Usability and User Experience, Vol. 39, 2022, 175–184
https://doi.org/10.54941/ahfe1001706

Understanding the User Experience of Battery Electric Vehicles: A Perspective Based on Big Data Text Mining Techniques

Quan Gu, Shengqing Huang, Zhang Jie, Ying Zhang, and Yue Cui

School of Art Design and Media, East China University of Science and Technology, Shanghai 200237, China

ABSTRACT

Battery electric vehicle (BEV) is the core innovation of low-carbon travel transformation, but there are still few evaluation studies on user experience. To more accurately understand the relatively real user experience of BEV, this paper uses text mining and natural language processing based on the big data text of BEV user experience and proposes a method for collecting, drawing, and analyzing these user experiences. Through this method, the user experience of the real scene can be restored to some extent. The content includes the following: First, obtain user comments on the typical BEV Model 3 on the online review website through crawler software, and use natural language processing technology to pre-process the data; secondly, based on the constructed stop word database, the text in the text Stop words are eliminated; then, the number of common occurrences between two adjacent words is counted, and a co-occurrence matrix is generated; finally, word frequency statistics, improved TFIDF keyword extraction, etc., are performed, and word frequency statistics are visualized. Keyword word cloud visualization, centrality analysis visualization, multi-scale keyword analysis visualization, etc., Compared with the traditional research focusing on individual user experience, this research explores the possibility of a research method of user experience evaluation in the context of big data. This will provide a certain theoretical reference for the research of the user experience evaluation system, and help the product user experience team of related BEV to get closer to the truth to understand the user's experience scenarios, behaviors, and real feelings.

Keywords: Big data, User experience, User experience visualization, Battery electric vehicle, Data visualization

INTRODUCTION

Developments in Electric Vehicles

With the EU's enacted goal of achieving carbon neutrality by mid-century, the development of clean energy is one of the key ways to achieve its goal, and all sectors are working towards new energy. It is inevitable that the automotive industry, as one of the main players, will vigorously develop new energy vehicles. 2.65 million new electric vehicles were sold worldwide in the first half of 2021, an increase of 68% compared to 2020, a proportion that illustrates

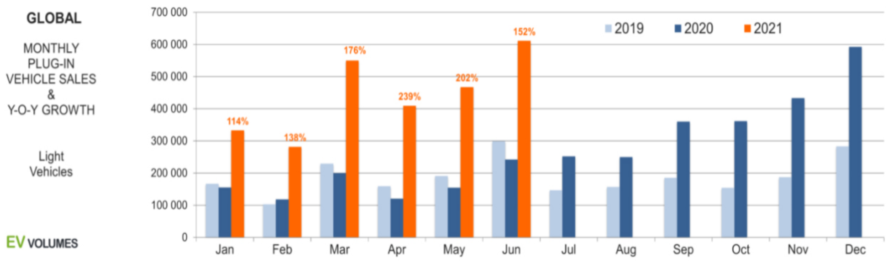

Figure 1: Global monthly sales growth of plug-in vehicles in light vehicles year-on-year.

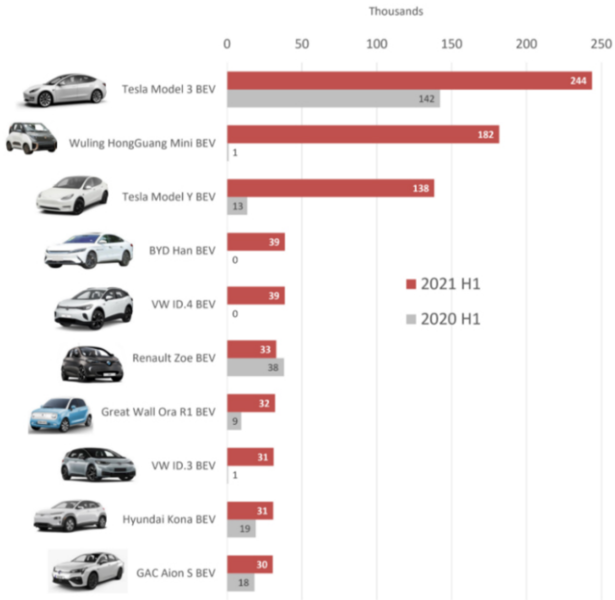

Figure 2: Electric vehicles - global delivery comparison (H1 2021 and H1 2020).

