses [10-12]. Thus, fake reviews must be identified in website and business management. Online platforms strive to prevent or detect fake reviews [9]. For example, the Yelp website developed algorithms to detect and filter fake reviews. Researchers also endeavored to develop optimal techniques to discern fake reviews.

Linguistic cues are widely used to distinguish fake reviews from genuine reviews because they are dissimilar in linguistic features [13]. Compared with the extensive efforts in optimizing linguistic analysis algorithms, the linguistic discrepancies based on psychological processes are under-researched [6]. Knowledge about the psychological mechanism underlying deception will complement the algorithm approach, and hence enhance the robustness of detection tools [6]. Therefore, this research aims to investigate the impact of linguistic characteristics (i.e., analytical thinking, authenticity) on review fakeness. Specifically, this research examines whether (1) the level of analytical thinking is lower for fake (vs. genuine) reviews (hypothesis 1); and (2) the level of authenticity is lower for fake (vs. genuine) reviews (hypothesis 2).

2. Theoretical Background and Hypotheses Development

2.1 Online Consumer Reviews and Computerized Text Analysis

Computational linguistics attempts to provide an efficient and effective method for studying the emotional, cognitive, and structural components present in a person's writing [14]. In fact, computational linguistic has been applied into different studies for understanding people's psychological activities, because the functional and emotional words that people use provide important psychological implications for their thinking processes, emotional states, intentions and motivations [15]. Based on the writing text or speaking words, people's underlying state of mind can be inferred and thus their personal characteristics can be summarized. Among the most commonly used computational linguistics software, Linguistic Inquiry and Word Count (LIWC) has been shown to be a useful tool in online review research. LIWC is a text analysis software that calculates the frequency and proportionality of words falling into specific linguistic categories [16]. Compared to other computerized text analysis tools, LIWC stands out for its user-friendly nature and affordable cost, making it appealing to service firms looking to analyze customer feedback [17]. It utilizes a language processing component and over 100 built-in dictionaries of words to analyze text data [15]. Each dictionary comprises a collection of words that define a particular category. LIWC compares each word in text data to the predetermined list of words in the LIWC dictionaries and counts the number of words falling into each category [18]. This information will be presented as the percentage of words within a specific text [17]. The categorization of the words provides insights into linguistic patterns (e.g., use of pronouns and emotional words) and psychological states. For example, the positive and negative emotion indicators in LIWC were also adopted for the review helpfulness forecast [19]. Additionally, sentiment scores from LIWC were employed in understanding the relationship between sentimental tone and review helpfulness [20].

Recently, computerized text analysis tools such as LIWC, IBM Watson, DICTION and T-lab, have gained popularity as powerful research instruments with which to conduct customer research [21]. Some of these tools are AI-based, which means it requires a large volume of text in order to complete an effective analysis [22]. A further type of computerized text analysis relies on the use of built-in dictionaries, where a piece of text is analyzed by comparing words contained in the text with those in a particular predefined dictionary [23]. Dictionary-based text analysis tools are effective with smaller datasets, such as online customer reviews, where each review typically comprises only a small number of words [24]. Making use of a computerized text analysis tool with predefined dictionaries, allows for the creation of specific summary variables designed to identify psychological states, thinking styles, emotional tone and social concerns [14]. Additionally, it is suitable for the analysis of multilevel language use, as per SAT, offering service firms an alternative tool with which to measure and predict customers' evaluation of their interactions with a firm, using verbatim online reviews.

2.2 Analytical Thinking

Analytical thinking is representative of the extent to which individuals use words that suggest the use of formal and hierarchical thinking patterns [25]. Higher analytical thinking is associated with more formal and logical text. People low in analytical thinking tend to write and think using language in a more narrative, informal, and personal ways. For example, analytical thinking is most apparent in an individual's use of articles which typically signals concepts and prepositions which conveys the relationships between the concepts [26]. Individuals naturally differ in the extent to which they engage in analytical thinking, which is typically formalized and associated with more formal settings, as opposed to intuitive thinking.

Research on automated linguistic analysis has highlighted that deceiving individuals lack the support of real experiences and memory, so they tend to communicate in a language that lacks complexity, detail and omits specific, analytical language [27]. The same study on deceptive language emphasized that deceptive senders use more informality in their messages than their respective receivers, including more typographical errors in written messages [27]. Deceivers use less analytical information and less clear and complete messages to manipulate content, not adding much detail and relevant information to their responses [28, 29]. There are significant differences in authentic communication, analytical writing style and text formalism as a function of the type of review [30]. Accordingly,

H1. The level of analytical thinking is lower for fake (vs. genuine) reviews.