the high rate of growth. The share of BEV+PHEV in global light vehicle sales increased from 3% in the first half of 2020 to 6.3% this year (see Figure 1). The top 10 best-selling electric vehicles are now all BEVs, with Tesla holding the number one spot (see Figure 2). Judging from the available data and related policies, the scale of new energy vehicles will continue to expand.

Current Status of Research

Global electric vehicles have been industrialized since 2010, it has maintained a high growth rate for more than a decade, and a number of enterprises such as Tesla, Wuling Hongguang and Weilai have emerged. But compared with the development of traditional fuel vehicles, the time is after all relatively short. The new energy vehicle market has been constantly transformed and upgraded, with rapid product iterations, and in addition to technological breakthroughs, the user needs and experiences have changed essentially. In addition to technological breakthroughs, user needs and experiences have changed in essence and need further study. User needs and experiences have always been a kind of "fuzzy" existence because of the emotional experiences of people, and are difficult to quantify scientifically. However, user feelings and emotions are of great significance in user experience, and enhancing user services and experience is of great value in further improving user experience in the manufacturing industry.

The current literature review on text analysis focuses on the mining of short texts, with three main research methods: first, text analysis based on word frequency statistics; second, comparative research on text similarity based on clustering algorithms; and third, text analysis based on semantics. The research on text visualisation of product reviews is mainly focused on the visual analysis of emotions in product reviews and the extraction of product features in product reviews, while the visualisation of the empirical knowledge embodied in product reviews is currently lacking. Questions and answers in Zhihu are a form of long text that expresses the user's experience of using the product. By analyzing the text, it is possible to explore more of the actual user experience and restore the real user scenario.

User Experience and its Visualization

Initially, UX design research focused on the psychological exploration of groups and individuals in order to obtain information and explore needs from the user's point of view, but in UX it is not only the exploration of needs that needs to be achieved, but also the transformation of functions. Therefore, the accuracy of user experience analysis is particularly important. Traditional user experience research is mainly based on qualitative research, but at present, quantitative research on user needs is relatively mature, and quantitative research on user experience is rarely mentioned. It is particularly important for product manufacturers, product designers and potential purchasers to be able to read the experiences of users of products more effectively. Visualization techniques using big data are a viable approach. Traditional non-visualization techniques require specialist statisticians to extract and summaries data in advance and cannot be dynamically adjusted in real time, which is time-consuming and costly. Text visualization of commodity reviews is a method of discovering specific information through the analysis of commodity review texts and presenting it graphically using computer technology, which is one of the important branches of information visualization. From the perspective of data visualization, information is conveyed more simply with the help of software processing to facilitate analysis by researchers. Therefore, this paper takes the Tesla model 3 as an example, and carries out big data text mining, processing and analysis on it to establish a real user experience scenario and provide a different approach from the traditional user experience research.

RESEARCH DESIGN AND METHODOLOGY

Research Tools

Data mining, also known as Knowledge Discover in Database (KDD), is the process of extracting potentially useful information and knowledge from a large amount of incomplete, noisy, fuzzy or even random application data, which is not known beforehand. Firstly, data mining is carried out by crawling the comments of model 3 from ZhiHu, which is a step before the data is studied. Due to the unstructured nature of the 'comments' section, data collection and pre-processing steps are required before the appropriate techniques can be used. In the current work, the convenient and easy-to-use Octopus software (https://www.bazhuayu.com) was used to crawl the comments.