2.3 Authenticity

Authenticity is primarily concerned with credibility and trustworthiness. Communication research shows that when people reveal themselves in an authentic or honest way, they are more personal, disclosed, and vulnerable [14]. Their writings are likely to establish a relationship between themselves and their stories, to take responsibilities for their behavior. Conversely, people involved in deceptive communication tends to distant themselves by using fewer first and third person singular pronouns, more negative words, etc. [31]. People with an authentic writing style would be considered to have high credibility and be generally trusted by others [32, 33]. The summary variable for authenticity was developed following the prior research that investigated the linguistic features that differentiate true and false stories [31]. Their research identified that those misrepresenting the truth (i.e., being inauthentic) typically showed a lower cognitive complexity, used less self-reference and made use of more words presenting negative motions. Accordingly,

H2. The level of authenticity is lower for fake (vs. genuine) reviews.

3. Method

3.1 Data Collection

An online survey was conducted, where actual hotel service consumers were asked to create both genuine and fake reviews under the same conditions. To conduct this user-generated data collection, a general population sample is employed from a common crowdsource platform, Amazon Mechanical Turk (MTurk) [34]. Because of low cost and convenient access, human subject samples collected using online surveys such as MTurk have been widely used in research for social sciences [35]. Yet, MTurk and other crowdsourcing websites have raised concerns about data quality. For instance, researchers have expressed concerns that MTurk participants self-select into experiments, which is not random sampling. Some of these participants participate in experiments too frequently, which makes them too familiar with common experimental paradigms [36]. Nevertheless, there are a lot of studies showing that MTurk data are generally comparable to data collected in traditional controlled labs if care is taken to screen the participants properly [37, 38].

Using the survey on MTurk, 400 participants were obtained. The final sample (N = 400) was composed of 215 females (53.8%) and 185 males (46.3%) who ranged in age from 18 to 75 years (mean = 35.65, SD = 11.04). In terms of ethnicity, 68.5% of the sample Caucasian origin, 16.3% Asian, 7.0% Hispanic, 5.5% African American, and 2.8% Others. The majority of the respondents had a college or university degree (75.9%), and 8.3% of the respondents had a high school graducation only (7.8%) and less than high school graduation (0.5%), and 16.0% had a postgraduate degree. Regarding the annual household income, 17.5% of the respondents reported income of less than $25,000; 29.0% fell within an income range of $25,000 to $49,999; 22.5% were in the $50,000 to $74,999 range; 15.0% were in the $75,000 to $99,999 range; 8.0% were in the $100,000 to $124,999 range; 3.8% were in the $125,000 to $149,999 range; 1.8% were in the $150,000 to $174,999 range; and 2.6% reported income of more than $175,000.

3.2 Procedure and Measures

Following the procedure adopted by previous research [39], the survey in this research presented both review writing requests, asking a participant to write a genuine positive review as well as a fake positive review. This important and energy-consuming task was placed early in the survey to collect high-quality reviews. In these tasks, this research focused on the hotel business, given that it has been widely used in detecting deceptive opinion spam in both computational linguistics and various text-mining applications in marketing [39].

Review fakeness (the dependent variable) was operationalized through a binary division into genuine and fake reviews. Following the previous research [8], genuine reviews were coded as 0 and fake reviews as 1. In this research, indicators revealing fake reviews are identified: analytical thinking and authenticity. To identify analytical thinking and authenticity, this research adopted the Linguistic Inquiry and Word Count (LIWC) 2022 because it is a widely used software system for text analysis and considered a reliable measure of the linguistic cues [14, 40, 41]. Analytical thinking is measured in LIWC by identifying formal, logical and hierarchical thinking patterns in text based on function words (e.g., pronouns, articles) and grammar words (e.g., prepositions, conjunctions) [13, 25]. Authenticity analyzes whether individuals communicate honestly, based on research showing that consumers are more personal, disclosed and vulnerable when authentic [13]. Thus, authenticity in LIWC detects whether people express in an authentic or honest way [31].