Unstructured text document analysis is often thought of as text mining, a crossover area between search engines and data mining. The purpose of text mining is to mine the implicit useful knowledge from a large number of document knowledge sources such as journals, publications, reports, Web pages, etc. The goal of this paper is to mine the text of Model 3 in Web pages. In order to improve the accuracy and processing efficiency of the mining results, it is necessary to find the way to obtain and use the semantic knowledge of the text processing, and the co-occurrence analysis method can provide semantic support. In general text data mining includes the following steps: Step 1: data collection; Step 2: data screening; Step 3: data extraction; Step 4: data analysis; Step 5: data visualization. Therefore, a software called KH Coder, developed by Japanese scholar Koichi Higuchi (2014), is used for text data mining (Text Mining), which not only collates text information, but also performs word frequency and lexical analysis, automatic clustering and generates co-occurrence relationship graphs.

Research Methodology

This paper considers that the idea of co-occurrence analysis for text information mining is based on internal and external association of text information mining. Co-occurrence analysis mines text information by mining author information in text content, comment content of model 3, etc., extracts keywords in text content through co-occurrence analysis, analyzes the association between different keywords and expresses them in the form of visualization. Combining the basic process of co-occurrence analysis and text knowledge mining, this paper summarizes the general application process of text knowledge mining using co-occurrence analysis as follows (see Figure 3):

The first step was to collect and organize the text information. Using the Octopus collector, we selected Zhihu as the target and collected the user experience comments of model 3. After the collection was completed, the irrelevant data were deleted and aggregated for further screening.

In the second step, the data is filtered (data noise reduction pre-processing). As the content of the collected text is a natural language description of the usage experience, the information is pre-processed with KH Coder. ① Setting of deactivation words and special words. ② Feature word extraction. Extract keywords from the text and adjust them by TF-IDF feature words to count the frequency of keyword occurrences in the filtered text. (iii) Co-occurrence matrix. A co-occurrence network of these words is formed, and the proximity of the network nodes reflects the affinity of the subject content for data analysis.

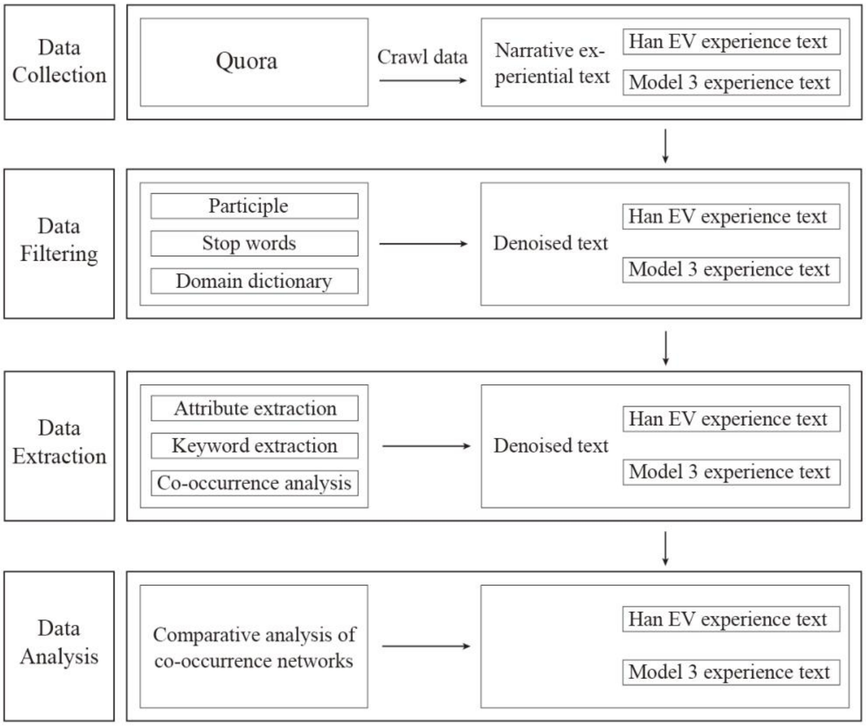

Figure 3: Research design methodology process.