3.3 Data Analysis

Unstructured online review text data is widely recognized as a rich source of information that reflects consumers' evaluations of products or services [42, 43]. The seemingly overwhelming volume of online data poses a challenge for firms to read them one by one, yet with the help of computerized text analysis tools, firms can quantify and predict consumer behavior [17]. Given that prior studies have called for more unstructured textual content and semantic analysis of online reviews [44, 45], this research performed a content and semantic analysis using LIWC-22 to convert user-generated content into numerical scales, based on existing psychometrically tested scales and algorithms that include analytical thinking [25] and authenticity [31]. LIWC, one of the computerized text analyses, can predict consumers' evaluations of their customer-firm interactions even without a quantitative measurement, for example, a star-rating system. This tool utilizes a text analysis module and pre-developed dictionaries to count words and determine their distributions along various linguistic pathways, such as emotions and cognitive processes [15]. The reliability and validity of LIWC in tourism and hospitality research has been tested and confirmed by numerous studies [40, 46]. LIWC was chosen on the basis that it offers its users a user-friendly and inexpensive means through which to analyze unstructured data, thereby aligning with the research purpose. There is strong empirical evidence to support the use of LIWC to discover meaningful linguistic insights relating to social relationships, emotionality, thinking styles and individual-level differences, based on natural language use [47]. In addition, to test the hypotheses 1 and 2, both one-way ANOVA and binary logistic regression analysis were performed.

4. Results

This research adopted binary logistic regression to analyze the collected data because the dependent variable is dichotomous. Considering that logistic regression is sensitive to multicollinearity issues, the correlation and multicollinearity tests were conducted in advance. Results show that all correlation coefficients are far below the threshold value of 0.8 (the highest value is the correlation between authenticity and review length (word count), which records 0.178) and the VIF scores are lower than the threshold value of 5 (the highest value is review length, which records 1.059). Thus, multicollinearity is not a concern for the logistic regression analysis.

Prior to the logistic regression analysis, a one-way ANOVA was conducted to examine whether the means of predictors significantly differ between genuine and fake reviews. Table 1 shows that compared with genuine reviews, fake reviews have (1) fewer words ($M_{genuine}$ = 72.61, SD = 48.33, M_{fake} = 57.91, SD = 38.72; F(1, 798) = 22.55, p = .000), (2) lower analytical thinking ($M_{genuine}$ = 56.69, SD = 26.23, M_{fake} = 47.21, SD = 28.57; F(1, 796) = 23.87, p = .000), and (3) lower authenticity ($M_{genuine}$ = 62.43, SD = 31.28, M_{fake} = 56.73, SD = 33.77; F(1, 790) = 6.090, p = .014).

Table 2 shows the results of the logistic regression analysis. First, the control variable (review length: the number of the words in the review) was entered into the Model 1 based on previous findings. Model 2 added the analytical thinking and authenticity. Nagelkerke R^2 increased from 0.034 in Model 1 to 0.068 in Model 2, and Cox and Snell R^2 increased from 0.025 in Model 1 to 0.051 in Model 2. The increased R^2 suggests that Model 2 has better goodness of fit than Model 1. Specifically, in terms of goodness of fit, Model 2 explains 6.8% of the total variance (Nagelkerke R^2 = .068) and correctly classified 61.6% of the sample respondents. Finally, the log-likelihood test, 2logλ = 1055.022 is significant at the .001 level. Consequently, the results indicate that the binary regression model has a good fit to the data. As shown in Model 2 in Table 2, fake (vs. genuine) reviews are found to be significantly lower in analytical thinking (Wald χ^2 = 18.08; p = .000; Exp(B) = 0.989) and authenticity (Wald χ^2 = 4.748; p = .029; Exp(B) = 0.995). Hence, H1 and H2 are supported, suggesting that both analytical thinking and authenticity are negatively associated with review fakeness.

Table 1. Results of one-way ANOVA

	Genuine Review		Fake Review		F-value	p-value
	Mean	SD	Mean	SD		
(1) Review length	72.61	48.33	57.91	38.72	22.55	.000
(2) Analytical thinking	56.69	26.23	47.21	28.57	23.87	.000
(3) Authenticity	62.43	31.28	56.73	33.77	6.090	.014

Table 2. Results of binary logistic regression

	Model 1					Model 2				
	B	SE	Wald	Sig.	Exp(B)	B	SE	Wald	Sig.	Exp(B)
Control var.										
Review length	-.008	.002	18.057	.000	.992	-.006	.002	10.702	.001	.994
Independent var.										
Analytical Thinking (H1)						-.012	.003	18.08	.000	.989
Authenticity (H2)						-.005	.002	4.748	.029	.995

5. Conclusion

The present research investigates the impact of linguistic characteristics (i.e., analytical thinking, authenticity) on review fakeness. Specifically, this research examines whether (1) the level of analytical thinking is lower for fake (vs. genuine) reviews (hypothesis 1); and (2) the level of authenticity is lower for fake (vs. genuine) reviews (hypothesis 2). This research analyzed user-generated hotel reviews (genuine reviews, fake reviews) collected from MTurk. This research adopted LIWC-22 to code review contents, and the hypotheses were tested using logistic regression. In support of the hypotheses 1 and 2, the results indicate that (1) analytical thinking is negatively associated with review fakeness; and (2) authenticity is negatively associated with review fakeness. The findings provide important implications to identify fake reviews based on linguistic characteristics.

Specifically, from the theoretical perspective, this research contributes to fake online review research in the fields of hospitality and tourism. This research is built on psychology theory about linguistic discrepancies between fake and genuine reviews. From the perspective of the psychological process, this research provides empirical evidence on the importance of considering linguistic characteristics in fake reviews. From the practical perspective, the research findings provide important managerial implications to identify fake reviews based on linguistic characteristics. For online review websites, especially those without a sophisticated algorithm for detecting fake reviews, the managers can improve the detection ability by analyzing the linguistic characteristics of reviews. Based on the research findings, online review websites can identify the reviews with a low level of analytical thinking and authenticity as the potential fake reviews. This research made use of LIWC software. Although LIWC analyzes text at the word level and might misinterpret the word's contextual meaning, it can efficiently analyze massive volumes of unstructured data available online and generate quantitative insights into linguistic patterns. Besides, LIWC is suitable in the service marketing industry due to its price accessibility, time efficiency, and ease of use requiring no coding or development skills. Thus, managers can use LIWC to identify the potential fake reviews and then report to the review websites. In practice, hoteliers and restaurant operators can establish an automated process that could perform the content analysis of online reviews and use a review fakeness checker, usually integrated into the browser, which calculates the score of analytical thinking and authenticity. Through this, hoteliers and restaurant operators can reduce the damage to their reputation due to fake comments. Regarding the consumers, they may not be willing to spend time and effort in identifying fake reviews, but they should be wary of reviews with a low level of analytical thinking and authenticity if they want to make a proper purchase decision.

Several limitations in this research should be noted. First, the samples were restricted to hotels. Future research should be replicated with reviews about other types of hospitality/tourism sectors, products or services to generalize the findings. Second, there is also potential interest in examining this topic on different online platforms (e.g., retailers, etailers, review aggregators, social media platforms, etc.). Different online platforms attract diverse demographics and might have unique linguistic features, which provides different perspectives and insights. Future research should broaden the scope of the research by exploring many different platforms and gathering data from additional sources in order to understand how linguistic characteristics influence review fakeness. Third, while this research primarily focused on cognitive aspects of linguistic characteristics in the context of positive reviews, future studies should consider the dynamics of various linguistic characteristics including a variety of cognitive and emotional cues and examine the role of review valence, in response to the call for combining linguistic cues to identify review fakeness [13].

Acknowledgement

This work was supported by Hankuk University of Foreign Studies Research Fund of 2023.

References

[1] J. Chevalier and D. Mayzlin, "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, Vol. 43, No. 3, pp. 345-354, August 2006. DOI: https://doi.org/10.1509/jmkr.43.3.345

[2] Y. Liu, "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing*, Vol. 70, No. 3, pp. 74-89, July 2006. DOI: https://doi.org/10.1509/jmkg.70.3.074

[3] L. Kwok, K.L. Xie, and T. Richards, "Thematic Framework of Online Review Research," *International Journal of Contemporary Hospitality Management*, Vol. 29, No. 1, pp. 307-354, January 2017. DOI: https://doi.org/10.1108/IJCHM-11-2015-0664

[4] I.C.C. Chan, L.W. Lam, C.W. Chow, L.H.N. Fong, and R. Law, "The Effect of Online Reviews on Hotel Booking Intention: The Role of Reader-Reviewer Similarity," *International Journal of Hospitality Management*, Vol. 66, pp. 54-65, September 2017. DOI: https://doi.org/10.1016/j.ijhm.2017.06.007