The third step is data analysis. The visualization method of co-occurrence networks is an important research tool in the field of co-occurrence analysis, which helps to reveal the association of concepts, helps researchers to understand the focus quickly and facilitates the analysis and use of clustering results. In this study, the KH Coder software was used to generate the topic network, using keywords as nodes of the topic network, with larger nodes indicating a higher frequency of occurrence and smaller nodes indicating a lower frequency of occurrence; there will be connecting lines between the nodes, with shorter lines closer together indicating higher correlation and longer lines further away indicating lower correlation. This is how we can analyse the characteristics of the Tesla model 3 user experience in depth.

After collection, a total of 1582 samples were collected, imported into an Excel sheet and pre-processed with KH Coder to have a certain amount of text, then KH Coder could be used for text mining.

RESEARCH FINDINGS AND ANALYSIS

Keyword Extraction for Comment Content

After using the feature word extraction in KH Coder and adjusting the TF-IDF feature words, and then counting the high-frequency words (see Figure 4), we found that the typical labels of Model 3 users are engineers,

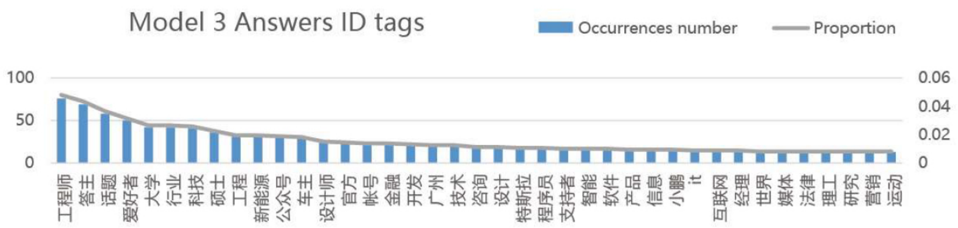

Figure 4: Model 3 Respondent ID self-description tag.

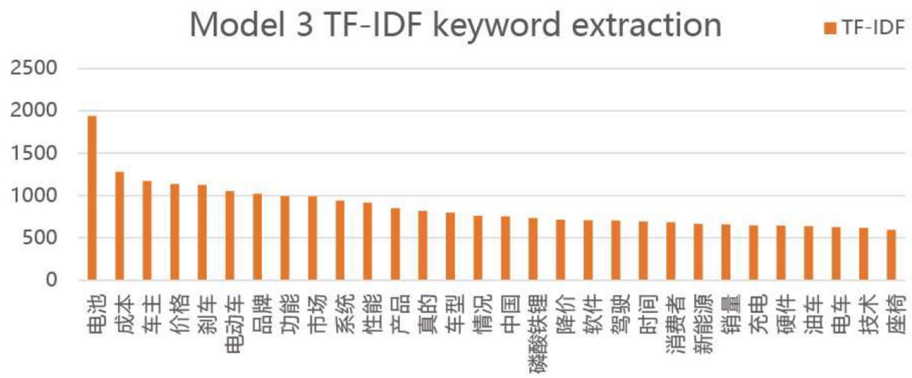

Figure 5: Model 3TF-IDF keyword extraction.

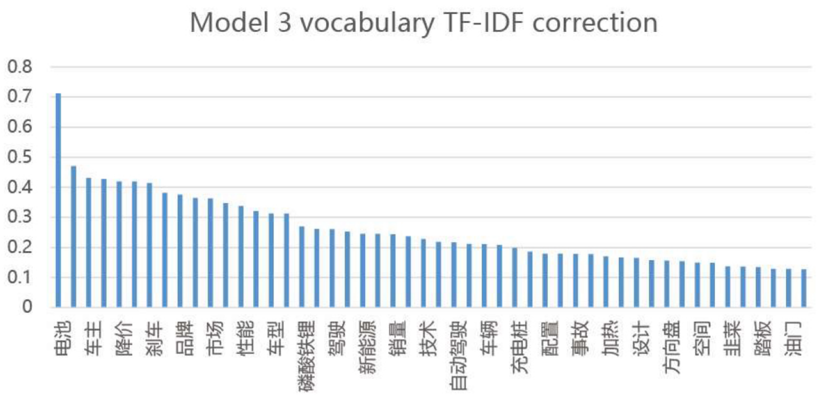

Figure 6: Model 3 vocabulary TF-IDF correction.

excellent topic responders, technology enthusiasts, finance, Internet, managers, etc. The users are willing to label themselves as "Tesla" and "car owners", which indicates that they have a certain sense of self-identity.

According to the TF-IDF algorithm and the "normalization" process, the keywords of the review content of Model 3were corrected (see Figure 5), as shown the typical labels of the review content of Model 3 are battery, cost, owner, price, brake, etc. (see Figure 6), and these keywords were categorized into three types: hardware, software and subjective experience. From these three aspects a more comprehensive analysis of the user experience can be made.

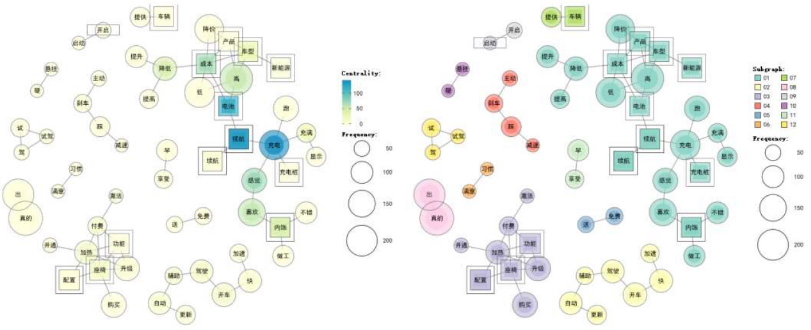

Figure 7: User experience in terms of hardware.

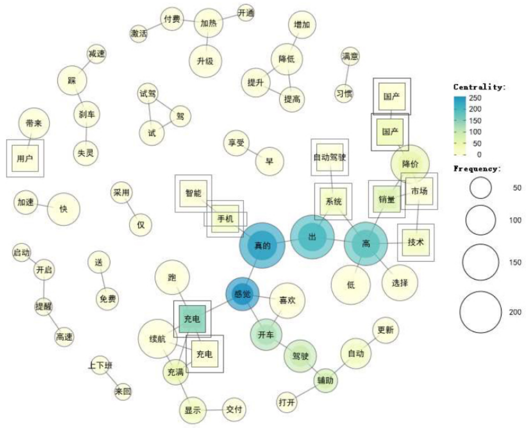

Figure 8: User experience in software.

User Experience Analysis

According to the relationship of the co-occurrence diagram (see Figure 7), it can be seen that Model 3 users' experiences are mainly focused on charging, range, battery, cost, and interior. In terms of hardware, Model 3 users are more concerned about cost and features, and their positive experience feelings are mainly focused on assisted driving, automatic updates, fast acceleration, active braking, hard suspension, like the interior, product price reduction and low cost. As shown in Figure 8, in terms of software, Model 3 users focus on driving mode, charging mode, brand, performance, etc. Their positive experience feelings are mainly on automatic driving mode and assisted driving function, price reduction, feeling of charging, etc. As in Figure 9, in terms of subjective experience, Model 3 users' experience is focused on driving, charging, brand, mode, price, and owner. Positive experience is mainly focused on price reduction, autonomous driving, fast acceleration and good charging experience, brake failure brings negative experience, and hardware and software upgrades and payment bring unclear feeling.