[5] M. Schuckert, X. Liu, and R. Law, "Hospitality and Tourism Online Reviews: Recent Trends and Future Directions," *Journal of Travel and Tourism Marketing*, Vol. 32, No. 5, pp. 608-621, August 2015. DOI: https://doi.org/10.1080/10548408.2014.933154

[6] Y. Wu, E.W.T. Ngai, P. Wu, and C. Wu, "Fake Online Reviews: Literature Review, Synthesis and Directions for Future Research," *Decision Support Systems*, Vol. 132, 113280, May 2020. DOI: https://doi.org/10.1016/j.dss.2020.113280

[7] Y. Xu, Z. Zhang, R. Law, and Z. Zhang, "Effects of Online Reviews and Managerial Responses from a Review Manipulation Perspective," *Current Issues in Tourism*, Vol. 23, No. 17, pp. 2207-2222, 2020. DOI: https://doi.org/10.1080/13683500.2019.1626814

[8] L. Li, K.Y. Lee, M. Lee, and S.B. Yang, "Unveiling the Cloak of Deviance: Linguistic Cues for Psychological Processes in Fake Online Reviews," *International Journal of Hospitality Management*, Vol. 87, 102468, May 2020. DOI: https://doi.org/10.1016/j.ijhm.2020.102468

[9] M. Luca and G. Zervas, "Fake It Till You Make It: Reputation, Competition and Yelp Review Fraud," *Management Science*, Vol. 62, No. 12, pp. 3412-3427, January 2016. DOI: https://doi.org/10.1287/mnsc.2015.2304

[10] U. Chakraborty, "Perceived Credibility of Online Hotel Reviews and Its Impact on Hotel Booking Intentions," *International Journal of Contemporary Hospitality Management*, Vol. 31, No. 9, pp. 3465-3483, August 2019. DOI: https://doi.org/10.1108/IJCHM-11-2018-0928

[11] R. Filieri, F. Acikgoz, V. Ndou, and Y. Dwivedi, "Is TripAdvisor Still Relevant? The Influence of Review Credibility, Review Usefulness and Ease of Use on Consumers' Continuance Intention," *International Journal of Contemporary Hospitality Management*, Vol. 33, No. 1, pp. 199-223, January 2021. DOI: https://doi.org/10.1108/IJCHM-05-2020-0402

[12] R. Law, D. Leung, and I.C.C. Chan, "Progression and Development of Information and Communication Technology Research in Hospitality and Tourism: A State-of-the-Art Review," *International Journal of Contemporary Hospitality Management*, Vol. 32, No. 2, pp. 511-534, February 2020.