Figure 9 presents the narrative relationship between the model 3's user experience based on three dimensions: hardware, software and subjective

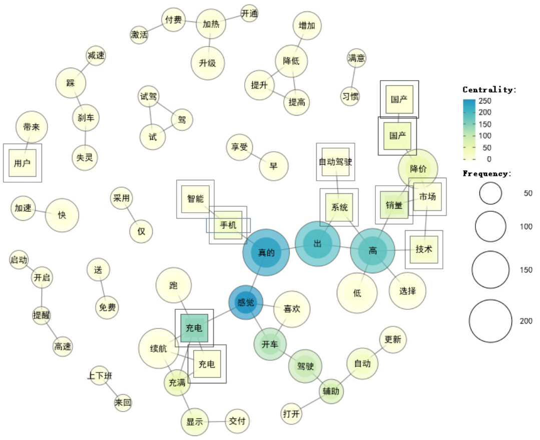

Figure 9: Subjective experience User experience in terms of.

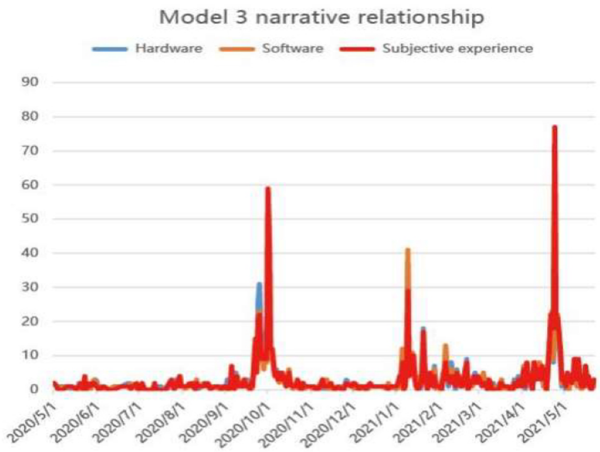

Figure 10: User experience in the time dimension.

experience, based on the time dimension. There are three large peak fluctuations between 1 May 2020 and 1 May 2021, with more moderate fluctuations until 1 October 2020 and a range of fluctuations after 1 January 2021. model 3 users generate significant attention in hardware in September-October 2020, concentrated attention in software in January-February 2021, and subjective experience The focus of attention in terms of subjective experience is concentrated in October 2020, January and May 2021. The reasons for this focus may be due to major events and iterations of the model 3, but the relationship diagram can only reveal a simple relationship between user experience and the exact reasons for this.

CONCLUSION

In this paper, software such as text mining is used to help understand user experiences by providing a perspective of real user experiences. The Octopus data collector and KH Coder text data mining software were used to obtain a network diagram of feature words and co-occurrence relationships for model 3. The results revealed that: i) on the hardware side, Model 3 focused more on functionality and cost. Secondly, in terms of software, the Model 3 focuses on charging, marketing and performance. Third, in terms of subjective experience, the Model 3 focuses more on owner, price and consumer and model concerns. The experience concerns can be clearly reflected in the co-occurrence diagram, but some of the deeper experience feelings are not yet visible as negative or positive experiences.

This is a hypothetical simulation method, in which the simulated experience of a group of users is used to represent the actual experience of users. Another typical qualitative research method is in-depth interview studies, such as customer journey maps, which are used to recreate the real experience of an individual user.

Based on the theory of big data and text mining, this paper proposes a user experience research method based on text mining and natural language processing, which is based on big data technology and excludes the errors caused by subjective factors such as questionnaires. The TF-IDF algorithm and "normalization" process are used to calculate the weight of each feature term. Of course, this method also has some limitations, as the researcher only conducted data mining on one platform, there are objective reasons such as insufficient data collection and incomplete setting of adjustment parameters, and the subjective thinking of the researcher is also one of the subjective factors. In the future, the author will continue to combine more updated data mining methods with visualization to obtain more accurate and clear user behavioral preferences and experiences, and to promote the in-depth development of the user experience industry.

ACKNOWLEDGMENT

This study was supported by the key project of national social science fund China (Grant No. 21zd215).

REFERENCES

Axsen, J., Bailey, J., & Castro, M. A. (2015). Preference and lifestyle heterogeneity among potential plug-in electric vehicle buyers. Energy Economics, 50, 190–201.

Bai Min, Zhang Minyu & Wang Xiaochao. (2021). Development policies and trends of the global new energy vehicle industry under the background of carbon neutrality. Information Technology and Standardization (12), 13–20.

Chen, H. (2001). Knowledge management systems: a text mining perspective. Knowledge Computing Corporation.

Hardman, S., Jenn, A., Tal, G., Axsen, J., Beard, G., Daina, N., ... & Witkamp, B. (2018). A review of consumer preferences of and interactions with electric vehicle charging infrastructure. Transportation Research Part D: Transport and Environment, 62, 508–523.

He, X., Zhan, W., & Hu, Y. (2018). Consumer purchase intention of electric vehicles in China: The roles of perception and personality. Journal of Cleaner Production, 204, 1060–1069.

Hidrue, M. K., Parsons, G. R., Kempton, W., & Gardner, M. P. (2011). Willingness to pay for electric vehicles and their attributes. Resource and energy economics, 33(3), 686–705.

Hu, M., & Liu, B. (2004, July). Mining opinion features in customer reviews. In AAAI (Vol. 4, No. 4, pp. 755–760).

Kang Bo, Liu Shengqiang. (2013). Internet business user experience management based on big data analysis. Telecommunications Science (03), 32-35.

Keim, D. A., Mansmann, F., Stoffel, A., Ziegler, H. (2008). Visual analytics.

Kim, S. G., & Kang, J. (2018). Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews. Information Processing & Management, 54(6), 938–957.

Li Xiang , Yang Fan. (2020). Research on Visualization Methods in User Experience Design. Design Art Research (02), 35–39.

Ma, S. C., Fan, Y., Guo, J. F., Xu, J. H., & Zhu, J. (2019). Analysing online behaviour to determine Chinese consumers' preferences for electric vehicles. Journal of Cleaner Production, 229, 244–255.

Morton, C., Anable, J., & Nelson, J. D. (2016). Exploring consumer preferences towards electric vehicles: The influence of consumer innovativeness. Research in transportation business & management, 18, 18–28.

Saif, H., He, Y., & Alani, H. (2012, November). Semantic sentiment analysis of twitter. In International semantic web conference (pp. 508–524). Springer, Berlin, Heidelberg.

Skippon, S. M., Kinnear, N., Lloyd, L., & Stannard, J. (2016). How experience of use influences mass-market drivers' willingness to consider a battery electric vehicle: a randomised controlled trial. Transportation Research Part A: Policy and Practice, 92, 26–42.

Tan Hao, You Zuo and Peng Shenglan. (2020). Overview of big data-driven user experience design. Packaging Engineering (02), 7–12+56.

Wu Xiaoqiu , Lu Na. (2012). Research on hot spot analysis method based on the frequency of keyword co-occurrence. Intelligence Theory and Practice (08), 115–119.

Wang Yuefen, Song Shuang & Xiong Minghui. (2007). Research on text knowledge mining method based on co-occurrence analysis. Library and Information Work (04), 66–70+79.

Wang Zhuoyu, Yuan Lei and Zhang Wenchao. (2020). Comparison of Chinese and Japanese STEM education research models based on KH Coder text data mining. Modern Distance Education Research (02), 56–63.

Wang Zhuoyu, Yuan Lei and Zhang Wenchao. (2020). Comparison of Chinese and Japanese STEM education research models based on KH Coder text data mining. Modern Distance Education Research (02), 56–63.

Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. Journal of Business Research, 69(5), 1562–1566.

Xu Zhiqiang, Xu Jinyu. (2019). User portrait construction and user experience optimization strategy based on big data. China Publishing (06), 52–56.

Zarrinkalam, F., Kahani, M., & Bagheri, E. (2018). Mining user interests over active topics on social networks. Information Processing & Management, 54(2), 339–357.