DOI: https://doi.org/10.1108/IJCHM-07-2018-0586
[13] D. Plotkina, A. Munzel, and J. Pallud, "Illusions of Truth—Experimental Insights into Human and Algorithmic Detections of Fake Online Reviews," *Journal of Business Research*, Vol. 109, pp. 511-523, March 2020. DOI: https://doi.org/10.1016/j.jbusres.2018.12.009
[14] J.W. Pennebaker, R. Boyd, K. Jordan, and K. Blackburn, *The Development and Psychometric Properties of LIWC 2015*, 2015.
[15] Y.R. Tausczik and J.W. Pennebaker, "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods," *Journal of Language and Social Psychology*, Vol. 29, No. 1, pp. 24-54, 2010. DOI: https://doi.org/10.1177/0261927X09351676
[16] J.W. Pennebaker, M.R. Mehl, and K.G. Niederhoffer, "Psychological Aspects of Natural Language Use: Our Words, Our Selves," *Annual Review of Psychology*, Vol. 54, pp. 547-577, 2003. DOI: 10.1146/annurev.psych.54.101601.145041
[17] C. Ferreira, J. Robertson, R. Chohan, L. Pitt, and T. Foster, "The Writing is on the Wall: Predicting Customers' Evaluation of Customer-Firm Interactions Using Computerized Text Analysis," *Journal of Service Theory and Practice*, Vol. 33, No. 2, pp. 309-327, February 2023. DOI: https://doi.org/10.1108/JSTP-04-2022-0100
[18] R.L. Boyd, A. Ashokkumar, S. Seraj, and J.W. Pennebaker, *The Development and Psychometric Properties of LIWC-22*, University of Texas at Austin, 2022.
[19] T. Ngo-Ye, A. Sinha, and A. Sen, "Predicting the Helpfulness of Online Reviews Using a Scripts-Enriched Text Regression Model," *Expert Systems with Applications*, Vol. 71, pp. 98-110, April 2017. DOI: https://doi.org/10.1016/j.eswa.2016.11.029
[20] A. Agnihotri and S. Bhattacharya, "Online Review Helpfulness: Role of Qualitative Factors," *Psychology & Marketing*, Vol. 33, No. 11, pp. 1006-1017, October 2016. DOI: https://doi.org/10.1002/mar.20934
[21] A. Humphreys and R. Wang, "Automated Text Analysis for Consumer Research," *Journal of Consumer Research*, Vol. 44, No. 6, pp. 1274-1306, April 2018. DOI: https://doi.org/10.1093/jcr/ucx104
[22] B. Balducci and D. Marinova, "Unstructured Data in Marketing", *Journal of the Academy of Marketing Science*, Vol. 46, pp. 557-590, June 2018. DOI: https://doi.org/10.1007/s11747-018-0581-x
[23] C. Pitt, C. M. Mulvey, and J. Kietzmann, "Quantitative Insights from Online Qualitative Data: An Example from the Healthcare Sector," *Psychology & Marketing*, Vol. 35, No. 12, pp. 1010-1017, November 2018. DOI: https://doi.org/10.1002/mar.21152
[24] J. Robertson, C. Ferreira, and J. Paschen, "Reading Between the Lines: Understanding Customer Experience with Disruptive Technology through Online Reviews," *Australasian Marketing Journal*, Vol. 29, No. 3, pp. 215-224, March 2021. DOI: https://doi.org/10.1177/1839334921999487
[25] J.W. Pennebaker, C.K. Chung, J. Frazee, G.M. Lavergne, and D.I. Beaver, "When Small Words Foretell Academic Success: The Case of College Admissions Essays," *PloS One*, Vol. 9, No. 12, 110, December 2014. DOI: https://doi.org/10.1371/journal.pone.0115844
[26] K. Jordan, J. Sterling, J.W. Pennebaker, and R. Boyd, "Examining Long-Term Trends in Politics and Culture through Language of Political Leaders and Cultural Institutions," *Psychological and Cognitive Sciences*, Vol. 116, No. 9, pp. 3476-3481, February 2019. DOI: 10.1073/pnas.1811987116
[27] U. Zhou, J.K. Burgoon, D.P. Twitchell, T. Qin, and J.F. Jr. Nunamaker, "A Comparison of Classification Methods for Predicting Deception in Computer-Mediated Communication," *Journal of Management Information Systems*, Vol. 20, No. 4, pp. 139-165, 2004. DOI: https://doi.org/10.1080/07421222.2004.11045779
[28] J.K. Burgoon, D.B. Buller, A.S. Ebesu, C.H. White, and P.A. Rockwell, "Testing Interpersonal Deception Theory: Effects of Suspicion on Communication Behaviors and Perceptions," *Communication Theory*, Vol. 6, No. 3, pp. 243-267, August 1996. DOI: https://doi.org/10.1111/j.1468-2885.1996.tb00128.x
[29] J.R. Carlson, J.F. George, J.K. Burgoon, M. Adkins, and C.H. White, "Deception in Computer-Mediated Communication," *Group Decision and Negotiation*, Vol. 13, No. 1, pp. 5-28, January 2004. DOI: https://doi.org/10.1023/B:GRUP.0000011942.31158.d8
[30] U. Chakraborty and S. Bhat, "The Effects of Credible Online Reviews on Brand Equity Dimensions and Its Consequence on Consumer Behavior," *Journal of Promotion Management*, Vol. 24, No. 1, pp. 57-82, 2018.

DOI: https://doi.org/10.1080/10496491.2017.1346541
[31] M.L. Newman, J.W. Pennebaker, D.S. Berry, and J.M. Richards, "Lying Words: Predicting Deception from Linguistic Style," *Personality and Social Psychology Bulletin*, Vol. 29, No. 5, pp. 665-675, May 2003. DOI: https://doi.org/10.1177/0146167203029005010
[32] S. Chaiken, A. Liberman, and A.H. Eagly, "Heuristic and Systematic Information Processing Within and Beyond the Persuasion Context," *Unintended Thought*, Vol. 212, pp. 212-252, 1989.
[33] S. Chen and S. Chaiken, "The Heuristic-Systematic Model in Its Broader Context," *Dual-Process Theories in Social Psychology*, Vol. 15, pp. 73-96, 1999.
[34] D. Fernandes, "The 1/N Rule Revisited: Heterogeneity in the Naïve Diversification Bias," *International Journal of Research in Marketing*, Vol. 30, No. 3, pp. 310-313, September 2013. DOI: https://doi.org/10.1016/j.ijresmar.2013.04.001
[35] S. Yang and M. Lynn, "More Evidence Challenging the Robustness and Usefulness of the Attraction Effect," *Journal of Marketing Research*, Vol. 51, No. 4, pp. 508-513, 2014. DOI: https://doi.org/10.1509/jmr.14.0020
[36] J. Chandler, P. Mueller, and G. Paolacci, "Nonnaïveté among Amazon Mechanical Turk Workers: Consequences and Solutions for Behavioral Researchers," *Behavior Research Methods*, Vol. 46, pp. 112-130, 2014. DOI: https://doi.org/10.3758/s13428-013-0365-7
[37] J. Kees, C. Berry, S. Burton, and K. Sheehan, "An Analysis of Data Quality: Professional Panels, Student Subject Pools, and Amazon's Mechanical Turk," *Journal of Advertising*, Vol. 46, No. 1, pp. 141–155, 2017. DOI: https://doi.org/10.1080/00913367.2016.1269304
[38] S.M. Smith, M. Scott, C.A. Roster, L.L. Golden, and G.S. Albaum, "A Multi-Group Analysis of Online Survey Respondent Data Quality: Comparing a Regular USA Consumer Panel to MTurk Samples," *Journal of Business Research*, Vol. 69, No. 8, pp. 3139-3148, August 2016. DOI: https://doi.org/10.1016/j.jbusres.2015.12.002
[39] S. Moon, M-Y. Kim, and D. Iacobucci, "Content Analysis of Fake Consumer Reviews by Survey-Based Text Categorization," *International Journal of Research in Marketing*, Vol. 38, No. 2, pp. 343-364, June 2021. DOI: https://doi.org/10.1016/j.ijresmar.2020.08.001
[40] M. Lee, S.A. Lee, and Y. Koh, "Multisensory Experience for Enhancing Hotel Guest Experience: Empirical Evidence from Big Data Analytics," *International Journal of Contemporary Hospitality Management*, Vol. 31, No. 11, pp. 4313-4337, July 2019. DOI: https://doi.org/10.1108/IJCHM-03-2018-0263
[41] S.W. Litvin and R.N. Sobel, "Organic versus Solicited Hotel Tripadvisor Reviews: Measuring Their Respective Characteristics," *Cornell Hospitality Quarterly*, Vol. 60, No. 4, pp. 370-377, 2019. DOI: https://doi.org/10.1177/19389655188112
[42] Z. Xiang, Q. Du, Y. Ma, and W. Fan, "A Comparative Analysis of Major Online Review Platforms: Implications for Social Media Analytics in Hospitality and Tourism," *Tourism Management*, Vol. 58, pp. 51-65, February 2017. DOI: https://doi.org/10.1016/j.tourman.2016.10.001
[43] J. Zhan, H.T. Loh, and Y. Liu, "Gather Customer Concerns from Online Product Reviews—A Text Summarization Approach," *Expert Systems with Applications*, Vol. 36, No. 2, pp. 2107-2115, March 2009. DOI: https://doi.org/10.1016/j.eswa.2007.12.039
[44] J. Büschken and G.M. Allenby, "Sentence-Based Text Analysis for Customer Reviews," *Marketing Science*, Vol. 35, No. 6, pp. 953-975, July 2016. DOI: https://doi.org/10.1287/mksc.2016.0993
[45] S. Tirunillai and G.J. Tellis, "Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation," *Journal of Marketing Research*, Vol. 51, No. 4, pp. 463-479, 2014. DOI: https://doi.org/10.1509/jmr.12.0106
[46] M. Yu, M. Cheng, L. Yang, and Z. Yu, "Hotel Guest Satisfaction during COVID-19 Outbreak: The Moderating Role of Crisis Response Strategy," *Tourism Management*, Vol. 93, 104618, December 2022. DOI: https://doi.org/10.1016/j.tourman.2022.104618
[47] V. Manchaiah, D. Swanepoel, and R. Bennett, "Online Consumer Reviews on Hearing Health Care Services: A Textual Analysis Approach to Examine Psychologically Meaningful Language Dimensions," *American Journal of Audiology*, Vol. 30, No. 3, pp. 669-675, September 2021. DOI: https://doi.org/10.1044/2021_AJA-20-00223